THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEURE DE

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE Sciences de la Vie et de la Santé

Biologie Cellulaire et Physiopathologie

Par Souad BENSAADA

ETUDE DE L'EFFET PERTURBATEUR ENDOCRINIEN D'ISOFLAVONES VEGETALES,

APPROCHES CLINIQUE ET MECANISTIQUE.

REDUCTION DE CES SUBSTANCES DANS L'ALIMENTATION HUMAINE.

Sous la direction de : Laurent AZEMA

Soutenue le 24/06/2024

Membres du jury :

M. JACQUOT, Yves, Professeur des universités, Université Paris Cité	Rapporteur
Mme MAGY-BERTRAND, Nadine, PU-PH, CHU de Besançon	Rapporteure
Mme DUMAY, Hélène, Maitresse de Conférence, Université de Bordeaux	Examinatrice
Mme BERQUE BESTEL, Isabelle, PU, Université de Bordeaux	Présidente

Membres invités :

M. AZEMA, Laurent, Maitre de Conférence, Université de Bordeaux	Directeur de thèse
Mme BENNETAU-PELISSERO, Catherine, Professeure, Bordeaux Sciences Agro	Co-encadrante

Titre : Etude de l'effet Perturbateur Endocrinien d'Isoflavones Végétales, Approches Clinique et Mécanistique. Réduction de ces Substances dans l'Alimentation Humaine.

Résumé :

La consommation de soja, connait une croissance notable en France. Cependant, le soja contient des isoflavones (IFs), des composés phytochimiques aux effets perturbateurs endocriniens à des doses élevées. Cette thèse s'attache à évaluer les risques associés aux IFs et à l'Entérolactone (ENL), autre phyto-estrogène notable, et à proposer des solutions favorisant une consommation de soja sûre.

Objectifs :

1. Évaluer l'exposition des Français·es aux IFs et valider de nouveaux outils d'estimation.
2. Étudier les effets endocriniens des IFs et de l'ENL chez l'humain, en se focalisant sur le lupus érythémateux disséminé (LED) et le cancer du sein triple négatif.
3. Développer des techniques pour réduire les teneurs en IFs des aliments à base de soja.

Méthodologie :

1. Exposition aux IFs et à l'ENL :
   - Analyses d'aliments et validation de questionnaires alimentaires spécifiques pour évaluer la consommation de soja et l'exposition aux IFs et aux lignanes en France.
   - Dosages des IFs dans les fluides biologiques (sang, urine) et les cheveux.
2. Effets endocriniens des IFs et de l'ENL :
   - Étude cas-témoin sur le LED, comparant l'exposition aux IFs et à l'ENL de patientes et de volontaires saines
   - Études *in vitro* de l'effet des IFs sur la prolifération de cellules de cancer du sein triple négatif, ainsi que des interactions des IFs avec le récepteur des estrogènes GPER.
3. Réduction des teneurs en IFs:
   - Développement de procédés (pré)industriels pour réduire les IFs des fèves de soja.
   - Étude des recettes traditionnelles asiatiques de soja et de leurs effets sur la réduction des IFs.

Résultats :

1. Exposition aux IFs et à l'ENL :
   - L'exposition aux IFs en France est significative, même chez les non-consommateur·ices de soja, à cause du « soja caché » présent dans les aliments ultra-transformés.
   - Les questionnaires alimentaires sont validés et les IFs et l'ENL sont dosés avec succès dans les cheveux.
2. Effets endocriniens des IFs et ENL :
   - Les résultats préliminaires suggèrent une association entre l'ENL et une réduction du risque de LED mais l'effectif limité de l'étude ne permet pas de conclure sur l'impact des IFs.
   - Les IFs et leurs conjugués circulants pourraient induire la prolifération des cellules cancéreuses du sein triple négatives, via une interaction avec le GPER.
3. Réduction des teneurs en IFs :
   - Des rinçages industriels permettent de réduire les IFs des fèves de soja de 50%.
   - Les recettes traditionnelles asiatiques et les rinçages domestiques réduisent significativement les IFs des aliments à base de soja.

Discussion :

La consommation de soja nécessite une vigilance particulière en raison des effets endocriniens des IFs. Des recherches complémentaires sont nécessaires pour comprendre ces effets et proposer des recommandations pour une consommation sûre du soja.

Conclusion :

Si la consommation de soja est traditionnelle en Asie et présente un intérêt nutritionnel certain, elle soulève des questions quant à ses répercussions sur la santé humaine à cause de la présence d'IFs estrogéniques et antithyroïdiennes. Cette thèse apporte des éléments de réponse sur l'exposition aux IFs en France, leurs effets potentiels sur le LED et les cancers du sein triple négatifs, et propose des solutions pour réduire les teneurs en IFs des aliments à base de soja.

Title : Study of the Endocrine Disrupting Effect of Plant Isoflavones, Clinical and Mechanistic Approaches. Reduction of these Substances in Human Food.

Abstract:

The consumption of soy, a valuable source of plant-based protein, has seen a steady rise in France. However, soy contains isoflavones (IFs) which are phytoestrogens with endocrine-disrupting effects at high doses. This thesis investigates the risks associated with IFs and enterolactone (ENL), another notable phytoestrogen, and explores strategies for promoting safe soy consumption.

Objectives:

1. Assess French exposure to IFs and ENL, and validate tools for accurate estimation.
2. Investigate the endocrine effects of IFs and ENL in humans, focusing on systemic lupus erythematosus (SLE) and triple-negative breast cancer.
3. Develop techniques to reduce the IFs content of soy-based food products.

Methodology:

1. IF and ENL Exposure Assessment:
   - Food analyses and dietary intake assessment to quantify the French exposure to IFs and lignans.
   - Development and validation of food questionnaires to evaluate soy and lignans consumption.
   - Measurements of IFs and ENL in biological fluids (blood, urine) and hair.
2. Endocrine effects of IFs and ENL:
   - Case-control study on SLE, comparing IFs and ENL exposure in patients and in healthy controls.
   - *In vitro* studies on triple-negative breast cancer cells, analyzing the effect of IFs on cell proliferation and their interactions with the GPER, a membrane estrogen receptor.
3. Reduction of IF levels:
   - Development of (pre)industrial rinsing processes to decrease IFs content in soybeans.
   - Investigation of traditional Asian recipes and domestic water rinsing techniques for reducing IFs in soy-based foods.

Results:

1. IFs and Lignans exposure:
   - French exposure to IFs is significant, even for non-soy consumers, due to “hidden soy” in ultra-processed foods.
   - The developed food questionnaires are validated, and a method for IFs and ENL analysis in hair is presented.
2. Endocrine effects of IFs and ENL:
   - Preliminary findings from the SLE cohort suggest a potential association between ENL and a healthier status. However, the limited sample size necessitates further investigation to definitively exclude an impact of IFs.
   - IFs and their circulating conjugates might promote triple-negative breast cancer cells proliferation through the GPER.
3. Reducing IF levels:
   - Industrial rinsing processes can effectively reduce IFs content in soybean by 50%.
   - Traditional Asian preparation methods and domestic rinsing significantly decrease IFs in soy-based products.

Discussion: Soy consumption warrants specific attention due to the endocrine effects of IFs. Further research is crucial to gain a more comprehensive understanding of these effects and formulate recommendations for safe soy consumption.

Conclusion: While soy consumption holds nutritional value and boasts a long tradition in Asia, its high isoflavone content raises questions about its potential impact on human health. This thesis contributes preliminary insights into French exposure to IFs, their possible effects on SLE and triple-negative breast cancer, and proposes solutions for mitigating IFs levels in soy-based foods

---

Unité de recherche

[ARNA : Régulations Naturelle et Artificielle, Inserm U1212 - CNRS UMR 5320 - Université de Bordeaux, Bat BBS Campus Carreire - 2, Rue du Dr Hoffmann Martinot, 33000 Bordeaux].

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude aux Pr. Nadine MAGY-BERTRAND et Pr. Yves JACQOT pour avoir accepté de siéger au jury de ma soutenance de thèse et d'avoir fait le déplacement malgré la distance. Je suis particulièrement honorée de vos présences et de l'intérêt que vous avez porté à mes travaux. Je remercie également la Pr. Isabelle BERQUE BESTEL d'avoir accepté de présider ce jury avec enthousiasme et bienveillance. Je suis infiniment reconnaissante envers la Dr. Hélène DUMAY-ODELOT pour ton soutien constant et ta disponibilité. Ton calme, ton expertise et ta sagesse m'ont permis de traverser les moments difficiles avec sérénité. Merci de me faire l'honneur d'être dans mon jury de thèse.

Je tiens également à remercier chaleureusement la Pr. Catherine BENNETAU-PELISSERO de m'avoir donné l'opportunité de travailler sur ce projet de recherche passionnant et de m'avoir transmis ton intérêt pour les isoflavones. Ton encadrement scientifique, tes encouragements et ta confiance ont été précieux pour moi. Je remercie chaleureusement le Dr. Laurent AZEMA pour avoir pris le relais de mon encadrement en fin de thèse. Tes conseils avisés et ta bonne humeur (oui oui) m'ont permis de finaliser mon projet dans les meilleures conditions. Un merci tout particulier pour l'inoubliable crevette.

Je suis également reconnaissante envers la Dr. Jeanne CHAIN LEBLOND pour ton soutien indéfectible et ta gentillesse. Tu as su créer une ambiance d'équipe conviviale et inclusive qui m'a permis de m'épanouir au sein du laboratoire. Merci pour toutes les soirées TAMS, meilleure équipe du monde ! Je tiens à remercier tout particulièrement Aurore GUÉDIN-BEAUREPAIRE pour ta rigueur, ton amitié et ton aide précieuse. Sans ton intervention, j'en serais encore à pleurer sur mes *western*. J'ai de la gratitude pour tous les moments partagés, tant sur le plan professionnel que personnel. Un grand merci à vous trois d'avoir fait l'effort de m'accompagner en tant qu'équipe au moment et après le changement de direction de thèse, malgré l'éloignement de mon sujet par rapport à vos thématiques. Je remercie également Clémentine AUBRY pour ton écoute et ta solidarité. C'était précieux de savoir que j'avais une alliée dans le *game*. Merci aussi à Jean DESSOLIN pour ton aide et ta disponibilité.

Je suis reconnaissante envers Gabriele PERRUZZI et plus largement Biopress/Berkem pour avoir financé ces trois années de recherche. Votre confiance et soutien ont été indispensables à la réalisation de mon projet de thèse.

A ma mère, merci pour les relectures minutieuses et les suggestions éclairées. Ce tapuscrit aurait été différent sans tes interventions, and *so do I*. A mon frère, merci pour tes conseils que je n'écoute pas toujours, merci d'être là pour moi. J'aurais aimé que vous soyez là.

A Aurore, de m'avoir réconciliée avec R. A Juliet, pour l'amour, les plans sur la comète, la radio, les vacances, la vaisselle et tellement d'autres choses... A Morgane, pour les séances de psy gratuites au gout de menthe et de soleil, pour la joie. A Sara, de grandir avec moi. A Sarah, Samia, Céleste, Delphine et Ariane, pour la force. A Claire, Julie, Steph, Axelle et Nolwenn, de me faire aimer Bordeaux et les balades à vélo. A Kafui, pour l'inspiration et le courage de se lancer. A Wiame, Hayat, Amina, Narimane et Feriel, je ne serai pas celle que je suis aujourd'hui si je ne vous avais pas dans ma vie. Alors, merci !

Préambule

La présente thèse de doctorat est le fruit d'un parcours académique et professionnel qui m'a orientée vers l'étude des effets perturbateurs endocriniens des isoflavones végétales, notamment celles du soja. En tant que pharmacienne diplômée de l'Université Es-Senia d'Oran (Algérie) et titulaire d'un master en Plantes à valeur santé et biomolécules d'intérêt de l'Université de Bordeaux, j'ai développé un intérêt profond pour la santé humaine et ses enjeux.

Mon engagement pour la santé publique s'est matérialisé lors de mon stage au sein du laboratoire ARNA, sous la direction de la Professeure Catherine Bennetau-Pelissero. En participant à l'élaboration d'une procédure visant à réduire les composés phyto-estrogéniques dans des matrices alimentaires de soja, j'ai pris conscience des effets potentiellement néfastes des perturbateurs endocriniens sur la santé humaine. Mon choix de thèse a été renforcé par mon régime végétarien, qui m'a sensibilisée à la présence de ces substances dans l'alimentation.

J'ai entrepris cette thèse de doctorat sous la direction initiale de la Pr. Catherine Bennetau-Pelissero, une pionnière dans la recherche sur les effets perturbateurs endocriniens des isoflavones de soja en France. Cette thèse CIFRE, soutenue par l'entreprise Biopress/Berkem, spécialisée dans la production d'huiles végétales issues de l'agriculture durable et engagée pour la santé publique, avait pour objectif d'explorer les différentes facettes de l'impact des isoflavones végétales sur l'organisme humain.

Suite à des circonstances imprévues, la direction de la thèse a été transférée au Dr. Laurent AZEMA en milieu de parcours. Malgré ce changement, mon engagement et ma détermination pour ce sujet de recherche n'ont jamais faibli. J'ai eu la chance de bénéficier de l'expertise et du soutien de mes deux directeur·ices de thèse, qui m'ont guidée tout au long de ce projet ambitieux.

Table des Matières

Table des Matières 6

I. Introduction générale 13

II. Etat de l'art et problématique 15

II.1. Perturbateurs endocriniens 15

II.1.1. Généralités 15

II.1.1.1. Effet cocktail 17

II.1.1.2. Période de vulnérabilité 17

II.1.2. Phyto-estrogènes 18

II.2. Isoflavones 20

II.2.1. ADME (Absorption-Distribution-Métabolisme-Elimination) 20

II.2.2. Toxicologie 22

II.2.3. Mécanismes d'action 24

II.2.4. Effets biologiques 25

II.2.4.1. Maladies thyroïdiennes 29

II.2.4.2. Cancers hormono-dépendants 29

II.2.4.3. Fibromes utérins 33

II.2.4.4. Fertilité 34

II.2.4.5. Cycle menstruel 35

II.2.4.6. Ménopause et ostéoporose 36

II.2.4.7. Métabolisme 36

II.3. Exposition aux phyto-estrogènes, enjeux de santé et réduction 38

II.3.1. Exposition aux perturbateurs endocriniens en France 38

II.3.1.1. Isoflavones 38

II.3.1.2. Exposition à d'autres perturbateurs endocriniens 39

II.3.2. Impact de la transformation alimentaire sur la teneur en isoflavones des aliments 40

II.3.2.1. Transformations traditionnelles à l'échelle domestique 41

II.3.2.2. Transformations à l'échelle industrielle 42

II.3.3. Evaluation de l'exposition aux phyto-estrogènes 42

II.3.3.1. Enquêtes alimentaires 42

II.3.3.2. Dosages dans les échantillons biologiques 43

II.4. Lupus Erythémateux Disséminé 43

II.4.1. Maladies auto-immunes 43

II.4.2. Lupus Erythémateux Disséminé 44

II.4.2.1. Généralités 45

II.4.2.2. Prévalence du LED 45

II.4.2.3. Facteurs de risque 46

II.4.2.4. Physiopathologie 47

II.4.3. Isoflavones et LED 48

II.5. Cancer du sein triple négatif (TNBC) 49

II.5.1. Le cancer 49

II.5.2. Evolution 50

II.5.3. Les cancers du sein 50

II.5.3.1. Classification 50

II.5.3.2. Prévalence et facteurs de risque 51

II.5.3.3. Physiopathologie 52

II.5.3.4. Résistance aux traitements 53

II.5.4. GPER 54
II.5.4.1. Données générales 54
II.5.4.2. Distribution 54
II.5.4.3. Site actif/agonistes/antagonistes 55
II.5.4.4. Fonctions physiologiques 55
II.5.4.5. Implication dans les cancers du sein 56
II.5.4.6. Mécanismes d'action 58
II.5.4.7. Complexité et controverse 66
II.5.5. Isoflavones et GPER 67

III. Méthodologie, matériels et méthodes 69

III.1. Evaluation de l'exposition des consommatrices et consommateurs et étude ISOLED 69
III.1.1. Exposition aux phyto-estrogènes et aux pesticides 69
III.1.1.1. Phyto-estrogènes 69
III.1.1.2. Produits phytosanitaires 70
III.1.2. Questionnaires alimentaires 71
III.1.2.1. Enquêtes d'habitudes alimentaires 72
III.1.2.2. Rappels des 24 et 48 heures 72
III.1.2.3. Etablissement de scores 73
III.1.3. Analyses des biomarqueurs 74
III.1.3.1. Extractions des isoflavones 74
III.1.3.2. Mise en œuvre de l'analyse 75

III.2. Diminution des isoflavones dans les aliments 76
III.2.1. A l'échelle du laboratoire 76
III.2.2. A l'échelle pilote (préindustrielle) 77
III.2.3. A l'échelle domestique à partir des recettes traditionnelles 78

III.3. Etudes cellulaires 78
III.3.1. Cultures cellulaires 78
III.3.2. Test de prolifération 78

III.4. Analyses moléculaires 79
III.4.1. Western Blot 79
III.4.1.1. Extraction et dosage des protéines 79
III.4.1.2. Migration sur gel 80
III.4.1.3. Transfert sur membrane 80
III.4.1.4. Détection 81
III.4.2. Immunocytochimie 82

III.5. Modélisation moléculaire 83
III.5.1. Construction par homologie 83
III.5.1.1. Etapes de construction du modèle 83
III.5.1.2. *Rescoring* 85
III.5.2. Ancrage moléculaire (*docking*) et notation (*scoring*) 86
III.5.2.1. Sélection des ligands à tester 86
III.5.2.2. *Docking* 88
III.5.3. *Rescoring* et consensus 88

III.6. Analyses statistiques 89
III.6.1. ISOLED 90
III.6.1.1. Analyse à n=57 90
III.6.1.2. Analyse après la fin de l'inclusion 91
III.6.2. PHYTOSECUR 1 et 2 92
III.6.3. *Docking* moléculaire 92

III.6.4. Etudes cellulaires 92

IV. Résultats et discussions *93*

IV.1. Exposition aux isoflavones en France 93

IV.1.1. Comparaison de l'exposition des Françaises et Français aux pesticides perturbateurs endocriniens et aux isoflavones du soja : article correspondant à la référence [213] 93

IV.1.2. Estimation de l'exposition des femmes de l'étude ISOLED aux phyto-estrogènes 105

IV.1.2.1. Estimation de l'exposition habituelle : validation d'un questionnaire d'habitude alimentaire par l'analyse de biomarqueurs capillaires. Article correspondant à la référence [214] 105

IV.1.2.2. Evaluation de la consommation récente : validation de questionnaires alimentaires par l'analyse de biomarqueurs urinaires et plasmatiques. Article correspondant à la référence [25] 123

IV.1.3. Exploration de l'influence de l'exposition à des phyto-estrogènes naturels sur la déclaration de poussées de lupus érythémateux disséminé 142

IV.1.3.1. Normalité des données 142

IV.1.3.2. Influence du statut sanitaire sur les scores et valeurs mesurées des concentrations d'IFs et d'ENL 144

IV.1.3.3. Corrélations entre dosages biologiques et scores alimentaires 146

IV.1.3.4. Analyse des données par ACP 150

IV.1.3.5. Discussion 154

IV.2. Réduction de l'exposition aux isoflavones : une nécessité pour la santé publique 157

IV.2.1. Réduction des isoflavones à l'échelle préindustrielle PHYTOSECUR 1 : article correspondant à la référence [49] 157

IV.2.2. Impact des traitements traditionnels et domestiques sur les taux d'IFs des aliments et conséquences sur l'estimation de l'exposition aux IFs en Asie : article correspondant à la référence [223] 175

IV.3. Etude cellulaire du GPER 195

IV.3.1. *Docking* 195

IV.3.1.1. Construction du modèle de récepteur 195

IV.3.1.2. Validation du modèle 197

IV.3.1.3. *Docking* des isoflavones et conjugués et *Rescoring* 200

IV.3.2. Etude de la prolifération 203

IV.3.2.1. Validation du modèle cellulaire 203

IV.3.2.2. Effet des isoflavones et implication du GPER 206

V. Discussion générale *208*

VI. Conclusion générale et perspectives *216*

VII. Références bibliographiques *218*

Annexes *239*

Liste des abréviations

AC : Adénylate cyclase

ACP : Analyse par composants principaux

AFSSA : Agence Française de Sécurité Sanitaire des Aliments

AFSSAPS : Agence Française de Sécurité Sanitaire des Produits de Santé

AMPc : adénosine monophosphate cyclique

Anses : Agence nationale de sécurité sanitaire, de l'alimentation, de l'environnement et du travail

APS : Ammonium persulfate

Arg : Arginine

BCA : *BiCinchoninic acid Assay* ou dosage de l'acide BiCinchoninique

BLAST : *Basic Local Alignment Search Tool* ou outil de recherche d'alignements locaux de base

BSA : Albumine sérique bovine

CAF : Fibroblastes associés aux cancers

CSD : *Cambridge Structural Database*

CT : Cholestérol total

Cys : Cystéine

DAI : Daidzéine

DMO : Densité Minérale Osseuse

DMSO : Diméthylsulfoxyde

EBV : Virus d'Epstein-Barr

ECR : Essais contrôlés randomisés

EGF(R) : *Epidermal Growth Factor (Receptor)* ou (récepteur aux) facteurs de croissance épidermiques

ELISA : *Enzyme-Linked Immunosorbent Assay* ou technique d'immunoadsorption par enzyme liée

EMT : Transition Epithélio-mésenchymateuse

ENL : entérolactone

ERK : *Extracellular Signal-Regulated Kinases* ou kinases régulées par le signal extracellulaire

FIV : Fécondation *in vitro*

FSH : *Folliculo-stimulating Hormone* ou hormone hypophysaire contrôlant la folliculogénèse gonadique

GEN : Génistéine

GEN-7-G : Génistéine -7-Glucuronide

Gln : Glutamine

Glu : Glutamate

Gly : Glycitéine

GnRH : *Gonadotrophin Releasing Hormone* ou gonadolibérine hypothalamique

GPER : *G Protein coupled Estrogen Receptor* ou récepteur aux estrogènes couplé à la protéine G, aussi désigné sous l'abréviation GPR30

GWAS : *Genome wide association studies* ou études d'association pangénomique

HbA1c : Hémoglobine glyquée

HB-EGF: *Heparin-Binding EGF* ou facteur de croissance épidermique lié à l'héparane

HDL : Lipoprotéines à haute densité

HOMA-IR : *Homeostasic model assessment of insulin resistance*

IC : Intervalle de confiance

IC_{50} : Concentration inhibitrice médiane

IFs : Isoflavones

Ile : Isoleucine

Kd : Constante de dissociation

LDL : Lipoprotéines à faible densité

LED : Lupus Erythémateux Dessiminé = Lupus Erythémateux Systémique

Leu : Leucine

LH : *Luteinising Hormone* ou hormone hypophysaire lutéinisante contrôlant l'ovulation

LOAEL (DMENO) : *Lowest-observed-adverse-effect level* ou dose minimale avec effet nocif observé

LPH : Lactase Phloridzine Hydrolase

Met : Méthionine

METABRIC : *Molecular Taxonomy of Breast Cancer International Consortium*

MMP : Métalloprotéinases de la matrice

NOAEL (DMSENO) : *No Observable Adverse Effect Level* ou dose maximale sans effet néfaste observable

NTP : *National Toxicology Programm*

OR : *Odd Ratio*

PAF: Paraformaldéhyde

PCB: Polychlorobiphényles

PDB : *Protein Data Bank*

PE : perturbateur endocrinien

Phe : Phénylalanine

PI3K : phosphoinositide 3-kinase

PFAS : perfluoroalkylés et polyfluoroalkylés

PhytoE : Phyto-estrogènes

Pro : Proline

PSA : Antigène prostatique spécifique

PST : Protéines de soja texturées

PTI : Purpua Thrombopénique Immunologique

PXR : *Peroxisome X Receptor*

RBA : affinités de liaison relatives

RMN : Résonnance magnétique nucléaire

RMSD : *Root Mean Square Deviation* ou Écart quadratique moyen

ROC : *Receiver Operating Characteristic* ou fonction d'efficacité du récepteur

SDS : Sodium dodecylsulfate

Ser : Sérine

SERD : *Selected Estrogen receptor*

SERM : *Selected Estrogen receptor modulator*

SRE : Eléments de réponse sériques

Tam : Tamoxifène, médicament anti-estrogène

TBS : *Tris buffer saline* ou tampon tris salin

TBST : TBS avec du tween 20

TCGA : *The Cancer Genome Atlas*

TGS : Tris-Glycine-SDS

TEMED : N,N,N′,N′-tétraméthylethylene-diamine

TG : Tris-glycine

Trp : Tryptophane

Tyr : Tyrosine

Val : Valine

WB : *Western blot*

Liste des Figures

Figure 1. Liste des phyto-estrogènes principaux 18
Figure 2. Structure du 17 β estradiol 19
Figure 3. Pourcentage de produits alimentaires contenant des légumineuses dans les différentes catégories proposées par le fournisseur de restauration collective (D'après Lee *et al*. [117]). 39
Figure 4. Incidence mondiale des cancers en 2020 chez (A) les hommes et (C) les femmes. Distribution des cas et décès pour les 10 cancers les plus communs en 2020 chez les hommes (B) et les femmes (D) (Sung *et al*. [149]). 49
Figure 5. Distribution des cancers du sein selon la présence des récepteurs aux estrogènes (ER), progestérone (PR) et GPER dans des spécimens de biopsie (D'après Hall et Fillardo [42]) 51
Figure 6. Voies de signalisation du GPER, d'après Girgert *et al*. [171] 59
Figure 7. Boites à moustaches représentant l'influence du statut sanitaire des volontaires (LED = femmes atteintes de LED, n= 20; PTI = femmes atteintes de PTI, n= 12 ; SAINES groupe témoin non malade, n= 31) sur l'ensemble des variables de l'étude ISOLED. 144
Figure 8. Boites à moustache représentant l'influence du statut sanitaire des volontaires (LED = femmes atteintes de LED, n=19 ; PTI = femmes atteintes de PTI, n= 11 ; SAINES groupe témoin non malade, n= 25) sur les variables IFs de l'étude ISOLED sans les valeurs extrêmes 145
Figure 9. Boites à moustache représentant l'influence du statut sanitaire des volontaires (LED = femmes atteintes de LED, n= 20 ; PTI = femmes atteintes de PTI, n= 12 ; SAINES groupe témoin non malade, n= 31) sur l'ensemble des variables ENL de l'étude ISOLED 146
Figure 10. Boites à moustache représentant l'influence du statut sanitaire des volontaires (LED = femmes atteintes de LED, n=19 ; PTI = femmes atteintes de PTI, n= 11 ; SAINES groupe témoin non malade, n= 25) sur les variables ENL de l'étude ISOLED sans les valeurs extrêmes 146
Figure 11. Représentation de la distribution des dosages d'ENL plasmatiques selon le statut des volontaires 146
Figure 12. Contributions de chaque dimension à l'explication de la variance 150
Figure 13. Cercle de corrélation entre les variables initiales et les composantes 151
Figure 14. Diagramme de corrélation des contributions des anciennes variables aux nouvelles dimensions 152
Figure 15. Nuage des individus répartis par statut et ellipses de concentration 153
Figure 16. Représentation des individus par groupe sanitaire selon les 3 principales dimensions 153
Figure 17. Comparaison des structure 3D des formes construites active (jaune) et inactive (rouge) du GPER vues sous 2 angles différents 198
Figure 18. Résidus du GPER_activé impliqués dans l'interaction avec les agonistes du *training-set* 199
Figure 19. Construction de la structure 3D du GPER avec représentation de sa surface avec en bleu les parties hydrophiles (compartiments extra- et intra-cellulaires) et en marron la partie hydrophobe (intramembrannaire); small-box = grille rouge 200
Figure 20. Poses de *docking* du G4G avec les résidus impliqués dans l'interaction avec le GPER. 203
Figure 21. Interactions G4G_GPER 203
Figure 22. Résultat du *western Blot* pour le GPER. 204
Figure 23.résultats de la détection du GPER par immunocytochimie dans (A) MDA-MB-436, (B) MDA-MB-468, (C) SKBR3 et (D) MCF7 205
Figure 24. Résultats du *western blot* pour ER. 205
Figure 25. Résultats de prolifération des cellules MDA-MB-468, 24 et 48 heures après traitement à la GEN-4-Glucuronide et DAI-7-Glucuronide. 207

Liste des Tableaux

Tableau 1. Résumé d'études épidémiologiques évaluant l'impact des isoflavones sur la santé humaine. 28
Tableau 2. Tableau clinique du LED (D'après Falten *et al.* [139]) 45
Tableau 3. Voies immunitaires clés impliquées dans la pathogénicité du LED (D'après Conelly *et al.* [143]) ... 48
Tableau 4. GPER et cancer du sein (complété de Hsu *et al* [172]) 65
Tableau 5. Conditions expérimentales (facteurs et niveaux) pour la réduction des isoflavones à l'échelle du laboratoire 77
Tableau 6. Liste des anticorps primaires et de leurs usages correspondants lors de l'analyse par *Western Blot* 81
Tableau 7. Liste des molécules du *training_set*, avec leurs activités correspondantes vis à vis du GPER 84
Tableau 8. Liste des isoflavones et conjugués utilisées pour le test de *docking* 88
Tableau 9. Résultats des tests de normalité de l'ensemble des données de l'étude ISOLED (n=63) 144
Tableau 10. Corrélations de Spearman de l'étude ISOLED 149
Tableau 11. Comparaison des résultats d'ISOLED avec d'autres études utilisant le dosage des IFs ou de l'ENL dans les fluides biologiques comme marqueur d'exposition. 154
Tableau 12. Résultats des calculs d'effectifs par groupe selon plusieurs scénarii. 156
Tableau 13. Résultats de l'analyse des structures 3D des protéines homologues au GPER 196
Tableau 14. Résidus cystéine conservés entre CXCR2 et GPER 197
Tableau 15. Fonctions de score retenues pour le consensus 200
Tableau 16. Classement des IFs et conjugués selon les scores obtenus avec les fonctions de score les plus pertinentes 202

I. Introduction générale

La consommation croissante de soja en France, en tant que source alternative et riche en protéines, s'inscrit dans un contexte de réduction de la consommation de produits carnés, motivée par des préoccupations environnementales et de santé publique. Le soja est également riche en fibres et en micronutriments, ce qui en fait un atout précieux pour une alimentation équilibrée.

Cependant, il contient aussi des isoflavones (IFs) qui sont des phyto-estrogènes (phytoE). Ces molécules, étant les xéno-estrogènes les plus semblables à l'estradiol (E_2) [1] une hormone sexuelle féminine, présentent, comme ce dernier, des effets bénéfiques ou néfastes sur la santé humaine, selon le stade physiologique des personnes exposées et les doses ingérées.

Parmi les effets bénéfiques potentiels majeurs, il convient de noter : la réduction des symptômes de la ménopause, tels que les bouffées de chaleur ; la protection contre certains cancers hormono-dépendants, comme le cancer du sein (dans des conditions très particulières) ; la préservation de la santé osseuse et la réduction du risque d'ostéoporose [2,3].

A contrario, une consommation excessive d'IFs a été associée à : des troubles de la fertilité chez les femmes et les hommes [3–5] ; l'augmentation du risque de certaines maladies hormono-dépendantes, comme le cancer du sein et le fibrome utérin ; la perturbation du système thyroïdien [3] et l'altération de la fonction immunitaire, pouvant augmenter le risque d'infections et de maladies auto-immunes (MAI). Des effets reprotoxiques et carcinotoxiques ont aussi été rapportés pour une IF du soja dans des études toxicologiques de références chez l'animal.

En particulier, des données épidémiologiques récentes suggèrent un lien potentiel entre les IFs et la croissance de tumeurs du sein triple négatives (TNBC), un type de cancer agressif et difficile à traiter [6]. Certaines de ces cellules cancéreuses présentent un récepteur aux estrogènes GPER pour lequel les IFs ont une grande affinité [7]. En outre, les IFs, ayant des propriétés estrogéniques, elles réduisent la synthèse des hormones gonadotropes hypophysaires et peuvent aussi être anti-androgéniques. Elles ont donc la capacité d'influencer les niveaux hormonaux et les réponses immunitaires, ce qui pourraient avoir un impact sur le développement et l'évolution de maladies estrogéno-sensibles telles que le Lupus Erythémateux Disséminé (LED).

Bien que des études aient signalé des effets positifs des IFs, en particulier chez des populations asiatiques ou chez des femmes souffrant d'un déficit estrogénique [2], l'existence d'effets potentiellement nocifs a également été soulevée dans d'autres recherches scientifiques. Cette dualité d'effets complique l'évaluation précise de leur impact réel sur la santé humaine, suscitant ainsi des débats dans le domaine. Par conséquent, il est impératif d'approfondir notre compréhension des mécanismes d'action des IFs et de leurs conséquences sur la santé humaine.

Les effets des IFs sur la santé humaine ont fait l'objet d'avis da la part des agences sanitaires françaises et de recherches internationales, partiellement financées par l'industrie agro-alimentaire. Dans leur rapport de 2005 [8], l'AFSSA (Agence Française de Sécurité Sanitaire des Aliments) et l'AFSSAPS (Agence Française de Sécurité Sanitaire des Produits de Santé) recommandaient de limiter l'exposition aux IFs, notamment chez les nourrissons et les enfants de moins de 3 ans, les femmes enceintes

et les femmes ayant des antécédents personnels et familiaux de cancers du sein. Ces recommandations incluaient l'interdiction des formules infantiles à base de soja et la limitation de la consommation d'IFs chez les adultes à 1 mg/kg/jour. Cette limite a été établie en tenant compte des estimations de l'exposition des populations japonaises, telles que publiées à l'époque, et qui ne semblaient pas associées à des effets néfastes. Les agences avaient également préconisé l'étiquetage des aliments contenant du soja pour informer les consommateur·ices de leur teneur en IFs. Malheureusement, ces deux dernières mesures n'ont jamais été mises en place.

La complexité des effets des IFs sur la santé humaine et les incertitudes qui persistent autour de leur innocuité motivent cette thèse doctorale.

Son objectif principal est de contribuer à une meilleure compréhension des mécanismes d'action et impacts potentiels des IFs de soja sur la santé humaine.

Il en découle les objectifs spécifiques suivants :

- quantifier de manière exhaustive l'exposition des consommateur·ices Français·es aux IFs et la comparer à celle d'autres perturbateurs endocriniens (PEs) ;
- développer et valider de nouveaux outils performants pour quantifier l'exposition aux IFs ;
- examiner le lien potentiel entre l'exposition aux IFs et les poussées de LED;
- étudier les effets des IFs sur la prolifération des cellules du TNBC et le rôle du récepteur GPER ;
- élaborer des stratégies domestiques et industrielles visant à diminuer la teneur en IFs des produits à base de soja, tout en préservant leurs bienfaits nutritionnels, en s'appuyant sur des pratiques de transformation modifiées ou optimisées fondées sur des preuves scientifiques validées;
- soutenir les autorités sanitaires à formuler des recommandations pertinentes de comportements alimentaires.

Dans ce contexte, les études envisagées dans le cadre de cette thèse visent à répondre aux questions de recherche clés énoncées ci-dessus en s'articulant autour des trois axes principaux interdépendants suivants.

1- Evaluation de l'exposition aux IFs de soja en France : il est envisagé d'une part de quantifier l'exposition des consommateur·ices aux IFs et la comparer à celles d'autres PEs et, d'autre part, de développer et valider de nouveaux outils utilisés dans ce but. A cet effet, il faut identifier les sources alimentaires majeures d'exposition aux IFs en France, en tenant compte des produits à base de soja, des compléments alimentaires et du « soja cachés ». L'hypothèse ici est que cette exposition est actuellement sous-estimée en France. En particulier, elle est souvent négligée dans l'évaluation de l'exposition aux PEs, alors qu'elle pourrait être significative. Afin de pallier à l'imprécision liée aux variabilités individuelles et temporelles, il y a lieu de mesurer l'exposition, aussi bien ponctuelle qu'à plus long terme, aux phytoE, notamment IFs, de sujets sains et malades pour mieux comprendre leur distribution et améliorer les scores de corrélation. Pour ce faire, l'analyse des IFs capillaires comme biomarqueurs est envisagée, en complément des analyses de l'alimentation. En effet, les cheveux accumulent d'autres substances stéroïdiennes de structure chimique similaire à celles des IFs et pour lesquelles des méthodes de dosage existent.

2- Exploration des effets perturbateurs endocriniens des IFs chez l'humain : il s'agit d'examiner, par le biais d'une approche combinant les aspects clinique et mécanistique, si une corrélation peut être établie entre l'exposition aux IFs (et autres phytoE) et les poussées de LED, d'une part, et, d'autre part, d'étudier l'impact des IFs et de leurs métabolites conjugués sur la prolifération de modèles cellulaires TNBC et le rôle du récepteur GPER. En effet, l'étude des interactions des PEs avec le GPER en est à ses balbutiements et les travaux de cette thèse visent aussi à attirer l'attention sur ce récepteur encore trop négligé par la communauté scientifique. L'hypothèse sous-jacente est que l'exposition aux phytoE pourrait potentiellement impacter le développement de maladies hormono-dépendantes telles que le LED et les TNBC.

3- Développement de stratégies pour réduire l'exposition aux IFs, dans la continuité des travaux de l'équipe de recherche : on se propose de contribuer à la mise en œuvre de techniques de production de produits à base de soja à faible teneur en IFs, à l'échelle pilote, pour une transposition faisable et efficace à l'échelle industrielle, en s'appuyant sur une meilleure compréhension des traitements traditionnels. L'hypothèse qui soutient cette ligne de recherche présume que les méthodes traditionnelles (trempage, fermentation, germination) de préparation du soja réduisent significativement les IFs des aliments.

Cette thèse s'inscrit dans un domaine de recherche d'actualité en santé publique, à savoir les effets des PEs sur la santé humaine. Elle est la première à étudier l'exposition aux IFs en France de manière exhaustive ainsi que le lien potentiel entre l'exposition aux IFs et les poussées de LED. Elle ambitionne également d'élucider le rôle du récepteur GPER dans les effets des IFs sur les cellules TNBC.

Les résultats de cette recherche devraient contribuer à une meilleure compréhension des risques liés aux IFs et à la mise en place de mesures de santé publique pour protéger la population des effets potentiellement néfastes de ces composés et au développement de produits à base de soja à faible teneur en IFs.

Le tapuscript est organisé en 3 chapitres (Etat de l'art et problématique ; Méthodologie, matériels et méthodes ; Résultats et discussions) encadrés par une introduction générale et une discussion générale, précédant une conclusion générale qui clôture le document. Le document est complété par 5 annexes.

II. Etat de l'art et problématique

II.1. Perturbateurs endocriniens

II.1.1. Généralités

Les perturbateurs endocriniens (PEs) sont des molécules exogènes capables d'interagir avec les systèmes endocriniens des animaux, entraînant ainsi des effets délétères pour l'individus et/ou ses descendant·es et les populations. Ces substances, essentiellement d'origine anthropique, sont présentes dans notre environnement quotidien et incluent notamment des polluants, des médicaments, des produits alimentaires et cosmétiques. Les perturbations induites par les PEs peuvent se manifester à différents niveaux : via l'hormone elle-même (synthèse, métabolisme, élimination, etc.), via l'expression de son récepteur ou en agissant sur sa cascade d'effecteurs **[9]**. Selon le rapport OMS-PNUE-2012 [10], près de 800 substances ont des effets de perturbation endocrine avérés ou suspectés, avec une production annuelle mondiale dépassant une tonne.

Santé Publique France, classe les PEs selon la liste suivante :

- pesticides organochlorés, fongicides, herbicides ;
- plastifiants (phtalates, Bisphénol A), retardateurs de flamme, revêtements (PFAs) ;
- médicaments comme le Distilbène (utilisé en prévention des fausses couches de 1940 à 1977), anti-douleurs (paracétamol, AINS, aspirine), antidépresseurs (Fluoxétine) ;
- produits issus de combustions incomplètes provenant des pots d'échappement, d'incinérateurs, de l'industrie métallurgique et sidérurgique et à la pratique de l'écobuage des végétaux (dioxines, furanes, PCB) ;
- produits d'hygiène (Triclosan) et cosmétiques (parabènes) ;
- phyto-estrogènes (phytoE).

L'expression « perturbateurs endocriniens » a été inventée en juillet 1991 lors de la conférence de Wingspread, aux Etats-Unis d'Amérique, qui a réuni 21 scientifiques de 15 disciplines différentes, à l'initiative de Theo Colborn [11], spécialiste en santé environnementale. Son équipe a publié la première liste de PEs cette même année, incluant les produits dérivés du soja ainsi que les produits alimentaires pour animaux de laboratoire et de compagnie. Cependant, l'attention des chercheur·es s'est portée sur les PEs dès les années 1950, lorsque les pygargues à têtes blanches, aigles emblématiques des Etats-Unis, ont commencé à disparaitre progressivement en raison d'un empoisonnement par les PCBs polluant les grands lacs d'Amérique du Nord. L'affaire du distilbène dans les années 1970, touchant les femmes enceintes et leurs bébés, a ensuite propulsé le sujet sur la scène scientifique et médiatique.

Une des particularités des PEs réside dans la diversité des formes dose-réponse. Dans certains cas, leurs effets peuvent suivre une courbe en cloche, impliquant des effets néfastes à doses médianes, puis diminuant. Dans d'autres cas, les effets des PEs présentent un profil d'activité en U, se manifestant à très faibles doses et à fortes doses, mais sans effet notable à doses médianes. Ceci est décrit comme une courbe non monotone (évolution non continue en fonction de la dose) [12]. Ces effets peuvent être différés dans le temps et se transmettre sur plusieurs générations [13]. Les mécanismes de toxicité non « classiques » des PEs font également l'objet de nombreuses discussions.

Les effets délétères des PEs peuvent intervenir sur toutes les fonctions physiologiques impliquant une signalisation hormonale, notamment le métabolisme, la reproduction, la croissance, la fonction thyroïdienne, l'induction et la progression des cancers hormonaux et le neurodéveloppement. Etant donné que les effets des PEs concernent en grande partie les axes reproducteurs, aussi bien masculin que féminin, il a d'abord été supposé qu'ils étaient médiés par les récepteurs des stéroïdes sexuels, à savoir les récepteurs des estrogènes (ER) et des androgènes (AR), qui agissent comme facteurs de transcription nucléaires. Mais il est rapidement apparu que d'autres récepteurs devaient également être impliqués dans l'action de PEs. Par exemple, le bisphénol A, qui présente une affinité bien plus faible pour ERα que l'E_2, a été associé à un autre récepteur nucléaire orphelin, ERRγ (ou Estradiol Related Receptor), qui ne se lie pas à l'E_2. Plusieurs études suggèrent l'implication d'un autre récepteur, le récepteur aux estrogènes couplé à la protéine G (GPER), dans les effets de nombreux PEs [9].

II.1.1.1. Effet cocktail

L'effet cocktail des PEs est un phénomène caractérisé par l'augmentation potentielle de la nocivité lors du mélange de plusieurs de ces substances. En effet, des chercheur·es ont découvert que certains PEs, bien que faiblement actifs individuellement, peuvent *in vitro* se fixer simultanément au même récepteur cellulaire et l'activer de manière synergique. C'est le cas par exemple de l'éthynylestradiol (principe actif de certaines pilules contraceptives) et du trans-nonachlor (un pesticide organochloré) qui peuvent se fixer ensemble au *Peroxisome X Receptor* (PXR), à des doses bien plus faibles que chacune de ces molécules individuellement. Ainsi les xénobiotiques provoquent la dérégulation des diverses fonctions physiologiques contrôlées par PXR. En effet, les deux substances semblent se stabiliser mutuellement dans la poche de liaison du récepteur et le ligand supramoléculaire ainsi constitué induit la transcription des gènes [14].

Il est important de noter que l'effet cocktail étant un sujet complexe, la toxicité d'un mélange de PEs peut dépendre de nombreux facteurs, en sus des concentrations spécifiques des différentes substances, comme leur biodisponibilité dans les fluides biologiques, leur bioaccumulation dans les tissus notamment adipeux et la manière dont elles interagissent entre elles [15].

L'étude des effets néfastes potentiels des PEs chez l'humain pose des défis analytiques et éthiques majeurs. En effet, il est impossible de soumettre intentionnellement des individus à des substances potentiellement nocives dans le cadre d'études d'intervention. En conséquence, les connaissances actuelles sur les PEs proviennent de trois sources principales : études de populations observationnelles, études animales expérimentales, et travaux *in vitro* sur des cellules ou des tissus biologiques. Toutefois, elles présentent des biais qui peuvent potentiellement entacher d'incertitude les résultats obtenus comme décrit ci-dessous.

- Études de populations : elles présentent des difficultés à établir des liens causaux directs et sont soumises à l'influence de facteurs confondants, etc. Aussi, des corrélations établies entre la présence d'une ou de plusieurs substances avec des effets physio-pathologiques peuvent omettre des composés importants et notamment des phytoE.
- Études animales : les modèles présentent des différences physiologiques et métaboliques avec les humains. Ils sont soumis à des expositions chroniques sur plusieurs générations qui n'existent pas forcément chez l'humain et à des doses pharmacologiques. Tous ces éléments peuvent conduire à une extrapolation incertaine des résultats.
- Travaux *in vitro* : les systèmes biologiques sont excessivement simplifiés et sans prise en compte de l'environnement global des cellules. Ils permettent l'utilisation de doses aberrantes au plan physiologique qui peuvent atteindre les limites de solubilité.

II.1.1.2. Période de vulnérabilité

Les effets des PEs sur la santé varient vraisemblablement selon l'âge et l'état de santé des individus exposés. La période de vulnérabilité ou fenêtre de susceptibilité fait référence à des moments spécifiques de la vie où l'exposition aux PEs peut avoir des effets particulièrement néfastes.

Dans l'organisme adulte, l'effet d'une hormone ou d'un PE est proportionnel à sa présence, diminuant progressivement avec son élimination. Cependant, pendant les phases de développement (*in utero*, petite enfance et première enfance) l'exposition à des hormones ou à des PEs peut entraîner des séquelles durables si elle a lieu alors que les mécanismes de détoxification sont immatures ou bien lors du développement d'un tissu spécifique. Il faut noter que des préjudices, notamment ceux résultant d'interactions épigénétiques, peuvent rester latents et ne se manifester que plusieurs décennies plus tard. C'est un phénomène connu sous le nom de programmation développementale. Étant donné que certains tissus, comme le système nerveux central et reproductif, continuent de se développer bien après la naissance, leur période de sensibilité peut s'étendre, parfois pendant plusieurs mois après la naissance. Pour cette raison, la période des « 1000 jours », qui s'étend du développement prénatal à la petite enfance, est particulièrement vulnérable aux effets des PEs et doit faire l'objet d'une attention particulière [16]. Les femmes enceintes seraient, elles-aussi, plus sensibles à l'action des PEs.

II.1.2. Phyto-estrogènes

Les êtres humains peuvent être exposés à des substances estrogéniques, qu'elles soient d'origine naturelle ou synthétique. Cette dernière catégorie englobe diverses substances telles que les produits pharmaceutiques, les pesticides (comme le dichlorodiphényltrichloroéthane (DDT), l'atrazine, l'alachlore ou les pyréthrines et l'hexachlorobenzène), les conservateurs, les plastifiants (comme le bisphénol A ou les phtalates) ainsi que les biphényles polychlorés. Les estrogènes naturels quant à eux se répartissent en trois groupes : les estrogènes endogènes, les phytoE et les myco-estrogènes (produits par des champignons).

Les phytoE sont, comme leur nom l'indique, des molécules aux propriétés estrogéniques présentes naturellement dans certains végétaux. Il s'agit de métabolites secondaires produits par les plantes vasculaires. Elles sont divisées en deux familles en fonction de leur structure chimique : les flavonoïdes et les non-flavonoïdes. Les structures des flavonoïdes partagent une caractéristique commune : deux cycles

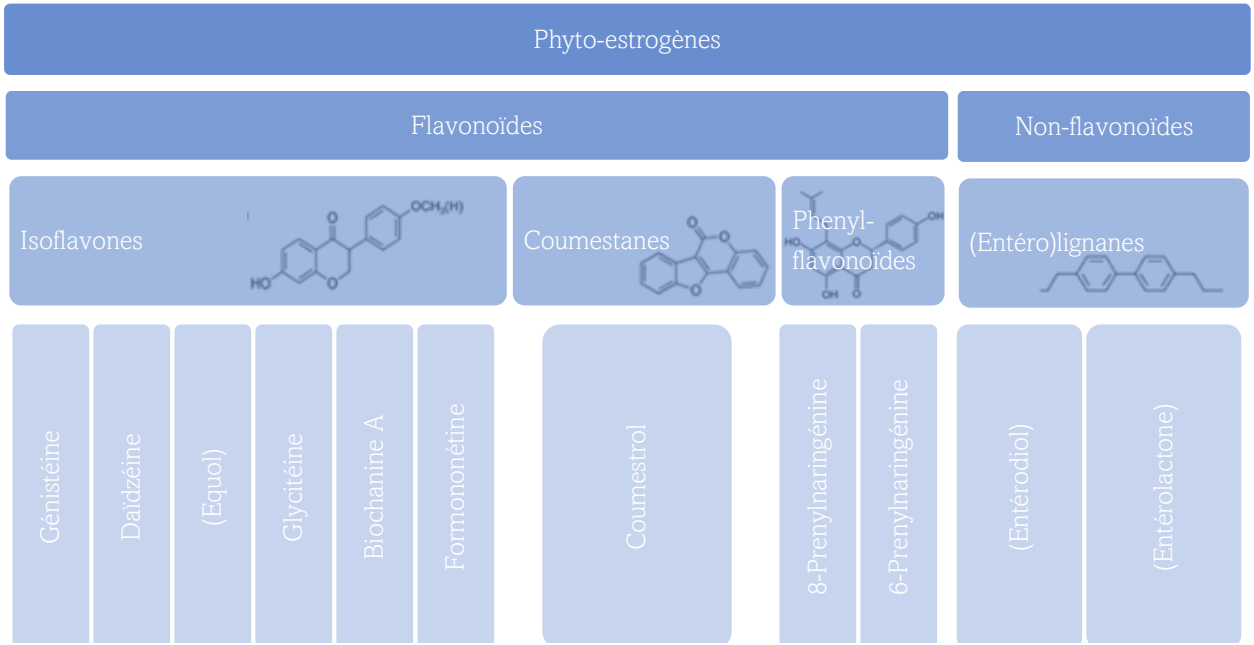


Figure 1. Liste des phyto-estrogènes principaux

aromatiques reliés par trois atomes de carbone en chaine, qui forment souvent un hétérocycle oxygéné hexa ou pentagonal [17]. Il est signalé que les molécules produites par les bactéries de la flore intestinale après transformation d'une isoflavone ou d'un lignane sont aussi considéré ici comme des phytoE. Elles sont notées entre parenthèses dans la Figure 1.

Les propriétés estrogéniques des phytoE sont dues à une similarité structurale avec le 17β E_2 représenté dans la Figure 2.

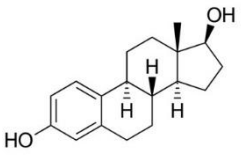


Figure 2. Structure du 17 β estradiol

La littérature mentionne également des propriétés anti-estrogéniques, anti-androgéniques, anti-FSH, anti-LH et anti-thyroïdiennes associées à ces molécules [18–20].

Les phytoE jouent un rôle crucial dans la physiologie végétale. Parmi elles, les isoflavones (IFs) sont des acteurs importants de la symbiose entre les légumineuses et les micro-organismes, en particulier dans la formation de nodules racinaires. Ainsi, les légumineuses s'associent à des bactéries fixatrices d'azote (rhizobium), elles même liées à des champignons mycorhiziens, pour former des nodules racinaires. Cette interaction mutuellement bénéfique repose sur un échange de ressources essentiel pour la croissance des deux partenaires. D'un côté, la rhizobie fournit des formes d'azote réduites à la plante hôte, tandis que la plante, en retour, met à profit les produits de la photosynthèse pour nourrir les bactéries [21,22]. Cette régulation spécifique est orchestrée par l'influence de l'expression des gènes nod bactériens, qui se lient à un seul facteur de transcription commun. Par exemple, la daidzéine (DAI) et la génistéine (GEN) agissent comme des inducteurs de gènes nod chez *Bradyrhizobium japonicum* (le principal symbionte du soja *Glycine max* (L.) Merr.) [22] mais inhibent son expression chez *Sinorhizobium meliloti* [21]. La fixation de l'azote atmosphérique par les légumineuses leur permet de synthétiser proportionnellement plus de protéines que les autres végétaux.

Les IFs sont également connus pour jouer un rôle majeur dans les défenses naturelles des plantes contre la prédation par les insectes et les herbivores. En effet, ils agissent probablement en modifiant l'appétence de la plante, en réduisant sa valeur nutritive ou sa digestibilité et en agissant comme agents toxiques [21]. Par exemple, DAI et GEN ont la capacité de repousser certains nuisibles comme les nématodes [21].

Enfin, les IFs augmentent la disponibilité des éléments nutritionnels dans le sol. Par exemple, les racines de *Medicago Sativa L.* excrètent les IFs qui augmentent la disponibilité du fer et du phosphate [21].

Cependant, il est important de noter que la présence des IFs dans les plantes est souvent sous forme conjuguée, principalement glycosidique, et que leurs niveaux peuvent varier en fonction de divers facteurs tels que le moment de la récolte, les conditions climatiques, l'environnement et les pratiques culturales et de stockage [23,24].

Les IFs estrogéniques se trouvent essentiellement dans des légumineuses comme la luzerne, le trèfle et le soja. La source principale d'IFs dans l'alimentation humaine, notamment en France, reste le soja et ses produits dérivés [25]. Le coumestrol un autre composé phytoE, se retrouve principalement dans les germes de trèfle et de luzerne, les pois cassés secs, les haricots mungo et d'autres légumineuses. Les doses sont

faibles dans les légumineuses utilisées en alimentation humaine et ne sont pas préoccupantes [26]. Les lignanes, et plus particulièrement le secoisolariciresinol (SECO), principal précurseur de l'entérolactone (ENL), se trouvent dans de nombreuses sources alimentaires telles que fruits, légumes, noix, céréales et boissons. Outre le SECO, le matairésinol, le laricirésinol, le pinorésinol, le syringarésinol, le médiorésinol et certains de leurs métabolites et glucosides hydroxylés peuvent être métabolisés en ENL et en entérodiol [27]. Il est important de noter que l'activité estrogénique de l'ENL est faible [28] et que son taux de conversion varie en fonction de la microflore individuelle. Par conséquent, l'exposition aux précurseurs de l'ENL ne conduit pas nécessairement à un effet estrogénique [29].

II.2. Isoflavones

Les IFs représentent la classe principale des phytoE flavonoïdes du fait de leur présence et de leur activité dans l'environnement humain. On les trouve dans de nombreuses plantes, essentiellement dans la famille des Fabacées ou légumineuses. Toutes les IFs ne sont pas estrogéniques mais c'est dans cette famille que se trouvent les composés estrogéniques les plus préoccupants.

GEN et DAI sont les principales IFs du soja. La glycitéine (Gly) est quant à elle, présente dans de plus faibles proportions [23,30] et son pouvoir estrogénique est moindre.

Un grand nombre d'études cliniques ont été conduites pour évaluer les impacts de ces molécules sur la santé humaine, cependant, les résultats obtenus ne sont pas toujours concluants. Cette incertitude pourrait être attribuée à la variabilité interindividuelle concernant la biodisponibilité et le métabolisme des IFs. En effet, le métabolisme modifie la structure chimique d'un composé parent et peut ainsi modifier considérablement son activité biologique. Par ailleurs, il a été montré que les taux circulants de GEN et DAI pouvaient varier du simple au décuple dans les fluides biologiques de femmes ménopausées soumises à un apport d'IFs contrôlé [31].

II.2.1. ADME (Absorption-Distribution-Métabolisme-Elimination)

Pour être actif et exercer ses effets physiologiques, un xénobiotique doit atteindre le compartiment des fluides biologiques et donc être absorbé par l'organisme.

Absorption

L'absorption intestinale se présente comme le mécanisme prédominant d'assimilation des IFs suite à leur ingestion, que ce soit à travers la consommation d'aliments ou de compléments alimentaires.

Dans le cas général des polyphénols, cette absorption dépend de leur structure, notamment de la proportion moléculaire conjuguée / non conjuguée. Il a donc été considéré qu'en raison de leurs propriétés amphiphiles [32] et de leur poids moléculaire plus faible [33], les IFs aglycones devraient être absorbées plus facilement et plus rapidement que leurs homologues conjugués [34]. Cette hypothèse a été vérifiée sur des modèles cellulaires mais ne correspond pas à la réalité physiologique pour laquelle la biodisponibilité des IFs est identique [35] voire meilleure après ingestion de glycosides, par rapport aux aglycones [36,37]. Ce fait pourrait s'expliquer par le fait que les IFs sont prises en charge de façon quantitative par la Lactase-Phloridzine Hydrolase (LPH) présente sur la bordure en brosse des entérocytes, et ce, tout le long du tube digestif. La LPH joue un rôle crucial dans l'hydrolyse des

glycosides, permettant la libération des aglycones et leur absorption subséquente [38]. Ainsi, il est établi que les glycosides de DAI et GEN ingérés en tant que composés purs ou provenant d'aliments complexes ne sont pas retrouvés sous cette forme dans le plasma ou l'urine [23]. De plus, les aglycones se trouvent dans de très faible proportion dans les différentes sources alimentaires d'IFs, à l'exception des aliments de soja fermentés [36].

Plusieurs facteurs influencent l'absorption des glycosides d'IFs, notamment ceux présentés ci-dessous.

- La structure du glycoside : la position et le type de sucre conjugué affectent l'hydrolyse par la LPH.
- Le microbiote intestinal : certaines bactéries intestinales peuvent également produire des enzymes β-glucuronidases, contribuant à l'hydrolyse des glycosides.
- L'activité de la LPH : elle peut varier entre les individus et être influencée par des facteurs génétiques, l'âge et la consommation d'aliments riches en polyphénols.

Métabolisme et distribution

Une fois passées dans le compartiment entérocytaire, les IFs sont en partie prises en charges par des enzymes de phase II capables de les conjuguer à des résidus glucuronides ou des sulfates [13]. Ce mécanisme a également été montré chez le rat [16]. C'est un mélange d'IFs conjugués et aglycones qui atteint ensuite le foie via la circulation porte entéro-hépatique. Au niveau de cet organe, des enzymes de phase I à cytochrome P 450 sont capables d'hydroxyler les IFs et donc de transformer les IFs métoxylées en position 4', comme la Biochanine A et la Formononétine du trèfle, respectivement en GEN et DAI [39]. Ces dernières molécules sont plus estrogéniques car elles possèdent comme l'E_2 deux groupements hydroxyles en positions opposées sur l'édifice moléculaire à une distance de 10 Ångströms. Dans le sang les IFs circulent donc essentiellement sous forme de GEN et DAI glucuronidées et sulfatées. Seuls 1 à 20% de ces composés se trouvent en réalité sous forme aglycone, mais cette proportion peut varier selon le composé considéré [40,41]. Il est important de souligner que cette observation s'applique également aux hormones stéroïdes, en particulier à l'estrogène endogène E_2. En effet, moins de 5% de cette hormone circule sous forme libre [42]. En plus du mécanisme qui se produit continuellement dans l'ensemble du tube digestif, des transformations sont également effectuées par la flore intestinale. Ainsi, selon leurs équipements bactériens, certains consommateur·ices sont capables de transformer la DAI en equol. Ce métabolite est une isoflavane possédant un carbone asymétrique et semble essentiellement produit sous sa forme S. Selon les études, la proportion de producteurs d'équol dans la population humaine varie de 12 à 60%. Cette classification fait l'objet de discussions parmi les experts car elle dépend non seulement de la sensibilité des techniques d'analyse mais aussi de la plausibilité d'activité des concentrations d'équol mesurées [23,43]. L'équipe du Pr. Bennetau qui utilise depuis plus de 20 ans la technique d'immunoadsorption par enzyme liée ou *Enzyme Linked Immunosorbant Essay* (ELISA) pour mesurer les IFs et l'équol a, à plusieurs reprises, évalué à environ 60% la proportion de producteurs d'équol dans des cohortes de sujets Français.

En outre, selon Setchell et ses collaborateur·ices [39], la bio-transformation microbienne de DAI peut aboutir à l'O-desméthylangolensine, un produit aux propriétés estrogéniques réduites. De manière générale, GEN et DAI peuvent être

retrouvées dans le plasma des consommateur·ices Français·es à des doses allant jusqu'à 800 ng/mL soit environ 1,5 µM. Leur Tmax se situe entre 8 et 9 heures [23,43]. Parfois, un premier pic d'absorption est observé à 5 heures dans le sang, correspondant à une absorption duodénale [41]. Celle-ci peut être majeure si la quantité d'IFs ingérée est faible.

Le coefficient de répartition tissu/sérum pour l'aglycone de GEN s'est avéré plus élevé que celui pour la GEN totale (GEN et ses conjugués), confirmant des expositions tissulaires significatives à l'aglycone, forme estrogénique active prédominante [13]. La facilité avec laquelle l'aglycone traverse les membranes cellulaires peut être attribuée à sa liposolubilité.

Le métabolisme microbien n'est pas le seul à influer sur les formes circulantes et le potentiel estrogéniques des IFs. En effet, la glucurono- et sulfo-conjugaison de ces molécules n'aboutit pas toujours à des métabolites inactifs. Dans le cas de la DAI, par exemple, le conjugué 7-O-sulfate possède une activité beaucoup plus élevée que celle du composé parent lui-même et il en va de même pour les sulfoglucuronides [41]. Les glucuronides sont les métabolites circulants principaux pour GEN et DAI [39,44]. Selon Soukup *et al.* et Islam *et al.* [41,44], dans le plasma comme dans l'urine, moins de 5 % de ces espèces se trouvent sous forme d'aglycones.

La moitié de l'équol (49,7%) circule sous forme libre contrairement aux autres IFs[40].

GEN et DAI ont des volumes de distribution faibles [45] et sont rapidement éliminées de l'organisme. En conséquence, seule une consommation récurrente de plus de 30 à 40 mg d'équivalent aglycone/jour pour les adultes devrait pouvoir exercer un effet endocrinien significatif. En effet, la consommation quotidienne de deux ou plus de produits à base de soja, qui expose à ce type de doses, permet d'atteindre un plateau pharmacocinétique. Ce phénomène est particulièrement notable dans le cas des compléments alimentaires.

Il est important de signaler que la biodisponibilité des phytoE et d'autres PEs par rapport au 17β-E_2 peut être significativement plus élevée en raison de leur faible capacité à interagir avec la globuline liant les hormones sexuelles (SHBG). Les concentrations sériques postprandiales de DAI, par exemple, peuvent dépasser de 100 à 1000 fois les concentrations sériques préovulatoires de 17β-E_2 ce qui pourrait compenser la plus faible affinité de cette IF pour les ERs [42,46].

Elimination

La demi-vie d'élimination des IFs varie entre 12 et 30 heures, en fonction du composé considéré et des individus. En effet des phénomènes de recirculation peuvent s'observer chez certaines personnes entre le foie, la vésicule biliaire et le duodénum [43]. L'élimination est majoritairement urinaire.

II.2.2. Toxicologie

La première mention d'effets délétères des IFs remonte à 1946 quand Bennets *et al.* ont rapporté des problèmes utérins et d'infertilité chez des brebis broutant des pâturages riches en trèfle rouge ou en trèfle souterrain [47].

Le *National Toxicology Program* (NTP) des États-Unis d'Amérique a démontré dans deux études toxicologiques multigénérationnelles que l'exposition à la GEN de rats Wistar provoquait des tumeurs mammaires et de l'hypophyse chez les femelles, ainsi

que des troubles de la reproduction chez les animaux des deux sexes [13,48]. La première est une étude sur six générations de rats (F0 à F5) menée pour évaluer les effets de la GEN sur la reproduction et le développement. Les rats ont été exposés à 0, 1, 5, 100 ou 500 ppm de GEN par voie alimentaire à l'exception des générations F4 et F5 non exposées ; 500 ppm de GEN équivalant à 35 mg/kg et 51 mg/kg de poids corporel chez les mâles et les femelles respectivement. Les résultats les plus significatifs de cette étude sont exposés ci-dessous.

- Une diminution du poids à la naissance n'a pas été systématiquement observée, par contre une réduction de la prise de poids pendant l'allaitement a été notée sur plusieurs générations à partir de 500 ppm de GEN.
- Le poids au sevrage était significativement plus bas chez les ratons mâles et femelles de la génération F1 à F4 à 500 ppm, et à 5 et 100 ppm pour les F1.
- La taille des portées a tendance à diminuer avec l'augmentation de la dose de GEN, avec une différence significative à 500 ppm en F2.
- Une diminution de la distance ano-génitale a été observée chez les mâles et femelles F1 (500 ppm), et uniquement chez les femelles F2 (500 ppm) et F3 (100 ppm).
- L'ouverture vaginale a été observée plus tôt chez les femelles exposées à 500 ppm (3 jours d'avance en F1 et F2 ; 1,3 jour en F3) et à 5 ppm (F3).
- Le poids à l'ouverture vaginale était plus faible à 500 ppm (F1, F2, F3) et à 5 ppm (F1).
- Des perturbations des cycles œstraux ont été observées à partir de 500 ppm (F1, F2, F3), à 100 ppm (F4) et 5 ppm (F2).
- Un retard de la descente testiculaire a été observé chez les fœtus mâles de F3 (500 ppm).
- Des hyperplasies des glandes mammaires et du rein ont été observées chez les mâles adultes de F0 à F3 (500 ppm).
- Des tubules rénaux calcifiés ont été observés chez les mâles F1 et F2 (100 et 500 ppm).

En résumé, l'exposition à la GEN à des doses de 100 et 500 ppm a des effets néfastes sur la reproduction et le développement des rats. Aucun effet significatif n'a été observé à des doses de 1 à 100 ppm. Certains de ces effets néfastes ont été transmis aux générations non exposées. Ils sont similaires à ceux observés avec l'E_2 et surviennent à des doses comparables à une consommation humaine modérée de produits de soja industriels (deux portions par jour) [13].

Cette exposition alimentaire quotidienne efficace de 35 mg/kg de poids corporel, permet de définir les niveaux toxiques préliminaires chez les animaux. La toxicologie classique admet que pour transposer une dose toxique minimale d'une substance, de l'animal à l'être humain, il faut la diviser par le produit de divers facteurs de sécurité allant chacun de 1 à 10 [49]. Dans le cas de la GEN, trois facteurs de sécurité doivent être pris en compte. Le premier est un facteur d'extrapolation d'un LOAEL (ou DMENO : niveau le plus bas d'effet nocif observable, qui est disponible pour la GEN) à un NOAEL (ou DMSENO : dose minimale sans effet nocif observable). Le deuxième est un facteur d'extrapolation entre les animaux et les humains. Le troisième est un facteur tenant compte des différences de sensibilité entre les êtres humains. Si l'on applique les valeurs communément admises, on obtient un total de 300, même si

Kodell et Galore [50] ont avancé des arguments statistiques en faveur d'un facteur moins élevé de 184 mais néanmoins suffisamment sûr.

Comme le soja contient également de la DAI qui peut être métabolisé en équol, potentiellement plus actif, il semble plus prudent de considérer la toxicité des IFs totales, *i.e.* GEN et DAI. Cependant, les données toxicologiques fiables sur la DAI sont limitées. Pour pallier à cette carence, il semble prudent d'appliquer le facteur de sécurité le plus élevé aux mesures des IFs totales, c'est-à-dire 300. Mais comme mentionné précédemment, un facteur de 184 pourrait être suffisant. Ainsi, la limite d'IFs pour un adulte de 60 kg serait de 11,4 mg/jour, et pour un enfant de 20 kg de 5,7 mg/jour. L'Anses devrait prochainement proposer une valeur toxique de référence basée sur les travaux des expert·es.

Une étude toxicologique sur des souris [51] a proposé une NOAEL de la DAI supérieure à 5 000 mg/kg, après un traitement de 28 jours. Aucun changement dans les paramètres hématologiques, la biochimie clinique et les paramètres de la fonction rénale n'a été rapportés. Mais les résultats de cette étude doivent être interprétés avec prudence en raison de ses limites méthodologiques. Elle a en effet été menée avec un nombre restreint d'animaux, seulement 3 souris par groupe (soit 12 au total). Ce nombre est insuffisant pour garantir la fiabilité des résultats. Aussi, la durée d'observation de 14 jours seulement ne permet pas d'évaluer la toxicité à long terme. Par ailleurs, il est connu qu'à la dose administrée sur plusieurs cycles chez la souris, les paramètres de la reproduction sont altérés. Cependant, ces derniers n'ont pas été examinés.

II.2.3. Mécanismes d'action

Bien que l'activité estrogénique des IFs soit 10^2 à 10^4 fois plus faible que celle de l'E_2 endogène *in vitro*, par l'intermédiaire des récepteurs nucléaires canoniques de l'E_2 (ERα, ERβ), elle reste significative. La raison est probablement liée aux concentrations circulantes plus élevées d'IFs [30]. Ces dernières ont à la fois des activités estrogéniques et anti-estrogéniques en fonction de l'équilibre entre leur concentration et celle de l'estrogène endogène [32,52]. L'hypothèse admise est qu'ils ont une activité anti-estrogénique dans un environnement à forte teneur en estrogènes, mais une activité estrogénique dans le cas inverse. Le premier cas peut s'expliquer par une compétition avec l'E_2 pour la liaison aux récepteurs estrogéniques [30] et donc, en raison de leur activité plus faible, l'effet global est anti-estrogénique. Les IFs se lient aux ERs avec une affinité moindre que celle de l'E_2 mais plus élevée pour les ERβ que pour les ERα [18]. Cependant, si les concentrations physiologiques d'E_2 sont finement ajustées, les concentrations d'IFs varient considérablement après une seule ingestion. Par conséquent, un effet anti-estrogènique dû à la compétition au niveau des récepteurs ne peut être observé qu'en cas d'ingestion d'IFs suffisante et régulière.

La capacité de la GEN à inhiber la croissance a été étudiée dans différentes lignées cellulaires, cancéreuses et non cancéreuses, et ce en fonction de la dose. Il a été observé que l'activation préférentielle de ERβ par la GEN est perdue à des concentrations élevées (300 nM). Par conséquent, l'effet final de la GEN sur l'expression des gènes et le destin cellulaire dépend de la dose administrée et de la capacité différentielle des complexes ligand-ER à recruter des modulateurs sur les sites de liaison à ER des promoteurs des gènes régulés par les hormones [53].

A des doses pharmacologiques (supérieures à 10 µM), l'activité antiproliférative de la GEN pourrait être médiée par l'inhibition des tyrosine kinases (PTK) [54].

ERα et ERβ interagissent avec le même élément de réponse aux estrogènes (ERE) sur l'ADN, sous la forme d'homodimères ou d'hétérodimères α/β. L'homodimère ERα est plus efficace que les homodimères ERβ ou l'hétérodimère ERα/β pour promouvoir la transcription des gènes contrôlant la croissance et la différenciation de l'utérus et du sein. En revanche, ERβ, lorsqu'il est co-exprimé avec ERα, peut contrecarrer ses effets prolifératifs sur les cellules cancéreuses du sein [54].

Les IFs se lient également au récepteur membranaire de l'E_2, le GPER, avec une affinité seulement 10 fois inférieure à celle de l'E_2, soit 0,01 µM. Il a été démontré que le GPER est présent dans de nombreux tissus impliqués dans la physiologie de la reproduction, tels que ceux du vagin, de l'utérus, de la glande mammaire, du cerveau ou des testicules. Les IFs peuvent induire la transcription des gènes via ces récepteurs ainsi que la prolifération des cellules cancéreuses portant ces récepteurs [55].

D'autres mécanismes d'action des IFs, basés sur des études *in vitro*, pourraient être l'inhibition des tyrosine kinases, la liaison et l'activation des régulateurs des proliférateurs de peroxysomes, l'activité antioxydante, l'inhibition des enzymes de biosynthèse de certains stéroïdes, l'influence sur les cellules tueuses naturelles et l'inhibition des métastases [18]. Cependant, tous ces résultats ont été enregistrés *in vitro* seulement et en utilisant des doses d'IFs 50 à 100 fois supérieures à celles possiblement présentes dans le plasma ou les tissus humains. De tels effets qui sont recherchés en chimiothérapie peuvent éventuellement être impliqués dans la protection contre le cancer, mais ne sont pas observés à des concentrations physiologiques. Les IFs ont également été signalés comme étant des inhibiteurs compétitifs de la peroxydase thyroïdienne (IC_{50} = 1 µM, ce qui est physiologique).

Les effets anti-estrogéniques des IFs font quant à eux encore l'objet de controverses. D'un côté, certains arguments semblent en faveur d'effets anti-proliférateurs comme ceux qui suivent.

Des études *in vitro* ont montré que des doses élevées d'IFs (≥40 µM pour équol, > 5 µM pour GEN et > 40 µM pour DAI) peuvent inhiber la croissance des cellules cancéreuses mammaires et déclencher l'apoptose [56–58]. Cependant, ces effets pourraient être obtenus par l'injection locale ou l'exposition à de très fortes doses d'IFs, mais pas par la consommation de soja.

D'un autre côté, les effets anti-estrogéniques des IFs dépendraient de la dose, du type de cellule et du contexte biologique et restent controversés *in vivo*. Ainsi, les effets des estrogènes sur la croissance cellulaire ne sont pas de purs effets estrogéniques [56]. La différence d'affinité entre l'E_2 et les IFs pour les ERs pourrait expliquer les effets anti-estrogéniques. Mais la compétition entre l'E_2 endogène et les IFs n'est possible que lorsque les concentrations respectives sont dans des proportions spécifiques, ce qui peut être maintenu *in vitro* mais n'est que transitoire *in vivo* [2].

II.2.4. Effets biologiques

Le tableau ci-après résume les études cliniques principales concernant le lien entre IFs et différentes pathologies. Les études cliniques, études de cas, méta-analyses et études de cohorte sont priorisées dans cette partie du tapuscrit.

Etude	# Cas	Sexe, ethnicité, âge, statut	Conception de l'étude	Quantité d'IFs	Forme d'IFs	Durée	Facteur étudié	Résultats	Limites
Prospective [6]	76442	Femmes Françaises ménoP[1]	QA[2] : consommation de compléments de soja, tous les 2-3 ans	3,7 – 37 mg/ cp[3]	Glucoside dans extrait de soja	11 ans	BC[4] (risque)	Supplémentation en IFs: plus petit risque (ER+) – plus grand risque (ER-). Précaution pour les femmes avec ATCD[5] familiaux de BC	Auto-déclaration ; pas d'apport précis en IFs ; pas de données sur la consommation parallèle de soja;
Randomisée vs placebo [59].	140	Femme, BC précoce (Caucasiennes en majorité)	IFs sériques, changement d'expression des gènes (tissue tumoral avant et après-traitement)	51,6 g Pst[6]/j	NS[7]	7 – 30 j	BC (expression des gènes)	Signature de l'expression génétique de la consommation de soja (GEN plasmatique élevée) : surexpression des gènes qui pilotent les voies du cycle cellulaire et de la prolifération	Étude courte; taux plasmatique évalué 2 fois; dose élevée d'IFs ; pas d'évaluation de l'impact clinique des modifications de l'expression génétique
Méta-analyse (5 cohortes) [60]	11206	Femmes après diagnostic BC (Chine et États-Unis)	Impact de la consommation de PST en fonction de l'origine ethnique, de l'état ménopausique et des doses.	2 – 13 g/j	NS	9 mois –7,3 ans	BC (survie)	La consommation d'aliments à base de soja post-diagnostic est associée à une réduction de la mortalité (HR 0,85, IC[8] à 95% 0,77 0,93) et de la récurrence (HR 0,79, IC à 95% 0,72 0,87	Pas d'évaluation qualitative des études; Statut ER manquant; pas d'informations sur le statut GPER ; pas de quantification des IFs ni d'informations sur la consommation parallèle de soja.
Méta-analyse (31 cas-témoins +2 cohortes) [61]	9222 cas 11413 contrôles	Femmes Chinoises (différentes régions)	Estimation de l'OR et de l'IC à 95 % (impact de la consommation de soja et de fruits)	NS	NS	NS	BC (risque)	La consommation de soja et de fruits était significativement associée à un risque plus faible de BC, avec un OR résumé de 0,65 (IC 95 % : 0,43-0,99 ; I2=88,9 %, P<0,001 ; N=13) et 0,66 (IC 95 % : 0,47-0,91 ; I2=76,7 %, P<0,001 ; N=7), respectivement.	Apport alimentaire non clairs; 10 études basées uniquement sur des QA ; aucune information sur le statut ER. ; confusion entre soja et IFs (pas d'information sur la source des aliments)
Prospective [62]	Sein : 244 cas sur 941 Colorectal : 221 sur 886 prostate : 204 sur 812	Européen·nes âgés de 40 à 79 ans	ATCDs médicaux et enregistrement du régime alimentaire sur 7 j. 509 aliments analysés (consommation de phytoE). OR calculé.	-	NS	7 j	Risque BC, colorectal, et prostate	Plus de BC avec SECO et lignanes (surtout en cas d'ATCDs familiaux de BC), association inverse entre cancer colorectal et ENL ; risque de cancer prostate positivement associé à une plus grande consommation d'entérolignanes.	Petit nombre de sujets et pas d'ajustement de la valeur p lors des comparaisons multiples.

[1] MénoP : ménopause
[2] QA : questionnaire alimentaire
[3] Cp : comprimé
[4] BC : cancer du sein
[5] ATCD : antécédent
[6] PST : protéines de soja texturées
[7] NS : non spécifié
[8] IC: Intervalle de confiance

Etude	# Cas	Sexe, ethnicité, âge, statut	Conception de l'étude	Quantité d'IFs	Forme d'IFs	Durée	Facteur étudié	Résultats	Limites
Étude d'intervention contrôlée [63]	15	Femmes en bonne santé, non végétariennes, préménoP	Régime alimentaire et niveau d'activité constants. Statut hormonal mesuré sur 2 ou 3 cycles menstruels	60-28 g/j Pt; 50g/j miso	C[9]: 45 ou 23mg, nonC: 25mg	9 m dont 4 à 6 contrôlés	Cycle menstruel	60 g de PST ont augmenté de manière significative ($P < 0,01$) la durée de la phase folliculaire et retardé le pic de concentration de progestérone.	Durée de l'étude ; limites de précision des méthodes de dosage HPLC en 1995.
Intervention [64]	20	Femmes préménoP en bonne santé	10 asiatiques, 10 non asiatiques	32 mg/j (aliments à base de soja)	NS	7 m (3 m intervention)	Fonction ovarienne	Le niveau d'E_2 lutéal est significativement plus bas chez les femmes asiatiques ; le SHBG sérique est statistiquement plus élevé chez les non-asiatiques mais plus bas chez les asiatiques.	Biais dans la collecte d'urine domestique ; différences entre aliments à base de soja consommés par les Asiatiques et les non-Asiatiques, nombre réduit de participants de chaque groupe ; manque de preuves de l'absence de consommation parallèle de soja.
Étude d'intervention diététique contrôlée [65]	6	Femmes préménoP en bonne santé et non végétariennes	Régime de contrôle surveillé> 1 cycle (avant et après), ajout de soja pendant un cycle complet - statut hormonal et IFs urinaires mesurés	45 mg/j (dans 60 g de PST)	DAI et GEN	4-6 m	Cycle menstruel	Phase folliculaire plus longue et/ou saignement retardé ; suppression significative des poussées de LH et de FSH en milieu de cycle ; augmentation de E_2 plasmatique en phase folliculaire.	Nombre de sujets ; pas de groupe de contrôle ; étude relativement courte.
Étude de cohorte [66]	1553	Jeunes femmes afro-américaines en bonne santé	Données du QA sur le soja pendant la petite enfance	Estimée	NS	NS	Fibromes	Aucune association entre les formules à base de soja et la prévalence des fibromes. Mais lien significatif entre formules à base de soja et diamètre des fibromes chez les femmes atteintes de fibromes.	Biais de sélection (exclusion des femmes diagnostiquées pour un fibrome) ; jeune âge des participantes (23-35 ans) (pas de détection d'un lien potentiel plus prévalent apparaissant lorsque les femmes sont plus âgées).
Randomisée, double-aveugle vs placebo [4]	376	Femmes ménoP en bonne santé dont l'utérus est intact	Consommation journalière pendant 5 ans. Biopsies endométriales au début et à la fin de l'étude.	150 mg/j	40–45% GEN, 40– 45% DAI, 10–20% glycitéine	5 ans	Histologie de l'endomètre	Aucun cas de tumeur maligne. La fréquence de l'hyperplasie endométriale était significativement plus élevée dans le groupe traité (3,37 % contre 0 %).	Pas d'information sur la consommation parallèle d'aliments à base de soja ; le dosage des comprimés est plus élevé que celui d'un complément alimentaire.

[9] C : Conjugués

Etude	# Cas	Sexe, ethnicité, âge, statut	Conception de l'étude	Quantité d'IFs	Forme d'IFs	Durée	Facteur étudié	Résultats	Limites
Etude de cas [20]	3	Femmes (35-56 ans)	Recherche de saignements utérins anormaux avec pathologie endométriale	40 g/j ; quantité excessive de soja/j	NS	3, 5 et 21 ans	Saignement utérin	La réduction/suppression du soja a fait disparaître les symptômes. Une consommation excessive de phytoE peut être nocive.	Nombre réduit de cas ; quantité excessive d'IFs. Quantité d'aliments à base de soja NS dans une étude.
Cohorte rétrospective [67]	248 (formule infantile à base de soja)	Jeunes adultes	Etat de santé général, maturation pubertaire et ATCDs menstruels et reproductifs de jeunes adultes nourris avec des formules infantiles à base de soja.	Estimée : 4,2 à 9,4 mg/kg /j	NS	6 mois	Voir conception de l'étude	Seule une légère augmentation de la durée des règles (0,37 jours ; IC à 95 % , 0,06-0,68) et une plus grande gêne (1,77 ; IC à 95 %, 1,04-3,00) ont été observées.	Auto-évaluation, étude courte, quantité de phytoE non mesurée
Cohorte [5]	99	Partenaires masculins caucasiens dans couples subfertiles	Évaluation de la consommation de soja/IFs	5,4 mg/j en moyenne	NS	3 mois	Qualité sperma-tique	Tendance statistiquement significative à la réduction de la concentration de spermatozoïdes avec l'augmentation de la consommation d'aliments à base de soja.	Auto-évaluation, étude relativement courte
Méta-analyse (15 études) [68]	-	Hommes et femmes	En bonne santé ou non Préménopause et postménopause	NS	IF, PST, DAI, GEN	<3 mois >3 mois	Thyroïde	Pas de modification significative des taux de T3 et T4. Changements significatifs des niveaux de TSH avec un traitement > 3m et une hypothyroïdie subclinique préexistante.	Nombre de sujets, courte durée de traitement pour certaines études

Tableau 1. Résumé d' études épidémiologiques évaluant l'impact des isoflavones sur la santé humaine.

II.2.4.1. Maladies thyroïdiennes

L'association entre les IFs et les pathologies thyroïdiennes est un sujet débattu. Des études épidémiologiques ont rapporté des résultats contradictoires, soulignant la nécessité d'une analyse approfondie et nuancée.

Certaines études ont mis en évidence un lien potentiel entre la consommation d'IFs et des effets néfastes sur la thyroïde, notamment :

- une étude de 2002 a révélé une association entre la consommation d'IFs de soja et une altération subclinique des hormones thyroïdiennes chez des femmes ménopausées [69] ;
- un cas rapporté en 2001 a suggéré que les IFs interférent avec l'absorption d'hormones thyroïdiennes chez les personnes souffrant d'hypothyroïdie [70].

Mais d'autres études ne trouvent aucun lien :

- une étude de 2009 n'a trouvé aucun lien entre la consommation de soja et les pathologies thyroïdiennes chez des femmes enceintes saines [71] ;
- une étude de 2015 a rapporté que la consommation d'IFs n'affectait pas la fonction thyroïdienne chez les femmes pré- et post-ménopausées saines [72].

Après examen de ces investigations, il apparait que plusieurs facteurs, décrits plus bas, peuvent influencer l'impact des IFs sur la thyroïde.

- Le type de population étudiée : la plupart de ces études ont été menées sur des populations asiatiques. Ainsi, il n'est pas exclu que les effets des IFs pourraient varier selon les populations.
- La supplémentation en iode : celle-ci semble annihiler les effets délétères des IFs. Cela pourrait expliquer que la fonction thyroïdienne ne soit pas altérée de façon marquée chez certaines populations asiatiques habituées à consommer beaucoup de produits de la mer, notamment des algues [73].
- L'état thyroïdien préexistant : les patient·es souffrant d'hypothyroïdie préexistante semblent plus sensibles aux effets néfastes des IFs.

L'étude de cohorte de Otun *et al.* [68] confirme qu'il n'y a aucun effet négatif des IFs sur la fonction thyroïdienne des sujets sains mais la prudence doit être de mise avec des patient·es souffrant d'hypothyroïdie préexistante. D'ailleurs, les compléments alimentaires et les aliments contenant du soja sont contre-indiqués chez les patient·es sous Levotyrox® (lévothyroxine sodique, analogue hormonal préconisé dans le traitement des hypothyroïdies).

En outre, il a été démontré que les formules infantiles à base de soja provoquaient des goitres réversibles à l'arrêt du produit, en particulier chez les nourrissons présentant une hypothyroïdie préexistante [74,75]. Compte tenu du caractère indispensable des hormones thyroïdiennes au développement cérébral des enfants, les formules au soja sont contre-indiquées chez les enfants présentant une hypothyroïdie congénitale, bien que les formules infantiles à base de soja aient été supplémentées en iode dès les années 1960.

II.2.4.2. Cancers hormono-dépendants

Le lien entre les IFs et les cancers hormono-dépendants, notamment le cancer du sein, de l'endomètre et de la prostate, est complexe et controversé. La littérature scientifique

présente des résultats divergents en apparence, soulignant l'importance d'une analyse approfondie et nuancée.

II.2.4.2.1. Cancer du sein

On peut constater que les études épidémiologiques montrant une réduction du risque de cancer du sein ou une amélioration de la survie, en lien avec la consommation d'IFs, ont été réalisées, à notre connaissance, uniquement sur des femmes asiatiques. Or comme suggéré par Fernandez-Lopez *et al.* [75], l'exposition aux IFs telle qu'établie dans les études de population menées en Asie doit être questionnée. Dans les autres études, aucun impact positif significatif des IFs n'a été observé. Chez les femmes ménopausées occidentales, l'effet du soja sur le cancer du sein n'est pas clair [3]. Il est important de souligner ici que la plupart des études existantes ont négligé d'inclure les IFs « cachées » dans les aliments manufacturés, réduisant ainsi la puissance statistique de leurs analyses. Selon l'AFSSA et l'AFSSAPS, la prudence est de mise lorsque les femmes ont des antécédents familiaux, personnels ou des cancers hormono-dépendants **[8]**. Il est également important de noter qu'un effet positif des IFs pourrait être simplement le marqueur d'un mode de vie plus sain, leur impact pouvant difficilement être isolé des autres facteurs. Ces différences pourraient également provenir des dissemblances génétiques entre les 2 groupes de population (asiatiques et caucasiennes).

Une autre observation importante est que de nombreuses études ne spécifient pas le statut ER différentiel des sujets. Cette donnée est significative en raison du rôle probablement opposé des 2 récepteurs à estrogènes nucléaires dans le cancer du sein : l'activation de ERα induit la prolifération tandis que celle de ERβ l'inhibe [18]. Ainsi, l'effet positif présumé des IFs en raison de leur plus grande affinité avec l'ERβ ne peut être établi de manière convaincante que si les femmes présentent une majorité de ERβ dans leurs tissus. Par ailleurs, à ce jour et à notre connaissance, aucune étude de population n'a mentionné le statut GPER des patientes suivies.

Plusieurs publications témoignent d'un double effet des IFs sur les cellules sensibles aux estrogènes, stimulant la croissance cellulaire à de faibles concentrations (0,1-10 µM pour GEN) et provoquant une inhibition à des concentrations plus élevées [76].

D'un point de vue mécanistique, les IFs ont un effet dual.

D'un côté, elles sembleraient protectrices contre les cancers du sein pendant la phase de promotion. Cet effet, basé sur une plus faible incidence des cancers du sein en Asie par comparaison à la situation occidentale, pourrait passer par des mécanismes épigénétiques, protégeant les femmes asiatiques exposées dès l'enfance à une alimentation contenant des IFs. Ainsi, il a été montré que la GEN et la DAI ont des effets méthylant inhibant l'expression des gènes lorsqu'ils sont administrés à doses alimentaires à des souris modèles agouti [77]. Ces marquages épigénétiques de méthylation ont aussi été observés sur la sphère vaginale de jeunes filles exposées quand elles étaient bébé à des IFs via des formules infantiles à base de soja [78]. Cette protection est en accord avec les résultats des investigations faites par Lamartiniere [79] sur des rats exposés au Dimethyl Benz(a)anthracen (DMBA), dans lesquelles l'effet préventif de la GEN apparaît pendant la phase de promotion. Cette prévention peut être obtenue en réduisant par méthylation l'expression précoce des pro-oncogènes, tels que MDM2. Certain·es auteur·es rapportent *in vitro* une augmentation d'onco-suppresseurs, tels que BrCa1 [2]. Pour l'instant, alors que l'expression de BrCa1 a été

observée *in vivo*, même si elle n'est pas liée aux IFs, l'action de GEN sur l'expression des gènes BrCA 1 et 2 et sur la réduction de MDM2 n'a été démontrée qu'*in vitro* et à des doses élevées. Il serait donc nécessaire de confirmer cette action à faibles doses sur des cellules normales pendant la phase d'initiation d'un cancer.

Il est évident que la réduction de l'émergence de cellules cancéreuses par n'importe quel mécanisme lors de la phase initiale entraînerait globalement une baisse de l'incidence du cancer. Mais dans les études de populations asiatiques, certain·es auteur·es ont associé la consommation de soja à celle de choux, de thé et de poissons qui apportent des glucosinolates, des anti-oxydants et des oméga 3. Ces substances naturelles sont aussi connues pour leur propriété préventive vis-à-vis des cancers. Compte tenu des incertitudes sur l'estimation de l'exposition aux IFs en Asie et des biais possibles liés à d'autres facteurs alimentaires, l'effet protecteurs des IFs sur les cancers du sein doit être considéré avec circonspection.

D'un autre côté, les IFs ont également montré des effets prolifératifs sur des cellules cancéreuses estrogéno-dépendantes, *in vitro* et chez les souris implantées *in vivo,* à des concentrations physiologiquement pertinentes [80,81]. Ces résultats ont été confirmés par 3 études cliniques menées dans les années 1990s sur des femmes pré-ménopausées saines [82–84], où un effet prolifératif a été observé à des doses comprises entre 45 et 50 mg/jour. Par ailleurs, une étude, sur des femmes atteintes d'un cancer du sein estrogéno-dépendant, a montré que la GEN induit l'expression de gènes impliqués dans la prolifération cancéreuse cellulaire [85]. Cependant, les études du NTP ont montré qu'à faibles doses, ces IFs n'ont pas d'incidence néfaste sur les animaux sains et que l'action de la GEN sur les rats varient selon qu'ils soient exposés à l'âge adulte ou *in utero* [13].

Ainsi, les effets des IFs sur les cancers du sein varient en fonction de plusieurs facteurs, notamment :

- la dose d'IFs consommée ;
- le moment de l'exposition aux IFs ;
- le statut hormonal de la femme ;
- le type de cancer du sein.

De plus, les différences interindividuelles dans la réponse aux IFs peuvent grandement influencer leur impact sur la cancérogénèse mammaire, sans compter le rôle d'autres facteurs de risque tels que la génétique et l'environnement

Selon Yuan *et al.* [86], la GEN stimule la croissance des cellules cancéreuses du sein (MCF-7 et T47D) de manière dose-dépendante à des concentrations relativement plus faibles que celles qui activent les ERs classiques. Toutefois, elle inhibe leur croissance à des concentrations plus élevées, d'ordre pharmacologique. Elle est largement métabolisée en glucuronides par les MCF-7 et en sulfates par les cellules T47D, alors qu'elle l'est faiblement par les cellules saines MCF-10A. La GEN-7-Glucuronide (GEN-7-G) présente une faible activité de stimulation dans les différentes cellules. La GEN-7-G subi un métabolisme important dans les cellules T47D et MCF-10A, mais pas dans les cellules MCF-7. Les effets prolifératifs de la GEN-7-G sur les cellules MCF-7 et T47D sont associés à son hydrolyse intracellulaire en GEN. En revanche, la GEN-7-S et la GEN-4′-S ne sont pas métabolisés par ces trois lignées cellulaires et n'ont aucune incidence sur leur croissance. Les auteur·es en ont conclu que la production de métabolites de phase II n'a pas d'impact sur l'effet prolifératif de la GEN sur les

cellules MCF-7 et T47D. La déconjugaison a été corrélée aux effets prolifératifs apparents de la GEN-7-G sur les cellules cancéreuses du sein [86].

II.2.4.2.2. Cancer de l'endomètre

L'impact des IFs sur le cancer de l'endomètre reste lui aussi objet à discussions.

Une méta-analyse récente a montré une augmentation de l'épaisseur de l'endomètre pour des doses d'IFs ingérées supérieures à 20 mg/jour, sans lien avec le cancer de cet organe. Cet effet est attribué à la stimulation de la prolifération des cellules endométriales par les IFs, imitant l'action des estrogènes [87,88]. Une dose plus élevée d'IFs (>54 mg/jour) a quant à elle été associée à une réduction de l'épaisseur de l'endomètre chez les femmes ménopausées nord-américaines contrairement aux femmes asiatiques pour lesquelles la conséquence inverse est observé [88]. Malgré cet impact, les études épidémiologiques ne parviennent pas à établir un lien clair entre la consommation d'IFs et le risque de cancer de l'endomètre. La complexité pourrait provenir des interactions des IFs avec les ERs (ERα et ERβ présents dans la sphère endométriale), des différences entre les femmes pré-ménopausées et post-ménopausées (chez qui les ERs disparaissent avec le temps) ainsi que des autres facteurs de risque du cancer de l'endomètre. En effet, certaines études n'ont observé aucune association significative entre IFs et cancer de l'endomètre [89], alors que d'autres ont suggéré une possible augmentation du risque chez les femmes ménopausées utilisant des traitements hormonaux substitutifs (THS) et consommant des IFs. Une étude a, au contraire, rapporté qu'une consommation d'IFs à partir d'aliments à base de soja et de légumineuses était associée à une réduction de 22% du risque de ce cancer chez les femmes asiatiques *vs* 18% chez les non-asiatiques [90]. Ce dernier résultat n'a cependant pas été observé quand les études de cohortes étaient considérées [91].

L'utérus, comme les ovaires et le vagin, exprime majoritairement ERβ, ce qui pourrait expliquer l'absence d'effet néfaste clair sur le cancer de l'endomètre. Ce récepteur, pour lequel les IFs ont une meilleure affinité, est impliqué dans la différenciation cellulaire et pourrait contrer l'effet prolifératif de ERα [87]. Chez les femmes post-ménopausées, il est possible d'observer une influence des IFs sur la sphère vaginale, tandis qu'un impact sur l'endomètre n'a jamais été clairement prouvé malgré de nombreux essais cliniques [87].

II.2.4.2.3. Cancer de la prostate

Plusieurs études ont exploré l'impact des IFs, en particulier la GEN, sur le cancer de la prostate.

- Etudes *in vitro* et animales :

Des investigations *in vitro* ont montré que les IFs peuvent inhiber la croissance des cellules des cancers de la prostate par l'induction de l'apoptose et l'arrêt du cycle cellulaire, à des doses supra-physiologiques [92]. Zhang *et al.* [93] ont également suggéré que les IFs pourraient réduire la sécrétion d'antigène prostatique spécifique (PSA), un marqueur de ce cancer. Toutefois, la reconnaissance de ce biomarqueur des tumeurs est aujourd'hui controversée et une diminution de PSA ne signifie pas forcément une réduction des cancers prostatiques. Par ailleurs, les IFs à des concentrations élevées, pourraient stimuler les cellules cancéreuses [93]. L'equol a quant à lui été associé à un effet protecteur sur différents modèles i*n vitro* [94].

Des essais sur des rats et des souris génétiquement modifiés pour développer un cancer de la prostate ont permis de réduire la taille et la dissémination des tumeurs après l'administration de GEN à doses pharmacologiques [92].

- Etudes cliniques chez l'homme :

Les études cliniques menées chez l'homme n'ont pas permis de démontrer de manière concluante l'efficacité des IFs dans la prévention du cancer de la prostate [92–94].

La diminution de la sécrétion de PSA observée *in vitro* n'a pas été confirmée par tous les essais cliniques [93]. Ce résultat pourrait s'expliquer par le fait que les concentrations d'IFs utilisées dans les études *in vitro* sont souvent bien supérieures à celles présentes dans les fluides biologiques lors d'une alimentation courante.

Une association négative a été démontrée entre consommation d'IFs et risque de cancer de la prostate dans les populations asiatiques. Les effets des IFs sur l'incidence de la pathologie diffèrent entre les populations occidentales et asiatiques. Cette disparité semble concerner la progression du cancer, puisque la survenue du cancer, analysée par diagnostic *post-mortem*, a montré des fréquences d'occurrence similaires entre les deux populations [3]. Le type de récepteurs de l'E_2 (ER) impliqué est ici crucial. En effet, les IFs peuvent être protectrices sur les tumeurs porteuses de ERβ et néfastes sur les tumeurs porteuses du variant ERβ2 [87]. En outre, une augmentation de l'expression d'ERα dans le cancer de la prostate de haut grade a été rapportée et pourrait favoriser la croissance des cancers prostatiques [94].

Il est cependant important de noter que l'incidence du cancer de la prostate est plus faible en Asie qu'en Occident, ce qui suggère un lien possible avec l'alimentation et d'autres facteurs environnementaux, ne pouvant être réduits exclusivement aux IFs. En outre, selon certaines études, l'effet protecteur des IFs sur les niveaux d'hormones sexuelles et le risque de cancer de la prostate serait principalement déterminé par la capacité de métabolisation de la DAI en equol qui pourrait varier entre les deux populations étudiées [94].

II.2.4.3. Fibromes utérins

La sphère utéro-vaginale exprime des récepteurs aux estrogènes (ERα, ERβ et GPER) [95] ce qui rend plausible l'effet des phytoE sur ces organes. De plus, les effets utérotrophiques des IFs du trèfle et du soja sont bien documenté [96].

Une méta-analyse a établi un lien entre la consommation de soja, à la fois dans la petite enfance et à l'âge adulte, et un risque accru de fibromes utérins [97]. Une autre étude a montré que la consommation importante de soja, comparable à la prise de pilule contraceptive, augmentait le risque de léiomyomes utérins [98].

Chandrareddhy *et al.* [20] ont rapporté plusieurs cas de femmes adultes ayant développé des myomes, des fibromes et, dans deux cas, des signes d'endométriose suite à une surconsommation chronique et excessive de soja. Par ailleurs, une corrélation positive significative a été observée entre le diamètre de fibromes préexistants et la consommation de produits dérivés de soja dans la petite enfance [66]. Également, une étude informelle indique que plus de 70 mg d'IFs par jour peuvent être associés à des saignements endométriaux chez les femmes ménopausées [99].

En conclusion des 2 sections précédentes, il apparait que l'impact des IFs sur la santé utérine est complexe. Des études supplémentaires sont requises pour mieux

comprendre les effets à long terme de la consommation d'IFs sur les risques d'occurrence de fibromes, de léiomyomes et de cancers de l'utérus, en tenant compte de la dose, de la récurrence, du moment de l'exposition et des facteurs individuels.

II.2.4.4. Fertilité

L'association entre les IFs et les troubles de la fertilité chez les mammifères a fait l'objet de plusieurs rapports. Cependant, les données concernant les humains restent limitées, en raison des difficultés éthiques liées à la conduite d'études d'intervention depuis l'identification d'un risque potentiel.

Les IFs ont une activité hormonale et peuvent perturber plusieurs mécanismes endocriniens, notamment réduire la sécrétion de la FSH (hormone folliculo-stimulante) et de la LH (hormone lutéinisante), hormones hypophysaires contrôlant la reproduction.

II.2.4.4.1. Chez les femmes

Une étude menée auprès de 11 688 Américaines du Nord et âgées entre 30 et 50 ans, dont plus de la moitié étaient végétariennes et ayant une hygiène de vie saine, a montré qu'une consommation élevée d'IFs (≥ 40 mg/jour) diminue de 3% la probabilité de donner naissance à un enfant avant 26 ans [99]. La même étude rapporte une augmentation de 13% du risque de nulligravidité à la ménopause en cas de consommation élevée d'IFs.

D'un autre côté, une étude randomisée a montré qu'une supplémentation en IFs de soja à fortes doses (1 500 mg par jour), en association à de la progestérone, peut être bénéfique dans le cadre d'un traitement de fécondation *in vitro* (FIV) [100]. A cette dose, des effets anti-FSH et anti-LH sont certains et pourraient expliquer ce résultat. Une autre étude rapporte que des concentrations d'IFs dans l'urine de l'ordre de la centaine de µg/L sont associées à une probabilité accrue de naissance vivante à l'issue d'un cycle de FIV [101]. Ces concentrations pourraient correspondre à une consommation d'un ou deux produits contenants « du soja caché » par jour.

En outre, la consommation de soja semble moduler l'effet du BPA sur la fertilité féminine. Une exposition plus élevée au BPA peut nuire aux taux de naissance vivante chez les femmes états-uniennes qui ne consomment pas de soja, mais cet effet semble atténué chez les consommatrices de soja et ce à des doses faibles de 3,4 mg par jour [102]. Cette dose correspond à une consommation de « soja caché ».

II.2.4.4.2. Chez les hommes

L'influence des IFs sur la qualité et la quantité du sperme demeure un sujet discuté. En effet, si les études de populations indiquent une réduction de la qualité et de la quantité spermatique chez les hommes exposés à de fortes doses d'IFs (un à deux produits de soja par jour), le lien entre la concentration de référence des spermatozoïdes dans le fluide séminal et la fertilité n'est pas direct ni toujours dose-dépendant. En fait, il existe une zone de concentration correspondant à un plateau de fertilité. C'est lorsque la concentration en spermatozoïdes viables diminue en dessous d'un certain seuil (hypospermie) que les problèmes de fertilité apparaissent. Considérant les doses qui semblent entraîner cette réduction de fertilité, ce phénomène pourrait être lié à une perturbation des FSH et LH hypophysaires, elles même responsables de la synthèse des androgènes chez l'homme.

Pourtant, d'autres études, notamment d'intervention, n'ont trouvé aucun lien significatif. Un point de vigilance doit nous alerter quant à l'analyse de ces études. En effet, elles devraient prendre en considération le fait que la spermatogenèse chez les hommes est un long processus (au moins 76 jours [2]), donc l'exposition devrait durer plus de 3 mois pour s'assurer que l'absence de lien n'est pas due à un biais temporel. De plus, les variations individuelles, et notamment le temps depuis la dernière éjaculation, devraient être prises en compte. Ce n'est malheureusement pas toujours le cas [103].

Même pour l'étude citée dans le tableau 1 ([5]) qui montre un lien entre IFs et une réduction du nombre de spermatozoïdes, plus significative avec des doses plus élevées d'IFs, il est impossible de dissocier l'effet des IFs de celui d'autres PEs potentiels [104,105]. Il faudrait donc raisonner sur la base d'un effet cocktail.

Ainsi, comme les PEs sont ubiquitaires et qu'il peut être difficile de les réduire à court terme, la réduction des IFs dans l'environnement peut représenter une solution à ne pas négliger.

II.2.4.5. Cycle menstruel

Ce que les investigations regroupées dans le tableau 1 montrent, c'est qu'il existe un petit allongement du cycle menstruel pour les femmes prenant des doses nutritionnelles quotidiennes d'IF (45 mg/jour). Cet allongement même s'il ne représente que quelques dizaines d'heures indique pourtant un effet endocrinien. Notons ici que dans plusieurs études, réalisées en Asie, les chercheurs rapportent une durée des cycles de 30 jours. Dans celle de Watanabe [106] qui demandait à des étudiantes japonaises de prendre en plus de leur régime habituel, 50 mg d'IFs dans du jus de soja, les cycles menstruels sont passés de 30 à 32 jours. Il est important de mentionner ici que les taux circulant d'IFs mesurés chez ces étudiantes au début de l'étude sont du même ordre que ceux que l'on mesure chez des femmes occidentales qui auraient consommé un steak de soja ou une tasse de « lait » de soja en France. Cet allongement du cycle est dû à un retard dans l'apparition du pic de LH impliqué dans l'ovulation, ce qui conduit à un allongement de la phase folliculaire du cycle menstruel [19].

Des adultes nourris dans la petite enfance avec des formules infantiles à base de soja ont expérimenté des cycles irréguliers [67]. Dans le cas de ces formules, les concentrations plasmatiques élevées d'IFs semble indiquer qu'elles sont bien absorbées par les nourrissons [107], même si le moment où un nourrisson acquière le microbiote impliqué dans la métabolisation n'est pas clair. De plus, leur système enzymatique de détoxification hépatique n'est pas complètement développé ce qui peut entrainer une exposition exclusive, quantitativement importante et continue à des IFs (aglycones). Des préoccupations peuvent être soulevées concernant l'impact de l'exposition aux estrogènes des nourrissons à un âge où la différenciation sexuelle n'est pas encore achevée alors que les synthèses de FSH et LH sont effectives lors de la mini-puberté [108]. Il est cependant important de noter que les délais d'évaluation des effets potentiels sur les adultes n'est probablement pas suffisant pour ce genre d'exposition.

II.2.4.6. Ménopause et ostéoporose

Les impacts des phytoE sur les symptômes de la ménopause font l'objet de nombreuses études suivies de controverses. Ainsi, diverses méta-analyses ont été réalisées pour évaluer leur efficacité.

Une méta-analyse comparant divers traitements naturels et médicamenteux a révélé que les plus efficaces sont ceux à base d'IFs et d'actée à grappes noires. Les IFs ont un *odd ratio* (OR) et un intervalle de confiance (IC) à 95% de 0,62 (0,44-0,67), tandis que pour l'actée à grappes noires, ces valeurs sont de 0,4 (0,16-0,90). Une autre méta-analyse axée sur les essais cliniques portant sur ces mêmes composés, a observé une réduction de 25,2 % des bouffées de chaleur après avoir éliminé l'effet placebo. Il est important de noter que l'effet placebo représente 57% de la réduction des symptômes par le traitement de référence : l'E_2. L'étude a également révélé qu'un traitement de 13,4 semaines était nécessaire pour que les IFs de soja produisent la moitié de leurs effets maximaux, et qu'il fallait au moins 48 semaines pour obtenir 80 % de ces effets. Malgré la confirmation de ces effets par ces méta-analyses, l'impact des phytoE sur les symptômes de la ménopause reste débattu [87] pour plusieurs raisons :

- une grande variabilité interindividuelle ;
- des effets différents selon le statut péri- ou post-ménopausique ;
- un fort effet placebo ;
- des études principalement basées sur des auto-déclarations.

Concernant l'ostéoporose, deux méta-analyses s'accordent à dire que les IFs peuvent aider à maintenir la densité minérale osseuse (DMO) chez les femmes ménopausées, mais seulement à des doses supérieures à 80 mg/jour [87]. La première, portant sur 1 240 femmes ménopausées et analysant des essais contrôlés randomisés (ECR) d'une durée de six à douze mois, a révélé un maintien de la DMO au niveau de la colonne vertébrale uniquement, avec une dose quotidienne moyenne de 82 mg d'IFs de soja. En revanche, la seconde, incluant des ECR d'une durée minimale d'un an, n'a pas mis en évidence d'effet significatif sur la DMO avec une dose quotidienne moyenne de 87 mg d'IFs. Néanmoins, des doses plus élevées semblent contribuer au maintien de la DMO au niveau des vertèbres lombaires et de la hanche [87].

Il est important de noter ici que les IFs n'ont jamais permis la reconstruction d'un squelette atteint d'ostéoporose chez l'humain ou les rongeurs [87].

Les données actuelles suggèrent que les IFs peuvent avoir des effets bénéfiques sur les symptômes de la ménopause et la DMO chez les femmes ménopausées, mais à des doses élevées et après un traitement prolongé.

II.2.4.7. Métabolisme

Une méta-analyse récente, regroupant 46 essais cliniques, confirme une légère diminution du cholestérol LDL (environ 3,2 %) et du cholestérol total (CT) (environ 2,8 %) après une consommation quotidienne de 25 g de protéines de soja pendant 6 semaines [109]. Cette réduction, bien que modeste est significative et concorde avec l'effet hypocholestérolémiant observé avec d'autres légumineuses à faible teneurs en IFs. Ceci suggère un effet non spécifique aux IFs [3].

Concernant le risque de syndrome métabolique, plusieurs études ont établi une corrélation, dépendant généralement du sexe, avec la consommation d'IFs [3,110].

Des études randomisées et contrôlées, présentées ci-dessous, ont apporté des éléments plus probants.

- Une supplémentation en GEN à 54 mg/jour (équivalente à un apport alimentaire) pendant un an chez des femmes ménopausées atteintes du syndrome métabolique a amélioré le profil lipidique, réduit la glycémie à jeun, l'insuline et l'indice HOMA-IR, et augmenté l'adiponectine. Les pressions artérielles systolique et diastolique ont également été réduites dans le groupe traité [87]. Cependant, les données sur l'évolution du régime alimentaires des volontaires pendant l'étude manquent alors qu'elle ne peut être exclue.
- Une supplémentation similaire à celle de la précédente étude pendant six mois chez des femmes ménopausées italiennes a révélé une amélioration légère, mais significative, de la dilatation médiée par le flux sanguin, un marqueur de la santé vasculaire, ainsi qu'une réduction du CT, des triglycérides, de l'homocystéine et de la visfatine [111]. Cet effet pourrait être relié à l'action de la GEN sur l'activité de la NO-Synthase connue pour participer au relâchement de la contraction des cellules musculaires lisses de la paroi vasculaire [112].
- Une comparaison entre la consommation de graines de soja (soja trempé et grillé) et de protéines de soja chez des femmes ménopausées atteintes d'un début de syndrome métabolique a révélé une amélioration du profil lipidique dans le groupe « graines de soja » et une réduction du CT, de l'insuline et du stress oxydatif dans le groupe « protéines de soja » [110]. Les graines de soja contenaient 47,6 mg DAI, 60,2 mg GEN et 9,45 mg glycitéine ; tandis que dans les protéines, les teneurs en IFs étaient de 38,5 mg DAI, 48,8 mg GEN et 8,9 mg glycitéine.
- Une étude sur la consommation d'équol, à raison de 10 mg/jour, a montré une amélioration modeste de l'HbA1c, du LDL-cholestérol et de l'index vasculaire cardio-jambier chez des sujets japonais en surpoids et obèses non producteurs d'équol [87].

Une méta-analyse réalisée par Simental-Mendía *et al.* [113] a en effet montré qu'il n'y avait pas d'effet des IFs de soja sur les concentrations plasmatiques de la lipoprotéine a (Lpa), une lipoprotéine de basse densité associée à un risque cardiovasculaire accru en raison de ses propriétés pro-thrombotiques et athérogènes.

Dans une autre méta-analyse portant sur un total de dix cohortes prospectives et sept études cas-témoins (incluant 508 841 participant·es et 17 269 événements cardiovasculaires), il a été observé que la consommation d'aliments à base de soja, mais pas celle d'IFs, était associée à un risque plus faible de maladies cardiovasculaires totales (y compris les maladies coronariennes et les accidents vasculaires cérébraux) [91].

Finalement, la consommation de soja, et en particulier de GEN et d'équol, peut contribuer à la réduction de certains facteurs de risque du syndrome métabolique, notamment le profil lipidique, la glycémie à jeun et l'insuline. Des études supplémentaires restent toutefois nécessaires pour déterminer les doses optimales, les populations les plus à même de bénéficier de ces effets et les interactions avec d'autres facteurs de risque.

II.3. Exposition aux phyto-estrogènes, enjeux de santé et réduction

II.3.1. Exposition aux perturbateurs endocriniens en France

II.3.1.1. Isoflavones

La source principale d'exposition aux phytoE, notamment les IFs reste l'alimentation ou les compléments alimentaires. Cette exposition se fait principalement par l'ingestion de soja et dans une moindre mesure de quelques autres légumineuses comme le haricot mungo ou le lupin. Son évaluation est essentielle pour comprendre les effets potentiels sur la santé humaine.

En France, les premières estimations de l'exposition aux IFs (GEN et DAI) datent de 2003 [8]. L'AFSSA avait estimé l'apport moyen en IFs à 0,151 mg/jour, en soulignant une différence importante entre la population générale et les consommateur·ices de produits à base de soja. Dans cette dernière catégorie, l'exposition moyenne était de 54 mg/jour et pouvait atteindre jusqu'à 109 mg/jour en cas de consommation quotidienne de plus d'un produit, avec un minimum de 8,3 mg/jour. L'AFSSA a également estimé les apports en IFs aglycones chez les non consommateur·ices de soja à 0,026 mg/jour en moyenne, en considérant les IFs des autres légumineuses [8]. Cependant, cette estimation est susceptible d'avoir sous-estimé l'exposition réelle aux IFs pour plusieurs raisons :

- la proportion de la population consommant des aliments à base de soja a augmenté de manière significative, en 2015, elle était d'environ 10 % et a atteint 60 % en 2020 ;
- l'AFSSA n'a pas pris en compte l'exposition aux IFs provenant des produits industrialisés contenant du « soja caché » [75].

A titre comparatif, des études plus récentes ont montré une exposition aux phytoE bien plus importante. En 2012, Zamora-Ros *et al.* [114] ont rapporté une exposition moyenne de 18 mg/jour chez les femmes britanniques ménopausées, prenant soin de leur santé. Leur exposition totale incluant les lignanes, et les compléments alimentaires à base de phytoE et de soja s'élevait à 21 mg/jour. Aux États-Unis, l'exposition quotidienne aux IFs estrogéniques a été estimée à environ 2,7 mg au moins sur la même période. Quant à celle chez les populations asiatiques, elle a été évaluée à entre 15 et 50 mg par jour [115].

Plusieurs études ont mis en évidence l'importance du « soja caché » dans l'exposition aux IFs [116]. En 2017, Lee *et al.* [117] ont montré que 41% des produits alimentaires transformés en France contenaient des légumineuses et environ 12% contenaient du soja. Leur analyse a concerné également les produits disponibles via un des principaux fournisseurs de cantines et restaurants. La Figure 3 montre la proportion d'aliments contenant des légumineuses, notamment du soja, parmi 10 catégories. Il est à noter, par exemple, que 100% des boulettes de viande et des burgers proposés dans les cantines contenaient du soja [117].

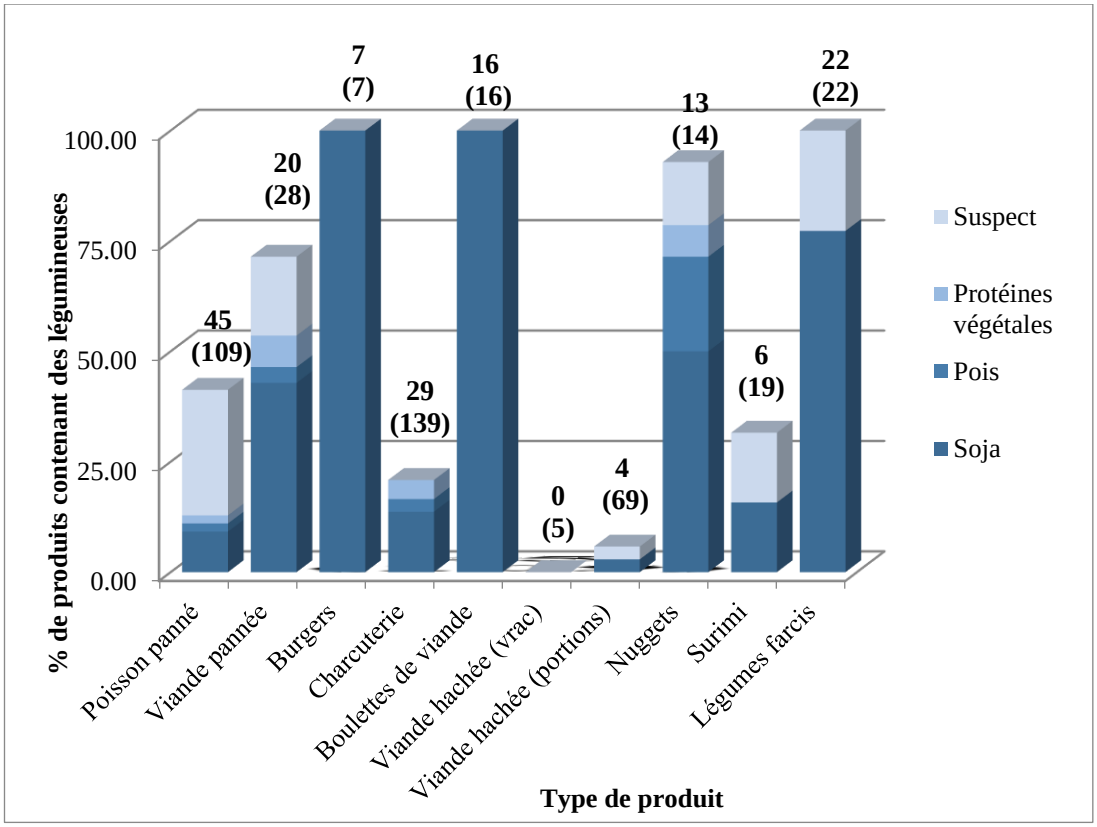


Figure 3. Pourcentage de produits alimentaires contenant des légumineuses dans les différentes catégories proposées par le fournisseur de restauration collective. Les chiffrents indiquent le nombre d'articles alimentaires qui en contiennent et, entre parenthèses, les nombres totaux. (D'après Lee et al. [117]*).*

Lee *et al.* ont révélé dans cette même étude que 12 % des femmes pré-ménopausées interrogées consommaient plus de 9 portions de soja par semaine, les exposant à plus de 50 mg d'IFs par jour.

Par conséquent, l'évaluation de l'exposition aux IFs basée uniquement sur la consommation de produits à base de soja est susceptible de sous-estimer l'apport réel. Il est crucial, dans les pays occidentaux, de prendre en compte le « soja caché » présent dans les produits transformés pour obtenir une estimation plus précise de l'exposition aux phytoE.

En conclusion, l'estimation de l'exposition aux IFs en France n'est ni simple ni directe et nécessite de prendre en compte plusieurs éléments pour obtenir une évaluation précise :

- la consommation de produits à base de soja ainsi que la présence de « soja caché » dans les produits transformés ;
- les teneurs différentes en IFs du soja déclaré et du « soja caché » (à base de concentrés ou d'isolats protéiques de soja) ;
- la consommation de lignanes ;
- la prise de compléments alimentaires à base de phytoE.

II.3.1.2. Exposition à d'autres perturbateurs endocriniens

Les PEs se retrouvent dans de nombreux produits du quotidien (produits ménagers, cosmétiques, aliments) et contaminent les milieux naturels (eaux, sols, air). Cette ubiquité, combinée à la diversité des substances chimiques regroupées sous le terme PE et aux multiples voies d'exposition (orale, cutanée et respiratoire), rend l'estimation de l'exposition globale de la population extrêmement difficile.

Par ailleurs, l'établissement de liens de causalité entre l'exposition aux PEs et leurs effets néfastes sur la santé représente un défi majeur. Cette complexité s'explique par plusieurs facteurs qui sont regroupés ci-dessous.

- Les mécanismes d'action des PEs sont souvent multiples et non linéaires. Ils peuvent interférer avec le système hormonal de différentes manières, à des doses très faibles, et leurs effets peuvent être non monotones, c'est-à-dire ne pas suivre une relation dose-effet proportionnelle.
- Les pathologies liées aux PEs sont généralement multifactorielles. Elles peuvent être causées par une combinaison de facteurs génétiques, environnementaux et de mode de vie, ce qui rend difficile l'attribution d'un effet spécifique à un seul PE.
- L'effet cocktail ou l'exposition simultanée à plusieurs PEs peut avoir des effets additifs, synergiques ou antagonistes, ce qui complique encore davantage l'évaluation des risques.
- Des effets différés et transmissibles sur plusieurs générations en raison de modifications épigénétiques sont aussi à prendre en compte.
- L'absence de population témoin à laquelle comparer la population exposée.

L'exemple des pesticides à faibles doses illustre bien la complexité de l'évaluation des risques liés aux PEs. Les consommateur·ices sont exposé·es à des mélanges complexes de pesticides qui peuvent avoir des effets additifs ou antagonistes [118]. L'effet global dépend de la concentration de chaque substance, de sa pharmacocinétique et de sa pharmacodynamique ainsi que des interactions avec d'autres PEs présents dans le mélange, y compris les agents d'enrobage, les retardateurs de flamme ou les IFs. Il a notamment été démontré que le glyphosate, à de faibles concentrations, avait des effets additifs avec la GEN *in vitro* sur la prolifération des cellules MCF-7 [119].

McKinlay *et al.* [120] ont rapporté que l'exposition aux pesticides des enfants issus de familles d'agriculteur·ices ou vivants à proximité de cultures de soja ou de maïs était cinq fois supérieure à celle des enfants de familles non agricoles. L'étude souligne également l'influence du régime alimentaire sur l'exposition aux pesticides. Les personnes consommant une proportion importante d'aliments biologiques sont moins exposées aux résidus de pesticides. En revanche, les consommateur·ices de poisson gras, de viande et de produits laitiers risquent d'être exposés à des doses plus élevées de pesticides lipophiles persistants, comme les organochlorés, en raison de la bioaccumulation. Le cas des nourrissons et des enfants est particulièrement préoccupant. En raison de leur régime alimentaire moins diversifié (composé de quelque fruits, céréales et légumes) et de leur rapport « poids corporel / consommation » plus élevé, ils sont exposés de manière disproportionnée à certains organochlorés et aux nouveaux pesticides moins persistants d'origine alimentaire [120].

Ainsi, l'évaluation de l'exposition aux PEs est un domaine complexe et en constante évolution. Il est important de développer des stratégies pour réduire cette exposition, en particulier pour les groupes les plus vulnérables.

II.3.2. Impact de la transformation alimentaire sur la teneur en isoflavones des aliments

La comparaison entre la consommation traditionnelle asiatique de soja et celle, plus récente, des pays occidentaux a permis de dévoiler des différences très importantes [18,75]. Contrairement aux procédés modernes, ces recettes traditionnelles réduisent la teneur en IFs. L'évaluation de l'exposition aux IFs des consommateur·ices asiatiques doit donc tenir

compte de l'origine des produits qu'ils consomment, qui ne sont pas nécessairement industriels.

Il est en effet important de rappeler que les agences sanitaires internationales peinent à fixer une dose d'IFs maximale autorisée, en partie à cause des contradictions entre les données toxicologiques et les données d'exposition en Asie. En effet, les données toxicologiques (notamment, du NTP) suggèrent que des apports quotidiens moyens d'IFs de 40 à 60 mg devraient entrainer des troubles, notamment de la reproduction et de la fonction thyroïdienne, et un risque accru de cancers estrogéno-dépendants. Ces doses correspondent à la consommation moyenne estimée des populations asiatiques [119]. Cependant, ces prédictions ne concordent pas avec les données épidémiologiques en Asie. Bien que les données de fertilité globales soient plus faibles dans certains pays asiatiques développés (Japon, Corée du sud, Taïwan, …), sans lien avec les taux de pénétration de la contraception dans ces populations, les données scientifiques ne confirment pas une association systématique entre la consommation de soja et d'IFs et des troubles de la fertilité. De plus, la consommation de soja en Asie remonte à plusieurs millénaires sans que des effets néfastes des IFs sur la santé n'aient été clairement établis.

L'absence apparente d'effets néfastes sur la reproduction chez les populations asiatiques malgré une consommation élevée de soja suggère que l'exposition aux IFs dans ces populations pourrait être plus faible que ce qui est actuellement estimé. Cette divergence pourrait être lié au fait que l'exposition aux IFs est estimée sur la base de mesures faites sur des aliments commerciaux, majoritairement industriels. Cependant, le soja est encore préparé selon les méthodes traditionnelles en Asie, notamment dans les régions rurales [255]. Ces méthodes impliquent généralement des traitements à l'eau tels que le trempage et l'ébullition ainsi que la fermentation après ébullition. Or, si ces traitements sont minimisés dans l'industrie il a été démontré que la cuisson dans l'eau réduisait de façon très significative les taux d'IFs dans les graines de soja, en raison de l'hydrosolubilité des IFs glycosylées [75].

II.3.2.1. Transformations traditionnelles à l'échelle domestique

Le soja, longtemps reconnu par les anciennes civilisations asiatiques pour sa faible digestibilité à l'état cru, a d'abord été utilisé comme engrais vert et pour la nutrition animale. Son incorporation dans l'alimentation humaine s'est faite graduellement, poussée par les famines et les guerres. Les Chinois, pionniers de sa consommation alimentaire, ont compris la nécessité de le faire bouillir pour pallier à ses effets digestifs défavorables. En effet, le soja, dédaigné par les rongeurs et les oiseaux, a toujours nécessité une transformation pour devenir comestible, en raison des facteurs antinutritionnels qu'il contient.

Les écrits chinois les plus anciens mentionnent des cuissons dans l'eau, suivies de fermentations pour améliorer le goût des bouillies de soja. Les préparations coagulées de soja, comme le tofu sont plus tardives et semblent dater de la fin de la dynastie Han, vers 200 après J.-C.

En Asie, où la consommation de soja est traditionnelle, les recettes se transmettent au sein des familles et une part non négligeable de la population continue de recourir à des préparations domestiques. Tofu, miso, feuilles de soja, pâte de soja, tempeh et sauce soja font partie des préparations les plus communes [18,75].

II.3.2.2. Transformations à l'échelle industrielle

Le soja est un ingrédient polyvalent présent dans une large gamme de produits alimentaires. On le retrouve sous forme entière ou transformée.

Les produits à base de soja peu transformés incluent le tofu, le tempeh, le lait, les yaourts et la crème de soja. Ils sont généralement produits à partir de graines de soja entières ou dépelliculées qui ont subi un minimum de traitement, tel que le trempage, la germination, la fermentation ou la coagulation.

Ceux plus transformés comprennent les protéines de soja texturées (PST), les concentrés protéiques et les isolats protéiques de soja. Ils sont obtenus à partir de graines de soja qui subissent plusieurs étapes de transformation dont le déroulé est le suivant :

- Pressage : les graines de soja entières ou dépelliculées sont pressées pour en extraire l'huile.
- Extrusion ou broyage : le tourteau restant après l'extraction de l'huile est soit broyé pour obtenir une farine déshuilée ou extrudé pour obtenir des PST.
- Traitements physiques : la farine déshuilée peut subir divers traitements physiques, (extrusion, micronisation, chauffage), pour obtenir des propriétés fonctionnelles spécifiques. Des traitements chimiques peuvent également être utilisés pour modifier les propriétés des protéines de soja, comme la solubilité ou le pH.

En Europe et aux États-Unis, les principaux aliments à base de soja consommés, comme la farine, les isolats de protéines, les protéines de soja concentrées et texturées, sont souvent des ingrédients d'autres spécialités alimentaires manufacturées [75,121].

L'analyse précise et reproductible des IFs dans les aliments nécessite une connaissance approfondie du type de produit à base de soja et de son procédé de fabrication. Cette information est essentielle pour interpréter correctement les résultats car la teneur en IFs peut varier considérablement en fonction de :

- l'origine agronomique du soja, car les conditions culturales affectent de manière importante la teneur en IFs des graines de soja ;
- la variété du soja ;
- le procédé de transformation ;
- les autres ingrédients dans le produit alimentaire peuvent également influencer la teneur finale de l'aliment en IFs ;
- la taille de la portion.

II.3.3. Evaluation de l'exposition aux phyto-estrogènes

Avant la publication des recherches dirigées par la Pr. Bennetau-Pelissero, les données récentes sur l'exposition aux phytoE de soja en France étaient inexistantes. La consommation de soja, en particulier du « soja caché », était difficile à évaluer. Le présent travail vise à apporter des données inédites sur ce sujet dans le contexte français, sachant qu'il existe des publications régulières aux États-Unis d'Amériques, dans plusieurs pays européens et en Asie.

II.3.3.1. Enquêtes alimentaires

En nutrition, l'évaluation des consommations alimentaires et leur corrélation avec des paramètres de santé constituent une approche classique. Les agences de sécurité sanitaire des aliments, comme l'Anses en France, publient régulièrement les résultats d'enquêtes alimentaires, à l'instar de l'étude INCA, menée tous les 7 ans. Ces données permettent d'élaborer des recommandations nutritionnelles pour l'ensemble de la population [122,123].

Les enquêtes alimentaires peuvent cibler une population spécifique, comme dans le cadre de cette thèse, ou viser à représenter l'ensemble de la population. La taille de l'échantillon doit être adaptée et représentative de la population cible. La méthodologie de l'enquête est cruciale : elle doit être pertinente, concise et privilégier des questions ouvertes pour éviter d'influencer les réponses.

L'identification adéquate des aliments d'intérêt s'avère un facteur crucial pour garantir la pertinence d'une enquête alimentaire. Concernant les IFs spécifiquement, Lee *et al.* [117] ont analysé plus de 12 000 étiquettes d'aliments distribués en France. Cette analyse a permis d'identifier les aliments contenant du soja, y compris le « soja caché » dans les produits transformés industriels. Des pourcentages de produits contenant du soja ont été établis par catégorie de produit et selon les lieux d'achat (grandes surfaces ou centrales d'achat pour la restauration collective). Parmi ces produits figurent des burgers, des boulettes de viande, des saucisses, des kebabs, des poissons pannés, des pâtés et des rillettes de poissons, par exemple.

II.3.3.2. Dosages dans les échantillons biologiques

Lorsque le suivi de marqueurs spécifiques à des aliments d'intérêt est possible, les analyses biologiques constituent un complément précieux pour valider les déclarations d'expositions. Un domaine de la métabolomique appliqué à la nutrition se développe actuellement, visant à identifier des marqueurs spécifiques d'aliments, tels que des vitamines [124] ou des polyphénols [125], et à relier ces profils analytiques aux déclarations des consommateur·ices.

Le cas du soja est particulièrement propice aux analyses de consommation car il représente la principale source d'IFs estrogéniques dans l'alimentation. Hormis les compléments alimentaires et les pousses de trèfles ou de luzerne, il n'existe pas d'autres sources significatives d'IFs en France.

Toutefois, l'analyse des échantillons biologiques doit tenir compte de la pharmacocinétique des composés recherchés. Cette considération a été prise en compte dans ce travail, ce qui n'a pas toujours été le cas dans des recherches antérieures [126,127].

Les choix méthodologiques des études présentées dans cette thèse ont été guidés par ces principes et seront détaillés dans le chapitre « Méthodologie, matériels et méthodes ».

II.4. Lupus Erythémateux Disséminé

II.4.1. Maladies auto-immunes

Les maladies auto-immunes (MAIs) se caractérisent par une perte de tolérance immunitaire envers les constituants du soi, conduisant à une attaque par le système immunitaire des tissus sains de l'organisme. Cette rupture de tolérance est à l'origine d'une large variété de pathologies, dont les mécanismes précis ne sont pas encore entièrement compris.

D'après l'Inserm [128], 5 à 10 % de la population mondiale est touchée par une MAI et les femmes en représentent 80 %. Cette prévalence accrue chez les femmes suggère l'influence probable d'hormones et de facteurs psychosociologiques [129,130]. Parmi les facteurs étudiés liés au sexe, les plus documentés [131] sont listés ci-dessous.

- Hormones sexuelles et histoire reproductive : les fluctuations hormonales au cours du cycle menstruel, de la grossesse et de la ménopause peuvent jouer un rôle dans le déclenchement ou l'aggravation des MAIs. Les estrogènes par exemple peuvent influencer la fonction

immunitaire, notamment la maturation, l'activation et la synthèse des lymphocytes, des anticorps et des cytokines.

- **Microchimérisme fœtal** : la présence et la persistance de cellules fœtales chez la mère après une grossesse pourrait contribuer à l'activation du système immunitaire et à la survenue de MAIs.
- **Gènes liés aux chromosomes sexuels** : le chromosome X contient un nombre important de gènes liés à l'immunité. Des perturbations dans le nombre de copies de certains de ces gènes, ou bien même dans la régulation de l'expression du chromosome entier, peuvent être à l'origine de la prédisposition des femmes aux MAIs. Par exemple, le syndrome de Turner, est associé à un risque accru de MAIs. Aussi, l'inactivation aléatoire d'un des deux chromosomes X chez les femmes pourrait influencer la réponse immunitaire et la susceptibilité aux MAIs.
- **Réponse immunitaire** : les femmes produisent plus d'anticorps et ont un plus grand nombre de lymphocytes circulants, ce qui pourrait également contribuer à leur prédominance dans les MAIs.
- **Microbiote** : des différences dans la composition du microbiote intestinal entre les sexes peuvent influencer la susceptibilité aux MAIs [132,133]. Le microbiote interagit avec le système immunitaire et peut moduler les réponses inflammatoires.
- **Facteurs psychosociaux** : les femmes sont plus susceptibles d'être exposées à des situations de stress chronique et à des événements traumatiques. Ces facteurs peuvent perturber le système immunitaire et augmenter la susceptibilité aux MAIs. De plus, les rôles sociaux et les expositions professionnelles peuvent jouer un rôle dans la sur-représentation des femmes dans certaines MAIs [134,135].

Il est important de noter que ces observations restent sujettes à caution et que les données scientifiques sont souvent contradictoires. Des recherches supplémentaires sont nécessaires pour mieux comprendre les facteurs contribuant à la prédominance féminine dans les MAIs.

Deux catégories principales de MAIs peuvent être distinguées :

- les MAIs spécifiques d'organes, pathologies localisées ciblant un organe ou groupe d'organes, comme le diabète de type 1 (pancréas) ;
- les MAIs systémiques qui affectent plusieurs organes ou systèmes simultanément, comme le lupus érythémateux disséminé (LED).

II.4.2. Lupus Erythémateux Disséminé

Le LED, également appelé lupus érythémateux systémique ou lupus érythémateux aigüe disséminé, est une MAI complexe et multifactorielle. Elle se caractérise par des poussées et des rémissions, et touche principalement les femmes. La symptomatologie est variable, allant de signes légers résorbables à des atteintes graves pouvant mettre en jeu le pronostic vital.

Le LED affecte majoritairement les femmes en préménopause, représentant 90% des cas. Ce ratio est l'un des plus marqués parmi les MAIs [131]. Les causes exactes du LED restent méconnues, mais des facteurs génétiques et environnementaux semblent jouer un rôle.

L'influence des hormones sexuelles sur le LED est un sujet de recherche important. Les estrogènes ont été associés à une augmentation de l'incidence et de la sévérité de la maladie, tandis que la testostérone et la progestérone semblent avoir un effet protecteur.

Les phytoE, dont les IFs de soja présentes dans de nombreux aliments transformés modernes, pourraient également influencer le LED. En raison de leurs propriétés estrogéniques et anti-androgènes, les IFs peuvent affecter la sécrétion de la LH et modifier le taux de progestérone dans le plasma. Par conséquent, les IFs peuvent être considérées comme des facteurs de risque environnementaux potentiels pour le LED et ses poussées.

Cependant, les données scientifiques concernant l'impact des IFs sur le LED chez l'humain sont encore limitées. Des études chez les rongeurs transgéniques ont montré des effets potentiels [136,137] mais des recherches cliniques chez les jeunes femmes sont nécessaires pour corroborer cette hypothèse.

II.4.2.1. Généralités

Le LED est une maladie inflammatoire chronique et auto-immune qui se caractérise par une grande complexité. De nombreux organes peuvent être ciblés par la maladie amenant à un tableau clinique très hétérogène résumé dans le Tableau 2.

Organe cible	Manifestation clinique
Peau et cheveux	Photosensibilité, rash malaire, alopécie
Poumons	Douleur thoracique pleurétique, inflammation, pneumonie, pleurite
Cœur	Douleurs thoraciques, souffle cardiaque, hypertension pulmonaire, insuffisance cardiaque
Reins	Inflammation, insuffisance rénale
Sang	Complexes immuns, caillots sanguins, anémie
Muscles et articulations	Gonflement des articulations, arthrose, myalgie
Bouche	Ulcères
Cerveau	Dépression, dysfonctions cognitives, fièvre

Tableau 2. Tableau clinique du LED (D'après Ameer et. [138]*)*

En plus de ces manifestations ciblées, la fatigue touche entre 67 et 90% des patient·es atteint·es de LED et a été décrite comme le symptôme le plus invalidant de la maladie en raison de son impact important sur la qualité de vie [139].

II.4.2.2. Prévalence du LED

L'incidence du LED est estimée entre 1 et 10 cas pour 100 000 personnes par an, avec une prévalence de 20 à 150 cas pour 100 000 personnes. Ces estimations varient considérablement d'une région du monde à l'autre [140]. L'Amérique du Nord enregistre les taux d'incidence et de prévalence les plus élevés, tandis que l'Afrique et l'Australie affichent les taux les plus bas.

Des facteurs tels que l'âge, le sexe et l'appartenance ethnico-sociale influencent considérablement l'évolution clinique et la gestion de la maladie. Par exemple, bien que le LED soit plus fréquent chez les femmes, il progresse plus rapidement et de manière plus sévère chez les hommes, conduisant à un pronostic défavorable.

L'incidence annuelle est actuellement de 6,73 nouveaux cas pour 100 000 personnes chez les Caucasiens et de 31,4 nouveaux cas pour 100 000 personnes chez les Afro-Américains. La prévalence est de 517 pour 100 000 personnes dans la population noire américaine,

comparativement à 134 pour 100 000 chez les Caucasiens américains et les Européens [138]. Ces disparités peuvent être attribuées à des facteurs génomiques et environnementaux, notamment des situations socioéconomiques plus défavorables dans la population afro-américaine (moins d'années d'éducation formelle, statut professionnel, revenu des ménages, plus grande pauvreté et inadéquation/absence d'assurance maladie) [130].

En Asie-Pacifique, le nombre de cas de LED fluctue entre 4,3 et 45,3 pour 100 000 personnes. Les malades atteints de LED et présentant un taux de mortalité élevé sont ceux ayant des niveaux de pauvreté plus élevés [130]. La fréquence de la maladie, souvent plus sévère, est 2 à 4 fois supérieure chez les individus non caucasiens, y compris les Asiatiques et les Australiens autochtones [141].

La prédominance féminine est de 9 : 1, ce qui fait du LED l'une des MAIs les plus différenciées en fonction du sexe [138]. Le LED est principalement observé chez les femmes en âge de procréer, entre 15 et 44 ans. Bien que la maladie puisse se déclarer à tout âge, elle apparaît le plus souvent à la fin de la deuxième décennie de vie des patient·es et au début de la troisième [140]. Chez les femmes, le LED figure parmi les 10 à 20 principales causes de mortalité, en fonction de leur âge [142,143].

II.4.2.3. Facteurs de risque

De nombreux facteurs contribuent à la susceptibilité au LED, et peuvent être regroupés en quatre catégories principales : génétiques, épigénétiques, environnementaux et socioculturels [130,140] :

- Facteurs génétiques :

  Les **polymorphismes génétiques** jouent un rôle crucial dans la susceptibilité à l'auto-immunité et peuvent expliquer les différences ethniques dans la présentation et la gravité des maladies [138,140]. Les études d'association pangénomique (GWAS) ont permis d'identifier de nouveaux variants génétiques associés au LED, bien que ces associations génétiques ne concernent qu'une minorité de patient·es et ont donc une signification clinique limitée [131]. De plus, plusieurs MAIs partagent des **variantes génétiques**, suggérant des voies pathogéniques communes [131].

- Facteurs épigénétiques :

  Les **modifications de l'ADN et des histones** peuvent influencer l'expression des gènes et jouer un rôle dans le développement du LED. De plus, L'exposition à certains facteurs **environnementaux** peut influencer les modifications épigénétiques et le risque de LED.

- Facteurs environnementaux et écologiques :

  Des études épidémiologiques ont montré l'association de l'incidence du LED avec l'exposition à la **silice**, aux **contraceptifs oraux estrogéniques**, à certains **médicaments** (certains antidépresseurs et antibiotiques) et à l'infection par le **virus d'Epstein-Barr** (EBV).

  D'autres facteurs environnementaux ont été associés au développement du LED. Les **particules fines** issues de la **pollution atmosphérique**, notamment celles provenant des gaz d'échappement diesel en milieu urbain, semblent jouer un rôle important. L'exposition à certains **métaux lourds**, comme l'uranium, le cadmium et le plomb, a également été corrélée à une augmentation du risque de LED. De plus, des **PEs** tels que le **bisphénol A** et les **phtalates** pourraient favoriser l'apparition de la maladie [144–146].

Le **tabagisme** affecte l'immunité innée et adaptative et peut influencer la production d'anticorps spécifiques dans des conditions telles que la polyarthrite rhumatoïde et le LED.

Les **ultraviolets B** sont connus pour affecter les fonctions du système immunitaire de la peau. Chez les individus prédisposés, ils peuvent activer des voies inflammatoires pathogènes, conduisant à des troubles auto-immuns. Dans d'autres cas, les UV peuvent être utilisés comme photothérapie pour désactiver l'agression immunitaire pathogène dans la peau. Ces effets divergents sont le résultat de la capacité des UV à ioniser les molécules de la peau et à modifier sa composition chimique. Dans certains cas, cela peut entraîner une blessure potentielle, comme la liaison de bases pyrimidiques adjacentes dans l'ADN. Les cellules peuvent subir l'apoptose pour éviter la transformation maligne, mais également déclencher des réponses immunitaires auto-réactives en exposant les antigènes nucléaires de la cellule.

Les mécanismes biologiques potentiels de ces associations environnementales comprennent l'augmentation du stress oxydatif, l'augmentation des cytokines inflammatoires, l'inflammation systémique et les modifications épigénétiques [138].

- Facteurs socioculturels :

Les **stress physiques** et **psychologiques** sont également rapportés comme facteurs de risque du LED.

Les facteurs de risque comprennent **l'accès aux soins** médicaux, l'observance, l'autogestion et les capacités d'adaptation, ainsi que le soutien social. Potentiellement modifiables, ils ont été identifiés comme très importants pour tous les indicateurs de résultats du LED [147].

En somme, la génétique joue un rôle important dans la susceptibilité à l'auto-immunité, mais d'autres facteurs tels que l'environnement et le mode de vie peuvent également influencer le risque de développer cette maladie. Les recherches en cours visent à mieux comprendre ces interactions et à développer de nouvelles stratégies de prévention et de traitement. Il est important de noter que ces facteurs de risque ne sont pas nécessairement indépendants. Ils peuvent interagir entre eux de manière complexe pour influencer le développement de la maladie.

II.4.2.4. Physiopathologie

L'étiologie exacte du LED n'est pas entièrement comprise. Cependant, il a été établi que la maladie résulte d'une interaction complexe entre facteurs génétiques et environnementaux. Cette interaction perturbe le système immunitaire et conduit à une réponse inappropriée, caractérisée par divers éléments décrits ci-dessous [138].

- Production d'auto-anticorps anti-nucléaires : ces anticorps attaquent les noyaux des cellules saines, provoquant une inflammation et des lésions tissulaires. La perte de tolérance aux antigènes nucléaires est un élément central du LED.
- Dérégulation des cytokines inflammatoires : les cytokines sont des protéines qui régulent l'inflammation. Leur production excessive dans le LED contribue à l'inflammation chronique et aux symptômes de la maladie.
- Production de complexes immuns : ces complexes se forment lorsque les auto-anticorps se lient à des antigènes nucléaires abondant dans le sang et les tissus. Ils peuvent ensuite se déposer et provoquer des lésions.

Le LED se caractérise également par des perturbations de l'immunité innée et adaptative. Ces perturbations affectent plusieurs fonctions immunitaires, notamment : la production de

cytokines, la clairance des cellules apoptotiques, l'immunité des lymphocytes B et la signalisation des lymphocytes T [140]. Elles sont synthétisées dans le Tableau 3.

Voies immunitaires	Anormalités liées au LED
Mécanisme de clairance immunitaire	Clairance altérée des débris apoptotiques qui stimule la réponse immunitaire innée et fournit une source d'antigènes nucléaires.
Interféron de type 1	Voie immunitaire clé impliquée dans le LED. Plus de la moitié des patientes ont une signature interféron (une surexpression des transcrits des gènes induits par l'interféron avec des conséquence auto-immunes et auto-inflammatoires multiples.
Neutrophiles	Stimulation de l'immunité innée par la formation de pièges extracellulaires à neutrophiles ainsi que par des dommages cellulaires et une stimulation de la synthèse de cytokines.
Cytokines et chémokines	Beaucoup de cytokines et chémokines sont mobilisées avec un déséquilibre entre les réponses pro- et anti-inflammatoires.
Cellules T	Altération des voies de signalisation, prolifération et production des cytokines, augmentation de la présentation d'auto-antigènes, présentation et provision de cellules B *helpers* combinée à une réponse altérée des cellules Treg.
Cellules B	Cellules B hyperactives avec des irrégularités dans les signaux, la migration, la sélection et l'activation, ainsi que des niveaux élevés de cytokines stimulant les cellules B.
Auto-anticorps	Des auto-anticorps pathogènes causent des dégâts tissulaires directement et via la formation de complexes immuns.

Tableau 3. Voies immunitaires clés impliquées dans la pathogénicité du LED (D'après Conelly et al. [141]*)*

En conclusion, l'étiologie du LED est multifactorielle et implique une interaction complexe entre facteurs génétiques et environnementaux. Cette interaction perturbe le système immunitaire et conduit à une réponse inappropriée, caractérisée par la production d'auto-anticorps, la dérégulation des cytokines et la formation de complexes immuns. Ces éléments contribuent à l'inflammation chronique, aux lésions tissulaires et aux symptômes du LED.

II.4.3. Isoflavones et LED

Il est admis, aujourd'hui, que l'E_2 aggrave la symptomatologie du LED, en particulier les lésions rénales, le nombre d'auto-anticorps ainsi que l'auto-réactivité des cellules B [145]. L'effets des phytoE, en revanche, est plus ambigüe.

De nombreuses études *in vivo* décrivent chez l'animal modèle, des effets aggravants des lésions rénales lors de l'exposition aux phytoE, notamment une accélération de la protéinurie, une élévation des concentration sériques de créatinine et une réduction de la clairance de cette dernière [136,137,148].

Quelques autres, exclusivement sur des modèles murins, rapportent une amélioration des symptômes du LED après exposition à des doses alimentaires de phytoE, notamment 20 mg d'IFs /kg/jour [149], 20 mg de DAI/kg/jour et une supplémentation de 0,01% en Coumestrol [140].

Une hypothèse possible pour expliquer ces disparités réside dans l'affinité plus élevée des phytoE, en particulier les IFs, pour le récepteur ERβ. Ce dernier pourrait médier l'amélioration de la sévérité du LED, contrairement au récepteur ERα dont les agonistes aggravent la maladie humaine [140].

Il est crucial de souligner l'absence flagrante d'études épidémiologiques ou cliniques susceptibles d'évaluer l'effet potentiel des IFs sur la maladie.

II.5. Cancer du sein triple négatif (TNBC)

II.5.1. Le cancer

Le cancer est défini comme la prolifération incontrôlée de cellules et est associée à certaines des principales causes de décès dans le monde. Il était la cause d'environ 10 millions de décès en 2020 alors que son incidence, la même année, était d'environ 19,3 millions de nouveaux cas [150].

Chez les hommes, la principale incidence du cancer par pays montre une certaine hétérogénéité (Figure 4 A). Cependant, chez les femmes, à l'exception de quelques rares pays, le cancer du sein prévaut comme le montre Sung *et al.* [150] sur la Figure 4 C.

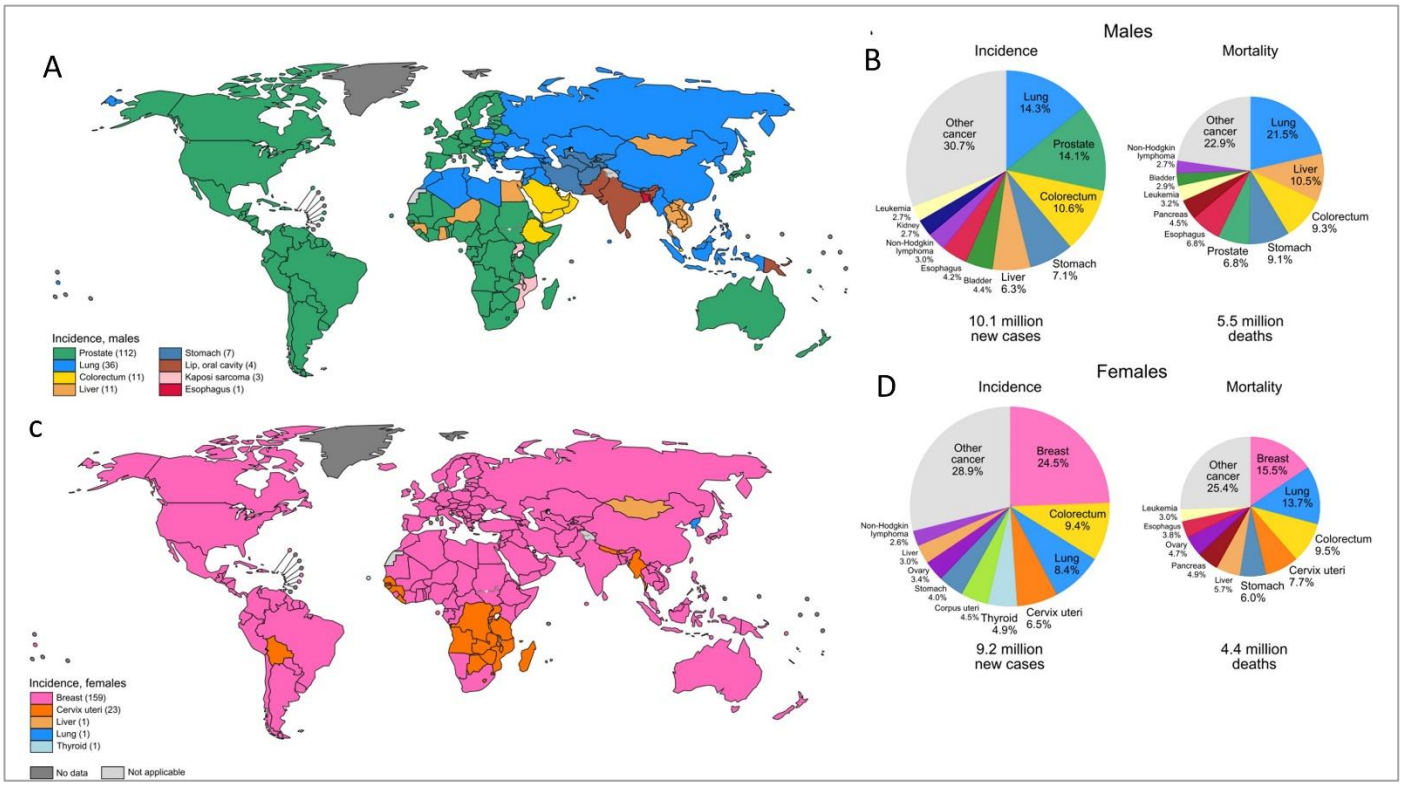


Figure 4. Incidence mondiale des cancers en 2020 chez (A) les hommes et (C) les femmes. Distribution des cas et décès pour les 10 cancers les plus communs en 2020 chez les hommes (B) et les femmes (D) (Sung et al. [150]*).*

Selon les estimations de GLOBOCAN pour 2020, le cancer du sein a supplanté le cancer du poumon et du côlon en tant que cancer le plus diagnostiqué à l'échelle mondiale, avec 2,3 millions de cas recensés, soit 11,7% de l'ensemble des nouveaux cas de cancer. Le Centre international de recherche sur le cancer (CIRC) de l'Organisation mondiale de la santé (OMS) a enregistré 685 000 décès imputables au cancer du sein en 2020. Le CIRC prévoit une augmentation considérable de ces chiffres, atteignant 3 millions de nouveaux cas et 1 million de décès d'ici 2040. Il est crucial de souligner l'augmentation rapide de l'incidence du cancer du sein dans les pays en transition, où les taux de mortalité associés à cette maladie demeurent élevés [42,150].

II.5.2. Evolution

Bien qu'avec de légères différences, l'évolution de la plupart des cancers peut être classés en 4 stades (de I à IV), un stade 0 peut également être décrit dans certains cas [151,152].

- Stade 0 : lésion précancéreuse ou début de cancer localisé, ou *in-situ*, les cellules anormales restent confinées à leur site d'origine, et sont donc facilement traitable par ablation chirurgicale.
- Stade I : cancer à un stade précoce limité en générale à l'organe d'origine, mais pouvant commencer à envahir les tissus voisins. La tumeur est généralement petite et peu agressive ; les cellules cancéreuses n'ont pas atteint de ganglion. Les options de traitement incluent la chirurgie, la radiothérapie ou la chimiothérapie.
- Stade II : croissance tumorale plus importante et une invasion plus profonde des tissus environnants. Les ganglions lymphatiques adjacents peuvent également être touchés. La combinaison de la chirurgie, de la radiothérapie et de la chimiothérapie est souvent nécessaire pour traiter ces cancers.
- Stade III : cancer dont le développement s'est propagé dans les tissus avoisinants, et aux ganglions lymphatiques régionaux. Les tumeurs sont plus volumineuses et plus agressives. La chirurgie, la radiothérapie et la chimiothérapie sont souvent utilisées en combinaison comme traitement.
- Stade IV : cancer métastatique ou de stade avancé. Le cancer s'est propagé à des organes distants (métastases). Les options de traitement visent à ralentir la progression, à soulager les symptômes et à améliorer la qualité de vie. La chimiothérapie, l'immunothérapie, les thérapies ciblées et la radiothérapie sont couramment utilisées, en combinaison.

Il est essentiel de noter que chaque type de cancer peut avoir ses propres systèmes de classification et de stadification, notamment, la classification TNM (Tumeur, Ganglions lymphatiques, Métastases) ou la classification FIGO (pour les tumeurs gynécologiques) [153].

II.5.3. Les cancers du sein

Chez les femmes, le cancer du sein représente environ 25% de tous les nouveaux cas de cancer [42,150] et 60 à 75% de ces cancers sont hormonodépendants [154,155].

II.5.3.1. Classification

La classification des cancers du sein repose sur plusieurs approches complémentaires [156]. En anatomopathologie, on distingue les carcinomes canalaires ou lobulaires localisés (*in situ*) et infiltrants, selon que les cellules cancéreuses restent confinées à leur site d'origine ou envahissent les tissus adjacents. La classification TNM définit le stade des tumeurs par rapport à leur taille (T), l'atteinte des ganglions (N) et la présence de métastases (M).

D'un point de vue clinique, une classification basée sur la sensibilité aux hormones était historiquement utilisée. Cette approche distinguait deux catégories : les cancers sensibles ou non aux estrogènes. Mais la mise en évidence du GPER a conduit à une révision de cette classification [156].

De nos jours, on distingue 3 sous-types moléculaires principaux de cancers du sein, établis à partir du matériel génétique des tumeurs [157] :

- luminaux A et B caractérisés par l'expression du récepteur estrogène ER+ et de gènes associés à la voie ER ;
- HER2-like caractérisés par soit une surexpression du gène HER2, protéine appartenant à la superfamille des récepteurs aux facteurs de croissance épidermiques ou *Epidermal Growth Factor* (EGF) soit une expression des récepteurs aux estrogènes et de la progestérone.
- phénotype basal dits Triples Négatifs (TNBC) définis par l'absence de récepteurs des estrogènes ou de la progestérone et d'amplification du gène HER2.

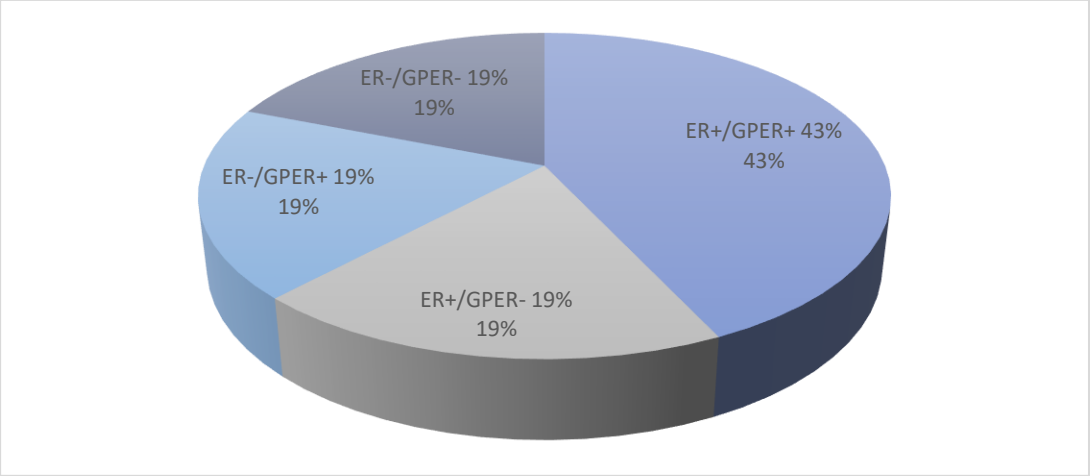


Figure 5. Distribution des cancers du sein selon la présence des récepteurs aux estrogènes (ER), progestérone (PR) et GPER dans des spécimens de biopsie (D'après Hall et Fillardo [42]*)*

Environs 20% des cancers du sein expriment le GPER sans exprimer ER (Figure 5), dont une grande majorité est TNBC [158]. Comparés aux autres types de cancers du sein, les TNBC sont associés à un âge de survenue plus précoce (pendant lequel la synthèse ovarienne d'E_2 est toujours en place), une agressivité accrue, des taux de rechute plus élevés et un pronostic globalement plus défavorable [159]. Malheureusement, l'évaluation du statut GPER des patientes n'est pas encore systématiquement réalisé en clinique.

II.5.3.2. Prévalence et facteurs de risque

De nombreux facteurs de risque sont associés à la genèse et au développement du cancer du sein. Parmi les plus largement étudiés, nous pouvons citer les suivants [160,161].

- **Facteurs hormonaux** : l'hormonothérapie ménopausique, en particulier la combinaison d'estrogènes et de progestérone, l'âge précoce des premières menstruations (< 12 ans) et la ménopause tardive (> 55 ans) sont associés au risque de cancer du sein. L'exposition prolongée aux hormones estrogènes et progestérone favorise la prolifération cellulaire et peut contribuer à la carcinogenèse.
- **Facteurs génétiques** : les porteuses des mutations des gènes BrCA1 et BrCA2 (qui participent à la réparation des lésions que l'ADN) ont un risque significativement accru de développer un cancer du sein et de l'ovaire [162].
- **Antécédents familiaux** : les antécédents de cancer du sein chez la mère et/ou la sœur augmentent le risque. Une histoire familiale positive est un indicateur important pour évaluer la susceptibilité individuelle.
- **Antécédents personnels** : les femmes ayant déjà eu un cancer de l'ovaire ou de l'endomètre présentent un risque élevé de cancer du sein.
- **Nulliparité** (absence de grossesse) ou première grossesse tardive (> 35 ans).

- Obésité : en particulier après la ménopause.
- Diabète de type 2 : il augmente le risque de cancer du sein et de décès par cette pathologie.
- Consommation d'alcool : même modérée, elle augmente le risque de survenue de ce cancer [163].

A l'opposé, l'allaitement est un facteur protecteur vis-à-vis du cancer du sein. Une pratique régulière d'activité physique réduit également le risque de cancer [164].

Il est important de noter que le risque de cancer du sein est multifactoriel et que la présence d'un ou plusieurs facteurs de risque ne signifie pas nécessairement qu'une femme développera un cancer. De même, l'absence de facteurs de risque ne garantit pas l'absence de cancer.

II.5.3.3. Physiopathologie

II.5.3.3.1. Carcinogenèse

Les principales étapes de la carcinogenèse du sein sont les suivantes [165].

- Initiation : altération de l'information génétique contenue dans l'ADN, conférant à la cellule la capacité d'échapper aux mécanismes de contrôle et de réparation naturels. Transmise aux cellules filles, elle peut avoir différentes origines : embryonnaire, héréditaire ou induite par des agents exogènes comme le tabac, l'alcool, les rayons ionisants, etc. L'impact des agents initiateurs est irréversible, modifiant de manière permanente l'information génétique de la cellule. En revanche, les agents promoteurs ne font qu'encourager la croissance et la prolifération des cellules déjà porteuses de mutations initiales. C'est pourquoi l'on qualifie la phase d'initiation de réversible, tant que l'altération n'a pas été fixée par un agent initiateur.
- Promotion : la prolifération des cellules initiées, lesquelles bien que porteuses de mutations génétiques, ne sont pas encore considérées comme cancéreuses à ce stade. Plusieurs agents comme les facteurs de croissance et cytokines, y contribuent. Cette phase est longue et réversible. C'est l'acquisition de nouvelles fonctionnalités qui les définit comme cancéreuses : prolifération autonome, échappement à la régulation, invasion des tissus voisins, angiogenèse, réplication illimitée, résistance à la mort cellulaire, dérégulation du métabolisme énergétique et échappement au système immunitaire.
- Progression : mise en place d'un environnement favorable à l'installation et au développement des cellules cancéreuses appelé micro-environnement tumoral. Un élément crucial de cette phase complexe est la néo-angiogenèse tumorale c'est à dire la création de nouveaux vaisseaux sanguins pour soutenir la croissance et la dissémination des cellules cancéreuses. Ces dernières acquièrent également des capacités d'invasion et de migration. Elles peuvent ainsi circuler dans le sang ou la lymphe, échappant au système immunitaire et colonisant d'autres organes.
- Métastases : migration et fixation de cellules cancéreuses vers d'autres sites métastatiques qui ne sont pas aléatoires. La migration se fait généralement vers les ganglions lymphatiques du même côté que la tumeur, ainsi qu'aux régions adjacentes comme l'aisselle, la clavicule et le sternum. Les principaux sites de métastases à distance sont le foie, les os et les poumons.

II.5.3.3.2. Formes anatomopathologiques

Le cancer du sein se présente sous différentes formes anatomopathologiques, chacune ayant ses propres caractéristiques et son pronostic associé. Les principales formes sont présentées dans cette section [165].

- Carcinome *in situ* : les cellules cancéreuses se limitent aux canaux ou aux lobules mammaires et n'ont pas franchi la membrane basale. À ce stade, il n'y a pas de métastases possibles. Il s'agit d'une lésion précancéreuse confinée à son site d'origine et généralement associée à un bon pronostic.
- Carcinomes infiltrants : le tissu mammaire avoisinant est envahi avec la capacité de se métastaser. Ces tumeurs plus agressives représentent la majorité des cancers du sein diagnostiqués et sont associées à un pronostic plus variable en fonction de l'étendue de l'invasion et d'autres facteurs
- Carcinomes inflammatoires : des phénomènes inflammatoires qui se manifestent cliniquement par une rougeur, une chaleur et un gonflement du sein. Il s'agit d'adénocarcinomes de mauvais pronostic. Leur diagnostic précoce est crucial pour une prise en charge efficace.
- Maladie de Paget du mamelon : un aspect eczématiforme du mamelon et la présence de cellules cancéreuses dans et autour de lui sont caractéristiques de cette forme rare de cancer du sein.
- Sarcome mammaire : le tissu conjonctif mammaire est affecté. Ce type de cancer, rare mais qui peut apparaitre à la naissance, est généralement plus agressif que les carcinomes et est associé à un pronostic plus sombre.

II.5.3.4. Résistance aux traitements

L'absence d'expression de ERα dans la plupart des cancers de type II a suscité l'hypothèse que ces tumeurs pourraient être « indépendantes des estrogènes », alors que les traitements anti-estrogènes sont largement utilisés. Ces derniers se divisent en deux catégories principales :

- modulateurs sélectifs des récepteurs aux estrogènes (SERM) comme le tamoxifène (Tam) et le raloxifène ;
- régulateurs négatifs sélectifs des récepteurs aux estrogènes (SERD) comme l'ICI 182 780 ou fulvestrant.

Cette particularité biologique des tumeurs de type II suggère que les traitements par anti-estrogènes, qu'il s'agisse de SERM ou de SERD, pourraient être inefficaces. Cette supposition est corroborée par un ensemble d'essais cliniques qui ont montré une absence de réponse significative au traitement par Tam chez les patientes atteintes de tumeurs de type II ERα-négatives [166].

En fait, le traitement prolongé par des SERM peut augmenter l'incidence du cancer de l'endomètre, en particulier des types histologiques à haut risque, réduisant ainsi significativement la survie globale [166].

Le GPER joue un rôle dans l'hormono-résistance du cancer du sein et dans l'augmentation du risque de celui de l'endomètre chez les femmes prenant du Tam [166]. Le Tam induit la prolifération *in vitro* des cellules MCF-7 résistantes au Tam par une voie dépendante du GPER. Cet effet prolifératif a été bloqué par le *knockdown* du GPER ou par le co-traitement avec l'antagoniste sélectif du GPER, G15. Le Tam induit également, par une voie GPER-dépendante, la prolifération (ou l'activation de facteurs prolifératifs) d'autres types de cellules, y compris les cellules cancéreuses de l'endomètre et de la thyroïde [167].

L'augmentation de la signalisation des facteurs de croissance, notamment EGF, peut également jouer un rôle important dans le développement de la résistance au Tam. Les facteurs de croissance et leurs voies de signalisation intracellulaires peuvent stimuler

l'activité des ERs en l'absence d'un ligand, ce que l'on appelle l'activation « indépendante du ligand ». Il a été proposé que ce phénomène soit associé à la surexpression d'EGFR et d'HER2 et soit particulièrement médié par l'activation de la voie de signalisation MAPK par les estrogènes [168].

De plus, le traitement à long terme avec Tam augmente la sensibilité des cellules à l'E_2 et conduit à une augmentation de la synthèse du GPER stimulée par l'E_2. Le Tam accroit aussi le recrutement et la translocation du GPER du réticulum endoplasmique vers la surface cellulaire. La surexpression de GPER transforme Tam en stimulateur de croissance par sa capacité à agir en tant qu'agoniste de ce récepteur [169]. Le dialogue entre GPER et EGFR est induit par la phosphorylation de MAPK et Akt. Ces derniers peuvent stimuler la transcription de différents gènes, conduisant à la prolifération cellulaire. Le blocage de la signalisation GPER / EGFR est une cible potentielle pour surmonter la résistance au Tam dans les cellules cancéreuses du sein [168].

Un mécanisme similaire de résistance au Tam a été proposé pour l'ERα membranaire [170].

En outre, l'existence d'une corrélation négative entre l'expression de l'ER et de l'EGFR/HER2 dans le cancer du sein remet en question le rôle potentiel de la communication croisée entre l'ERα membranaire et l'EGFR dans les mécanismes de résistance au Tam. En effet, l'ERα membranaire semble agir comme un répresseur de l'EGFR. En revanche, des résultats récents ont montré que l'expression du GPER pourrait être corrélée positivement avec l'expression de HER-2 [168].

En résumé, la résistance au traitement hormonal du cancer du sein est un phénomène complexe et multifactoriel. De nombreux mécanismes y contribuent, dont l'activation des ERs « indépendante du ligand », l'augmentation de la signalisation des facteurs de croissance et l'expression de GPER.

II.5.4. GPER

II.5.4.1. Données générales

Le GPER est un récepteur aux estrogènes faisant partie de la grande famille des récepteurs liés à la protéine G (GPCR). Il est constitué de 7 domaines transmembranaires. La mise en évidence de l'existence du GPER, de sa liaison aux estrogènes et ses effets non-génomiques rapides remontent à une 20^{aine} d'années [42].

Les antagonistes habituels de l'ER (Tam, fulvestrant, ...) se sont révélés être des agonistes du GPER. Ainsi, la résistance de certaines lignées cancéreuses du sein au Tam pourrait être expliquée par l'expression de GPER [171], comme mentionné précédemment.

II.5.4.2. Distribution

La présence du GPER est avant tout liée à la membrane cytoplasmique. Pour autant, elle a également été décrite dans plusieurs autres localisations subcellulaires, y compris dans les compartiments membranaires internes [172], tels que le réticulum endoplasmique et l'appareil de Golgi. Une localisation nucléaire a également été rapportée en tant que protéine de liaison de la chromatine [9].

Pupo *et al.* [173] ont mis au jour un polymorphisme nucléotidique unique relativement fréquent qui conduit à l'expression d'un variant du GPER associé à une localisation nucléaire. En effet, le GPER avec la mutation P16L ne subit pas de glycosylation, vraisemblablement en raison d'un effet conformationnel sur les sites de glycosylation

avoisinants. Par conséquent, le GPER P16L perd sa fonction de signalisation membranaire mais agit en revanche comme un facteur de transcription stimulé par les estrogènes.

La localisation intracellulaire du GPER, notamment dans le réticulum endoplasmique et l'appareil de Golgi, suggère qu'il s'agit d'étapes intermédiaires de la synthèse protéique avant son arrivée à sa destination finale, la membrane cellulaire. Au sein de la membrane plasmique, le GPER forme un complexe avec une guanylate kinase associée à la membrane et une protéine d'ancrage de la kinase 5, ce qui atténue la production d'AMP cyclique (AMPc) et maintient le récepteur à la membrane plasmique [17].

Certain·es auteur·es ont identifié le GPER dans la membrane mitochondriale et l'ont associé à la régulation de l'activité mitochondriale par l'E_2 [174]. Cette localisation semblerait être impliquée dans le rôle des estrogènes sur la reperfusion tissulaire après une ischémie.

II.5.4.3. Site actif/agonistes/antagonistes

Le site actif du GPER est une région flexible et adaptable capable de lier une grande variété de ligands. Cette particularité confère au GPER une grande plasticité fonctionnelle et lui permet de jouer un rôle dans de nombreux processus physiologiques et pathologiques.

Des études de *docking* aveugle (ancrage moléculaire sur l'ensemble la protéine) ont montré que les sites actifs des IFs et de l'E_2 sont similaires, avec des affinités de liaison comparables (-8,9 ; -8,8 ; -8,6 et -8,3 kcal/mol pour l'équol, la GEN, la DAI, et l'E_2, respectivement) [7]. Cependant, il est important de noter que dans cette publication d'Ariyani *et al*, le *docking* proposé a été testé dans la partie basse (intracellulaire) du récepteur, sans justification explicite. Cette approche soulève des interrogations quant à la pertinence de cette localisation spécifique pour l'interaction ligand-récepteur.

Des travaux menés par Mendez-Luna *et al*. [175] ont suggéré que les cavités du GPER peuvent accepter des ligands à grand-volume grâce à des interactions π - π, hydrophobes et des liaisons hydrogène. L'analyse des motifs de liaison, basée sur la dynamique moléculaire, a révélé des similitudes entre l'antagoniste G15 et l'agoniste G1, laissant supposer l'implication des résidus Cys107 et Glu275 dans l'activation de la protéine. Les auteur·es ont conclus que le GPER pouvait être sujet à des changements structurels significatifs, expliquant ainsi sa capacité étendue à accepter divers ligands. De plus, leurs travaux soulignent que le même ligand pourrait être reconnu dans différentes poses (positions de liaison) selon les conformations structurelles du récepteur.

Par ailleurs, Périan *et al.* [9] ont indiqué dans leur étude de modélisation et d'ancrage moléculaires que le GPER pouvait accueillir ces ligands volumineux dans plusieurs cavités, ce qui pourrait expliquer le grand nombre de composés capables de s'y lier.

Ces résultats mettent en lumière la complexité des interactions ligand-récepteur et soulignent l'importance de considérer la conformation tridimensionnelle du GPER pour mieux comprendre sa spécificité de liaison aux ligands et son rôle dans la signalisation cellulaire.

II.5.4.4. Fonctions physiologiques

Le GPER est un récepteur membranaire qui joue un rôle crucial dans divers aspects de la signalisation cellulaire, y compris la libération de calcium, la transactivation de l'EGFR et la régulation génique.

Il est impliqué dans la physiologie dépendante de l'E_2 des cellules mammaires épithéliales non cancéreuses. Des études ont montré qu'il régule la prolifération, la migration et l'apoptose de ces cellules [159].

Il est également exprimé dans d'autres tissus et organes, notamment l'utérus, les ovaires, le cerveau, les os et le système cardiovasculaire et donc impliqué dans une variété de fonctions physiologiques, y compris la reproduction, le métabolisme, la cognition et la santé des os [176,177].

L'un des défis dans l'étude du GPER est son chevauchement fonctionnel et de spécificité de ligand avec les ERα / β [178]. Le 17β-E_2, le 4-hydroxyTam et l'ICI 182 780 se lient tous au GPER et aux ERα / β, mais avec des effets agonistes et antagonistes distincts. Le 17β-E_2 est un agoniste du GPER et un agoniste ERα, tandis que le 4-hydroxyTam est un SERM et l'ICI 182 780 est un antagoniste pur de l'ERα [178].

Les études *in vivo* menées sur des souris *knock down* pour le GPER montrent qu'il serait non essentiel à la reproduction, en contraste avec des résultats de pharmacologie indiquant un rôle dans la prolifération des cellules utérines épithéliales. D'autres travaux *in vivo* ont montré une implication du GPER dans le métabolisme du glucose et des lipides, l'accrétion osseuse, la pigmentation de la peau ainsi que dans la régulation de la fréquence cardiaque et des systèmes nerveux et immunitaire [9].

Certains travaux notent que l'affinité du GPER pour l'E_2 *in vitro* est environ 10 à 60 fois inférieure à celle de l'ERα [179]. Cependant, comme l'ont rappelé Rouhimoghadam *et al.* [158], les affinités de liaison relatives (RBA) de l'ERα, de l'ERβ et du GPER ne peuvent être directement comparées. Cela s'explique par le fait que ces récepteurs présentent des niveaux d'expression et des environnements physico-chimiques distincts. Les ERs sont isolés dans des homogénats cytosoliques sans détergent, tandis que le GPER est enrichi dans des préparations de membrane plasmique riches en lipides. Par conséquent, la valeur de constante de dissociation (Kd) plus faible mesurée pour l'E_2 dans les tests de liaison des ERs par rapport aux tests de liaison du GPER ne signifie pas forcément que l'E_2 a une affinité plus élevée pour les ERs que pour le GPER.

De plus, la position membranaire du GPER pourrait lui permettre d'accéder à des quantités de 17β-E_2 supérieures à celles atteintes par les ERs intracellulaires. Cependant, il est également possible que l'action du 17β-E_2 sur le GPER ne soit déclenchée qu'en présence d'une concentration élevé d'hormone, dans des conditions physiologiques particulières telles que la période pré-ovulatoire chez les femmes ou la grossesse [179]. Cette question mérite d'être explorée plus en profondeur.

Par ailleurs, GPER semble jouer un rôle dans les cancers hormono-dépendants. Les travaux existants [9] ont notamment montré une corrélation positive entre l'expression de GPER et les pronostics des cancers du sein et de l'endomètre. Il parait en outre être impliqué dans la promotion de caractères cancéreux tels que la migration de TNBC, la résistance à l'hypoxie, la réponse pro-angiogénique et l'activation des CAF (Fibroblastes associés aux cancers).

II.5.4.5. Implication dans les cancers du sein

Le récepteur des estrogènes GPER est présent dans environ 50 à 60% des cas de cancer du sein. Dans la plupart des cellules cancéreuses mammaires, l'expression du GPER est nettement plus élevée par rapport aux tissus sains [180].

Nous rappelons que les TNBC représentent 15 à 20% des cancers du sein et se distinguent par l'absence d'ER, de récepteurs aux progestatifs (PR) et à HER2 [171]. Contrairement aux autres types de cancer du sein, ils ne répondent pas aux traitements hormonaux classiques [42]. Autre rappel : le GPER est exprimé dans la majorité des TNBC et serait associé à un âge plus précoce au diagnostic, à des tumeurs plus grandes et à un stade clinique plus avancé [181]

Le GPER serait impliqué dans la cancérogénèse à travers plusieurs mécanismes [159] succinctement décrits ici.

- Action sur la **prolifération** : via l'activation et la régulation positive de certains gènes cibles.
- Action sur la **migration** de cellules cancéreuses mammaires : via l'activation du GPER et potentiellement à travers des interactions avec des facteurs tels que CTGF, Cycline E, voie notch et CXC-R-1.
- Promotion de **l'invasion** de cellules **inflammatoires** cancéreuses ainsi que l'adhésion des cellules cancéreuses aux cellules endothéliales, ce qui suggère une implication dans le processus **métastatique**.
- Action sur **l'angiogenèse** induite par hypoxie : l'expression du GPER est régulée positivement par l'hypoxie et peut contribuer à l'angiogenèse tumorale.

Des études ont montré que le GPER était responsable de la prolifération, de l'invasion et de la migration de cellules cancéreuses mammaires MCF-7 traitées à l'E_2 à des doses physiologiques. De plus, l'E_2 favoriserait l'expression du GPER, créant une boucle de rétroaction positive qui stimule la croissance [182].

Ces découvertes soulignent l'importance du GPER dans le contexte du cancer du sein et ouvrent des perspectives pour de futures recherches et des approches thérapeutiques ciblées [171].

Actuellement, toutefois, les rôles du GPER dans le cancer font l'objet d'un débat croissant, avec des études rapportant des effets pro- et anti-tumoraux. Cette hétérogénéité de résultats (voir Tableau 4) peut s'expliquer par plusieurs mécanismes non exclusifs.

- **Localisation subcellulaire** du GPER : la distribution cellulaire du GPER varierait en fonction des espèces, des tissus et des types de cellules et pourrait influencer ses fonctions et le pronostic. Des études récentes suggèrent que le GPER nucléaire pourrait avoir des fonctions distinctes de celles du GPER membranaire, notamment dans la régulation de la prolifération cellulaire, de la migration et de l'apoptose [42]. Par exemple, une faible expression nucléaire du GPER ne semble pas être corrélée à l'agressivité du cancer du sein.
- **Contexte tissulaire** : les fonctions du GPER varient selon les types de cellules et de tissus où elles sont observées. Dans le pancréas et le foie, les effets anti-inflammatoires du GPER, dans les cellules non cancéreuses, pourraient contrebalancer sa capacité à stimuler la croissance tumorale. En revanche, dans le cancer du sein, le GPER semble jouer un rôle pro-tumoral, favorisant la prolifération, la migration et l'invasion cellulaire [167].
- **Stade du développement tumoral** : le rôle du GPER peut évoluer au cours de la progression tumorale. Dans un modèle murin de cancer du sein, le GPER ne semble pas jouer de rôle dans l'initiation du cancer, mais contribue à la formation de métastases. De plus, l'expression du GPER peut être corrélée à différents stades de la maladie. Par exemple, une surexpression du GPER a été observée dans les tumeurs de grande taille (>2 cm) et métastatiques [183].

- **Interaction avec d'autres facteurs** : le GPER pourrait avoir des rôles différents en fonction de l'expression de facteurs interagissant. Par exemple, il stimule la croissance des cellules SKBR3 ER(-), mais inhibe celle des cellules MCF7 ER(+). De plus, son effet stimulant sur les cellules cancéreuses ovariennes dépend de l'EGFR. L'effet inhibiteur du GPER sur des cellules ER(+) pourrait provenir d'une confusion possible entre l'effet GPER et l'effet ERα membranaire.
- **Réponses calciques différentielles** : Une hypothèse récente propose que les effets du GPER sur les cellules cancéreuses ER(+) et ER(-) pourraient être liés à des réponses calciques transitoires ou prolongées et impliquant des canaux calciques différents [184]. Dans leur étude, Ariazi *et al.* [184] ont rapporté que E_2 et G1, à une concentration de 1 µM, induisaient une augmentation de la concentration intracellulaire de calcium dans les cellules SKBR3 (ER-) plus importante et transitoire que dans les cellules MCF7 (ER+).

En conclusion, les effets du GPER dans le cancer sont complexes et multidimensionnels. La compréhension des mécanismes sous-jacents à ses actions pro- et anti-tumorales est essentielle pour le développement de nouvelles thérapies ciblées [9].

II.5.4.6. Mécanismes d'action

Le GPER est un récepteur membranaire qui lie les estrogènes et d'autres stéroïdes sexuels. Contrairement aux récepteurs d'estrogènes nucléaires classiques (ERα et ERβ), qui agissent directement comme facteurs de transcription, le GPER initie des cascades de signalisation intracellulaires rapides et variées. Ses effets sur la prolifération cellulaire, la survie et la migration dépendent du contexte tumoral et de la concentration de ligand.

II.5.4.6.1. Activation du GPER et voies de signalisation

La liaison de ligands, tels que l'E_2 ou les xéno-estrogènes, au GPER déclenche l'activation de diverses voies de signalisation (voir Figure 6), notamment ceux présentés ci-dessous [42,168,172,180].

- Voie du **récepteur EGFR** : L'activation du GPER entraîne la libération du facteur de croissance épidermique lié à l'héparane (HB-EGF) via l'activation de la Src kinase et des métalloprotéinases matricielles (MMPs). L' HB-EGF active à son tour l'**EGFR** (*Epidermal Growth Factor Receptor*). Cette transactivation déclenche les voies de signalisation **MAPK** et **PI3K/Akt**, impliquées dans la prolifération, la survie et la migration des cellules cancéreuses. Des études sur les lignées cellulaires de cancer du sein ont démontré l'importance de cette voie dans la réponse au GPER. Dans la lignée SKBR3 (ERα et β négatives), une multiplication par 6 de la phosphorylation **ERK** a été rapportée suite à 5 min d'exposition à l'E_2. Dans la lignée MDA-MB-231 GPER (-), il n'y a pas d'augmentation de la phosphorylation d'ERK après traitement par EGFR mais la transfection de ces cellules avec le gène du GPER permet l'augmentation de cette phosphorylation en impliquant la **sous-unité βγ** [172].
- Voie du **calcium** : L'activation du GPER suite à l'action de l'E_2 induit une augmentation rapide du calcium intracellulaire, ainsi que la mobilisation des stocks cellulaires de calcium [172], contribuant potentiellement à la prolifération cellulaire et à la migration. La sous-unité Gα active la **phospholipase C** (PLC), qui augmente la concentration intracellulaire de calcium et active d'autres enzymes et voies de signalisation. Cependant, Xu *el al.* [180] ont observé que l'activation du GPER par G1 à une concentration inférieure à 1 µM, induisait une mobilisation du calcium intracellulaire découlant, de manière inattendue, de la transactivation de l'EGFR et non de l'activation de la voie Gα classique. Le dimère libre de la sous-unité Gβγ joue également un rôle en déclenchant la libération de calcium intracellulaire. Cela entraîne le clivage protéolytique dépendant de la kinase de la

famille Src Shc et la libération consécutive de HB-EGF de la face exoplasmique de la membrane plasmique [42]. Il a été observé que G1 à une concentration de 1 µM induirait l'action de deux canaux calciques différents : dans les cellules MCF-7, il agirait par l'intermédiaire des récepteurs de l'inositol triphosphate (IP3R), alors qu'il agirait par l'intermédiaire des récepteurs de la ryanodine (RyR) dans les cellules SKBR3 [184].

- Voie de l'**AMPc** : Dans certaines circonstances, en particulier à des concentrations pharmacologiques de ligand, par exemple G1 à une concentration supérieure à 1 µM, le GPER lié à son ligand active la **sous-unité Gα**, ce qui entraîne l'activation de **l'adénylate cyclase** (AC) et l'augmentation de la production d'**AMPc.** Cette cascade de signalisation conduit à l'activation de la **protéine kinase A** (PKA) et la CREB, ce qui influence la transcription de gènes impliqués dans la progression du cycle cellulaire et son arrêt.

En plus de ces voies bien caractérisées, le GPER peut également influencer les voies IκB, Hippo et HOTAIR, contribuant à la complexité de ses effets dans le cancer du sein.

Les voies de signalisation intracellulaires du GPER favorisant la carcinogenèse induite par les estrogènes, induisent par ailleurs une série d'effets [42] regroupés ici :

- perte de polarité des cellules épithéliales et invasion de la membrane basale associée à la transition épithélio-mésenchymateuse (EMT) ;
- invasion des tissus locaux par l'activation des MMPs et d'autres protéinases de la surface cellulaire ;
- échappement à l'apoptose et la survie en dehors des limites de la membrane basale ;
- synthèse d'une matrice provisoire de fibronectine ;
- production et libération de facteurs solubles qui modifient l'immunité locale et périphérique ;
- activation de la signalisation dans un environnement hypoxique et pour activer les CAF qui conduisent au développement d'un stroma cancéreux réactif ;
- activation de signaux qui favorisent le recrutement d'un apport vasculaire pour réoxygéner le tissu « blessé », ce qui peut finalement entraîner une dissémination systémique des cellules tumorales.

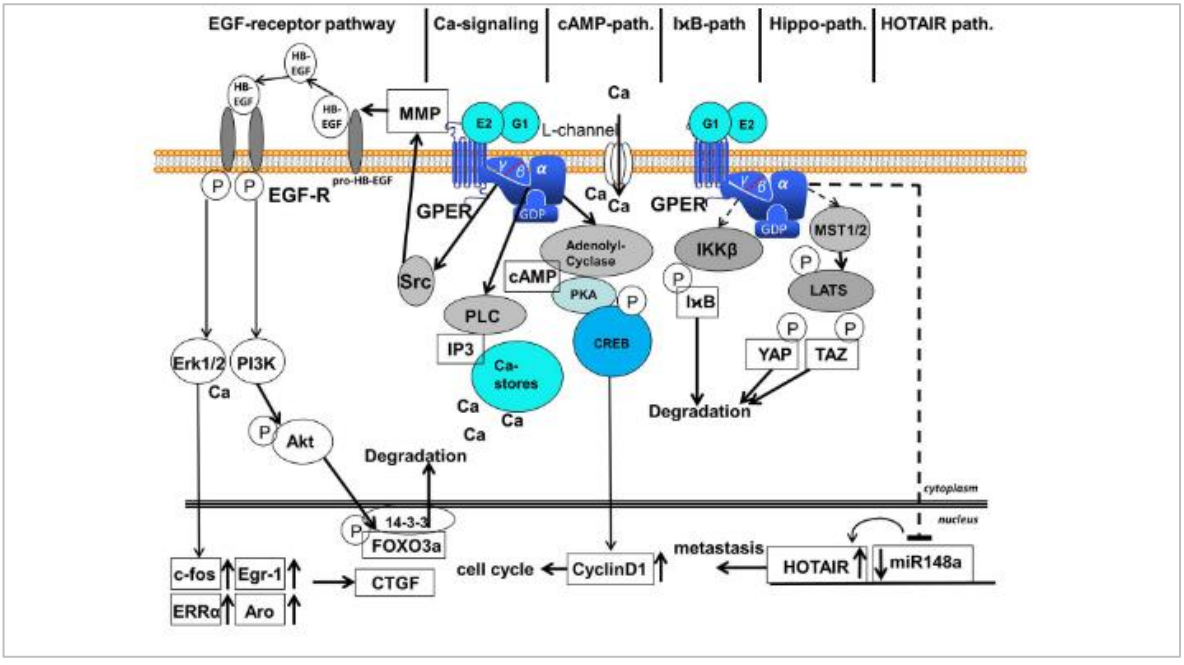


Figure 6. Voies de signalisation du GPER, d'après Girgert et al. [171]

Les travaux de Filardo *et al.* [185], ont apporté un éclairage important sur le mécanisme par lequel l'activité des ERK-1/2 est rapidement restaurée aux niveaux basaux après stimulation par les estrogènes. Les auteur·es ont proposé que cette augmentation puis atténuation de l'activité d'ERK-1/2 se produirait via une double signalisation dépendante du GPER, aux effets opposés. D'abord, l'activation transitoire et rapide de la voie EGF et des MAPK comme exposé précédemment. Ensuite, la stimulation de l'AC par le GPER et la signalisation dépendante de l'AMPc. Ce résultat a été rapporté après exposition de cellules MCF-7 GPER et ERs (+) et SKBR3 GPER (+) et ERs (-) à 1 nM d'E_2. Cependant, l'utilisation d'un milieu sans sérum dans leur étude a probablement permis d'augmenter considérablement la quantité d'E_2 disponible, augmentant ainsi la probabilité d'une activation de la voie de l'AMPc par l'E_2.

Ces résultats démontrent la complexité de la régulation de l'activité d'ERK-1/2 par les estrogènes et mettent en lumière le rôle crucial du récepteur GPER dans ce processus. Ils suggèrent que l'E_2 peut moduler la voie Raf-vers-ERK et la voie EGF de manière transitoire et opposée, permettant un contrôle fin de la prolifération cellulaire et de la différenciation [185]. Cependant, il faut noter qu'au sein d'un organisme et contrairement aux observations *in vitro*, les concentrations d'E_2 ne présentent pas de stabilité et fluctuent continuellement en fonction des rythmes de synthèse, de conjugaison, de transport et d'élimination hormonaux.

II.5.4.6.2. Approches pharmacologiques

Des approches pharmacologiques utilisant des agonistes spécifiques du GPER, comme le G1, et des antagonistes, comme le G36 et G15, permettent de mieux comprendre son rôle dans la prolifération cellulaire. Le G36 est un meilleur antagoniste de GPER que le G15. Il est en effet plus spécifique du GPER et n'a qu'une affinité très faible avec les ERα. L'affinité du GPER avec G1 est similaire à celle de E_2 (Kd de 10 et 6 nM, respectivement). Le G1 n'a pas d'affinité avec ER en dessous de 10μM. [180].

L'antagoniste spécifique du GPER G15 présente une affinité 0,5 fois supérieure à celle de G1 et une sélectivité 1 000 fois supérieure à celle d'ERα et d'ERβ [180]. Le G15 se lie au GPER avec une affinité d'environ 20 nM. En comparaison, l'affinité pour le G1, en utilisant le même test, est d'environ 7 nM et est similaire à l'affinité précédemment rapportée du G1 pour le GPER recombinant. Celle-ci serait de 11 nM et les affinités rapportées pour le 17β-estradiol entre 3 et 6 nM. Des études supplémentaires de liaison compétitive pour évaluer les interactions avec ERα et ERβ ont révélé que, comme G1, G15 se lie peu à ERα ou ERβ à des concentrations allant jusqu'à 10 μM, où l'estrogène entre en compétition avec un Ki d'environ 0,3-0,5 nM. Ces résultats révèlent que G15, comme G1, présente une forte affinité pour GPER avec une liaison minimale à ERα et ERβ (Ki > 10 μM) [178].

Effets anti-prolifératifs de G1 dans les cellules cancéreuses

Dans le cancer de la prostate, G1 (1 μM) a montré une capacité à inhiber la prolifération des cellules PC-3 *in vitro* et *in vivo* (xénogreffes chez la souris). Cet effet semble lié à l'arrêt du cycle cellulaire en phase G2 et à la phosphorylation des régulateurs du point de contrôle G2, NF-YA [186]. De même, dans le cancer de l'ovaire, G1 (2 μM) a induit une suppression de la prolifération et une augmentation de l'apoptose en ciblant la tubuline dans les lignées cellulaires IGROV-1 et SKOV-3 [187].

Plusieurs études décrivent un effet suppresseur de la prolifération de G1 à 1 µM sur des cellules cancéreuses mammaires MCF7 (ER+) [184,188] ou TNBC [189–191] (voir Tableau 4). Un effet inhibiteur a également été observé sur des cellules SKBR3 après traitement avec G1 à 10 µM.

Il faut noter que ces doses, qui sont 100 fois plus concentré que la dose active d'E_2 n'ont peut-être pas de pertinence physiologique.

Effets pro-prolifératifs et pro-métastatiques de G1

Par contraste, G1 (100 nM, soit 10 fois plus concentré que la dose active d'E_2, et donc en accord avec les ratio d'affinité des deux substances pour le GPER) a stimulé la croissance des cellules de cancer du sein MCF-7 résistantes ou non au Tam, soulignant son rôle potentiel dans la résistance aux traitements endocriniens [169]. De plus, G1 (100 nM) a augmenté la prolifération et la migration des cellules TNBC [181].

Etrangement, G1 à une concentration de 1µM semble également responsable d'une prolifération et adhésion de CAF dans les travaux de Luo *et al.* [192,193]. Néanmoins, dans le cadre de ces expériences, les cellules étaient synchronisées dans un milieu sans sérum 24 heures avant le traitement. Or, à ce jour, aucune donnée n'est disponible, à notre connaissance, concernant la liaison potentielle du G1 aux protéines sériques. Si une telle liaison était possible, l'absence de sérum induirait une plus grande disponibilité des composés, susceptibles d'agir sur les cellules, comparativement à d'autres études. En outre, ces travaux soulèvent une question relative aux seuils d'activation de la voie AMPc : pourraient-ils différer dans les CAFs ?

L'effet du G1 sur la prolifération et l'apoptose des cellules cancéreuses semble donc dépendre de manière critique de la concentration utilisée. Des concentrations faibles de G1 pourraient stimuler la prolifération cellulaire, tandis que des concentrations élevées pourraient induire l'apoptose. Ce phénomène pourrait s'expliquer par l'activation de différentes voies de signalisation intracellulaires via GPER, sous différentes configurations du récepteur, selon la dose de G1.

La configuration spatiale particulière du GPER induite par une sur-concentration de G1 pourrait potentiellement permettre l'entrée de plusieurs molécules dans le site de liaison du récepteur. Cette hypothèse soulève l'importance de considérer les résultats de la littérature en fonction des doses d'agonistes ou d'antagonistes utilisées. En effet, une transposition des effets *in vitro* à la situation physiologique implique nécessairement de prendre en compte la question de la dose. Il est crucial de s'interroger sur les concentrations physiologiques des ligands du GPER ou sur celles susceptibles d'être atteintes lors d'un traitement médicamenteux.

De plus, l'interaction du GPER avec d'autres récepteurs, tels que les ERs et les récepteurs aux androgènes (AR), peut moduler son action suite à la stimulation par le G1. Cette interaction complexe peut expliquer les résultats parfois contradictoires observés dans la littérature scientifique, où l'effet du G1 sur la croissance tumorale varie selon le type de cancer et la présence ou l'absence de récepteurs stéroïdiens nucléaires dans les cellules cancéreuses.

L'approche thérapeutique ciblant l'inhibition de la croissance des TNBC via le GPER présente une limite majeure au niveau clinique : la nécessité de développer des antagonistes « suicides » qui se lient de manière irréversible au GPER et bloquent son activité en permanence. Cette stratégie médicamenteuse ne peut être envisagée que

si elle est extrêmement précise et sélective, afin d'éviter tout effet délétère sur d'autres organes, en particulier en s'assurant qu'elle ne traverse pas la barrière hématoencéphalique, et ce en raison du rôle essentiel du GPER dans le fonctionnement d'autres tissus, notamment les neurones.

Référence	Modèle cellulaire	Méthode	Localisation cellulaire	Concentrations testées	Commentaire
			Effet promoteur du cancer		
Ignatov 2010 [168]	MCF-7, MCF-7 Tam R[10]	WB[11]	Membrane, RE[12]	E_2 10nM, Tam 1 µM Fulvestrant 1 µM	Effet attendu (voie EGFR)
Ignatov 2011 [168]	TC[13]	IHC [14]	Membrane, cytoplasme, noyau	NA	Effet attendu (voie EGFR)
Mo 2013 [194]	TC, MCF-7 Tam R, Xénogreffe de souris Tam R	IHC, Immuno-fluorescence, WB, RT-PCR[15],	Membrane, cytoplasme	Tam 1 µM	Effet attendu (voie EGFR)
Chen 2014 [195]	MCF-7, SKBR3	WB, qRT-PCR[16]	NA[17]	E_2 10 nM Fulvestrant 5 µM	Effet attendu (voie EGFR)
Catalano 2014 [196]	MCF-7, MCF-7 Tam R, SKBR3, CAF	RT-PCR, WB	NA	E_2 10 nM, Tam 1 µM D4Androstene-dione 0,5 µM	Effet attendu (voie EGFR)
Lappano 2010 [197]	MCF-7, SKBR3	RT-PCR, WB	NA	E_2 1 nM, E_3 1 nM Fulvestrant 1 µM	Effet attendu (voie EGFR)
Girgert 2012 [198]	MDA-MB-435, HCC1806	RT-PCR, WB	NA	E_2 1 et 10 nM 4-OH-Tam 0,1 et 1 µM	Effet attendu (voie EGFR)
Steiman 2013 [199]	TC	IHC	NA	NA	Observation de cohorte. Association à récurrence
Yu 2014 [181]	MDA-MB-463, -436 et -231, MCF-7, TC	IHC, IF[18], RT-PCR, WB	Cytoplasme	E_2 100 nM, G1 100 nM Tam 100 nM	Effet attendu (voie EGFR activée avec100 nM de G1 et de Tam)

[10] R : résistante
[11] WB : Western Blot
[12] RE : réticulum endoplasmique
[13] TC : tissues cancéreux
[14] IHC: Immunohistochimie
[15] PCR : réaction en chaîne par polymérase
[16] qRT-PCR : Reverse Transcriptase PCR quantitative
[17] NA : Non applicable
[18] IF : immunofluorescence

Référence	Modèle cellulaire	Méthode	Localisation cellulaire	Concentrations testées	Commentaire
Zhou 2016 [200]	SKBR3 MDA-MA-231	WB, IF	NA	E_2 1-100 nM 4-OH-Tam 1 µM	Effet attendu (voie EGFR)
Albanito 2008 [201]	SKBR3 BT20	RT-PCR, WB, IF	Intracellulaire	E_2 100 nM EGF 50 ng/mL	Effet attendu (voie EGFR)
Luo 2014 [193], **2016** [192]	CAFs d'échantillons BC primaires	IHC, RT-PCR, WB	Cytoplasme	E_2 10-100 nM G1 1 µM, Tam10 nM	Effet attendu (voie EGFR)
Ignatov 2011 [189]	TC	IHC	Membrane, cytoplasme, noyau	NA	Effet promoteur (si Tam) Effet attendu
Effets multiples					
Ariazi 2010 [184]	Microréseaux d'expression génique de carcinomes mammaires ER(+) et (-), MCF-7, SKBR3	qRT-PCR, imagerie du Ca^{2+}, WB	NA	G1 1 µM, diethylstilbestrol 1 nM, E_2 1-1000 nM, Fulvestrant 1 µM	Effet suppresseur (ER+) et promoteur (ER-) Effet attendu pour ER(-) Co-expression du ER peut modifier la réponse
Effet suppresseur du cancer					
Broselid 2013 [202]	TC ER(+), T47D MCF-7 ± GPER *knock-down,* HEK ± GPER	ICH, WB, IF	NA	NA	GPER est surtout présent dans des cellule ER(+), PR(+). Effet suppresseur (ER+)
Poola 2008 [203]	TC ERα(+) et (-)	qRT-PCR	NA	NA	Statut ménopausique inconnu Absence de validation du modèle cellulaire (expression du GPER non vérifiée, Statut HER2 inconnu) Pertinence du test statistique ?
Kuo 2007 [204]	TC ERα (+) et (-)	qPCR	NA	NA	Statut ménopausique inconnu Statut PR et HER2 inconnu, Pas de vérification de l'expression du GPER Pertinence du test statistique ? Cellules ER(+)/PR(+) majoritaires

Référence	Modèle cellulaire	Méthode	Localisation cellulaire	Concentrations testées	Commentaire
Weißenborn 2014 [188]	MCF-7, SKBR3	RT-PCR, WB, cytométrie de flux, MSP-PCR[19]	NA	G1 1 µM,	Hypothèse questionnable : effets prolifératifs observés avec des agonistes non spécifiques (E_2 et TAM) mais inhibition de la croissance avec des agonistes spécifiques G1
Weißenborn 2014 [205]	MDA-MB-231 et -468	RT-PCR, WB, MSP-PCR	NA	G1 1 µM	Effet attendu G1 à 1 µM active la voie AMPc qui inhibe la voie EGFR.
Chen 2016 [190]	TC (TNBC), MDA-MB-231, xénogreffes de MDA-MB-231 (souris)	IHC, WB, qRT-PCR	NA	G1 1 µM	Effet attendu (idem)
Liang 2016 [191]	MDA-MB-231et -468, Microréseaux de tissus TNBC, Xénogreffe de souris	qRT-PCR, ELISA, *Tissue microarray*, ICH, WB, *scratch test in vitro*	NA	G1 1 µM	Pas de vérification de l'expression du GPER
Okamoto 2016 [206]	SKBR3	qRT-PCR, WB	NA	E_2 10 µM G1 10 µM	Pas de vérification de l'expression du GPER Concentrations testées élevées

Tableau 4. GPER et cancer du sein (complété de Hsu et al [171]*).*

[19] MSP-PCR : PCR spécifique à la méthylation

II.5.4.6.3. Signalisation génomique

Le GPER exerce ses effets transcriptionnels de manière indirecte en utilisant les protéines G hétérotrimériques. Ainsi, comme vu précédemment, les effets du GPER sont initiés par l'activation d'enzymes associées à la membrane plasmique (telles que l'AC et les MMPs) ; ces actions pré-génomiques conduisent finalement à la transactivation des gènes [42]. Cette dernière médiée par GPER a été établie par des expériences qui ont démontré la médiation de la transcription par c-fos, un facteur de transactivation génique ubiquitaire, par le 17β-E_2 ou la GEN dans des cellules de cancer du sein humaines SKBR3 ERs(-), GPER(+) ou dans des cellules humaines MDA-MB-231 ERs(-) suite à la reconstitution de GPER. Une autre preuve de concept démontrant ce mécanisme spécifique du GPER a été apportée par des expériences supplémentaires utilisant l'agoniste sélectif du GPER, G1, et un plasmide rapporteur de c-fos dans le carcinome ovarien humain BG1 ER(+) et GPER(+). Bien que l'ensemble des gènes dont la transcription est médiée par la transactivation de l'EGFR dépendante de GPER n'ait pas encore été entièrement déterminé, il est bien connu que la signalisation de l'EGFR peut conduire à l'activation de facteurs de transcription qui se lient à des éléments de réponse au sérum (SRE) [42]. Une étude récente [207] analysant les transcriptomes de patientes atteintes d'un cancer du sein ER(-) à partir des ensembles de données de The Cancer Genome Atlas (TCGA) et de Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) a donné une vision plus large des signatures génétiques qui peuvent être attribuées à la signalisation du GPER. Cette étude a montré une corrélation dans l'expression de GPER avec des gènes qui s'alignent sur des voies pro-métastatiques.

II.5.4.7. Complexité et controverse

L'étude du GPER est compliquée par plusieurs facteurs émergeants de différentes sources.

- **Interaction avec d'autres récepteurs aux estrogènes** : la signalisation initiée par la membrane (MISS) impliquant d'autres récepteurs aux estrogènes, notamment ERα-membranaire, peut interagir avec les effets du GPER, rendant difficile l'isolement de ses actions spécifiques [180].
- **Effets contextuels** : les effets du GPER sur la prolifération cellulaire et la progression tumorale peuvent varier en fonction du contexte tumoral, de la concentration de ligand et de l'interaction avec d'autres voies de signalisation.
- **Rôle d'oncogène ou de suppresseur de tumeur** : bien que certaines études suggèrent un rôle suppresseur de tumeur pour le GPER, des preuves cliniques et pharmacologiques indiquent plutôt son association avec la progression de la maladie et une survie réduite dans certains cancers.

La « preuve » que GPER est un antioncogène est extrapolée à partir d'études qui montrent que le promoteur de GPER est hyper-méthylé dans certaines tumeurs. Cependant, comme le précisent Hall et Filardo [42], ces études n'ont pas évalué la méthylation globale due à l'instabilité du génome, ni le statut de méthylation d'autres gènes, y compris ER. Bien que le GPER puisse être considéré comme un suppresseur de tumeur en fonction de ces critères, il est important de noter que cela contredit les données cliniques montrant son association avec la progression de la maladie et une survie réduite dans les cancers du sein et gynécologiques. Les preuves pharmacologiques soutenant l'activité antioncogène du GPER proviennent d'études démontrant que les agonistes sélectifs du GPER à doses pharmacologiques peuvent

induire des événements cellulaires conduisant à la mort cellulaire, tels que la rupture de la tubuline, l'arrêt mitotique G2/M et la régulation à la hausse de p53, p21 et c-myc. Dans le contexte de cet argument, on peut noter que l'antagoniste sélectif du GPER, le G36 (à une dose de 11 mg pellets) inhibe la croissance du cancer de l'endomètre [166] et le G15 (1,46 mg/kg) celle du poumon non à petites cellules.

Ainsi, il serait plus approprié de reconnaître que l'exposition totale aux estrogènes, qu'ils soient d'origine naturelle ou environnementale, constitue un stimulus chronique favorisant l'action des estrogènes, indépendamment du type de récepteur. L'action complémentaire de ces récepteurs, à l'exception de celle de certains variants du ERβ, est associée aux tumeurs malignes induites par les estrogènes. Par exemple, les xéno-estrogènes abondants, tels que les bisphénols, les phtalates et les IFs, ont été négligés en grande partie en raison de leurs faibles RBA pour ERα, bien qu'ils soient nombreux à présenter des RBA nettement plus élevées pour GPER. Par exemple, le RBA du 4-hydroxyTam pour ER est de 175-257 % contre 0,021 % et 6,8 % pour la GEN pour ERα et ERβ, respectivement. Tandis que les RBA de la GEN et du Tam vis-à-vis de GPER sont similaires (2-5 % et 0,3-6 %, respectivement), et les deux composés fonctionnent comme des agonistes de GPER [42]. Le Tam et la GEN présentent donc des RBA similaires pour GPER. Cependant, leurs RBA pour ERα et ERβ diffèrent considérablement.

La preuve que le GPER agit indépendamment des ERα et ERβ est apportée par le fait que l'action rapide des estrogènes ne doit pas nécessairement être corrélée à l'expression des protéines ERα et ERβ et peut se produire dans les cellules cancéreuses du sein SKBR3 dépourvues d'ARNm pour les deux récepteurs [168]. L'existence d'agonistes et d'antagonistes spécifiques permettent aussi de séparer les deux types de voies de signalisation.

L'exploration du rôle du GPER implique nécessairement la prise en compte du rôle potentiel des récepteurs aux estrogènes membranaires dans la signalisation cellulaire. Un corpus croissant de preuves suggère que ces récepteurs membranaires jouent un rôle crucial dans la signalisation stéroïdienne initiée par la membrane (MISS), régulant ainsi divers événements physiologiques. Cependant, il est crucial de souligner que les modèles actuels ne permettent pas d'exclure définitivement l'influence de ces récepteurs dans leur interaction avec le GPER. De plus, l'interaction complexe entre le GPER et les actions non génomiques d'autres stéroïdes sexuels engendre des résultats parfois contradictoires d'une publication à l'autre [42,172].

II.5.5. Isoflavones et GPER

La variabilité des résultats observés lors de l'évaluation de l'activité de la GEN ou d'autres phytoE, dans le cancer du sein, via le GPER peut s'expliquer non seulement par les différentes concentrations utilisées, mais aussi par la quantité et le répertoire des récepteurs aux estrogènes présents dans les différents modèles cellulaires. Ces derniers déterminent des interactions pouvant conduire à une amplification ou une inhibition de l'effet. Ces phénomènes sont similaires à ceux décrits pour les récepteurs aux estrogènes classiques, et les interactions entre les différents types de récepteurs jouent un rôle important dans la physiologie et la pathologie des cellules mammaires [17]. Il est également important de considérer les interactions entre les co-cultures de cellules MCF-7 (positives pour les récepteurs ERα et GPER) et les cellules souches issues de MDA-MB-231 (négatives pour ERα et GPER) stimulées par la GEN (à 40 nM et 2μM), ce qui induit une différenciation importante des cellules souches du cancer

du sein [208]. La coexistence de plusieurs formes de récepteurs aux estrogènes dans les cellules cancéreuses indique une communication complexe entre GPER, ERα et ERβ ainsi que les nouveaux variants d'épissage non nucléaires ERα36 et ERα46, exprimés dans de nombreux tissus, y compris les cellules du cancer du sein.

Ariyani *et al.* [40] ont rapporté dans une étude de 2019 des effets du S-équol sur des neurones et dendrocytes via le GPER. Ce dernier est responsable de l'augmentation de l'arborisation dendritique et d'une prolifération des astrocytes observés avec des concentrations physiologiques de 1 à 10 nM de S-équol. Il est également impliqué dans la migration des astrocytes induite par l'équol. Une des autrices montre dans une publication ultérieure [7] que E_2 et les IFs possèdent une position de liaison similaire dans l'ancrage moléculaire proposé par l'équipe. Cette dernière a également observé que des concentrations de 1, 10 et 100 nM d'IFs induisaient la migration des cellules gliales principalement via le GPER, cet effet étant supprimé par l'antagoniste sélectif G15 ou des siARN anti-GPER.

Vivacqua *et al.* [209] ont démontré que le GPER médie la prolifération des cellules cancéreuses thyroïdiennes (WRO, FRO et ARO) induite par la GEN, via la voie MAPK. Cet effet a également été observé avec l'E_2 et l'hydroxy-TAM. La GEN a notamment entraîné une augmentation de l'expression des gènes c-fos, cycline A et cycline D1. Ces effets stimulants ont été inhibés par le *silencing* GPR30.

Sur un modèle d'endométriose, Lin *et al.* [210] ont rapporté que la GEN à 100 nM augmentait la transcription du facteur de transcription SF-1 et induisait la prolifération cellulaire dans les cellules humaines primaires H-38 exprimant des niveaux élevés de GPER sans expression détectable de ERα ou ERβ, ainsi que dans les cellules cancéreuses de l'endomètre Ishikawa exprimant le GPER. Ces effets prolifératifs étaient dépendants du GPER.

Des effets prolifératifs des IFs ont donc été rapportés dans la littérature sur des modèles cellulaires différents, mais le possible effet prolifératif des isoflavones, notamment sous forme conjuguée, sur des cellules TNBC exprimant le GPER n'a à notre connaissance pas été exploré.

III. Méthodologie, matériels et méthodes

III.1. Evaluation de l'exposition des consommatrices et consommateurs et étude ISOLED

III.1.1. Exposition aux phyto-estrogènes et aux pesticides

Au commencement de cette recherche, l'évaluation de l'exposition aux substances toxiques, en France, faisait l'objet de plusieurs études en cours. Les travaux de l'Anses incluaient les phytoE, les mycotoxines ainsi que des PEs d'origine anthropique, tels que les migrants d'emballage (Bisphénol A ou phtalates), les pesticides (Carbamates, Imidazoles, Pyretroïdes…) et d'autres substances d'origine industrielle (PBB, PFOS, Dioxines…). Ces données sont régulièrement compilées par l'Anses dans les études EAT [26,211,212]. D'autres travaux, financés notamment par des projets européens, restaient focalisés sur les PEs d'origine anthropiques. Dans les deux cas ces estimations reposaient sur de nombreuses mesures coûteuses réalisées dans des sources alimentaires et des échantillons d'eau. Cependant, la représentativité d'un échantillon, en particulier à l'échelle d'un pays, peut s'avérer variable. En effet, les denrées alimentaires ne proviennent pas uniquement d'une production locale contrôlée mais aussi d'importations dont le suivi exhaustif est difficile à mettre en œuvre. De plus, l'évaluation individualisée de l'exposition de volontaires à des PEs présente des défis particuliers, car sa pertinence repose sur une prise en compte exhaustive des habitudes de vie, notamment alimentaires, et sur des analyses biologiques approfondies. En effet, la compréhension des effets cocktails, qui impliquent des interactions complexes entre divers polluants, exige une analyse métabolomique complète des profils de polluants afin de saisir l'ensemble de l'exposome. Cet exposome englobe à la fois des polluants persistants susceptibles de se bioaccumuler dans les tissus et de laisser des traces à long terme, et des substances plus facilement éliminables mais dont l'activité potentielle ne saurait être négligée.

III.1.1.1. Phyto-estrogènes

Les PEs estrogéniques sont ubiquitaires dans l'environnement. Parmi eux, les IFs, souvent négligées, présentent une activité estrogénique *in vitro* supérieure à celle de nombreux xéno-estrogènes et sont présentes à concentrations significatives (milligrammes) dans l'alimentation, contrairement à d'autres PEs généralement présents à l'échelle du microgramme ou du nanogramme.

L'évaluation de l'exposition aux phytoE a reposé sur l'exploitation de données de dosage issues de produits alimentaires français, élargies dans le cadre de cette thèse [213]. Ces données ont également été complétées par une estimation initiale de la consommation d'IFs en France, obtenue par l'analyse d'étiquettes de produits alimentaires provenant des sites web des grandes surfaces et du site Mercuriales dédié à la restauration collective [117]. Cette précédente étude menée au sein de l'équipe a permis d'identifier les produits en France susceptibles de contenir du « soja caché » et a montré que les consommateur·ices fréquentant les cantines présentent un risque accru d'exposition aux IFs. Trois types de produits ont ainsi été retenus pour la présente étude : les produits à base de soja, ceux contenant du « soja caché » et ceux ne contenant pas de soja. Le « soja caché » est généralement présent sous forme de protéines texturées, de concentrés ou d'isolats protéiques de soja, dont la teneur en IFs diffère de celle des produits de soja traditionnels. Ce paramètre a été pris en compte dans l'évaluation de l'exposition. Celle au « soja caché » a été affinée en croisant les

données des étiquettes obtenues précédemment avec les déclarations des sites professionnels concernant l'évolution des ventes de protéines végétales, et plus particulièrement des protéines de soja.

Les détails de la méthode de calcul sont présentés dans l'article intitulé : *Endocrine disruptors on and in fruits and vegetables : Estimation of the potential exposure of the French population*, avec la base de données associée [213].

III.1.1.2. Produits phytosanitaires

En l'absence de données exhaustives sur les concentrations de pesticides dans l'alimentation française, notre étude s'est inspirée d'une approche théorique pratiquée à l'Anses pour estimer l'exposition aux produits phytosanitaires PEs. L'accès aux données de vente de pesticides en France étant restreint pour des raisons de confidentialité, la méthodologie adoptée a été la suivante.

1. **Sélection des pesticides pertinents** : parmi les 1426 produits phytosanitaires autorisés en Europe, 426 étaient homologués en France. Cette liste a ensuite été réduite aux 379 fongicides, insecticides, acaricides, herbicides, et substances de croissances pour lesquelles une limite maximale de résidus (LMR) était définie en France.
2. **Identification des PEs** : une recherche bibliographiques croisée avec les terme « endocrine », « disruption », « hormones », « disruptors » a permis d'identifier 70 PEs parmi les produits phytosanitaires utilisables en France. Tous ont été considérés, que leurs effets aient été démontrés *in vivo* ou *in vitro*.
3. **Détermination des fruits et légumes exposés** : l'analyse des utilisations autorisées pour chaque substance retenue a permis de dresser une liste de 64 végétaux susceptibles d'être contaminés par les PEs sélectionnés.
4. **Estimation de la consommation alimentaire** : les volumes de consommation des Français·es ont été estimés à partir des données des syndicats de producteurs. Une consommation individuelle quotidienne moyenne a été déduite, en divisant la consommation totale par 365 jours et 63 millions d'habitant·es, en excluant les populations les plus jeunes et les plus âgées sur la base des données de l'INSEE.
5. **Validation de l'estimation de la consommation individuelle** : les chiffres obtenus ont ensuite été confrontés aux données Inca produites par l'Anses à partir de questionnaires alimentaires validés.
6. **Estimation de l'exposition aux PEs** : les LMR des produits phytosanitaires PEs pour chaque fruit et légume ont été croisées avec les quantités consommées quotidiennement.
7. **Prise en compte des choix des agriculteurs** : le nombre total de formules phytosanitaires disponibles pour chaque application et chaque fruit ou légume a été intégré afin d'estimer la probabilité d'utilisation d'une formule contenant un PE. A titre d'illustration, pour les pommes il existe 400 formules d'insecticides disponibles sur le marché et 341 d'entre-elles contiennent une substance PE. Un agriculteur a donc 85,25% de chance d'utiliser un insecticide contenant un PE.

L'analyse menée a permis d'estimer une exposition limitée aux produits phytosanitaires autorisés en France, en tenant compte des choix statistiques des agriculteur·ices. Cette estimation repose sur l'hypothèse que l'ensemble des produits

alimentaires disponibles contiennent une quantité de produits phytosanitaire égale à la LMR. Cette hypothèse, bien que simplificatrice, permet d'établir une borne supérieure de l'exposition potentielle de la population française, bien qu'elle ne tienne pas compte des produits agricoles importés qui pourraient contenir des PEs non autorisés en France.

III.1.2. Questionnaires alimentaires

Afin d'évaluer le rôle potentiel des phytoE comme facteurs de risque du LED, une étude clinique nommée ISOLED (CHUBX 2017/18 ou NCT03421184) a été menée. Il s'agit d'une étude observationnelle, monocentrique et préliminaire, menée au CHU de Bordeaux, ayant été approuvée par le comité d'éthique local sous le numéro 2017T2-29. Compte tenu des concentrations de phytoE auxquelles les consommateur·ices sont généralement exposés, deux familles de composés nous ont semblé pertinentes : les IFs du soja et les lignanes. Il est important de rappeler que les lignanes estrogéniques ne sont pas présents en tant que tels dans les aliments, mais qu'ils peuvent être transformés en entérolignanes (entérodiol et ENL hormonaux-actifs), par la flore intestinale compétente.

L'étude ISOLED a débuté le 26 novembre 2018 et visait à inclure 100 participantes réparties en 3 groupes.

- Groupe LED : femmes pré-ménopausées, ayant au moins 18 ans, avec une manifestation de LED aigu, ayant donné leur consentement éclairé et couvertes par une assurance sociale.
- Groupe autre MAIs : femmes pré-ménopausées d'âge correspondant atteintes d'une autre MAI (polyarthrite rhumatoïde, thrombocytopénie auto-immune, etc.), ayant donné leur consentement éclairé, et couvertes par une assurance sociale.
- Groupe contrôle (volontaires saines) : femmes pré-ménopausées d'au moins 18 ans, ayant donné leur consentement éclairé, et couvertes par une assurance sociale.

Les critères d'exclusion étaient :

- séropositivité pour le virus de l'immunodéficience humaine (VIH), le virus de l'hépatite C (VHC) ou le virus de l'hépatite B (VHB) ;
- femmes enceintes ou allaitantes ;
- femmes ménopausées ;
- et pour le groupe LED, les patientes en rémission de phase quiescente de la pathologie.

Les données suivantes ont été collectées pour chaque participante.

- Données biologiques : prélèvements sanguins, urinaires et capillaires pour dosage des IFs du soja (GEN, DAI, et équol) et de l'ENL.
- Données sur les habitudes alimentaires : questionnaire des habitudes alimentaires détaillé et rappel des 24 -48 heures précédant le prélèvement.
- Données sociodémographiques : âge, niveau d'éducation, profession, etc.
- Données médicales : contraception, pathologie en cours, traitements hormonaux, rang de naissance, gémellarité, FIV, IVG, etc.

Compte tenu que l'étude ISOLED cible les femmes pré-ménopausées, l'estimation de l'exposition aux phytoE déterminée dans ce travail est limitée à cette catégorie de

population. Néanmoins, elle pourrait être étendue à d'autres sous-populations en s'appuyant sur les données recueillies dans l'étude de Lee *et al.* [117].

Au début de nos travaux, aucun outil validé pour estimer l'exposition aux phytoE des consommateur·ices n'était disponible en France. Or, les outils d'évaluation de l'exposition aux substances alimentaires présentent une portée limitée au niveau national, en raison de la diversité des régimes alimentaires et de la variabilité de la disponibilité des aliments d'un pays à l'autre. Par conséquent, deux types de questionnaires alimentaires ont été élaborés : un visant à cerner les habitudes alimentaires des participantes et un autre portant sur leur consommation récente (24-48 heures) susceptible d'influencer les teneurs en biomarqueur dans les échantillons biologiques prélevés. L'objectif était d'affiner l'évaluation de l'exposition des consommateur·ices basée initialement sur la compilation de données, par l'analyse de biomarqueurs biologiques et l'établissement de corrélation avec des données alimentaires. L'élaboration de ces questionnaires s'est appuyée sur l'expérience acquise lors de l'étude de Lee *et al.* [117] et de l'étude sur les PEs dans et à la surface des fruits et légumes [213].

III.1.2.1. Enquêtes d'habitudes alimentaires

La conception des questionnaires d'enquête s'est articulée autour des aliments pourvoyeurs d'IFs, en distinguant les aliments au soja clairement identifiés et les aliments ultra-transformés contenant du « soja caché ». Pour les questionnaires d'habitudes alimentaires, des listes d'aliments présélectionnés ont été proposés aux enquêteur·ices, qui les ont ensuite présentées aux participantes et ont rempli les cases correspondant aux fréquences de consommation et aux tailles des portions. Afin de consolider la fiabilité des réponses, des questions plus générales ont été incluses dans le questionnaire, portant sur la fréquence habituelle des repas, les lieux de repas habituels, le type d'aliments (faits maison ou industriels), les types de régimes (omnivore, végétarien, flexitarien) et le mode de cuisson habituel.

Pour la première fois dans ce domaine de recherche, des dosages d'IFs dans les cheveux ont été développés afin de valider ces questionnaires d'habitudes alimentaires. Les cheveux constituent une matrice biologique particulière capable de fixer des biomarqueurs sur le long terme. Une consommation régulière et historique de soja peut donc laisser des traces dans les cheveux sous la forme d'IFs extractibles et dosables.

III.1.2.2. Rappels des 24 et 48 heures

Des enquêtes alimentaires ont été menées afin d'évaluer plus précisément le rôle des phytoE dans les poussées lupiques. Ces enquêtes ont été réalisées auprès de patientes hospitalisées pour une poussée de LED dans les heures suivant leur admission. Elles visaient à identifier d'éventuelles expositions aux IFs et à expliquer les teneurs mesurées dans les fluides biologiques des patientes.

Pour ce faire, les aspects pharmacocinétiques des phytoE et le temps écoulé entre la consommation alimentaire et le prélèvement sanguin ou urinaire sont pris en compte. Cette approche a conduit à l'élaboration de scores alimentaires intégrant à la fois le temps écoulé et la quantité de phytoE ingérée.

Les enquêtes, de type « rappels des 24 et 48 heures », étaient ouvertes et demandaient aux patientes de décrire leur consommation alimentaire sur les 2 jours précédents leur

hospitalisation. Les questions portaient sur les consommations aux repas et hors repas, en enregistrant les heures de prise alimentaire et les quantités consommées. Le lieu de consommation, l'origine et, si possible, la marque des produits consommés étaient également enregistrées.

Comme la pharmacocinétique des IFs dans le sang et l'urine humaine est bien connue, il est possible de retracer la consommation de soja jusqu'à 48 heures par l'analyse des urines et jusqu'à 24 heures par l'analyse du plasma.

III.1.2.3. Etablissement de scores

III.1.2.3.1. Score pour les habitudes alimentaires

La validation du questionnaire d'habitudes alimentaires visant à estimer l'exposition aux IFs, présenté précédemment, a été réalisée par comparaison avec des dosages d'IFs dans les cheveux.

Développement de scores alimentaires

Des scores alimentaires ont été appliqués pour tenir compte à la fois de la quantité d'IFs apportée par un aliment et de la probabilité de cet apport. Ainsi, une galette de soja apporte plus d'IFs qu'un yaourt et les boulettes de viande achetées en grande surface contiennent plus souvent du soja que les charcuteries (80 vs 30%). Ces scores ont été construit en prenant en considération ces deux paramètres.

Score capillaire

Un score capillaire a également été construit à partir d'information sur des facteurs pouvant influencer la présence d'IFs dans la matrice capillaire. Ces facteurs incluent la fréquence et la qualité des shampoings, des coupes, des colorations ou des permanentes. Compte tenu des données de la littérature, des coefficients de pondération ont été appliqués au score alimentaire strict pour tenir compte de ces facteurs spécifiques. La longueur des cheveux a également été considérée couplée à la durée pendant laquelle le régime alimentaire déclaré a été suivi.

Les détails de la méthode de calcul sont présentés dans l'article intitulé : *Development of an Assay for Soy Isoflavones in Women's Hair* [214].

III.1.2.3.2. Score pour la consommation récente

Les enquêtes de rappels des 24 et 48 heures ont permis de recueillir des données détaillées sur la consommations d'aliments pourvoyeurs d'IFs dans les deux jours précédant le prélèvement des échantillons biologiques. Afin de refléter plus précisément l'exposition aux IFs, des scores ont été établis en tenant compte de la pharmacocinétique de ces composés.

En effet, les IFs n'apparaissent dans le plasma qu'environ 8 heures après ingestion. Cela signifie qu'un bol de lait de soja consommé au petit déjeuné, à 7 heures du matin, ne sera pas détectable dans une prise de sang réalisée à 9 heures du matin. En revanche, les IFs d'une portion de tofu consommé la veille au soir seront encore présentes dans le sang et, surtout, dans l'urine le lendemain matin à 9 heures.

Par conséquent, les scores développés ont intégré une pondération de la dose prévisible d'IFs ou de lignanes liée à la prise alimentaire en fonction de l'évolution cinétique de ces composés dans le fluide biologique considéré. Cette approche, inédite

à notre connaissance, permet une estimation plus précise de l'exposition aux IFs en prenant en compte le délai d'absorption et d'élimination de ces composés.

III.1.3. Analyses des biomarqueurs

L'analyse des IFs a été effectuée par dosage immuno-enzymatique (ELISA), méthode inventée dans les années 1960 par Peter Perlmann et Eva Engvall. Cette méthode est basée sur le principe d'une caractérisation immunologique, utilisant une enzyme au lieu de la radioactivité comme biomarqueur [215]. Elle a été validée par un test circulaire international en comparaison avec des méthodes physico-chimiques bien établies [75,216]. Grâce à sa technique d'extraction, elle s'est avérée être l'une des plus précises et des plus sensibles pour le dosage des IFs.

L'équipe du Pr. Bennetau-Pelissero a développé dans les années 2000 des dosages ELISA des IFs du soja : GEN, DAI et du métabolite équol [217,218]. En 2008, elle a publié un dosage ELISA de glycitéine [219] et en 2014 un dosage d'ENL [220].

III.1.3.1. Extractions des isoflavones

La procédure d'extraction varie selon la matrice analysée.

III.1.3.1.1. A partir des aliments

Pour cette expérience, le protocole validé d'extraction utilisé peut être résumé par les étapes décrites ci-dessous [30,221].

- **Préparation de la matrice alimentaire** : dissolution de 1 g de matière à base de soja broyé dans 50 mL d'eau déminéralisée (dilution 1 : 50), sous agitation pendant 20 min à température ambiante, puis 10 min à 100°C. Après refroidissement à température ambiante, 3 échantillons de 500 µL sont prélevés sous agitation, pour assurer l'homogénéité.
- **Digestion enzymatique des conjugués** : selon l'Anses et d'autres expert·es [30,221], les concentrations d'IFs doivent être exprimées sous forme d'aglycones. Libérer l'aglycone des formes glycosidiques avant le dosage permet de répondre à cette exigence sans calculs supplémentaires. Ainsi, 2 mL de tampon acétate (0,1 M, pH=5, contenant des antibiotiques) sont ajoutés à chaque tube à essai du triplicat, suivis de 10 µL de l'enzyme de déconjugaison βglucuronidase arylsulfatase d'*Helix pomatia* (Roche Diagnostic). La réaction enzymatique est menée à 37°C pendant 12 heures sous agitation lente. La durée de digestion a été établie sur la base de tests effectués à différentes durées.
- **Extraction** : elle est réalisée au moyen de 3 cycles d'extraction liquide-liquide à l'acétate d'éthyle acidifié suivi d'une agitation au vortex. La phase organique est ensuite séparée par centrifugation (à 4°C, 10 min, 2000 RPM) et congélation. Les échantillons sont évaporés à sec sous vide. Les résidus sont redissous dans un tampon phosphate salin (PBS) avec 0,05% de Tween 20 et 1% de DMSO (PBS-T-DMSO) puis stockés à -20°C.
- **Evaluation des taux d'extraction** : le taux de récupération de l'extraction est évalué en utilisant la génistine (forme glycosidique) comme étalon externe. Une solution de 0,01 mg/mL de génistine est préparée et soumise à la même digestion enzymatique et aux mêmes conditions d'extraction que les échantillons. Le pourcentage de récupération est calculé comme le rapport entre la concentration moyenne en génistéine (forme aglycone) obtenu et sa concentration théorique si l'extraction était de 100%. Ce contrôle est réalisé pour chaque extraction, en

triplicat. En outre, un échantillon de lait de soja est extrait en triplicat et en parallèle. Il permet de standardiser la digestion, l'extraction et le dosage.

Avant dosage, les échantillons d'aliment sont dilués de 250 à 500 000 fois.

III.1.3.1.2. A partir des cheveux

Les dosages des phytoE dans les cheveux est une technique innovante développée dans le cadre de cette étude. La résistance de cette matrice, qui devait pourtant être hydrolysée pour libérer les substances qu'elle contenait, a nécessité le développement d'un protocole d'extraction spécifique. La difficulté majeure résidait dans l'extraction efficace de ces molécules dans des conditions drastiques tout en préservant leur intégrité structurale. En s'appuyant sur la littérature existante, notamment sur les méthodes de dosages des stéroïdes dans les cheveux, plusieurs protocoles d'extraction ont été testés. Ces protocoles combinaient des digestions enzymatiques douces (pronase, kératinase) et/ou des hydrolyses en solvants plus forts (acide chlorhydrique, soude) pour décomposer la matrice et libérer les composés recherchés. La méthodologie optimisée, sélectionnée sur la base de ses performances en termes d'efficacité d'extraction et de préservation des analytes, est décrite plus en détails dans la publication [214] présentée dans le chapitre « Résultats et Discussions ».

III.1.3.1.3. A partir des fluides biologiques

La principale différence entre les dosages des IFs dans les aliments et les fluides biologiques réside dans le traitement des échantillons lors de l'extraction. Les fluides biologiques (sang et urine) sont traités directement, sans dilution préalable, par la β–glucuronidase arylsulfatase qui transforme les composés glucuronidés et sulfatés en composés aglycones. Ces composés hydrophobes sont ensuite extraits par de l'acétate d'éthyle en 3 étapes successives.

Pour les dosages, les échantillons de plasma sont dilués de 5 à 20 fois et les échantillons d'urine de 5 à 200 fois.

III.1.3.2. Mise en œuvre de l'analyse

La synthèse des haptènes a été le fruit d'une collaboration continue entre des chimistes organiciens et les biologistes de l'équipe et a donné lieu à plusieurs publications [217–220]. Les haptènes couplés à l'albumine sérique bovine (BSA) ont été injectés à des lapins qui ont générés des anticorps polyclonaux. Les mélanges polyclonaux permettent d'obtenir une très grande spécificité. Ceci est particulièrement important ici car les IFs estrogéniques sont très proches structuralement les unes des autres. Les dosages en eux-mêmes se font par compétition pour un anticorps spécifique entre un haptène fixé sur une plaque et un antigène libre dans la gamme d'étalonnage ou dans l'échantillon spécifique [217,218].

Le dosage comporte les étapes successives suivantes.

III.1.3.2.1. Saturation et incubation primaire

- Fixation des antigènes : les puits des plaques sont recouverts avec un couplage d'haptène de phytoE à la thyroglobuline de porc à 0,17 µg/mL pour GEN et 0,2 µg/mL pour DAI, laissées pendant une nuit à 4°C.
- Saturation : les plaques sont saturées avec 200 µl de PBS-T- DMSO contenant 0,2 % de BSA et 0,9 % de NaCl, pendant 30 minutes à 37 °C. Ils sont ensuite lavés 3 fois dans du PBS-T- DMSO.

- **Préparation de la gamme d'étalonnage et des échantillons** : des étalons de référence et les échantillons sont dilués dans du PBS-T-DMSO. La courbe standard est calculée sur 10 points en duplicata qui correspondent à des concentrations du standard de référence allant de 0,96 à 500 ng/mL [222]. Tous les échantillons ont été testés trois fois sur différentes plaques.
- **Mise en contact avec l'anticorps primaire** : 100 µL des standards de référence et des échantillons sont distribués par puits, puis mis en contact avec100 µl de la solution d'anticorps primaire spécifique de l'IF à doser, diluée dans le PBS-T-BSA-DMSO. Les solutions d'anticorps ont été diluées à 1 : 40 000 pour l'anti-GEN et à 1 : 50 000 pour l'anti-DAI. Après 2 heures à 37°C, les plaques sont lavées 3 fois dans du PBS-T- DMSO.

III.1.3.2.2. Incubation secondaire et révélation

- **Mise en contact avec l'anticorps secondaire couplé à la peroxydase** : après que les plaques aient été vidées, 200µL d'une solution d'IgG de chèvre anti-lapin (Amdex® - GE HealthCare RPN4301) diluée à 1/10 000 dans du PBS-T-BSA-DMSO sont ajoutés par puits et laissés pendant 30 minutes à 37°C.
- **Lavages** : les plaques sont lavées 3 fois avec du PBS-T-DMSO.
- **Révélation** : l'anticorps secondaire est mis en contact avec un substrat. 200 µL d'ortho-phénylène diamine (OPD), 0,5 M, dans un tampon citrate-phosphate (0,15 M, pH 5), sont ajoutés par puits. La réaction progresse à 37°C pendant 20 min pour DAI et 30 min pour GEN. L'OPD se colore alors sous l'effet de l'oxydation de manière proportionnelle à la concentration d'anticorps fixé. La réaction est arrêtée en ajoutant 50 µL de H_2SO_4 4 M dans chaque puits.
- **Lecture** : 10 minutes après la révélation, les plaques sont lues à 490 nm.

Tout au long de l'analyse, certaines erreurs classiques en ELISA ont été soigneusement évitées, notamment, en utilisant des tubes à hémolyse et des cônes de pipette siliconés pour éviter l'adhérence des IFs aux récipients en plastique [222].

III.2. Diminution des isoflavones dans les aliments

Les données scientifiques accumulées sur les effets des IFs du soja, notamment leurs effets estrogéniques, ainsi que celles issues de nos propres analyses sur l'exposition des consommateur·ices français·es, convergent vers la nécessité de limiter l'exposition aux IFs. Cette position rejoint celle des agences sanitaires françaises exprimée en 2005 [8].

III.2.1. A l'échelle du laboratoire

Cette étape, menée en partenariat avec la société Biopress/Berkem, est importante pour une transposition ultérieure à l'échelle pré-industrielle. Ainsi, les conditions optimales au pré-traitement du soja pour minimiser la teneur en IFs (GEN et DAI) dans les produits dérivés du soja, destinés à l'industrie agroalimentaire ou à l'alimentation humaine, doivent être identifiés.

Cinq facteurs influençant potentiellement la teneur en IFs ont été sélectionnés pour cette étude, avec cinq niveaux distincts (voir Tableau 5).

Un plan expérimental de Tagushi a été utilisé pour optimiser le nombre d'essais nécessaires tout en évaluant l'effet de chaque facteur sur la réduction des IFs. Ce plan a permis de tester 25 combinaisons parmi les 3125 possibles, réduisant ainsi considérablement le nombre d'expériences à réaliser.

Niveaux	Température (°C)	Ratio volumétrique soja : eau	Durée (min)	Vitesse d'agitation (RPM)	Matrice de soja
1	20	1 : 2	0	0	Graines
2	45	1 : 2,5	5	50	Graines dépelliculées
3	55	1 : 3	10	100	Graines dépelliculées broyées
4	65	1 : 4	20	150	PST
5	75	1 : 5	30	200	Tourteaux

Tableau 5. Conditions expérimentales (facteurs et niveaux) pour la réduction des isoflavones à l'échelle du laboratoire

La méthodologie détaillée du prétraitement du soja est décrite dans l'article *Improved Food-Processing Techniques to Reduce Isoflavones in Soy-Based Foodstuffs,* dans le chapitre « Résultats et Discussions » de ce tapuscrit.

III.2.2. A l'échelle pilote (préindustrielle)

L'échelle pilote, qui opère sur des quantités de quelques kilogrammes, représente une étape intermédiaire cruciale entre l'échelle du laboratoire (quelques grammes) et l'échelle industrielle (plusieurs tonnes). Elle permet de transposer les résultats obtenus en laboratoire à des conditions plus proches de celles rencontrées en production industrielle, tout en limitant les risques et les coûts associés à une mise en œuvre à grande échelle.

Cette étude, développée dans le cadre du projet PHYTOSECUR 1, financé par la Région Nouvelle Aquitaine, vise à explorer les méthodes de traitement du soja dépelliculé, à l'échelle pilote, dans le but de minimiser la teneur en IFs d'une matière première destinée à l'industrie de transformation du soja. Ce projet a associé deux plateformes technologiques : l'Institut de la Filtration et des Techniques Séparatives (IFTS) et l'AgroTech d'Agen, ainsi qu'un industriel produisant des PST (Biopress/Berkem) et l'Université de Bordeaux.

Deux types de traitement ont été comparés : un traitement à l'eau simple et un autre précédé d'un douchage. Les conditions de traitement ont été basés sur les résultats obtenus à l'échelle du laboratoire, tout en étant adaptés aux contraintes pré-industrielles.

L'efficacité du traitement a été évaluée en comparant les concentrations finales d'IFs dans la matière sèche aux valeurs rapportées dans la littérature comme ayant des effets délétères sur la santé humaine.

La méthodologie détaillée du traitement du soja dépelliculé est également présentée dans l'article *Improved Food-Processing Techniques to Reduce Isoflavones in Soy-Based Foodstuffs,* dans le chapitre « Résultats et Discussions » de ce tapuscrit.

III.2.3. A l'échelle domestique à partir des recettes traditionnelles

Dans le cadre d'un projet intitulé PHYTOSECUR 2, plusieurs aliments à base de soja ont été préparés en suivant des protocoles traditionnels, en incluant un contrôle des IFs à chaque étape de préparation. Ces valeurs ont ensuite été comparées aux teneurs en IFs issues de produits industriels équivalents. On rappelle que l'objectif ici est de démontrer que les traitements traditionnels et domestiques peuvent réduire de manière significative la teneur des IFs dans les aliments à base de soja.

La méthodologie appliquée est décrite en détails dans l'article intitulé *Traditional and Domestic Cooking Dramatically Reduce Estrogenic Isoflavones in Soy Foods* [223], dans le chapitre « Résultats et Discussions » de ce tapuscrit.

III.3. Etudes cellulaires

III.3.1. Cultures cellulaires

Les lignées cellulaires MCF-7, MDA-MB-436, MDA-MB-468 et MDA-MB-231 sont maintenues dans du milieu RPMI 1640 Glutamax (61870-010, Gibco) supplémenté avec 10% de Sérum de Veau Fétal (SVF) (SV30160.03, Cytiva) et 1% de Pénicilline/streptomycine (15140-122, Gibco). Les SKBR3 sont maintenues dans du milieu DMEM (12491-015, Gibco) supplémenté de manière identique. Les cellules sont mises en culture avec changement de milieu tous les deux à trois jours jusqu'à atteindre la confluence souhaitée (~80%). L'ensemble des cellules est maintenu dans des conditions standardisées : 37°C, 5% CO_2 et atmosphère saturée en humidité.

Le choix des modèles cellulaires s'est basé sur la littérature. MDA-MB-436 et MDA-MB-468 ont été sélectionnées parce que considérés comme bons modèles de TNBC exprimant le GPER [180,181]. Les MDA-MB-231 TNBC n'exprimant pas ou très peu le GPER ont été retenus comme modèle de contrôle négatif. Les MCF7, positives pour l'ensemble de ces récepteurs (ER+, PR+, HER2+, GPER+) étaient considérés comme notre modèle contrôle positif. Les SKBR3 sont ER(-), PR(-), GPER(+) et HER2(+). L'objectif était de comparer l'impact des IFs, dans leurs formes circulantes, sur différents modèles cellulaires, et d'étudier l'implication potentielle du GPER.

III.3.2. Test de prolifération

A 80% de confluence, les cellules sont ensemencées dans des plaques 96 puits, en utilisant seulement les 60 puits internes afin d'éviter un effet de bord. Pour les lignées cellulaires étudiées, à l'exception des MDA-MB-468, 5 000 cellules ont été distribuées dans 100 µL de milieu de culture par puits de plaque. Les plaques ont ensuite été incubées à 37°C dans une atmosphère à 5% de CO_2. Dans le cas des cellules MDA-MB-468, en raison de leur temps de doublement plus court, 3 000 cellules seulement ont été distribuées par puits de plaque. Les conditions d'incubation étaient identiques à celles des autres lignées cellulaires. Une fois les cellules adhérentes, le milieu est remplacé par du milieu sans rouge phénol (11835-030, Gibco) et supplémenté avec 10% de SVF déplété en hormones par un traitement sur charbon actif. Les cellules sont ensuite traitées 24 heures après le changement de milieu et maintenues en culture jusqu'à $J_{traitement+5}$. Le milieu (traitement et contrôle) est remplacé tous les deux jours. Une lecture est effectuée quotidiennement.

Traitement des cellules

Les molécules testées sont dissoutes dans du DMSO ou directement dans de l'eau ultrapure stérile, pour obtenir des solutions stock/mères à la limite de solubilité. Ces

solutions sont diluées progressivement dans de l'eau ou du DMSO si nécessaire puis finalement dans du milieu sans rouge phénol complété avec du SVF déplété, jusqu'à l'obtention des concentrations souhaitées (1, 10, 100 nM et 1µM). La concentration finale de DMSO dans les puits est <0,3% dans tous les cas. La même concentration de DMSO est ajoutée aux puits contrôles (ne contenant que du milieu).

Dans les puits le nécessitant, un pré-traitement de 30 minutes avec 1µM de G15 est effectué.

Mesure de la prolifération

La prolifération est évaluée par le test PrestoBlue® (A13262, Invitrogen). Après le traitement des cellules, 10 µL de réactif PrestoBlue® sont ajoutés à chaque puits, en remplaçant 10 µL de milieu de culture. La plaque est ensuite incubée 10 minutes à 37°C après l'avoir secoué légèrement. La fluorescence est mesurée à 560 nm à l'aide d'un lecteur de micro-plaques Infinite M1000 Pro (TECAN).

III.4. Analyses moléculaires

III.4.1. Western Blot

Le western blot (WB), également connu sous le nom d'immunotransfert, est une méthode de séparation, d'identification et de dosage (semi-quantitatif) des protéines. Il repose sur la séparation des protéines par électrophorèse sur gel en fonction de leur masse moléculaire, suivie de leur transfert sur une membrane et de leur détection par des anticorps spécifiques.

III.4.1.1. Extraction et dosage des protéines

Les échantillons utilisés correspondent aux protéines totales des cellules étudiés après dénaturation.

Extraction des protéines totales

Les cellules sont récoltées, centrifugées et rincées plusieurs fois avec du PBS froid (14190-144, Gibco). Le culot cellulaire est re-suspendus dans du tampon RIPA (Pierce RIPA Buffer 89901, Thermo Scientific) supplémenté avec un cocktail d'inhibiteurs de protéases (Pierce Proteases Inhibitor Tablets A32963, Thermo Scientific). Cette étape permet la lyse des composants cellulaires et la conservation des protéines. L'échantillon est homogénéisé au vortex puis au sonicateur pour s'assurer de la rupture des membranes et organites. L'échantillon est ensuite placé 15 min sur glace, en mélangeant régulièrement. La suspension est centrifugée à 15 600 RPM à 4°C pendant 20 min. Le surnageant contenant les protéines totales est récupéré.

Dosage des protéines

L'extrait est dosé par la méthode BCA en utilisant le protocole suggérée par le fournisseur (Pierce BCP Protein Assay Kit 23227, Thermo Scientific).

Une courbe d'étalonnage est tracée en utilisant des solutions de BSA de concentrations connues.

L'extrait protéique est dilué au 1/5ème et 1/10ème, pour s'assurer que les densités optiques obtenues se situent dans le domaine de linéarité de la courbe d'étalonnage.

Le dosage des protéines est effectué en mélangeant l'extrait dilué avec le réactif BCA et en mesurant l'absorbance à 562 nm. La concentration en protéines est déterminée en comparant l'absorbance de l'échantillon à la courbe d'étalonnage.

Dénaturation et linéarisation des protéines

Les protéines sont, par la suite, linéarisées en ajoutant, à l'extrait protéique, du tampon Laemmli 2x (1610737, Bio-Rad) complété avec 5% de β-Mercaptoéthanol dans un rapport de 1:1 v:v. Le mélange est chauffé à 95°C pendant 5 min. Ce traitement permet de rompre les ponts disulfures et autres liaisons intramoléculaires, tout en alourdissant l'échantillon pour faciliter sa mise en place dans les puits du gel. Le SDS présent dans le tampon Laemmli permet quant à lui de conférer une charge négative uniforme aux protéines et d'empêcher ainsi leur précipitation. Les échantillons sont stockés à -80°C.

III.4.1.2. Migration sur gel

Préparation du gel

1. **Préparation de la solution de polyacrylamide :** une solution à 10% d'acrylamide est préparée à partir d'acrylamide : bisacrylamide 29 : 1 40% (EU077-B, Euromdex), de solution 4x Tris/SDS pH 8,8 et d'eau ultra-pure pour le gel de séparation. Une autre solution à 4% est également préparée à partir d'acrylamide : bisacrylamide 29 : 1 40% et de solution 4x Tris/SDS pH 6,8.
2. **Polymérisation du gel** : un initiateur (APS pour Ammonium persulfate 10%) et un catalyseur (TEMED pour N,N,N′,N′-tetramethylethylene-diamine, 17919, Thermo Scientific) sont ajoutés à la solution d'acrylamide juste avant le coulage du gel. La solution doit être mélangée délicatement et le gel coulé rapidement, la polymérisation du gel étant rapide.
3. **Coulage du gel de séparation** (10%) : dans un support à gel. Après polymérisation, un gel de concentration (4%) est ensuite coulé sur le gel de séparation pour synchroniser le début de la séparation.

Séparation des protéines

La séparation se fait en cuve dans un tampon de migration à 25 mM Tris, 192 mM Glycine et 0,1% SDS (TGS) et grâce à un courant électrique.

Après avoir rempli la cuve de migration de tampon TGS, les puits sont chargés avec les échantillons de protéines préalablement dosées et linéarisées. Les échantillons chargés correspondent à une quantité de protéines totales de 20 µg, et sont accompagnés d'une échelle de poids moléculaire *Protein Ladder plus* (EuroMedex, 06P-0211) à raison de 5 µL dans le premier puits.

La migration est effectuée à 100 V pendant environ 2h.

III.4.1.3. Transfert sur membrane

Le transfert permet de rendre accessible à la détection ultérieure les protéines séparées sur gel, en les décalquant sur une membrane en nitro-cellulose (NC) 0,45 µm (Amersham Protran Supported 0,45 µm NC, 1060018, Cytiva). Ce processus utilise l'électrophorèse pour déplacer les protéines du gel vers la membrane sous l'effet d'un champ électrique.

Le transfert s'effectue par la méthode du transfert humide, où le gel et la membrane sont immergés dans un tampon de transfert. Ce tampon est constitué de tampon Tris-glycine (TG) préparé dans de l'eau froide et complété avec 20% d'éthanol. La fixation des protéines sur la membrane est favorisée par ses propriétés de liaisons non spécifiques, incluant des interactions hydrophobes et ioniques.

Lors de l'assemblage du système de transfert, il est crucial d'éliminer soigneusement toutes les bulles d'air et de respecter le sens du courant électrique. Étant donné que les protéines portent une charge négative, elles migreront du gel (proche de la cathode) vers la membrane (plus proche de l'anode). Le transfert est réalisé à une tension de 90V sur glace pendant une durée de 2h.

A l'issue du transfert, la membrane est colorée au Rouge Ponceau afin de visualiser et de confirmer ainsi le transfert efficace des protéines.

III.4.1.4. Détection

III.4.1.4.1. Saturation de la membrane

Afin de bloquer les sites de liaison non spécifiques sur la membrane et d'éviter la liaison non spécifique des anticorps, la membrane est saturée avec une solution de lait écrémé dilué à 5% dans du TBS contenant 0.05% de Tween 20 (TBST). La saturation est réalisée pendant 1 heure à température ambiante, sous agitation douce.

III.4.1.4.2. Incubation avec l'anticorps primaire

La membrane saturée est incubée avec un anticorps primaire spécifique à la protéine d'intérêt. Le Tableau 6 résume les anticorps utilisés, leur dilution et les protéines ciblées. Tous les anticorps sont dilués dans le tampon de blocage : TBST-lait 5%.

Anti-corps	Référence	Dilution	Protéine recherchée
Anti-GPER (lapin)	Ab-39742 (Abcam)	1 : 600	55 KDa mais observée à 44 KDa
Anti-ERα (souris)	SC-8002 (Santa Cruz)	1 : 200	66 ou 54 kDa
Anti-ßactin (souris)	Ab-8226 (Abcam)	1 : 1000	42 KDa

Tableau 6. Liste des anticorps primaires et de leurs usages correspondants lors de l'analyse par Western Blot

Quelques précautions à prendre sont recommandées.

- Jeter le contenu du tube et bien l'égoutter pour ne pas diluer les anticorps primaires.
- Incuber la membrane pendant la nuit avec l'anticorps primaire dilué à 4°C sous agitation douce.
- Laver la membrane 3 fois au TBST pendant 5 minutes sous agitation douce à chaque lavage.

III.4.1.4.3. Incubation avec l'anticorps secondaire et visualisation

- Préparer les anticorps secondaires : anti-souris (sc-516102, Santa Cruz Biotechnology) et anti-lapin (sc-2357, Santa Cruz Biotechnology) dilués au 1 : 8000 dans du TBST-lait.
- Incuber chaque membrane avec l'anticorps secondaire correspondant pendant 1 heure à température ambiante sous agitation douce.
- Éliminer l'anticorps secondaire non lié et laver la membrane 3 fois au TBST pendant 6 min à chaque lavage.
- Préparer le substrat ECL selon les instructions du kit (Clarity Western ECL Substrate, 170-5060, Bio-Rad). Préparer 800 µL en mélangeant 1 : 1 v : v les deux solutions du kit.

- Déposer la membrane sur un film transparent, appliquer les 800µL de substrat ECL sur un coin et refermer la pochette transparente pour chasser les bulles d'air.
- Attendre 5 minutes, puis égoutter la membrane sur du papier absorbant et la placer au centre du Chemidoc (Bio-Rad). Vérifier l'absence de bulles d'air.
- Réaliser une première lecture colorimétrique puis en chimiluminescence et ajuster le temps de révélation si nécessaire.

III.4.1.4.4. Contrôles et normalisation

Pour normaliser la quantité de protéines dans les différents échantillons, une protéine de ménage, également appelée témoin de charge, est utilisée comme contrôle interne. Elle sert à confirmer que la quantité de protéines déposée est identique sur l'ensemble des pistes du gel, ce qui permet de normaliser le niveau de protéines détectées. La protéine de ménage choisie est la βactine, car son niveau d'expression est généralement constant quel que soit le type de tissu ou de cellule et quelles que soient les conditions expérimentales.

Pour se faire, la même membrane est tout d'abord rincée de manière prolongée au TBST pendant 48 heures en changeant le TBST régulièrement (au moins 5 changements).

Les opérations d'incubation primaire, secondaire et visualisation pour la β-actine sont répétées (incubation 1 heure à température ambiante, rinçages puis incubation secondaire 1 heure à température ambiante, rinçages et révélation au substrat ECL).

Afin de valider l'expression du récepteur couplé aux protéines G (GPER) dans les modèles cellulaires sélectionnés, une protéine recombinante GPER a été synthétisée par transfection d'un plasmide codant pour le GPER dans des cellules Expi293®. L'extrait protéique purifié a ensuite servi de contrôle positif pour confirmer la spécificité de la détection du GPER dans les cellules.

III.4.2. Immunocytochimie

L'immunocytochimie est une technique de biologie médicale d'immunofluorescence permettant de visualiser la présence, l'emplacement et la distribution de protéines spécifiques ou d'antigènes dans les cellules. Le protocole opératoire est décrit ici.

Jour 1 : Culture cellulaire

Placer des lamelles stérilisées dans une plaque 24 puits. Ensuite ensemencer les cellules cancéreuses dans les puits selon les densités cellulaires suivantes à raison de 500µL par puit, en triplicat :

- MDA-MB-468 : 30 000 cellules/mL
- MDA-MB-436 : 50 000 cellules/mL
- MCF7 : 50 000 cellules/mL
- SKBR3 : 100 000 cellules/mL

Incuber à 37°C et 5% de CO_2 pendant 24 heures.

Jour 2 : Marquage et détection

- Retirer le milieu de culture des puits et rincer 2 fois avec du PBS (500 µL par rinçage).

- Fixer les cellules avec du Paraformaldéhyde 4% pendant 10 minutes.
- Laver les lamelles deux fois avec du PBS et les conserver dans du PBS jusqu'au marquage.
- Déposer une goutte de PBS (20 µL) sur un parafilm.
- Déposez les lamelles sur les gouttes de PBS, en veillant à ce que la face avec les cellules soit immergée.
- Préparer la solution de perméabilisation (PBS avec 0,1% de Triton X-100 (X100, Sigma Aldrich)), retourner les lamelles et déposer des gouttes de cette solution sur les lamelles. Incuber les lamelles dans la solution de perméabilisation pendant 2 minutes, ensuite laver 3 fois au PBS.
- Saturer les lamelles avec du tampon de blocage (1% BSA, 2% SVF dans du PBS) pendant 5 minutes et préparer la solution d'anticorps primaire anti-GPER (dilution 1 : 250 dans du PBS).
- Déposer 20 µL de solution anti-GPER sur chaque lamelle et incuber 1 heure à température ambiante dans l'obscurité.
- Préparer le mélange d'anticorps secondaire anti-lapin (dilution 1 : 100 dans du PBS) et déposez 20 µL du mélange sur chaque lamelle.
- Incuber les lamelles avec l'anticorps secondaire pendant 1 heure à température ambiante dans l'obscurité.
- Laver les lamelles à l'eau, les déposer sur une lame, cellules vers le bas et monter le système avec du liquide de montage sans créer de bulles d'air.

Enfin, lire les résultats au microscope confocal (x63).

III.5. Modélisation moléculaire

III.5.1. Construction par homologie

L'absence de structure cristalline pour le GPER a nécessité la construction d'un modèle de récepteur par analogie de structure. Ce principe repose sur l'hypothèse que les structures tridimensionnelles sont mieux conservées que les séquences d'acides aminés [224].

La référence uniprot Q99527 correspond au GPER.

III.5.1.1. Etapes de construction du modèle

Recherche de structures homologues

La première étape consiste à effectuer une recherche par BLAST (alignement de séquences) sur la plateforme *Uniprot* afin d'identifier des protéines homologues à la protéine d'intérêt. Les résultats obtenus sont ensuite classés par ordre de similarité décroissante. Seules les protéines vérifiées par les paires, ayant une structure 3D connue, idéalement déterminées par RMN, humaines et présentant le meilleur score *e-value,* sont conservées. Ce score représente la probabilité attendue d'obtenir un alignement au moins aussi bon que celui observé par pur hasard. Autrement dit, plus la valeur e est faible, plus l'alignement trouvé par BLAST est statistiquement significatif.

Vérification et préparation des structures

Chaque structure sélectionnée est ensuite vérifiée en la comparant au fichier FASTA de la protéine d'intérêt, et en utilisant le logiciel Discovery Studio® (version 3.1) pour générer un rapport complet (*protein_report*). Le format FASTA est un format de fichier

texte couramment utilisé pour stocker des séquences d'acides nucléiques et d'acides aminés.

Construction des modèles de GPER actif et inactif

Les modèles de GPER actif et inactif sont construits à partir des structures sélectionnées précédemment, en utilisant le logiciel Modeler®. Ce dernier permet de reconstruire les chaines latérales de la protéine d'intérêt tout en conservant le squelette de la structure de référence.

Pour se faire, la séquence de référence est modifiée en ajoutant les gaps de l'alignement avec le GPER. La structure est ensuite alignée à la nouvelle séquence de la protéine et les résidus supplémentaires sont supprimés. Les boucles manquantes du GPER par rapport à la référence sont ensuite complétées en sélectionnant la structure dont l'énergie est la plus faible afin d'éviter des contacts entre résidus voisins.

Pour chaque conformère, 20 structures sont obtenues. Elles sont classées par niveau d'énergie et par écart quadratique moyen (RMSD) entre le modèle et la structure obtenue. Un RMSD ≥2 correspond au seuil empirique de différence entre deux protéines.

Établissement d'un ensemble d'apprentissage

Un ensemble d'apprentissage (*training_set)* est établi pour valider le modèle de récepteur et sélectionner le conformère à tester. Pour cela, les molécules liant le GPER sont extraites de la base ChEMBL et de la littérature. Ces molécules sont ensuite minimisées et leurs structures corrigées si nécessaire. La minimisation consiste à déplacer les atomes d'une molécule afin d'obtenir la conformation la plus stable possible (qui possède une énergie plus basse que le point de départ). Le *training_set* retenu comprend les molécules regroupées dans le Tableau 7.

Molécule	Activité	Référence
17β_estradiol (E_2)	Agoniste complet	[225,226]
3_ACS_MCL_	Inactif	[227]
9_ACS_MCL_	Agoniste	[227]
1_ACS_MCL_	Agoniste	[227]
4_ACS_MCL_1	Inactif	[227]
5_ACS_MCL_	Agoniste	[227]
2_ACS_MCL_1	Inactif	[227]
G1	Agoniste complet	[228]
Fulvestrant	Agoniste complet	[225]
Tamoxifène	Agoniste complet	[225]
4OH_tamoxifène	Agoniste	[226]
Raloxifène	Agoniste complet	[166]
G15	Antagoniste	[178]
G36	Antagoniste	[229]
2MeO_estradiol	Agoniste	[230]
4_ACS_MCL_2	Inactif	[227]
2_ACS_MCL_2	Inactif	[227]
17α_estradiol	Inactif	[228]

Tableau 7. Liste des molécules du training_set, avec leurs activités correspondantes vis à vis du GPER

Validation du modèle de récepteur

La dernière étape consiste à valider le modèle de récepteur en testant le *training_set* sur tous les modèles de GPER construits précédemment et à sélectionner le plus favorable. Ce choix est effectué par traitement statistique afin de déterminer les meilleurs conformères de récepteurs et les fonctions de score les plus pertinentes.

D'abord des domaines d'ancrage moléculaire (*docking*), appelés *box*, sont définis dans la protéine. La *big-box* correspond à la totalité du récepteur pour le *docking* aveugle, testant les interactions moléculaires des ligands sur l'ensemble de la protéine et permettant d'identifier les sites actifs. Les coordonnées de la *small-box* sont déduites après identification de la zone des sites actifs potentiels. La *small-box* doit être la plus resserrée possible tout en étant suffisamment large pour permettre de retrouver les poses obtenues grâce à la *big-box*.

Le *docking* proprement dit est effectué à l'aide du logiciel AutoDock Vina® (version 1.1.2) avec un ligand flexible et une exhaustivité de 64 (nombre d'exécutions aléatoires initiales pour la recherche et l'optimisation de la conformation du ligand).

III.5.1.2. *Rescoring*

Le *rescoring* consiste à appliquer différentes fonctions de score pour évaluer la qualité des poses prédites par le *docking*. Ces fonctions mathématiques permettent d'estimer l'affinité de liaison entre un récepteur et un ligand, dans une pose donnée. Trois grandes familles de fonctions de score sont communément utilisées [231].

- Champ de force : représenté par la fonction Vina, utilisée par défaut par AutoDock Vina®. Le calcul du score prend en compte l'énergie interne du ligand ainsi que l'énergie d'interaction entre lui et le récepteur. Cette fonction est exprimée en kcal/mol [232].

- Fonctions empiriques : représentées par le groupe de fonctions X-Score avec HPScore, HMScore, HSScore et AV. Elles sont conçues pour reproduire des données d'affinité obtenues expérimentalement. Elles comprennent des termes tenant compte de l'interaction de Van der Waals, des liaisons hydrogènes, d'une pénalité de déformation par torsion et de l'effet hydrophobe. Trois algorithmes différents ont été mis en œuvre pour calculer le terme de l'effet hydrophobe, ce qui se traduit par trois fonctions de score parallèles (HPS, HMS, HSS), combinées en une fonction consensuelle (AV) [233].

- Potentiels statistiques : représentés par le groupe de fonctions *drugscore* générées avec des ligands de la *Protein Data Bank* (PDB) et de *Cambridge Structural Database* (CSD). Chaque sous-famille possède trois fonctions différentes : distance entre paire d'atomes (*pair*), surface accessible au solvant (*surf*) et un consensus atome et surface (*pair-surf*). Elles sont conçues pour reproduire des données expérimentales de structure et non d'énergie de liaison, à la différence des fonctions empiriques. Ces fonctions sont basées sur l'analyse statistique de complexes dont la structure est connue, permettant de calculer la fréquence d'occurrence des contacts entre paires d'atomes, afin d'extraire des potentiels de force moyenne [224]. Ce modèle tient compte de manière implicite des effets de solvatation et d'entropie, tout en reproduisant correctement la géométrie des interactions directionnelles. Il reste indépendant des hypothèses concernant les

états de protonation dans la mesure où il ne considère que les contacts avec les atomes non hydrogènes [234].

Chaque fonction de score a ses propres avantages et inconvénients. Pour pallier aux limitations des fonctions individuelles, l'approche du consensus propose de les combiner, augmentant ainsi la probabilité d'identifier des molécules ayant réellement une affinité pour la cible [224].

Analyse des résultats

Les données obtenues après le *rescoring* sur le *training-set* sont combinées et analysées statistiquement.

- Analyse ROC (*Receiver Operating Characteristic* ou Fonction d'efficacité du récepteur) : permet d'identifier les conformères capables de différencier les ligands selon leur activité. Une codification binaire est appliquée aux activités du *training_set* : 1 pour actif et 0 pour inactif.
- Analyse en composantes principales (ACP) : permet d'identifier les fonctions de score les plus fortement corrélées aux composantes principales pour les récepteurs sélectionnés. Cette méthode a démontré sa valeur ajoutée dans l'identification de ligands d'intérêt [231].

Limites de l'étude

Il est à noter que pour l'ensemble des calculs effectués pour cette étude, la dynamique naturelle des processus d'arrimage ligand/récepteur, où les deux parties sont susceptibles de subir des changements conformationnels importants, n'a été que partiellement prise en considération. En effet, de la flexibilité sur le ligand a été autorisée dans les calculs, permettant d'anticiper les changements conformationnels possibles de ce dernier. Pour ce faire, les rotations sur les liaisons chimiquement valides sont autorisées pour le calcul des énergies de *docking*. Cependant, la protéine a été considérée comme rigide, constituant ainsi une des limites de ces calculs, bien que cette configuration ligand flexible-protéine rigide soit la plus largement utilisée aujourd'hui, pour gagner du temps de calcul [224].

III.5.2. Ancrage moléculaire (*docking*) et notation (*scoring*)

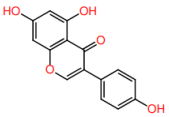
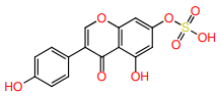
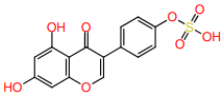
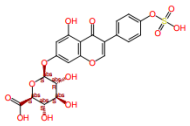
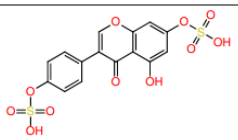
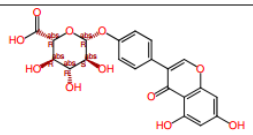
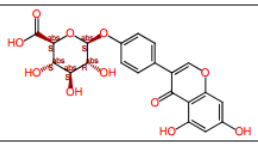
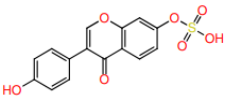
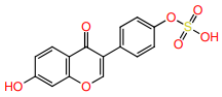
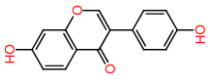
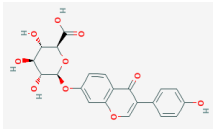
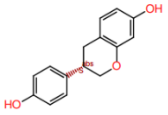
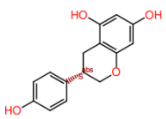
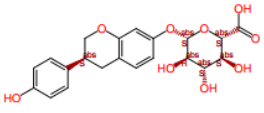
Après sélection des conformères les plus discriminants pour les ligands actifs, l'étape de *docking* proprement dite peut commencer. L'automatisation des calculs est possible grâce à des scripts dédiés.

III.5.2.1. Sélection des ligands à tester

La base de données *Life Chemicals* a été interrogée afin de sélectionner 5 340 composés présentant une similitude d'au moins 30% avec GEN parmi 541 996 composés naturels. La similarité a été évaluée par l'indice de Tanimoto selon les structures 2D.

De plus, 19 IFs et conjugués ont été ajoutés à la liste des molécules à tester. Elles sont listées dans le Tableau 8.

Les molécules pour lesquelles il n'y avait pas de structure 3D ont été reconstruites puis minimisées.

Substances		Pubchem CID
GEN		5280961
Génisteine 7-Sulfate		10291508
Génisteine 4'-Sulfate		53997099
Génisteine 4'-sulfate 7-glucuronide		54020950
Génisteine 4',7-Disulfate		44602471
Génisteine-7-Glucuronide		124934066
Génisteine 4'-O-Glucuronide		45782816
Daidzéine 7-Sulfate		12114465
Daidzéine 4'-Sulfate		12114463
DAI		5281708
Daidzéine 7-O-Glucuronide		23930394
Equol		91469
5-OH-équol		9795113
Equol 7-O-Glucuronide		29979359

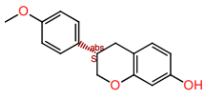
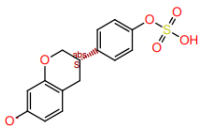
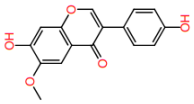
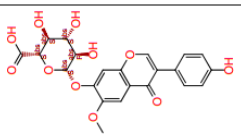
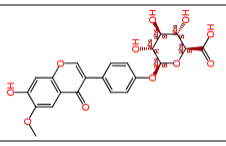
4-O-Méthyl équol		6454212
Equol 4'-Sulfate		156960885
Glycitéine		5317750
Glycitéine 7-O-Glucuronide		45782894
Glycitéine 4'-O-Glucuronide		124202362

Tableau 8. Liste des isoflavones et conjugués utilisées pour le test de docking

III.5.2.2. *Docking*

La zone d'amarrage a été confinée à la *small-box* et les calculs de positionnement ont été effectués à l'aide du logiciel AutoDock Vina® avec ligand flexible et une exhaustivité de 64.

III.5.3. *Rescoring* et consensus

Les méthodes de *docking* et de *scoring* présentent plusieurs limitations, parmi lesquelles peuvent être citées les suivantes.

- Manque de flexibilité des protéines : nous rappelons que les modèles utilisés pour représenter les protéines dans les simulations de *docking* ne prennent souvent pas en compte leur flexibilité conformationnelle, ce qui peut conduire à des prédictions inexactes des poses de liaison des ligands.
- Traitement insuffisant de la solvatation : l'effet du solvant sur les interactions entre les ligands et les protéines est souvent négligé dans les calculs de *scoring*, ce qui peut altérer la précision des évaluations de liaison.
- Simplicité de la fonction énergétique : les fonctions énergétiques employées pour évaluer les interactions entre les ligands et les protéines sont souvent simplifiées et ne capturent pas l'intégralité des forces intermoléculaires en jeu, ce qui peut limiter la fiabilité des classements des poses de liaison.
- Coût computationnel élevé : les simulations de *docking* et de *scoring* peuvent être gourmandes en ressources computationnelles, ce qui restreint le nombre de composés pouvant être testés et la précision des explorations conformationnelles.

Ces limitations inhérentes aux méthodes actuelles de *docking* et de *scoring* conduisent à des inexactitudes dans l'évaluation et le classement des poses de liaison des ligands, affectant ainsi la performance des approches de criblage virtuel. Celles-ci doivent donc être considérée comme des compléments aux études cellulaires *in vitro*.

Approches de *scoring* par consensus

L'utilisation d'approches de *scoring* par consensus a été proposée pour pallier aux limitations des méthodes individuelles et améliorer la précision des prédictions de liaison des ligands. Cette stratégie consiste à combiner les résultats de plusieurs fonctions de *scoring* distinctes afin d'obtenir un score global plus robuste et plus fiable.

Cependant, le développement d'approches efficaces de *scoring* par consensus se heurte à plusieurs défis : les scores de liaison calculés par différentes fonctions de notation sont exprimés dans des unités et des échelles différentes, les fonctions de *scoring* utilisées proviennent souvent de catégories distinctes (champ de forces, fonctions empiriques et potentiels statistiques) et enfin, la relation linéaire entre de nombreuses fonctions de notation ne permet pas d'atteindre des taux d'enrichissement optimaux [231]. Le taux d'enrichissement, un indicateur clé de performance du criblage virtuel, mesure sa performance à identifier les vrais positifs (ligands actifs) parmi l'ensemble des candidats testés. Il s'agit du rapport entre le nombre de vrais positifs identifiés et le nombre total de candidats testés.

Il existe trois manières d'obtenir un consensus.

- Classement par le nombre : cette méthode calcule une moyenne des scores de liaison obtenus pour chaque ligand par toutes les fonctions de *scoring*. Cependant, cette stratégie quantitative est peu applicable en raison du manque de cohérence entre les unités utilisées par les différentes fonctions.
- Classement par le rang : cette approche calcule la moyenne des rangs obtenus par chaque ligand pour chaque fonction de *scoring*. Elle permet de s'affranchir des problèmes d'unités mais peut lisser les variations d'énergie importantes qui pourraient être révélatrices de la distinction entre des ligands très affins et d'autres qui le sont moins.
- Classement par vote : cette méthode réalise une intersection mathématique des têtes de listes issues de chaque fonction de *scoring*, en attribuant des points aux ligands qui figurent en tête de liste pour plusieurs fonctions [224].

Dans le cadre de ce travail, c'est l'approche par vote qui a été retenue. En pratique, après sélection des 3 meilleurs fonctions de *scoring* (une par catégorie), les composés figurant en tête de liste pour chacune d'entre elles ont été identifiés. Seules les molécules communes aux 3 listes ont été retenues pour le consensus. Cette approche se fonde sur l'hypothèse que ces molécules présentent une probabilité plus élevée d'être réellement capables de se lier à la cible que celles identifiées par une seule fonction de *scoring* [231,235]. Cette hypothèse a été étayée par plusieurs études et pourrait s'expliquer par une réduction du nombre de faux positifs en tête de liste [224].

III.6. Analyses statistiques

Le choix des méthodes d'analyse statistique a été dicté par la nature des données et questions explorées. Cette section détaille aussi le choix des méthodes employées.

Les analyses statistiques ont été réalisées avec le logiciel RStudio (R version 4.2.2 « Innocent and Trusting » © 2022 The R Foundation for Statistical Computing Platform ; R core Team, 2015), sauf si spécifié autrement.

L'ensemble du code R utilisé se trouve en Annexe 1.

III.6.1. ISOLED

Le premier objectif de cette étude a été de développer des outils afin d'estimer des scores alimentaires de IFs et ENL dans le plasma, les urines et les cheveux. Ces variables sont les suivantes.

- IHS (Isoflavone Hair Score) : score capillaire d'IFs (corrigé selon les traitements, la longueur des cheveux et la durée du régime).
- IFS (Isoflavone Food Score) : score d'habitudes alimentaires non corrigé.
- Score_urin : score urinaire de l'exposition aux IFs dans les 24-48 heures.
- Score_urin_G : score urinaire de l'exposition à la GEN dans les 24-48 heures.
- Score_urin_D : score urinaire de l'exposition à la DAI dans les 24-48 heures.
- Score_urin_IFs : score urinaire de l'exposition aux IFs (µg/L) dans les 24-48 heures.
- Score_urin_ENL : score urinaire de l'exposition aux lignanes dans les 24-48 heures.
- Score_plasm_G : score plasmatique de l'exposition à la GEN dans les 24-48 heures.
- Score_plasm_D : score plasmatique de l'exposition à la DAI dans les 24-48 heures.
- Score_plasm_IFs : score plasmatique de l'exposition aux IFs (µg/L) dans les 24-48 heures.
- Score_plasm_ENL : score plasmatique de l'exposition aux lignanes dans les 24-48 heures.

Pour tester la validité de ces outils, nous avons comparé les scores calculés à partir des résultats des questionnaires aux dosages biologiques réellement observés chez 57 participantes à l'étude ISOLED dans une première étape, puis 63 participantes à sa clôture. Les variables dosées sont les suivantes.

- Hair_G_D : concentration de GEN et de DAI dans les cheveux (ng/mg)
- Hair_IFs : concentration totale en IFs dans les cheveux (ng/mg)
- Urine_G : concentration urinaire de GEN (µg/L)
- Urine_E : concentration urinaire d'équol (µg/L)
- Urine_D : concentration urinaire de DAI (µg/L)
- Urine_IFs : concentration urinaire d'IFs totales (µg/L)
- Urine_ENL : concentration urinaire d'ENL (µg/L)
- Plasma_G : concentration plasmatique de GEN (ng/mL)
- Plasma_D : concentration plasmatique de DAI (ng/mL)
- Plasma_E : concentration plasmatique d'équol (ng/mL)
- Plasma_IFs : concentration plasmatique totales d'IFs (ng/mL)
- Plasma_ENL : concentration plasmatique d'ENL (ng/mL)

Des outils statistiques différents ont été utilisés pour ces deux étapes. Cette section est donc présentée en deux parties distinctes.

III.6.1.1. Analyse à n=57

L'étude était encore ouverte au recrutement au moment de la première analyse des données. Le test de corrélation paramétrique de Pearson a été utilisé pour valider les questionnaires alimentaires et les scores correspondants après transformation logarithmique des données (voir les publications *Development of an Assay for Soy Isoflavones in Women's Hair* et *Validation of ELISAs for Isoflavones and Enterolactone for Phytoestrogen Intake Assessment in the French Population* dans le chapitre « Résultats et Discussions »). La transformation logarithmique permet de s'approcher d'une distribution normale. Ce choix statistique a été motivé par la possibilité d'utiliser un plus grand nombre d'outils (applicables aux données normales). De plus, l'utilisation

de cet outil est justifiée par l'objectif de comparaison avec les études précédemment publiées par d'autres auteur·es.

III.6.1.2. Analyse après la fin de l'inclusion

L'étude ayant été clôturée, le statut sanitaire des participantes a pu être révélé. Les 63 participantes sont donc séparées en 3 groupes : femmes saines (SAINES), femmes atteintes de LED (LED) et femmes atteintes de Purpura Thrombopénique Immunologique (PTI), une autre MAI moins clairement liée au sexe.

Afin de choisir un test statistique approprié, l'étape préalable est de tester la normalité de la distribution des données. La distribution peut être testée de manière analytique et graphique et il est recommandé d'utiliser ces méthodes indépendantes [236].

Pour le test analytique, nous avons choisi celui d'Anderson-Darling [237] dont une valeur de p calculée inférieure à 0,05 permet d'assumer que la distribution n'est pas normalement distribuée. Ce test permet de déterminer, pour un échantillon donné d'une distribution spécifiée, s'il faut accepter ou rejeter l'hypothèse que l'échantillon suit une distribution normale (notée hypothèse nulle ou H0). Une valeur de $p > 0,05$ signifie qu'il n'y a pas suffisamment de preuves pour rejeter l'hypothèse nulle, et donc par extension que la variable étudiée suit une distribution normale.

Une limite des tests analytiques est que cette valeur p est affectée par la taille de l'échantillon. En effet, plus n est grand, plus la probabilité de détecter de petites différences augmente et avec elle le risque de rejeter l'hypothèse de normalité pour des écarts sans signification. Pour pallier ce biais potentiel lié à la taille de l'échantillon, nous avons procédé à un test graphique des diagrammes Quantile-Quantile (QQ-plots). Cette méthode de QQ-plots compare les quantiles de la distribution réelle (en ordonné) par rapport à ceux de la distribution normale (en abscisse). Si les données sont normales, tous les points se situent sur une droite. Plus les données s'écartent de la ligne, moins elles sont normalement distribuées.

Les tests de normalité des données ont révélé que le jeu de données ne présente pas une distribution normale pour la majorité des données (cf. chapitre « Résultats et discussions »). Par conséquent, les effets du traitement sur les variables testées ont été analysés en utilisant des tests non paramétriques ou en appliquant des transformations pour normer les données.

Pour tester l'influence du statut sanitaire des participantes sur les scores et les variables dosées, des tests de Kruskall-wallis ont été réalisés. En cas de différence significative, une analyse post-hoc par tests de Wilcoxon corrigés par la méthode de Bonferroni [238] a été réalisée groupe par groupe.

Pour évaluer la validité de l'outil développé, les scores calculés ont été comparés aux dosages biologiques avec des tests non paramétriques de corrélation de rang de Spearman. Ce coefficient varie entre -1 et 1 et permet de déterminer l'intensité de la corrélation ainsi que la direction. Un rho $r<0,1$ montre une absence de corrélation, entre 0,1 et 0,3 une faible corrélation, entre 0,3 et 0,5 une corrélation moyenne, entre 0,5 et 0,7 une corrélation élevée et entre 0,7 et 1 une corrélation très élevée.

Afin de répondre à l'hypothèse de l'étude ISOLED et estimer si la consommation de phytoE peut être associée au déclenchement de poussées du LED, une ACP a été

effectuée. Après centrage et réduction des données, cette ACP (normée) a permis d'identifier les variables et leur covariante contribuant le plus à la description du jeu de données.

III.6.2. PHYTOSECUR 1 et 2

Chaque expérience de ce travail a été effectuée en triplicata. Les données de cette partie sont présentées sous forme de moyennes ± erreur type. La comparaison des valeurs avant et après traitement a été réalisée à l'aide du test non paramétrique des rangs de Wilcoxon-Mann-Whitney. Ce test a été choisi car il ne présume pas de la normalité des données. Le seuil de significativité a été fixé à $p < 0,05$.

III.6.3. *Docking* moléculaire

Afin d'identifier les meilleures configurations récepteur/fonction de score, permettant de discriminer efficacement entre les ligands actifs et inactifs, une combinaison de deux tests statistiques a été employée :

- Analyse ROC : L'analyse ROC permet d'évaluer et de comparer les performances d'un modèle de classification en se basant sur sa sensibilité et son taux de faux-positifs, pour un seuil de décision donné. Elle se fonde sur le calcul de l'aire sous la courbe. Plus le seuil est restrictif, plus les chances de détecter un faux positif sont élevées. Le modèle de classification interrogé dans la présente étude est la capacité d'un conformère de récepteur de discriminer les ligands selon leur activité et le seuil utilisé était de 1. L'analyse ROC a été effectuée sur le logiciel Discovery Studio®.
- ACP : analyse par composante principale a été utilisée pour identifier les fonctions de score qui expliquent le mieux la variance des données des simulations de *docking*. L'ACP est une méthode statistique qui consiste à réduire un ensemble de données à un nombre inférieur de dimensions virtuelles, appelées composantes principales. Ces composantes principales sont des combinaisons linéaires des variables originales et permettent de capturer l'information la plus importante contenue dans les données. En d'autres termes, l'ACP permet de transformer un jeu de données complexe en un ensemble de variables plus simples et plus faciles à interpréter. Appliqué au cas du *docking*, l'ACP a été appliquée aux scores obtenus pour chaque fonction de score pour toutes les molécules testées. L'objectif était d'identifier les fonctions de score qui contribuent le plus à la variabilité des scores et qui, par conséquent, décrivent le mieux les interactions possibles entre les molécules testées et le récepteur cible.

III.6.4. Etudes cellulaires

Dans le cadre des études de prolifération, la significativité des différences entre les groupes a été évaluée à l'aide du test non paramétrique de Kruskall-Wallis, adapté aux petits échantillons et aux données non normales. En cas de différence significative entre les groupes ($P < 0,05$), un test de Wilcoxon a été réalisé par paires pour identifier les groupes spécifiques présentant des différences significatives. Le résultat a été corrigé par la méthode de Bonferoni pour réduire le risque d'erreur due aux nombreuses itérations.

IV. Résultats et discussions

Les résultats et les discussions des travaux conduits en adéquation avec différents volets de cette thèse (décrits dans le chapitre « Introduction Générale ») ont fait l'objet de publications, intégralement reproduites dans les sections correspondantes pertinentes. Ceux relatifs à l'analyse statistique étudiant le lien entre les phytoE et le LED, au *docking* moléculaire et à l'étude de prolifération, n'ayant pas encore été publiés, seront développés de manière usuelle.

IV.1. Exposition aux isoflavones en France

IV.1.1. Comparaison de l'exposition des Françaises et Français aux pesticides perturbateurs endocriniens et aux isoflavones du soja : article correspondant à la référence [213]

Cette section concerne les informations collectée afin d'évaluer la part des IFs dans l'exposition aux xéno-estrogènes environnementaux, notamment les PEs d'origine alimentaire présents dans et sur les fruits et légumes, en France [213].

Les nouvelles données obtenues justifient une attention accrue à ces substances et à leurs effets potentiels sur la santé. Ainsi notre équipe a pu montrer que l'exposition aux IFs ne pouvait être ignorée en France, justifiant l'étude ISOLED qui a été menée par la suite.

Food Chemistry 373 (2022) 131513

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Endocrine disruptors on and in fruits and vegetables: Estimation of the potential exposure of the French population

Alexandre Lee [a,b], Souad Bensaada [a,c], Valérie Lamothe [a,b], Melissa Lacoste [a,b], Catherine Bennetau-Pelissero [a,b,c,*]

[a] University of Bordeaux, 33070 Bordeaux France
[b] Bordeaux Sciences Agro, 33175 Gradignan France
[c] U1212 Inserm, UMR Inserm U1212, CNRS 5320, University of Bordeaux, 33070 Bordeaux France

ARTICLE INFO

Keywords:
Endocrine disruptors
Plant protection products
Isoflavones
Consumer exposure
Maximum residue limits

ABSTRACT

Fruits and vegetables are considered to be healthy compared to fats, carbohydrates, and meats. However, their production involves plant protection products (PPPs) or they can contain phytoestrogens which may exhibit endocrine effects. Thus, the exposure to the main PPPs and to phytoestrogens known as endocrine disruptors (EDs) is estimated. PPPs include fungicides, growth substances, herbicides, and insecticides authorised in France. ED-PPPs exposure is estimated from the maximum residue limits (MRLs) of 70 potential ED-PPPs used in France on 64 fruits and vegetables. The estimated exposure to potential ED-PPPs is 509 μg/d and involves agonist and antagonist substances in complex mixtures. Anti-androgens are preeminent, at 353 μg/d. Exposure to genistein and daidzein is calculated from 140 measurements in 9 categories of food-items containing soy. The global exposure to isoflavones in France is evaluated at 6700 μg/d. Phytoestrogen exposure is much higher than that of ED-PPPs. Their endocrine effects should be considered.

1. Introduction

Nowadays, consumers are advised to eat more fruits and vegetables and many studies showed a real benefit in consuming these products instead of a diet rich in meat, fat, and carbohydrates (Wallace et al., 2019). Indeed, fruits and vegetables prevent food-related diseases like diabetes, obesity, cardiovascular ailments, metabolic syndrome, and some cancers (Wallace et al., 2019). However, their production at industrial scales uses plant protection products (PPPs) that include pesticides, *i.e.*, fungicides, herbicides, insecticides, and growth substances. Several of these anthropic substances with biocide properties were also reported to be endocrine disruptors (EDs) (Rizzati et al., 2016). Additionally, some pulses with high nutritional value, including soybean, were shown to protect themselves from predation by producing anti-nutritional substances (Bennetau-Pelissero, 2019). Most of them are enzyme inhibitors or allergens, and molecules with hormone-like effects. These substances were removed from food by traditional cooking practices. Nowadays, the industrialisation of food processing has led to a dramatic increase in such hormone-like substances in certain foodstuffs (Lee et al., 2019). Therefore, the exposure to isoflavones has risen in France, and it was shown that 12% of premenopausal women inquired had more than 9 soy-portions per week and more than 50 mg of isoflavones per day, in 2017 (Lee et al., 2019). The effect of hormone-like substances depends on potency, exposure and organ-specificity. Basically, physiological effects were recorded in human-beings at doses of 20 to 60 mg/day depending on the expression of the doses in aglycones or glucosides. These substances can either be beneficial or deleterious to consumers (Nakamura et al., 2017; Nakamoto et al., 2018, Nakai et al., 2020; Messina et al., 2021), depending on their physiological status (infancy, childhood, pregnancy, pre-menopause, post-menopause, obesity, hormone dependent pathologies…), and hence should not be present in a diet intended for use by the general population. Indeed, on the one hand these natural substances may be used in food supplements for some physiological improvement, and on the other hand, because of potential deleterious effects in certain population categories (infants, children, pregnant women, patients with estrogen-dependent pathologies or thyroid-dependent pathologies), they can be considered as EDs, and can act synergically with other EDs (Thongprakaisang et al., 2013).

According to the European Union, "an endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine

* Corresponding author at: University of Bordeaux, ARNA, U 1212 Inserm, 146 Léo Saignat Street, BP12 33076 Bordeaux, France.
E-mail address: catherine.bennetau@u-bordeaux.fr (C. Bennetau-Pelissero).

https://doi.org/10.1016/j.foodchem.2021.131513
Received 28 May 2021; Received in revised form 25 October 2021; Accepted 29 October 2021
Available online 1 November 2021
0308-8146/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

system and consequently causes adverse health effects in an intact organism, or on its progeny, or in (sub)populations" (Andersson et al., 2018). This definition talks about exogenous substances without any reference to their natural or synthetic origin. Hormones are essential to physiological functions that are crucial for an individual's or a population's survival. Endocrine disrupting effects can alter reproduction (Chevrier et al., 2013), metabolism with obesogenic effects (Janer et al., 2005), development, and neurodevelopment (Requena et al., 2019). Conversely, when physiology leads to a decline in hormone production, like at menopause, hormone supplementation can help restoring some physiological conditions. This is essentially the case where hormones and estrogenic isoflavones can be useful for certain population categories (Gómez-Zorita et al., 2020). Consequently, this field of investigation is growing intensely and knowledge in this area is in constant evolution.

Hence, the present study aims at providing an overview of the current situation in France through the assessment of the exposure to synthetic and natural EDs that may be present in French consumers' foodstuffs due to industrialisation. This work was conducted with regard to consumer protection and the resulting data may be overestimated and/or non-exhaustive. Oestrogens, androgens, and thyroid disruptors or modulators present inside and outside fruits and vegetables (including soybean) were all considered here, since they represent the largest portion of identified EDs. The study is limited to the effects reported in the literature and released in scientific or regulatory documents at the time of writing. It is a methodological approach and provides a time point assessment of the situation to help scientists advice regulators and consumers regarding their everyday choices.

EDs were considered regardless of the actual European Food Safety Authority (EFSA) definition and criteria that essentially considers *in vivo* data (Andersson et al., 2018). This was done to cover a large spectrum of substances and thus anticipate future classification and take a broader protective approach. Two types of effects were considered: (1) those observed *in vivo* either in human populations or in animals and (2) those reported *in vitro*. Whenever possible, the distinction between those two situations is given. The authorisations regarding PPPs in France are managed by ANSES: the French Agency for Food, Environmental and Occupational Health & Safety. The study is based on data gathered and analysed in the Mendeley data repository at https://data.mendeley.com/datasets/dk8gsx2r4j/9. Raw data on PPPs are available at https://ephy.anses.fr/. The French exposure to natural substances with hormone-like effects, *i.e.*, genistein and daidzein, is calculated from 140 measurements on food found in France and from a specific inquiry performed in supermarkets on current soy-based foodstuffs.

2. Methods

2.1. Calculation of exposure to endocrine-disrupting pesticides

The EU pesticide database (https://ec.europa.eu/food/plant/pesticides/eu-pesticides-database/public/?event=homepage&language=EN) was screened for EDs. The data selection followed the process described in Fig. 1-PPP-selection-scheme of the Mendeley data repository associated with this study (https://data.mendeley.com/datasets/dk8gsx2r4j/9). On the 1426 entries, the 460 substances approved in France (https://ephy.anses.fr/) were retained. Next, 379 fungicides, insecticides, acaricides, herbicides, and growth substances and their Maximum Residue Limits (MRLs) were considered in various fruits and vegetables. They are allowed in 1666 formulas in France. See table "French PPP-ED" at https://data.mendeley.com/datasets/dk8gsx2r4j/9. Seventy EDs were found by combining the name of each substance with the word "endocrine" and/or "disruption" and/or "disruptor" on the PubMed database, Google Scholar database, and Google. Consumption data were obtained from the official French websites https://www.lesfruitsetlegumesfrais.com/ and https://agreste.agriculture.gouv.fr/agreste-web/disaron/!searchurl/searchUiid/search/ in March 2020, which describe the annual consumption of fruits and vegetables by the French population. The population data were obtained in March 2020 from the Official French statistic service, Insee (https://www.insee.fr/fr/statistiques/1892086?sommaire=1912926). EDs in the different categories of PPPs, *i.e.*, insecticides, fungicides, herbicides, and growth substances, were determined using the scientific literature, and the list of references for this classification can be found at https://data.mendeley.com/datasets/dk8gsx2r4j/9 in "French-PPP-ED".

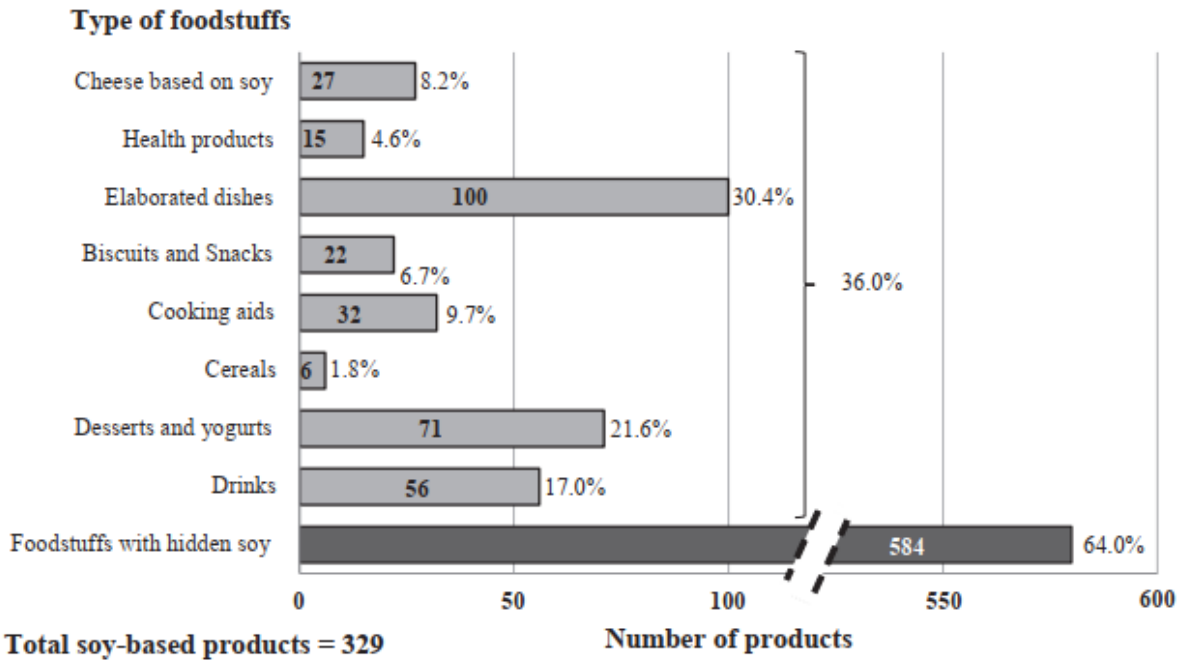


Fig. 1. Repartition of French soy-based foodstuffs in different categories according to a specific enquiry performed in March 2020 on the main 7 French supermarket websites. Foodstuffs containing hidden soy are separated.

For the calculation, 64 fruits and vegetables for which consumption data were available on the websites previously mentioned, were selected. The number of formulas containing ED-PPP substances allowed in France for these fruits and vegetables were entered into the table "PPP-French usage-ED" found at https://data.mendeley.com/datasets/dk8gsx2r4j/9.

For each of these 64 fruits and vegetables, the annual national consumption is reported per day and per capita in the table "Exposure-calculation-ED," found in the data repository at https://data.mendeley.com/datasets/dk8gsx2r4j/9. The annual consumption was reported for 64 million French consumers. This number includes all declared citizens according to the Insee website in 2019 and does not include children below 3-years-old (2,150,000 children) and elderly individuals over 89-years-old (912,000 individuals). For each foodstuff, the potential amounts of ingested pesticides were calculated from the legal MRLs obtained from the EU pesticide database and using a factor that takes into account the "popularity" of the substance for the considered use (*i.e.*, same plant and same use). The equation is:

$$\text{Pesticide daily intake} = \Sigma_{(1-64)}(\text{Gram of food per day} \times \text{MRL} \times \text{Popularity factor})$$

with 64 being the number of fruits and vegetables considered in this study
and Popularity factor = (N/Ntot) where
N: Number of formulas containing the substance
Ntot: Total number of formulas allowed for the same application (same plant, same use).

When a substance belonged to two categories of usage, Ntot is the sum of formulas allowed in the two categories. Therefore, the popularity factor reflects the global frequency of use of a substance for one plant and one application compared to the other substances possibly used in the same context. When a substance was allowed for one fruit or vegetable in France, all the formulas containing this substance were considered to be applicable. The calculation can be checked in the data repository https://data.mendeley.com/datasets/dk8gsx2r4j/9 "Exposure-calculation-ED". Finally, substances exclusively used for cereals, including wheat, oat, and rye, were not considered. For grapes, substances were not retained if they were specifically used for wine production. For corn, substances were retained if they were not for exclusive use on forages.

2.2. Isoflavone measurements

2.2.1. Materials

Genistein and daidzein measurements were performed on 140 foodstuffs from French supermarkets and allocated to 9 different categories (see file "Genistein and Daidzein values-ED" at https://data.mendeley.com/datasets/dk8gsx2r4j/9. For ELISA measurements, all salt reagents were from VWR-France. Antibiotics used for the digestion buffer, penicillin G (ref. P-3032) and streptomycin sulphate (ref. S-6501), were from Sigma-Aldrich. β-glucuronidase aryl-sulfatase (ref. 10127698001), thyroglobulin from swine (T1126), and o-phenylenediamine dihydrochloride (ref. P1526) were also from Sigma-Aldrich. Bovine serum albumin came from Euromedex (ref. 04-100-812-C) and secondary antibody goat anti-rabbit IgG antibody-Amdex (ref. RPN4301) came from GE Healthcare. Glassware was coated with silicone (Sigmacote ref. SL2).

2.2.2. Method

Genistein and daidzein were analysed in 140 foodstuffs based on their claimed compositions. For each foodstuff, three crushed fresh matter samples of 1 g each were suspended in 50 mL of distilled water under stirring. The extraction was performed for 20 min at room temperature and 10 min at 100 °C. A total of 500 µL of cooled extracts containing the glycosylated isoflavones were then hydrolysed into aglycones using β -glucuronidase aryl-sulfatase at 37 °C overnight, as described previously (Vergne et al., 2007). The aglycones were extracted 3 times with 3 mL of acidified ethyl-acetate and evaporated to dryness using a Speed-Vac (Thermo Electron Corporation). Each sample was then diluted in 500 µL of the assay buffer in a silicone-coated vial. The isoflavone analysis was performed using specific ELISAs as described previously (Vergne et al., 2007). Haptens functionalized on different carbon atoms were synthesized. Specific polyclonal antibodies were obtained for each isoflavone bound to bovine serum albumin and the best antibodies were retained for assay development. When cross-reactions were identified, harvesting of the antibody with the reacting compound allowed the removal of unwanted reactions. Standard curve preparations and sample dilutions were performed in silicone-coated glass vials. The ELISAs follow a competitive procedure with an immobilized competitor, the homologous hapten bound to swine thyroglobulin. The sensitivities varied between 0.08 ng/well and 0.4 ng/well; intra-assay variation was always below 7% and inter-assay variation was below 17%. For the foodstuffs, final dilutions varied between 1/250 and 1/40,000. Each food item was assayed in triplicate using three extracts and on three different microtitration plates. All values are given in aglycone equivalent and reflect isoflavone that can potentially reach the consumer plasma.

2.3. Isoflavone exposure estimation

A survey of soy-based foodstuffs available in France was performed in spring 2020 by looking at the seven main French supermarket websites, *i.e.*, Carrefour®, Auchan®, Leclerc®, Intermarché®, Super U®, Lidl®, and Monoprix®, as well as a few specialized shops. The entire list of products registered during this survey is given under "Phyto-Product-Enquiry-ED", at https://data.mendeley.com/datasets/dk8gsx2r4j/9. Each product was recorded with its composition only once, even if it was found on several websites. This survey gave a general picture of the percentages of the different types of products sold and consumed in France (Fig. 1). To obtain the exposure to isoflavones for each type of foodstuff, a mean dose, determined from isoflavone measurements, was calculated (Table 1). See "Genistein-Daidzein values-ED" at https://data.mendeley.com/datasets/dk8gsx2r4j/9 to find the individual genistein and daidzein measurements in the foodstuffs and the mean and SEM calculations. According to the Sojaxa and GEPV websites, the current human consumption of soy in France is 185,000 tons, broken into 118,400 tons used as hidden soy in elaborated dishes and 66,600 tons used to produce well-identified soy-based foodstuffs. According to GEPV (2019) the number of products containing hidden soy on the French market represent 64% of the total soy product consumption, *i.e.*, 584 products. Meanwhile, well-identified soy-based foodstuffs accounted for 329 food products, *i.e.*, 36% of soy consumption. The isoflavones (genistein and daidzein) exposure was calculated from the mean of isoflavones measurements (µg/g) in nine categories of soy foodstuffs at https://data.mendeley.com/datasets/dk8gsx2r4j/9 in "Genistein-Daidzein values-ED". A percentage of consumption was assigned to each soy food category and a mean percentage of soy in the final products was also calculated. The exposure calculations consider the mean amount of soy consumed per day and per capita in France, which is 8 g/d (https://data.mendeley.com/datasets/dk8gsx2r4j/9), the mean amount of isoflavones per gram of the different food type (Table 1) and, for better accuracy, the percentage of each food category present on the French market (Fig. 1). Two exposures were estimated: one for soy food consumers (60% of the population) and one for soy food non-consumers (40% of the population) who only ingest isoflavones through foodstuffs with hidden soy.

Table 1
Mean isoflavones contents in French soy-based foodstuffs and isoflavones per gram of soy proteins based on labelling and measurements.

Soy-based foodstuff	Portion size*	Isoflavones content µg/portion			Isoflavones content (µg/g of soy-protein) **			Number of measurements
		Mean	±	SEM	Mean	±	SEM	
Cheese based on soy	50 g	23,838	±	7,468	1542.9	±	483.3	4
Health soy products	83 g	37,631	±	9,086	3434.8	±	109.5	13
Elaborated dishes	120 g	29,203	±	2,751	839.2	±	79.0	40
Snacks and cakes	72.5 g	44,627	±	14,512	1832.0	±	200.2	8
Cooking aids	63 g	19,030	±	4,832	1706.6	±	433.4	17
Cereals	111 g	44,656	±	13,113	2272.9	±	667.4	6
Dessert, cream, yogurt	106 mL	9,805	±	1,663	1423.0	±	241.4	17
Soy-based drinks	250 mL (1 mug)	24,715	±	4,302	1373.0	±	239.0	9
Foodstuffs with hidden soy	*150 g*	*5,988*	±	*1,269*	*307.1*	±	*65.1*	*26*

* Portion size is the mean of portions proposed by manufacturers.
** Isoflavones per gram of claimed soy-protein. Differences are mainly due to food processing.

3. Results

3.1. Estimation of exposure to synthetic endocrine disruptors used as PPP

As mentioned previously, the full calculation can be found in "Exposure calculation-ED" at https://data.mendeley.com/datasets/dk8gsx2r4j/9. A total of 64 foodstuffs were retained because (1) they were cited either in the European or the French website for PPP applications and (2) consumption data were found on reliable French websites. The calculation led to a final fruits and vegetables mean consumption of 503 g/d, including rice.

Seventy substances were detected as potential EDs as mentioned in "French PPP-ED" at https://data.mendeley.com/datasets/dk8gsx2r4j/9. Some EDs were *de facto* not allowed in France since all the formulas containing them are forbidden. Among the selected substances, 52 had an effect *in vivo* and 18 had an effect only *in vitro*. The overall exposure to potential EDs was estimated at 509 µg/d, as shown at https://data.mendeley.com/datasets/dk8gsx2r4j/9 in "Exposure calculation-ED". According to this estimation reported at "Exposure summary-ED", the products reaching French citizens include 147 µg/d of potential oestrogens, 118 µg/d of potential anti-oestrogens, 3.2 µg/d of potential androgens, 353 µg/d of potential anti-androgens, 113 µg/d of potential pro-thyroid substances, and 73.8 µg/d of potential anti-thyroid substances. Finally, there were several substances with unclassified effects, like anti-PPARγ, anti-cholesterase, anti-FSH, or anti-LH effects. They account for 87.3 µg/d of potential ED exposure. Adding these figures leads to an exposure higher than 509 µg/d since some PPPs exhibit more than one effect. For instance, myclobutanyl seems to have anti-oestrogen, anti-androgen, and anti-thyroid effects. This estimation shows that the main contributors to potential ED-PPP exposure (greater than 20 µg/d) are, in decreasing order: apples, potatoes, rice, soybeans, and lettuces. Next come citrus fruits (Fig. 2A).

When analysis is performed by PPP categories, fungicides are those most used globally (Fig. 2B). They are also the most represented, with 1438 formulas authorized in France. A total of 9 fungicides are among the 17 highest contributors to potential ED exposure i.e. exposure at and over 4 µg/d. There are also 6 insecticides and only 2 herbicides. The most prevalent is glyphosate, with a daily intake evaluated at 88.82 µg. Glyphosate's ED effects are reported in several publications, some of which can be found in the "French PPP-ED" file in the data repository at https://data.mendeley.com/datasets/dk8gsx2r4j/9.

3.2. Exposure to isoflavones with hormone-like action

In the French diet, genistein and daidzein, with hormone-like activities, are essentially introduced by soybean and soy-based foodstuffs. In March 2020, 329 soy-based foodstuffs were found on French supermarket websites (see https://data.mendeley.com/datasets/dk8gsx2r4j/9, "Phyto-Product-Inquiry-ED"). Steaks and "elaborated dishes" containing meat-like products based on soy were the most popular soy-based foodstuffs, followed by desserts, creams, and drinks (Fig. 1). Based on the French consumption data, it was found that the global population had 8 g of soy/d/capita in 2020. Soy-consumers had 10 g/d/capita while the soy non-consumers had 5 g/d/capita through foodstuff with hidden soy. The mean content of genistein and daidzein in each soy-based food category is given in Table 1. The mean concentration of isoflavones per gram of soy is calculated from the measurements made for this study and gathered in the data repository at https://data.mendeley.com/datasets/dk8gsx2r4j/9 in "Genistein and Daidzein values-ED". The measurements showed that, for hidden soy, the amount of isoflavones per gram of soy is lower than in soy-based foodstuffs due to different processing. Following the genistein and daidzein measurements performed in the food investigated in this study, 3 values can be given: (1) global amount of isoflavones per gram of soy used for French foodstuffs (845 µg/g); (2) mean isoflavones content per gram of soy in identified soy-based foodstuffs (1800 ± 275 µg/g); and (3) mean isoflavones amount per gram of soy in foodstuffs with hidden soy (310 ± 65 µg/g). See Fig. 2-Isoflavone-food-content-2020 at https://data.mendeley.com/datasets/dk8gsx2r4j/9 for more details. These values were obtained by combining isoflavones measurement and claimed soy percentage in foodstuffs. Associating these data with the total soy consumption by soy consumers (60% of the French population) and the proportion of products with hidden soy (64%), an isoflavone daily exposure was calculated for each soy-based food category (Fig. 3). This estimated daily exposure of soy-consumers to genistein + daidzein was found to be 6915 ± 1220 µg/d, with 4590 ± 855 and 2320 ± 470 µg/d of genistein and daidzein expressed in equivalent-aglycone, respectively. Equivalent-aglycone is the amount of aglycone compounds contained in the total glycosidic forms measured. Glycosidic forms can be glucoside, acetyl or malonyl forms. The glycoside moieties been hydrolysed by β-glycosidase (lactase-phloridzin-hydrolase) only aglycone compounds enter the enterocytes. The daily exposure of soy non-consumers to genistein + daidzein was estimated at 1930 ± 410 µg/d, with 1370 ± 140 and 560 ± 10 µg/d of genistein and daidzein, respectively. Fig. 3 also shows the contribution of each food category to the global intake. It appears that hidden soy is a significant contributor to isoflavones exposure in the French population.

4. Discussion

This study essentially focussed on potential EDs that can be found in or on fruits and vegetables. It did not consider wrapping agents that can be found at very low levels in such products. Also, it did not mention drugs with hormonal effects used either for contraception or treatments of hormone-dependent pathologies. These substances are very powerful but are fortunately used at low doses under medical supervision and in framed contexts, although they can be found at very low doses in the environment. They might potentially be in contact with edible plants, but no data are currently available in France on this route of exposure.

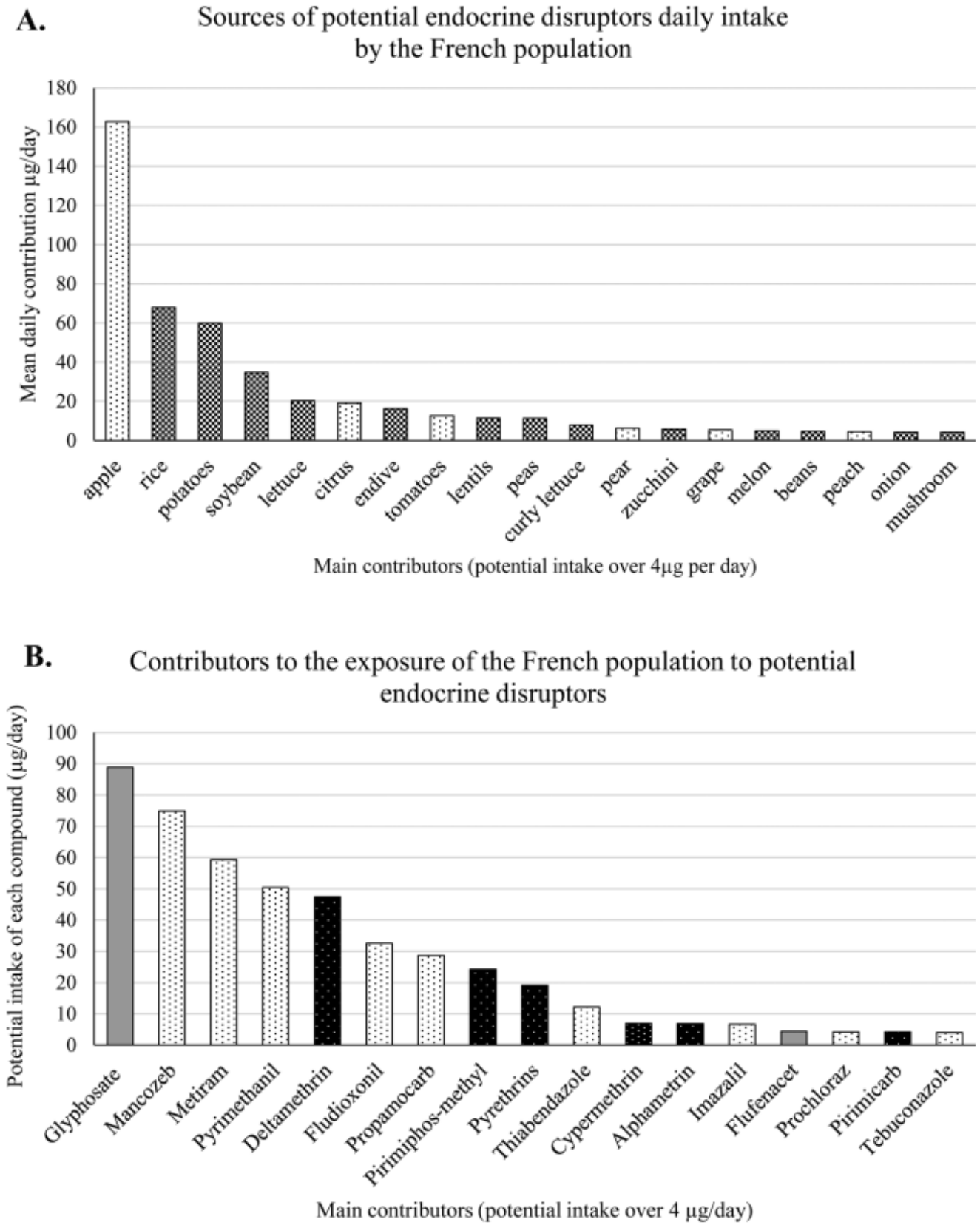


Fig. 2. Main contributors to the daily intake of plant protection products (PPPs) with endocrine disruptor (ED) properties in the French population (at and over 4 µg/d). Data collected in March 2020 are expressed in µg/d. They result from a calculation involving the daily consumption of each food contributor, the maximum residue limit (MRL) for each treating-PPP with ED property, and a "popularity factor" that considers the number of formulas containing the substance reported compared to the total formulas available in France for the same application (plant and pest). A. Plant contributors, B. Substance contributors.

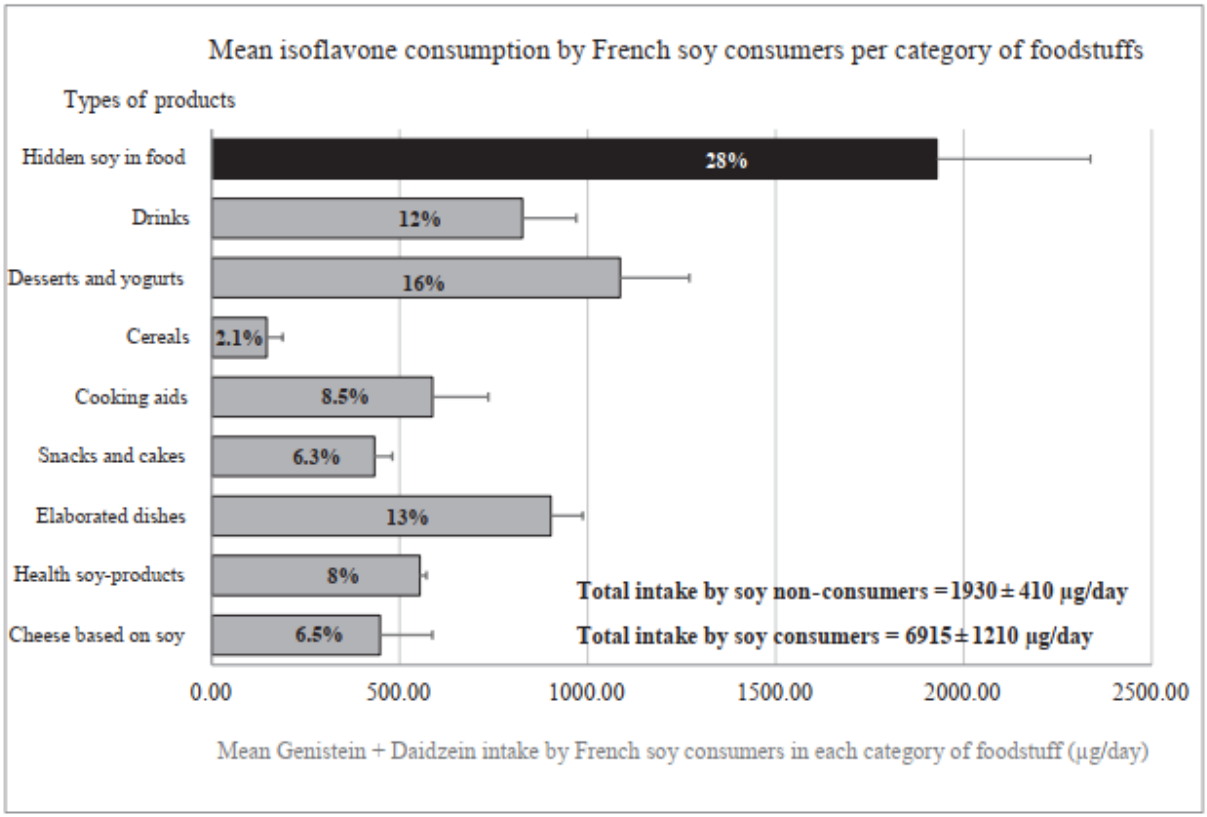


Fig. 3. Calculated exposure to isoflavones of French consumers in March 2020. Data are means ± SEM of 140 measurements. All concentration data result from assays performed by ELISAs (Table 1). Data on soy food consumption result from a specific enquiry performed using French supermarket websites. Data on hidden soy consumption result from calculations based on manufacturers data (Sojaxa, 2020; GEPV, 2019).

4.1. Exposure to ED pesticides and exposure route

4.1.1. Checking for fruit and vegetable consumption.

To check the approach, the calculated figures for fruits and vegetables consumption were compared to data published in 2017 in the INCA3 survey (ANSES, 2017). According to INCA3, vegetable consumption in France including rice and pasta is 517 g/d. Pasta consumption accounts for 22 g/d according to UNAFPA, the European Union of Pasta Manufacturers (http://www.pasta-unafpa.org/fr-statistics4.htm). ANSES differentiates children, adolescents, and adults up to 79 years old. Adding the consumption of each subpopulation according to the number of citizens in each category, a global fruit and vegetable consumption of 517 g/d was obtained. This figure should be compared to the 503 g/d obtained here, not including pasta. Therefore, the calculation in the present study seems reliable. Here, elderly individuals from 80 to 89 years old were considered as adults. The present EDs evaluation omits cereals because they are usually eaten after several transformation steps to get bread, pasta, or plant proteins. Such industrial processes can deeply modify the initial level of contamination. Therefore, MRLs would not reflect the final exposure. Finally, in the present study, only 64 fruits and vegetables were considered. Thus, ancient or exotic foods for France, such as scorsonera, mango, coconut, lychee, and sweet potatoes, for instance, were not included.

4.1.2. Reliability of the estimation of ED-PPPs exposure.

The pesticides that were selected here are only potential EDs since they do not match the criteria that were recently defined by OECD and EFSA - ECHA (Andersson et al., 2018). Our definition includes more substances than those potentially retained following the recent definition. Certain compounds exert direct effects through hormone receptors or on the bioavailability of the hormones via an interaction with liver or kidney enzymes (see "French-PPP-ED" for details at https://data.mendeley.com/datasets/dk8gsx2r4j/9). Several times, *in vivo* deleterious effects on organs involved in endocrine processes were found, while the mechanisms of these effects were not always understood (Andersson et al., 2018). Because alterations of such tissues are likely to induce deleterious endocrine effects, those substances were retained. Their dosages, when tested *in vitro* or *in vivo* in the scientific literature, can be far higher than those found here. However, it can be seen that many different substances can potentially reach consumers during a year, inducing a cocktail effect. This cocktail effect is very difficult to forecast since it depends on the potency and concentration of the individual compounds. Moreover, the substances in the mixtures can have agonistic or antagonistic effects. They also can act on different cell or tissue pathways resulting in complex phenomena. Additionally, sometimes and because of a large distribution volume, relatively high levels can be achieved *in vivo*, due to bio-accumulation phenomena, particularly in fat tissues. The calculation shows a predominance of potential anti-androgens, although this could come from the current lack of knowledge on other compounds and other effects. Also, no data are presently available on the substances' bioavailability once ingested by Humans. Therefore, saying that the French population is presently affected by an overall anti-androgenic contamination should still be confirmed.

4.1.3. Validity of the MRL approach

No assay of pesticides was performed here. Indeed, specific assays would only have been indicative of a defined situation but may not have reflected the overall one. Therefore, although the MRLs can be thought to be higher than the real concentration in food, they may be as reliable as measurements while estimating an exposure. In addition, there can be variations from one harvested plant to another, and fruits, vegetables and foodstuffs coming from countries with different pesticide

regulations were not considered. Also, taking MRLs as doses in plant-food can lead to an overestimation of the exposure, while missing a plant product can lead to its underestimation. Finally, a popularity factor has been introduced in the calculation of PPPs exposure, which reduces the final value. This popularity factor was adopted since it was not possible to get sale data on the PPPs selected in this study. This factor reflects the chance that a farmer has to choose a preparation containing the ED substance, among other preparations that can be used for the same plant and the same application. This popularity factor assumes that all preparations are equally available and known by farmers. Such a popularity factor can be adjusted in future studies if a better solution is found.

4.1.4. Discussion on the effects of complex mixtures

The cocktail exposure appears clearly here and interactions with other EDs, including wrapping agents, flame retardants, or isoflavones, should also be considered. For instance, glyphosate, at low concentrations, was shown to have additive effects with genistein *in vitro* on the proliferation of MCF-7 cells (Thongprakaisang et al., 2013). The cocktail effect of many different compounds occurring at low doses is difficult to forecast. However, some population studies recently showed correlations between pesticide exposures and health alterations (Mostafalou & Abdollahi, 2017). For instance, pesticides have long been implicated in adverse reproductive outcomes (Mnif et al., 2011). However, there are limited data on pesticide effects on the human reproduction while animal studies showed alterations of ovarian function by organochlorine pesticides. Malathion, still used in Europe, increased apoptosis in goat granulosa cells (1 and 10 nM) (Bhardwaj and Saraf, 2016). Fertility was shown to be significantly decreased in women by organochlorine pesticides via an increasing delay in conception (Chevrier et al., 2013). In some studies, an earlier menopause age could be observed (Rattan et al., 2017). In males, according to Brehm et al. (2019), several pesticides that are no longer authorized in Europe or in the US but still remain in soils, can increase the risk of testis and prostate alterations in rats. Human sperm characteristics can also be altered by pesticides (Brehm & Flaws, 2019), and this can negatively impact fertility. Sometimes, deleterious effects in rats could be recorded at the third generation due to epigenetic modification of gene transcription (Mostafalou & Abdollahi, 2017). Finally, Requena et al. (2019) found a significant association between high use of pesticides and human thyroid dysfunction. The study focussed on pesticide mixtures. Hypothyroidism was the most impacted by pesticides and women were significantly more affected.

Conversely, fruits and vegetables consumption, is still considered to prevent many chronic diseases (Wallace et al., 2019). Such an apparent discrepancy raises the question of the exposure route. It may be possible that fruits and vegetables consumption remains positive for health even when PPPs are used, especially at doses lower than MRLs. However, the usage of such products and the neighbouring exposure via lung absorption during the spreading of substances, are probably more deleterious (Mostafalou & Abdollahi, 2017). Moreover, such substances are adverse for the environment, for fish, birds, and earthworm species (Andersson et al., 2018).

4.1.5. Interest of the method for historical comparisons

When the present analysis was made, it applied an approach similar to that used 6 years earlier in France (data not shown). By comparison, several substances found to be EDs in 2013 are no longer allowed or used in France. They are acetochlor, aminotriazole, bifenthrin, chlorpropham, chlorpyrifos-methyl, cyfluthrin, dimethoate, flusilazole, iprodione, linuron, oil, propiconazole, thiacloprid, thiourea, thiram, and zinc. In addition, some new EDs have appeared on the list, either because they are newly used in France or because their ED effects were recently documented. They are: bromoxynil, buprofezin, cyromazine, etridiazole, flubendiamide, fluopyram, flutianil, folpet, geraniol, malathion, MCPB, metam, metamitron, metiram, metobromuron, myclobutanil, oxyfluorfen, profoxydim, pyridate, quizalofop p-ethyl, spiromesifen, tebufenpyrad, terbuthylazine, tetraconazole, thymol, triflumizole, zoxamide, 1-naphthylacetamide (1-NAD), 2-phenylphenol (including sodium salt orthophenylphenol), and 8-hydroxyquinoline (including oxyquinoleine). The evolution observed from 2013 to 2020 demonstrates that substances are regularly tested and evaluated by the safety agencies based on evolving knowledge and availability. This sustains the transient nature of the present study and also its interest for further comparisons. As an example, glyphosate is currently under particular scrutiny in France prior to a possible ban, and it was considered here as an ED, based on publications that show direct oestrogenic effects *in vitro* on oestrogen receptors and on MCF-7 cells (Thongprakaisang et al., 2013). It also exhibits deleterious effects *in vivo* on animal development with NOAEL (No Observable Adverse Effect Level) at 50 mg/kg bw/d on reproductive and developmental criteria (Romano et al., 2012). Although this substance use has been progressively reduced in France, with 690 pre-existing formulas that are no longer authorized, it is still potentially the highest contributor to endocrine disruption in the French population, with 32 formulas still authorized as of March 2020 for ubiquitous treatments.

4.2. Exposure to natural hormone-like substances from food

4.2.1. Reliability and limits of the estimation

Enterolignans that are oestrogenic only after gut flora transformation were not included in this study. Indeed, to evaluate oestrogenic enterolignan exposure, gut flora competence characterisation at the population level should be known. Such data are not available in France or in any other country. Also, enterolignans act more as SERMs (selected oestrogen receptor modulators) rather than EDs (Carreau et al., 2008) and, when tested in rodents, enterolignans antagonise the proliferative effects of isoflavones on breast cancer cells (Power et al., 2006). The present analysis estimates the isoflavones intake by French soy-consumers at 6915 ± 1220 μg/d, *i.e.*, 13.5 times higher than the estimated exposure to ED-PPP. In soy non-consumers, the exposure is estimated at 1930 ± 410 μg/d, which is 3.8 times higher than the ED-PPP exposure estimated here. To check the exposure calculation, another approach based on the global amount of soy consumed per year by the French population can be performed. According to the GEPV, soy proteins consumed by French citizens account for 185,000 tons a year, *i.e.*, 8 g of soy per day/capita. Here, the isoflavones measurements in French foodstuffs lead to the estimate that each gram of soy used in French food contains a mean of 845 μg of isoflavones. Considering these figures, the global exposure to isoflavones can be estimated at 6700 μg/d. This is of the same order of magnitude as the present estimation, which considers different food categories. Still, considering this global approach, the exposure of French soy non-consumers is found to be 1110 μg/d.

4.2.2. Evolution of the exposure during the last years

Note that, in 2005, AFSSA estimated the isoflavones exposure of French citizens not eating soy at 26 ± 24 μg/d. AFSSA, (2005) did not consider hidden-soy but only isoflavones from other legumes. AFSSA mentioned a great difference between the general and the vegetarian population eating soy-based foodstuff. It mentioned that while the intake was close to 0 in the population not eating soy, the intake could be of 10,000 μg /d for a daily consumer and up to 109,000 μg/d (mean 54,000 μg/d) for a consumer eating more than one soy-product daily. In 2005 the proportion of people who claimed having soy-based food at least occasionally was about 10% while in 2020 Sojaxa claims that 60% of the French population is eating soy-based foodstuffs at least occasionally. According to the Oleopro website, the human-food market increased by 20% per year between 2001 and 2010 with a French production from 5600 tons in 2001 to 65 000 tons in 2010. In addition, the mean growth rate of soy-based foodstuffs volume in France is 19% per year since 2000. Although in France, soy for human food is not only from local production, these figures sustain the progressive increase of isoflavones exposure in the French population. Still, soybean consumption

in France is not homogenous. The general population is only exposed to small amounts of isoflavones through hidden soy incorporated as soy-meal, soy protein concentrates, or soy protein isolates in transformed dishes. In parallel, some consumers can be exposed to high isoflavones amounts. Recently, Lee et al. (2019) showed that 12% of French pre-menopausal women had a daily intake over 50 mg in aglycone equivalent (see https://data.mendeley.com/datasets/dk8gsx2r4j/9 and Fig. 3 from Lee et al., 2019). The figures show that the soy-intake is mostly a recent practice in France and that the main week-intake is 1 to 3 portions. However, a significant proportion of women has more than 9 portions per week (2.2% has more than 25 portions /week) and this leads to a significant proportion of women being exposed to 50 mg isoflavone per day or more, expressed in aglycone equivalent. The highest exposure recorded in this study was 120 mg/day. Such doses are of the same order of magnitude as the highest dosages in food supplements and are active in women. In France, the dosages of isoflavones in food supplements were, until recently, far higher than those found in food (Vergne et al., 2007). However, nowadays, the food industry produces increasing amounts of soy-food based on textured soy-proteins. These ingredients are more concentrated in isoflavones than soybean seeds and, hence, soy-based food can contain high levels of genistein and daidzein (Table 1). This is a recent development due to industrialisation of soy-food processing. Indeed, Asian recipes that traditionally involved prolonged simmering in water (2 to 4 h), were able to dramatically reduce all antinutritional factors including isoflavones (Bennetau-Pelissero, 2019). Modern processing, lead the intake of isoflavone in food to be equivalent to that in dietary supplements which are not dedicated to children.

4.2.3. Effects of isoflavones in humans

Isoflavones can exhibit deleterious or beneficial effects in animal models and sometimes in Humans. However, the scientific evidences in Humans are not strong and many beneficial effects are refuted by inefficient studies (Gómez-Zorita et al., 2020). These beneficial effects are mainly documented in people over 50-years-old. Thus, many food supplements were developed based on isoflavones-rich extracts for menopausal women or to prevent prostate cancer evolution. The effect on bone is probable because most studies showed a reduction of bone resorption markers in urine with isoflavones intakes over 2 months. However, osteoporosis prevention is still questionable (Gómez-Zorita et al., 2020) because bone evolves at a slow kinetic and a drug effect can only be seen on prolonged clinical interventions. Additionally, osteoporosis prevention not only requires bone density preservation but also bone architecture preservation. Consequently, studies that focus on the positive effects of isoflavones usually mention some controversial physiological issues indicating that isoflavones should not be spread toward all human sub-populations like infants or pregnant women (Gómez-Zorita et al., 2020, Messina et al., 2021). In parallel, other academic works tented to analyse the deleterious effects of isoflavones demonstrating that many studies found no potential concern (Messina et al., 2021). However, dosages can be too low to induce an effect and sometimes the relevant biological markers are not adequately analysed. Meanwhile, due to modern soy-processing, an increasing proportion of consumers in Western or Asian countries face raising isoflavones exposures, comparable to that of dietary supplements. Some studies report heart disease prevention, diabetes improvement, cholesterol lowering, or immune function regulation (Nakai et al., 2020, Messina et al., 2021). These effects may require pharmacological experimental exposures (>100 mg isoflavones/day). Then, as with drugs, isoflavones intake over 20 to 60 mg/d (depending on the glycoside or aglycone status) should be targeted to specific populations to avoid endocrine effects in physiological contexts that are not compatible with hormone-like effects (infancy, pregnancy, malignancy…etc).

In fact, isoflavones at 45 mg/day (glycoside or aglycone) in women can exhibit potentially adverse effects at specific physiological stages, lowering FSH, and LH synthesis and release (Cassidy et al., 1994). They can also bind to oestradiol receptors, *i.e.*, the canonical receptors ERα and ER β (Kuiper et al., 1998) and GPER (Petrie et al., 2013). Isoflavones also interact with thyroid hormone synthesis (Nakamura et al., 2017) and receptor binding at doses achievable dietarily (https://data.mendeley.com/datasets/dk8gsx2r4j/9). These endocrine effects can be beneficial or deleterious, depending on the physiological status of the consumer. Hence, in some cases isoflavones can be qualified as EDs or as drugs, and should be used under medical supervision. The National Toxicology Program (NTP) in the USA demonstrated in multigenerational studies the oestrogenic and reprotoxic effects of genistein and its carcinotoxicity on the mammary and pituitary glands in Wistar rats (National Toxicology Program, 2008a; National Toxicology Program, 2008b). The doses of genistein tested in these studies are in line with human dietary intake of genistein. Therefore, it is not surprising if several studies correlated high genistein and daidzein levels in body fluids with altered characteristics of human semen (Yuan et al., 2019).

Some authors report that EDs can be more deleterious at low doses (Vandenberg et al., 2019). The explanations that are currently given, rely on the activation of different pathways by different doses of the same substance. For instance, isoflavones can activate GPER at low doses and can inhibit protein tyrosine kinase (inducing cell apoptosis) at high doses. Moreover, *in vivo* and in humans there is no evidence of any adverse effects of low doses of isoflavones. It is rather doses over 20–40 mg/d (aglycone - glycosidic forms) that can exert physiological activities, either beneficial or deleterious, depending on the status of the consumer (Cassidy et al., 1994).

4.3. Comparison of effects and exposure of the two types of substances

4.3.1. Comparison of in vitro effects

When ED effects are studied *in vitro*, it is found that genistein and daidzein are usually more potent oestrogens and thyroid disruptors than synthetic compounds that are now banned in European countries. See "Table for comparisons" in the Mendeley data repository at https://data.mendeley.com/datasets/dk8gsx2r4j/9. This table gathers data obtained in different *in vitro* systems by four different authors and genistein and oestradiol can be considered as reference compounds. In all cases except for 4-nonyl-phenol, an alkyl-phenol found in cleaning products with strong anti-thyroid activity, the effect of genistein was greater than that recorded for other synthetic compounds.

4.3.2. Comparison of NOAELs and reference doses (RfD)

Comparisons of NOAELs and LOAELs obtained *in vivo* mainly in rats are given in the file "Table for comparisons" at https://data.mendeley.com/datasets/dk8gsx2r4j/9 in the Mendeley data repository. RfD values in this table derive from the NOAELs, by a reduction factor of 100. When only a LOAEL is available, the RfD values in obtained applying a reduction factor of 300. NOAEL and LOAEL can vary greatly according to the target organ and function. Inter-sex and interspecies differences are also observed. When different safety values were available, the reproductive and/or developmental effects were considered preferentially. This is because this study focusses on endocrine disrupting effects and essentially on reproductive hormones. However, sometimes, hepatotoxicity occurred at lower dosages than reprotoxicity.

Based on their chemical composition, the major classes of pesticides include: (1) organochlorines (*e.g.*, DDT, lindane and methoxychlor with respective NOAELs equal to 0.05 mg/kg bw/d; 5 mg/kg bw/d and 5.01 mg/kg bw/d) that have been banned in Europe, (2) organophosphates (e.g., malathion and glyphosate with respective NOAELs on reproduction or development in rat equal to 2 mg/kg bw/d and 50 mg/kg bw/d), (3) carbamates (*e.g.*, carbetamide, fenoxycarb, indoxacarb, pirimicarb and propamocarb with respective NOAELs on reproduction or development in rat equal to 6–8 mg/kg bw/d; 6–10 mg/kg bw/d; 1.5 mg/kg bw/d; 20 mg/kg bw/d and 10 mg/kg bw/d), (4) pyrethroids (*e.g.*, cypermethrin, deltamethrin and permethrin with respective NOAELs on reproduction or development in rat equal to 10 mg/kg bw/d; 50 mg/kg

bw/d and 180 mg/kg bw/d), and (5) triazines (*e.g.*, cyromazine and terbuthylazine with respective NOAELs on reproduction or development in rat equal to 2 mg/kg bw/d and 5 mg/kg bw/d). When several NOAELs were given in the same species for the same substance and still on reproduction or development, the lowest was retained. The references used for NOAELs are at https://data.mendeley.com/datasets/dk8gsx2r4j/9 in "Table for comparisons". Besides, the LOAEL of genistein (35 mg/kg bw/d) recorded by the NTP in rat, may be derived in a NOAEL of the same order of magnitude that safety values obtained on the same model for other isoflavones and also for diethylstilbestrol, zearalenone and β-sistosterol. Besides, p,p' DDE, nitrofen, lindane or methoxychlor also present reprotoxicity at NOAEL close to those that can be deduced for genistein. For daidzein, the NOAEL is proposed in hen. Although it may be relevant to humans, the NOAEL obtained in birds can sometimes be derived to an RfD in Human using a safety factor higher than 100.

Hence, the RfD deduced from the NOAEL or LOAEL are also given in the file "Table for comparisons" at https://data.mendeley.com/datasets/dk8gsx2r4j/9 in the Mendeley data repository. Considering the NOAEL of the ED-PPP studied here, from 2 to 180 mg/kg bw/d the individual RfD can be either of the same order of magnitude than genistein or greater. It is noted that theses NOAEL are generally higher than those of the PPPs banned from use in Europe (DDT, Lindane, methoxychlor…). The RfD for genistein derives from the LOAEL of 35 mg/kg.bw/d in rat (National Toxicology Program, 2008a) and for daidzein from a study performed in hen. All the scientific articles taken into consideration for this *in vitro* and *in vivo* comparison are listed in the file "Table for comparisons" in the Mendeley data repository at https://data.mendeley.com/datasets/dk8gsx2r4j/9.

4.3.3. Comparisons of exposure to synthetic and natural EDs

It can be observed in Table 2 based on substances measurements that natural compounds are present in raw materials at higher concentrations than any other synthetic EDs. Therefore, consumption of plants containing high doses of natural hormone-like substances should be considered. Anyway, as already mentioned pesticides at low doses reach consumers as complex mixtures that can have either additive or antagonizing effects (Rizzati et al., 2016). The final overall effect depends on concentrations, pharmaco-kinetics and pharmaco-dynamics.

For most of the ED-PPPs listed in table 10 at https://data.mendeley.com/datasets/dk8gsx2r4j/9, the exposure data in France are not available, especially because, as mentioned earlier, the substances have been withdrawn from the market or are no longer authorized. It is likely that the exposure to persistent toxic compounds is progressively declining since their ban in France. Therefore, their safety concern is also declining. Globally, looking at the exposures estimated here for the potential ED-PPPs listed in the previous paragraph, it can be seen that the doses that potentially reach the consumers are individually all much lower than the RfD that can be derived from their NOAEL on reproduction and development. This does not imply that there is no effect especially because of the recurrence of exposure and of the complex mixtures achieved.

Besides, isoflavones can be found in great amount in several food-items that are available to current consumption. As a consequence, Lee et al., (2019) found that 12% of premenopausal women enquired in 2017 had more than 50 mg/ day of isoflavones expressed in aglycone equivalent. This is roughly 75 mg/day of isoflavones in their glycosidic forms and between 0.83 and 1 mg/kg bw/d in aglycone equivalent. Hence, cocktail-effects associating EDs with isoflavones at high and low doses should be investigated. Indeed, additive effects were recorded with glyphosate (Thongprakaisang et al., 2013) and other compounds, like vinclozolin or bisphenol A (Lee et al., 2019). Moreover, several authors, including Omoruyi and coworkers (Omoruyi et al., 2013), looked for oestrogens in the consumer's plate using different extraction procedures and different screening techniques like YES or CALUX tests. Sometimes, the extraction procedure was not the best to recover glycosylated isoflavones, however, all agreed in saying that when soy was present, genistein was by far the most potent oestrogen that could be recorded. The authors considered that such natural EDs should be taken into consideration while evaluating human EDs exposure. For natural compounds and according to our last study (Lee et al., 2019) which showed an increase and a large variation in the exposure to isoflavones in the French population, only genistein and daidzein can be absorbed at doses higher than the RfD.

Finally, contrary to the majority of ED-PPPs, isoflavone bioavailability is well known. Genistein and daidzein can be found in the plasma of French consumers at doses up to 800 ng/mL (i.e. about 1.5 μM) (data not shown). Their T_{max} range is between 8 and 9 h. Their elimination half-life can be lengthened by recirculation phenomena, ranging from 12 to 24 h (Shinkaruk et al., 2012). Eighty percent of ingested compounds can be excreted in the urine over a 48 h-period, showing good intestinal absorption. Moreover, daidzein can be metabolized by the gut flora, into equol which is more active than its parent compound. This occurs in about 40% of the European population (Shinkaruk et al., 2012). Genistein and daidzein are rapidly eliminated from the body and only a recurrent intake over 20 to 30 mg of aglycone equivalent /day in adults is expected to exert a significant endocrine effect (Shinkaruk et al., 2012). This may be observed in developed Asian countries, where soybean is now processed using industrial techniques that no longer remove isoflavones (Nakamoto et al., 2018). Even if industrial soy-based food prevents the main food-induced chronic diseases (diabetes, cardiovascular diseases, metabolic syndrome), a deleterious effect of isoflavones on human reproduction cannot be excluded, looking at fertility rates in industrial Asian countries (Index Mundi website, 2019). However, to demonstrate such an effect, control Asian populations not exposed to isoflavones should be compared to the regular population. Moreover, because of epigenetic interactions, an isoflavone effect, if real, would not disappear after a short period of abstention from genistein and daidzein.

4.4. Conclusion and perspectives

Potential EDs exposure in France can be assessed from global fruits and vegetables intake based on the primary consumed items. MRLs of pesticides can be used for each substance identified as ED, and their use can be considered using a "popularity factor". This leads to the assessment of global exposure to a complex mixture of substances, exerting pro- and antagonising effects, emphasising the need for more studies of cocktail effects. Anti-androgens deserve particular attention, since they are most represented in this study. Here, only PPP substances allowed in

Table 2
Comparison of active substances in and on fruits and vegetables.

Doses (μg/100 g)	Substances	Cereals	Traditional French fruits & vegetables	Soybean	Mung bean
Natural disruptors	Isoflavones	0–0.6[1]	0–6,900[4]	26,400–237,000[4]	tr – 2,650[4]
Synthetic disruptors	Pesticides	0.001–0.34[2]	3 – 78[5]	2.8 – 280[6]	1.8–16.5[6]
	Wrapping agent	0.04–0.173[3]	0.41–8.37[3]	0.07 [3]	No data

1. Thompson et al., 2006
2. Walorczyk and Drożdżyński, 2012
3. Cao et al., 2011
4. Eldridge, 1982
5. Schilirò et al., 2011
6. Li et al., 2012

France are considered, but other chemicals can be found from imports. And this may affect items consumed in France, such as tomatoes or citrus fruits, that largely come from Spain and North Africa. In this way, the present study may not be exhaustive. Furthermore, exposure to natural EDs, especially those from legumes, namely isoflavones, is far from negligible. Indeed, because of the recent significant reduction of meat consumption in the French population, the exposure to pulse-EDs has risen dramatically, reaching probably 4 to 12 times that of ED- PPPs. In addition, interactions at low doses exist between those natural EDs and other substances and effects can be additive, for example with bisphenol A, vinclozolin, or glyphosate on specific physiological targets. Such exposures should be considered when ED effects are assessed for consumers' health management. Exposure to isoflavones could be readily reduced if traditional Asian cooking practices would be adapted to industrial soy-based food processing. Soy isoflavones having oestrogenic, anti-androgenic, and anti-thyroid effects, this could largely reduce consumer's EDs exposure.

CRediT authorship contribution statement

Alexandre Lee: was involved in the methodology choices, data curation and data formal analysis. **Souad Bensaada:** was involved in investigation, formal analysis, data curation, and writing. **Valérie Lamothe:** was involved in the investigation, formal analysis and validation of the manuscript. **Melissa Lacoste:** was involved in the investigation and data curation. **Catherine Bennetau-Pelissero:** is responsible for the conceptualization of the study, the project administration, the funding acquisition, a part of the formal analysis and data curation, the supervision and validation of the study as well as the writing and reviewing. None of the authors have any conflict of interest to declare.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Inserm and Bordeaux Sciences Agro. Assays were possible thanks to the Regional Council of Nouvelle Aquitaine. A. L. helped with the exposure calculation and manuscript writing. S.B. helped with foodstuff assays and manuscript writing. V. L. was involved in the isoflavone assays and checked the manuscript. M. L. was involved in the PPPs analysis identifying endocrine disruptors in the scientific literature. C. B-P. initiated, coordinated, and supervised the work. She was involved in the assays, PPP analyses, isoflavone exposure calculation, and manuscript writing. None of the authors have any conflict of interest to declare. The work was funded by the Aquitain Region. The grant number is : 2020.809.CP.

References

Andersson, N., Arena, M., Auteri, D., Barmaz, S., Grignard, E., Kienzler, A., et al. (2018). Guidance for the identification of endocrine disruptors in the context of Regulations (EU) No 528/2012 and (EC) No 1107/2009. *EFSA Journal, 16*(6), 5311. https://doi.org/10.2903/j.efsa.2018.5311

Anses, 2017. Étude individuelle nationale des consommations alimentaires 3 (INCA 3). June 2017. Scientific Edition. 566 pages. ISBN - 979-10-286-0170-6.

Bennetau-Pelissero, C. (2019). Plant Proteins from Legumes. In J. M. Mérillon, & K. Ramawat (Eds.), *Bioactive Molecules in Food. Reference Series in Phytochemistry* (pp. 1–43). Cham: Springer. https://doi.org/10.1007/978-3-319-54528-8_3-1.

Bhardwaj, J. K., & Saraf, P. (2016). Transmission electron microscopic analysis of malathion-induced cytotoxicity in granulosa cells of caprine antral follicles. *Ultrastructural Pathology, 40*(1), 43–50. https://doi.org/10.3109/01913123.2015.1088908

Brehm, E., & Flaws, J. A. (2019). Transgenerational Effects of Endocrine-Disrupting Chemicals on Male and Female Reproduction. *Endocrinology, 160*(6), 1421–1435. https://doi.org/10.1210/en.2019-00034

Cao, J., Guo, L.-H., Wan, B., & Wei, Y. (2011). In vitro fluorescence displacement investigation of thyroxine transport disruption by bisphenol A. *Journal of Environmental Sciences, 23*(2), 315–321. https://doi.org/10.1016/s1001-0742(10)60408-1

Carreau, C., Flouriot, G., Bennetau-Pelissero, C., & Potier, M. (2008). Enterodiol and enterolactone, two major diet-derived polyphenol metabolites have different impact on ERalpha transcriptional activation in human breast cancer cells. *Journal of Steroid Biochemistry and Molecular Biology, 110*(1–2), 176–185. https://doi.org/10.1016/j.jsbmb.2008.03.032

Cassidy, A., Bingham, S., & Setchell, K. D. (1994). Biological effects of a diet of soy protein rich in isoflavones on the menstrual cycle of premenopausal women. *American Journal of Clinical Nutrition, 60*(3), 333–340. https://doi.org/10.1093/ajcn/60.3.333

Chevrier, C., Warembourg, C., Gaudreau, E., Monfort, C., Le Blanc, A., Guldner, L., et al. (2013). Organochlorine pesticides, polychlorinated biphenyls, seafood consumption, and time-to-pregnancy. *Epidemiology, 24*(2), 251–260. https://doi.org/10.1097/EDE.0b013e31827f53ec

Eldridge, A. C. (1982). High-performance liquid chromatography separation of soybean isoflavones and their glucosides. *J Chromatography A, 234*(2), 494–496. https://doi.org/10.1016/S0021-9673(00)81894-4

Gómez-Zorita, S., González-Arceo, M., Fernández-Quintela, A., Eseberri, I., Trepiana, J., & Portillo, M. P. (2020). Scientific Evidence Supporting the Beneficial Effects of Isoflavones on Human Health. *Nutrients, 12*(12), 3853. https://doi.org/10.3390/nu12123853

Janer, G., Leblanc, G. A., & Porte, C. (2005). Androgen metabolism in invertebrates and its modulation by xenoandrogens: A comparative study. *Annals of the New York Academy of Sciences, 1040*, 354–356. https://doi.org/10.1196/annals.1327.060

Kuiper, G. G., Lemmen, J. G., Carlsson, B., Corton, J. C., Safe, S. H., van der Saag, P. T., et al. (1998). Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinology, 139*(10), 4252–4263. https://doi.org/10.1210/endo.139.10.6216

Lee, A., Beaubernard, L., Lamothe, V., & Bennetau-Pelissero, C. (2019). New Evaluation of Isoflavone Exposure in the French Population. *Nutrients., 11(10). pii*, E2308. https://doi.org/10.3390/nu11102308

Li, X., Pham, H. T., Janesick, A. S., & Blumberg, B. (2012). Triflumizole is an obesogen in mice that acts through peroxisome proliferator activated receptor gamma (PPARγ). *Environmental Health Perspectives, 120*(12), 1720–1726. https://doi.org/10.1289/ehp.1205383

Messina, M., Blanco Mejia, S., Cassidy, A., Duncan, A., Kurzer, M., Nagato, C., et al. 2021. Neither soyfoods nor isoflavones warrant classification as endocrine disruptors: a technical review of the observational and clinical data. Crit Rev Food Sci Nutr 1-57. 10.1080/10408398.2021.1895054.

Mnif, W., Hassine, A.I.H., Bouaziz, A., Bartegi, A., Thomas, O., & Roig, B. 2011. Effect of endocrine disruptor pesticides: a review. Int J Environ Res Public Health 8(6):2265-2303. https://doi.org.10.3390/ijerph8062265.

Mostafalou, S., & Abdollahi, M. (2017). Pesticides: An update of human exposure and toxicity. *Archives of Toxicology, 91*(2), 549–599. https://doi.org/10.1007/s00204-016-1849-x

Nakai, S., Fujita, M., & Kamei, Y. (2020). Health Promotion Effects of Soy Isoflavones. *J Nutr Sci Vitaminol (Tokyo), 66*(6), 502–507. https://doi.org/10.3177/jnsv.66.502

Nakamoto, M., Shuto, E., Nakamoto, A., Hata, A., Aki, N., Shikama, Y., et al. (2018). Soy product and isoflavone intake associations with allergic diseases in Japanese workers: Rhinitis, dermatitis and asthma. *Asia Pac J Clin Nutr, 27*(6), 1277–1285. https://doi.org/10.6133/apjcn.201811_27(6).0015

Nakamura, Y., Ohsawa, I., Goto, Y., Tsuji, M., Oguchi, T., Sato, N., et al. (2017). Soy isoflavones inducing overt hypothyroidism in a patient with chronic lymphocytic thyroiditis: A case report. *Journal of Medical Case Reports, 11*(1). https://doi.org/10.1186/s13256-017-1418-9

National Toxicology Program. (2008a). Toxicology and carcinogenesis studies of genistein (Cas No. 446–72-0) in Sprague-Dawley rats (feed study). *Natl Toxicol Program Tech Rep Ser, 545*, 1–240. PMID: 18685716.

National Toxicology Program. (2008b). Multigenerational reproductive study of genistein (Cas No. 446–72-0) in Sprague-Dawley rats (feed study). *Natl Toxicol Program Tech Rep Ser, 539*, 1–266. PMID: 18685713.

Omoruyi, I. M., Kabiersch, G., & Pohjanvirta, R. (2013). Commercial processed food may have endocrine-disrupting potential: Soy-based ingredients making the difference. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess, 30*(10), 1722–1727. https://doi.org/10.1080/19440049.2013.817025

Petrie, W. K., Dennis, M. K., Hu, C., Dai, D., Arterburn, J. B., Smith, H. O., et al. (2013). G protein-coupled estrogen receptor-selective ligands modulate endometrial tumor growth. *Obstet Gynecol Int., 2013*, 1–17. https://doi.org/10.1155/2013/472720

Power, K. A., Saarinen, N. M., Chen, J. M., & Thompson, L. U. (2006). Mammalian lignans enterolactone and enterodiol, alone and in combination with the isoflavone genistein, do not promote the growth of MCF-7 xenografts in ovariectomized athymic nude mice. *International Journal of Cancer, 118*(5), 1316–1320. https://doi.org/10.1002/ijc.21464

Rattan, S., Zhou, C., Chiang, C., Mahalingam, S., Brehm, E., & Flaws, J. A. (2017). Exposure to endocrine disruptors during adulthood: Consequences for female fertility. *Journal of Endocrinology, 233*(3), R109–R129. https://doi.org/10.1530/JOE-17-0023

Requena, M., López-Villén, A., Hernández, A. F., Parrón, T., Navarro, A., & Alarcón, R. (2019). Environmental exposure to pesticides and risk of thyroid diseases. *Toxicology Letters, 315*, 55–63. https://doi.org/10.1016/j.toxlet.2019.08.017

Rizzati, V., Briand, O., Guillou, H., & Gamet-Payrastre, L. (2016). Effects of pesticide mixtures in human and animal models: An update of the recent literature. *Chemico-Biological Interactions, 254*, 231–246. https://doi.org/10.1016/j.cbi.2016.06.003

Romano, M. A., Romano, R. M., Santos, L. D., Wisniewski, P., Campos, D. A., de Souza, P. B., et al. (2012). Glyphosate impairs male offspring reproductive development by disrupting gonadotropin expression. *Archives of Toxicology, 86*(4), 663–673. https://doi.org/10.1007/s00204-011-0788-9

Schilirò, T., Porfido, A., Longo, A., Coluccia, S., & Gilli, G. (2011). The E-screen test and the MELN gene-reporter assay used for determination of estrogenic activity in fruits and vegetables in relation to pesticide residues. *Food and Chemical Toxicology, 62*, 82–90. https://doi.org/10.1016/j.fct.2013.07.067

Shinkaruk, S., Durand, M., Lamothe, V., Carpaye, A., Martinet, A., Chantre, P., et al. (2012). Bioavailability of glycitein relatively to other soy isoflavones in healthy young Caucasian men. *Food Chemistry, 135*(3), 1104–1111. https://doi.org/10.1016/j.foodchem.2012.03.135

Thompson, L.U., Boucher, B.A., Liu, Z., Cotterchio, M. & Kreiger, N. 2006. Phytoestrogen content of foods consumed in Canada, including isoflavones, lignans, and coumestan. Nutr Cancer 54(2):184-201. https://doi.org:10.1207/s15327914nc5402_5.

Thongprakaisang, S., Thiantanawat, A., Rangkadilok, N., Suriyo, T., & Satayavivad, J. (2013). Glyphosate induces human breast cancer cells growth via estrogen receptors. *Food and Chemical Toxicology, 59*, 129–136. https://doi.org/10.1016/j.fct.2013.05.057

Vandenberg, L. N., Hunt, P. A., & Gore, A. C. (2019). Endocrine disruptors and the future of toxicology testing - lessons from CLARITY-BPA. *Nature Reviews Endrocrinology, 15*(6), 366–374. https://doi.org/10.1038/s41574-019-0173-y

Vergne, S., Titier, K., Bernard, V., Asselineau, J., Durand, M., Lamothe, V., et al. (2007). Bioavailability and urinary excretion of isoflavones in humans: Effects of soy-based supplements formulation and equol production. *Journal of Pharmaceutical and Biomedical Analysis, 43*(4), 1488–1494. https://doi.org/10.1016/j.jpba.2006.10.006

Wallace, T. C., Bailey, R. L., Blumberg, J. B., Burton-Freeman, B., Chen, C. O., Crowe-White, K. M., et al. (2019). Fruits, vegetables, and health: A comprehensive narrative, umbrella review of the science and recommendations for enhanced public policy to improve intake. *Critical Reviews in Food Science and Nutrition, 3*, 1–38. https://doi.org/10.1080/10408398.2019.1632258

Walorczyk, S., & Drożdżyński, D. (2012). Improvement and extension to new analytes of a multi-residue method for the determination of pesticides in cereals and dry animal feed using gas chromatography-tandem quadrupole mass spectrometry revisited. *Journal of Chromatography A, 1251*, 219–231. https://doi.org/10.1016/j.chroma.2012.06.055

Yuan, G., Liu, Y.u., Liu, G., Wei, L., Wen, Y., Huang, S., et al. (2019). Associations between semen phytoestrogens concentrations and semen quality in Chinese men. *Environment International, 129*, 136–144. https://doi.org/10.1016/j.envint.2019.04.076

IV.1.2. Estimation de l'exposition des femmes de l'étude ISOLED aux phyto-estrogènes

Pour rappel, les IFs constituant la catégorie de xéno-estrogènes à laquelle les consommateur·ices sont le plus exposés en France, l'étude ISOLED visait à explorer ce lien potentiel en évaluant l'exposition aux phytoE des femmes atteintes de LED et en comparant leurs niveaux d'exposition à ceux de femmes saines et de femmes atteintes d'autres MAIs

IV.1.2.1. Estimation de l'exposition habituelle : validation d'un questionnaire d'habitude alimentaire par l'analyse de biomarqueurs capillaires. Article correspondant à la référence [214]

Le développement et la validation de la méthode inédite de dosage capillaire des IFs (conçue pour déterminer l'exposition habituelle des participant·es à l'étude ISOLED afin de pallier au temps de résidence limité des IFs dans les fluides biologiques) ont abouti à des données biologiques, lesquelles complétées par une enquête sur les habitudes alimentaire, ont fait l'objet de la publication [214] présentée ici.

Les dosages réalisés ont permis d'établir qu'environ 80% de la population analysée était exposée aux IFs et qu'environ 60% des personnes exposées étaient productrices d'équol. Ce résultat est conforme à une étude antérieure menée par notre équipe en 2006 [31].

nutrients

MDPI

Article

Development of an Assay for Soy Isoflavones in Women's Hair

Souad Bensaada [1,2], Isabelle Raymond [1,3], Malena Breton [1], Isabelle Pellegrin [1,4], Jean-François Viallard [1,3] and Catherine Bennetau-Pelissero [1,2,5,*]

[1] Carreire Campus, Sciences and Technology Department, Pharmacy Faculty, University of Bordeaux, 33076 Bordeaux, France
[2] ARNA, U1212 Inserm, 5320 CNRS, Pharmacy Faculty, 33076 Bordeaux, France
[3] CHU Bordeaux, USN B0-Hôpital Haut Lévêque, 33604 Pessac, France
[4] CHU Bordeaux, Laboratory of Immunology and Immunogenetics, Resources Biological Center (CRB), 33000 Bordeaux, France
[5] Bordeaux Sciences Agro, 33175 Gradignan, France
* Correspondence: catherine.bennetau@u-bordeaux.fr; Tel.: +33-(0)-557-571-271

Abstract: Soy isoflavones, at adequate dosages, have estrogenic and anti-thyroidal effects in animals and humans, which can either be beneficial or adverse, depending on the consumer's physiological status. Hence, this study presents an assay of soy isoflavones in hair, aiming to give new information about a person's exposure to isoflavones, when health issues related to estrogenic or thyroidal effects are observed. Aqueous or organic extraction procedures following acidic, basic, or enzymatic digestions were tested on 60 hair samples (from volunteers) from a hairdresser, and a clinical trial 2017T2-29. The acidic digestion method was the most efficient regarding isoflavones. A specific inquiry was developed to assess the dietary habits of French consumers based on the analysis of 12,707 food labels from France. It was used to check for the reliability of the new assay method. A score for the consumer exposures to isoflavones was built considering, among other parameters, soy-based diets and foodstuff containing soy as an ingredient, i.e., "hidden-soy". The correlation between this score and isoflavone measurements in hair reached 0.947; $p < 0.001$. Therefore, providing that relevant data are considered to assess isoflavone exposure, hair that smoothens daily isoflavone intake variations, is a relevant tissue to assess human isoflavone exposure for subsequent health analyses.

Keywords: soy isoflavones; consumer exposure; hair samples; dietary inquiry; digestion method

check for updates

Citation: Bensaada, S.; Raymond, I.; Breton, M.; Pellegrin, I.; Viallard, J.-F.; Bennetau-Pelissero, C. Development of an Assay for Soy Isoflavones in Women's Hair. *Nutrients* **2022**, *14*, 3619. https://doi.org/10.3390/nu14173619

Academic Editor: Maria Dolores del Castillo

Received: 30 May 2022
Accepted: 29 August 2022
Published: 1 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Soybean has been used as a dietary source in Asia for likely more than 3000 years [1]. Basically, soybean, as a pulse, is rich in protein (over 40% of crude composition), polyunsaturated fatty acids (ω3 and ω6), and fiber. In addition, soy is one of the pulses best suited for animal feeding due to its good amino acid profile. This exceptional composition gives soy one of the best nutritional values among all legumes. However, besides its fabulous composition, raw soy, as other legumes, contains many anti-nutritional factors that reduce its palatability and digestibility [2]. This developed defensive biochemical arsenal consists mainly of phytic acids, tannins, oligosaccharides, haemagglutinins, lipoxygenases, anti-trypsin factors (Bowman–Birk and Kunitz), allergens, and isoflavones with hormonal activities [2]. However, the latter is not the least as they were shown to reduce the herbivore's fertility in both breading and wild conditions [3,4]. Moreover, some beneficial effects have been reported on human health [5,6].

While traditional Asian cooking recipes were able to remove estrogenic isoflavones from soy, thanks to prolonged cooking and simmering in water, modern industrial preparations no longer involve these water steps [7]. Consequently, the current human exposure to estrogenic isoflavones is probably the highest it has ever reached in human history. In addition, the environment is a source of other endocrine disruptors, including wrapping agents, pesticides, plastic stabilizers, flame retardants, and other anthropogenic substances [8].

This study was performed in the frame of a clinical investigation: ISOLED 2017T2-29, which aims to analyze the potential link between isoflavones exposure and the occurrence of systemic lupus erythematosus (SLE) crises. SLE is an autoimmune disease known to exhibit a highly unbalanced sex ratio occurrence [9]. Ninety percent of the SLE subjects are women and the current hypotheses link this occurrence to genetic factors bearded by the X chromosome and to estrogen impregnation. Indeed, SLE progression is not observed in pre-adolescent girls or post-menopausal women. The role of environmental factors in SLE occurrences is nowadays under debate [10]. Among other environmental factors are estrogens and phytoestrogens from soy. Little is known about this potential association; this may be due to the lack of reliable data on exposure to estrogenic isoflavones.

To correctly determine soy isoflavone's health effects on humans, it is crucial to assess their exposure in the best possible way. Some studies have tackled this issue by attempting to correlate dietary intakes to isoflavones concentrations in human urine or serum and plasma samples as biomarkers of exposure. In this context, a comparative review [11] highlighted that in many circumstances the correlation was shown to be poor. Consequently, developing an assay on hair samples seemed of interest to improve the knowledge of human exposure to soy isoflavones. The sustaining hypothesis is that hair is a better candidate as it smooths daily isoflavones intake variations as well as the variations of biological fluid concentrations. Therefore, even though no one has attempted (until now) to assay soy isoflavones in hair or use this measurement as a biomarker of exposure, both issues are described here. Such knowledge may open up new perspectives in analyzing the health effects of isoflavones. However, hair is a very specific matrix since it presents great resistance to biochemical, chemical, and physical treatments. Therefore, the challenge is to achieve efficient hair digestion to recover the largest portion of substances of interest without altering them.

In this paper, several methods were used to digest hair samples and extract isoflavones and are presented alongside their respective isoflavones recovery rates. A dietary inquiry was set up to assess chronic isoflavones exposure and dietary hair scores were derived from it, taking into consideration some hair parameters. A Pearson correlation coefficient was obtained between these scores and isoflavones measurements in hair and compared to the data reported in the literature.

2. Materials and Methods

2.1. Materials

Hair samples were collected in two locations. The first set (of 8 samples) was anonymously collected from female volunteers aged 35 to 85 at a hairdresser (with full consent from volunteers). These women accepted answering the food habit inquiry. The second set (of 52 hair samples) was anonymously collected from premenopausal women with or without autoimmune diseases, including SLE, at the hospital, in the frame of the 2017T2-29 clinical trial (ISOLED study). In this trial, in addition to hair—blood, and urine samples were collected in healthy premenopausal women and patients with SLE or other autoimmune diseases. The recruitment is ongoing, with the objective of reaching a minimum of 72 volunteers: 33% healthy, 33% SLE patients, and 33% patients with other autoimmune diseases, with a less unbalanced sex ratio. The only other autoimmune disease currently diagnosed among the volunteers is immune thrombocytopenic purpura. The aim of the clinical study is to examine a potential difference in isoflavones exposure between SLE patients and the other volunteers, by assessing isoflavones in all biological samples. Urine and plasma samples were not used in this study as it focuses on the new IFs assay in hair samples. All volunteers agreed to answer the dietary habit inquiry designed for this study and signed a full informed consent. The volunteers were on their casual diets and no selection was made based on soy consumption habits. Hair sampling was standardized as in most scientific works. It was performed at the back of the skull, at the nape of the neck [12].

2.1.1. Direct Extraction Tests

Hair samples were crushed on a Minilys® apparatus from Bertin Technologies (France) for 4 cycles of 1 min each at full power. For this operation, samples were placed in Precellys® Lysing kits (P000917-LYSK0-A). Isoflavones were extracted directly using either methanol, ethanol, or ethyl-acetate. All solvents were from Sigma-Aldrich™.

2.1.2. Digestion Tests

Hair digestion was implemented by chemical or biochemical reactions on two types of samples. Tests were performed on 50 mg hair either pre-cut or pre-cut then crushed (Minilys®). Unless otherwise mentioned, all reagents were from Sigma-Aldrich™. The chemical digestion was carried out using either NaOH 1 N or HCl 0.1 N. The reagents were diluted in ultrapure water from an Elga Veolia® water instrument. The biochemical digestion was tested using Pronase (SIGMA, Sigma Aldrich Chimie S.a.r.l, France, P5147) or Keratinase (SIGMA, K4519) solutions. Both enzymes were dissolved in a digestion buffer, which was Tris buffer pH 8.25 0.05 M with 100 units penicillin G/mL (SIGMA, P-3032), 1 mM streptomycin, and 1 mM EDTA. Stock solutions were kept frozen (−22 °C) in black storing vials. The Keratinase stock solution was at a concentration of 180 Units/mL in water. The Pronase stock solution was at a concentration of 20 mg/mL in a phosphate buffer pH 7.4 0.01 M with $(CH_3CO_2)Ca$ 10 nM. For each digestion test, 2 mL of digestion solution was used.

2.1.3. Digestion and Extraction of the Conjugated Isoflavones

The digestion of isoflavone conjugates (glucuronides and sulfates) was performed with β-glucuronidase-aryl sulfatase from *Helix pomatia* (Roche, 10127698001) diluted (5 µL/mL) in sodium acetate buffer 0.1 M, 0.14 M EDTA, 100 UI/mL penicillin G (SIGMA, P-3032), and 0.1 mg/mL streptomycin (SIGMA, S-6501) pH 5. The extraction of aglycone compounds was then performed using acidified ethyl-acetate (500 µL HCl 38% per L).

2.1.4. ELISA Measurements

All salt reagents were from VWR-France. The protocol also involved thyroglobulin from swine (SIGMA, T1126), bovine serum albumin (EUROMEDEX, 04-100-812-C), and secondary antibody goat anti-rabbit IgG antibody-Amdex (SIGMA, RPN4301). The revelation steps required o-phenylenediamine dihydrochloride –OPD– (SIGMA, P1526). Stock solutions were prepared at 10 mg/mL in water and stored in black vials at −22 °C. Glassware in contact with unconjugated isoflavones was coated with Sigmacote® (SIGMA, SL2). The sample, washing, and assay buffers were phosphate buffer saline 0.01 M, 0.9% NaCl, 0.2% Tween, 1% DMSO, and pH 7.3. To obtain the saturation and antibody buffers, 1.6 mg/L of bovine serum albumin was added to the latter. The revelation buffer was citrate-phosphate buffer 0.15 M, pH 5 with 0.05 mg OPD/mL. The stop solution was H_2SO_4 4 M. The primary antibodies obtained on various functionalized isoflavones were selected based on their specificity and the sensitivity obtained by applying the ELISA procedure [13–15]. For each hapten initially synthesized, 2 rabbits were injected and the best polyclonal antibody was retained for further use in the assay. These antibodies were harvested if required to avoid undesired cross-reactions with the unwanted antigen.

2.1.5. Control Solutions for Digestion Tests

Several pure compounds were used as control solutions to assess the digestion and extraction recovery. The first was Genistin from EXTRASYNTHESE™ (Ref. 1325 S) and was dissolved at 1 mg/mL in DMSO as the stock solution. Then, genistein-O-7-sulfate, Ref. G350045, and genistein-7-b-D-O-glucuronide, Ref. G350015, were from TRC Canada. Their stock solutions were dissolved at a concentration of 2 mM in DMSO and were stored at −22 °C.

2.2. Methods

2.2.1. Dietary Inquiry

The dietary inquiry used in this study is given in the Supplementary Material S1. It was drawn out to focus on potential isoflavones exposure. It was adapted to the French food market offer and the French diet habits. From a previous study, in which more than 12,700 foodstuff labels were analyzed [16], it was possible to determine in which foodstuff and transformed dishes soy could be incorporated. The occurrence of soy as an ingredient, i.e., "hidden-soy", was known to be more frequent in collective restaurants than in foodstuff bought in supermarkets [16]. Therefore, the inquiry addressed the usual location of food intake. It also recorded the following items:

- The dietary specificity, such as vegetarian, Asian, or gluten-free diets;
- The eater's status, i.e., as big, medium, or small;
- The snacking practice;
- The main cooking practices;
- The intake frequency of soy-based foodstuffs but also of food that contain "hidden-soy".

An estimation of the food portion size was also collected, although a retrospective estimation could never be of high reliability. However, such data could validate the eater status given elsewhere. As the inquiry was used to validate the hair measurements, some parameters were also documented in line with the specificity of these samples, such as the frequency of haircuts and shampooing. Finally, the duration of a given described diet was related to the hair length to determine whether the hair sample has always been exposed to the same diet (see Supplementary Material S1 for details).

2.2.2. Isoflavones Food and Hair Scores

A food score was attributed to foodstuff considering the average isoflavones content [17] and intake frequency by the volunteers (see Supplementary Material S2). This was based on the known isoflavones content of 240 French foodstuffs. The average content combined both mean isoflavones level measured in a food category and the occurrence of food containing soy in this given category. In Lee et al. [17], the concentration values of genistein (GEN) and daidzein (DAI) expressed in mg/100 g of food or mg/portions were obtained using ELISA and, therefore, were expressed in aglycone equivalents. The calculation of both the Isoflavones Food Score (IFS) and Isoflavones Hair Score (IHS) can be found in Supplementary Material S2. Basically, the IFS was composed of the sum of the food scores multiplied by a factor considering, on the one hand, the intake quantity, and on the other hand, some hair parameters, such as cutting and treatments, e.g., shampoo, coloring, perms, and their frequencies, which are shown to marginally influence the xenobiotics measurements. Meanwhile, the IHS was adjusted from the IFS using a potential dilution factor considering the hair length and diet duration. See Supplementary Material S2 for details on the calculation.

2.2.3. Sample Preparation

Hair samples collected either at the hairdresser or at the hospital and duly identified, were stored in plastic vials at −22 °C until processing, at the Bordeaux Hospital biobank CRB (Bordeaux Biothèques Santé Centre de Ressources Biologiques, Bordeaux University Hospital).

Hair samples were first cut into small pieces of less than 1 mm long. Preliminary assay tests were performed directly on 50 mg of such samples (soda, acid, pronase, and keratinase digestions). However, this cutting was found insufficient and was then associated with Minilys® crushing for which 50 mg of hair was distributed into 3 different Precellis® tubes containing 6 metal beads. Crushing was performed on a 4-cycle basis, each cycle lasting 1 min. The crushed samples were transferred into 40 mL screwed top glass vials, coated with silicone, in which the digestion was performed at the relevant temperature and duration in 2 mL of the corresponding solution (see Table 1 for procedure summary).

The samples were then brought to 100 °C for 10 min and cooled to room temperature to deactivate the proteases if the latter were used for the digestion.

Table 1. Test of digestions, procedure details.

Digestion Reagent	Solution	Incubation Time	Incubation Temperature
Soda digestion	2 mL NaOH 1 N	1 h	100 °C
Acid digestion	2 mL HCl 0.1 N	8 h	40 °C
Pronase digestion	2 mL 1 mg/mL Tris buffer pH 8.25	24 h	40 °C
Keratinase digestion	2 mL 1.8 UI/mL Tris buffer pH 8.25	2 h	40 °C

After pH neutralization, and when relevant (see Table 2), 2 mL of 0.1 M of acetate buffer was added to the vials with 10 µL of b-glucuronidase aryl-sulfatase solution. The samples were then incubated overnight at 37 °C under 200 rpm shaking to release the aglycone compounds [18]. The latter were extracted directly in the glass vials using 3 mL of acidified ethyl-acetate. Namely, they were extracted 3 times and evaporated to dryness using a Speed-Vac (Thermo-Electron^{TM} Corporation, Fisher Scientific, Illkirch, France). At each step, the vials were vortexed for 30 seconds, centrifuged at 500 g (Jouan^{TM} CR3, Fisher Scientific, Illkirch, France) for 10 min at 4 °C, and stored for 1.5 h at −22 °C to allow phase separation. The aqueous phase froze while the ethyl-acetate phase which contained the isoflavones in aglycone forms remained liquid. The three ethyl-acetate phases of the same sample were collected and evaporated to dryness in the same hemolyzed tube precoated with Sigmacoat®. The conjugate hydrolysis—either by b-glucuronidase aryl-sulfatase or acid digestion—was monitored using genistein-7-sulfate as a control reagent in an external standard run in parallel to the samples. This hydrolysis performance was always between 97 and 101%. After final evaporation, samples were diluted in 1 mL of assay buffer and when required sonicated using an ultrasound probe (LCD Series Ultrasonic Processor). Samples were stored at −22 °C until assay processing.

2.2.4. ELISA Measurements in Hair

GEN, DAI, and equol were assayed in the hair samples using specific ELISAs as described previously [15]. Haptens functionalized on different carbon atoms were previously synthesized [13,14]. Specific polyclonal antibodies were obtained from rabbits, for each isoflavone bound to bovine serum albumin and the best antibodies were retained for assay development. Standard curve preparations and sample dilutions were performed in silicone-coated glass vials. The ELISAs followed a competitive procedure with an immobilized competitor which was the homologous hapten bound to swine thyroglobulin. The sensitivities varied between 0.08 ng/well and 0.4 ng/well; intra-assay variation was always below 7% and inter-assay variation was below 17%. For the hair samples, final dilutions varied between 1:5 and 1:20. Each hair sample was assayed in triplicate using three extracts and on three different microtitration plates. All values are given in aglycone equivalents.

2.2.5. Statistical Treatments

Each concentration value was always a mean and standard deviation of at least triplicate measurements except when repeatability was assessed. In that case, the number of replicates was given. Due to the large cohort analyzed, data followed a normal distribution. Therefore, the correlation between IFS or IHS and isoflavones measurements as well as their significance were calculated following the Pearson procedure. The significance was obtained by comparing the ν value with data from Student significance tests.

Table 2. Comparison of the GEN + DAI recovery following different digestion and extraction methods.

Test Samples	Soda Digestion	Pronase Digestion	Keratinase Digestion	Methanol Extraction	Ethanol Extraction	Acid Digestion
Genistein-7-sulfate	0.00% ± 0.00% ***	50.97% ± 0.56% ***	90.39% ± 3.56% ***	54.00% ± 15.00% ***	83.00% ± 8.6%	98.32% ± 16.26%
Genistein-7-β-D-glucuronide	0.00% ± 0.00% ***	61.59% ± 7.85% ***	47.26% ± 6.20% ***	26.00% ± 3.12% ***	8.02% ± 7.56% ***	0.0% ± 0.00%
Genistein	94.00% ± 13.00% ***	96.00% ± 15.00%	97.00% ± 12.00%			56.62% ± 12.12%
Additional treatment	β-glucuronidase digestion					
Positive hair sample	15.82% ± 18.41% ***	74.06% ± 38.62%	55.04% ± 59.37% **			100% ± 10.09%
Negative hair sample	5.86% ± 4.26% ***	106.0% ± 38.00%	24.00% ± 54% ***			100% ± 42.00%
	Minilys^{TM} crushing					
Mixture of hair	0.00% ± 0.00% ***	56.42% ± 2.63% ***	64.74% ± 14.07 **	41.34% ± 7.26% ***	18.72% ± 3.08% ***	100% ± 24.86%
References	[19]	[20]	[21]	[22]	[23]	[24]

Significance is given compared to the acid treatment. * $p < 0.5\%$; ** $p < 0.1\%$; *** $p < 0.05\%$.

3. Results

3.1. Tests of the Extraction Methods for Isoflavones from Hair Samples

3.1.1. Tests on Pure Control Substances

The results of several extraction and digestion tests performed on pooled hair samples are reported in Table 2. The higher recovery rate for genistein-7-sulfate was recorded with the acid digestion using HCl 0.1 M for 8 h at 40 °C. However, this digestion was not able to recover the genistein-7-β-D-glucuronide metabolite. This could indicate that HCl most probably hydrolyzed the isoflavone moiety and destroyed it. Moreover, it can be observed that soft digestion processes using pronase or keratinase were able to preserve the conjugates, i.e., both genistein-7-sulfate and genistein-7-β-D-glucuronide, although the best overall recovery was obtained using keratinase. Direct methanol and ethanol extractions were not very efficient compared to acid treatment. Finally, the use of genistein as recovery control led to singular results. Indeed, NaOH treatment, leading to hair sample dissolution, was more efficient on hair digestion than any other treatment tested, and this includes the enzymatic and acid treatments. However, it did not significantly retrieve isoflavones sulfo- and glucurono-conjugates from pure solutions or hair samples (see Table 2).

3.1.2. Test on Hair Samples

Interestingly, although the acid treatment was not able to retrieve the glucuronide conjugates, it was the treatment that led to the highest dosages whatever the hair samples considered (see Table 2). Next to the acid treatment, NaOH treatment gave highly variable values not significantly different from 0. Pronase and keratinase digestions led to significant isoflavones assays although lower than those determined using the acid treatment, except for pronase on the low-concentrated hair sample (see Table 2). The latter was a mixture of several samples collected on soy non-eaters. The high-concentrated hair sample was a mixture of several samples collected on soy-eaters. The mixture sample was a pool of three samples: one came from a high soy-eater, the second from a medium soy-eater, and the third from a non-soy-eater. Crushing using Minilys^{TM} reduced the inter-assay variations (see Table 2). This could be seen on all treated samples. Likewise, the inter-sample variation is higher for the low-concentrated sample compared to the high-concentrated one, independently of the digestion treatment used. These low-concentrated samples were close to the quantification limit.

3.2. Characterization of the Isoflavones Assay in Hair

3.2.1. Sensitivity

The specificity of the primary antibodies allowed to assay hair samples at dilutions as low as 1:5 without a pre-washing procedure. Avoiding pre-washing steps allowed to minimize losses of desirable molecules. The sensitivity of GEN assay was based on the IC_{50}—half maximal inhibitory concentration of the standard—curve, i.e., 20 ng/mL. Then, since assay samples were 50 mg each, the sensitivity was 2 ng/mg of hair for GEN. However, the quantification and the detection limits for GEN were 0.8 ng/mg and 0.5 ng/mg of the hair samples, respectively. Moreover, the sensitivity of DAI assay, based on the IC_{50} of the standard curve was 13 ng/mL. Then, the sensitivity was 1.3 ng/mg for DAI. However, the quantification and the detection limits for DAI were 0.41 ng/mg and 0.12 ng/mg of the hair samples, respectively. Finally, the assay sensitivity for equol, based on the IC_{50} of the standard curve was 6.8 ng/mL. Then, its sensitivity was 0.7 ng/mg. However, the assay and the detection limits for equol were 0.1 ng/mg and 0.08 ng/mg of the hair samples, respectively.

3.2.2. Specificity

As for other immunological tests, the specificity relies on the antibodies used. In ELISA tests several dilutions of the same sample are currently done in order to obtain values as close as possible to the IC_{50} of the standard curve. If non-specific bindings had occurred, it would have been noticed since the different dilutions would give rise to highly different

concentrations. Here, most of the samples had to be diluted at 1:5 and 1:10, increasing, by the way, the potential occurrence of cross-reactions with undesired compounds. However, there were good similarities between the concentrations results obtained at those different dilutions. This indicates that undesired cross-reactions were very limited.

3.2.3. Repeatability and Reproducibility

Tests were performed repeatedly on different samples assayed on different microtitration plates. Therefore, the results given in Table 3 are inter-assay variation coefficients. As mentioned in this table, data were from the same extraction for samples 26 and 28 while they were from 2 different extractions for the soy-eaters hair sample. The coefficient obtained on the same extraction, expressed the repeatability, while the data obtained from two different extractions gave the reproducibility.

Table 3. Repeatability and reliability tests of the GEN and DAI ELISAs.

		Inter-Assay Variation Coefficients				
		Same Extraction		**Different Extractions**		
				1st Extraction	**2nd Extraction**	**Overall**
GEN	Sample	Sample 28 (n = 6)	Sample 29 (n = 8)	Soy-eater-1 (n = 6)	Soy-eater-2 (n = 6)	Soy-eaters (n = 12)
	Mean ± SD	2.25 ± 0.46	1.18 ± 0.17	3.47 ± 0.35	4.88 ± 0.92	4.18 ± 0.76
	Variation coefficient	20.44%	14.59%	10.12%	18.85%	18.25%
DAI	Sample	Sample 28 (n = 6)	Sample 29 (n = 7)	Soy-eater-1 (n = 6)	Soy-eater-2 (n = 6)	Soy-eaters (n = 12)
	Mean ± SD	1.67 ± 0.28	2.27 ± 0.35	7.86 ± 0.38	6.30 ± 1.53	7.08 ± 1.29
	Variation coefficient	16.63%	15.52%	4.89%	24.34%	18.18%

GEN: genistein in aglucone form; DAI: daidzein in aglucone form.

In all cases, the inter-assay variation coefficients were always below 25% and seemed acceptable. However, it should be mentioned that nearly 35% of the 60 samples assayed were close to the quantification limit and 18% were at the detection limit.

3.3. Correlation with the Dietary Data

In order to validate this new extraction and assay, it was decided to check its reliability based on specific dietary habit inquiries. The results are shown in Figure 1.

As can be seen in Figure 1a. A large amplitude was observed in isoflavones measurements with the minimum level of GEN + DAI being 0.11 ng/mg and the maximum being 11.18 ng/mg. This reflected the IHS which varied from 1.39–23.32. Figure 1b gives the correlation between the IFS and the IHS. The Pearson regression coefficient r between IHS and isoflavones levels in hair was 0.947 based on 60 sets of data ($p < 0.001$). Its significance was good with less than 1 chance over 1000 to have no correlation. The r coefficient was also 0.914 for GEN alone with $p < 0.001$ and 0.91 for DAI alone with $p < 0.001$ for Pearson correlations (see supplementary Material S4). Moreover, the Pearson correlation between IFS and IHS was also significant and high when based on 60 pairs of values with an r value of 0.898 ($p < 0.001$).

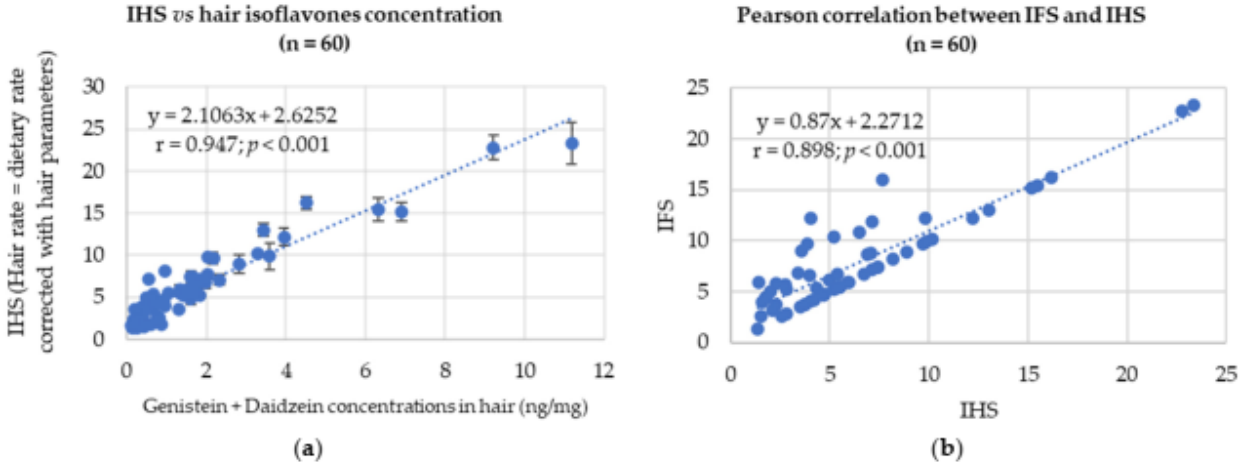


Figure 1. Pearson correlations between HIS and either total hair isoflavones or IFS, calculated for the whole 60 volunteers ($p < 0.001$): (**a**) Pearson correlation between IHS and hair isoflavones (GEN+DAI) concentration; (**b**) Pearson correlation between IFS and HIS.

3.4. Detection of Equol

Equol was assayed for the first time in 55 of the hair samples and detected in 34 of them. This represented 61.82% of the population. However, levels were consistently low and for 67% of the equol positive samples, the concentration value was close to the detection limit.

4. Discussion

Isoflavone contents were measured in human hair samples and applied to the assessment of isoflavones exposure for the first time. So far, urine, serum, and plasma samples have been the only biological materials used to detect isoflavones as biomarkers of exposure [11].

4.1. Reliability of Isoflavones Assessment in Hair Compared to Other Biological Samples

Table 4 provides an overview of different studies, using either plasma or urine biomarkers, allowing a comparison with the present work. As can be seen, the correlation coefficient obtained between soy intake and measurements of GEN + DAI is one of the highest among those reported so far in the literature.

4.1.1. Comparison with Urine Samples

Urine, as it basically concentrates the blood substances, can be used to explore low isoflavones intakes. Compared to isoflavone concentrations in hair, urine-associated levels can be 100 times higher. However, the measurement can only reflect an exposure that occurred within the last 48 h. Indeed, in some of the studies reported in Table 4, as urine was collected on a 24- or 48-h period, a highly reliable exposure evaluation was obtained. However, such a sampling may not always be possible in large cohorts. Other than that, the r values, correlating urine isoflavones measurements and dietary intake recording soy, are as low as 0.27 [25]. Indeed, spot urine samples have given less accurate results compared to 24- or 48-h urine collection but can be used with a certain reliability provided that they were adjusted on urine creatinine levels. On another side, the highest r value (0.97), obtained from a Spearman calculation, in fact corresponds to a small number of subjects (14) for whom it was hence possible to estimate the isoflavones intake and biomarker intensity with high accuracy. When the number of subjects is larger than 50, the Pearson correlation coefficients obtained range between 0.27 and 0.52. This could be explained first by the possible large isoflavone–daily-intake variations and second by the fact that isoflavones concentrations in urine depend greatly on the timespan after ingestion.

As pointed out by Shinkaruk et al. [15], after an isoflavone challenge, the urine C_{max}—maximum concentration in body fluid after ingestion—was observed between 8 and 16 h, depending on the isoflavone type measured. However, a daily intake could also lead to pharmacokinetic steady-state levels [26]. However, apart from experimental situations, this latter case has been only observed with food-supplement intakes or for chronic heavy-soy-consumers. Hence, depending on the urine sampling time, the measurements could yield very different results, even for identical isoflavone exposures. Consequently, if timespan between intake and sample collection is not recorded, then the correlation based on unadjusted data could be falsely low. Finally, urine samples, such as hair samples, are usually obtained in a non-invasive way. This aspect constitutes an advantage over blood samples.

Table 4. Overview of different studies on blood or urine isoflavones as biomarkers of dietary exposures and present work.

Subjects	Nature of Samples	Biomarkers	Dietary Data (mg/day)	Correlation	References
80 British volunteers	Plasma	GEN DAI	7 days food diaries	GEN: r = 0.80; $p < 0.001$ DAI: r = 0.78; $p < 0.001$	[27]
360 women	2 × overnight urine (48 h apart)	ISO	2 days 24 h recall DAI (µg): 5.0–6.4 GEN (µg): 7.3–9.3	ISO: r = 0.52; $p = 0.001$ FFQ: r = 0.29; $p < 0.01$	[28]
77 volunteers	Plasma	GEN DAI	FFQ	GEN: r = 0,53; $p < 0.001$ DIAD: r = 0,45; $p < 0.001$	[29]
284 volunteers 333 volunteers	Urine Serum	ISO	7 days food diaries	Urine r = 0.27; $p < 0.001$ Serum r = 0.31; $p < 0.001$	[25]
203 male volunteers	Serum	ISO	FFQ	ISO: r = 0.27; $p < 0.001$	[30]
256 premenopausal women	12 h urine	ISO	FFQ Low: 0.1–2.3 High: 49.8–74.6	ISO: r = 0.51; $p < 0.001$	[31]
14 adults (14% men)	24 h urine Plasma	ISO	24 h food record: 11.0 24 h recall: 12.3	Urine: r = 0.97; $p < 0.001$ Plasma: r = 0.92; $p < 0.001$	[32]
51 Japanese women 18 Caucasian women	24-h urine	GEN DAI	48-h dietary recall	GEN: r = 0.54; $p < 0.001$ DAI: r = 0.58; $p < 0.001$	[33]
24 pubertal girls	12 h urine	ISO	3 days 24 h recall ISO: 3.0–13.3	ISO: r = 0.72; $p < 0.001$	[34]
100 healthy women	12 h urine	ISO	24 h recall	ISO: r = 0.460; $p < 0.001$	[35]
60 French women	Hair	ISO GEN DAI	French dietary habit questionnaire	ISO: r = 0.947; $p < 0.001$ GEN: r = 0.914; $p < 0.001$ DAI: r = 0.911; $p < 0.001$	Present study

ISO: isoflavones; GEN: genistein; DAI: daidzein.

4.1.2. Comparison with Plasma or Serum Samples

In the literature (see Table 4), when plasma or serum samples are measured, they usually correspond to spot sampling. Sometimes, the time of sampling is standardized, for instance between 7:00 am and 8:00 am, after overnight fasting. Some other times, it is arbitrary and even not mentioned, although, as already referred to, the pharmacokinetic of soy isoflavones orally administrated is well documented [15]. For GEN, the T_{max}—time at a maximum concentration in body fluid after ingestion—observed on 18 male volunteers, was shown to occur 7.33 h ± 2.38 h after ingestion with an elimination half-life $T_{1/2}$—time required for the elimination of half concentration of a given substance from plasma—of 22.2 h ± 13.38 h due to individual recirculation phenomena. For DAI, the T_{max} was shown to occur 8.78 h ± 2.84 h after ingestion with an elimination half-life $T_{1/2}$ of 11.2 h ± 4.51 h. Consequently, if the time of sampling is not standardized, high

variability in isoflavone measurements is likely to happen even for identical food intakes, resulting in weak correlations. It is thus not surprising that the published correlation coefficients between soy-food intake and serum measurements are as low as 0.27 [30]. Concerning the highest r values observed for plasma correlation, it is possible to point out some objective reasons for these situations. For instance, the value 0.92 [32] leans on the particularity of the study as even though performed on only 14 volunteers, this small cohort size allowed an accurate isoflavones intake assessment made by weighted food diaries. The other study reporting a high r value (0.80) for plasma correlations is proposed by Verkasalo et al. [27]. The authors performed a Spearman correlation analysis since the 80 consumer profiles used for their study were distributed into 4 groups only, according to their soy-food intake. They used a statistically unsatisfactory process as only 4 mean isoflavones blood levels—4 values for GEN and 4 values for DAI—were deduced and used as a basis to obtain the correlation, instead of the whole 80 data. In fact, apart from these two interesting yet questionable studies (Ritchie's and Verkasalo's), the coefficient r correlating plasma isoflavones levels and dietary records is always below 0.55, as shown in Table 4. Conversely, the daily isoflavones variations in biological fluids were invisible in hair, by the way giving insights about long-term exposure. Finally, if 24- or 48-h dietary surveys could be suitable for isoflavones assessment in biological fluids, a new dietary questionnaire recording global French food habits focusing on soybean intakes had to be built to validate hair measurements.

4.1.3. The Importance of the Food Intake Estimation

Of course, a good correlation between isoflavone measurements in body tissues and food intake relies on the quality of isoflavone measurements and the reliability of the food inquiry. Past food records were not as accurate as newer ones primarily because the isoflavones concentrations in soy ingredients incorporated into processed foods were underestimated. For instance, in 2006, correlations between isoflavones and dietary records or food databases in the USA ranged from 0.41–0.55 [36]. Although these coefficients were in the same range as those published later, the initial database essentially focused on soy-based products, under-considering products containing "hidden-soy". These could be, for instance, bread in which soy flour was incorporated to whiten the breadcrumbs. In the present study, the food inquiry was mainly focused on soy-based food but also on French foodstuff containing "hidden-soy". These were primarily elaborated dishes or foodstuff based on meat or fish that were previously crushed and restructured. In France, it was shown [16] that soy was not used to whiten breadcrumbs, but broad bean meals could be used in industrial soft bread. Contrary to soybean, broad-bean flour contains only traces of GEN and DAI. Thus, soft bread was not retained in this inquiry. More globally, the recorded items choice was based on the study by Lee et al. [16]. This previous analysis of the French foodstuffs' labels determined which could possibly contain soy as an ingredient. The comprehensive analysis of 12,707 food labels included transformed foodstuff containing minced ingredients [16]. The food categories that were investigated included a large panel of products. For instance, 2 dairy products out of the 2146 analyzed were shown to contain soy. So, dairy products were not assessed in the present inquiry. Moreover, 65 out of 154 labels of minced beef–meat were found to contain soy and, therefore, frozen minced beef portions were analyzed in the present study. Hence, as can be seen in Supplementary Materials S1 and S2, the inquiry recorded: soy juice, soy yogurt, soy cream and ice cream, soy cheese and tofu, other plant-based juices, other soy products (miso, tempeh . . .), meat-based dishes, fish-based dishes, vegetables based dishes, pizzas, elaborated dishes, canned recipes, nuggets, hamburgers, kebabs, meatballs, escalope cordon bleu, stuffed vegetables (tomatoes, cabbages), frozen minced steaks, industrial biscuits, other legumes, chickpeas, soy seeds, soybean sprouts, alfalfa or clover sprouts. In the score-determination-tool, the portion size was only addressed through a question dealing with the eater status, i.e., big–eater, medium–eater, or small–eater. Questions were addressed to refine the consumption habits. Some dealt with the consumption mode, i.e., vegan,

vegetarian, lacto-vegetarian, flexitarian or omnivorous. Others addressed the type of the main meals, i.e., Asian, gluten-free, lactose-free or other. Indeed, it was shown [16] that gluten-free French products could more frequently contain soybean than conventional ones. Moreover, some questions assessed the main cooking practices. For instance, wok or oven have slight effects on isoflavones concentrations. Conversely, boiling in water can reduce isoflavones content significantly. Then, the main meal-place—canteens, fast-foods, restaurants, home, . . . —was recorded, as it was shown [16] that meals taken in canteens or fast-food restaurants could bring more isoflavones than those prepared at home. Finally, the IFS was equal to the sum of all food scores multiplied by a factor considering hair treatments and eater status. The IHS was obtained by multiplying the IFS by a factor including the duration of the described diet and the hair length. The latter represents a dilution factor introduced because the overall hair sample length was analyzed in the assays. The regularly high correlation factors obtained in this study suggest that this dietary approach is correct and takes sufficiently into account the isoflavone intake, including the one provided by "hidden-soy".

4.2. Positive Innovative Elements Brought in This Work

4.2.1. The Assay

Even if 35% of samples were close to the quantification limit in this panel of volunteers, the correlation between isoflavones measurements and the IHS is good (Figure 1). Because soy consumption is still low in France, it was not expected to find high levels of isoflavones in all hair samples. Nevertheless, only 11.6% of the samples assayed were considered as completely free of isoflavones. This reinforces the idea that soy and isoflavones intakes are significant in the French population [17] even if the level of consumption is not always of concern. The assay was also able to detect abnormalities in dietary declarations as will be seen below. Finally, it is also the first time that equol is measured in hair. The proportion of equol producers was found to be 61.82% of 55 volunteers. A similar proportion was found in 2006 in a panel of 60 French women, using the same assay technique [26]. Indeed, it is important to consider the sensitivity of the assay technique while assessing the equol–producer percentages in a population. Hence, in the present study, even if ELISA has a great sensitivity, a large proportion of samples were close to the detection limits and those samples might have gone undetected by other assay methods.

4.2.2. The Food Survey

Here the measurements were validated by the food inquiry which was itself validated by the hair measurements. As can be seen in the Supplementary Materials S1 and S2, the inquiry was targeted to specific foodstuffs containing soy and isoflavones [16,17]. Thus, the inquiry did not take too much time by itself and could be proposed to the volunteers. Moreover, as mentioned earlier, the food intake gave information data, but these had to be adjusted to obtain an IHS which reflected better the isoflavones hair content. Hence, the hair basic color and treatments applied to hair were considered including washing, cutting, coloring as well as perm frequency. The duration of the claimed diet was also very important to consider since, here, the all hair-length were used for measurements. Indeed, due to a mean growth of about 1 cm/month, the last 6 months consumption was only reflected into the 6 cm of hair proximal to the skull. Hence, if the diet was changed for soy, only recently, the trace of the consumption may be diluted if the volunteers wore long hair. Such a phenomenon was initially observed while collecting the first data in this study and led to the corresponding adjustment.

4.3. Challenges of the Present Analytical Method

4.3.1. Digestion Procedure

As a first attempt to assay isoflavones in hair, this work presented some difficulties. The first one is linked to the digestion method retained. NaOH 1 M treatment described in Gaillard et al. [19] was not adequate for isoflavones. This treatment destroyed the

isoflavone-conjugates used as controls except for Genistin. Because Genistin is only present in plants and not in human biological fluids, it is not expected to be present in hair. On hair samples it retrieved isoflavones at levels not significantly different from 0. From multiple tests, it clearly appeared that acid digestion using HCl 0.1 N had given the best recovery results on hair samples. However, as it was observed with Genistein-7-β-D-Glucuronide as control, this digestion process destroyed the glucuronide conjugates, including those potentially present in hair. Interestingly, while the enzymatic procedures better preserved these conjugates, they were not able to retrieve more isoflavones from hair samples. Indeed, as mentioned in Table 2, the pronase and keratinase hair digestions were followed by β-glucuronidase-aryl-sulfatase hydrolysis. The latter was performed to hydrolyze the conjugates into aglucones, if any were present in the hair samples. This step, which generated free GEN, DAI and equol allowed a 98% recovery of the external standards via the ethyl-acetate liquid-liquid extraction. Therefore, it was expected that the enzymatic method preserving the glucuronide conjugates would lead to higher isoflavones values in hair samples. However, it was not the case, when comparing the recovery percentages obtained on hair samples with keratinase or acid digestion. It was assumed that the keratinase treatment was not able to digest the hair matrix as much as the acid treatment and therefore that the isoflavones and metabolites could not be available for β-glucuronidase digestion. Moreover, pronase gave a better recovery percentage on genistein-7-β-D-glucuronide than acid (61.59% ± 7.85% vs. 00.00%) but a lower genistein-7-sulfate recovery (50.97% ± 0.56% vs. 98.32% ± 16.26%). However, on hair samples, the isoflavone recovery was not better than the one obtained with the acidic treatment. Additionally, it was not retained because the inter-sample variation was higher, especially when applied on concentrated hair samples. The overall recovery observed by enzymatic digestion was never significantly different from the acid digestion. In addition to that, the digestion with NaOH 1 N is not adapted to isoflavones as the recovery rates of both conjugates were null. On the one hand, the hair digestion may not have been efficient enough to release all isoflavones and metabolites. On the other hand, if pronase digestion—compared to the acid one—did not lead to higher values in hair samples, it may be because genistein–glucuronide conjugates were not present. Indeed, looking at the literature, it was observed that the only glucuronide form having ever been found in the hair so far, is ethyl-glucuronide used to assess ethanol addiction [37]. Ethanol being a small molecule, may follow a different process to enter the hair compared to isoflavones. The latter are physico-chemically much closer to steroids. Then, looking at steroid measurements in hair samples, many relevant data were found in the literature. They dealt with cortisol, DHEA, testosterone, androgens and even estrogens. These steroid hormones were mainly detected in hair as either free or sulfate-conjugates [38]. According to Pragst and Balikova [37], the basic and lipophilic character of a substance is important for its incorporation into hair. However, Genistein-7-β-D-glucuronide is a rather strong acid substance with a pKA of 2.74. Meanwhile, Genistein-7-O-sulfate is a strong acid substance with an acid pKA of −2.6. Additionally, free GEN is essentially under a neutral form at physiological pH with a pKA of 7.31 [39]. This property together with the lipophilic nature of GEN may help its incorporation into hair-follicle cells. However, free GEN is far from being the major compound in plasma [40] and if there was no conjugate hydrolysis in hair cells' vicinity, this would explain why the ELISAs' sensitivity is challenged while assaying hair samples. DAI and daidzein conjugates follow the same patterns of physicochemical properties. Daidzein-7-β-D-glucuronide pKA is 2.88 and Daidzein-4′-sulfate pKA is -2.4 indicating that both compounds are acidic and strongly acidic respectively in a neutral environment. Meanwhile, free DAI in plasma is lipophilic with poor water solubility and has a pKA of 7.51. Therefore, DAI incorporation in hair can follow the same process as GEN even though the GEN to DAI ratio between plasma and hair may not be completely conserved. Finally, when equol is considered its pKA is 9.63 showing that it acts as a basic compound. It is also lipophilic when not conjugated and these physicochemical characteristics may help its passage into the hair. Conversely, equol-4′-sulfate and equol-7-O-glucuronide have pKA of −2 and 3.27 showing

that they are strongly acidic and mildly acidic respectively. Both of them being soluble in water, they are not likely to be incorporated passively in hair bulbs and, therefore, in hair. Nevertheless, the isoflavones concentrations obtained here i.e., between 0.11 and 11 ng/mg of hair, are in the range of other drug levels in hair [37].

4.3.2. Limit Linked to the Assay Procedure

The ELISA procedure does not allow the use of internal standards since it relies on specific antibodies. The recovery is always assessed by parallel external standards runs. Moreover, the ELISAs used in this study were based on the recognition of isoflavones aglucones by specific antibodies. The latter could not be used to assay the isoflavone-conjugates directly and required hydrolysis prior to measurements. Thus, the presence of the isoflavones conjugates in hair could still be addressed, by other methods. Such a question is important, since glucuronide–conjugates represent by far the highest proportion of isoflavones forms in plasma or serum [40]. If effectively glucuronides are not able to enter the hair, this could explain the low level of isoflavones found in this matrix. Additionally, as mentioned previously, a large proportion of the measurements of hair isoflavones were close to the quantification limit. Therefore, this technique may be relevant to identify frequent soy consumers but may not be able to distinguish low and very low soy consumers. Nevertheless, only frequent and high soy consumption, exposing people to high isoflavone intakes, have so far been involved in health impairments [41,42]. Therefore, despite these limitations the hair measurements described here may be of help in determining the consumers' exposure to isoflavones and its health consequences.

4.3.3. Limit Linked to the Dietary Inquiry

Similar to all other dietary surveys, the one developed here for monitoring isoflavones intake relied on the consumer's declaration. To avoid misinterpretations, several questions were linked to assess the accuracy of nutrition claims and the meals' locations were recorded. Then, when it was reported that meals were essentially taken outdoors, and if the place was not a gastronomic restaurant, the score attributed to food intake was slightly increased. Despite this, when testing the dietary inquiry at the hairdresser, a large discrepancy was found between the IHS and the hair isoflavones concentration of one of the volunteers who showed some resistance to the study. She was in fact opposed to isoflavones potential negative effects since she belongs to a traditionally soy–eating family. Even though she claimed to have her meals at home, she also claimed to be a non-soy-eater. The Pearson regression r value jumped from 0.633 to 0.989 when her data were not included in the calculation, see Supplementary Material S3 for details. This finding confirms the notion that an answer can be psychologically influenced by the initial question. In our case, the volunteers were always informed about the reasons behind the inquiry, and this may have influenced their answers [43,44]. Finally, as can be seen in Supplementary Material S2, a score is attributed to each foodstuff recorded, depending on the isoflavones level measured, but also on the frequency of soy incorporation if the foodstuff is ultra-processed. For instance, not all nuggets found in supermarkets were shown to contain soy as an ingredient [16] but all contained it in canteens. Thus, the score is only valid in the French population based on recent labels' analyses. Finally, although this inquiry was applied to 60 people so far, some foodstuffs with high isoflavones concentrations—such as clover sprouts—were never mentioned by any volunteer. Hence, the scores attributed to them are not completely validated.

5. Conclusions

The present study reports, for the first-time, measurements of isoflavones in hair and shows that they could be well correlated with dietary records especially built for French consumers. Thanks to acid digestion, the isoflavones–sulfate could be preserved during hair extraction but not glucuronide–conjugates, leaving the question of the presence of conjugated forms of isoflavones in hair without a firm answer. However, this should not

prevent the use of hair as a provider of isoflavones exposure biomarkers. This is partly because the use of hair samples to assess a significant exposure of dietary soy isoflavones has proven possible and partly because hair has a remarkable advantage over biological fluids: it can be used to track dietary habits on a long-term basis and, thus, give a more accurate exposure evaluation. Therefore, isoflavone assays in hair samples may be used to monitor certain patient populations at risk of excessive exposure to soy–isoflavones, in particular for pathologies related to estrogens or thyroid hormones. As a result, appropriate dietary advice could be provided to avoid health impairments linked to the risks of interactions between high exposure to isoflavones and pathologies or drug treatments.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/nu14173619/s1, Supplementary Material S1: Dietary inquiry proposed to volunteers, Supplementary Material S2: Matrix used to calculate IFS and HIS, Supplementary Material S3: Evidencing a questionable answer to the food questionnaire, Supplementary Material S4: Correlation regression graphs between IHS and GEN or DAI.

Author Contributions: S.B. participated in the hair isoflavone analysis as well as manuscript writing. I.R. supervised the sample collection, the inquiry process with volunteers, and manuscript writing. M.B. was involved in the digestion tests and results analysis. I.P. was in charge of the sample management and conservation and participated in the manuscript writing. J.-F.V. was the main investigator for the clinical portion of this trial and also participated in the manuscript writing. C.B.-P. initiated and supervised the analytical work. She was involved in the assays, inquiry creation, data analyses, and manuscript writing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no specific external funding.

Institutional Review Board Statement: The ISOLED Study The study was conducted in accordance with the Declaration of Helsinki, and approved by Committee for the Protection of Persons (CPP) (protocol code 2017T2-29 given in June 2017).

Informed Consent Statement: An informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available in Supplementary Materials and raw data can be sent upon request.

Acknowledgments: We thank the hairdresser "Sophie Coiffure" in Camblanes et Meynac, Gironde, for participating in this study and introducing the volunteers to the present topic, and Bordeaux Biothèques Santé BBS-BRC for their excellent technical assistance. This work was supported by Inserm, by Bordeaux Science Agro, and by the New Aquitaine Region. The assays were possible thanks to the Regional Council of Nouvelle Aquitaine.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Lee, G.A.; Crawford, G.W.; Liu, L.; Sasaki, Y.; Chen, X. Archaeological Soybean (Glycine Max) in East Asia: Does Size Matter? *PLoS ONE* **2011**, *6*, e26720. [CrossRef]
2. Bennetau-Pelissero, C. Plant Proteins from Legumes. In *Bioactive Molecules in Food. Reference Series in Phytochemistry*; Mérillon, J.M., Ramawat, K., Eds.; Springer: Cham, Switzerland, 2019; pp. 223–265.
3. Bennetts, H.W.; Uuderwood, E.J.; Shier, F.L. A Specific Breeding Problem of Sheep on Subterranean Clover Pastures in Western Australia. *Aust. Vet. J.* **1946**, *22*, 2–12. [CrossRef]
4. Stott, P.; Harris, S.; Wight, N. Fertility and Infertility in the European Hare in Australia. In *Lagomorph Biology*; Alves, P.C., Ferrand, N., Hackländer, K., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 225–240.
5. Chadha, R.; Bhalla, Y.; Jain, A.; Chadha, K.; Karan, M. Dietary Soy Isoflavone: A Mechanistic Insight. *Nat. Prod. Commun.* **2017**, *12*, 627–634. [CrossRef]
6. Mikulić, M.; Krstonošić, M.; Sazdanić, D.; Cvejic, J. Health Perspectives on Soy Isoflavones. In *Phytochemicals in Soybeans*; CRC Press: Boca Raton, FL, USA, 2022.
7. Fernandez-Lopez, A.; Lamothe, V.; Delample, M.; Denayrolles, M.; Bennetau-Pelissero, C. Removing Isoflavones from Modern Soyfood: Why and How? *Food Chem.* **2016**, *210*, 286–294. [CrossRef] [PubMed]

8. Ismanto, A.; Hadibarata, T.; Kristanti, R.A.; Maslukah, L.; Safinatunnajah, N.; Kusumastuti, W. Endocrine Disrupting Chemicals (EDCs) in Environmental Matrices: Occurrence, Fate, Health Impact, Physio-Chemical and Bioremediation Technology. *Environ. Pollut.* **2022**, *302*, 119061. [CrossRef]
9. Nusbaum, J.S.; Mirza, I.; Shum, J.; Freilich, R.W.; Cohen, R.E.; Pillinger, M.H.; Izmirly, P.M.; Buyon, J.P. Sex Differences in Systemic Lupus Erythematosus: Epidemiology, Clinical Considerations, and Disease Pathogenesis. *Mayo Clin. Proc.* **2020**, *95*, 384–394. [CrossRef] [PubMed]
10. Barbhaiya, M.; Costenbader, K.H. Environmental Exposures and the Development of Systemic Lupus Erythematosus. *Curr. Opin. Rheumatol.* **2016**, *28*, 497–505. [CrossRef]
11. Jang, H.H.; Lee, Y.M.; Choe, J.S.; Kwon, O. Validation of Soy Isoflavone Intake and Its Health Effects: A Review of the Development of Exposure Biomarkers. *Nutr. Res. Pract.* **2021**, *15*, 1–11. [CrossRef] [PubMed]
12. Avino, P.; Lammardo, M.; Petrucci, A.; Rosada, A. Weekly and Longitudinal Element Variability in Hair Samples of Subjects Non-Occupationally Exposed. *Appl. Sci.* **2021**, *11*, 1236. [CrossRef]
13. Le Houérou, C.; Bennetau-Pelissero, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Syntheses of Novel Hapten-Protein Conjugates for Production of Highly Specific Antibodies to Formononetin, Daidzein and Genistein. *Tetrahedron* **2000**, *56*, 295–301. [CrossRef]
14. Bennetau-Pelissero, C.; Le Houérou, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Synthesis of Haptens and Conjugates for ELISAs of Phytoestrogens. Development of the Immunological Tests. *J. Agric. Food Chem.* **2000**, *48*, 305–311. [CrossRef] [PubMed]
15. Shinkaruk, S.; Durand, M.; Lamothe, V.; Carpaye, A.; Martinet, A.; Chantre, P.; Vergne, S.; Nogues, X.; Moore, N.; Bennetau-Pelissero, C. Bioavailability of Glycitein Relatively to Other Soy Isoflavones in Healthy Young Caucasian Men. *Food Chem.* **2012**, *135*, 1104–1111. [CrossRef] [PubMed]
16. Lee, A.; Beaubernard, L.; Lamothe, V.; Bennetau-Pelissero, C. New Evaluation of Isoflavone Exposure in the French Population. *Nutrients* **2019**, *11*, 2308. [CrossRef] [PubMed]
17. Lee, A.; Bensaada, S.; Lamothe, V.; Lacoste, M.; Bennetau-Pelissero, C. Endocrine Disruptors on and in Fruits and Vegetables: Estimation of the Potential Exposure of the French Population. *Food Chem.* **2022**, *373*, 131513. [CrossRef]
18. Vergne, S.; Titier, K.; Bernard, V.; Asselineau, J.; Durand, M.; Lamothe, V.; Potier, M.; Perez, P.; Demotes-Mainard, J.; Chantre, P.; et al. Bioavailability and Urinary Excretion of Isoflavones in Humans: Effects of Soy-Based Supplements Formulation and Equol Production. *J. Pharm. Biomed. Anal.* **2007**, *43*, 1488–1494. [CrossRef]
19. Gaillard, Y.; Vayssette, F.; Pépin, G. Compared Interest between Hair Analysis and Urinalysis in Doping Controls: Results for Amphetamines, Corticosteroids and Anabolic Steroids in Racing Cyclists. *Forensic Sci. Int.* **2000**, *107*, 361–379. [CrossRef]
20. Offidani, C.; Strano Rossi, S.; Chiarotti, M. Improved Enzymatic Hydrolysis of Hair. *Forensic Sci. Int.* **1993**, *63*, 171–174. [CrossRef]
21. Gupta, S.; Singh, R. Hydrolyzing Proficiency of Keratinases in Feather Degradation. *Indian J. Microbiol.* **2014**, *54*, 466–470. [CrossRef]
22. Balagova, L.; Jezova, D. Importance of Methodological Details in the Measurement of Cortisol in Human Hair. *Endocr. Regul.* **2018**, *52*, 134–138. [CrossRef]
23. Gao, W.; Kirschbaum, C.; Grass, J.; Stalder, T. LC–MS Based Analysis of Endogenous Steroid Hormones in Human Hair. *J. Steroid Biochem. Mol. Biol.* **2016**, *162*, 92–99. [CrossRef]
24. Balíková, M.A.; Habrdová, V. Hair Analysis for Opiates: Evaluation of Washing and Incubation Procedures. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2003**, *789*, 93–100. [CrossRef]
25. Grace, P.B.; Taylor, J.I.; Low, Y.-L.; Luben, R.N.; Mulligan, A.A.; Botting, N.P.; Dowsett, M.; Welch, A.A.; Khaw, K.-T.; Wareham, N.J.; et al. Phytoestrogen Concentrations in Serum and Spot Urine as Biomarkers for Dietary Phytoestrogen Intake and Their Relation to Breast Cancer Risk in European Prospective Investigation of Cancer and Nutrition-Norfolk. *Cancer Epidemiol. Biomarkers Prev.* **2004**, *13*, 698–708. [CrossRef] [PubMed]
26. Mathey, J.; Lamothe, V.; Coxam, V.; Potier, M.; Sauvant, P.; Bennetau-Pelissero, C. Concentrations of Isoflavones in Plasma and Urine of Post-Menopausal Women Chronically Ingesting High Quantities of Soy Isoflavones. *J. Pharm. Biomed. Anal.* **2006**, *41*, 957–965. [CrossRef] [PubMed]
27. Verkasalo, P.K.; Appleby, P.N.; Allen, N.E.; Davey, G.; Adlercreutz, H.; Key, T.J. Soya Intake and Plasma Concentrations of Daidzein and Genistein: Validity of Dietary Assessment among Eighty British Women (Oxford Arm of the European Prospective Investigation into Cancer and Nutrition). *Br. J. Nutr.* **2001**, *86*, 415–421. [CrossRef]
28. Atkinson, C.; Skor, H.E.; Fitzgibbons, E.D.; Scholes, D.; Chen, C.; Wähälä, K.; Schwartz, S.M.; Lampe, J.W. Overnight Urinary Isoflavone Excretion in a Population of Women Living in the United States, and Its Relationship to Isoflavone Intake. *Cancer Epidemiol. Biomarkers Prev.* **2002**, *11*, 253–260, Erratum in *Cancer Epidemiol. Biomarkers Prev.* **2022**, *11*, 1511.
29. Frankenfeld, C.L.; Patterson, R.E.; Kalhorn, T.F.; Skor, H.E.; Howald, W.N.; Lampe, J.W. Validation of a Soy Food Frequency Questionnaire with Plasma Concentrations of Isoflavones in US Adults. *J. Am. Diet. Assoc.* **2002**, *102*, 1407–1413. [CrossRef]
30. Heald, C.L.; Bolton-Smith, C.; Ritchie, M.R.; Morton, M.S.; Alexander, F.E. Phyto-Oestrogen Intake in Scottish Men: Use of Serum to Validate a Self-Administered Food-Frequency Questionnaire in Older Men. *Eur. J. Clin. Nutr.* **2006**, *60*, 129–135. [CrossRef]
31. Morimoto, Y.; Beckford, F.; Franke, A.A.; Maskarinec, G. Urinary Isoflavonoid Excretion as a Biomarker of Dietary Soy Intake during Two Randomized Soy Trials. *Asia Pac. J. Clin. Nutr.* **2014**, *23*, 205–209. [CrossRef]
32. Ritchie, M.R.; Morton, M.S.; Deighton, N.; Blake, A.; Cummings, J.H. Plasma and Urinary Phyto-Oestrogens as Biomarkers of Intake: Validation by Duplicate Diet Analysis. *Br. J. Nutr.* **2004**, *91*, 447–457. [CrossRef]

33. Huang, M.H.; Harrison, G.G.; Mohamed, M.M.; Gornbein, J.A.; Henning, S.M.; Go, V.L.W.; Greendale, G.A. Assessing the Accuracy of a Food Frequency Questionnaire for Estimating Usual Intake of Phytoestrogens. *Nutr. Cancer* **2000**, *37*, 145–154. [CrossRef]
34. Kim, J.; Kim, H.J.; Joung, H.; Park, M.K.; Li, S.; Song, Y.; Franke, A.A.; Paik, H.Y. Overnight Urinary Excretion of Isoflavones as an Indicator for Dietary Isoflavone Intake in Korean Girls of Pubertal Age. *Br. J. Nutr.* **2010**, *104*, 709–715. [CrossRef] [PubMed]
35. Chávez-Suárez, K.M.; Ortega-Vélez, M.I.; Valenzuela-Quintanar, A.I.; Galván-Portillo, M.; López-Carrillo, L.; Esparza-Romero, J.; Saucedo-Tamayo, M.S.; Robles-Burgueño, M.R.; Palma-Durán, S.A.; Gutiérrez-Coronado, M.L.; et al. Phytoestrogen Concentrations in Human Urine as Biomarkers for Dietary Phytoestrogen Intake in Mexican Women. *Nutrients* **2017**, *9*, 1078. [CrossRef] [PubMed]
36. Horn-Ross, P.L.; Barnes, S.; Lee, V.S.; Collins, C.N.; Reynolds, P.; Lee, M.M.; Stewart, S.L.; Canchola, A.J.; Wilson, L.; Jones, K. Reliability and Validity of an Assessment of Usual Phytoestrogen Consumption (United States). *Cancer Causes Control* **2006**, *17*, 85–93. [CrossRef] [PubMed]
37. Pragst, F.; Balikova, M.A. State of the Art in Hair Analysis for Detection of Drug and Alcohol Abuse. *Clin. Chim. Acta* **2006**, *370*, 17–49. [CrossRef]
38. Cho, S.H.; Choi, M.H.; Sim, W.Y.; Lee, W.Y.; Chung, B.C. Metabolic Alterations of DHEA and Cholesterol Sulphates in the Hair of Patients with Acne Measured by Liquid Chromatography-Mass Spectrometry. *Exp. Dermatol.* **2010**, *19*, 694–696. [CrossRef]
39. Mielczarek, C.; Pająk, W. Analysis of Acid-Base Properties of Flavonoid Genistein. *J. Appl. Spectrosc.* **2013**, *80*, 737–744. [CrossRef]
40. Soukup, S.T.; Al-Maharik, N.; Botting, N.; Kulling, S.E. Quantification of Soy Isoflavones and Their Conjugative Metabolites in Plasma and Urine: An Automated and Validated UHPLC-MS/MS Method for Use in Large-Scale Studies. *Anal. Bioanal. Chem.* **2014**, *406*, 6007–6020. [CrossRef]
41. Chandrareddy, A.; Muneyyirci-Delale, O.; McFarlane, S.I.; Murad, O.M. Adverse Effects of Phytoestrogens on Reproductive Health: A Report of Three Cases. *Complement. Ther. Clin. Pract.* **2008**, *14*, 132–135. [CrossRef]
42. Imai, H.; Nishikawa, H.; Suzuki, A.; Kodama, E.; Iida, T.; Mikura, K.; Hashizume, M.; Kigawa, Y.; Tadokoro, R.; Sugisawa, C.; et al. Secondary Hypogonadism Due to Excessive Ingestion of Isoflavone in a Man. *Intern. Med.* **2022**. [CrossRef]
43. Satya-Murti, S.; Lockhart, J. Recognizing and Reducing Cognitive Bias in Clinical and Forensic Neurology. *Neurol. Clin. Pract.* **2015**, *5*, 389–396. [CrossRef]
44. Kumar, V. The Power of Expectations: Undermining the Role of Pygmalion Effect in Mentoring. *CHRISMED J. Health Res.* **2019**, *6*, 126–127. [CrossRef]

IV.1.2.2. Evaluation de la consommation récente : validation de questionnaires alimentaires par l'analyse de biomarqueurs urinaires et plasmatiques. Article correspondant à la référence [25]

L'évaluation de l'exposition aux phytoE a mis en jeu des questionnaires inédits, développés à cet effet. Ils ont été administrés aux volontaires de l'étude ISOLED et leur validité a été confirmée par l'analyse de biomarqueurs sanguins et urinaires. Ces travaux font l'objet de la publication ci-après [25].

Dans le but d'apporter un éclairage supplémentaire sur certains points, des éléments de notation des catégories d'aliments sont exposés ici.

- IFs élevés : jus de soja ; tofu et « fromage » de soja ; aliments au soja.
- IFs moyens : yaourts au soja.
- Transformé ++ : produits contenant fréquemment du soja (ex : boulettes de viande, burgers, …)
- Transformé + : Produits contenant du soja occasionnellement (nuggets, kebabs, …).
- Transformé : produits pouvant contenir du soja de manière incertaine (pâtisseries industrielles, ...)
- Transformé - : produits très peu susceptibles de contenir du soja (ex : produits laitiers).
- Cantine : produits utilisés contiennent du soja plus fréquemment que ceux préparés à domicile.

Malgré les effectifs de volontaires relativement faibles, l'étude représente la première tentative en France d'estimer les concentrations d'IFs et d'ENL chez des femmes non sélectionnées sur leur régime alimentaire.

Il est crucial de souligner que les résultats obtenus corroborent l'idée selon laquelle l'étude des effets des entérolignanes sur la santé ne peut être menée qu'à travers l'analyse des fluides biologiques. En effet, les entérolignanes étant produits par la flore intestinale, seuls les dosages dans le fluides biologiques permettent d'établir une corrélation entre la présence d'entérolignanes et des effets biologiques. Les résultats suggèrent également des pharmacocinétiques différentes entre IFs et ENL, avec une clairance rénale probablement plus élevée pour l'ENL. Ceci pourrait expliquer le fait que les effets des entérolignanes chez les consommateur·ices humains semblent être peu préoccupants.

nutrients

MDPI

Article

Validation of ELISAs for Isoflavones and Enterolactone for Phytoestrogen Intake Assessment in the French Population

Souad Bensaada [1,2], Isabelle Raymond [1,3], Isabelle Pellegrin [1,4], Jean-François Viallard [1,3] and Catherine Bennetau-Pelissero [1,2,5,*]

[1] Department Sciences and Technology for Health, Campus Carreire, University of Bordeaux, 33076 Bordeaux, France
[2] ARNA, U1212 Inserm, 5320 CNRS, Pharmacy Faculty, 33076 Bordeaux, France
[3] Centre Hospitalo Universitaire Bordeaux, USN B0-Hôpital Haut Lévêque, 33604 Pessac, France
[4] Centre Hospitalo-Universitaire Bordeaux, Laboratory of Immunology and Immunogenetics, Resources Biological Center (CRB), 33000 Bordeaux, France
[5] Bordeaux Sciences Agro, 33175 Gradignan, France
* Correspondence: catherine.bennetau@agro-bordeaux.fr

Abstract: Phytoestrogens are dietary compounds with low estrogenic activity. The two main categories in the French diet are isoflavones from pulses and enterolignans metabolized by the gut flora from various lignans found in fruits, vegetables, grains, and beverages. Isoflavones and lignans have different effects on human physiology and can antagonize each other. Comprehensive lists of phytoestrogen sources were constructed based on measurements and literature data. The 24 h and 48 h dietary recalls were proposed to the volunteers of the ISOLED cohort (NCT03421184). Urine and plasma samples from these volunteers were assayed for genistein, daidzein, equol, and enterolactone. A dietary score was constructed considering the pharmacokinetic characteristics of these compounds. Correlation analyses were applied to fluid concentrations associated with dietary scores. Pearson correlations reached 0.921 ($p < 0.001$) for $urine_{IF}$, 0.900 ($p < 0.001$) for $plasma_{IF}$, 0.764 ($p < 0.001$) for $urine_{ENL}$, and 0.723 ($p < 0.001$) for $plasma_{ENL}$. ELISAs associated with careful intake assessments proved to be good tools for phytoestrogens' exposure estimation.

Keywords: phytoestrogens; soy isoflavones; enterolactone; plasma; urine; dietary inquiry; exposure assessment

check for updates

Citation: Bensaada, S.; Raymond, I.; Pellegrin, I.; Viallard, J.-F.; Bennetau-Pelissero, C. Validation of ELISAs for Isoflavones and Enterolactone for Phytoestrogen Intake Assessment in the French Population. *Nutrients* **2023**, *15*, 967. https://doi.org/10.3390/nu15040967

Academic Editor: Adam Matkowski

Received: 14 January 2023
Revised: 7 February 2023
Accepted: 9 February 2023
Published: 15 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Phytoestrogens (PHYTOs) are naturally occurring polyphenols that have endocrine effects at relevant doses [1]. Sources of estrogenic isoflavones (IF), i.e., genistein (GEN), daidzein (DAI), and glycitein (Figure 1A), in the human diet are essentially pulses including alfalfa, clover, and soy. However, nowadays, only soybean is a significant source of IFs in the French diet [2,3]. Conversely, enterolactone (ENL) is produced by the gut flora [4,5] and its precursors were shown to be numerous in the diet. Indeed, previous authors showed that besides seccoisolariciresinol di-glucoside (SECO or SDG), which is the main ENL-precursor, matairesinol (MAT), lariciresinol (LAR), pinoresinol (PINO), syringaresinol (SYR), medioresinol (MED), and some of their hydroxylated metabolites and glucosides can be metabolized into ENL [6] (see Figure 1C,D). The conversion rate was shown to vary with the individual microflora. Therefore, exposure to ENL-precursors was not necessarily associated with an estrogenic effect [7]. Moreover, ENL-precursors were shown to be present in a large panel of plant foodstuffs, including fruits, vegetables, nuts, grains, and beverages (see Table S1).

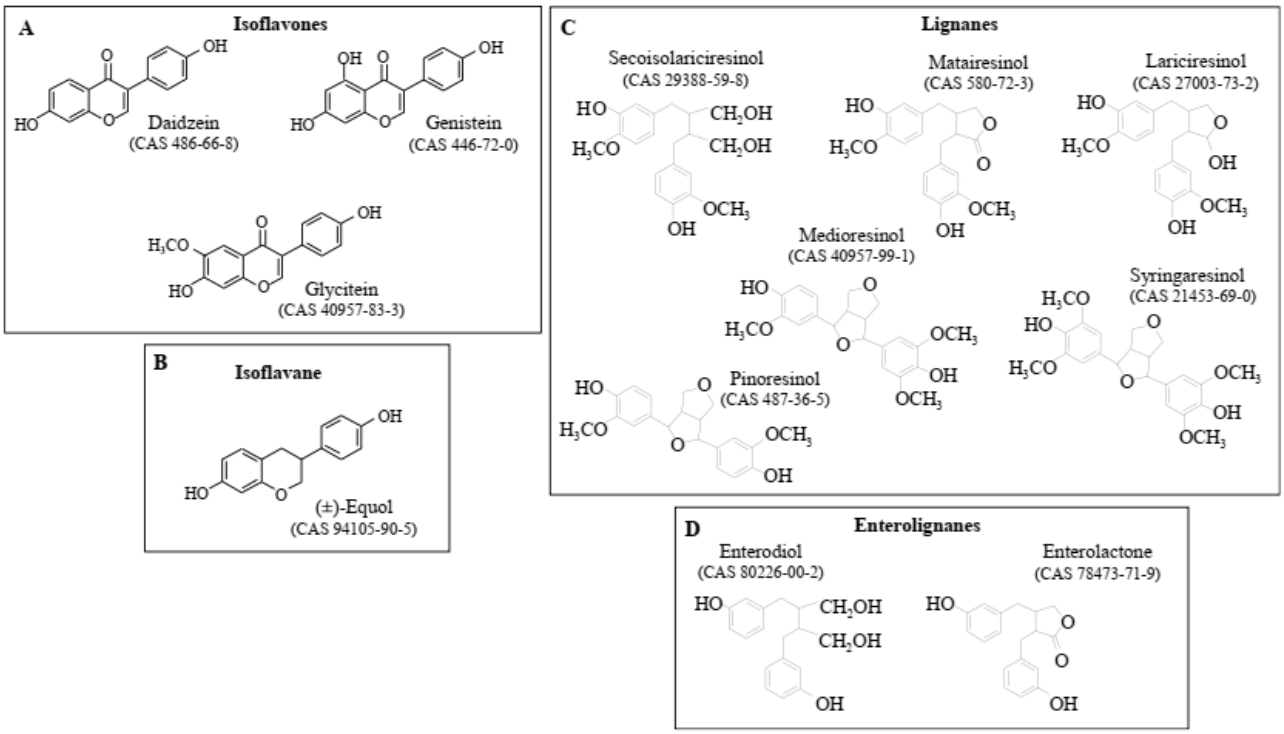


Figure 1. (**A**) Isoflavones, (**B**) the isoflavane metabolite equol, (**C**) the enterolactone-precursors considered in this study, and (**D**) the enterolignans known to be produced by the human gut flora.

Within the fruit category, berries, plums, and citrus were consistently shown to contain significant amounts of ENL-precursors, while in the vegetables group, cabbages, squashes, soy, and legumes were consistently shown to contain significant amounts of ENL-precursors.

Therefore, while the assessment of IFs intake could simply rely on the determination of soy and "hidden soy" intakes [2,3], ENL-precursors' intake assessment should consider many plant foodstuffs (see Table S1, and associated references).

The estrogenic effects of IF and ENL could either be beneficial, when an estrogen deficiency was diagnosed (i.e., at menopause), or adverse when estrogens were not required. This is the case in infancy, during childhood, during pregnancy, in males, or when an estrogen-dependent pathology is feared or diagnosed [8]. However, ENL was shown to have a much lower affinity for the estradiol receptors (ERs) than IFs [9]. It is also present at low doses in plasma [10] and is a selective ER modulator (SERM) whose estrogenic effect essentially relies on the activating function 2 (AF-2) domain of ERα [11]. This property probably explains the anti-proliferation effect of ENL on human breast cancer cells when present in food together with IFs in the nude athymic mice model [12] and in human population approaches [13]. Meanwhile, IFs were the subject of many debate as their in vitro estrogenic activity was found to be the same as estradiol [14]. Their proliferative effects on estrogen breast cancer cells were shown in many situations in vitro [15], in animals [16], and in humans [17]. Conversely, epidemiological studies showed that dietary soy IFs exposure can reduce the incidence of breast or colorectal cancers [18,19]. One plausible explanation would be that IFs can act as growth factors on estrogen-dependent cancer cells, while they can prevent cancerization in healthy cells. The two mechanisms may occur based on different biochemical properties of IFs.

To better understand the potential health actions of PHYTOs in the Western diet, a better estimation of the consumers' exposure is required. Human body fluids can provide biomarkers of exposure, but their reliability should be challenged using dietary inquiries. This is what was proposed in this study where 24 h and 48 h dietary recall inquiries were

proposed to the 57 volunteers enrolled so far in the ISOLED study. ISOLED analyzes the potential link between PHYTOs exposure and the occurrence of systemic lupus erythematosus (SLE) crises. SLE as an autoimmune disease which is known to exhibit a highly unbalanced sex-ratio occurrence [20]. The current hypotheses link this occurrence to genetic factors bearded by the X chromosome and to the estrogen impregnation. Indeed, SLE progression is neither observed in pre-adolescent girls nor in postmenopausal women, and the role of the environmental estrogen in SLE crises' occurrence is under debate [21].

The use of body fluids as biomarker providers of PHYTOs exposure has led us to considering the substances monitored pharmacokinetics. That is why, in this study the dietary score calculated considered the delays of the appearance and disappearance of each compound in the body fluids (see Table S2A–D). The pharmacokinetic parameters were obtained in previous studies [7,22]. Correlations were calculated between fluid concentrations and dietary scores to determine the reliability of the exposure estimation. These correlation coefficients were compared to those obtained in previous studies.

2. Materials and Methods

2.1. Materials

2.1.1. Volunteers

The volunteers were women included in the ISOLED clinical trial registered at clinicaltrial.gov PRS under the registration number NCT03421184, which is still in its inclusion phase. The trial was approved by the local ethics committee under the name ISOLED and the number 2017T2-29. The consort diagram flow is presented in the Supplementary Material as Figure S1. The health status of the volunteers is unknown due to the blind procedure. The protocol provided the recruitment of at least 100 premenopausal women: 30 being healthy controls without autoimmune diseases, 30 being diagnosed with SLE, and 40 having another autoimmune disease where the link with estrogens is not so obvious. This study was planned to be a first exploration, to check for the biological sample analyses and for the food inquiry feasibility. If a tendency would be detected, it was planned to open it up to a multicentric recruitment in order to find more volunteers with autoimmune diseases. The volunteers were welcomed in the Internal Medicine Service of the Haut-Lévêque Hospital in Bordeaux. Their health status was assessed by a medical questionnaire performed by a nurse and by biological analyses. The women with an autoimmune disease were already followed in the medical service and their health status was well-known. The included women were all in a premenopausal status. Their contraceptive treatment was registered. They should be healthy except for their autoimmune disease, if any. The exclusion criteria were infection with VIH, Hepatitis B and C, pregnancy or breast-feeding, menopause, and autoimmune pathology in remission or in quiescent form. All volunteers presenting a liver or a kidney disease were excluded. The health questionnaire included the gut, liver, and kidney status. They were asked to answer to a 24 h and 48 h dietary recall inquiry which was held by a nurse recruited in the medical service. No dietary recommendations were provided, and the premenopausal women recruited in the control group represented an unbiased sample of the normal French premenopausal population. The details of the open inquiries are available in the Supplementary Materials as Document 1 (SD1). Volunteers were always recruited between 9:30 am and 2:30 pm and the biological samples were all collected during this period. The time of sampling was recorded and used subsequently for the dietary score calculation. Finally, an example of the informed consent form which was translated from the French original version is available as Supplementary Document 2 (SD2).

2.1.2. Sample Treatments

- Digestion and extraction of the conjugated isoflavones

The digestion of PHYTOs conjugates (glucuronides and sulfates) was performed with β-glucuronidase-aryl sulfatase from *Helix pomatia* (Roche, 10127698001), diluted (5 $\mu L.mL^{-1}$) in sodium acetate buffer 0.1 M, 0.14 M EDTA, 100 $UI.mL^{-1}$ penicillin G

(SIGMA, P-3032), 0.1 mg.mL^{-1} streptomycin (SIGMA, S-6501), pH 5. The extraction of aglycone compounds was then performed using acidified ethyl-acetate (500 µL HCl 38% per L of ethyl-acetate).

- ELISA measurements

All salt reagents were from VWR-France. The protocol also involved thyroglobulin from swine (SIGMA, T1126), bovine serum albumin (EUROMEDEX, 04-100-812-C), and secondary antibody goat anti-rabbit IgG antibody-Amdex (SIGMA, RPN4301). The revelation steps required o-phenylenediamine dihydrochloride (OPD) (SIGMA, P1526). Stock solutions were prepared at 10 mg.mL^{-1} in water and stored in black vials at −22 °C. Glassware in contact with unconjugated PHYTOs was coated with Sigmacote® (SIGMA, SL2). The sample, washing, and assay buffers were phosphate-buffered saline 0.01 M, 0.9% NaCl, 0.2% Tween, and 1% DMSO, pH 7.3. To obtain the saturation and antibody buffers, 1.6 mg.L^{-1} of bovine serum albumin was added to the latter. The revelation buffer was citrate-phosphate buffer 0.15 M, pH 5, with 0.05 mg of OPD.mL^{-1}. The stop solution was H_2SO_4 4 M. The primary antibodies were selected according to previous works [23–25]. They were harvested if required to avoid undesired cross-reactions with the unwanted antigen.

2.1.3. Creatinine Measurements

Creatinine was measured in urine samples using the LTA Creatinine kit (LIBIOS ref. Creatinine-600). It was designed for colorimetric determination of creatinine in biological liquids without deproteinization.

2.2. Methods

2.2.1. Dietary Inquiry and Dietary Scores

Both the 24 and 48 h dietary inquiries are open questionnaires with several items cited to avoid forgetting (see Supplementary document SD1 for details on the first 24 h recall). The inquiries were analyzed and translated into a dietary score. The latter considered the level of PHYTOs in foodstuffs that were documented either in [3] for IFs in French foodstuffs or by other investigators for ENL-precursors (see Table S1, and associated references). In those studies, the concentration values were all expressed in µg/100 g of fresh food and in aglycone equivalents. The score calculation considered the average content of PHYTOs by foodstuff combined with the occurrence of food containing PHYTOs in this category. For enterolignans-precursors, subcategories of vegetables and fruits were determined based on their average content in ENL-precursors and on common French diet habits.

The foodstuffs considered were mainly soy-based foodstuffs and transformed foodstuffs known to contain "hidden soy". They were soy juice, tofu/soy-based cheese, soy-food, soy-based yogurt, and transformed food (see Table S2A,B). These foodstuffs were graded according to their estimated concentrations of IFs. For instance, soy-based yogurt had a lower score than tofu or soy-based food since their IFs content was previously shown to be lower [2]. In addition, within each food category, different proportions of transformed food items were known to contain soy proteins with IFs. Indeed, within the frozen burgers category, it is known that more than 50% of items contained up to 30% of soy proteins when bought by the consumers in supermarkets. Conversely, within the frozen nuggets category, only 30% were known to contain soy proteins when bought in the same place. However, in canteens, the proportion of transformed food containing soy was shown to be higher than in equivalent food found in supermarkets. This is why when the volunteers claimed to have their meals in a canteen, an additional score was applied (see Table S2A–D).

More importantly, because both IFs and ENL were known to only be present transiently in body fluids, the scores considered the pharmacokinetic behaviours of GEN, DAI and ENL in plasma and in urine after ingestion. The pharmacokinetic parameters originated from Shinkaruk et al. study [22] for IFs and Kuijsten et al. [7] for ENL (See Table S2A–D for details). Concretely, for establishing those accurate dietary scores, six determination scales were applied to each volunteer's consumption claims. Thus, the scales were adapted: (1) to the correlation with $urine_{GEN}$ concentrations, (2) to the correlation with $plasma_{GEN}$

concentrations, (3) to the correlation with $urine_{DAI}$ concentrations, (4) to the correlation with $plasma_{DAI}$ concentrations (5) to the correlation with $urine_{ENL}$ concentrations (6) to the correlation with $plasma_{ENL}$ concentrations (See Table S2A–D).

2.2.2. ELISA Measurements in Body Fluids

ENL, GEN, DAI, and equol (EQ)—a DAI metabolite produced by the gut flora of EQ-producers-were assayed in the urine and plasma samples using specific ELISAs, as described previously [22,23]. Haptens functionalized on different carbon atoms were previously synthesized [23–25]. Specific polyclonal antibodies were obtained for each PHYTO bound to bovine serum albumin and the best antibodies were retained for assay development. Standard curve preparations and sample dilutions were performed in silicone-coated glass vials. The ELISAs followed a competitive procedure with an immobilized competitor, which was the homologous hapten bound to swine thyroglobulin, except for ENL [23]. The sensitivities varied between 0.08 ng/well and 0.4 ng/well, depending on the substance assayed, and intra-assay variation was always below 7% and inter-assay variation was below 17%. For the plasma samples, final dilutions varied between 1:5 and 1:20. For urine samples, the final dilutions varied between 1:5 and 1:4000 due to various levels of natural dilution of the urine samples. To take this phenomenon into account, the measured concentrations were corrected based on a creatinine excretion measurement performed on the same urine samples.

2.2.3. Statistical Treatments

Each concentration value was always a mean and standard deviation of at least triplicate measurements. Due to the large cohort analyzed, the data followed a normal distribution. Therefore, the correlations between PHYTOs concentrations in fluids and dietary scores as well as their significance were calculated following the Pearson procedure. The significance was obtained by comparing the *t* value with data from Student's significance tests at two degrees of freedom.

3. Results

3.1. Biological Fluid Measurements

3.1.1. Isoflavones (IFs)

Results of IFs assayed in the urine and plasma samples of the 57 volunteers are summarized in Table 1. In urine, based on 54 samples and after correction by the creatinine excretion, the mean concentration of GEN + DAI + EQ was 8484.50 ± 61,775.14 $\mu g.L^{-1}$. Three urine samples were below the detection limit and therefore were not considered for the mean, median, and standard deviation. The median concentration was 124.32 $\mu g.L^{-1}$ with a minimum at 9.53 $\mu g.L^{-1}$ and a maximum at 454,384.51 $\mu g.L^{-1}$. Based on a urine concentration threshold at 150 $\mu g.L^{-1}$, 38.60% of the volunteers were considered to have a significant IFs exposure. EQ, was detected in 26 volunteers at a urinary concentration over 20 $\mu g.L^{-1}$. Additionally, EQ was undetectable in 11 urine samples and thus the parameters presented in Table 1 were obtained on 46 volunteers. This means that, based on urine detection, 56.52% of the cohort analyzed was able to significantly produce EQ. Besides, when EQ was detectable in the samples, the mean, median, minimum and maximum concentration values were 72.99 ± 219.72 $ng.mL^{-1}$, 21.63 $ng.mL^{-1}$, 3.37 $ng.mL^{-1}$, and 1469.81 $ng.mL^{-1}$, respectively.

In plasma samples, IFs concentrations (i.e., GEN + DAI + EQ) were much lower, and in 23 samples out of 57, they were not all detectable. Therefore, the summarizing data presented in Table 1 were based on 26 subjects only, for whom GEN and DAI were over the quantification limits. For IFs plasma concentrations the mean, median, minimum and maximum values were 74.95 ± 147.40 $ng.mL^{-1}$, 21.21 $ng.mL^{-1}$, 729.60 $ng.mL^{-1}$ and 1.17 $ng.mL^{-1}$, respectively. Only 14 volunteers out of 57 had a plasma concentration of GEN + DAID + EQ over 100 $ng.mL^{-1}$. EQ was detectable in 28 plasma samples collected from the 57 volunteers, but only 12 of them showed EQ concentrations over

2 ng.mL^{-1}. This corresponds to 49.12% of EQ-producers and only 29.82% of significant EQ-producers. The mean and median EQ concentrations in plasma were 4.18 ± 8.86 ng.mL^{-1} and 1.72 ng.mL^{-1}, respectively. The minimum and maximum concentrations were 0.45 ng.mL^{-1} and 46.57 ng.mL^{-1}, respectively.

Table 1. Summary of data obtained on phytoestrogens' concentrations for the different analyses.

	Plasma$_{IFs}$ (*n* = 26)	**Urine$_{Ifs}$ * (*n* = 54)**	**Urine$_{EQ}$ * (*n* = 46)**	**Plasma$_{ENL}$ (*n*= 57)**	**Urine$_{ENL}$ * (*n* = 57)**
Mean (µg.L^{-1})	74.95	8484.94	72.99	17.89	802.53
Standard deviation (µg.L^{-1})	147.40	61,775.14	219.72	11.71	1438.85
Median (µg.L^{-1})	21.21	124.32	21.63	13.68	304.73
Min (µg.L^{-1})	1.17	9.53	3.37	1.94	6.76
Max (µg.L^{-1})	729.60	454,384.51	1469.81	57.165	8076.13
Producer characteristics (%)	-	-	56.52 [a]	7.01 [b]	10.53 [b]
Significant consumers (%) **	29.82	38.60	-	40.35	36.84

IF: isoflavones (genistein + daidzein); EQ: equol; ENL: enterolactone. * Corrected by creatinine excretion, ** the thresholds for significant consumers' determination are presented in the text. [a.] Significant EQ producers, urine concentrations > 20 µg.L^{-1}. [b.] Low ENL producers, plasma concentration < 10 ng.mL^{-1} for dietary scores > 10, and urine concentration < 500 µg.L^{-1}.

3.1.2. Enterolactone (ENL)

Results of ENL assay in the 57 urine and plasma samples are summarized in Table 1. The concentrations were corrected by the creatinine excretion to take into account the urine dilution at sampling. The mean, median, minimum and maximum urine concentrations measured were 802.53 ± 1438.85 µg.L^{-1}, 304.73 µg.L^{-1}., 6.76 µg.L^{-1} and 8076.13 ng.L^{-1}, respectively. When dietary scores were associated with urine$_{ENL}$ levels, it was shown that in 6 cases (10.53%), the urine levels were low and did not reflect the dietary intake of sources of ENL-precursors. These 6 volunteers, whose urine$_{ENL}$ concentrations were all below 25 µg.L^{-1} (mean: 15.36 ± 7.45 µg.L^{-1}) all had dietary scores around 9 (mean: 9.28 ± 2.23). Therefore, they were assigned as low-ENL-producers. Finally, 21 volunteers had urine$_{ENL}$ concentrations over 500 µg.L^{-1} and therefore were considered as significantly exposed to ENL-precursors.

Considering plasma$_{ENL}$ levels, they were much lower than the urine$_{ENL}$ concentrations and were generally assayed at a 1/5 and 1/10 dilutions. However thanks to the good sensitivity of the assay (mid-point of the assay curve at 2 ng.mL^{-1} of plasma), ENL was detected in all samples. The mean, median, minimum and maximum plasma concentrations were 17.89 ± 11.71 ng.mL^{-1}, 13.68 ng.mL^{-1}, 1.94 ng.mL^{-1}, and 57.165 ng.mL^{-1}, respectively. Based on ENL plasma levels over 15 ng/mL^{-1}, 43.86% of the volunteers could be considered as consumers significantly exposed to ENL. In addition, 4 volunteers out of 57 had low ENL plasma levels (<10 ng.mL^{-1}) despite a dietary score over 10. They were considered as low-ENL-producers and represented 7.01% of the population.

3.2. Dietary Inquiry and Scores

A summary of the results obtained on the dietary scores is provided in Table 2. Briefly, the urine$_{ENL}$ mean dietary score was 12.11 ± 5.94, and the median was close to the mean and equal to 11.40. The minimum and maximum score were 3.20 and 42.80, respectively. Considering urine$_{ENL}$, 28.07% of the analyzed population was considered high-consumers of ENL-precursors, i.e., with a dietary score over 15. For the plasma$_{ENL}$ dietary scores, the mean, median, minimum and maximum dietary score were 12.38 ± 5.57, 11.73, 3.2 and 38.35 respectively. Therefore, the median was close to the mean. Based on a dietary score over 13, 40.35% of the volunteers were considered exposed to significant amounts of lignans per day.

Table 2. Dietary scores for each phytoestrogen measured.

	Dietary Scores			
	Plasma$_{IFs}$ (n = 26)	**Urine$_{IFs}$ (n = 54)**	**Plasma$_{ENL}$ (n = 57)**	**Urine$_{ENL}$ (n = 57)**
Mean	3.92	3.11	12.38	12.11
Standard deviation	3.43	2.85	5.57	5.94
Median	2.62	2.40	11.73	11.40
Min	0.63	0.40	3.2	3.20
Max	16.56	16.90	38.35	42.80
Significant consumers (%) *	42.11	38.60	40.35	59.65

IFs: isoflavones (genistein + daidzein); ENL: enterolactone precursors. * The thresholds for significant consumers' determination are in the text.

For IFs, the results obtained on the urine$_{IFs}$ dietary scores analysis can be found in Table 2. The mean, median, minimum and maximum scores obtained on 54 volunteers were found to be 3.11 ± 2.85, 2.4, 0.4, and 16.90, respectively. Based on a threshold of 3, 38.60% of the volunteers were considered to have a significant IFs exposure over 48 h. Besides, in Table 2, the 26 plasma$_{IFs}$ dietary scores had a mean, median, minimum and maximum of 3.92 ± 3.43, 2.62, 0.63 and 16.56 respectively. Based on a urine$_{IFs}$ dietary score over 1, 42.11% of the volunteers were considered as significant IFs consumers.

3.3. Correlations between Dietary Scores and Measurements in Biological Fluids

Considering IFs, i.e., GEN + DAI + EQ urine concentrations, the Pearson correlation coefficient r for dietary scores reached 0.921 ($p < 0.001$) (n = 54). Besides, when the plasma concentrations of GEN + DAI + EQ (n = 26) were associated with their corresponding dietary scores, the Pearson correlation coefficient r was found to reach 0.900 ($p < 0.001$). The missing data correspond either to samples where IFs were undetectable or at the limit of detection or to scores which were only based on transformed foodstuff in which the occurrence of soy was not systematic. The correlations are presented in Figure 2A,B for the urine and the plasma concentrations, respectively.

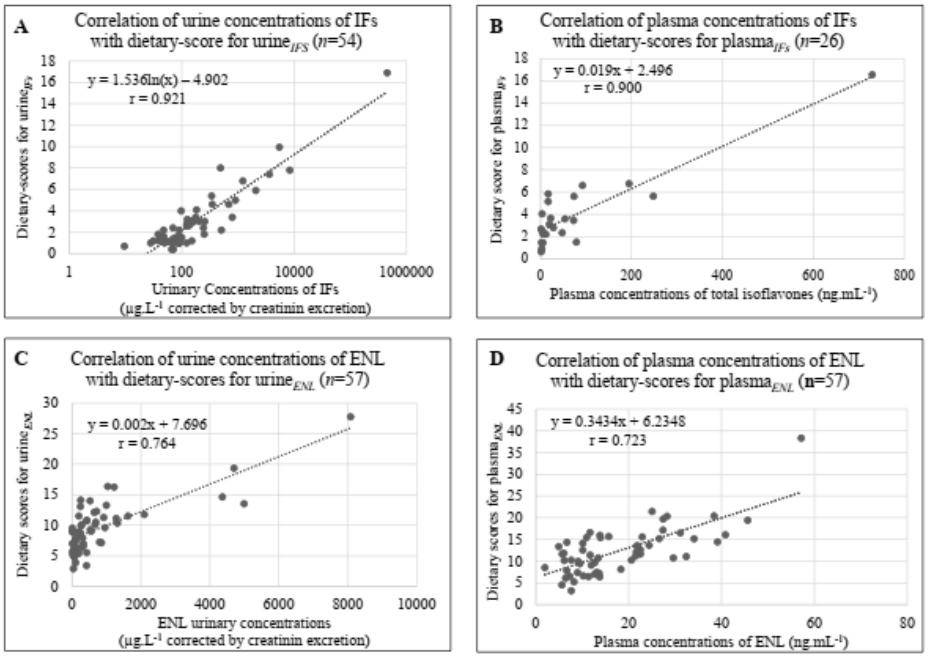


Figure 2. Correlations between IFs (genistein + daidzein + equol) or ENL (enterolactone) in biological fluids and their respective dietary scores.

The correlations between urine concentrations of GEN and DAI considered individually with their respective dietary scores of urine$_{GEN}$ and urine$_{DAI}$ are available in Figure S2A,B. Briefly, r for urine$_{GEN}$ was 0.885 and r for urine$_{DAI}$ was 0.919. In addition,

correlations between plasma concentrations of GEN and DAI considered individually with their respective dietary scores are available in Figures S3A,B. Indeed, r for $plasma_{GEN}$ and $plasma_{DAI}$ were 0.921 and 0.846, respectively.

Considering $urine_{ENL}$ and $plasma_{ENL}$ dietary scores ($n = 57$), the Pearson correlation coefficient r were found to be 0.764 ($p < 0.001$), and 72.3 ($p < 0.001$), respectively. The correlation graphs are presented in Figure 2C,D.

4. Discussion

4.1. Concentrations of PHYTOs in the Body Fluids of French Volunteers

This study is the first in which IFs and ENL have been assayed in biological fluids from French volunteers following their own habitual diet. Even though these data cannot be strictly compared to previous data from a French population, they can be compared between them and with previous data obtained on Western populations. All were obtained after enzymatic hydrolysis and were expressed in aglycone equivalent.

4.1.1. IFs Concentrations

Tables 1 and 2 showed that the standard deviation around the mean and the differences between means and medians were huge. This indicated that there was a large variability in IFs intake and consequently in measured urine and plasma concentrations. Such an observation was already previously made [2,26] by Van der Velpen and co-workers, who considered that body fluids' concentrations of IFs could not be used to assess dietary intakes of IFs. However, the opposite was concluded in this study, possibly because one of the main variability factors, i.e., the pharmacokinetics of the ingested IFs, was taken into consideration. Moreover, only 3 urine samples among 57 were below the quantification limit of IFs. Therefore, this showed that the vast majority (95%) of the volunteers consumed IFs in the past 72 h. Details on GEN and DAI scores can be obtained from Table S2. Besides, one vegan consumer among 57 showed a chronic consumption of soy-based products and presented abnormally high concentrations of IFs in her biological fluids. Indeed, these levels reached 454,384.51 $\mu g.L^{-1}$ for GEN + DAI + EQ, with 231,158.12 $\mu g.L^{-1}$ and 222,912.86 $\mu g.L^{-1}$ for GEN and DAI, respectively. This represents an excretion of 856.14 and 870.75 nM for GEN and DAI respectively. Besides, 38% of the volunteers were significant consumers ($urine_{IFs} > 150$ $\mu g.L^{-1}$ (i.e., > 0.5 µM)). Their median concentration of IFs was 347.6 $\mu g.L^{-1}$ (1.3 µM). This population was recruited between 2018 and early 2022. This period partially covers the COVID-19 pandemic and the diet transition to vegan practices in France [27]. The present assessment of the exposure to IFs was in line with previous studies [2,3]. In addition, the subpopulation exposed to only low doses of IFs exhibited a median urinary excretion of 0.15 µM. Conversely, when plasma levels were examined, the highest concentration values were observed in the "high-soy" vegan consumer. For this specific volunteer, the IFs plasma level was 729.60 $ng.L^{-1}$, with 468.15 (1.73 µM) and 253.17 $ng.L^{-1}$ (0.99 µM) for GEN and DAI, respectively. It should be noted here that upon chronic exposure to soy IFs, GEN is usually higher in plasma than DAI due to a longer elimination half-life [28]. Looking at the significant IFs consumers, defined as persons with IFs plasma levels over 15 $ng.mL^{-1}$, the median concentration of IFs was 66.82 $ng.L^{-1}$, with GEN at 36.38 $ng.mL^{-1}$ (0.135 µM) and DAI at 24.99 $ng.L^{-1}$ (0.097 µM). It might be possible to compare these results to those previously published on Western postmenopausal women. Indeed, in [28], a study performed in the USA on post-menopausal women from the WHI cohort, the IFs plasma levels were less variable and lower than in the present study. Indeed, in [29], the mean concentration of GEN was 12.2 ± 4.3 nM, while it was 225 ± 410 nM here, when the vegan volunteer and only the detectable concentrations were included. Moreover, the mean concentration of DAI in [29] was 6.9 ± 3.6 nM, while it was 180.5 ± 241.4 nM in the present study when the vegan volunteer and only the detectable concentrations were included. When the vegan consumer was not considered, the mean for $plasma_{GEN}$ was 130.8 ± 136.9 nM and the mean $plasma_{DAI}$ was 130.0 ± 125.9 nM. When all data were included in the calculation, even those below the quantification limits and therefore being

null, and without the high-soy-consumer, the mean plasma concentrations of GEN and DAI were in the range of those reported in [29] (plasma$_{GEN}$: 10.09 ± 25.14 nM and plasma$_{DAI}$: 9.51 ± 22.67), showing a huge variability. Finally, the plasma concentrations recorded in this study in soy consumers were between 0.1 and 1 µM, and therefore, in the low range of the IFs concentrations usually tested in vitro on cultured cells.

4.1.2. Comparisons between Urine and Plasma IFs Concentrations

Looking at Tables 1 and 2, it can be observed that the proportion of significant consumers was different when plasma or urine data were considered. This could be explained because: (1) the IFs concentrations were much lower in plasma and not always detectable using the specific ELISAs, and (2) plasma data essentially reflected a consumption of IFs which occurred within the last 24 h, while the urine data reflected an IFs exposure which occurred during the 18 to 48 previous hours. This is why for better accuracy, the dietary scores took into consideration the pharmacokinetic data of the IFs assayed. To sustain this view, the Pearson correlation coefficient *r* between IFs urinary and plasma concentrations based on the detectable plasma concentrations was only of 0.794 ($n = 26$; $p < 0.001$) (Figure S4). Such discrepancies indicated that IFs intakes were mostly occasional and most probably not performed on an everyday basis.

4.1.3. Equol Production

When urine and plasma concentrations were compared, the percentage of EQ-producers was different. Indeed, based on measurements in the urine samples, 56.52% of the volunteers had detectable EQ concentrations, while based on plasma samples, 49.12% were defined as EQ-producers. These proportions were lower than those established in [30] on the same population and based on EQ measurements in hair (61.82%). This discrepancy was most probably explained because hair keeps a trace of previous exposures and metabolic transformations. However, in a previous study [28], performed on 60 postmenopausal women supplemented with 100 mg of IFs per days, the proportion of EQ-producers was 59.1%, using the same assay method. In comparison, in [29], based on 192 American women, the proportion of EQ-producers was 27.5%. Equally, in [31], based on 89 postmenopausal women, the proportion of EQ-producers was 25%. Besides, in the ISOHEART study [32], which was a European multicentric clinical trial, the EQ-producers' proportion was determined in volunteers from Germany, Denmark, the UK, and Italy. Indeed, out of 117 volunteers supplemented with 50 mg of soy IFs, the proportion of EQ producers were 50, 21.4, 25;6 and only 10% in Germany, Denmark, UK and Italy respectively. All urine samples were assayed by the AutoDELFIA assay, which is a fluorescent immunoassay. Finally, it can be said that the proportion of EQ-producers varies between 20% and 60% in Europe depending on the number of assays and the method.

4.1.4. The Case of the Heavy Consumer of Soybean

Heavy consumers of soybean are not rare [33,34], but usually their IFs concentrations in body fluids are not reported. Here, thanks to the 24 and 48 h dietary recalls, it was possible to link soy-food intake to body fluid concentrations. The high-soy-consumer was a healthy young vegan woman of 25 years-old who adopted a vegan diet 4 years ago but started eating soy and soy-based products only 1.5 years ago. She used to cook her own meals at home and used soy juice and soy cream for food preparations as well as chickpea meal to replace eggs in pies and cakes. She also frequently ate tofu and used soy-textured proteins in her recipes. In addition, she had five meals a day, with one collation in the morning and one in the afternoon. Briefly, in this study, she had soy juice (one mug) in the morning before sampling. On the evening preceding sampling, she had a portion of a vegan omelet with soy cream, soy juice, and chickpeas, as well as a tofu portion. At lunch of the previous day, she had a vegan quiche based on chickpea meal, soy juice, and soy cream. During the day, for collation, she had soy juice and cream in a home-made cake and a vegan quiche. Two days before the inquiry and sampling, she had soy juice in the

morning and a soy-based vegan quiche for the evening meal. For collations, she had a portion of vegan omelet based on chickpea meal and soy juice and cream and a home-made cake based on soy juice. Therefore, her soy consumption was continuous and probably partly reflected her dietary habits. As can be deduced from Tables 1 and 2, her plasma and urine IFs concentrations were 729.60 ng.mL^{-1} (about 1.7 µM GEN and 0.98 µM DAI) and 454,384.51 µg.L^{-1} (850 µM of GEN and DAI), respectively. Based on previous data obtained on French volunteers with the same assay technique [28,35,36], it seemed that the urine$_{IFs}$ concentrations were very high when considering the plasma$_{IFs}$ concentrations. This may indicate either a genetic predisposition to readily eliminate IFs from the blood or a kind of protection mechanism that was progressively developed with time and which would enhance the excretion of IFs when the chronic ingestion is high. This putative defense adaptation would result in the reduction of plasma$_{IFs}$ and of their potential physiological effect. Indeed, although this young woman never tried to be pregnant, she claimed having menstrual cycles typically lasting between 30 and 32 days.

4.1.5. ENL Plasma and Urine Concentrations

Looking at Table S1, the intake of ENL-precursors compared to IFs appeared modest [3]. In addition, the plasma$_{ENL}$ levels were generally lower than plasma$_{IFs}$ but the urine excretion was generally higher. The plasma$_{ENL}$ concentrations were in accordance with previous studies [10,37,38]. As in these reports, the interindividual variations in plasma$_{ENL}$ levels were high, reflecting the habitual diet. Moreover, it appeared here, that the Western diet, based on potatoes (fried or boiled), hamburgers, pizzas, kebabs, etc., was associated with low intake of ENL-precursors and low fluids' levels of ENL. On the other hand, diets rich in vegetables, berries, and wholegrain bread were associated with high intake of ENL-precursors and high body fluids' concentrations of ENL. In the population inquired, both dietary profiles were encountered. Moreover, ENL appeared later than IFs in body fluids because of its colic production by the gut flora, while IFs were mainly absorbed in the duodenum and jejunum. This allowed a progressive accumulation of ENL in the urine compartment, and according to [7], ENL can be measured in human urine samples up to two days after a unique ingestion of ENL-precursors. However, plasma$_{ENL}$ concentrations (nM range) were much lower than those usually tested in vitro to decipher the mechanisms of action of ENL. This rose the question of the relevance of these in vitro data since the cell pathways potentially triggered by ENL in vitro could be different when the concentrations vary.

4.2. Discussion on Correlations

4.2.1. Calculations of Dietary Scores

As can be seen in Table S2A–D, the dietary scores considered the level of food intake pondered by a coefficient based on the pharmacokinetic profiles of GEN, DAI, and ENL. For GEN and DAI, the data were from [22], and for ENL they were from [7]. Indeed, all samples were collected in the morning between 9:30 am and 2:30 pm. At that time, PHYTOs that were ingested at breakfast could not be seen in the IFs or ENL urine samples and were pondered by a factor of 0 in the urine dietary scores. Conversely, if foods containing IFs were ingested the previous evening, they were close to their maximum levels (urinary T*max* between 14 and 18 h) in the following morning urine samples. Thus, the ponderation of the IFs intake was the highest for GEN and DAI urine dietary scores calculation. Since the IFs urine levels were known to progressively decrease after 12 h, the urine$_{IFs}$ dietary scores decreased with time from sampling. For plasma, the IFs C_{max} were known to occur around 8 h after intake. Therefore, C_{max} was not reached after a soy intake at breakfast and the plasma$_{IFs}$ concentrations were already declining at sampling time after an IFs intake the previous evening. Thereafter, the plasma concentrations were known to progressively decrease, and thus the plasma dietary scores ponderations followed the same tendencies [22]. Finally, for ENL, the C_{max} occurred in the urine between 24 and 48 h after the ENL-precursors intake in women [7]. Thus, the meals eaten 36 to 48 h before

sampling were the main source of $urine_{ENL}$ concentrations and were attributed the highest ENL-urinary scores. Consequently, the highest ENL dietary scores for urine were seen for intakes occurring two days before sampling. If ENL-precursors were ingested on the previous evening, the $urine_{ENL}$ were not at their maximum levels and the urine scores for dietary intake were lower. Similarly, C_{max} for $plasma_{ENL}$ was found to occur 18 to 24 h after intake in women [7,37,38]. Therefore, when ENL-precursors were ingested at lunch or dinner, before sampling, the ENL was not entirely passed in the plasma and the dietary scores applied to food intake were lower. It should be noted that, to our knowledge, this study is the first in which pharmacokinetic parameters were considered when addressing the correlations reliability between vegetable biomarkers in plasma and urine and dietary inquiries. As far as we know, such calculation was not performed in [39,40]. Producing such scores allowed to strongly link biomarkers in biological fluids and dietary intakes recorded over the last 48 h. Therefore, contrary to what was concluded by other authors, IFs and ENL in biological fluids could be advantageously and accurately used as biomarkers for assessing vegetables and legumes' intakes.

4.2.2. Comparisons of Correlation between Dietary Scores and IF in Urine or in Plasma

Like in [27,39,41], $urine_{IFs}$ concentrations were found to be 50 to 100 times higher than $plasma_{IFs}$. Therefore, technically, it could be easier to use urine concentrations as biomarkers of IFs exposure. Additionally, urine sampling is not invasive, even if the conservation of urine samples may be more complicated than for plasma ones, due to potential oxidation at collection. Here, IFs plasma levels were sometimes below the detection limits, and this is why only 26 values could be used for the correlation with diet data. The sample size allowed the calculation of a Pearson coefficient of correlation between dietary scores and $plasma_{IFs}$ levels. Additionally, Figure S3 showed that the Pearson correlation coefficient between $plasma_{DAI}$ and its corresponding dietary score was lower than for $plasma_{GEN}$. Briefly, for $plasma_{DAI}$, r was 0.846 ($p < 0.001$), vs. 0.921 ($p < 0.001$) for $plasma_{GEN}$. This may be because a part of DAI was transformed into EQ in some volunteers, decreasing the link between dietary records and $plasma_{DAI}$ levels. Another explanation, may be that the DAI plasma residence time is notably lower than that of GEN, and thus it varied more rapidly. Consequently, the scores used here, might have been of lower accuracy for DAI than for GEN due to interindividual variations.

4.2.3. IFs and ENL as Dietary Biomarkers

It appeared that IFs and ENL can both be used as biomarkers of vegetables' intakes (see Table S1). However, IFs were much more specific for a defined category of food, i.e., pulses, while ENL-precursors were present mainly in grains and grain-enriched foods, carob, berries, cabbages, or squashes (Table S1). Additionally, in the French diet, the main source of IFs was soybean, which appeared in soy-based foodstuffs and as "hidden soy" in transformed foodstuffs [2]. Therefore, IFs intake could also be a biomarker of transformed food consumption.

Considering the different occurrences of IFs and ENL-precursors in foodstuffs, two different dietary scales were constructed to record the two types of PHYTOs, (see Table S2A–D). Finally, it was not surprising to show that the Pearson correlation coefficient r between IFs and ENL urine concentrations was only 0.310 ($p < 0.001$) ($n = 56$), while the Pearson correlation coefficient r between $urine_{IFs}$ and $urine_{ENL}$ dietary scores was only 0.285 ($p < 0.001$) ($n = 56$) (Figure S5A,B). These scores were calculated excluding the highest consumer of IFs. Similarly, the correlation between $plasma_{IFs}$ and $plasma_{ENL}$ levels was low, as was the correlation between their respective dietary scores. These correlations were not sufficient for prediction (data not shown), and it could be deduced from these figures that, in the French diet, IFs intake was not significantly correlated with ENL-precursors' intake.

4.2.4. Comparisons with Other Published Data

Table 3 provides an overview of the correlations found in the literature between dietary inquiries and PHYTOs concentrations in biological fluids. The studies were performed either on spot urine samples, as in this study, or on urine samples collected at home over 12 or over 24 h. It should be expected that samples collected over a defined period would be more accurate, providing that no collection was missed during the period. This seemed to be confirmed comparing the correlation coefficients obtained by Grace et al. [39] for IFs spot urine levels, with those obtained by other authors that collected urine samples over a longer period. However, the sample collection is not the only source of uncertainty, and the dietary records could also be questionable. Indeed, for the previous correlations between dietary data and urine concentrations of IFs, details on the food items considered to assess the IFs exposure were not clearly disclosed. It might be hypothesized that those IFs intake records were not exhaustive because the part of "hidden soy" was usually underestimated [29]. This led to a poor correlation coefficient, lower than those observed in the present study. In addition, these correlation coefficients were quite variable from one study to another.

Table 3. Comparisons with other studies connecting PHYTOs in body fluids and estimated dietary intake of PHYTOs.

Subjects	Nature of Samples	Biomarkers	Dietary Data ($mg.day^{-1}$)	Correlation	References
80 British volunteers	Plasma	GEN DAI	7-day food diaries	GEN: $r = 0.80$; $p < 0.001$ DAI: $r = 0,45$; $p < 0.001$	[42]
77 volunteers	Plasma	GEN DAI	FFQ	GEN: $r = 0,53$; $p < 0.001$ DAI: $r = 0,45$; $p < 0.001$	[29]
14 adults (14% men)	Plasma	IFs	24 h food record: (11.0) 24 h recall: (12.3)	IFs: $r = 0.92$; $p < 0.001$	[41]
333 volunteers	Serum	IFs	7-day food diaries	IFs $r = 0.31$; $p < 0.001$	[39]
203 male volunteers	Serum	IFs	FFQ	IFs: $r = 0.27$; $p < 0.001$	[43]
26 French women	Spot plasma	IFs	24 h and 48 h dietary recall	IFs: $r = 0.900$; $p < 0.001$ GEN: $r = 0.921$; $p < 0.001$ DAI: $r = 0.846$; $p < 0.001$	Present study
51 Japanese women 18 Caucasian women	24 h urine	GEN DAI	48 h dietary recall	GEN: $r = 0.54$; $p < 0.001$ DAI: $r = 0.58$; $p < 0.001$	[44]
360 women	2 overnight urines	IFs	Twice, 24 h recall DAI (µg): (5.0–6.4) GEN (µg): (7.3–9.3)	IFs: $r = 0.52$; $p = 0.001$ FFQ: $r = 0.29$; $p < 0.01$	[45]
284 volunteers	Spot urine	IFs	7-day food diaries	IFs $r = 0.27$; $p < 0.001$	[39]
14 adults (14% men)	24 h urine	IFs	24 h food record: (11.0) 24 h recall: (12.3)	IFs: $r = 0.97$; $p < 0.001$	[41]
26 premenopausal Canadian women	24 h urine	IFs	Habitual record 24 h recall	IFs: $r = 0.64$, $p < 0.001$ IFs: $r = 0.54$, $p = 0.004$	[40]
24 pubertal girls	12 h urine	IFs	3-day 24 h recall ISO: 3.0–13.3	IFs: $r = 0.72$; $p < 0.001$	[46]
256 premenopausal women	12 h urine	IFs	FFQ Low: 0.1–2.3 High: 49.8–74.6	IFs: $r = 0.51$; $p < 0.001$	[47]
100 healthy women	12 h urine	IFs	24 h recall	IFs: $r = 0.460$; $p < 0.001$	[48]
57 French women	Spot urine	IFs	24 and 48 h dietary recall	IFs: $r = 0.921$; $p < 0.001$ GEN: $r = 0.885$; $p < 0.001$ DAI: $r = 0.919$; $p < 0.001$	Present study
284 British women of the EPIC-Norfolk study	Spot urine	ENL	Fiber intake over 7-day recall	ENL: $r = 0.29$; $p < 0.001$	[39]
26 premenopausal Canadian women	24 h urine	ENL	Habitual record 24 h recall	ENL: $r = 0.46$, $p = 0.02$ ENL: $r = 0.40$, $p = 0.05$	[40]
100 apparently healthy Mexican women	12 h urine	ENL	24 h recall	ENL: $r = 0.067$; $p = 0.580$	[48]
57 healthy premenopausal French women	Spot urine Spot plasma	ENL	24 and 48 h dietary recalls	ENL: $r = 0.764$; $p < 0.001$ ENL: $r = 0.723$; $p < 0.001$	Present study

Considering plasma analyses, Table 3 only lists spot samples, although the dietary records could be different. Indeed, some authors performed dietary recall over 24 or 48 h or applied food frequency questionnaires (FFQ) to assess diet habits. Again, such data collections could be more or less accurate because they relied on claims from the volunteers and because the foodstuffs containing the substance of interest may not be known as accurately as they should be. Finally, if only Pearson's correlation coefficients were analyzed, i.e., excluding the studies [41,42], the correlations obtained here were greater than those previously published, and they characterized a significant correlation. This

shows that considering the pharmacokinetic parameters is important to establish relevant dietary scores.

As seen in Table 3, the correlation coefficients between dietary records and urine$_{ENL}$ levels obtained by previous authors were not as good as those recorded here. Analyzing the previous studies, it was seen that sometimes the urine samplings were not appropriate considering the pharmacokinetic parameters of ENL. For instance, in [48], the urine collection was only organized over the previous 12 h and could not reflect the last 24 h intakes due to the urine T_{max} of ENL. Additionally, in [39], the main uncertainty could come from the approximation made on the lignan intakes. Basically, the authors considered that dietary fibers well-reflected the ENL-precursors' intakes. They did not consider specific distributions of ENL-precursors in certain fruits, vegetables, or grain categories, as was performed here (see Table S2C,D, based on Table S1). This may explain why the present study appeared to yield more accurate correlation coefficients. Again, contrary to the present studies, the previous analyses did not provide details on the foodstuffs recorded to assess ENL-precursor intakes. Therefore, it may be difficult to conclude on the relevancy of these dietary records.

4.3. Limits of the Study

4.3.1. Phytoestrogens Analyzed

All PHYTOs were analyzed in their aglycone forms after enzymatic digestion by β-glucuronidase aryl-sulfatase, which hydrolyzed the glucurono- and sulfo-conjugates before liquid–liquid extraction by ethyl-acetate. Glycitein was not assayed in this study as it is known to be present only at a 10% ratio in soy-based products and it was usually undetectable in the plasma samples of volunteers. Moreover, its estrogenic activity is the lowest of the IFs series [49]. Additionally, ENL is not the only estrogenic metabolite of SDG, SECO, MAT, LAR, PINO, SYR, and MEDIO. Besides ENL, the other metabolite is enterodiol (Figure 1D), which was also shown to be slightly estrogenic but for which plasma and urine concentrations were found to be lower than ENL [7]. The lack of a specific ELISA for enterodiol and the specificity of the polyclonal antibody raised against ENL which has no significant recognition of enterodiol, explain why it was not assayed in this study. The ELISA for ENL was sensitive enough to assay ENL in plasma and was 10 times more sensitive than the IFs-ELISAs. This is most probably because it was a heterologous ELISA, using an immobilized ligand which was different from the hapten on which the specific polyclonal antibody was raised. Such a sensitivity was associated with a great specificity since the polyclonal antibody did not significantly recognize enterodiol nor other ENL-precursors [23]. However, urine samples were much easier to collect, and showing higher ENL concentrations when compared to plasmas ([7] and the present study), they could be better biomarkers of vegetable intakes.

4.3.2. Accuracy of the Dietary Scores

One of the main advantages of this study is that it took into consideration the pharmacokinetic parameters, i.e., C_{max}, T_{max}, and $T_{1/2}$, for each biomarker and included them in the dietary score. In addition, great attention was paid to the selection of the sources of biomarkers in the diet, to analyze the dietary records in the most relevant way. Indeed, the sources of IFs were derived from a previous exposure analysis of the French population and based on more than 12,000 foodstuffs labels assessment [2], while the selection of ENL-precursors was based on data collected from previous analytical studies by other authors and mentioned in Table S1. Contrary to some studies, six lignans were considered in this study, having been shown to be ENL-precursors. Indeed, besides SECO, which is the most popular ENL-precursor in countries from Northern Europe, this study also examined vegetables and grains containing significant concentrations (>100 µg/100 g of fresh weight) of other ENL-precursors. Likewise, flaxseeds are not usually consumed in France; thus, it was important to also consider foodstuffs containing MAT, LAR, PINO, MEDIO, and SYR. In doing so, it was possible to associate certain vegetable or fruit groups

with higher ENL-intake coefficients. However, the correlation obtained between $urine_{ENL}$ concentrations and the $urine_{ENL}$ dietary score was only 0.764, while it reached 0.921 for IFs. This may be due to a lack of accurate data on the ENL-precursors' concentrations in all foodstuffs and especially to the lack of data on the impact of food transformation (cooking, grilling, boiling, etc.) on ENL-precursors' resulting concentrations. Plasma concentrations of ENL were detectable in all samples, but a large variability was observed that could be related mainly to cabbages, berries, and wholegrain bread. As for the $urine_{ENL}$ correlation, the $plasma_{ENL}$ correlation *r* was lower than that recorded for IFs, being only 0.723.

Besides, $plasma_{IFs}$ concentrations were sometimes too low to be detected, and thus the correlation was not calculated using all the samples. As mentioned previously, it was shown here that IFs exposure is highly variable in the French population [2,3]. Additionally, the non-soy-eater category represented about half of the population in [2,3], and this proportion was represented here with 26 out of 57 volunteers having detectable $plasma_{IFS}$ concentrations.

4.3.3. Dietary Score Uncertainties

Regarding IFs intake, the larger uncertainties were due to the consumption of ultra-transformed meals. Indeed, based on previous data [2], four categories of transformed foodstuffs were identified: "Transformed++", "Transformed+", "Transformed", and "Transformed–" (see Table S2A,B). The first subcategory contained transformed foodstuffs known to very frequently contain a high proportion of IFs, such as soy-containing burgers, meatballs, or stuffed vegetables, especially when consumed in canteens. On the other hand, the "Transformed–" category contained foodstuffs that definitely did not contain soy, such as dairy products, for instance. The two intermediate categories were defined according to their potential IFs content and based on the occurrence frequency of soy in such food groups. For instance, nuggets are known to contain soy only occasionally, but this frequency was higher in canteens. These ultra-transformed foodstuffs were the highest source of uncertainty when a dietary score was calculated for IFs.

4.3.4. Uncertainties on Correlations

Another source of uncertainty came from the case of the vegan woman who had soy and IFs several times a day before sampling, and who exhibited abnormally high levels of IFs, GEN and DAI in her plasma and urine samples. The data collected on this person largely increased the variation around the mean of all concentrations or scores for IFs, GEN, and DAI. Nevertheless, because there are only few data published so far on biological samples collected on "high-soy-consumers", it was decided to maintain these data for the analysis. As expected, considering this case impacted the correlation coefficients obtained. Indeed, when the $urine_{IFs}$ concentration values and $urine_{IFs}$ dietary scores of this volunteer were removed from the calculation of the Pearson correlation coefficient, *r* decreased from 0.921 ($p < 0.001$) to 0.864 ($p < 0.001$). Anyway, the latter coefficient was still significant and higher than those available in the literature. For plasma, based on 25 values only (without the "high-soy-consumer"), the correlation coefficient *r* between plasma concentrations and $plasma_{IFs}$ dietary scores was 0.653 ($p < 0.001$), vs. 0.900 ($p < 0.001$) when the "high-soy-consumer" was included. These changes in correlation coefficient and dispersion may be partly due to the low values recorded in some samples and to the presence of equol producers among the volunteers included so far.

5. Conclusions

Here, specific ELISAs for GEN, DAI, EQ, and ENL were used on plasma and urine samples collected at the hospital on premenopausal women who were either healthy volunteers or women with an autoimmune disease. No dietary intervention was performed, and this analysis reflects the current French diet. An open questionnaire inquiring the previous consumptions over the last 24 h and 48 h was proposed to the volunteers. It was shown that, when the pharmacokinetic parameters of the biomarkers were considered, quite

accurate dietary scores could be calculated providing that the food intake was reasonably assessed. Knowing the IFs and the ENL-precursors' content of foodstuffs was a prerequisite for such an approach, and accurate databases have been used to address this question. Indeed, knowing the global dietary habits of the French population helped to build the dietary questionnaires and scores, and those retained for French consumers may not be adapted to other populations. IFs concentrations in plasma and urine of a vegan woman overconsuming soy were reported here. The results showed a major excretion of IFs and addressed the question of a potential physiological protection phenomenon leading to lowering IFs plasma concentrations. This study also showed that nearly 60% of the French volunteers had significant EQ concentrations in their urine samples. Urine samples seemed to be good providers of pulses and ENL-precursors' biomarkers. They can help to determine PHYTOs intake and to predict their health effects. Since natural estrogenic compounds present or deriving from food may be active in humans, they should be considered, while accurately counting endocrine disruptors in the environment. When estrogen-dependent diseases are encountered, the exposure to PHYTOs should be taken into account for more accurate health recommendations.

Supplementary Materials: Supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/nu15040967/s1, Table S1: Sources of ENL-precursors according to previous analytical studies. Table S2: Dietary scales used to calculate dietary scores for DAI, GEN, and ENL. Table S2A: Scales for DAI and GEN associated with urine determination. Table S2B: Scales for DAI and GEN associated with plasma determination. Table S2C: Scales for ENL associated with urine determination. Table S2D: Scales for ENL associated with plasma determination. Figure S1: Consort diagram of the recruitment of volunteers at the time of the study and samples collected and inquiry completed. Figure S2: Correlation between urinary DAI and GEN and their respective dietary scores. Figure S2A: Correlation between urinary DAI concentrations and the corresponding dietary scores. Figure S2B: Correlation between urinary GEN concentrations and the corresponding dietary scores. Figure S3: Correlation between plasma DAI and GEN and their respective dietary scores. Figure S3A: Correlation between plasma DAI concentrations and the corresponding dietary scores. Figure S3B: Correlation between plasma GEN concentrations and the corresponding dietary scores. Figure S4: Correlation between urine and plasma IFs concentrations. Figure S5: Correlation between ENL and IFs data. Figure S5A: Correlation based on PHYTOs urine concentrations. Figure S5B: Correlation based on PHYTOs dietary scores. Document SD1: Open questionnaire proposed for the last 24 h recall. Document SD2: Example of the informed consent form (translated in English). References [50–60] are cited in the supplementary materials.

Author Contributions: Conceptualization, C.B.-P., J.-F.V. and I.R.; methodology, C.B.-P., J.-F.V. and I.R.; validation, C.B.-P., I.R. and I.P.; formal analysis, C.B.-P. and S.B.; resources, I.P.; data curation, S.B. and C.B.-P.; writing—original draft preparation, C.B.-P. and S.B.; writing—review and editing, I.R., I.P. and J.-F.V.; supervision, J.-F.V. and C.B.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was approved by the local ethic comity under the name ISOLED and the number 2017T2-29. The protocol was approved in august 2017.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. The local ethical committee approved this study under the registration number: CHUBX2017/08, and the name ISOLED.

Data Availability Statement: Raw data are available upon request to the corresponding author.

Acknowledgments: The nurses of the Intern Service of the Haut-Levêque Hospital are thanked here for their help in the volunteers' management and for the health and dietary inquiries at inclusion. The CRB BBS is also acknowledged for the accurate management of the biological samples.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bennetau-Pelissero, C. Natural Estrogenic Substances, Origins, and Effects. In *Bioactive Molecules in Food, Reference Series in Phytochemistry*; Mérillon, J.-M., Ramawat, K.G., Eds.; Springer International Publishing AG: Berlin/Heidelberg, Germany, 2018. [CrossRef]
2. Lee, A.; Beaubernard, L.; Lamothe, V.; Bennetau-Pelissero, C. New Evaluation of Isoflavone Exposure in the French Population. *Nutrients* **2019**, *11*, 2308. [CrossRef]
3. Lee, A.; Bensaada, S.; Lamothe, V.; Lacoste, M.; Bennetau-Pelissero, C. Endocrine disruptors on and in fruits and vegetables: Estimation of the potential exposure of the French population. *Food Chem.* **2022**, *373*, 131513. [CrossRef] [PubMed]
4. Smeds, A.I.; Eklund, P.C.; Sjöholm, R.E.; Willför, S.M.; Nishibe, S.; Deyama, T.; Holmbom, B.R. Quantification of a Broad Spectrum of Lignans in Cereals, Oilseeds, and Nuts. *J. Agric. Food Chem.* **2007**, *55*, 1337–1346. [CrossRef] [PubMed]
5. Lagkouvardos, I.; Kläring, K.; Heinzmann, S.S.; Platz, S.; Scholz, B.; Engel, K.-H.; Schmitt-Kopplin, P.; Haller, D.; Rohn, S.; Skurk, T.; et al. Gut metabolites and bacterial community networks during a pilot intervention study with flaxseeds in healthy adult men. *Mol. Nutr. Food Res.* **2015**, *59*, 1614–1628. [CrossRef] [PubMed]
6. Moreno-Franco, B.; García-González, A.; Montero-Bravo, A.M.; Iglesias-Gutiérrez, E.; Úbeda, N.; Maroto-Núñez, L.; Adlercreutz, H.; Peñalvo, J.L. Dietary Alkylresorcinols and Lignans in the Spanish Diet: Development of the Alignia Database. *J. Agric. Food Chem.* **2011**, *59*, 9827–9834. [CrossRef] [PubMed]
7. Kuijsten, A.; Arts, I.C.W.; Vree, T.B.; Hollman, P.C.H. Pharmacokinetics of Enterolignans in Healthy Men and Women Consuming a Single Dose of Secoisolariciresinol Diglucoside. *J. Nutr.* **2005**, *135*, 795–801. [CrossRef]
8. Bennetau-Pelissero, C. Risks and benefits of phytoestrogens: Where are we now? *Curr. Opin. Clin. Nutr. Metab. Care* **2016**, *19*, 477–483. [CrossRef]
9. Mueller, S.O.; Simon, S.; Chae, K.; Metzler, M.; Korach, K.S. Phytoestrogens and Their Human Metabolites Show Distinct Agonistic and Antagonistic Properties on Estrogen Receptor α (ERα) and ERβ in Human Cells. *Toxicol. Sci.* **2004**, *80*, 14–25. [CrossRef]
10. Travis, R.C.; A Spencer, E.; E Allen, N.; Appleby, P.N.; Roddam, A.W.; Overvad, K.; Johnsen, N.F.; Olsen, A.; Kaaks, R.; Linseisen, J.; et al. Plasma phyto-oestrogens and prostate cancer in the European Prospective Investigation into Cancer and Nutrition. *Br. J. Cancer* **2009**, *100*, 1817–1823. [CrossRef]
11. Carreau, C.; Flouriot, G.; Bennetau-Pelissero, C.; Potier, M. Enterodiol and enterolactone, two major diet-derived polyphenol metabolites have different impact on ERα transcriptional activation in human breast cancer cells. *J. Steroid Biochem. Mol. Biol.* **2008**, *110*, 176–185. [CrossRef]
12. Power, K.A.; Saarinen, N.M.; Chen, J.M.; Thompson, L.U. Mammalian lignans enterolactone and enterodiol, alone and in combination with the isoflavone genistein, do not promote the growth of MCF-7 xenografts in ovariectomized athymic nude mice. *Int. J. Cancer* **2006**, *118*, 1316–1320. [CrossRef] [PubMed]
13. Touillaud, M.S.; Thiébaut, A.C.M.; Fournier, A.; Niravong, M.; Boutron-Ruault, M.-C.; Clavel-Chapelon, F. Dietary Lignan Intake and Postmenopausal Breast Cancer Risk by Estrogen and Progesterone Receptor Status. *J. Natl. Cancer Inst.* **2007**, *99*, 475–486. [CrossRef] [PubMed]
14. Carreau, C.; Flouriot, G.; Bennetau-Pelissero, C.; Potier, M. Respective contribution exerted by AF-1 and AF-2 transactivation functions in estrogen receptor α induced transcriptional activity by isoflavones and equol: Consequence on breast cancer cell proliferation. *Mol. Nutr. Food Res.* **2009**, *53*, 652–658. [CrossRef] [PubMed]
15. Uifălean, A.; Schneider, S.; Ionescu, C.; Lalk, M.; Iuga, C.A. Soy Isoflavones and Breast Cancer Cell Lines: Molecular Mechanisms and Future Perspectives. *Molecules* **2015**, *21*, 13. [CrossRef] [PubMed]
16. Allred, C.D.; Allred, K.F.; Ju, Y.H.; Virant, S.M.; Helferich, W.G. Soy diets containing varying amounts of genistein stimulate growth of estrogen-dependent (MCF-7) tumors in a dose-dependent manner. *Cancer Res.* **2001**, *61*, 5045–5050. [PubMed]
17. Shike, M.; Doane, A.S.; Russo, L.; Cabal, R.; Reis-Filo, J.; Gerald, W.; Cody, H.; Khanin, R.; Bromberg, J.; Norton, L. The Effects of Soy Supplementation on Gene Expression in Breast Cancer: A Randomized Placebo-Controlled Study. *J. Natl. Cancer Inst.* **2014**, *106*, dju189. [CrossRef]
18. Messina, M. Impact of Soy Foods on the Development of Breast Cancer and the Prognosis of Breast Cancer Patients. *Complement. Med. Res.* **2016**, *23*, 75–80. [CrossRef]
19. Jiang, R.; Botma, A.; Rudolph, A.; Hüsing, A.; Chang-Claude, J. Phyto-oestrogens and colorectal cancer risk: A systematic review and dose–response meta-analysis of observational studies. *Br. J. Nutr.* **2016**, *116*, 2115–2128. [CrossRef]
20. Nusbaum, J.S.; Mirza, I.; Shum, J.; Freilich, R.; Cohen, R.E.; Pillinger, M.; Izmirly, P.; Buyon, J. Sex Differences in Systemic Lupus Erythematosus: Epidemiology, Clinical Considerations, and Disease Pathogenesis. *Mayo Clin. Proc.* **2020**, *95*, 384–394. [CrossRef]
21. Barbhaiya, M.; Costenbader, K.H. Environmental exposures and the development of systemic lupus erythematosus. *Curr. Opin. Rheumatol.* **2016**, *28*, 497–505. [CrossRef]
22. Shinkaruk, S.; Durand, M.; Lamothe, V.; Carpaye, A.; Martinet, A.; Chantre, P.; Vergne, S.; Nogues, X.; Moore, N.; Bennetau-Pelissero, C. Bioavailability of glycitein relatively to other soy isoflavones in healthy young Caucasian men. *Food Chem.* **2012**, *135*, 1104–1111. [CrossRef] [PubMed]
23. Shinkaruk, S.; Pinot, E.; Lamothe, V.; Schmitter, J.-M.; Baguenard, L.; Bennetau, B.; Pelissero, C.B. Design and validation of a novel immunological test for enterolactone. *Talanta* **2014**, *119*, 116–124. [CrossRef] [PubMed]

24. Le Houérou, C.; Bennetau-Pelissero, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Syntheses of Novel Hapten–Protein Conjugates for Production of Highly Specific Antibodies to Formononetin, Daidzein and Genistein. *Tetrahedron* **2000**, *56*, 295–301. [CrossRef]
25. Bennetau-Pelissero, C.; Le Houérou, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Synthesis of Haptens and Conjugates for ELISAs of Phytoestrogens. Development of the Immunological Tests. *J. Agric. Food Chem.* **2000**, *48*, 305–311. [CrossRef] [PubMed]
26. van der Velpen, V.; Hollman, P.C.; van Nielen, M.; Schouten, E.G.; Mensink, M.; Veer, P.V.; Geelen, A. Large inter-individual variation in isoflavone plasma concentration limits use of isoflavone intake data for risk assessment. *Eur. J. Clin. Nutr.* **2014**, *68*, 1141–1147. [CrossRef] [PubMed]
27. Salomé, M.; Huneau, J.F.; Le Baron, C.; Kesse-Guyot, E.; Fouillet, H.; Mariotti, F. Substituting Meat or Dairy Products with Plant-Based Substitutes Has Small and Heterogeneous Effects on Diet Quality and Nutrient Security: A Simulation Study in French Adults (INCA3). *J. Nutr.* **2021**, *151*, 2435–2445. [CrossRef] [PubMed]
28. Mathey, J.; Lamothe, V.; Coxam, V.; Potier, M.; Sauvant, P.; Pelissero, C.B. Concentrations of isoflavones in plasma and urine of post-menopausal women chronically ingesting high quantities of soy isoflavones. *J. Pharm. Biomed. Anal.* **2006**, *41*, 957–965. [CrossRef]
29. Frankenfeld, C.L.; McTiernan, A.; Tworoger, S.S.; Atkinson, C.; Thomas, W.K.; Stanczyk, F.Z.; Marcovina, S.M.; Weigle, D.S.; Weiss, N.S.; Holt, V.L.; et al. Serum steroid hormones, sex hormone-binding globulin concentrations, and urinary hydroxylated estrogen metabolites in post-menopausal women in relation to daidzein-metabolizing phenotypes. *J. Steroid Biochem. Mol. Biol.* **2004**, *88*, 399–408. [CrossRef]
30. Bensaada, S.; Raymond, I.; Breton, M.; Pellegrin, I.; Viallard, J.-F.; Bennetau-Pelissero, C. Development of an Assay for Soy Isoflavones in Women's Hair. *Nutrients* **2022**, *14*, 3619. [CrossRef]
31. Atkinson, C.; Newton, K.M.; Stanczyk, F.Z.; Westerlind, K.C.; Li, L.; Lampe, J.W. Daidzein-metabolizing phenotypes in relation to serum hormones and sex hormone binding globulin, and urinary estrogen metabolites in premenopausal women in the United States. *Cancer Causes Control* **2008**, *19*, 1085–1093. [CrossRef]
32. Talbot, D.C.; Ogborne, R.M.; Dadd, T.; Adlercreutz, H.; Barnard, G.; Bugel, S.; Kohen, F.; Marlin, S.; Piron, J.; Cassidy, A.; et al. Monoclonal Antibody-Based Time-Resolved Fluorescence Immunoassays for Daidzein, Genistein, and Equol in Blood and Urine: Application to the Isoheart Intervention Study. *Clin. Chem.* **2007**, *53*, 748–756. [CrossRef] [PubMed]
33. Chandrareddy, A.; Muneyyirci-Delale, O.; McFarlane, S.I.; Murad, O.M. Adverse effects of phytoestrogens on reproductive health: A report of three cases. *Complement. Ther. Clin. Pract.* **2008**, *14*, 132–135. [CrossRef] [PubMed]
34. Imai, H.; Nishikawa, H.; Suzuki, A.; Kodama, E.; Iida, T.; Mikura, K.; Hashizume, M.; Kigawa, Y.; Tadokoro, R.; Sugisawa, C.; et al. Secondary Hypogonadism due to Excessive Ingestion of Isoflavone in a Man. *Intern. Med.* **2022**, *61*, 2899–2903. [CrossRef] [PubMed]
35. Vergne, S.; Titier, K.; Bernard, V.; Asselineau, J.; Durand, M.; Lamothe, V.; Potier, M.; Perez, P.; Demotes-Mainard, J.; Chantre, P.; et al. Bioavailability and urinary excretion of isoflavones in humans: Effects of soy-based supplements formulation and equol production. *J. Pharm. Biomed. Anal.* **2007**, *43*, 1488–1494. [CrossRef]
36. Vergne, S.; Bennetau-Pelissero, C.; Lamothe, V.; Chantre, P.; Potier, M.; Asselineau, J.; Perez, P.; Durand, M.; Moore, N.; Sauvant, P. Higher bioavailability of isoflavones after a single ingestion of a soya-based supplement than a soya-based food in young healthy males. *Br. J. Nutr.* **2008**, *99*, 333–344. [CrossRef]
37. Mazur, W.M.; Uehara, M.; Wähälä, K.; Adlercreutz, H. Phyto-oestrogen content of berries, and plasma concentrations and urinary excretion of enterolactone after a single strawberry-meal in human subjects. *Br. J. Nutr.* **2000**, *83*, 381–387. [CrossRef]
38. Hålldin, E.; Eriksen, A.K.; Brunius, C.; da Silva, A.B.; Bronze, M.; Hanhineva, K.; Aura, A.; Landberg, R. Factors Explaining Interpersonal Variation in Plasma Enterolactone Concentrations in Humans. *Mol. Nutr. Food Res.* **2019**, *63*, e1801159. [CrossRef]
39. Grace, P.B.; Taylor, J.I.; Low, Y.-L.; Luben, R.N.; Mulligan, A.A.; Botting, N.P.; Dowsett, M.; Welch, A.A.; Khaw, K.-T.; Wareham, N.J.; et al. Phytoestrogen Concentrations in Serum and Spot Urine as Biomarkers for Dietary Phytoestrogen Intake and Their Relation to Breast Cancer Risk in European Prospective Investigation of Cancer and Nutrition-Norfolk. *Cancer Epidemiol. Biomark. Prev.* **2004**, *13*, 698–708. [CrossRef]
40. French, M.R.; Thompson, L.U.; Hawker, G.A. Validation of a phytoestrogen food frequency questionnaire with urinary concentrations of isoflavones and lignan metabolites in premenopausal women. *J. Am. Coll. Nutr.* **2007**, *26*, 76–82. [CrossRef]
41. Ritchie, M.R.; Morton, M.S.; Deighton, N.; Blake, A.; Cummings, J.H. Plasma and urinary phyto-oestrogens as biomarkers of intake: Validation by duplicate diet analysis. *Br. J. Nutr.* **2004**, *91*, 447–457. [CrossRef]
42. Verkasalo, P.K.; Appleby, P.N.; Allen, N.E.; Davey, G.; Adlercreutz, H.; Key, T.J. Soya intake and plasma concentrations of daidzein and genistein: Validity of dietary assessment among eighty British women (Oxford arm of the European Prospective Investigation into Cancer and Nutrition). *Br. J. Nutr.* **2001**, *86*, 415–421. [CrossRef] [PubMed]
43. Heald, C.L.; Bolton-Smith, C.; Ritchie, M.R.; Morton, M.S.; E Alexander, F. Phyto-oestrogen intake in Scottish men: Use of serum to validate a self-administered food-frequency questionnaire in older men. *Eur. J. Clin. Nutr.* **2006**, *60*, 129–135. [CrossRef] [PubMed]
44. Huang, M.-H.; Harrison, G.G.; Mohamed, M.M.; Gornbein, J.A.; Henning, S.M.; Go, V.L.W.; Greendale, G.A. Assessing the Accuracy of a Food Frequency Questionnaire for Estimating Usual Intake of Phytoestrogens. *Nutr. Cancer* **2000**, *37*, 145–154. [CrossRef] [PubMed]

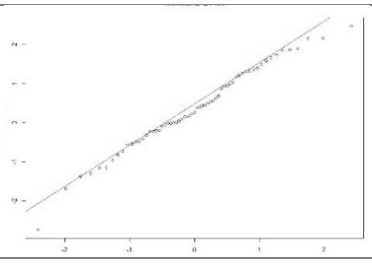
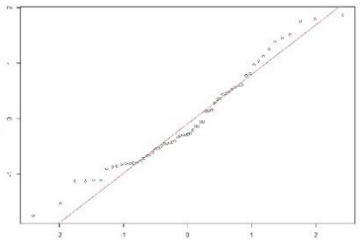
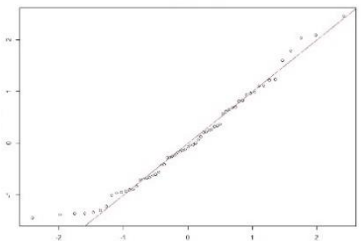
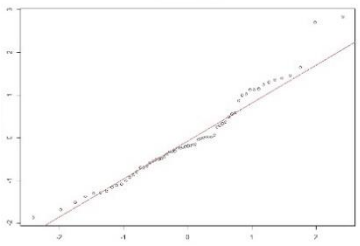
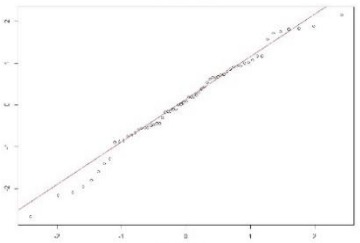
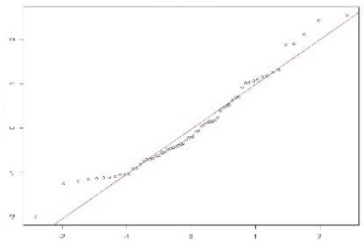
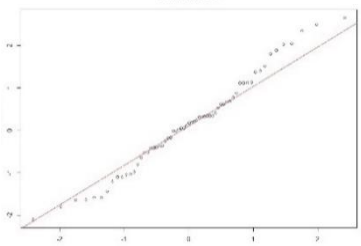
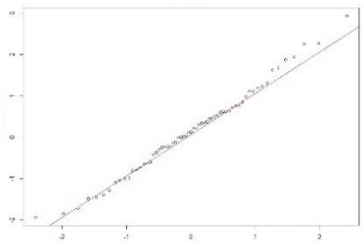
45. Atkinson, C.; Skor, H.E.; Fitzgibbons, E.D.; Scholes, D.; Chen, C.; Wähälä, K.; Schwartz, S.M.; Lampe, J.W. Overnight urinary isoflavone excretion in a population of women living in the United States, and its relationship to isoflavone intake. *Cancer Epidemiol. Biomark. Prev.* **2002**, *11*, 253–260.
46. Kim, J.; Kim, H.J.; Joung, H.; Park, M.K.; Li, S.; Song, Y.; Franke, A.A.; Paik, H.-Y. Overnight urinary excretion of isoflavones as an indicator for dietary isoflavone intake in Korean girls of pubertal age. *Br. J. Nutr.* **2010**, *104*, 709–715. [CrossRef]
47. Morimoto, Y.; Beckford, F.; A Franke, A.; Maskarinec, G. Urinary isoflavonoid excretion as a biomarker of dietary soy intake during two randomized soy trials. *Asia Pac. J. Clin. Nutr.* **2014**, *23*, 205. [CrossRef] [PubMed]
48. Chávez-Suárez, K.M.; Ortega-Vélez, M.I.; Valenzuela-Quintanar, A.I.; Galván-Portillo, M.; López-Carrillo, L.; Esparza-Romero, J.; Saucedo-Tamayo, M.S.; Robles-Burgueño, M.R.; Palma-Durán, S.A.; Gutiérrez-Coronado, M.L.; et al. Phytoestrogen Concentrations in Human Urine as Biomarkers for Dietary Phytoestrogen Intake in Mexican Women. *Nutrients* **2017**, *9*, 1078. [CrossRef]
49. Xiang, H.; Schevzov, G.; Gunning, P.; Williams, H.M.; Silink, M. A Comparative Study of Growth-Inhibitory Effects of Isoflavones and Their Metabolites on Human Breast and Prostate Cancer Cell Lines. *Nutr. Cancer* **2002**, *42*, 224–232. [CrossRef]
50. Angeloni, S.; Navarini, L.; Khamitova, G.; Sagratini, G.; Vittori, S.; Caprioli, G. Quantification of lignans in 30 ground coffee samples and evaluation of theirs extraction yield in espresso coffee by HPLC-MS/MS triple quadrupole. *Int. J. Food Sci. Nutr.* **2019**, *71*, 193–200. [CrossRef]
51. Horn-Ross, P.L.; Lee, M.; John, E.M.; Koo, J. Sources of Phytoestrogen Exposure among Non-Asian Women in California, USA. *Cancer Causes Control* **2000**, *11*, 299–302. [CrossRef]
52. Kuhnle, G.G.C.; Dell'Aquila, C.; Aspinall, S.M.; Runswick, S.A.; Mulligan, A.A.; Bingham, S.A. Phytoestrogen Content of Beverages, Nuts, Seeds, and Oils. *J. Agric. Food Chem.* **2008**, *56*, 7311–7315. [CrossRef]
53. Kuhnle, G.G.C.; Dell'Aquila, C.; Aspinall, S.M.; Runswick, S.A.; Mulligan, A.A.; Bingham, S.A. Phytoestrogen Content of Cereals and Cereal-Based Foods Consumed in the UK. *Nutr. Cancer* **2009**, *61*, 302–309. [CrossRef] [PubMed]
54. Kuhnle, G.G.; Dell'Aquila, C.; Aspinall, S.M.; Runswick, S.A.; Joosen, A.M.; Mulligan, A.A.; Bingham, S.A. Phytoestrogen content of fruits and vegetables commonly consumed in the UK based on LC–MS and 13C-labelled standards. *Food Chem.* **2009**, *116*, 542–554. [CrossRef]
55. Milder, I.E.J.; Feskens, E.J.M.; Arts, I.C.W.; de Mesquita, H.B.B.; Hollman, P.C.H.; Kromhout, D. Intake of the Plant Lignans Secoisolariciresinol, Matairesinol, Lariciresinol, and Pinoresinol in Dutch Men and Women. *J. Nutr.* **2005**, *135*, 1202–1207. [CrossRef]
56. Nurmi, T.; Heinonen, S.; Mazur, W.; Deyama, T.; Nishibe, S.; Adlercreutz, H. Lignans in selected wines. *Food Chem.* **2003**, *83*, 303–309. [CrossRef]
57. Peñalvo, J.L.; Haajanen, K.M.; Botting, N.; Adlercreutz, H. Quantification of Lignans in Food Using Isotope Dilution Gas Chromatography/Mass Spectrometry. *J. Agric. Food Chem.* **2005**, *53*, 9342–9347. [CrossRef]
58. Tetens, I.; Turrini, A.; Tapanainen, H.; Christensen, T.; Lampe, J.W.; Fagt, S.; Håkansson, N.; Lundquist, A.; Hallund, J.; Valsta, L.M. Dietary intake and main sources of plant lignans in five European countries. *Food Nutr. Res.* **2013**, *57*. [CrossRef]
59. Thompson, L.U.; Boucher, B.A.; Liu, Z.; Cotterchio, M.; Kreiger, N. Phytoestrogen Content of Foods Consumed in Canada, Including Isoflavones, Lignans, and Coumestan. *Nutr. Cancer* **2006**, *54*, 184–201. [CrossRef]
60. Valsta, L.M.; Kilkkinen, A.; Mazur, W.; Nurmi, T.; Lampi, A.-M.; Ovaskainen, M.-L.; Korhonen, T.; Adlercreutz, H.; Pietinen, P. Phyto-oestrogen database of foods and average intake in Finland. *Br. J. Nutr.* **2003**, *89*, S31–S38. [CrossRef]

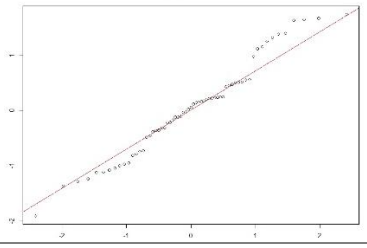
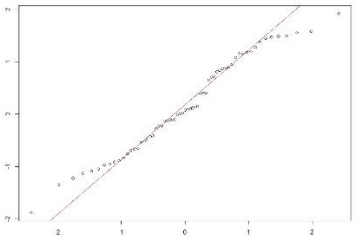
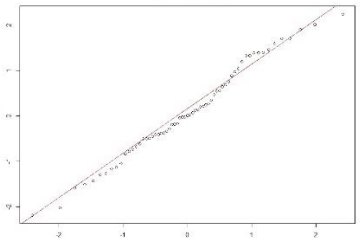
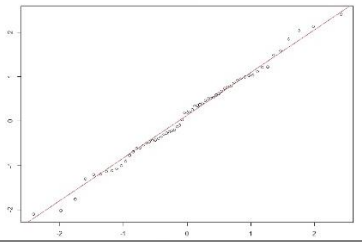
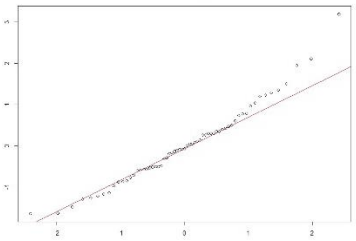
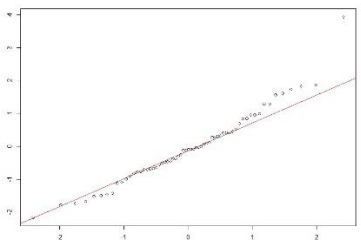
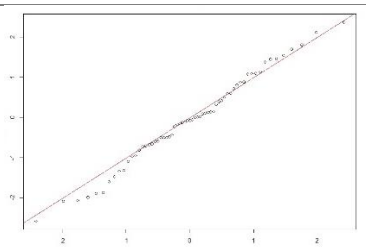
Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

IV.1.3. Exploration de l'influence de l'exposition à des phyto-estrogènes naturels sur la déclaration de poussées de lupus érythémateux disséminé

IV.1.3.1. Normalité des données

Le Tableau 9 regroupe les résultats des tests de normalité de l'ensemble des données de l'étude ISOLED (n=63)

Test analytique	QQplot	Test analytique	QQplot
	Hair_G_D		Hair_IFs
A = 3,4977 p = 7,471e-09		A = 3,2661 p = 2,768e-08	
	IFS		IHS
A = 2,9964 p = 1,301e-07		A = 2,1369 p = 1,696e-05	
	Plasma_D		Plasma_E
A = 13,754 p < 2,2e-16		A = 12,545 p < 2,2e-16	
	Plasma_ENL		Plasma_G
A = 1,7156 p = 0,0001903		A = 13,16 p < 2,2e-16	
	Plasma_IFs		Score_plasm_D

A = 12,332 p < 2,2e-16	
Score_plasm_ENL	
A = 0,89361 p = 0,02115	
Score_plasm_IFs	
A = 4,4719 p = 3,174e-11	
Score_Urin_ENL	
A = 0,73201 **p = 0,05354** Acceptation de H0	
Score_urin_IFs	
A = 4,073 p = 2,984e-10	
Urine_D	
A = 23,222 p < 2,2e-16	
Urine_IFs	
A = 23,283 p < 2,2e-16	

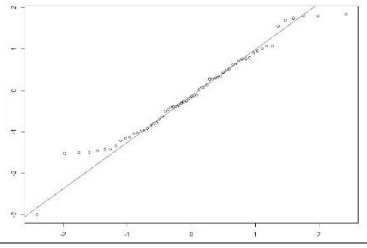
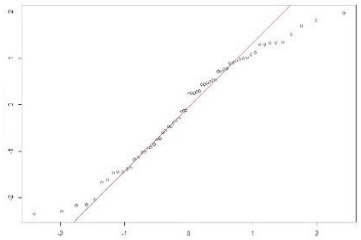
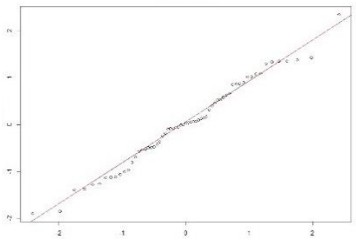
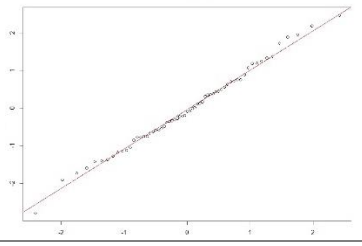
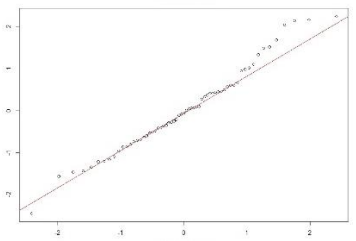
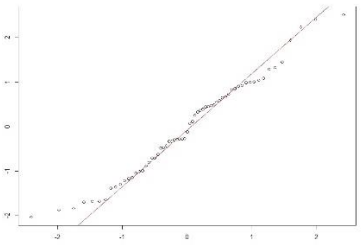
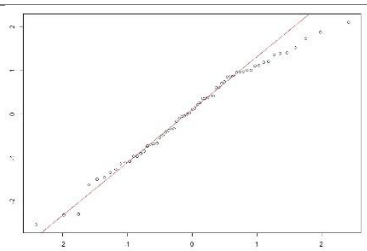
A = 4,9103 p = 2,723e-12	
Score_plasm_G	
A = 4,1462 p = 1,977e-10	
Score_urin_D	
A = 4,3566 p = 6,061e-11	
Score_urin_G	
A = 4,1119 p = 2,397e-10	
Urine_G	
A = 23,443 p < 2,2e-16	
Urine_E	
A = 15,927 p < 2,2e-16	
Urine_ENL	
A = 11,277 p < 2,2e-16	

Tableau 9. Résultats des tests de normalité de l'ensemble des données de l'étude ISOLED (n=63)

Les tests d'Anderson-Darling et QQ plots révèlent que la plupart des variables étudiées présentent un écart significatif par rapport à une distribution normale, à l'exception du score urinaire d'exposition aux lignanes sur les dernières 24-48 heures (Score_urin_ENL), dont la valeur p est légèrement au-dessus de 0,05. Par conséquent, le jeu de données a été considéré ne pas suivre la loi normale, et des tests non-paramétriques ont été réalisés pour la suite des calculs.

IV.1.3.2. Influence du statut sanitaire sur les scores et valeurs mesurées des concentrations d'IFs et d'ENL

L'observation des médianes et de la dispersion des valeurs de concentrations mesurées et les scores d'IFs selon les statuts des participantes sont exposés dans la Figure 7.

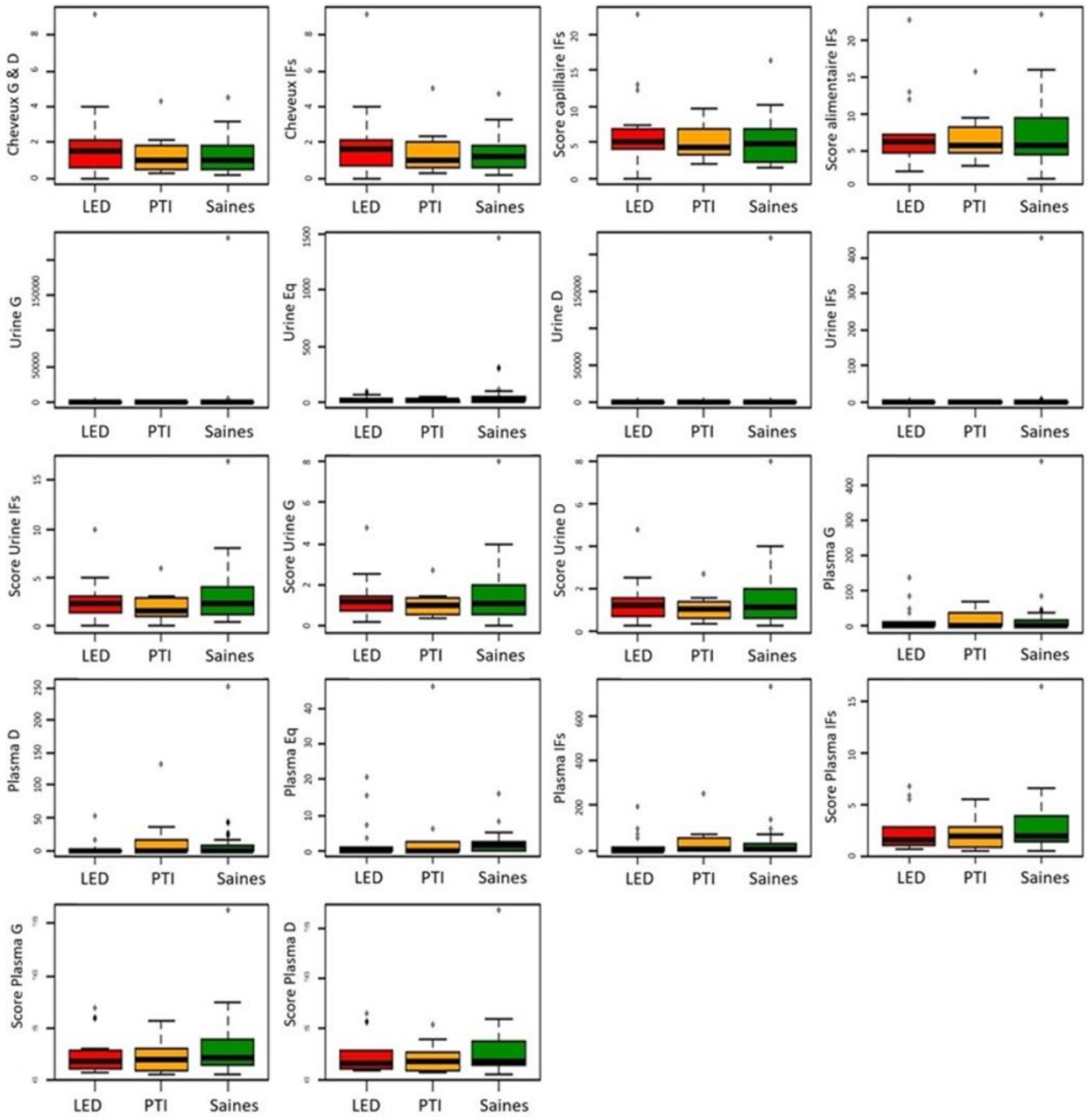


Figure 7. Boites à moustaches représentant l'influence du statut sanitaire des volontaires (LED = femmes atteintes de LED, n= 20; PTI = femmes atteintes de PTI, n= 12 ; SAINES groupe témoin non malade, n= 31) sur l'ensemble des variables de l'étude ISOLED.

Il apparait que certaines valeurs sont extrêmes (Figure 7) et pourraient avoir un poids disproportionné dans l'analyse. Pour évaluer leur impact, ces échantillons (6 Saines, 1 LED et 1 PTI) ont été éliminés dans la nouvelle analyse présentée en figure 8.

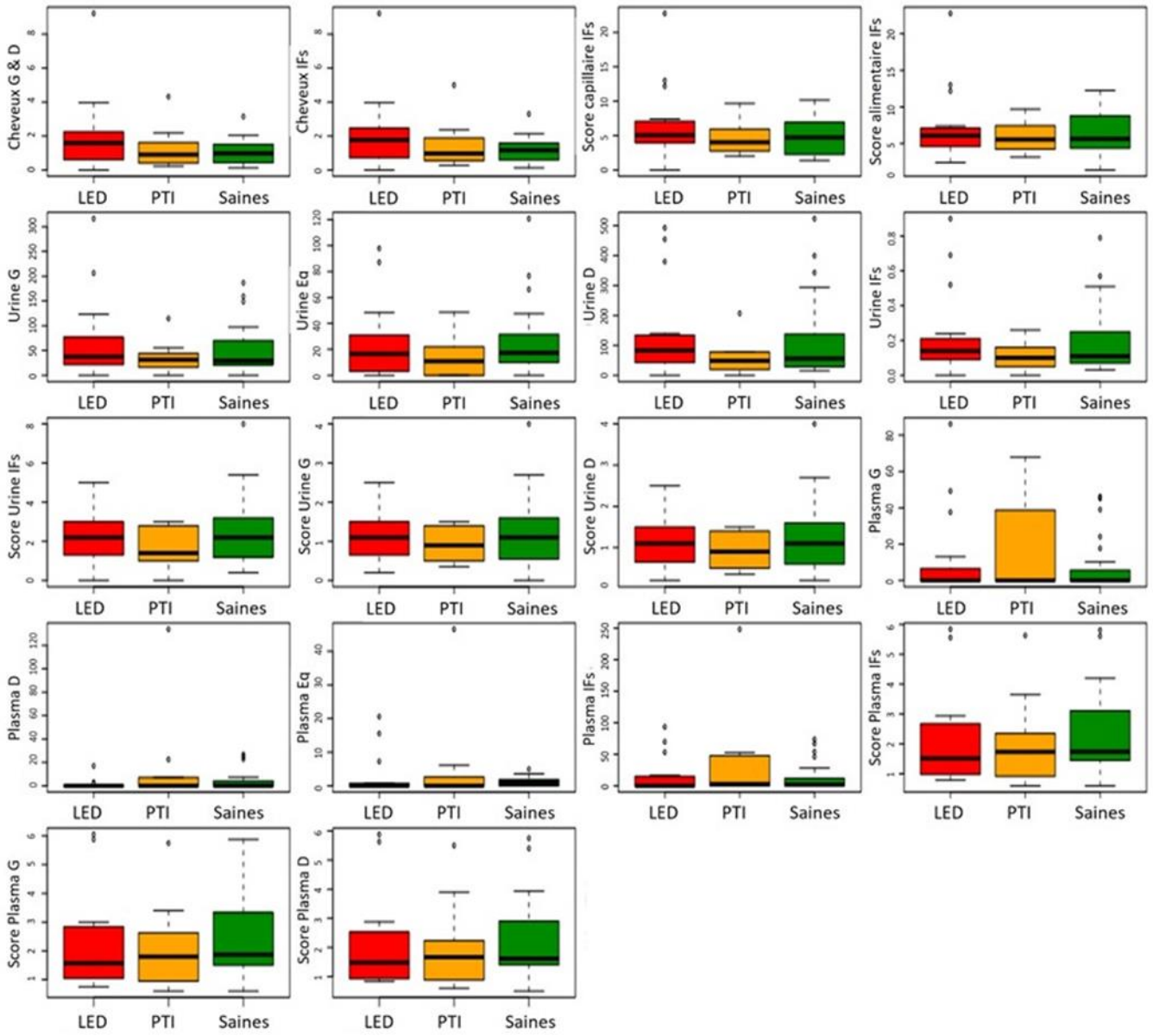


Figure 8. Boites à moustache représentant l'influence du statut sanitaire des volontaires (LED = femmes atteintes de LED, n=19 ; PTI = femmes atteintes de PTI, n= 11 ; SAINES groupe témoin non malade, n= 25) sur les variables IFs de l'étude ISOLED sans les valeurs extrêmes

Les tests de Kruskall-Wallis sur les jeux de données complets (Figure 7) ou allégés des valeurs extrêmes (Figure 8) ne révèlent aucune différence significative de l'influence du statut sanitaire sur aucune des variables IFs.

Les données obtenues sur l'ENL font l'objet des Figures 9 et 10.

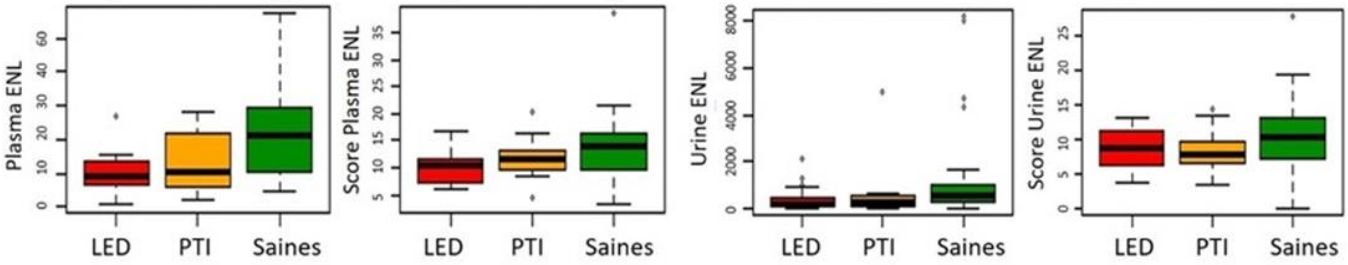


Figure 9. Boites à moustache représentant l'influence du statut sanitaire des volontaires (LED = femmes atteintes de LED, n= 20 ; PTI = femmes atteintes de PTI, n= 12 ; SAINES groupe témoin non malade, n= 31) sur l'ensemble des variables ENL de l'étude ISOLED

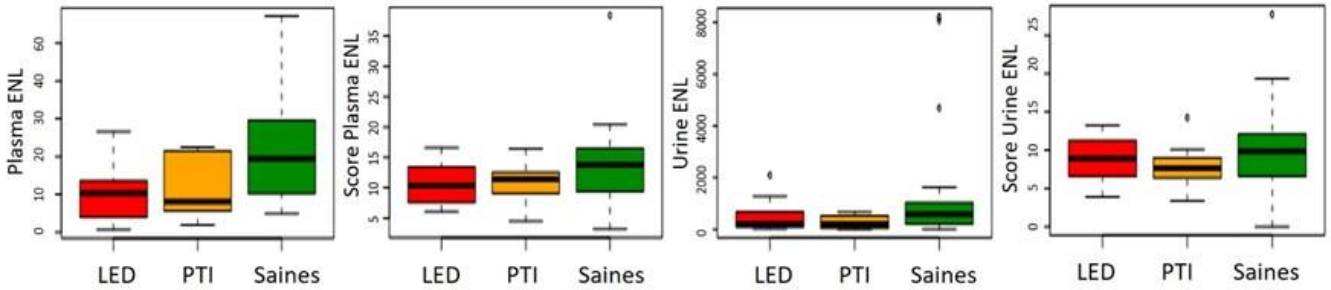


Figure 10. Boites à moustache représentant l'influence du statut sanitaire des volontaires (LED = femmes atteintes de LED, n=19 ; PTI = femmes atteintes de PTI, n= 11 ; SAINES groupe témoin non malade, n= 25) sur les variables ENL de l'étude ISOLED sans les valeurs extrêmes

Les tests de Kruskal-Wallis sur les variables ENL (Figures 9 et 10) ont révélé une différence significative des concentrations plasmatiques d'ENL selon le statut sanitaire des volontaires.

L'analyse par paire (Figure 11) a montré que les concentrations plasmatiques d'ENL des deux groupes de patientes (LED et PTI) étaient significativement inférieures à celle du groupe de volontaires saines ($p<0,01$ pour LED et $p<0,05$ pour PTI), sans que la différence entre les deux groupes de patientes ne soit statistiquement significative.

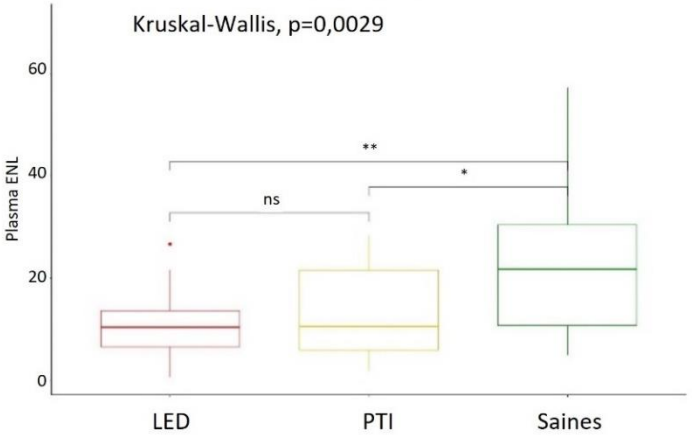


Figure 11. Représentation de la distribution des dosages d'ENL plasmatiques, selon le statut des volontaires

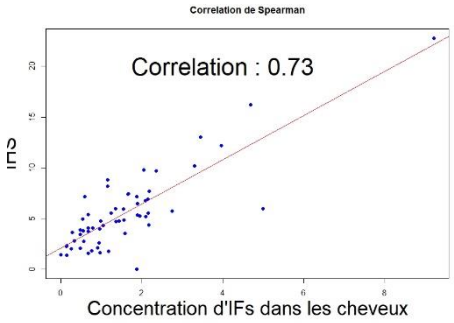
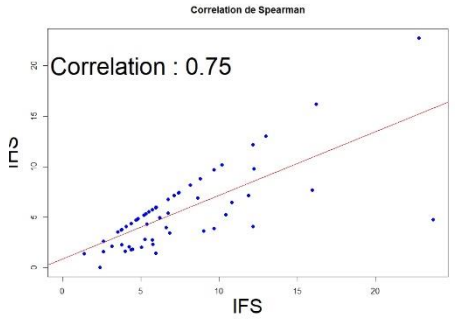
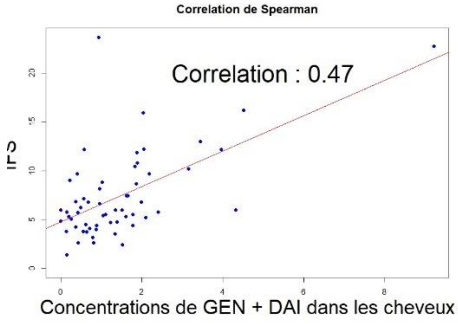
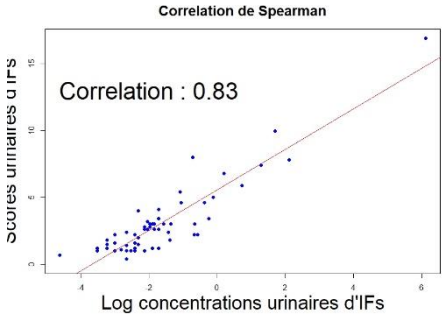
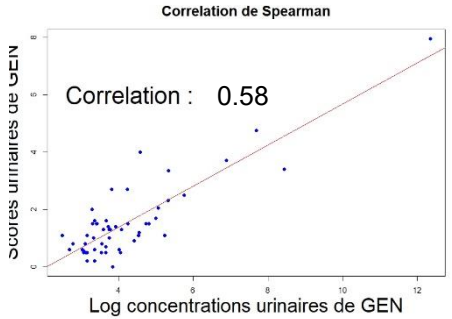
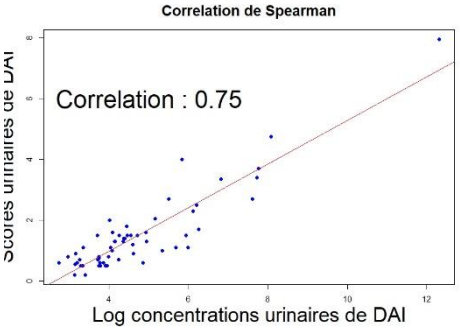
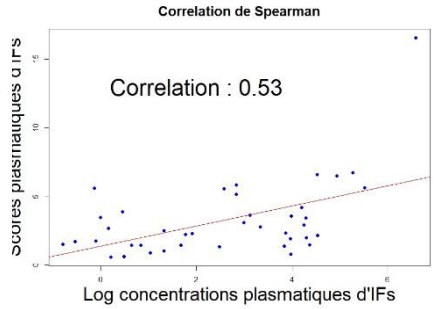
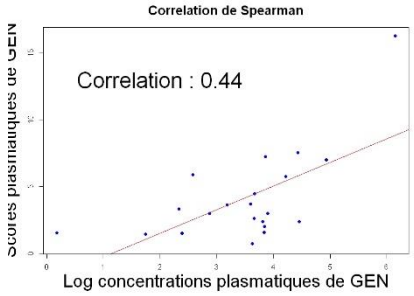
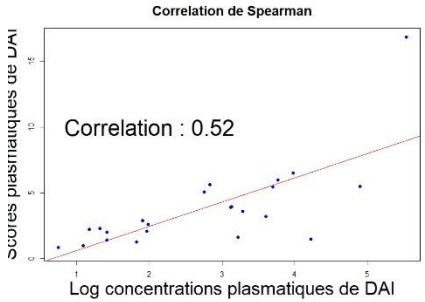
Par conséquent, la concentration plasmatique d'ENL constitue le seul indicateur permettant de discriminer statistiquement les patientes atteintes de LED ou PTI des volontaires saines.

IV.1.3.3. Corrélations entre dosages biologiques et scores alimentaires

Le Tableau 10 décrit les résultats des corrélations entre les dosages plasmatiques, urinaires et capillaires et leurs scores correspondants, obtenus selon le test non paramétrique de Spearman. Il est à noter que chacune des variables biologiques testée

est positivement corrélée, de manière statistiquement significative, à son score correspondant.

Les variables IFs capillaires et IHS ainsi que dosage urinaire des IFs et score urinaire des 24-48 heures sont très fortement corrélées. Les scores plasmatiques sont quant à eux fortement corrélés aux dosages correspondants. Les scores d'ENL sont également fortement corrélés aux dosages urinaires et plasmatiques correspondants.

Variables testées et résultats (rho de Spearman)	Représentation graphique	Variables testées et résultats (rho de Spearman)	Représentation graphique	Variables testées et résultats (rho de Spearman)	Représentation graphique
Hair_IFs et IHS sont très fortement corrélées (p <0,001)		IFS et IHS sont très fortement corrélées (p <0,001)		Hair_IFs et IFS sont moyennement corrélées (p <0,001)	
Urine_IFs et Score_urin_IFs sont très fortement corrélées (p <0,001)		Urine_G et Score_urin_G sont fortement corrélées (p <0,001)		Urine_D et Score_urin_D sont très fortement corrélées (p <0,001)	
Plasma_IFs et Score_plasm_IFs sont fortement corrélées (p <0,001)		Plasma_G et Score_plasm_G sont moyennement corrélées (p <0,001)		Plasma_D et Score_plasm_D sont fortement corrélées (p <0,001)	

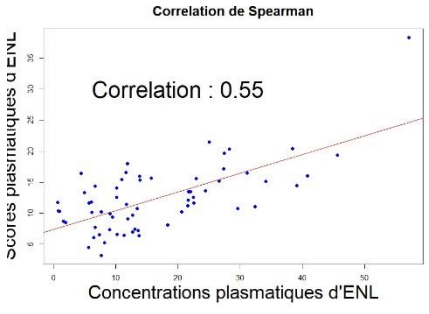
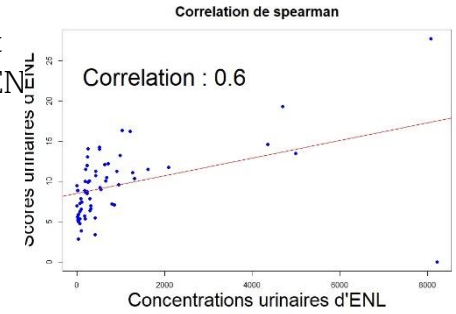
Variables testées et résultats (rho de Spearman)	Représentation graphique	Variables testées et résultats (rho de Spearman)	Représentation graphique	Variables testées et résultats (rho de Spearman)	Représentation graphique
Plasma_ENL et Score_plasm_ENL sont fortement corrélées (p <0,001)		Urine_ENL et Score_Urin_ENL sont fortement corrélées (p <0,001)			

Tableau 10. Corrélations de Spearman de l'étude ISOLED

IV.1.3.4. Analyse des données par ACP

Compte tenu du fait que la plupart des variables ne suivent pas une distribution normale, le jeu de données a été centré et réduit. Cette opération consiste à soustraire à chaque variable sa moyenne et à la diviser par son écart-type, ce qui revient à considérer que la moyenne de chaque variable est égale à 0 et l'écart-type à 1. Cette transformation permet de normaliser les données, ce qui est une condition préalable à l'utilisation de l'ACP.

La Figure 12 représente les contributions à la variance de chaque nouvelle dimension obtenue par l'ACP. Les deux premières dimensions suffisent à expliquer plus de 64% de la variance totale, et ce pourcentage augmente jusqu'à 77% en incluant la troisième dimension. Ce résultat indique une bonne qualité explicative globale de l'analyse.

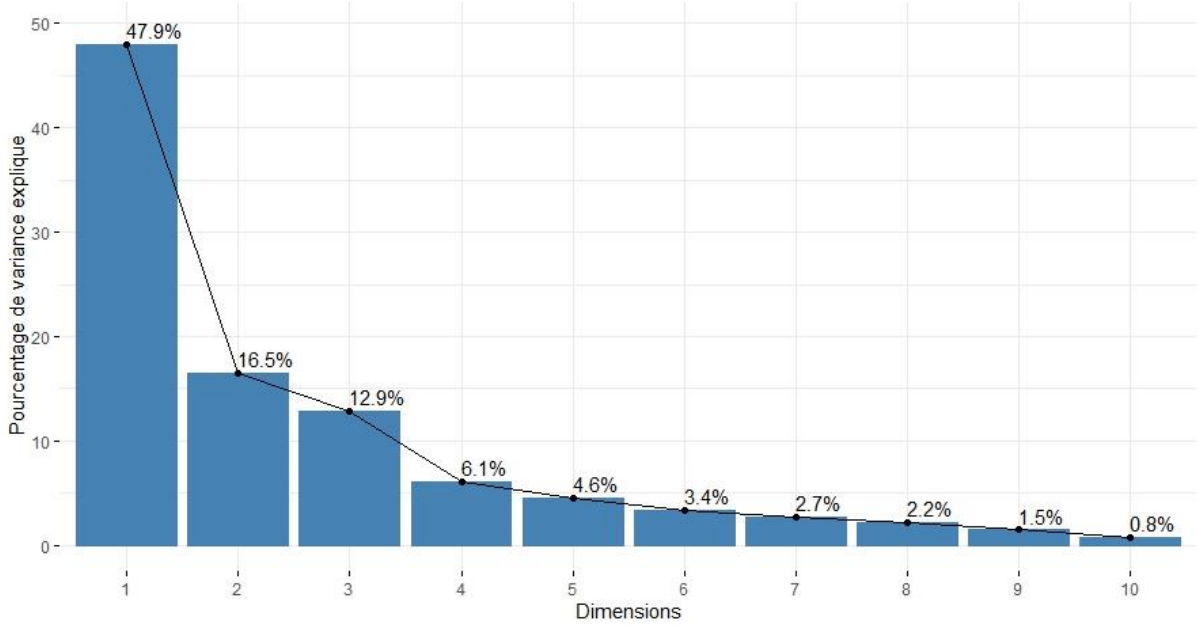


Figure 12. Contributions de chaque dimension à l'explication de la variance

Le cercle de corrélation présenté en Figure 13 permet d'observer que la plupart des variables, à l'exception de « Urine_E », « Plasma_E » et celles liées à l'ENL se situent à proximité du cercle de corrélation, indiquant une bonne représentation sur ce plan factoriel. Ce manque de corrélation pour l'équol et l'ENL pourrait s'expliquer par le fait que la présence de ces métabolites dépend de bactéries spécifiques de la flore intestinale qui ne semblent pas liées aux pathologies considérées.

Les angles fermés entre les variables « Hair_G_D », « Hair_IFs » et « IHS » indiquent une excellente corrélation entre elles. Elles sont également très bien représentées par la dimension 2 (axe vertical), car très proches de cet axe et très éloignées de l'autre. La variable « IFS » est relativement proche du cercle mais assez éloignées des deux axes, ce qui signifie qu'elle est mal représentée par les deux dimensions. Elle peut donc ne pas être considérée pour l'analyse ultérieure.

Les variables urinaires et plasmatiques des IFs se situent globalement à proximité de l'axe de la dimension 1 (axe horizontal), indiquant une bonne représentation par cette dimension. Les petits angles formés entre elles indiquent un bon niveau de corrélation.

Ce groupe de variables est indépendant du groupe formé par « Hair_G_D », « Hair_IFs » et « IHS » en raison du grand angle séparant leurs variables respectives.

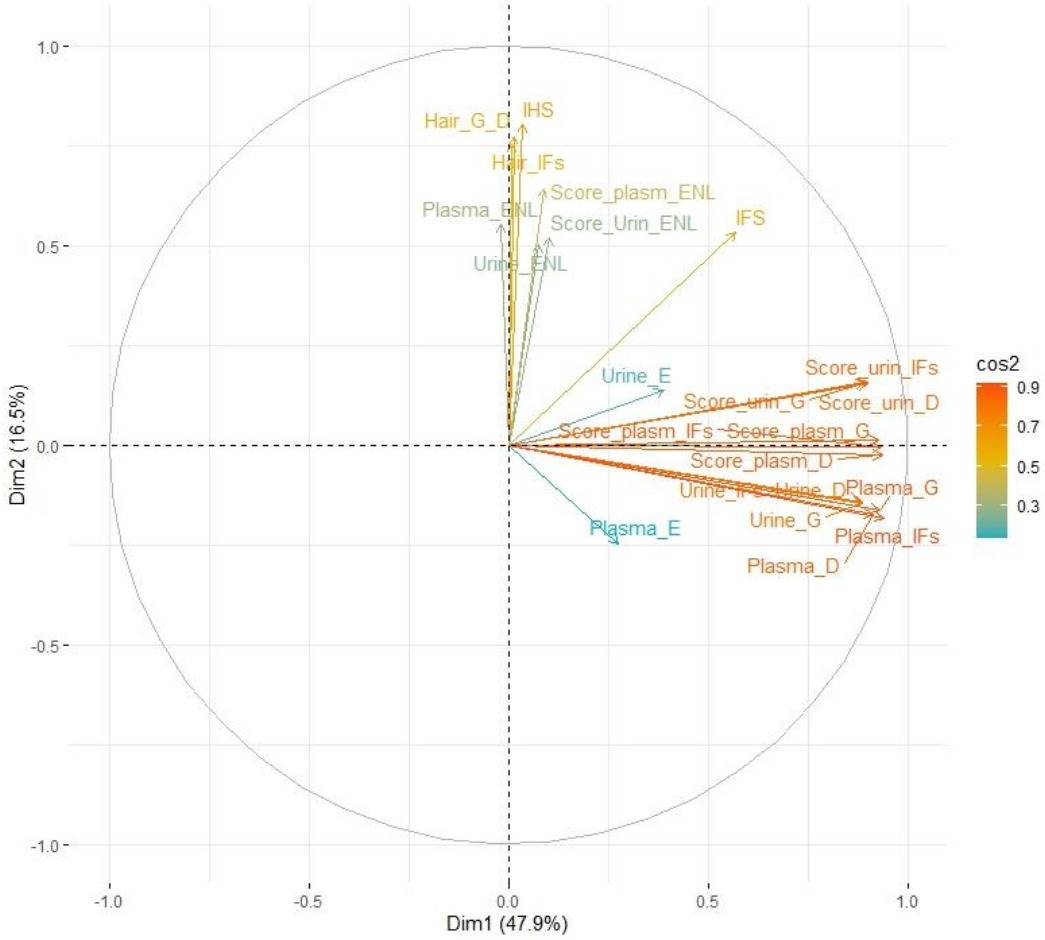


Figure 13. Cercle de corrélation entre les variables initiales et les composantes

Parmi les variables représentées par la dimension 1, il faut noter que celles liées au plasma (« Plasma_IFs », « Score_plasm_D », « Score_plasm », « Plasma_G », « Score_plasm_G », et « Plasma_D ») contribuent légèrement plus à la dimension 1 que celles liées aux urines. Une hypothèses peut être formulées pour expliquer au moins en partie cette observation. Il s'agit d'une meilleure estimation des concentrations biologiques réelles par le biais de mesures directes dans le plasma, comparativement à l'ajustement nécessaire des concentrations urinaires obtenues sur la base de la créatinine (correction pour tenir compte de la dilution potentielle des composés dans les urines).

L'analyse du diagramme de corrélation exposé en Figure 14 confirme que les variables liées à l'ENL ne sont pas représentées par les deux premières dimensions de l'ACP présentées dans la Figure 13. En effet, ces variables contribuent de manière plus importante à la dimension 3.

Ces observations permettent de déduire que :

- la dimension 1 représente principalement la consommation récente d'IFs ;
- la dimension 2 correspond aux habitudes de consommation d'IFs et, dans une moindre mesure, à la consommation récente de lignanes ;

- la dimension 3 représente une combinaison de la consommation récente de lignanes et de celle habituelle d'IFs, dans une moindre mesure.

L'analyse réalisée révèle une indépendance marquée entre les variables représentées par les dimensions 1 et 2. Cette observation s'accorde avec le contexte occidental, où la consommation de soja, principale source d'IFs, est généralement occasionnelle. En conséquence, les concentrations d'IFs mesurées dans les cheveux ne reflètent pas nécessairement l'apport en phytoE sur les 48 dernières heures.

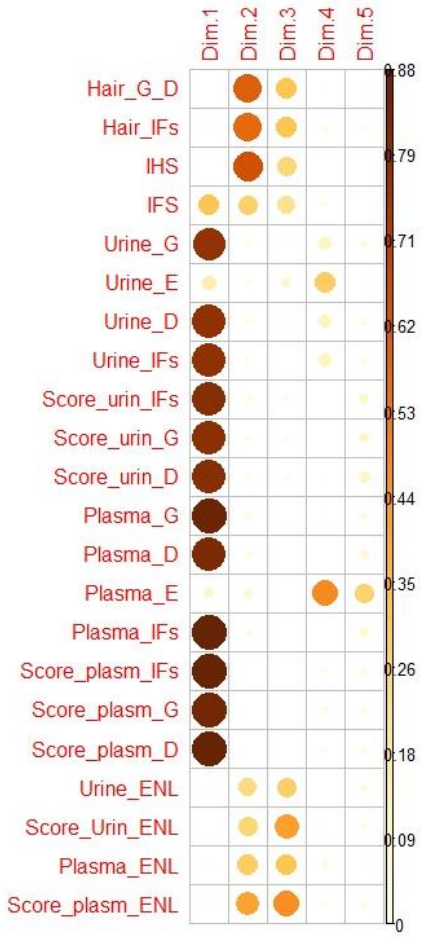


Figure 14. Diagramme de corrélation des contributions des anciennes variables aux nouvelles dimensions

La figure 15 représente la répartition des individus regroupés par statut selon les dimensions 1 et 2 de l'ACP. Il est notable que les aires des 3 groupes se superposent en grande partie. Ce résultat suggère que les deux premières dimensions de l'ACP ne permettent pas de distinguer clairement les individus selon leur statut de santé.

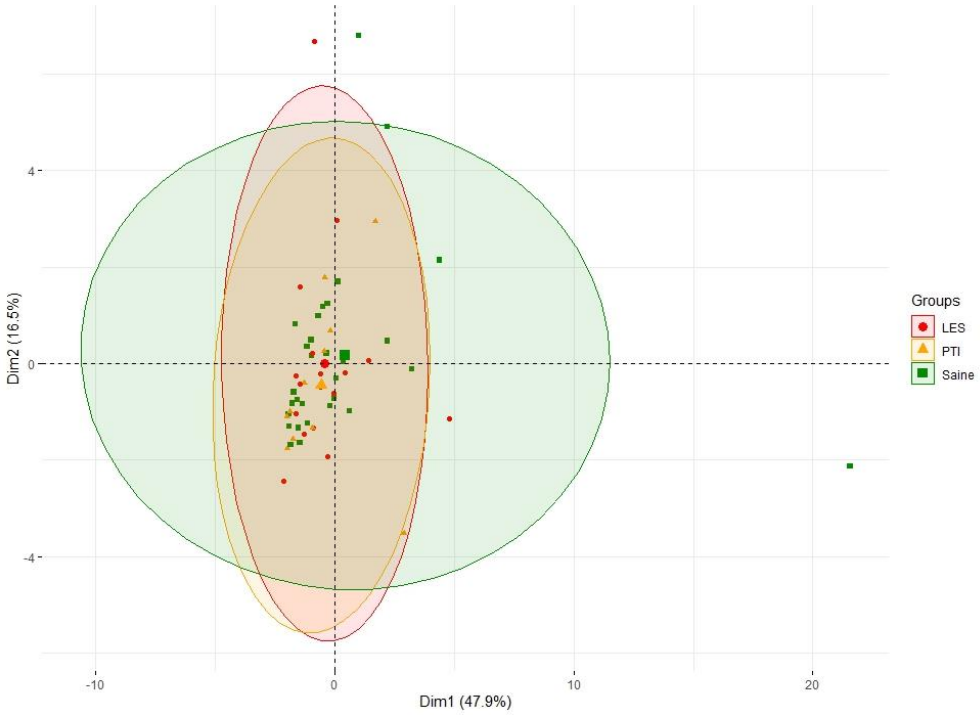


Figure 15. Nuage des individus répartis par statut et ellipses de concentration

Une représentation en 3D des individus selon les trois premières dimensions ne permet pas une meilleure distinction des groupes comme on peut le voir sur la Figure 16.

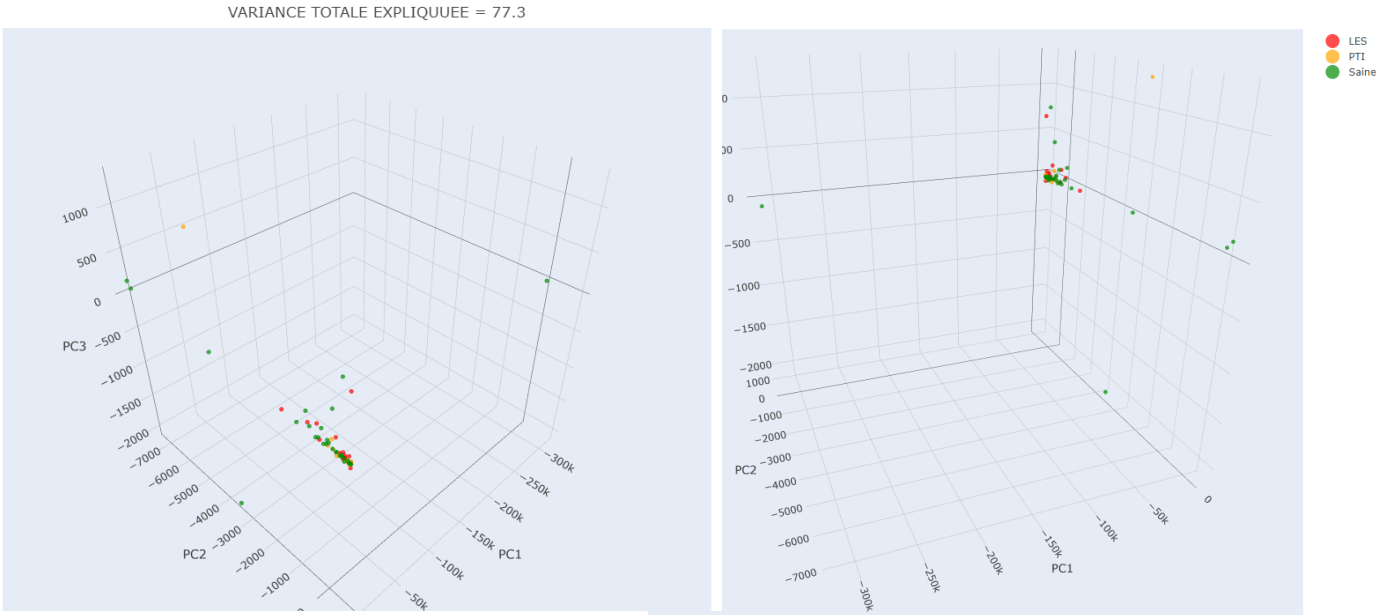


Figure 16. Représentation des individus par groupe sanitaire selon les 3 principales dimensions

On rappelle que l'étude ISOLED a été clôturée après inclusion de 63 participantes dont 31 étaient des volontaires saines, 20 souffraient de LED et 12 de PTI. Les effectifs réduits, particulièrement dans les groupes de patientes, ont limité la puissance statistique de l'étude et n'ont pas permis de dégager des résultats probants quant à la comparaison des groupes.

IV.1.3.5. Discussion

Pouvoir prédicteur des scores et comparaison aux autres études

Le Tableau 11 reprend les résultats de cette étude (n=63) en les comparant à ceux des études de Heald, French, Morimoto, Kim, Chávez et leurs collaborateur·ices [127,239–242]. L'ensemble des études comparés s'est basé sur le test de Spearman, jugé plus robuste.

Sujets	Nature des échantillons	Données alimentaires	Corrélations	Référence
		Biomarqueur = IFs		
26 femmes Canadiennes prémenoP[20]	Urine des 24h	Questionnaire d'habitude Rappel des 24h	rho = 0,64; p < 0,001 rho = 0,54; p = 0,004	[241]
203 hommes volontaires	Sérum	QFA[21]	rho = 0,27; p < 0,001	[127]
256 femmes préménoP	Urine des 12h	QFA Faible : 0,1 – 2,3 Fort : 49,8 – 74,8	rho = 0,51; p < 0,001	[239]
24 filles prépubaires	Urine des 12h	Rappel des 24h – 3 j IFs : 3,0 – 13,3	rho = 0,72; p < 0,001	[240]
100 femmes en bonne santé	Urine des 12h	Rappel des 24h	rho = 0,46; p < 0,001	[242]
63 femmes Françaises préménoP	Cheveux Urine ponctuelle Sérum	Questionnaire d'habitude Rappel des 48h et 24h	rho = 0,73; p < 0,001 rho = 0,83; p < 0,001 rho = 0,53; p < 0,001	Présente étude
		Biomarqueur = ENL		
26 femmmes Canadiennes prémenoP[22]	Urine des 24h	Questionnaire d'habitude Rappel des 24h	rho = 0,46; p = 0,02 rho = 0,40; p = 0,05	[241]
63 femmes Françaises préménoP	Urine ponctuelle Sérum	Rappel des 48h et 24h	rho = 0,60; p < 0,001 rho = 0,55; p < 0,001	Présente étude

Tableau 11. Comparaison des résultats d'ISOLED avec d'autres études utilisant le dosage des IFs ou de l'ENL dans les fluides biologiques comme marqueur d'exposition. Seules les études utilisant le test de corrélation de Spearman ont été sélectionnées.

Cette étude complète les résultats des deux précédents articles cités dans les sections IV.1.2.1 et IV.1.2.2. Alors qu'ici la taille de la cohorte est plus grande et les tests statistiques plus robustes, les conclusions antérieures sont confirmées : les questionnaires alimentaires développés dans cette investigation ont permis

[20] PréménoP : pré-ménopausées
[21] QFA : questionnaire de fréquence alimentaire
[22] MénoP : ménopausées

d'améliorer les corrélations entre les dosages des marqueurs biologiques (IFs et ENL) et les scores alimentaires correspondants.

Le dosage des IFs dans les cheveux permet d'atteindre une corrélation de rho = 0,73 avec les habitudes alimentaires, contre rho = 0,64 lorsque les urines de 24 heures sont utilisées comme source de biomarqueurs [241]. Même sur des cohortes plus importantes (plus de 100 individus), les corrélations obtenues avec les scores urinaires de cette étude restent supérieures (0,83 vs 0,27 [127] ; 0,51 [239] et 0,46 [242]).

Ces meilleures résultats peuvent s'expliquer par une estimation plus exhaustive des sources d'IFs dans les aliments et par la prise en compte des cinétiques (biodisponibilité et élimination) des phytoE dans le calcul des scores alimentaires.

Influence du statut sanitaire sur les variables mesurées

L'analyse statistique n'a pas révélé de différences significatives entre les groupes de patientes et de sujets sains pour la majorité des variables mesurées (test de Kruskal-Wallis et ACP). Seule la concentration plasmatique d'ENL a permis de discriminer efficacement les groupes sains et malades, ces derniers ayant moins d'ENL que les premiers.

Ce résultat pourrait indiquer que la consommation de précurseur d'ENL et la flore intestinale compétente pour la production d'ENL seraient associées à un meilleur état de santé et, dans ce cas, à une diminution du risque de déclarer un LED. En fait l'ENL peut être considéré comme un marqueur de la consommation de fruits et légumes [243,244] et il n'est donc pas surprenant de voir que les taux d'ENL sont plus élevées chez les femmes saines. Cependant, il apparaît important de noter ici que quelques travaux ont essayé de mettre en avant l'implication du microbiote dans le LED, notamment un rapport intestinal Firmicutes/Bacteroidetes significativement (2,5 fois) plus faible chez les patient·es atteint·es de LED, même en rémission, par rapport aux témoins sains [245]. D'autres auteur·es ont rappelé la dysbiose du microbiote associé au LED qui pourrait être impliquée dans sa pathogénèse notamment en favorisent le dépôt de complexes immunitaires, l'activation du complément et l'infiltration des macrophages [246]. D'autres études sur le lupus ont mis en évidence une perméabilité intestinale accrue, également appelée « hyperperméabilité intestinale », qui conduit à une altération de la fonction de barrière intestinale. Ces mêmes études, menées chez l'humain et la souris lupique, ont également observé une diminution des bactéries bénéfiques telles que Bifidobacterium et une augmentation des bactéries pathogènes telles que *Enterococcus gallinarum* et *Ruminococcus gnavus*, étroitement associées à la progression de la maladie [247]. L'implication de ces dérégulations du microbiote, dont la compréhension n'est que partielle, ne peut donc pas être exclue.

Par ailleurs, contrairement aux IFs qui sont des agonistes stricts des ER α et β, l'ENL est un SERM qui bloque l'activité ERα via AF-1 [248]. Si ce mécanisme était confirmé et s'avérait exact cela signifierait que le LED pourrait être lié à une action des estrogènes via ce récepteur et ce domaine de transactivation.

Limites techniques

Malgré des résultats insuffisamment probants, l'étude ISOLED a permis d'observer une différence significative de l'ENL plasmatique entre les individus des groupes LED, PTI et volontaires saines. Cependant, il n'a pas été possible de distinguer ces groupes en fonction de leur consommation globale de phytoE, notamment les IFs.

Plusieurs facteurs pourraient expliquer cette limitation. Tout d'abord, la collecte des données sur la consommation de phytoE via des questionnaires et des prélèvements urinaires ponctuels présente des biais méthodologiques inhérents. Les questionnaires reposent sur la mémoire des participants, ce qui peut conduire à des erreurs de rappel. Les prélèvements urinaires, quant à eux, sont ajustés à la créatinine sans tenir compte du volume total d'urines excrété, ce qui peut introduire des imprécisions dans les calculs. De plus, de nombreux échantillons se situaient proches de la limite de quantification de l'ELISA, ce qui peut affecter la fiabilité des résultats.

En second lieu, l'effectif réduit de l'étude, en particulier dans le groupe LED, pourrait limiter la robustesse de l'analyse statistique. Un échantillon plus large aurait permis de détecter des différences plus subtiles entre les groupes et de mieux comprendre les relations potentielles entre la consommation de phytoE et poussées lupiques.

Calcul de la puissance statistique

Afin d'estimer la taille d'échantillon requise pour détecter une différence significative entre les groupes, un calcul de puissance statistique a été réalisé en utilisant la formule suivante :

$$n = \frac{2 \times S^2}{\Delta^2}\left(Z_{\alpha/2} + Z_{\beta}\right)^2$$

Avec :

- n = taille d'échantillon par groupe ;
- écart-type (S) = 3,85 ; 2,4 ou 1,3 (cf. annexe 2_Résultats des dosages d'IFs_ISOLED) ;
- Δ (la différence attendue entre les scores des groupes) = 0,05 ou 0,7 correspondant respectivement aux écarts des scores urinaires ou plasmatiques entre groupe LED et groupe sain ;
- Z = valeur critique de la distribution normale ;
- $Z_{\beta} = 0{,}842$ si le risque β=20% ;
- $Z_{\alpha/2}$ =1,96 si α = 5%.

Les résultats du calcul sont résumés dans le tableau 12, selon plusieurs scénarii d'écart-type observés et de différence attendues. Par exemple, la ligne 2 du tableau signifie qu'il faudrait 185 personnes pour le groupe LED et 185 personnes pour le groupe contrôle, soit 370 personnes en totalité, pour un écart-type de 2,4 et une différence attendue de 0,7, qui correspondent tous les deux aux valeurs obtenues des scores plasmatiques d'IFs. La 3ème ligne du tableau correspond à l'effectif estimé en se basant sur la différence et l'écart-type des scores urinaires.

n	S	Δ
475	3,85	0,7
185	2,4	
55	1,3	
93073	3,85	0,05
36168	2,4	
10612	1,3	

Tableau 12. Résultats des calculs d'effectifs par groupe selon plusieurs scénarii. n : nombre de sujets, S : écart-type attendu ; Δ signification de la différence attendue

En appliquant cette formule à différentes valeurs d'écart-type (S) et de différence attendue (Δ), il a été déterminé que la taille d'échantillon requise par groupe pouvait varier de 185 à 10 612. Ces effectifs importants suggèrent qu'une étude multicentrique serait nécessaire pour obtenir une puissance statistique adéquate.

En conclusion, l'étude ISOLED, bien que limitée par sa taille d'échantillon, fournit des informations préliminaires sur l'association potentielle entre l'exposition aux phytoE et les poussées de LED. Le rôle préventif de l'ENL, bien que n'étant pas complètement établi par cette étude, est conforme à l'idée que la consommation de fruits et légumes, dont l'ENL est le marqueur, est synonyme de meilleure santé globale. A noter que l'effet des IFs se distingue, ici comme dans d'autres études, de l'effet des entérolignanes.

Des études multicentriques avec des effectifs plus importants sont nécessaires pour confirmer ou infirmer cette association.

IV.2. Réduction de l'exposition aux isoflavones : une nécessité pour la santé publique

IV.2.1. Réduction des isoflavones à l'échelle préindustrielle

PHYTOSECUR 1 : article correspondant à la référence [49]

Les travaux, présentés dans la publication intitulée *Improved Food Processing Techniques to Reduce Isoflavones in Soy-Based Foodstuffs*, ont porté sur des graines de soja destinées à la production de PST [49].

Le protocole tel que mis en œuvre à l'échelle pilote n'a pas pu être testé en phase industrielle. En effet, l'étape de texturisation, devant être mise en œuvre sur les graines dépelliculées déflavonoïdisées, ne pouvait être réalisée que sur des quantités de matière de l'ordre de la tonne. Un deuxième projet, n'ayant pas encore obtenu de financement, a été envisagé pour répondre à cette contrainte.

foods

MDPI

Article

Improved Food-Processing Techniques to Reduce Isoflavones in Soy-Based Foodstuffs

Souad Bensaada [1,2,3,*], François Chabrier [4], Pascal Ginisty [5], Carine Ferrand [1,6], Gabriele Peruzzi [3], Marc Valat [1,7] and Catherine Bennetau-Pelissero [1,2,8,*]

1 Campus Carreire, Pharmacy Faculty, Department Sciences and Technology, University of Bordeaux, 33076 Bordeaux, France; carine.ferrand@u-bordeaux.fr (C.F.); marc.valat@u-bordeaux.fr (M.V.)
2 ARNA (Nucleic Acids: Natural and Artificial Regulations), U1212 Inserm, UMR CNRS 5320, University of Bordeaux, 33076 Bordeaux, France
3 Biopress, 47400 Tonneins, France; gabriele.peruzzi@biopress.fr
4 Agrotec, Agropole, 47310 Estillac, France; f.chabrier@agrotec-france.com
5 IFTS (Institute of Filtration and Separative Techniques), 47510 Foulayronnes, France; pascal.ginisty@ifts-sls.com
6 BFP (Fruit Biology and Pathology), UMR CNRS 1332, INRA Bordeaux-Aquitaine, University of Bordeaux, 33140 Villenave-d'Ornon, France
7 I2M (Mechanical and Engineering Institute), UMR CNRS 5295, University of Bordeaux, 33405 Talence, France
8 Bordeaux Sciences Agro, 33175 Gradignan, France
* Correspondence: souad.bensaada@u-bordeaux.fr (S.B.); catherine.bennetau@u-bordeaux.fr (C.B.-P.)

Abstract: Soy is a growing protein source; however, the isoflavones it contains are of concern, as they exhibit estrogenic activities whose toxicological limits might be exceeded. Reducing their concentrations to safe levels while preserving nutritional quality in soy foodstuffs is therefore a matter of public health. The main objective of this paper is to develop at pilot scale a process for isoflavones' extraction from soybeans, and to show its feasibility and efficiency. The study was conducted by first optimizing the previously obtained laboratory treatment key factors. These data were then transposed to the pilot level. Finally, the process was adjusted to technical constraints which appeared at pilot scale: the mandatory use of drenching and the exploration of granulometry analysis. The involved steps were validated by monitoring the genistein and daidzein content variations through statistical analysis of the data of an ELISA and a Folin–Ciocalteu assay. Additionally, isoflavones' recovery from treatment waters for their valorisation and the water cleaning by means of filtration, centrifugation and resin adsorption were carried out. The results showed that the most successful pilot treatment developed involved soybean dehulling, drenching, washing and drying and almost halved isoflavones while preserving the main nutritional characteristics. A combination of techniques led to almost complete recovery of isoflavones from process waters.

Keywords: genistein; daidzein; consumer exposure; manufacturing process; soybean; isoflavone removal

check for updates

Citation: Bensaada, S.; Chabrier, F.; Ginisty, P.; Ferrand, C.; Peruzzi, G.; Valat, M.; Bennetau-Pelissero, C. Improved Food-Processing Techniques to Reduce Isoflavones in Soy-Based Foodstuffs. *Foods* **2023**, *12*, 1540. https://doi.org/10.3390/foods12071540

Received: 24 January 2023
Revised: 18 March 2023
Accepted: 28 March 2023
Published: 5 April 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Isoflavones (IFs) are polyphenols bearing both estrogenic and antiestrogenic traits, among other properties. It has been shown that they exhibit beneficial health effects, especially in people with an oestrogen deficiency, i.e., women in natural or artificial menopause [1,2]. Indeed, IFs at doses of 40 mg/day or more could prevent vasomotor symptoms associated with menopause. Some other meta-analysis suggested that phytoestrogens from soy could prevent bone loss in menopause via an estrogenic effect [3–5], and even more efficiently under aglycone forms [6].

However, alongside these potential beneficial effects, menstrual cycle impairments were observed in women exposed to IFs in doses of about 0.75 mg/kg bw (body weight)/day in aglycone forms [7,8]. More recently, IF intake at a level of 50 mg/day in women, i.e.,

about 0.8 mg/kg bw/day, was shown to be significantly correlated with luteal phase impairment and with the risk of nulliparity at 26 years old [9,10]. In addition, five clinical or population studies have so far reported the association of high soy IFs intake and their biomarkers in body fluids with a reduction in sperm count and quality [11–15]. When studies reported no effects on semen quality in men, it was at low doses [16] or for too short exposure times [17]. Overconsumption of soy has also been linked to several adverse clinical cases in women [18] and men [19]. In the former, on one hand, menstrual cycles were impaired and contraceptive pill treatments suffered from interference, and on the other hand, large uterus fibrosis and endometriosis features were observed. The effects regressed after soy arrest. In men, a secondary hypogonadism, which is the impairment of sexual behaviour and the feature of gynecomastia, was observed. The hypogonadism, characterised by dramatically reduced sex steroid concentrations in the blood, tended to return to normal levels after 3 months without soy. IFs have also been the centre of a great controversy regarding breast cancer, with several studies concluding in favour of a protective effect, mainly in Asian populations [20], while others reporting a proliferative impact [21]. It is though currently admitted, thanks to a better understanding of their mechanisms of action, that IFs have a dual effect: protective during the promotion phase (healthy cells), but proliferative on already declared oestrogen-dependent cancer cells. For instance, the activation of GPER, inducing cell proliferation, and the inhibition of protein tyrosine kinase, inducing cell apoptosis, were reported with low and high IF doses, respectively [22,23].

Nowadays, the main dietary source of IFs for humans is soy and derived soy by-products [24]. Modern western soy transformation processes sometimes expose consumers to high IF concentrations in food [25–27]. Considering the whole spectrum of effects recorded in humans, this potential dietary exposure may be considered a concern, and the precaution principle should be applied for people who exhibit a particular sensitivity to oestrogens, i.e., infants, children, pregnant women [20] and those with high risk of oestrogen-dependent breast or endometrial cancers. Moreover, hypothyroid patients taking medication, namely levothyroxine, should avoid soy IFs [28,29]. IFs should instead be kept for natural medication when proven useful. A perfect illustration of this observation can be found in the Chinese pharmacopeia, in which IFs are used for their medicinal properties only as extracted from Kudzu roots [30], with soy never being considered a health ingredient. This is likely because ancient recipes of Asian soy-based food included water soaking, cooking and simmering, a process that removes most IF content, thereby reducing any health impact.

Toxicological data have been available since 2008 on genistein (GEN), the most commonly studied phytoestrogen from soy. They were published by the US National Toxicology Program and dealt with the reprotoxicity [31] and carcinotoxicity [32] of this isoflavone. The no observable adverse effect level (NOAEL) could not be obtained from these studies, so a lowest observable adverse effect level (LOAEL) was alternatively proposed for both reproduction and carcinotoxic effects, which was 35 mg/kg bw/day in rats. Such a dose may help to determine a reference dose (RefD), i.e., a limit of exposure for humans. Additionally, as discussed by Lee et al. [26], even if endocrine disruptors can be more deleterious at low doses, due probably to the activation of different pathways, as mentioned above, there is no evidence of any adverse effects of low IF doses in vivo and in humans. Rather, it is doses over 20–40 mg/d (aglycone–glycosidic forms) that can exert physiological activities, either beneficial or deleterious, depending on the status of the consumer.

In this context, a previous work [27] has already demonstrated at the laboratory level the feasibility of a strong reduction in the IF content of soybeans by a specific water treatment. An important issue ensued, and may be phrased by the following question: how can Ifs' removal from soy matter be carried out on an industrial scale without degrading its nutritional quality and generating a heavy additional hurdle either? The purpose of the present investigation is (1) to provide a positive and convincing answer to this question by means of a pilot-scale study that could then be scaled up, (2) to mitigate the environmental

impact and supplementary costs through water and energy savings, and (3) to recover IFs from treatment water to feed the phytochemical industry. Indeed, as IFs are active compounds, they could and should be used for health applications in rigorously controlled supplementation. Therefore, collecting them in soy-based food factories would also reduce unwanted environmental exposure.

To do this, the approach utilised is as follows. (1) To determine the optimum experimental configuration of design factors those deemed relevant for IFs' removal [27] at the laboratory were first optimized by monitoring the content variations of the main soy IFs, GEN and daidzein (DAI) [33]. The individual impacts of these factors were evidenced by IF concentration measurements, taking advantage of the enzyme-linked immuno-sorbent assay (ELISA)'s high sensitivity and specificity. (2) To achieve improved treatment at pilot scale, the transposed data of the optimized laboratory factors were then readjusted. This was accomplished while exploring the impacts of unexpected (drenching) and relevant additional (granulometry, determined by laser diffraction and sieve analysis) factors. The whole process also implemented the supplementary constraint of the final step of drying the treated soybeans. (3) On another side, as the ELISA requires a rather long implementation, a faster assay method, i.e., Folin–Ciocalteu, was tested, since it appeared better suited to routine usage in an industrial environment. (4) The nutritional quality control of soy matter was performed by Phytocontrol® (Nimes, France) according to Dumas and Kjeldhal's methods for proteins and gravimetry for lipids. (5) To study the feasibility and quality of water recycling as well as IFS recovery, centrifugation, filtration and resin adsorption were tested on treatment waters. Appropriate statistical analyses were performed, i.e., signal to noise ratios for laboratory study and a Mann–Whitney test for pilot-scale data.

The results presented here are discussed under the realm of the potential exposure limit RefD, as defined above. The pilot-scale investigation resulted in the highlighting of a successful improved treatment that significantly reduced IF concentrations in soy.

2. Materials and Methods

2.1. Materials

All soy matter (whole beans, dehulled beans, pounded dehulled beans, soy cake and textured proteins) were supplied by Biopress® (Tonneins, France), except the dehulled beans for the laboratory-scale experiments, which were purchased in an organic store. The chemical reagents were acquired from MERK (Fontenay-sous-Bois, France) and Sigma-Aldrich (Saint Quentin Fallavier, France), unless otherwise stated. The samples and reagents' aqueous solutions were made in ultrapure water obtained from an Elga Veolia® (High Wycombe, UK) instrument.

2.2. Isoflavones Reduction at Laboratory Scale

Based on the conclusions of a previous work, the treatment of soy matter through water washing, the effectiveness of which depends upon five factors [27], was implemented. To adjust it, the latter were tested at five different levels, each completing a Taguchi L25 design of experiment [34,35]. For each attempt, 10 g of matter was added to the corresponding water volume needed to reach a soy matter:water volume ratio of 1:2, 1:2.5, 1:3, 1:4 or 1:5. The approximation that the soy matter density is 1 g/mL, regardless of the soy matrix, was made so that the sample volumes were easy to calculate. Treatments lasting 0, 5, 10, 20 and 30 min were carried out at room temperature (20–24 °C), 45 °C, 55 °C, 65 °C and 75 °C. The mixtures were stirred with either a magnetic or an orbital table stirrer at different speeds of 0, 50, 100 and 200 RPM (revolutions per minute). Each test was run in triplicate. The resulting soy samples were dried in a thermostatically controlled oven at 75 °C until a constant weight was reached, and then stored at −20 °C. The statistical analyses relied on the measurement of signal-to-noise ratios with XLSTAT software (Addinsoft_2020, XLSTAT, Paris, France).

2.3. Isoflavones Reduction at Pilot Scale

A batch of 6 kg of dehulled beans was washed using 24 L (volume ratio 1:4) of water at 20 °C for 10 min at 200 RPM in a 100 L tank. Another batch of the same weight was drenched before washing. For this purpose, the beans were spread out on a grid and exposed to 2 L of water sprayed through six nozzles for 13 s. The water used for both operations was kept close to 20 °C. As the batch which was both drenched and washed showed better results over the simply washed one, the corresponding assay was carried out three times.

The treated beans were then dried either by lick-drying (air temperature, velocity, and relative humidity equal to 55 °C, 2.25 m/s and 11%, respectively) or thick bed-drying (air temperature and velocity equal to 55 °C and 1.5 m/s, respectively).

To assess the achievement of satisfactory IF removal, the remaining IF concentrations in the treated dry soy matter were compared to values that have deleterious effects, as found in the literature.

2.4. Biochemical Analysis

2.4.1. ELISA

Sample Preparation

In order to carry out the IF analyses by ELISA, the samples were first subjected to the treatment described below. Thereby, 1 g of mortar-crushed soy-based material was dispersed into 50 mL of water by two-step stirring at room temperature for 20 min, and then at 100 °C for 10 min. After cooling down to room temperature, 500 µL samples were collected in triplicate under stirring to insure homogeneity. Washing- and drenching-water samples were also treated as 500 µL triplicates. Two millilitres of acetate buffer (sodium acetate 0.1 M; EDTA 0.14 M; 100 UI/mL penicillin G (Sigma, P-3032); and 0.1 mg/mL streptomycin (Sigma, S-6501)) at pH 5 were added to each sample vial, together with 10 µL of β-glucuronidase aryl-sulfatase from *Helix pomatia* (Roche®, 10127698001, Mannheim, Germany) to allow the digestion of glycosylated IFs. The samples were then incubated overnight at 37 °C under shaking [36]. Afterwards, the extraction of aglycone compounds was performed: 3 mL of acidified ethyl-acetate (500 µL HCl 38% per L) was added; the vial vortexed for 30 s, centrifuged at 500 g (Jouan™ CR3, Fisher Scientific, Illkirch, France) for 10 min at 4 °C and then stored at −22 °C to allow phase separation. The organic phase containing the IFs in aglycone forms was evaporated to dryness using a Speed-Vac (Thermo-Electron™ Corporation, Fisher Scientific, Bordeaux, France). Afterwards, each sample was diluted in 0.5 mL of assay buffer, i.e., phosphate-buffered saline (PBS) 0.01 M, 0.9% NaCl, 0.2% Tween, 1% DMSO, pH 7.3, and sonicated when required. Samples were stored at −22 °C until they were processed for IF analysis.

To assess the digestion and extraction recovery, the hydrolysis by β-glucuronidase aryl-sulfatase was monitored using genistin (EXTRASYNTHESE™, 1325 S, Genay, France), a pure compound used as a control reagent, in an external standard run in parallel to each measured sample. The compound was dissolved at 1 mg/mL in DMSO as the stock solution. In these operating conditions, the hydrolysis performance was always between 87 and 103%. Recoveries greater than 100% can be explained by the high accuracy of ELISA measurements and inter- and intra-assay variations.

Assay

GEN and DAI were assayed in soy matter and treatment water using an ELISA specific to each molecule, as explained by Shinkaruk et al. [37]. The primary antibodies were selected and obtained in previous works [38,39].

Characteristics of the ELISA

The soy matter samples were always diluted to 1/50 (*v*/*v*) for enzymatic digestion. The sensitivity of the GEN assay in food matter was 10 µg/mg. Similarly, sensitivity for DAI was 6.5 µg/mg. For water samples, the sensitivity was 100 ng/mL and 65 ng/mL

for GEN and DAI, respectively. Based on multiple tests, the intra-assay variation was never over 7%, and the inter-assay variation obtained on different microtitration plates was always below 17% [38,39]. Assays were not considered if the r of the sigmoid calibration curve was <98.5%. The final dilutions in the microtitration plates ranged from 1/50 to 1/800 for soy matter and 1/5 to 1/3200 for water samples. Each sample was assayed in triplicate, using three extracts and on three different microtitration plates. All values in this paper are given in aglycone equivalent.

2.4.2. Folin–Ciocalteu Assay

The phenolic compounds were assayed by the spectrophotometric Folin–Ciocalteu assay, within the wavelength range 750–765 nm [40]. The soy samples were treated to obtain the hydrophobic components. They were extracted twice for 2 h at 90 °C under stirring (magnetic stirrer, Ika Labortetechnik™, Staufen, Germany) with water:ethanol and then methanol:hydrochloric acid mixtures at 50:50 by volume. All resulting samples were then dried in a vacuum oven (Oven Memmert™, Schwabach, Germany) to obtain a dry matter reference which was weighed and then diluted back in water to obtain a final absorbance ranging between 0.5 and 1. The treatment water samples were centrifugated for 10 min at 10,000 RPM (Centrifuge Rotanta 460R™, Hettich®, Berlin, Germany) prior to assays. A standard curve was worked out using gallic acid (Sigma, 842649) at concentrations starting from 0 up to 250 µg/mL. For the measurements, 1000 µL of Folin–Ciocalteu reagent 1/10 (Supelco®, 109,001, Saint-Quentin-Fallavier, France) was added to 200 µL of either diluted soy matter extracts, water samples or standards. Each mixture was subsequently vortexed and then left to stand for 5 min away from light at room temperature, before adding 800 µL of Na_2CO_3 (75 g/L; Supelco®, 106,392, Saint-Quentin-Fallavier, France). All mixtures were shaken before being left to stand for 90 min away from light and at room temperature. Then, they were read at 760 nm with a spectrophotometer (Jenway™ 6305, Stone London, UK). The negative control did not contain phenolic compounds and was used as a 0 µg/mL sample to generate the standard curve.

2.5. Granulometric Testing Preparation

To assess the dependence of the IF removal process on the beans granulometric factor, dehulled beans were pounded in a cutter (Stephan, ProEx™, Hameln, Germany) to generate different particle sizes. The particles were sieved on a RETSCH™ (Eragny, France) AS200 sieve and subsequently weighed, before being separated into two batches which were washed for 5 min at 25 °C under magnetic or orbital table stirring at 200 RPM. The washing step in ultrapure water was performed at small scale in Erlen Meyer vials. The samples were kept frozen at −18 °C until they were required for further analysis.

2.6. Treatment Water Analysis and Processing

2.6.1. Quality Characterisation

Water samples were collected after the drenching-and-washing pilot process for IF analysis. Other parameters, namely dry matter, suspended matter (glass fiber membranes AP 40, Millipore, France)—according to NF EN 872—and turbidity (turbidimeter TL2360 Hach, Düsseldorf, Germany)—according to NF EN ISO 7027-1—were also measured. Particle size distributions were evaluated using laser diffraction (Cilas PSA 1190L Anton Paar™, Orleans, France) according to NF ISO 13320-1, with or without ultrasonic treatments.

2.6.2. Cleaning

Several tests were performed to elaborate the best operative protocol. First, samples of 200 mL each were subjected to centrifugation at 3000, 6000 and 9000 g for 1 min at 20 °C using a lab centrifuge (Hettich™, Tuttlingen, Germany). Based on the results obtained with the spin tests, 3.4 L of this water was centrifuged at 3000 g for 1 min, for solid–liquid separation, in order to maintain a material balance between the suspended matter and the polyphenols, among which are the IFs.

Another cleaning operation was then tested on 1 m^3 of the same water batch involving a 10-inch-high cartridge filter (PRECart PP 2, 499A110W071SP, PALL, New York, NY, USA) with a filtration area and a cutting of 0.5 m^2 and 1 µm, respectively. The pressure was fixed at a maximum of 5 bars, and the initial flow rate was set up at 300 kg/h/m^2. Cartridges were used in a BECO INTEGRA CART 1 carter (EATON Begerow® Product Line, Paris, France). The samples were characterized before and after the filtration operation.

2.6.3. Isoflavones Trapping and Recovery

Previous studies already showed the ability of synthetic resins to fix IFs [41]. Therefore, after cleaning of the pilot treatment water by PRECart PP 2 cartridge filtration, specific adsorption resins were tested for IF trapping and for water recycling.

As a pre-selection step, a jar test was performed, wherein 100 mL of water was exposed to 20 mg of resin under stirring for 2 h, before IFs were assayed. Ten different materials were tested to allow a preliminary selection: Lewatit® VPOC-1064 (Lanxess, Cologne, Germany), AD3008 (Jacobi, Paris, France), PAD900 (Jacobi, Paris, France), PAD950 (Jacobi, Paris, France), SEPLITE LXA838 (Sun Resin, Shaanxi, China), SEPLITE LXA94 (Sun Resin, Shaanxi, China), Relite EXA 90 (Resindion, Binasco, Italy) and Optipore™ SD-2, Amberlite™ XAD7HP, Amberlite™ FPX66, the last three supplied by Dupont® (Paris, France).

The two most promising resins were selected for open circuit column tests to determine the maximum volume of water that can be treated by the resin before performance loss (that is, the achievement of saturation), and to define the IFs' recovery conditions. For this final selection, 212 mL of the pre-selected resin was pre-wetted, then inserted into a 3 cm diameter column and gradually saturated with 4.2 L of water, representing 20 times the bed volume of the resin. Five samples were collected, each after passing 850 mL of filtrate through the column. The desorption function was checked using either demineralized water or ethanol, or mixtures of both in different volume proportions: 70:30, 50:50 and 30:70.

2.7. Statistical Treatments

For the laboratory scale testing, a Tagushi experimental design was chosen, as it allowed the evaluation of numerous variables with relatively few runs [34,35]. The statistical analyses relied on the measurement of signal-to-noise ratios with XLSTAT software (Addinsoft_2020; XLSTAT, Paris, France).

For pilot-scale experiments, each concentration corresponds to a mean value, and its standard deviation is derived from at least triplicate measurements. The significance was obtained by assessing the ν value according to a non-parametric Mann–Whitney test.

3. Results

3.1. Laboratory Scale

The results of the statistical analysis at laboratory scale can be found in Figure S1 in the Supplementary Materials, which exposes the five factors tested, and their values. It can be seen that only two factors, the temperature and the matrix form, seem to influence mainly the IF removal rate. As for the matter:water volume ratio, contact time and stirring speed, their variations did not result in significant differences in the IF removal rate ($p > 0.05$). The latter appears to be optimal at room temperature and for the smallest particles. With regard to the effect of the nature of soy matter, the data indicate that pounded dehulled beans give the best IF removal result, followed by soy cake (formed from crushed beans) and finally dehulled beans. However, the first two were not retained, because firstly, pounding was not seen as a manageable solution at pilot or industrial scale, and secondly, the treatment water in the case of soy cake was highly turbid, indicating that a large quantity of soluble matter had leaked into it.

As shown in Figure 1A, at laboratory scale, the mean IF removal from raw soybeans was 14.7 ± 5.23%, while it was 42.5 ± 7.56% for dehulled beans. This effect was obtained regardless of the other experimental conditions, in line with the statistical Taguchi analysis

(Figure S1), from which the type of soy matter was identified as one main influencing factor. Indeed, it was not possible to significantly reduce IF content in whole soybeans, while up to 60% of the initial GEN could be removed by working on dehulled beans at room temperature for 10 min. It appeared that the beans' envelope could limit extraction of IFs.

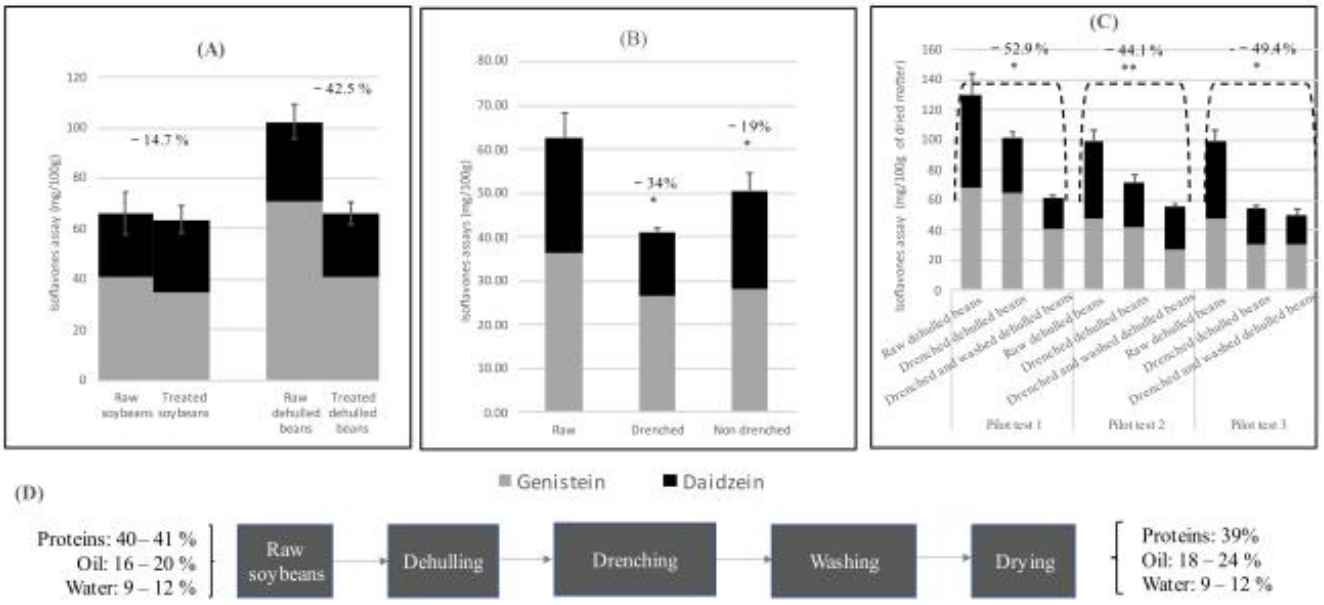


Figure 1. Results of isoflavone removal at laboratory and pilot scales. **(A)** Laboratory-scale average isoflavone removal from raw and dehulled soybeans. **(B)** Pilot-scale mean isoflavone removal from dehulled beans, with and without drenching before washing. **(C)** Pilot-scale mean isoflavone removal after application of the improved process on three pilot trials. Percentages correspond to the isoflavone reductions. Error bars are SEM. Significant differences for a risk α of 5% (*) and 1% (**) are also indicated. **(D)** Full improved pilot process overview with input and output matter characteristics.

3.2. *Pilot Scale*

As established by the laboratory investigation, the following were retained for pilot-scale operations: dehulled soybeans as a starting material, and room temperature, since no heating was required. The matter:water volume ratio had to be fixed at 1:4 due to a technical limitation leading to deficient homogenization for the least-diluted mixtures, knowing that the 1:5 case did not bring any improvement. The water/soy contact time was also fixed at 10 min, since 5 min was not sufficient to obtain the expected results. As for the stirring speed, it had to comply with technical constraints.

Besides, during the first transposition of the trial to pilot scale, the washing water was turbid due to dust generated by the dehulling step, and the IF elimination efficiency was less than expected based on laboratory tests. This is the reason that a drenching step was added before washing. Furthermore, the dehulling process was shown to generate heterogenous particle sizes. To check the influence of the particle size on the IF removal efficiency, a comprehensive analysis was performed.

Additionally, it should be noted here that drying is a mandatory step in the production of soy-based matter when water pre-treatment is applied, and consequently, it is crucial to restrict the corresponding allotted time as much as possible. Soy-based materials cannot be extruded or sold to a food manufacturer if their moisture levels are higher than 12%.

3.2.1. Effect of Drenching

Figure 1B shows the difference between IF removal rates from dehulled beans that were either drenched or not, prior to washing. Indeed, the soy lost almost twice as

many IFs (40.1 ± 7.23 mg/L instead of 25.8 ± 2.54 mg/L) if submitted to the additional step. The differences were significant for total IFs, GEN and DAI at 0.5%, 0.1% and 0.05%, respectively.

Figure 1C shows the result of the pilot trial performed three times. It can be seen that IFs were successfully reduced by 52.9, 44.1 and 49.4% (mean 48.8 ± 4.43%), respectively. The efficiency was better for DAI. The differences in IF removal between the three pilot batches are non-significant, thus validating the process of IF removal at pilot scale.

3.2.2. Effect of Granulometry

The industrial process of dehulling generated different types of particles. Their sorting by size gave the following proportions in number: 11.53 ± 1.09% whole beans and 88.47 ± 1.16% half-beans. When the proportion was given by mass, whole beans represented 14.35 ± 2.50%, dehulled beans 72.25 ± 3.44%, seed fragments 13.32 ± 4.92% and envelope residues 0.09%. The morphological analysis of the particle sizes generated by dehulling showed solid densities of 1.2–1.4 and 1.2–1.8 whether pounded or not, respectively. It was highest for the fractions 1–2.24 mm and 0.5–1 mm for pounded and unpounded beans, respectively. The details of the particle size sorting are given in Figure 2, with a 2D photograph for each fraction. It can be seen that the particles have an ellipsoidal shape which is more pronounced as the size decreases. Pounding made the particles' shapes more irregular. The particle sizes generated have dimensions either over 4 mm, between 4 and 2.24 mm, or between 2.24 and 0.5 mm. The 4–6.23 mm fraction corresponds to the most represented mass, i.e., 94% of the total mass of particles recovered on the sieves.

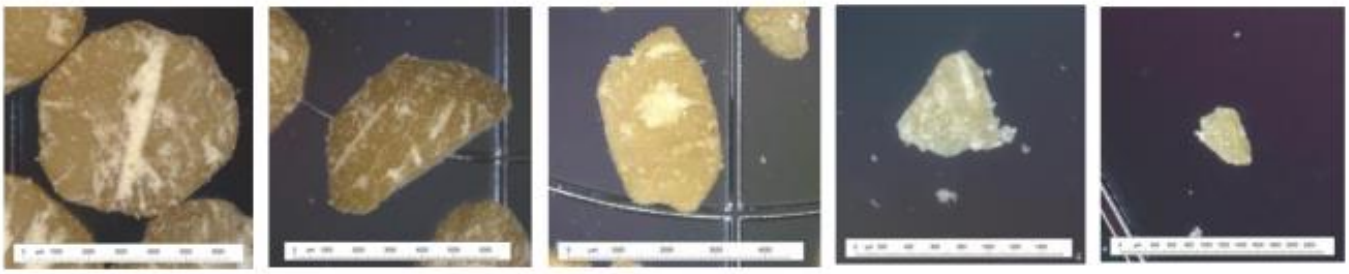


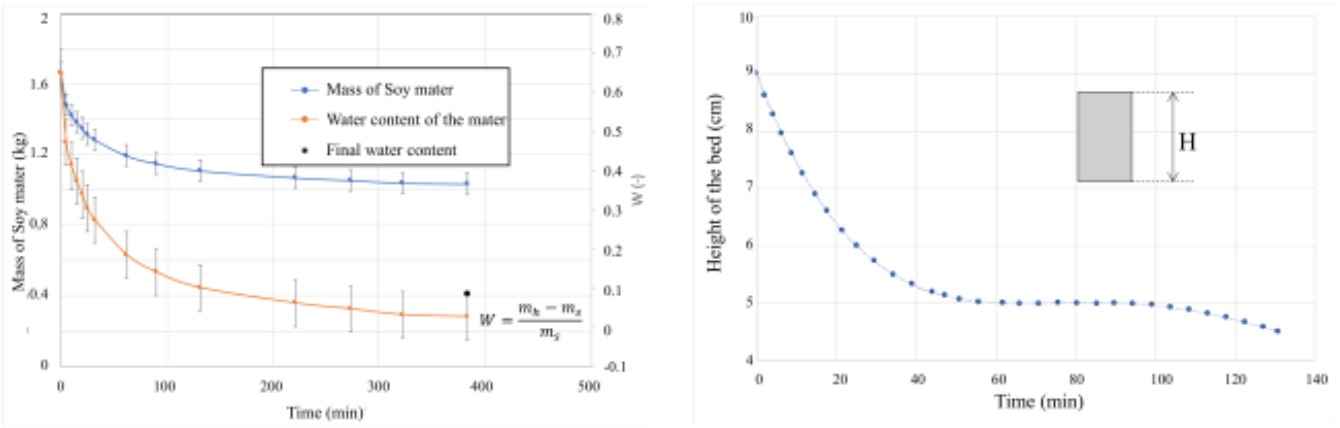
Figure 2. Granulometry of the fractions obtained by sieve analysis on dehulled beans with 2D photographs.

With regard to the impacts of granulometry on the process, the following correlations were observed. The washing effectiveness on deflavonoidisation was slightly better with magnetic than orbital stirring, but the difference was only significant with the smallest particle sizes. The IF content decrease between raw and washed matter was significant, regardless of the particle size. It was also noted that the smaller the particle size, the more efficient the deflavonoidisation, independent of the hydration of the samples, since the values were calculated according to the dry matter (see Figure S2 in the Supplementary Materials). However, in return, the drenching waters were all the more contaminated by solids, the smaller the particle size. This indicates a loss of water-soluble substances that are also present in soybeans, i.e., proteins, sugars, fibres, etc.

3.2.3. Drying of Treated Dehulled Soybeans

The required drying time ranged from 1 to 2.5 h, depending on the technique tested and the operating conditions. Thick bed-drying was more effective than lick-drying. The main results obtained on the thick bed are gathered in Figure 3, and also in Figure S5 in the Supplementary Materials. In this case, the mandatory 12% moisture could be achieved in about 1 h. Drenching and washing caused swelling of the dehulled beans by water which

was then reabsorbed (shrinkage) upon drying, causing the pile volume to decrease sharply (Figure 3B).



(**A**) (**B**)

Figure 3. (**A**) Water and mass evolutions during the thick bed-drying process. (**B**) Evolution of the bed height during drying. Operating conditions: air temperature 55 °C; air velocity 1.5 m/s; initial water activity (Wini) 65.0% and final water activity (Wfin) 8.3%.

3.2.4. Nutritional Quality

A sketch of the improved processing steps for IF removal at pilot scale is presented on Figure 1D. It can be seen that the protein, oil and water contents of dehulled beans were equivalent before and after being subject to the whole treatment. The protein content went from 40 to 39%. The oil proportion increased from 16–20% to 18–24% (with inter-batch variation), most probably because the removal of the envelope essentially reduced the seeds' carbohydrate content [42]. Additionally, the targeted water content of 12% was successfully achieved after the drying process, as described above.

In short, the whole process, including drenching and washing dehulled beans with water (20 °C, 10 min, 1:4 matter:water volume ratio and 200 RPM) and then drying them for 1 hour, from now on referred to as the improved process, succeeded in maintaining an important part of the nutritional quality of soybeans, especially the protein and lipid contents.

3.3. Comparison of ELISA and Folin–Ciocalteu Method

Figure 4A represents the correlation between IF concentrations obtained from soybean extracts by ELISA and the phenolic compound concentrations measured by the Folin–Ciocalteu method. It can be seen that the data points are completely dispersed, without any apparent correlation. In fact, it is impossible to quantify the reduction of IFs in soy matter using the Folin–Ciocalteu method. On the contrary, in Figure 4B, which represents the ELISA results plotted against the Folin–Ciocalteu data obtained from treatment water samples, the correlation is strong (R = 0.995). Interestingly, and confirming this result, the correlation between the decrease rate of the total IFs in dehulled beans determined by the ELISA and the polyphenol content in treatment water determined by the Folin–Ciocalteu method is significant, as shown in Figure 4C. Additionally, it is noted that using magnetic stirring, the R correlation coefficient is 0.946, while as mentioned previously, data obtained using orbital stirring are inconsistent.

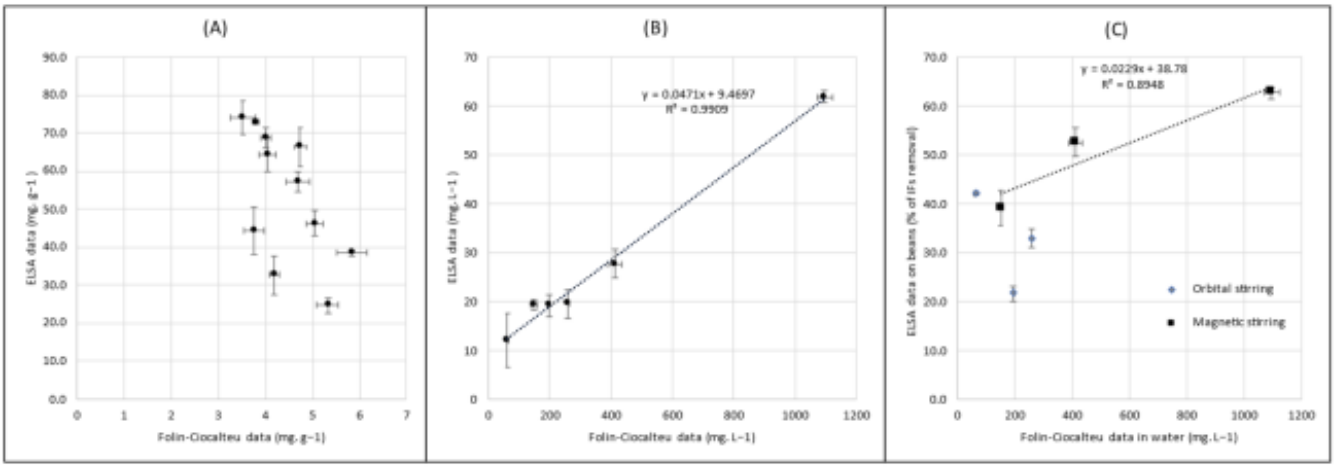


Figure 4. Comparisons between the ELISA and the Folin–Ciocalteu methods for isoflavone content determination: (**A**) in soybeans, a lack of correlation; (**B**) in water, the presence of correlation; (**C**) between IF removal from soybeans (ELISA) and polyphenols in water (Folin–Ciocalteu), the presence of correlation. Error bars are SEM.

3.4. Treatment-Water Analysis

The mixture made of the drenching and washing waters of dehulled beans contained 2158 mg/L of suspended matter with a turbidity of 4195 FNU (Formazin Nephelometric Unit). Laser diffraction analysis highlighted particle sizes ranging from 0.1 µm to 300 µm [d_{10}: 6.35 ± 2.40 µm; d_{50}: 106 ± 8.91 µm; d_{90}: 192 ± 24.1 µm; d_{max}: 400 ± 71 µm.], considering that the few large particles present masked the smallest ones, and most agglomerates were dissociable under mechanical stress. The results of the first centrifugations (200 mL samples) are presented in Figure S3 in the Supplementary Materials. The different suspended matter concentrations obtained at different speeds ranged from 309 mg/L to 217 mg/L in the supernatant. Therefore, the centrifugation step allowed the clarification of the aqueous phase. The pellets obtained after the second centrifugation (3.4 L sample) had a volume of 60 mL. The pellets contained soybean fragments, and therefore proteins and other insoluble soy particles. The percentage of IFs in the dried extract was found to be 13.02 ± 2.27%.

The quality of the drenching and washing water mixture was significantly improved by membrane filtration on a PRECart PP 2 cartridge. This test was performed on raw water without centrifugation pre-treatment. The initial suspended matter concentration was 2158 mg/L and was reduced by a factor of 100 to reach a residual value of 21 ± 11 mg/L after cartridge filtration. Considering the water sample tested, the cartridge filtration, for which the maximum pressure was set on 5 bars, started at a flow rate of 300 kg/h. The filtration flow was decreased progressively to 18 kg/h after 33 min. The amount of water that was filtrated during this period was 22.22 kg. Two filtrates were collected, one (filtrate 1) before pressure stabilization at 5 bars and the other (filtrate 2) when the pressure was kept at its maximum (Figure S4A in Supplementary Materials). The flow was maintained at a constant level of 300 kg/h for 6 min. A mass balance based on measurements performed before and after filtration was carried out for IF concentrations. The data seen in Figure S4B show that 83.33% of all IFs were trapped on the cartridge filter, representing 83.42% for DAI and 82.56% for GEN, specifically. In total, the 22.22 L of treatment water left 45.8 mg of IFs on the cartridge. In parallel, the suspended matter was reduced from 2158 mg/L to 25 mg/L, leaving a mass of 47,300 mg of solids on the filter. Considering these figures, the cartridge extract contained only 0.01% of all IFs.

3.5. Isoflavones Recovery

The pre-selection results, as described in Section 2.6.3, highlighted evidence of polyphenol retentions ranging from 30% for acrylic resins to 40–60% for conventional styrenic resins,

the performance depending on the pores' volumes, diameters, and specific surface areas, as shown in Figure S6 in the Supplementary Materials. Remarkable performances were obtained with functionalized resins recently developed by DUPONT™ and SUN RESIN™, with polyphenol reductions over 95%. The various resins tested were all effective in trapping IFs; some of them, such as VPOC resins 1064, Relite EXA 90 and Optipore SD-2, made it possible to lower the water's residual IF content below the ELISA's limits of quantification. They correspond to resins with pore diameters between 5 and 10 nm and a styrene-divinylbenzene (DVB) base. Unlike the first two references mentioned, the Optipore SD-2 resin also fixed almost all polyphenols, probably because of its functionalization with an amine group.

Consequently, the choice of the resin to be tested at pilot scale went naturally towards the one most effective for trapping organic compounds present in water, which was Optipore resin SD-2. However, it was likely that the adsorbed compounds would be difficult to desorb due to its functionalization. Thus, a second resin had to be selected in order to satisfy to both the most effective trapping and an adequate desorption. The candidate resin which appeared to fulfil these two constraints simultaneously was Amberlite® FPX 66. It is a non-grafted resin of the same base (styrene-DVB), from the same supplier, which ended up being very effective in trapping IFs (98.7%), even if it trapped only 50% of the total polyphenols.

All polyphenols were recovered from Amberlite® FPX 66 resin by a 50:50 water:ethanol mixture, whereas even with 100% ethanol, only 40% were recovered from the Optipore™ SD-2 resin. Concerning the total IFs, the trend is the same and it can be noted that GEN was less readily eluted than DAI, as shown in Figure 5B. A 70:30 water:ethanol mixture did not have sufficient eluting strength in both cases, and a minimum of 50% and 100% alcohol was required for Amberlite® FPX-66 and Optipore™ SD-2 resins, respectively.

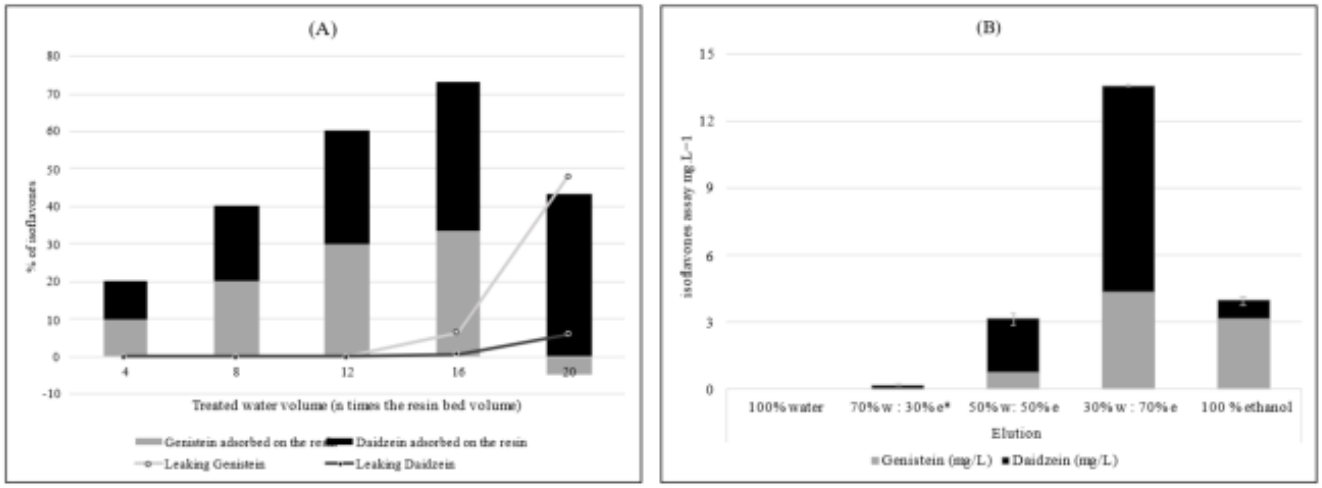


Figure 5. Isoflavones adsorption (**A**) and desorption (**B**) from Amberlite FPX 66® resin. The curves (**A**) represent progressive isoflavones leaking from the resin. * w and e refer to water and ethanol, respectively.

Tests on Optipore™ SD-2 resin showed that the polyphenol content of the first filtrate was low (4 mg/L), which confirms the IF-trapping yields obtained in the jar test. The exhaustion of the resin was gradual and only partial, since after filtration of 32 times the bed volume, the suspended matter was only 17 mg/L, which corresponds to a 62% reduction in total polyphenols. The Optipore™ SD-2 resin fixed 223 mg of polyphenols, i.e., a fixing capacity of 1.05 mg/g, without reaching its maximum saturation capacity. For the IFs (44% GEN, 56% DAI), a quantity of 5.73 mg (i.e., 0.027 mg/g) was fixed without reaching saturation.

Regarding the Amberlite® FPX66 resin, during the filtration of 20 bed volumes, the polyphenol content fluctuated between 20 and 25 mg/L, and the reduction yields for polyphenols ranged between 44 and 56%, in line with the results of the jar test. In total, the Amberlite® FPX 66 resin fixed 112 mg of polyphenol, i.e., a fixing capacity twice as low (0.52 mg/g) as Optipore™ SD-2 resin, without reaching its maximum saturation capacity. For IFs, the assay results showed that the Amberlite® FPX 66 resin was less selective, since resin exhaustion appeared after passage of, at most, 16 bed volumes for DAI and GEN. The maximum saturation capacity of the Amberlite® FPX 66 resin was therefore reached for IFs at 2.865 mg (i.e., 0.01 mg/g), which represented 73% of remaining IFs in the treatment water.

4. Discussion

4.1. Isoflavone Removal from Soybeans

Previous experiments have shown that a 60 min precooking of pounded dehulled soybeans at 90 °C successfully removed 54.32 ± 12.35% of the initial IFs [27]. The present process succeeded in proving that a treatment of only 10 min at 20 °C was sufficient, after drenching, to reach similar IF removal rates (48.8 ± 4.43%) from dehulled soybeans.

The drenching step of dehulled beans was added at pilot scale for cleaning the seeds before deflavonoidisation, as the dehulling and light grinding favoured the appearance of a superficial layer of fat, which trapped a large part of the dust and flour. The latter two, containing IFs, had to be eliminated before washing in order to improve the contact between water and the matrix surface, helping to achieve a better deflavonoidisation. Indeed, both GEN and DAI are mainly present in a glycosidic conjugated water-soluble form (logP < 1) in soybean seeds [43]. This finding was confirmed by IFs' assay in water. Drenching helped maintain a good concentration gradient between dehulled beans and water, and improved the effectiveness of IF leaching.

Surprisingly, IF removal was better at 20 °C than at higher temperatures. This could be partly explained by the significant decrease in soybean solid content when the temperature rises from 30 to 40 °C [44]. The hypothesis is that proteins and carbohydrate leakage could prevent IF removal. Another explanation may be that cooking at high temperatures could precipitate the most external proteins of the dehulled beans. Such precipitation could reduce the IFs' transfer to water. This has not been reported in the literature beforehand, thus additional experiments are needed to help understand this result.

4.2. Influence of Granulometry

The goal of these measurements was to show the influence of the seeds/the water exchange surface on IFs' extraction from dehulled beans. Indeed, the data indicate that the smaller the particle size, the larger the specific exchange surface and the greater the IF removal. However, this was true only for the magnetic stirring case. The difference observed between the two mixing techniques on the efficiency of IF removal could be explained by considering that the orbital shaking at 200 RPM was not able to ensure good contact between water and soy matter. Indeed, it is the orbital shaking that generated matter areas that remained dry in the testing bottle. Consequently, the IF measurements were not consistent when compared to those obtained using magnetic stirring.

However, this better IF removal from smaller particles was counterbalanced by a higher leakage of nutritive elements into the water (high turbidity with the smallest particles). Therefore, dehulled beans were considered a better starting material.

4.3. Comparison of ELISA and Folin–Ciocalteu Methods

As already mentioned, comparing the ELISA and the Folin–Ciocalteu assay makes sense given the time involved, including the extraction steps, in each case. Thus, after sample collection, the latter technique gives a result in a few hours, while the ELISA requires at least three whole days and one night.

The results of the correlation study showed that the Folin–Ciocalteu method was not able to assess the reduction of IFs in soy matter. The reason is undoubtedly linked to the presence of many insoluble polyphenols in soy products, and to the fact that the amount of water fixed on the material changes with the particle size (more water fixed on the smaller particles). The main discrepancy observed between the two methods is based on differences in their specificity. Indeed, the ELISA used in this study targets specifically GEN or DAI [38,39] and the extraction process triggers the water-soluble fraction, whereas the Folin–Ciocalteu method recognizes all polyphenols based on the chemical reaction of phenol rings with the assay reagents. The extraction performed prior to the analysis also includes non-water-soluble compounds. However, soybeans contain many phenolic compounds, some of which are soluble in water, some of which are not. Indeed, insoluble tannins, saponins, and also proteins, amino-acids, sugars and vitamins, among other food components, were reported to be detected with a Folin–Ciocalteu assay [45]. Thus, it was not really surprising to observe strong discrepancies between the two assay techniques.

On the contrary, the very good correlation between the ELISA and Folin–Ciocalteu results on water analysis was crucial, as it confirmed that the Folin–Ciocalteu method, by measuring water-soluble phenolic compounds, allowed an indirect evaluation of the total IF content in the treatment water samples. As the IF content in these waters is correlated with that in the corresponding treated soybeans, the Folin–Ciocalteu method seems adequate to simply and quickly evaluate the effectiveness of the IF reduction process.

Interestingly, and confirming this result, the correlation between the reduced concentration of IFs in dehulled beans found by ELISA and the amounts of polyphenols in treatment waters determined by the Folin–Ciocalteu assays was significant. A plausible explanation could be that after centrifugation, the water only contains soluble polyphenols, and so potential interference is reduced. Consequently, the Folin–Ciocalteu method could be advantageously used to quickly compare the efficiency of water treatments of different processes designed for IF removal.

It is important to acknowledge here that there is also a sensitivity difference between the two methods. The ELISA's sensitivity is 1 ng per well, and the detection limit is 0.48 ng per well. The 50% competition points vary from 8 to 50 ng/mL, depending on the assays. The Folin–Ciocalteu standard curve linearity lies between 35 and 250 µg/mL of gallic acid. The detection limit is 0.863 µg/mL, and the sensitivity is 0.0073 OD/µg/mL. These data are given in equivalent gallic acid.

4.4. Isoflavones Collection

The resin open-circuit column test with Amberlite® FPX66 allowed the trapping of almost 80% of the remaining IFs in pre-filtered treatment water. However, considering the number of IFs that could be present in general in treatment waters, the fixation capacity of the resin might not be sufficient alone.

Luckily, IFs and other flavonoids are known to adsorb to proteins in their native or glycosidic forms [46]. Therefore, it was not surprising to find them in the pellets after centrifugation. The percentages of IFs in the dried extract (13.02 ± 2.27%) rank within currently commercialized IF extracts that can be found on the pharmaceutical market (2, 10, 20 or 40%). The recovery of dried IF extracts still needs to be tested at an industrial scale, as it can represent a significant quantity. This was not the case for the extract obtained by cartridge filtration, as it contained only 0.01% of IFs. The difference between the two separation techniques (centrifugation and cartridge filtration) comes from their respective cutting rates. With cutting at 1 µm diameter, the filtration cartridge retained more matter than the centrifugation step. In addition, the affinity of the IFs for the PRECart PP 2 cartridge may be low, favouring the water solubility of the glycosylated IFs.

Thus, to obtain the richest possible IF extract, it might be advantageous to combine the pellets produced by the centrifugation with the IFs recovered from the resin treatment of the supernatant. This alternative route should potentially lead to full recovery of IFs from treatment water.

4.5. Safety Limits for Isoflavones

The US National Toxicology Program demonstrated in two multigenerational reproductive and carcinogenesis studies that exposure to GEN caused mammary and pituitary gland tumours in female rats, and reproductive impairments in both sexes [31,32]. The efficient daily dietary exposure was 35 mg/kg body weight, thereby defining the preliminary toxic levels in animals. It has been admitted in classical toxicology that in order to translate a minimal toxic dose of a substance from animals to humans, it has to be divided by the product of various safety factors that range from 1 to 10 [47–49]. In the case of GEN, three safety factors must be considered. The first is a factor of extrapolation from a LOAEL (lowest observable adverse effect level, which is available for GEN) to NOAEL (no observable adverse effect level). The second is a factor of extrapolation between animals and humans. The third is a factor counting for sensitivity differences among humans. If the commonly admitted values are applied, this brings out a total of 300, even if 184 could be safe enough, as argued by Kodell & Gaylor [48]. Regarding soybeans, the NTP assessed only GEN toxicity, thus the question regarding other estrogenic IFs remains unanswered. Because soybean also contains DAI that can be metabolised into a more potent estrogen, i.e., equal [50], it would be safer, in the authors opinion, to consider the toxicity of total IFs. Since no data are available on DAI toxicity, it seems adequate to apply the highest safety factor to total IF measurements, i.e., 300. Such a decision will bring the IF limit for an adult weighing 60 kg to 20 mg/day, and for a 20 kg child to 6.7 mg/day.

In previous works, multiple soy-based foods from France were analysed for their GEN and DAI contents [25,26]. Several of them, including roasted soybeans, soy milk, soy galettes, and soy cheese or tofu, could bring up to 50 mg or more per portion. Such amounts clearly justify the work presented here. In this study, we managed to reduce dehulled beans' IF levels from 90–100 mg/100 g to 40–52 mg/100 g. Dehulled beans are usually texturized in soy proteins. If it is assumed that texturization does not change the IF content of textured proteins, the latter could be further treated to achieve acceptable IF levels in final foodstuffs, if required. Luckily, this is starting to be the case in France, and at an industrial scale. Repeated water rinsing of textured soy proteins instead of soybeans can reduce their IF content by up to 90%. One plausible explanation for this high IF removal efficiency could be the protein conformational and molecular changes that occur during texturization, as the protein water absorption has been found to increase with denaturation, due to the formation of a protein matrix that is stabilized by hydrophobic interactions [51]. This water treatment of textured proteins was a preliminary result obtained during the implementation of a new industrial recipe, but it failed to reach the RefD. The implementation of drenching and washing of dehulled beans, on the one hand, and rinsing of textured proteins, on the other, would give rise to highly safe soy-based food that may be eaten without any restriction.

5. Conclusions

The present work reports mainly on the study of pilot processes tested to remove IFs from soybeans intended for the preparation of processed foods. The entirely developed process was found to be not only effective in reducing the IFs in soy without decreasing its nutritional benefits, but also water and energy efficient. Indeed, it mainly involves water rinsing and drenching of dehulled beans at room temperature. In addition, as the water used for these treatments should be cleaned up, it was shown that collecting IF-enriched extracts for pharmaceutical applications through cost-effective and easy-to-implement techniques is possible. The results obtained should contribute to the evolution of regulations on the toxicological values of estrogenic IFs. Consumers will thus benefit from the best nutritional values of soy without the potential undesirable effects. However, to fully reach this aim, the improved process developed here must be further tested in association with other water treatment steps on textured soy proteins. All these results remain to be validated at industrial scale.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/foods12071540/s1. The following supporting information are openly available in supplementary material. Figure S1. Results obtained on tests performed at laboratory scale following the Tagushi L25 experimental design. Figure S2. Test of isoflavone removal by water washing. Figure S3. Water-cleaning operation by centrifugation. Figure S4. Effect of the cartridge filtration on isoflavones' quantity in treatment water. Figure S5. Characterisation of the thick bed-drying process. Figure S6. Isoflavone measurements in water samples incubated for two hours with different resin types.

Author Contributions: S.B. took part in the experiment design, implemented and performed the ELISA isoflavones measurements and analysis and was in charge of the manuscript writing which included contributions from all collaborators. C.B.-P. initiated and supervised the analytical work and was involved in the assays, the design of experiment, the data analysis and manuscript writing. M.V. supervised the drying experiments and helped in the scientific consortium management. C.F. was in charge of the Folin–Ciocalteu analysis. P.G. designed and supervised all the experiments dealing with the water recycling. F.C. managed the soybeans treatments at pilot scale, i.e., the drenching and washing operations. G.P. was in charge of the funding acquisition, resources and writing of the original draft. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Nouvelle Aquitaine Region, grant number 2020.809.CP.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding authors. The data are not publicly available due to confidentiality.

Acknowledgments: We thank the Pôle Aliments et Santé, Bordeaux University (ARNA), BIOPRESS and its Agen subcontractors (Agrotec, I2M, IFTS, INRAE), INSERM and Bordeaux Science Agro. We also thank Gabriele Peruzzi for his contribution to the funding acquisition and valuable contribution to the work. This research was possible thanks to the support of the Regional Council of Nouvelle Aquitaine.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Chen, M.N.; Lin, C.C.; Liu, C.F. Efficacy of Phytoestrogens for Menopausal Symptoms: A Meta-Analysis and Systematic Review. *Climacteric* **2015**, *18*, 260–269. [CrossRef]
2. Franco, O.H.; Chowdhury, R.; Troup, J.; Voortman, T.; Kunutsor, S.; Kavousi, M.; Oliver-Williams, C.; Muka, T. Use of Plant-Based Therapies and Menopausal Symptoms: A Systematic Review and Meta-Analysis. *JAMA* **2016**, *315*, 2554–2563. [CrossRef] [PubMed]
3. Akhlaghi, M.; Ghasemi Nasab, M.; Riasatian, M.; Sadeghi, F. Soy Isoflavones Prevent Bone Resorption and Loss, a Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Crit. Rev. Food Sci. Nutr.* **2020**, *60*, 2327–2341. [CrossRef] [PubMed]
4. Jin, X.; Sun, J.; Yu, B.; Wang, Y.; Sun, W.J.; Yang, J.; Huang, S.H.; Xie, W.L. Daidzein Stimulates Osteogenesis Facilitating Proliferation, Differentiation, and Antiapoptosis in Human Osteoblast-like MG-63 Cells via Estrogen Receptor–Dependent MEK/ERK and PI3K/Akt Activation. *Nutr. Res.* **2017**, *42*, 20–30. [CrossRef]
5. Wei, P.; Liu, M.; Chen, Y.; Chen, D.C. Systematic Review of Soy Isoflavone Supplements on Osteoporosis in Women. *Asian Pac. J. Trop. Med.* **2012**, *5*, 243–248. [CrossRef]
6. Lambert, M.N.T.; Hu, L.M.; Jeppesen, P.B. A Systematic Review and Meta-Analysis of the Effects of Isoflavone Formulations against Estrogen-Deficient Bone Resorption in Peri- and Postmenopausal Women. *Am. J. Clin. Nutr.* **2017**, *106*, 801–811. [CrossRef]
7. Cassidy, A.; Bingham, S.; Setchell, K.D. Biological Effects of a Diet of Soy Protein Rich in Isoflavones on the Menstrual Cycle of Premenopausal Women. *Am. J. Clin. Nutr.* **1994**, *60*, 333–340. [CrossRef]
8. Nagata, C.; Oba, S.; Shimizu, H. Associations of Menstrual Cycle Length with Intake of Soy, Fat, and Dietary Fiber in Japanese Women. *Nutr. Cancer* **2006**, *54*, 166–170. [CrossRef]
9. Andrews, M.A.; Schliep, K.C.; Wactawski-Wende, J.; Stanford, J.B.; Zarek, S.M.; Radin, R.G.; Sjaarda, L.A.; Perkins, N.J.; Kalwerisky, R.A.; Hammoud, A.O.; et al. Dietary Factors and Luteal Phase Deficiency in Healthy Eumenorrheic Women. *Hum. Reprod.* **2015**, *30*, 1942–1951. [CrossRef]
10. Jacobsen, B.K.; Jaceldo-Siegl, K.; Knutsen, S.F.; Fan, J.; Oda, K.; Fraser, G.E. Soy Isoflavone Intake and the Likelihood of Ever Becoming a Mother: The Adventist Health Study-2. *Int. J. Women's Health* **2014**, *6*, 377. [CrossRef]

11. Chavarro, J.E.; Toth, T.L.; Sadio, S.M.; Hauser, R. Soy Food and Isoflavone Intake in Relation to Semen Quality Parameters among Men from an Infertility Clinic. *Hum. Reprod.* **2008**, *23*, 2584–2590. [CrossRef]
12. Toshima, H.; Suzuki, Y.; Imai, K.; Yoshinaga, J.; Shiraishi, H.; Mizumoto, Y.; Hatakeyama, S.; Onohara, C.; Tokuoka, S. Endocrine Disrupting Chemicals in Urine of Japanese Male Partners of Subfertile Couples: A Pilot Study on Exposure and Semen Quality. *Int. J. Hyg. Environ. Health* **2012**, *215*, 502–506. [CrossRef]
13. Xia, Y.; Chen, M.; Zhu, P.; Lu, C.; Fu, G.; Zhou, X.; Chen, D.; Wang, H.; Hang, B.; Wang, S.; et al. Urinary Phytoestrogen Levels Related to Idiopathic Male Infertility in Chinese Men. *Environ. Int.* **2013**, *59*, 161–167. [CrossRef]
14. Mumford, S.L.; Kim, S.; Chen, Z.; Barr, D.B.; Louis, G.M.B. Urinary Phytoestrogens Are Associated with Subtle Indicators of Semen Quality among Male Partners of Couples Desiring Pregnancy. *J. Nutr.* **2015**, *145*, 2535–2541. [CrossRef]
15. Yuan, G.; Liu, Y.; Liu, G.; Wei, L.; Wen, Y.; Huang, S.; Guo, Y.; Zou, F.; Cheng, J. Associations between Semen Phytoestrogens Concentrations and Semen Quality in Chinese Men. *Environ. Int.* **2019**, *129*, 136–144. [CrossRef]
16. Povey, A.C.; Clyma, J.A.; McNamee, R.; Moore, H.D.; Baillie, H.; Pacey, A.A.; Cade, J.E.; Cherry, N.M. Phytoestrogen Intake and Other Dietary Risk Factors for Low Motile Sperm Count and Poor Sperm Morphology. *Andrology* **2020**, *8*, 1805–1814. [CrossRef]
17. Beaton, L.K.; McVeigh, B.L.; Dillingham, B.L.; Lampe, J.W.; Duncan, A.M. Soy Protein Isolates of Varying Isoflavone Content Do Not Adversely Affect Semen Quality in Healthy Young Men. *Fertil. Steril.* **2010**, *94*, 1717–1722. [CrossRef]
18. Chandrareddy, A.; Muneyyirci-Delale, O.; McFarlane, S.I.; Murad, O.M. Adverse Effects of Phytoestrogens on Reproductive Health: A Report of Three Cases. *Complement. Clin. Pract.* **2008**, *14*, 132–135. [CrossRef]
19. Imai, H.; Nishikawa, H.; Suzuki, A.; Kodama, E.; Iida, T.; Mikura, K.; Hashizume, M.; Kigawa, Y.; Tadokoro, R.; Sugisawa, C.; et al. Secondary Hypogonadism Due to Excessive Ingestion of Isoflavone in a Man. *Intern. Med.* **2022**, *61*, 2899–2903. [CrossRef]
20. Nakai, S.; Fujita, M.; Kamei, Y. Health Promotion Effects of Soy Isoflavones. *J. Nutr. Sci. Vitam.* **2020**, *66*, 502–507. [CrossRef]
21. Shike, M.; Doane, A.S.; Russo, L.; Cabal, R.; Reis-Filho, J.S.; Gerald, W.; Cody, H.; Khanin, R.; Bromberg, J.; Norton, L. The Effects of Soy Supplementation on Gene Expression in Breast Cancer: A Randomized Placebo-Controlled Study. *J. Natl. Cancer Inst.* **2014**, *106*, dju189. [CrossRef]
22. Russo, M.; Russo, G.L.; Daglia, M.; Kasi, P.D.; Ravi, S.; Nabavi, S.F.; Nabavi, S.M. Understanding Genistein in Cancer: The "Good" and the "Bad" Effects: A Review. *Food Chem.* **2016**, *196*, 589–600. [CrossRef] [PubMed]
23. Duru, K.C.; Kovaleva, E.G.; Danilova, I.G.; Van der Bijl, P.; Belousova, A.V. The Potential Beneficial Role of Isoflavones in Type 2 Diabetes Mellitus. *Nutr. Res.* **2018**, *59*, 1–15. [CrossRef] [PubMed]
24. Franke, A.A.; Custer, L.J.; Cerna, C.M.; Narala, K. Rapid HPLC Analysis of Dietary Phytoestrogens from Legumes and from Human Urine. *Soc. Exp. Biol. Med.* **1995**, *208*, 18–26. [CrossRef] [PubMed]
25. Lee, A.; Beaubernard, L.; Lamothe, V.; Bennetau-Pelissero, C. New Evaluation of Isoflavone Exposure in the French Population. *Nutrients* **2019**, *11*, 2308. [CrossRef]
26. Lee, A.; Bensaada, S.; Lamothe, V.; Lacoste, M.; Bennetau-Pelissero, C. Endocrine Disruptors on and in Fruits and Vegetables: Estimation of the Potential Exposure of the French Population. *Food Chem.* **2022**, *373*, 131513. [CrossRef]
27. Fernandez-Lopez, A.; Lamothe, V.; Delample, M.; Denayrolles, M.; Bennetau-Pelissero, C. Removing Isoflavones from Modern Soyfood: Why and How? *Food Chem.* **2016**, *210*, 286–294. [CrossRef]
28. Conrad, S.C.; Chiu, H.; Silverman, B.L. Soy Formula Complicates Management of Congenital Hypothyroidism. *Arch. Dis. Child.* **2004**, *89*, 37–40. [CrossRef]
29. Sathyapalan, T.; Dawson, A.J.; Rigby, A.S.; Thatcher, N.J.; Kilpatrick, E.S.; Atkin, S.L. The Effect of Phytoestrogen on Thyroid in Subclinical Hypothyroidism: Randomized, Double Blind, Crossover Study. *Front. Endocrinol.* **2018**, *9*, 531. [CrossRef]
30. Wang, S.; Zhang, S.; Wang, S.; Gao, P.; Dai, L. A Comprehensive Review on Pueraria: Insights on Its Chemistry and Medicinal Value. *Biomed. Pharmacother.* **2020**, *131*, 110734. [CrossRef]
31. National Toxicology Program. Multigenerational Reproductive Study of Genistein (CAS No. 446-72-0) in Sprague-Dawley Rats (Feed Study). *Natl. Toxicol. Program Tech. Rep. Ser.* **2008**, *539*, 1–266.
32. National Toxicology Program. Toxicology and Carcinogenesis Study of Genistein (CAS No. 446-72-0) in Sprague-Dawley Rats (Feed Study). *Natl. Toxicol. Program Tech. Rep. Ser.* **2008**, *545*, 1–240.
33. Messina, M.J. Legumes and Soybeans: Overview of Their Nutritional Profiles and Health Effects. *Am. J. Clin. Nutr.* **1999**, *70*, 439S–450S. [CrossRef]
34. Wolf, A.; Henes, D.; Bogdanski, S.; Lutz, T.; Krämer, E. Statistical Analysis of Parameter Variations Using the Taguchi Method. *Notes Numer. Fluid Mech. Multidiscip. Des.* **2013**, *122*, 247–264. [CrossRef]
35. Cimbala, M.J. *Taguchi Orthogonal Arrays*; Penn State University: State College, PA, USA, 2014; pp. 4–6.
36. Vergne, S.; Titier, K.; Bernard, V.; Asselineau, J.; Durand, M.; Lamothe, V.; Potier, M.; Perez, P.; Demotes-Mainard, J.; Chantre, P.; et al. Bioavailability and Urinary Excretion of Isoflavones in Humans: Effects of Soy-Based Supplements Formulation and Equol Production. *J. Pharm. Biomed. Anal.* **2007**, *43*, 1488–1494. [CrossRef]
37. Shinkaruk, S.; Durand, M.; Lamothe, V.; Carpaye, A.; Martinet, A.; Chantre, P.; Vergne, S.; Nogues, X.; Moore, N.; Bennetau-Pelissero, C. Bioavailability of Glycitein Relatively to Other Soy Isoflavones in Healthy Young Caucasian Men. *Food Chem.* **2012**, *135*, 1104–1111. [CrossRef]
38. Bennetau-Pelissero, C.; Le Houerou, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Synthesis of Haptens and Conjugates for ELISAs of Phytoestrogens. Development of the Immunological Tests. *J. Agric. Food Chem.* **2000**, *48*, 305–311. [CrossRef]

39. Le Houerou, C.; Bennetau-Pelissero, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Syntheses of Novel Hapten-Protein Conjugates for Production of Highly Specific Antibodies to Formononetin, Daidzein and Genistein. *Tetrahedron* **2000**, *56*, 295–301. [CrossRef]
40. Singleton, V.L.; Rossi, J.A. Colorimetry of Total Phenolics with Phosphomolybdic-Phosphotungstic Acid Reagents. *Am. J. Enol. Vitic.* **1965**, *16*, 144–158.
41. Tran, T.N.T.; Bui, X.V.; Loan, N.T.T.; Anh, N.H.T.; Le, T.D.; Truong, T.M.H. Adsorption and Desorption Characteristics and Purification of Isoflavones from Crude Soybean Extract Using Macroporous Resins. *Pol. J. Food Nutr. Sci.* **2022**, *72*, 183–192. [CrossRef]
42. Cassales, A.; De Souza-Cruz, P.B.; Rech, R.; Záchia Ayub, M.A. Optimization of Soybean Hull Acid Hydrolysis and Its Characterization as a Potential Substrate for Bioprocessing. *Biomass Bioenergy* **2011**, *35*, 4675–4683. [CrossRef]
43. D'Archivio, M.; Filesi, C.; Di Benedetto, R.; Gargiulo, R.; Giovannini, C.; Masella, R. Polyphenols, Dietary Sources and Bioavailability. *Ann. Ist. Super. Sanita* **2007**, *43*, 348–361. [PubMed]
44. Guan, X.; Zhong, X.; Lu, Y.; Du, X.; Jia, R.; Li, H.; Zhang, M. Changes of Soybean Protein during Tofu Processing. *Foods* **2021**, *10*, 1594. [CrossRef] [PubMed]
45. Pico, J.; Pismag, R.Y.; Laudouze, M.; Martinez, M.M. Systematic Evaluation of the Folin–Ciocalteu and Fast Blue BB Reactions during the Analysis of Total Phenolics in Legumes, Nuts and Plant Seeds. *Food Funct.* **2020**, *11*, 9868–9880. [CrossRef]
46. Cao, H.; Wu, D.; Wang, H.; Xu, M. Effect of the Glycosylation of Flavonoids on Interaction with Protein. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2009**, *73*, 972–975. [CrossRef]
47. Dourson, M.L.; Stara, J.F. Regulatory History and Experimental Support of Uncertainty (Safety) Factors. *Regul. Toxicol. Pharmacol.* **1983**, *3*, 224–238. [CrossRef]
48. Kodell, R.L.; Gaylor, D.W. Combining Uncertainty Factors in Deriving Human Exposure Levels of Noncarcinogenic Toxicants. *Ann. N. Y. Acad. Sci.* **1999**, *895*, 188–195. [CrossRef]
49. Renwick, A.G. The Use of an Additional Safety or Uncertainty Factor for Nature of Toxicity in the Estimation of Acceptable Daily Intake and Tolerable Daily Intake Values. *Regul. Toxicol. Pharmacol.* **1995**, *22*, 250–261. [CrossRef]
50. Magee, P.J. Is Equol Production Beneficial to Health? *Proc. Nutr. Soc.* **2011**, *70*, 10–18. [CrossRef]
51. Webb, D.; Plattner, B.J.; Donald, E.; Funk, D.; Plattner, B.S.; Alavi, S. Role of Chickpea Flour in Texturization of Extruded Pea Protein. *J. Food Sci.* **2020**, *85*, 4180–4187. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

IV.2.2. Impact des traitements traditionnels et domestiques sur les taux d'IFs des aliments et conséquences sur l'estimation de l'exposition aux IFs en Asie : article correspondant à la référence [223]

L'étude, publiée sous le titre *Traditional and Domestic Cooking Dramatically Reduce Estrogenis Isoflavones in Soy Foods* [223], concerne l'impact des méthodes de préparation traditionnelles sur les teneurs en IFs des aliments à base de soja. Les données obtenues ont été examinées en rapport avec celles d'aliments industriels équivalents, disponibles en France. Les travaux ont porté sur trois spécialités : le miso, le tempeh (fermentés) et le tofu (non fermenté). En parallèle, l'impact de rinçages domestiques sur les teneurs en IFs d'aliments à base de soja non traditionnels (lait de soja et PST) « faits maison » a été étudié. Enfin, les effets du rinçage sur les teneurs en IFs d'aliments à base de soja du commerce ont été évalués.

foods

MDPI

Article

Traditional and Domestic Cooking Dramatically Reduce Estrogenic Isoflavones in Soy Foods

Souad Bensaada [1,2], Gabriele Peruzzi [2], Laurent Cubizolles [2], Muriel Denayrolles [3,4] and Catherine Bennetau-Pelissero [3,*]

[1] Univ. Bordeaux, CNRS, INSERM, ARNA, UMR 5320, U1212, F-33000 Bordeaux, France; souad.bensaada@u-bordeaux.fr
[2] Berkem, Rue Jean Duvert, 33290 Blanquefort, France; gabriele.peruzzi@biopress.fr (G.P.); laurent.cubizolles@berkem.com (L.C.)
[3] Feed & Food Department, Bordeaux Sciences Agro, 33175 Gradignan, France; muriel.denayrolles@agro-bordeaux.fr
[4] CBMN, UMR CNRS 5248, 33607 Pessac, France
* Correspondence: catherine.bennetau@laposte.net

Abstract: Soybean is a pulse which has considerable nutritional value due to its high protein, fibers and polyunsaturated fatty acid (PUFA) contents. It also contains phytoestrogenic compounds that definitely hinder its recommendation for general consumption. Contrary to ancient times, when soybeans were boiled, modern commercial soy foods can contain up to 150 mg/100g of estrogenic isoflavones. Interestingly, current estimations of isoflavone intake in the literature do not distinguish between the origins of soy food, i.e., whether it is homemade or commercial. As a result, the isoflavone exposure in Asian countries may well be overestimated. This study aims to demonstrate, based on step-by-step monitoring of isoflavones, that traditional and domestic treatments, leveraging isoflavones water-solubility, can indeed significantly reduce their content in soy foods. Indeed, when compared to commercial foods, the isoflavone content was found to be 20, 2.6, 4.5 and 9.8 times lower in "homemade" soy juice, tofu, tempeh and miso, respectively. Additionally, water soaking was found to reduce the isoflavones levels in soy-textured proteins by more than 70%. Hence, this simple process has the potential to help drastically reduce overall xenoestrogens exposure. This study could serve as a basis for establishing the isoflavones Reference Dose and issuing food safety guidelines.

Keywords: soy food; phytoestrogens; genistein; daidzein; cooking practices; water treatments; human exposure; health

check for updates

Citation: Bensaada, S.; Peruzzi, G.; Cubizolles, L.; Denayrolles, M.; Bennetau-Pelissero, C. Traditional and Domestic Cooking Dramatically Reduce Estrogenic Isoflavones in Soy Foods. *Foods* **2024**, *13*, 999. https://doi.org/10.3390/foods13070999

Academic Editor: Dapeng Peng

Received: 8 March 2024
Revised: 20 March 2024
Accepted: 21 March 2024
Published: 25 March 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Soybean (*Glycine max* (L.) Merr.) is recognized as the foodstuff with the highest concentration of isoflavones (IFs), which exhibit estrogenic activities. Compared to other xenoestrogens, IFs bear the closest resemblance to estradiol and are highly concentrated in food sources [1,2]. The effects of IFs can vary from beneficial to adverse, contingent on the physiological status of the consumer and the ingested quantity [3]. In 2008, the National Toxicology Program (NTP) of the USA classified genistein (GEN), the primary soy-isoflavone, as reprotoxic and carcinotoxic in rats at a dosage of 35 mg/kg/day [4,5]. Given the absence of a reliable NOAEL (No Observable Adverse Effect Level), Food Safety Agencies should propose a common Reference Dose for the estrogenic IFs that should not be exceeded. This limit should also consider the actual exposure in Asian countries. Estimating this exposure necessitates reliable measurements in foodstuffs, an issue this study aims to address.

Due to potential adverse effects in certain population categories, it may be beneficial to reduce the levels of estrogenic IFs in the diet intended for the general population, reserving IFs for specific applications and circumstances. This cautionary approach is often disputed,

with arguments that soy is traditional in Asia and that no adverse effects, or even beneficial ones, have been reported in the past [6]. However, does high soy consumption necessarily equate to high IFs exposure?

A close examination of soy's history in China reveals that written testimonies were frequently lost and reconstructed over the centuries, making it challenging to trace. The first reliable reports date from the Han dynasty (206-BCE to 220 CE). They described agricultural practices that could be dated back 2000 years before reporting, making it difficult to distinguish factual information from myths [7]. Although soy has been cultivated for ages, the destruction of written sources prevents the confirmation of a time period for soybean's use as a foodstuff. For instance, it was primarily used as a green manure before the Han dynasty, when the major crops cultivated were millet, wheat, rice, hemp, mung beans, and sesame [8]. In fact, soybean can be cultivated even in low-quality soils due to its ability to fix nitrogen in soils, enhancing subsequent cultures. As such, it was used for over 1000 years in quite complex crop rotations [9]. Soy was easily stored for up to three years, since it contains many antinutritional factors which reduced pest (rodents or birds) attacks in the fields or during seed storage. Initially, it was mainly used to feed animals and was considered a starvation and war crop around 100 BCE [7,10]. It was not highly appreciated due to its poor taste and the flatulencies and the stomach pains it induced. The most effective method to improve soybeans' digestibility at that time was to boil them. Early fermented soy foods were seemingly reported in 1200 BCE in Korea. Soy was then considered a potential food for human beings [9]. In China, according to Baraibar-Norberg and Deutsh [7], *shi* (fermented soybeans) was found in Han tombs dated from 200 to 150 BCE. Fermentation, which allowed better conservation and enhanced taste, required the use of clean and sterile beans before the development of specific molds, implying water boiling [9]. Soy curd like tofu seemed to appear during the late Han dynasty. Again, the traditional recipes included several cleanings and rinses as well as prolonged water boiling [11].

Such treatments applied to beans or dehulled beans have been reported to yield the inactivation of antinutritional factors and the removal of estrogenic IFs while preserving proteins [12], fibers, and polyunsaturated fatty acids (PUFAs) among ω3 and ω6 families [13]. In contrast, modern industrial processes primarily involve physical constraint treatments such as pressing and extrusion, which may lead to some concentration of IFs. Therefore, when dietary reports in the literature suggest that Asian populations are exposed to high levels of IFs, based solely on high soy food consumption, this appears to be a questionable assumption if the food origin is not specified.

Additionally, when data associating estimated IFs intakes and IFs in body fluids are gathered, a discrepancy emerges between Western and Asian populations. The same estimated exposure leads to 1 to 1.5 µM in Western people' blood, while it is about 0.2 to 0.3 µM for Asians. This was previously questioned by Vergne et al., in [14], when the pharmacokinetics of GEN and Daidzein (DAI) were assessed concurrently in French and Chinese students. The study revealed only a small difference in the C_{max} of DAI during chronic soybean challenge and it was hypothesized that, in Asian students, the gut flora may negatively influence the IFs bioavailability. However, this finding alone could not explain the significant difference observed in Asians and Westerners' body fluids concentrations of IFs. Thus, the hypothesis of a misestimation of IFs intakes remained.

This study was designed to address this issue. Its objective is to demonstrate that traditional recipes, and by extension the domestic ones inspired by them, reduce significantly IFs levels in Asian soy foodstuffs. Hence, GEN and DAI concentrations were measured on soy foods prepared under domestic conditions by ELISA. As the exact step for potential IFs removal was unknown, the processes were analyzed step by step to determine their impacts on the concentration of these two major estrogenic IFs. The results should indicate that traditional exposure to IFs in Asia is not as high as suggested by recent data [15,16] which established the IFs exposure at a median of 13.5 mg/day in China and at 21.4 mg/day in Japan.

The foodstuffs selected were tempeh and miso for fermented food and soy juice and tofu for unfermented food, these being among the most common traditional dishes. Soy

textured proteins were also treated under domestic conditions since they can be used by Western consumers. A comparison was then made with commercial industrial equivalent products analyzed here and in a previous study by the same assay technique. The health effects of estrogenic IFs are then discussed as well as the most probable consequences of this finding on the estimation of Asians' IFs exposure.

2. Materials and Methods

2.1. Materials

All soy matters, i.e., whole and dehulled soybeans and textured soy proteins, were supplied by the organic producer of soy proteins Biopress®/Berkem® (Tonneins, France). Raw matters were acquired from a local organic farmers' union: Bioprogress (Fourcès, Gers, France), which follows Demeter specifications. The soybeans farms were located in the Lot et Garonne and Gers departments. Both whole and dehulled soybeans were obtained from the ISIDOR cultivar, but from two different storage silos. Dehulled soybeans were sifted before use to remove soybean dust. Textured soy proteins were sorted in four different particle-size groups: large (12–20 mm), medium (8–12 mm), fine (5–8 mm) and extra-fine (2.5–5 mm). Rice-koji ferments, i.e., *Rhizopus oligosporus* for tempeh and *Aspergillus flavus* var *orizae* for miso, as well as miso starter were bought from an organic store (Biocoop, Bègles, France). The plain tofu and tempeh used in the rinsing experiment also came from the same local organic store. They both came from the "Soy" brand. The tofu was composed of water, 25.9% dehulled soybeans, gellant (calcium sulphate), and nigari. The tempeh was composed of 99.3 % dehulled soybeans precooked in water, cider vinegar and ferments. The commercial products analyzed as a basis for comparison given in Table 1 were purchased during the last 5 years from French supermarkets located around Bordeaux.

Table 1. Genistein and daidzein amounts per portion of different types of soy food collected on the French market.

Type of Product	Number of Measurements	Portion Size (g)	Mean GEN + DAI/ Portion (mg)	Standard Error of Mean/Portion	Range of Measurements *
Raw soybeans	5	100	84.6	8.3	72.5 to 100.6
Toasted soybeans	5	100	134.0	37.9	61.9 to 247.7
Edamame	2	120	49.2	14.1	35.1 to 63.3
Soybean flour	2	100	38.5	5.5	32.9 to 43.9
Protein isolate (sports)	3	165	52.2	12.3	39.8 to 64.5
Soy-based drinks	9	100	9.9	1.7	2.8 to 17.7
Soy-based dessert cream	11	150	12.3	2.8	2.5 to 33.4
Soy-based yogurt	14	125	15.9	3.8	5.1 to 31.5
Soy-based cream	12	50 mL	4.5	5.9	3.1 to 6.7
Soy-based Vegan steak	19	100	27.8	3.5	4.1 to 51.2
Soy sausages	4	90	24.8	7.3	11.1 to 44.1
Soy-based raw pasta	2	100	23.7	1.6	22.1 to 25.3
Soy-based cooked pasta	2	100	8.75	0.6	8.2 to 9.3
Soy-based infant formulas	6	4-month-old infants	25.2	2.9	15.7 to 34.3
Tempeh	4	100	27.8	4.5	15.5 to 34.3
Tofu	13	100	34.8	5.9	9.4 to 79.8
Soy-based cheese	8	50	23.8	7.4	2.9 to 36.2
Miso	2	100	14.7	0.98	13.7 to 15.6
Soy sauce	9	10 mL	0.16	0.004	0.16 to 0.17

* Figures calculated from [1] and isoflavones are genistein + daidzein concentrations in aglycone forms.

The chemical reagents were acquired from MERK (Fontenay-sous-bois, France) and Sigma-Aldrich (Saint Quentin Fallavier, France), unless otherwise stated. The samples and reagents aqueous solutions were made in ultrapure water obtained from an Elga Veolia® (High Wycombe, UK) instrument.

2.2. Methods

2.2.1. Traditional Recipes

Soy Juice

Soy juices were prepared from either 100 g of dehulled beans or whole beans. The first soaking step in 1L of water lasted 12 h, during which the water was renewed 0, 1, 2, 3 or 4 times. Then, the beans were mashed in 700 mL of either soaking water or clean water. The different mixtures obtained were individually filtered through a piece of fabric and the corresponding okara kept aside. For each type of juice collected, 350 mL was diluted twice with water to reach the commercial concentration of soymilk. After a boiling step to reduce all antinutritional factors, it was allowed to cool down to room temperature before storage at 4 °C or −32 °C. At each stage of the process, three samples were collected and analyzed individually.

Tofu

Tofu samples were prepared from the different soy juices obtained just before the last dilution step as previously described. After boiling, nigari ($MgCl_2$) was added at a final concentration of 5 g/L and the juice was left to curdle for 30 min. The curds were poured onto a tofu press and the whey discarded by pressing for 30 min at 25,500 Pa.

Tempeh

The recipe is based on a traditional process described by Fernandez-Lopez et al. [17]. Dehulled beans were rinsed 3 times with tap water before being cooked in boiling water for 20 min and then left to soak in cooking water for an extra 20 min. During this first cooking step, the foam that formed on top of the water was removed. These precooked beans in turn underwent the same 3 processing stages as the dehulled beans. After the second cooking, beans were lightly dried and spread on a baking sheet in an oven at 60 °C for 10 min. Finally, the beans were mixed with white vinegar and *Rhizopus oligosporus* ferment, reaching concentrations of 6 mL/100 g and 500 mg/100 g, respectively. The mixture was poured in zip-lock plastic bags, which were pierced with a fork and left in a heat chamber at 28 °C for 72 h.

Miso

The traditional recipe for miso was taken from The Japanese Lab [18], which describes how to make light-colored miso. First, 100 g of dehulled beans were cooked in boiling water for 3 h, during which the boiling water was renewed twice. They were cooled down to room temperature and mixed with the equivalent amount of rice-koji, i.e., rice pre-incubated with *Aspergillus flavus* var *oryzae* and 133 mg of miso starter. The obtained mixture was formed into balls which were deposited and pressed into a large-necked glass jar to avoid air pockets. Its surface was sprinkled with salt and then covered with a food-grade plastic film supporting a ballast. The jar was closed and the mixture was left to ferment for 8 months.

Hydrated Soy Proteins

Four sizes of textured soy-proteins were tested. In each case, 100 g of matter was soaked in 1 L of tap water to hydrate for 30 min, during which 0, 1, 2 or 3 water renewals were performed. In the case of 1 renewal, it occurred after 30 min; for 2 renewals, they occurred after 15 and 30 min; for 3 renewals, they occurred every 10 min. Two water temperatures were tested: 22 °C and 65 °C.

Commercial Tofu and Tempeh

A consumer portion, either 100 g of tempeh or 125 g of tofu, was immersed in 1 L of deionized water for a total of 15 min, with two water changes at 5 and 10 min, before which the foodstuff was dried on absorbent paper and weighted. Water was gently stirred for 1 min at each water renewal. Samples were collected at 0, 5, 10 and 15 min of soaking. IFs, i.e., GEN and DAI in aglycone forms, were assayed at each time set.

2.2.2. ELISA

Despite recommendations from the French Food Safety Agency in 2005 [19], declaration of isoflavones content in soy-based foodstuffs is still not mandatory in France. However, considering the functional value of these substances, ELISAs were developed by our team in the late 1990s and have since been used to analyze these polyphenols in foods, biological fluids [20], and recently in hair [21]. This technique was validated in an international ring test monitored by TNO (Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek, The Netherland) and results can be found in the work of Bennetau-Pelissero et al. [20]. Although this ring test showed considerable heterogeneity in isoflavones determination between laboratories, this was essentially due to inadequate extraction procedures in some cases. Nevertheless, data obtained using ELISA appeared reliable and comparable to the best chromatographic techniques challenged in this trial when aglycone forms were considered. Additionally, ELISAs have been validated in two studies designed to test food-frequency questionnaires assessing recent [22] and usual soy intakes [21]. To date, there is no legal technique for isoflavones determination in France.

Samples Preparation

To perform the IFs analyses via ELISA, the samples were first subjected to an extraction as described below. Thereby, 1 g of mortar-crushed soy-based material was dispersed into 50 mL of water by a two-step stirring at room temperature for 20 min and then at 100 °C for 10 min. After cooling down to room temperature, 500 µL samples were collected in triplicate under stirring to ensure homogeneity. Two milliliters of acetate buffer (sodium acetate 0.1 M; EDTA 0.14 M; 100 UI.mL^{-1} penicillin G (Sigma, P-3032); 0.1 mg.mL^{-1} streptomycin (Sigma, S-6501)) at pH 5 were added to each sample vial together with 10 µL of β-glucuronidase aryl-sulfatase from *Helix pomatia* (Roche®, 10127698001, Mannheim, Germany) to allow the digestion of glycosylated IFs. The samples were then incubated overnight at 37 °C with shaking [23]. Afterwards, the extraction of aglycone compounds was performed: 3 mL of acidified ethyl-acetate (500 µL HCl 38% per L) was added, and the vial vortexed for 30 s, centrifuged at 500 g (Jouan™ CR3, Fisher Scientific, Illkirch, France) for 10 min at 4 °C and finally stored at −22 °C to allow phase separation. The organic phase containing the IFs in aglycone forms was evaporated to dryness using a Speed-Vac (Thermo-Electron™ Corporation, Fisher Scientific, Bordeaux, France). Then, each sample was diluted in 0.5 mL of assay buffer, i.e., phosphate-buffered saline (PBS) 0.01 M, 0.9% NaCl, 0.2% Tween, 1% DMSO, pH 7.3, and sonicated when required. Samples were stored at −32 °C until processed for IFs analysis.

To assess the digestion and extraction recovery, the hydrolysis by β-glucuronidase aryl-sulfatase was monitored using genistin (EXTRASYNTHESE™, 1325 S, Genay, France), a pure compound used as a control reagent in an external standard run in parallel to each measured sample. The compound was dissolved at 1 mg.mL^{-1} in DMSO as the stock solution. In these operating conditions, the hydrolysis performance was always between 87 and 103%. In addition, a sample of chocolate soymilk was run similarly to the tested samples. It was considered to be the plate control. Recoveries greater than 100% can be explained by the high accuracy of ELISA measurements and inter- and intra-assay variations.

Assay

GEN and DAI were assayed in soy matter and treatment water using ELISA specific to each molecule, as explained by Shinkaruk et al. [24]. The primary antibodies were selected and obtained in previous works [25,26]. All glassware was coated with silicone (Sigmacoat®, Sigma-Aldrich, Saint Quentin Fallavier, France, Ref SL2) and all IFs solutions were handled with low retention tips.

Characteristics of the ELISA

The soy-matter samples were diluted to 1/50 (v/v) for enzymatic digestion. The sensitivities of the GEN and DAI assays in food matter were 10 µg.mg^{-1} and 6.5 µg.mg^{-1},

respectively. The intra-assay variation was never greater than 7% and the inter-assay variation obtained on different microtitration plates was always below 17% [25,26]. Assays were not considered if the "r" coefficient of the sigmoid calibration curve was <98.5%. The final dilutions in the microtitration plates ranged from 1/50 to 1/800. All values are given here in aglycone equivalent.

2.3. Statistical Treatments

All data correspond to a mean value and its standard error of the mean (SEM) derived from triplicate or quadruplicate measurements performed on three or four different microtitration plates. Due to the small number of values in each sample, a non-parametric statistical test, i.e., the Wilcoxon–Mann–Whitney test, was used. Sample values were sorted by rank, then the U parameter describing the intercalation of values of the two samples compared was obtained. The significance level (p) was determined using specific tables that provide this value according to the size of each sample. The final significance was half the value provided by the table for two-tailed tests. If the *p*-value was less than 0.05, it indicated that all the values of a triplicate or quadruplicate were lower than those of the other sample.

3. Results

Here, we present the IFs contents of tempeh and miso prepared according to traditional recipes. Data showing the impacts of water renewal during beans soaking on IFs contents for soy juice and tofu domestic preparations are given next. The impact of water rinsing on the IFs contents of textured proteins of different sizes is also shown. Finally, the results of rinsing tests of commercial tofu and tempeh are presented. The IFs levels in the okara samples collected during soy juices and tofu makings are presented in Supplementary Figure S1.

The concentrations of IFs were compared to those measured in commercial products bought in French supermarkets specifically for this study and also to those previously measured by our team for the same products using the same technique and published in [1]. The reference data issued in [1] are also presented in Table 1.

3.1. Impact of Traditional Process on Isoflavones Content in Tempeh and Miso

IFs levels measured at different stages of the traditional recipes of tempeh and miso are presented in Figure 1.

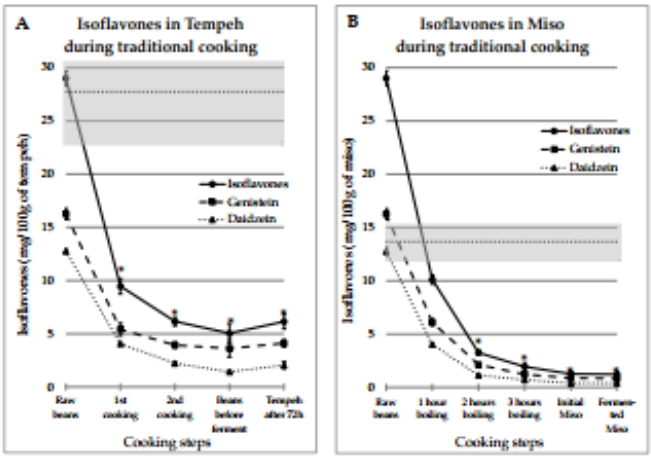


Figure 1. Evolutions of Genistein and Daidzein concentrations during the traditional preparations of tempeh (**A**) and miso (**B**). The gray zone covers the mean and SEM of the equivalent commercial products presented in Table 1. The stars indicate IFs values in food that were significantly different from the mean ± SEM of IFs content in commercial products ($p < 0.05$).

As mentioned above, tempeh and miso were both made from dehulled beans which contained 16.5 ± 0.52 mg GEN/100 g and 12.8 ± 0.17 mg DAI/100 g. The IFs concentrations measured at the different preparation steps decreased in an asymptotic way, as shown

in Figure 1A for tempeh and Figure 1B for miso. The gray zones on the figure cover the range of values observed in industrial-equivalent food products. The stars indicate IFs values in prepared food that were significantly different from the mean ± SEM of IFs contents in commercial products ($p < 0.05$). It can be seen that in both cases, the resulting IFs levels were consistently much lower in "homemade" food than in commercial products. Moreover, this effect is the most important as soon as the first preparation step is complete; additional steps results in only slight reductions.

3.2. *Impact of Water Renewals during Beans Soaking for Soy Juice and Tofu*

3.2.1. Process Tested on Whole Soybeans

IFs were assayed in soy juice and tofu prepared from whole beans, which contained 32.8 ± 0.9 mg GEN/100 g and 27.6 ± 0.12 mg DAI/100 g, soaked for 12 h with 0, 1, 2 or 3 water renewals. It should be noted that the two first soaking water batches contained dust and lipids, but the subsequent water batches were clear. The concentrations of aglycone IFs in mg/100 mL of soy juice and in mg/100 g of tofu are given in Figure 2. The amounts of IFs were always significantly lower ($p < 0.05$) than those measured in commercial equivalent products. After the fourth rinsing, the amount of IFs was reduced by 43% in the juice and by 33% in the tofu sample. Indeed, the envelope of the seed prevented IFs from leaking in the water during the soaking steps.

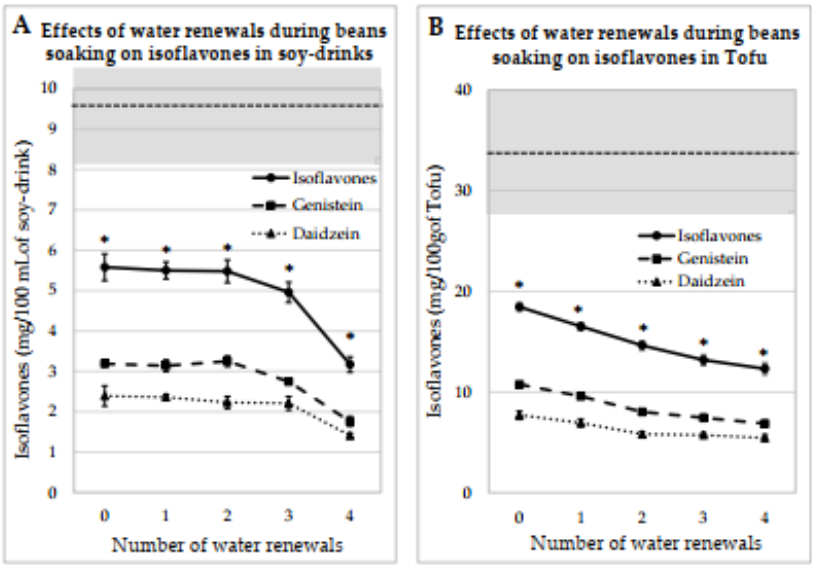


Figure 2. Evolutions of Genistein and Daidzein concentrations in soy juice (**A**) and in tofu (**B**) prepared from whole soybeans soaked for 12 h in water renewed several times (1 to 4) or not renewed (0). The gray zone covers the mean and SEM of the equivalent commercial products presented in Table 1. * indicates a significant difference from commercial products ($p < 0.05$).

The IFs levels measured in the okara produced during soy juice and tofu manufacturing are shown in Supplementary Figure S1. It can be seen that the IFs decreased asymptotically. When okara was obtained from whole beans, 40% of IFs were removed after three water renewals, while almost 70% were eliminated after four water renewals on dehulled beans. The IFs decrease was significant after the first water renewal.

3.2.2. Process Tested on Dehulled Beans

According to a traditional tofu-making experience in Debao Guangxi [11], it appears that the traditional recipe of tofu was based on dehulled beans. Therefore, the same soaking process as already described for whole beans was tested on dehulled beans. The concentrations of IFs, i.e., GEN + DAI, measured in each case are listed in Figure 3. As observed in the precedent case, the "homemade" juice and tofu always contain significantly

($p < 0.05$) less IFs than commercial equivalent products. When the number of water renewals increases, the IFs levels decrease in a quasi-asymptotic way.

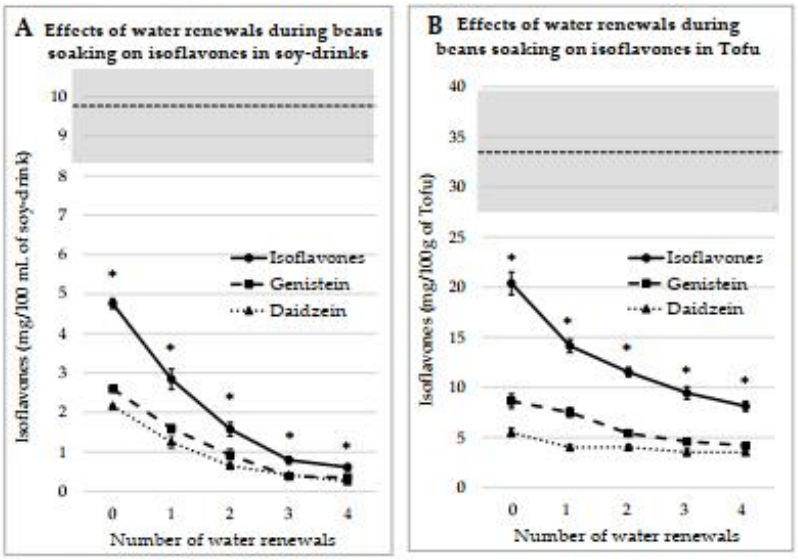


Figure 3. Evolutions of Genistein and Daidzein concentrations in soy juice (**A**) and in tofu (**B**) prepared from dehulled soybeans soaked for 12 h in water renewed 1 to 4 times or not renewed (0). The gray zone covers the mean and SEM of the equivalent commercial products presented in Table 1. * indicates a significant difference from commercial products ($p < 0.05$).

It can also be inferred by comparison between Figures 2 and 3 that the additional water renewals are more efficient at removing IFs when applied on dehulled beans than on whole beans. After four rinses, the IFs concentration in soy juice is reduced by 87%, while in tofu it was reduced by 60%. The reason is probably that the lack of a seed coat eases the IFs leakage into the water.

3.3. Impact of Water Renewals on the Levels of Isoflavones in Textured Proteins

The textured soy-proteins tested contained higher IFs concentrations than whole or dehulled soybeans. Indeed, although the different products did not come from the same soy matter, regular analysis of soybeans resulted in GEN + DAI concentrations of IFs aglycones around 60–70 mg/100g. In parallel, textured proteins were regularly measured at concentrations ranging from 90 to 110 mg/100g. This is because the matter used did not come from the same batches of beans. In addition, textured soy-proteins were obtained from soy defatted cakes resulting from soy seeds pressed for oil. As the extrusion process used did not significantly alter the IFs content and since oil represented about 20% of the mass of the initial seeds, the resulting textured soy-proteins were enriched in IFs.

3.3.1. Treatment with Water at 22 °C

The concentrations of IFs in textured proteins hydrated for 30 min at room temperature with different rinsing steps are presented in Figure 4A. Here, the initial matter contained: 73.4 ± 3.8 mg GEN/100 g and 35.2 ± 2.3 mg DAI/100 g; 71.5 ± 2.4 mg GEN/100 g and 35.1 ± 3.6 mg DAI/100 g; 70.4 ± 2.4 mg GEN/100 g and 38.3 ± 3.5 mg DAI/100 g; 70.7 ± 0.8 mg GEN/100 g and 36.5 ± 1.7 mg DAI/100 g, for large, medium, fine and extra-fine particles, respectively.

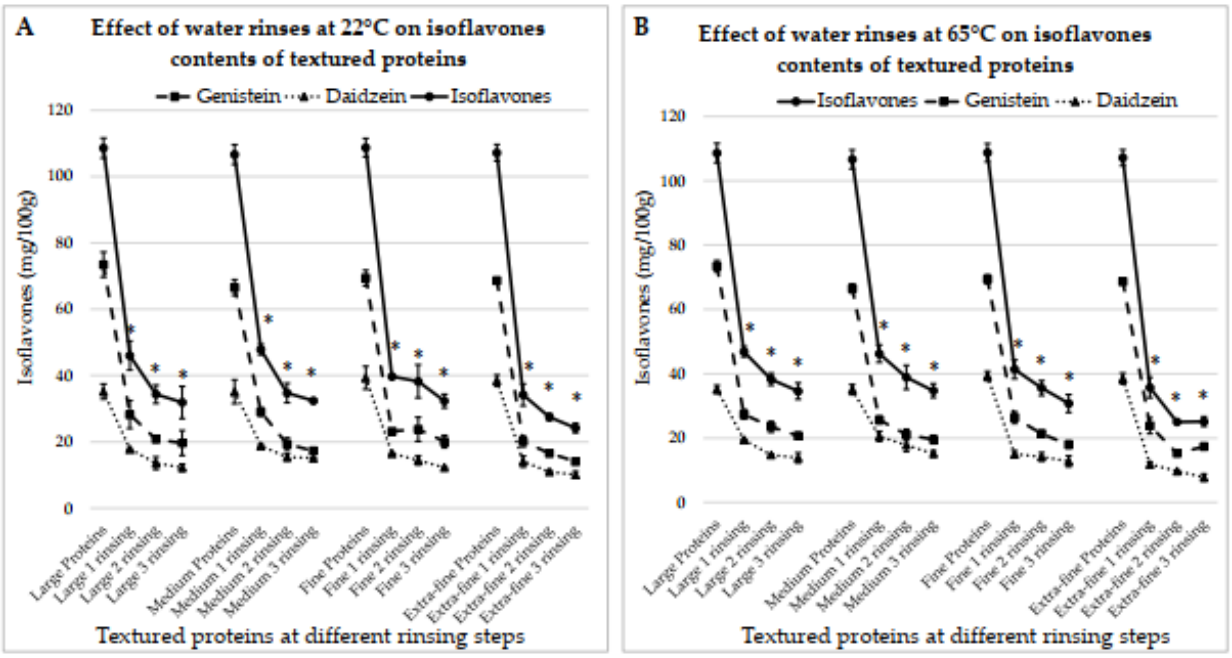


Figure 4. Effects of water rinses performed at 22 °C (**A**) and 65 °C (**B**) on the isoflavone contents of textured soy-proteins of different sizes, from large to extra-fine. * indicates a significant difference from initial IFs level ($p < 0.05$).

From Figure 4A, it can be seen that in all cases, the first rinsing results in the greatest drop in IF levels ($p < 0.05$). After the third rinsing, the concentrations of IFs are decreased by 70.6%, 69.9%; 70.3% and 77.4% for large, medium, fine and extra-fine proteins, respectively.

3.3.2. Treatment with Water at 65 °C

The results for the treatment with hot water are presented in Figure 4B. They are very similar to the previous case: after the first rinsing, the concentrations of IFs decreased significantly ($p < 0.05$); after the third rinsing, they decreased by 68.1%, 67.4%, 71.7% and 76.5% for large, medium, fine and extra-fine particles, respectively. Hence, IFs' removal efficiency is not significantly influenced by the soaking temperature. However, for both temperatures, the difference in the results is significant between the large and extra-fine proteins. At room temperature, after three rinsing steps, the total IFs contents are 31.9 ± 4.9 mg/100 g and 24.2 ± 1.5 mg/100 g in large and extra-fine textured proteins, respectively. At 65 °C, the results are 34.6 ± 2.6 mg/100 g and 25.1 ± 1.5 mg/100 g in large and extra-fine textured proteins, respectively.

3.4. Impact of Domestic Rinsing on Isoflavones Content of Commercial Tofu and Tempeh

A consumer may prefer to buy ready-to-use tofu or tempeh, and of course in that case the levels of IFs can be quite high, as shown in Table 1. Of the commercial products used here for the domestic rinsing test, the plain tofu contained 16.4 ± 0.92 mg GEN/100 g and 11.6 ± 1.0 mg DAI/100 g and the plain tempeh contained 19.3 ± 0.18 mg GEN/100 g and 17.2 ± 0.44 mg DAI/100 g. The IFs levels corresponding to the different rinsing steps are presented in Figure 5.

Rinsing with water seems to be effective for commercial products provided that their texture is strong enough to withstand such treatment. In this experiment, the products were not significantly impregnated with water as they conserved their firmness and weight before and after each rinsing step. However, as can be seen in Figure 5, after just 10 min of contact with water, the IFs levels significantly decreased ($p < 0.05$). Indeed, in tofu, they reduced to 12.3 ± 0.17 mg and 6.1 ± 0.7 mg for GEN/100 g and DAI/100 g, respectively. After 20 min of contact and two water renewals, they were reduced to 8.57 ± 0.41 mg

and 4.5 ± 0.6 mg of GEN/100 g and DAI/100 g, respectively. Similarly, in tempeh, after 10 min of rinsing, the IFs concentrations were reduced to 8.6 ± 0.6 mg and 5.7 ± 0.5 mg for GEN/100 g and DAI/100 g, respectively. At the end of the rinsing process, the amounts of IFs were reduced to 7.2 ± 0.03 mg and 4.3 ± 0.1 mg for GEN/100 g and DAI/100 g, respectively. As observed previously, the IFs removal followed an asymptotic scheme and the third rinsing step did not seem to be very efficient.

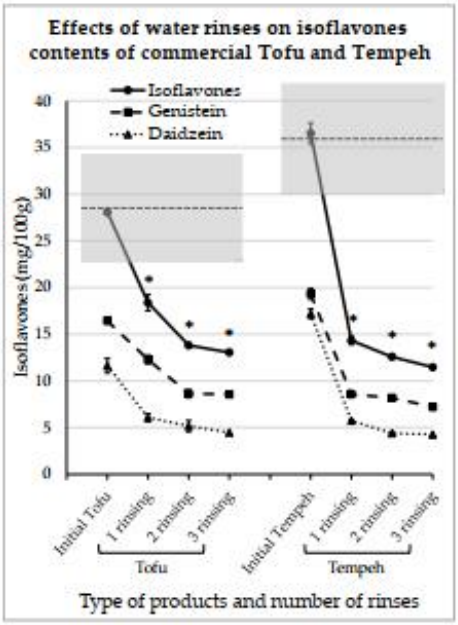


Figure 5. Reductions in isoflavones in commercial tofu and tempeh obtained after domestic water rinses. The stars indicate a significant difference in isoflavones concentration compared to the initial product ($p < 0.05$).

4. Discussion

Although raw soybean is the pulse with the highest content of proteins, it also contains phytic acid and tannins, which reduce mineral absorptions; oligosaccharides, which ferment in the gut, causing flatulencies; lipoxygenases, which decrease PUFAs quality; and saponins and sapogenins, which degrade the lipid membranes of enterocytes [27]. Soybeans also contain hemagglutinins, which can induce the coagulation of white and red blood cells if they enter the blood stream, and protease inhibitors, the Bowman–Birk and Kunitz factors, which reduce the proteins digestibility [27]. Finally, they also contain allergens and IFs [27]. This is why Baraibar Norberg and Deutsh [7] wrote that *"The same chemistry that protects the soybean against pests makes it partially indigestible for humans and conducive to flatulence and stomach pain"*. However, barely heating soybeans considerably reduces all antinutritional factors except IFs, which are thermoresistant. Fortunately, since IFs are in glyco-conjugated forms, they are soluble in water and may be reduced via water rinsing. This effect was tested in different contexts in this work.

4.1. Data Analysis

As seen in Table 1, industrial soy food currently found in supermarkets may contain quite high, although variable, amounts of IFs. Unfortunately, only very few have a mention of their IFs contents on the packaging. In France the Food Safety Agency advised in 2005 to limit IFs intake to 1 mg.kg^{-1}day^{-1} [19]. As will be seen later, this "Reference Dose" should most probably be lowered nowadays due to toxicological results subsequently obtained by the USA NTP. The French Agency's dose represents less than a portion of toasted soybeans for an adult of 60 kg.

In this study, it also appears that the initial isoflavone levels in raw matters may vary. Indeed, the whole and dehulled soybeans used for this work were sourced from

two different storage silos belonging to the same organic farmers' union. As a result, they might have originated from different cultivation areas in the South-West of France. It is well known that culture conditions, including climate, soil quality and even farm management practices, may influence the amount of isoflavones in soybeans [28]. Therefore, such differences in raw materials were not unexpected.

The experiments performed here show that IFs can easily be removed from soy matter via simple water treatments. This effect was already shown by several authors [29]. By implementing traditional recipes of tempeh and miso, in this work, it was possible to obtain foodstuffs containing almost 5 to 10 times less IFs than commercial equivalent products. This suggests that the ancient soy-consumption, traditionally in Asia, only provided small amounts of estrogenic IFs. This work also shows that simple water renewals during domestic beans soaking can dramatically reduce estrogenic IFs from subsequent preparations of soy juice or tofu. It appears that treating dehulled beans is more efficient for IFs removal than treating whole beans, as already observed [12]. Similarly, when water rinsing was tested on textured proteins of different sizes, IFs were more efficiently removed from fine or extra-fine proteins than from large or medium-sized ones. The difference in IFs levels between large and extra-fine proteins after water treatment was significant ($p < 0.05$). Moreover, using hot water did not bring additional efficiency to the process. These results are consistent with those previously published by Bensaada et al. [12]. Although this result might appear to be counterintuitive, it was observed several times by our team during the analysis of water treatments on both soybeans and soy-textured proteins. These observations suggest that in hot water, the proteins on the surface of a soy particle may coagulate faster than those from the center, thereby limiting isoflavones migration into the water. This hypothesis warrants further investigation using appropriate measurement techniques.

Regarding the proteins and lipids contents of soy products, they are not significantly reduced after water treatment [12]. Indeed, working on whole beans or dehulled beans allows the soy-cells integrity to be preserved. Thus, losses of large molecules are limited while small molecules such as IFs or vitamins and minerals can leak into the water when they are not adsorbed to larger biological cell constituents. In this study, vitamins and minerals were not analyzed based on the assumption that this issue should not constitute a major matter of concern, since this was already the case when soybeans were boiled and if the soy was a part of a diversified diet.

Finally, it may be difficult to perform these treatments at an industrial scale, and this is why the rinsing processes were also tested on commercial products and they proved to be efficient. Regardless, such water rinsing or soaking should be encouraged at domestic levels for most consumers who do not require estrogenic IFs intake. Food companies may also innovate, creating food products that could be easily treated at home.

4.2. *Traditional Treatments*

The ancient texts mentioning soybean and the actual soy-cooking practices, which derive from ancient traditions, both involve cooking with water. Such process was applied to reduce antinutritional factors, including IFs. This occurred when water was discarded. Soybean washing, soaking and boiling have been mandatory throughout the history of soy food [7]. In the past, boiling allowed to cook soybeans and to sterilize them before fermentation [9].

Considered as a low-value crop, soybean was initially used for animal feeding. However, under the Han dynasty, there was an increase in the Chinese population followed by a subsequent decline in livestock farming, and land was more often used to produce plant foods for humans instead. During this period, veganism developed in Asia and foods of animal origin were essentially based on fish, poultries and pork. During the territorial expansion wars of the Han period, soy became a war food, as it could be easily stored and transported. Because of its antinutritional factors, soy was not appreciated by human consumers and was always boiled. According to Lee et al. [9], fermentation was developed

quite early to improve the taste and preservation of soy, and most probably to improve the nutritional quality of boiled soybeans. Indeed, the first remains of fermented soy were found with the first earthenware produced in Korea [9]. This production appeared to occur during the Bronze age. Tempeh, miso and natto fermentations allow the taste of soybeans to be improved and allow an increase in the vitamins of the group B [30]. This could compensate the losses of these water-soluble micronutrients during water treatments. Interestingly, Puri et al. [31] mentioned that soybean sterilization was mandatory prior to inoculation of soy matter and prior to fermentation. They reported that boiling considerably reduced IFs in subsequent fermented soy food. Sterilization via boiling was the process traditionally applied in ancient times. Water contact with seeds favored the leakage of glycosylated IFs into the cooking water. Nowadays, to save time and energy, sterilization processes are mainly performed by steaming under pressure using an autoclave [30,32]. However, it appears that steaming reduces the contact between soy matter and water maintaining larger amounts of IFs in final foodstuffs [33]. Consequently, the estrogenic IFs levels are higher in steamed food than in boiled food. Additionally, while this effect may be beneficial when IFs are required and useful, it constitutes a risk when they reach excessive levels.

4.3. *Health Effects*

4.3.1. Traditional Pharmacopeia

Soy's polyphenols content has been studied thoroughly. Thus, soybean is known to contain many functional substances such as phenolic acids, flavonoids, IFs, stilbenes, saponins [30], or the peptide lunasin. According to Baraibar-Norberg and Deutsh [7] soybean first appeared in the traditional Chinese pharmacopeia. The oldest treatise reporting soy medicinal properties that is still available is the *Shiliao bencao* (*Compendium of Diet Therapy [Materia Dietetica]*), which was written by Meng Shen in 670 CE. It states that "*soybeans boiled into a liquid form, can eradicate poisons from the system and cure gastric fever, paralysis, pains, difficulty in passage of urine and other bladder troubles. It can also improve circulation of the blood, improper heart, liver, kidneys and stomach function and even remedy the chills*" [7,34]. It should be noted that none of these health effects rely on estrogenic properties. This observation suggests that IFs were not present in these preparations, confirming that traditional processes tended to eliminate them. By comparison, kudzu roots, which contain GEN, DAI and puerarin, among other substances, were known to induce some estrogenic and antipyretic effects. The latter are now attributed to IFs [35] and suggest that the traditional preparation of kudzu preserved GEN and DAI levels [35]. These facts put together with data obtained on IFs content of soy foods prepared following traditional recipes, strongly suggest that the exposure to IFs was low in the pre-industrialized period. It probably increased in human populations with the advent of industrial processing.

4.3.2. Toxic Effects and Reference Doses

The reprotoxic effects of soy IFs and their methoxylated parents from clover were first documented in the 1940s when ewes were affected by the clover disease [36]. Within four years, the herds' fertility reduced from 80% to less than 5% and many farmers stopped their activities. Veterinarians then came to understand that IFs reduced the pituitary hormones FSH and LH, which themselves were decreased due to an impairment of GnRH production at hypothalamic level [37]. Such a mechanism is known to be used pharmacologically in contraceptive pills based on synthetic estrogen: ethynyl-estradiol [38]. This impairment of pituitary hormones was then observed in premenopausal women who had 45 mg dietary IFs per day [39]. More recently, it was shown that exposure to 50 mg IFs per day significantly increases the risk of luteal deficiency [40] and infertility [41] in American women. An exposure of baby girls to soy IFs via soy-based infant formulas was shown to increase the incidence of fibrosis, endometriosis, menstrual pains and bleedings [42–45]. These impairments were also observed in women who engaged in overconsumption of soy-based products [46]. Moreover, a case of hypogonadism, gynecomastia and reduced libido was observed in a Japanese man who overconsumed soymilk for four years [47].

Finally, in five studies, soy IFs in biological fluids were associated with reduced quality and quantity of semen in men [48–52].

In addition, a controversy remains regarding the effects of IFs on breast cancers. An obvious discrepancy was observed between two data sources. On one hand, there are data obtained in vitro on recognized cell models [3,53], in animals implanted with these cancer cells [3,54], in healthy premenopausal women under IFs oral supplementation [3,55] and in women with breast cancer supplemented with soy and IFs [3,56]. On the other hand, there are epidemiological studies [3,57] and reviews of randomized controlled trials (RCTs) [3,58]. The population studies tend to show a reduction in breast cancer risk in Asian women in an Asian environment (including green tea polyphenols intake) but this protection has not been clearly observed in Western women in a Western environment [3,59]. Many biases, including genetic, dietary and environmental parameters, make it difficult to reach a definitive conclusion on the effects of IFs on breast cancer incidence in Western menopausal women. The reviews including RCTs concluded that there was no increased risk of breast cancer in women taking soy or food supplements based on soy IFs [3,60]. However, the subjects recruited for these RCTs were selected based on health criteria and did not represent the real population. Moreover, the RCTs were not designed to study the incidence of breast cancers, and thus the conclusions of these analyses should be considered cautiously.

Still, there is now sufficient evidence to indicate that soy IFs can interact with human thyroid function. On top of mechanistic arguments [3], clinical cases [61,62], observation studies [63,64] and intervention studies [65] confirm that soy IFs have slight anti-thyroid effects and can worsen the status of hypothyroid patients.

Beside the effects directly reported in human beings, the USA NTP showed that GEN was reprotoxic in Sprague Dawley rats in a multi-generational study published in 2008 [4]. The study showed that some reproductive parameters, including anogenital distance in pups and litter size, were altered at a GEN dose of 35 mg/kg/day in males. According to the general toxicology rules, the NTP study can be used to define a Reference Dose in humans by applying safety factors to the level defined. Indeed, an amount of 35 $mg.kg^{-1}day^{-1}$ was the LOAEL (Lowest Observed Adverse Effect Level) and no effects were recorded on reproductive parameters at a dose of 7 $mg.kg^{-1}.day^{-1}$. Usually, when a LOAEL is available, safety factors are required, including 10 for interspecific differences between rats and humans, 10 for intraspecific differences, i.e., interindividual differences between human beings; and 1.8 or 3 for conversion of LOAEL to NOAEL. If the last factor is fixed at 1.8, the Reference Dose for GEN appears to be about 20 $mg.day^{-1}$ for an adult weighing 60 kg. In addition, as GEN is usually present in food with DAI, both IFs should be considered. Unfortunately, there is neither LOAEL nor NOAEL for DAI in rats. Additionally, even if DAI is less estrogenic than GEN, it can be converted into Equol, which may be more active than GEN on certain tissues [3]. Consequently, it would be relevant to consider the limit concentration of both GEN and DAI to be set at 20 $mg.day^{-1}$ for an adult weighing 60 kg. The present study showed that the IFs levels reached in tofu, tempeh, miso and soy juice were below this limit after water treatments. Meanwhile, this was not the case in the corresponding industrial products. For textured proteins, the treatments allowed the IFs concentrations to decrease at levels from 35 to 22 mg/100 g depending on protein sizes. This is higher than the potential Reference Dose. However, textured proteins are usually used with other ingredients and at a percentage ranging from 45 to 20%. Therefore, the final concentration in a 100g portion is lower than 20 mg. This dose is about half that which was shown to have a physiological effect in humans and seems to correspond to exposure levels in Asian populations following traditional soy food preparation processes. Finally, the various doses discussed here correspond to median modern exposures in China and Japan [15,16]. To our knowledge, they did not induce deleterious effects in the past.

To conclude this part, excess IFs should be avoided by populations which should not be exposed to estrogens. These include infants, children, premenopausal women, pregnant women, hypothyroid patients and men. Concerning women with breast cancer, it seems

difficult to advise on soy consumption, considering the estrogenic effects of IFs. The latter may prevent breast cancer occurrence, but their actions on an established breast cancer dependent on estrogens are still a subject of debate.

4.3.3. Beneficial Health Effects for a Restricted Population

If IFs should be carefully monitored, it is because they seem to be active in human beings at dietary levels. Their beneficial estrogenic effects are restricted to certain categories of populations. Indeed, the most studied effect is the reduction in menopausal symptoms. The last meta-analysis which analyzed the effects of IFs on vasomotor symptoms and criteria showed only modest effects [66]. However, other meta-analyses reported a decrease in the occurrence of hot flushes [67]. In some studies, the effects appeared to be restricted to some natural compounds such as Equol [68], while in others, only some preparation types were found to be active [69]. Nevertheless, while the latest studies tend to show that vasomotor symptoms equally affect Asian and Western peri- and post-menopausal women [70] in accordance with an Asian exposure to IFs lower than 50 mg/day, IFs are the most popular substances used worldwide to reduce hot flushes.

Besides menopausal symptoms, there is evidence that IFs doses >80 mg /day can prevent bone mineral density (BMD) decrease in menopausal women [71]. Equol is also considered to have a specific beneficial effect on BMD during menopause when administered as supplement [72]. In the latter study, an RCT gathered 76 menopausal women between 50 and 55 years old. The treatment lasted one year and the supplement contained 80 mg IFs aglycone, 10 mg of Equol and 25 mg Resveratrol as an antioxidant. In these specific conditions, the BMD was preserved significantly in the treated group compared to the placebo group.

To conclude, there are sufficient data showing that IFs may be useful for menopausal women showing no signs of thyroid or estrogen-dependent diseases.

4.4. Consequences of this Finding on the Estimation of the Population's Exposure

This study has shown that simple water treatments like water rinsing, soaking or boiling can dramatically reduce the content of IFs in soy foods. It also gathered historical testimonies indicating that such water treatments were traditional when preparing soy food in Asia, as soybeans were eaten increasingly often. The historical data show that fermentation was developed quite early as a technique to improve the digestibility, taste and preservation of soy. Looking at actual recipes reported by Asian consumers, it appears that homemade soy is generally processed following family recipes, and thus using water treatments. However, nowadays, traditional boiling in water tends to be replaced by steaming, which reduces water-to-beans contact and maintains the IFs in soy foods.

Hence, the most significant result of this work is the large differences observed between water-treated products and those found on the market due to the limited water treatments performed in the industry to reduce energy and environmental costs. Thus, the importance of checking the origin of soy foods when estimating IFs intakes is highlighted. This is particularly true in Asia, where a larger percentage of people may be prompted to prepare soy at home according to family recipes. Table 2 below shows the discrepancies between IFs plasma levels, as found in different studies where IFs intakes were estimated.

Table 2. Review of isoflavones concentrations in human plasma in relation to estimated isoflavone intakes.

Location	Estimation Method	Estimated IFs Intake (mg)	IFs Plasma (nM)	Time of Plasma Collection	Study
Los Angeles	Hawaii Food Composition Database based on commercial items	23.72	30.1	No specific instruction	[73]
China (Shanghai)	USDA-ISU database based on commercial items	23.5 58.15 84.5 284.0	106.3 119.9 146.3 188.3	No specific instruction	[74]
Japan	From [75] (commercial items)	46.4	419.2	Overnight fasting	[76]
Japan	From [75] (commercial items)	40.8	411.0	After fasting > 5h	[77]
England	USDA-ISU database based on commercial items	33.7	510	Not precise	[78]
Japan	From [75] (commercial items)	32.5 34.8	553 * 605 *	No specific Instruction	[79]
The Netherlands	Strictly controlled soy-protein diet	48.5 100.1 104.2 93.9	2160 2530 2900 4800	Fasting samples after chronic intake	[80]
Hawaii	USDA-ISU database based on commercial items	96.0	5600	12h fasting and chronic intake	[81]

* Figures calculated from values expressed in ng.mL^{-1} from [79].

As can be seen in Table 2, the plasma IFs levels are much higher in Western people under an equivalent estimation of IFs exposure. The latter were generally assessed by assaying commercial soy foods found in the corresponding local markets. Of course, at least part of the differences may come from analytical techniques and the blood sampling time. However, in Asia, if IFs are ingested on a regular basis, e.g., every day, a steady-state level may be expected. Therefore, whatever the time of sampling, the blood levels should be fairly stable. On the other hand, in Western countries, where soy intake remains irregular, the time of sampling plays a major role on blood IFs concentrations, and this may explain the low blood levels recorded in [78]. In addition, in [14], it was shown that there are no fundamental differences in IFs metabolism and pharmacokinetics between Asian and Western consumers. Thus, the intake estimation should be assessed to explain the discrepancies appearing in Table 2.

While it is likely that in the West, soy food is mainly obtained from industrial sources, as there is no tradition of soy cooking, in Asia, it is customary to prepare soy at home according to family recipes. Therefore, the intake estimation may be over-evaluated in Asia if it is only based on industrial products, as described in the works of Kimira et al. [75] and Chan et al. [82]. This comment should be considered while discussing a Reference Dose for soy IFs.

5. Conclusions

In population studies, the distinction between domestically and industrially processed soy foods has not been made, potentially leading to an overestimation of soy consumers' overall exposure to IFs. This is particularly true in Asian contexts, where traditional soy-cooking practices prevail. A comprehensive review of traditional soy-processing methods revealed a consistent practice: soybeans were invariably boiled, either as a precursor to subsequent

fermentation steps or to enhance their digestibility. This study shows that traditional and domestic processing methods significantly reduce the content of IFs in soy-based foods compared to industrial methods. The application of two or three water treatments aids in achieving the potential "Reference Dose" for IFs, established according to the toxicological studies of the US National Toxicology Program. Furthermore, the results corroborate previous findings that water treatments can significantly decrease IFs in soy foods.

Given the functional activities of IFs, their use should be judicious. One approach to achieve this goal is by controlling dietary exposure levels through the introduction of water treatments at both industrial and domestic soy food preparation stages, thereby ensuring a safer democratization of their consumption. This simple rule could indeed result in a substantial shift in the global human exposure to estrogenic substances.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/foods13070999/s1. Figure S1. Effect of water renewals during beans soaking on the level of isoflavones in Okara. A. Okara made from entire beans. B. Okara made from dehulled beans. The stars indicate a significant difference with 0 renewal ($p < 0.05$)

Author Contributions: Conceptualization, C.B.-P.; methodology, C.B.-P.; G.P. and M.D.; formal analysis, S.B. and C.B.-P.; investigation, S.B. and C.B.-P.; writing—original draft preparation, C.B.-P. and S.B.; writing—review and editing, L.C., G.P. and M.D.; supervision, C.B.-P.; project administration, C.B.-P. and G.P.; funding acquisition, C.B.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: We acknowledge ANRT for its support with the S.B. funding. We also thank Nuraili Robert for her advice regarding traditional soy cooking.

Conflicts of Interest: S.B. is a PhD student whose financial support is partially insured by Biopress®/Berkem®. G.P. is employed by the Biopress®/Berkem® company. L.C. is employed by the Berkem company. C.B-P has no conflict of interest. She did not receive funds from Biopress®/Berkem® to perform this study. Biopress®/Berkem® had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. L.C. and G.P. were only informed of the work advancement and had the opportunity to read and correct the manuscript.

References

1. Lee, A.; Bensaada, S.; Lamothe, V.; Lacoste, M.; Bennetau-Pelissero, C. Endocrine disruptors on and in fruits and vegetables: Estimation of the potential exposure of the French population. *Food Chem.* **2022**, *373*, 131513. [CrossRef] [PubMed]
2. Gier, K.; Preininger, C.; Sauer, U. A Chip for Estrogen Receptor Action: Detection of Biomarkers Released by MCF-7 Cells through Estrogenic and Anti-Estrogenic Effects. *Sensors* **2017**, *17*, 1760. [CrossRef] [PubMed]
3. Canivenc-Lavier, M.C.; Bennetau-Pelissero, C. Phytoestrogens and Health Effects. *Nutrients* **2023**, *15*, 317. [CrossRef]
4. NTP. Multigenerational reproductive study of genistein (Cas No. 446-72-0) in Sprague-Dawley rats (feed study). *Natl. Toxicol. Program. Tech. Rep. Ser.* **2008**, *539*, 1–266.
5. NTP. Toxicology and carcinogenesis studies of genistein (Cas No. 446-72-0) in Sprague-Dawley rats (feed study). *Natl. Toxicol. Program. Tech. Rep. Ser.* **2008**, *545*, 1–240.
6. Cao, Z.H.; Green-Johnson, J.M.; Buckley, N.D.; Lin, Q.Y. Bioactivity of soy-based fermented foods: A review. *Biotechnol. Adv.* **2019**, *37*, 223–238. [CrossRef] [PubMed]
7. Baraibar Norberg, M.; Deutsh, L. The first soybean cycle (domestication to 900 CE). In *The Soybean Through World History. Lessons for Sustainable Agrofood Systems*, 1st ed.; Routledge Studies in Food, Society and the Environment Taylor & Francis: Abingdon, UK, 2023; pp. 22–56.
8. Zhao, Z. New Archaeobotanic Data for the Study of the Origins of Agriculture in China. *Curr. Anthropol.* **2011**, *52*, 295–306. [CrossRef]
9. Lee, C.-H.; Kim, M.L. History of Fermented Foods in Northeast Asia. In *Ethnic Fermented Foods and Alcoholic Beverages of Asia*; Tamang, J.P., Ed.; Spinger: New Delhi, India, 2016; pp. 1–16.

10. Lander, B.; DuBois, T.D.A. History of Soy in China: From Weedy Bean to Global Commodity; Ethnic Fermented Foods and Alcoholic Beverages of Asia. In *The Age of the Soybean: An Environmental History of Soy during the Great, Acceleration*; da Silva, C.M., de Majo, C., Eds.; The White Horse Press: Cambridgeshire, UK, 2022; pp. 29–47.
11. Traditional Tofu Making Experience in Debao Guangxi. Available online: https://twojadebowls.com/2013/08/16/traditional-tofu-making-experience-in-debao-guangxi/ (accessed on 27 February 2024).
12. Bensaada, S.; Chabrier, F.; Ginisty, P.; Ferrand, C.; Peruzzi, G.; Valat, M.; Bennetau-Pelissero, C. Improved Food-Processing Techniques to Reduce Isoflavones in Soy-Based Foodstuffs. *Foods* **2023**, *12*, 1540. [CrossRef]
13. McClure, T.; Cocuron, J.C.; Osmark, V.; McHale, L.K.; Alonso, A.P. Impact of Environment on the Biomass Composition of Soybean (*Glycine max*) seeds. *J. Agric. Food Chem.* **2017**, *65*, 6753–6761. [CrossRef]
14. Vergne, S.; Sauvant, P.; Lamothe, V.; Chantre, P.; Asselineau, J.; Perez, P.; Durand, M.; Moore, N.; Bennetau-Pelissero, C. Influence of ethnic origin (Asian v. Caucasian) and background diet on the bioavailability of dietary isoflavones. *Br. J. Nutr.* **2009**, *102*, 1642–1653. [CrossRef]
15. Zhu, J.; Zhao, Q.; Qiu, Y.; Zhang, Y.; Cui, S.; Yu, Y.; Chen, B.; Zhu, M.; Wang, N.; Liu, X.; et al. Soy Isoflavones Intake and Obesity in Chinese Adults: A Cross-Sectional Study in Shanghai, China. *Nutrients* **2021**, *13*, 2715. [CrossRef] [PubMed]
16. Murai, U.; Sawada, N.; Charvat, H.; Inoue, M.; Yasuda, N.; Yamagishi, K.; Tsugane, S. Soy product intake and risk of incident disabling dementia: The JPHC Disabling Dementia Study. *Eur. J. Nutr.* **2022**, *61*, 4045–4057. [CrossRef] [PubMed]
17. Fernandez-Lopez, A.; Lamothe, V.; Delample, M.; Denayrolles, M.; Bennetau-Pelissero, C. Removing isoflavones from modern soyfood: Why and how? *Food Chem.* **2016**, *210*, 286–294. [CrossRef] [PubMed]
18. The Japanese Lab. How Miso Was Made Traditionally. Available online: https://thejapanesefoodlab.com/traditional-japanese-miso#how/ (accessed on 27 February 2024).
19. Afssa, Afssaps. Sécurité et Bénéfices des Phyto-Estrogènes Apportés par L'alimentation—Recommandations Rapport D'expertise Mars 2005. Available online: https://www.anses.fr/fr/system/files/NUT-Ra-Phytoestrogenes.pdf (accessed on 23/03/24).
20. Bennetau-Pelissero, C.; Arnal-Schnebelen, B.; Lamothe, V.; Sauvant, P.; Sagne, J.; Verbruggen, M.; Mathey, J.; Lavialle, O. ELISA as a new method to measure genistein and daidzein in food and human fluids. *Food Chem.* **2003**, *82*, 645–658. [CrossRef]
21. Bensaada, S.; Raymond, I.; Breton, M.; Pellegrin, I.; Viallard, J.-F.; Bennetau-Pelissero, C. Development of an Assay for Soy Isoflavones in Women's Hair. *Nutrients* **2022**, *14*, 3619. [CrossRef] [PubMed]
22. Bensaada, S.; Raymond, I.; Pellegrin, I.; Viallard, J.F.; Bennetau-Pelissero, C. Validation of ELISAs for Isoflavones and Enterolactone for Phytoestrogen Intake Assessment in the French Population. *Nutrients* **2023**, *15*, 967. [CrossRef] [PubMed]
23. Vergne, S.; Titier, K.; Bernard, V.; Asselineau, J.; Durand, M.; Lamothe, V.; Potier, M.; Perez, P.; Demotes-Mainard, J.; Chantre, P.; et al. Bioavailability and Urinary Excretion of Isoflavones in Humans: Effects of Soy-Based Supplements Formulation and Equol Production. *J. Pharm. Biomed. Anal.* **2007**, *43*, 1488–1494. [CrossRef]
24. Shinkaruk, S.; Durand, M.; Lamothe, V.; Carpaye, A.; Martinet, A.; Chantre, P.; Vergne, S.; Nogues, X.; Moore, N.; Bennetau-Pelissero, C. Bioavailability of Glycitein Relatively to Other Soy Isoflavones in Healthy Young Caucasian Men. *Food Chem.* **2012**, *135*, 1104–1111. [CrossRef]
25. Bennetau-Pelissero, C.; Le Houérou, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Synthesis of Haptens and Conjugates for ELISAs of Phytoestrogens. Development of the Immunological Tests. *J. Agric. Food Chem.* **2000**, *48*, 305–311. [CrossRef]
26. Le Houérou, C.; Bennetau-Pelissero, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Syntheses of Novel Hapten-Protein Conjugates for Production of Highly Specific Antibodies to Formononetin, Daidzein and Genistein. *Tetrahedron* **2000**, *56*, 295–301. [CrossRef]
27. Bennetau-Pelissero, C. Plant Proteins from Legumes. In *Bioactive Molecules in Food Vol 1. Reference Series in Phytochemistry*; Mérillon, J.M., Ramawat, K., Eds.; Springer: Cham, Switzerland, 2018; pp. 223–266.
28. Veremeichik, G.; Grigorchuk, V.; Butovets, E.; Lukyanchuk, L.; Brodovskaya, E.; Bulgakov, D.; Bulgakov, V. Isoflavonoid biosynthesis in cultivated and wild soybeans grown in the field under adverse climate conditions. *Food Chem.* **2021**, *342*, 128292–128300. [CrossRef] [PubMed]
29. Kao, T.H.; Lu, Y.F.; Hsieh, H.C.; Chen, B.H. Stability of isoflavone glucosides during processing of soymilk and tofu. *Food Res. Int.* **2004**, *37*, 891–900. [CrossRef]
30. Chung, I.M.; Seo, S.H.; Ahn, J.K.; Kim, S.H. Effect of processing, fermentation, and aging treatment to content and profile of phenolic compounds in soybean seed, soy curd and soy paste. *Food Chem.* **2011**, *127*, 960–967. [CrossRef] [PubMed]
31. Puri, A.; Mir, S.R.; Panda, B.P. Effect of sequential bio-processing conditions on the content and composition of vitamin K2 and isoflavones in fermented soy food. *J. Food Sci. Technol.* **2015**, *52*, 8228–8235. [CrossRef] [PubMed]
32. Qiao, Y.; Zhang, K.; Zhang, Z.; Zhang, C.; Sun, Y.; Feng, Z. Fermented soybean foods: A review of their functional components, mechanism of action and factors influencing their health benefits. *Food Res. Int.* **2022**, *158*, 111575. [CrossRef] [PubMed]
33. Kuligowski, M.; Sobkowiak, D.; Polanowska, K.; Jasinska-Kuligowska, I. Effect of different processing methods on isoflavone content in soybeans and soy products. *J. Food Comp. Anal.* **2022**, *110*, 104535–104541. [CrossRef]
34. Shurtleff, W.; Aoyagi, A. *Early History of Soybeans and Soyfoods Worldwide (1024 BCE to 1899): Extensively Annotated Bibliography and Sourcebook*; Soyinfo Center: Lafayette, CA, USA, 2014.
35. Wang, S.; Zhang, S.; Wang, S.; Gao, P.; Dai, L. A comprehensive review on Pueraria: Insights on its chemistry and medicinal value. *Biomed. Pharmacother.* **2020**, *131*, 110734. [CrossRef] [PubMed]

36. Bennetts, H.W.; Underwood, E.J.; Shier, F.L. A specific breeding problem of sheep on subterranean clover pastures in Western Australia. *Aust. Vet. J.* **1946**, *23*, 2–13. [CrossRef]
37. Findlay, J.K.; Buckmaster, J.M.; Chamley, W.A.; Cumming, I.A.; Hearnshaw, H.; Goding, J.R. Release of luteinising hormone by œstradiol 17β and a gonadotrophin-releasing hormone in ewes affected with clover disease. *Neuroendocrinology* **1973**, *11*, 57–66. [CrossRef]
38. Cohen, B.L.; Katz, M. Further studies on pituitary and ovarian function in women receiving hormonal contraception. *Contraception* **1981**, *24*, 159–172. [CrossRef]
39. Cassidy, A.; Bingham, S.; Setchell, K.D. Biological effects of a diet of soy protein rich in isoflavones on the menstrual cycle of premenopausal women. *Am. J. Clin. Nutr.* **1994**, *60*, 333–340. [CrossRef] [PubMed]
40. Andrews, M.A.; Schliep, K.C.; Wactawski-Wende, J.; Stanford, J.B.; Zarek, S.M.; Radin, R.G.; Sjaarda, L.A.; Perkins, N.J.; Kalwerisky, R.A.; Hammoud, A.O.; et al. Dietary factors and luteal phase deficiency in healthy eumenorrheic women. *Hum. Reprod.* **2015**, *30*, 1942–1951. [CrossRef] [PubMed]
41. Jacobsen, B.K.; Jaceldo-Siegl, K.; Knutsen, S.F.; Fan, J.; Oda, K.; Fraser, G.E. Soy isoflavone intake and the likelihood of ever becoming a mother: The Adventist Health Study-2. *Int. J. Womens Health* **2014**, *6*, 377–384. [CrossRef] [PubMed]
42. Strom, B.L.; Schinnar, R.; Barnhart, K.T.; Sammel, M.D.; Macones, G.A.; Stallings, V.A. Exposure to soy-based formula in infancy and endocrinological and reproductive outcomes in young adulthood. *JAMA* **2001**, *286*, 807–814. [CrossRef]
43. Upson, K.; Harmon, Q.E.; Laughlin-Tommaso, S.K.; Umbach, D.M.; Baird, D.D. Soy-based Infant Formula Feeding and Heavy Menstrual Bleeding Among Young African American Women. *Epidemiology* **2016**, *27*, 716–725. [CrossRef] [PubMed]
44. Upson, K.; Adgent, M.A.; Wegienka, G.; Baird, D.D. Soy-based infant formula feeding and menstrual pain in a cohort of women aged 23–35 years. *Hum. Reprod.* **2019**, *34*, 148–154. [CrossRef] [PubMed]
45. Qin, H.; Lin, Z.; Vásquez, E.; Luan, X.; Guo, F.; Xu, L. High soy isoflavone or soy-based food intake during infancy and in adulthood is associated with an increased risk of uterine fibroids in premenopausal women: A meta-analysis. *Nutr. Res.* **2019**, *71*, 30–42. [CrossRef]
46. Chandrareddy, A.; Muneyyirci-Delale, O.; McFarlane, S.I.; Murad, O.M. Adverse effects of phytoestrogens on reproductive health: A report of three cases. *Complement. Ther. Clin. Pract.* **2008**, *14*, 132–135. [CrossRef]
47. Imai, H.; Nishikawa, H.; Suzuki, A.; Kodama, E.; Iida, T.; Mikura, K.; Hashizume, M.; Kigawa, Y.; Tadokoro, R.; Sugisawa, C.; et al. Secondary Hypogonadism due to Excessive Ingestion of Isoflavone in a Man. *Intern. Med.* **2022**, *61*, 2899–2903. [CrossRef]
48. Chavarro, J.E.; Toth, T.L.; Sadio, S.M.; Hauser, R. Soy food and isoflavone intake in relation to semen quality parameters among men from an infertility clinic. *Hum. Reprod.* **2008**, *23*, 2584–2590. [CrossRef]
49. Toshima, H.; Suzuki, Y.; Imai, K.; Yoshinaga, J.; Shiraishi, H.; Mizumoto, Y.; Hatakeyama, S.; Onohara, C.; Tokuoka, S. Endocrine disrupting chemicals in urine of Japanese male partners of subfertile couples: A pilot study on exposure and semen quality. *Int. J. Hyg. Environ. Health* **2012**, *215*, 502–506. [CrossRef]
50. Xia, Y.; Chen, M.; Zhu, P.; Lu, C.; Fu, G.; Zhou, X.; Chen, D.; Wang, H.; Hang, B.; Wang, S.; et al. Urinary phytoestrogen levels related to idiopathic male infertility in Chinese men. *Environ. Int.* **2013**, *59*, 161–167. [CrossRef] [PubMed]
51. Mumford, S.L.; Kim, S.; Chen, Z.; Boyd Barr, D.; Buck Louis, G.M. Urinary Phytoestrogens Are Associated with Subtle Indicators of Semen Quality among Male Partners of Couples Desiring Pregnancy. *J. Nutr.* **2015**, *145*, 2535–2541. [CrossRef]
52. Yuan, G.; Liu, Y.; Liu, G.; Wei, L.; Wen, Y.; Huang, S.; Guo, Y.; Zou, F.; Cheng, J. Associations between semen phytoestrogens concentrations and semen quality in Chinese men. *Environ. Int.* **2019**, *129*, 36–44. [CrossRef] [PubMed]
53. Schmitt, E.; Dekant, W.; Stopper, H. Assaying the estrogenicity of phytoestrogens in cells of different estrogen sensitive tissues. *Toxicol. Vitr.* **2001**, *15*, 433–439. [CrossRef]
54. Wu, Q.; Yang, Y.; Yu, J.; Jin, N. Soy isoflavone extracts stimulate the growth of nude mouse xenografts bearing estrogen-dependent human breast cancer cells (MCF-7). *J. Biomed. Res.* **2012**, *26*, 44–52. [CrossRef] [PubMed]
55. McMichael-Phillips, D.F.; Harding, C.; Morton, M.; A Roberts, S.; Howell, A.; Potten, C.S.; Bundred, N.J. Effects of soy-protein supplementation on epithelial proliferation in the histologically normal human breast. *Am. J. Clin. Nutr.* **1998**, *68* (Suppl. S6), 1431S–1435S. [CrossRef]
56. Shike, M.; Doane, A.S.; Russo, L.; Cabal, R.; Reis-Filo, J.; Gerald, W.; Cody, H.; Khanin, R.; Bromberg, J.; Norton, L. The effects of soy supplementation on gene expression in breast cancer: A randomized placebo-controlled study. *J. Natl. Cancer Inst.* **2014**, *106*, dju189. [CrossRef]
57. Van Die, M.D.; Bone, K.M.; Visvanathan, K.; Kyrø, C.; Aune, D.; Ee, C.; Paller, C.J. Phytonutrients and outcomes following breast cancer: A systematic review and meta-analysis of observational studies. *JNCI Cancer Spectr.* **2024**, *8*, pkad104. [CrossRef]
58. Messina, M. Impact of Soy Foods on the Development of Breast Cancer and the Prognosis of Breast Cancer Patients. *Forsch. Komplementmed.* **2016**, *23*, 75–80. [CrossRef]
59. Chen, M.; Rao, Y.; Zheng, Y.; Wei, S.; Li, Y.; Guo, T.; Yin, P. Association between soy isoflavone intake and breast cancer risk for pre- and post-menopausal women: A meta-analysis of epidemiological studies. *PLoS ONE* **2014**, *9*, e89288. [CrossRef] [PubMed]
60. Finkeldey, L.; Schmitz, E.; Ellinger, S. Effect of the Intake of Isoflavones on Risk Factors of Breast Cancer-A Systematic Review of Randomized Controlled Intervention Studies. *Nutrients* **2021**, *13*, 2309. [CrossRef] [PubMed]
61. Caprio, A.M.; Umano, G.R.; Luongo, C.; Aiello, F.; Iacono, I.D.; Palumbo, S.; del Giudice, E.M.; Grandone, A. Case report: Goiter and overt hypothyroidism in an iodine-deficient toddler on soy milk and hypoallergenic diet. *Front. Endocrinol.* **2022**, *13*, 927726. [CrossRef]

62. Fruzza, A.G.; Demeterco-Berggren, C.; Jones, K.L. Unawareness of the effects of soy intake on the management of congenital hypothyroidism. *Pediatrics* **2012**, *130*, e699–e702. [CrossRef]
63. Conrad, S.C.; Chiu, H.; Silverman, B.L. Soy formula complicates management of congenital hypothyroidism. *Arch. Dis. Child.* **2004**, *89*, 37–40. [CrossRef] [PubMed]
64. Fan, Y.; Qian, H.; Wu, Z.; Li, Z.; Li, X.; Zhang, Y.; Xu, Q.; Lu, C.; Wang, X. Exploratory analysis of the associations between urinary phytoestrogens and thyroid hormones among adolescents and adults in the United States: National Health and Nutrition Examination Survey 2007–2010. *Environ. Sci. Pollut. Res.* **2022**, *29*, 2974–2984. [CrossRef] [PubMed]
65. Sathyapalan, T.; Manuchehri, A.M.; Thatcher, N.J.; Rigby, A.S.; Chapman, T.; Kilpatrick, E.S.; Atkin, S.L. The effect of soy phytoestrogen supplementation on thyroid status and cardiovascular risk markers in patients with subclinical hypothyroidism: A randomized, double-blind, crossover study. *J. Clin. Endocrinol. Metab.* **2011**, *96*, 1442–1449. [CrossRef]
66. Guo, P.-P.; Li, P.; Zhang, X.-H.; Liu, N.; Wang, J.; Chen, D.-D.; Sun, W.-J.; Zhang, W. Complementary and alternative medicine for natural and treatment-induced vasomotor symptoms: An overview of systematic reviews and meta-analyses. *Complement. Ther. Clin. Pract.* **2019**, *36*, 181–194. [CrossRef]
67. Chen, M.N.; Lin, C.C.; Liu, C.F. Efficacy of phytoestrogens for menopausal symptoms: A meta-analysis and systematic review. *Climacteric* **2015**, *18*, 260–269. [CrossRef]
68. Daily, J.W.; Ko, B.-S.; Ryuk, J.; Liu, M.; Zhang, W.; Park, S. Equol Decreases Hot Flashes in Postmenopausal Women: A Systematic Review and Meta-Analysis of Randomized Clinical Trials. *J. Med. Food* **2019**, *22*, 127–139. [CrossRef]
69. Taku, K.; Melby, M.K.; Kronenberg, F.; Kurzer, M.S.; Messina, M. Extracted or synthesized soybean isoflavones reduce menopausal hot flash frequency and severity: Systematic review and meta-analysis of randomized controlled trials. *Menopause* **2012**, *19*, 776–790. [CrossRef] [PubMed]
70. Islam, R.M.; Bell, R.J.; Rizvi, F.; Davis, S.R. Vasomotor symptoms in women in Asia appear comparable with women in Western countries: A systematic review. *Menopause* **2017**, *24*, 1313–1322. [CrossRef]
71. Taku, K.; Melby, M.K.; Takebayashi, J.; Mizuno, S.; Ishimi, Y.; Omori, T.; Watanabe, S. Effect of soy isoflavone extract supplements on bone mineral density in menopausal women: Meta-analysis of randomized controlled trials. *Asia Pac. J. Clin. Nutr.* **2010**, *19*, 33–42. [PubMed]
72. Corbi, G.; Nobile, V.; Conti, V.; Cannavo, A.; Sorrenti, V.; Medoro, A.; Scapagnini, G.; Davinelli, S. Equol and Resveratrol Improve Bone Turnover Biomarkers in Postmenopausal Women: A Clinical Trial. *Int. J. Mol. Sci.* **2023**, *24*, 12063. [CrossRef]
73. Wu, A.H.; Yu, M.C.; Tseng, C.C.; Twaddle, N.C.; Doerge, D.R. Plasma isoflavone levels versus self-reported soy isoflavone levels in Asian-American women in Los Angeles County. *Carcinogenesis* **2004**, *25*, 77–81. [CrossRef] [PubMed]
74. Frankenfeld, C.L.; Lampe, J.W.; Shannon, J.; Gao, D.L.; Ray, R.M.; Prunty, J.; Kalhorn, T.F.; Wähälä, K.; E Patterson, R.; Thomas, D.B. Frequency of soy food consumption and serum isoflavone concentrations among Chinese women in Shanghai. *Public Health Nutr.* **2004**, *7*, 765–772. [CrossRef] [PubMed]
75. Kimira, M.; Arai, Y.; Shimoi, K.; Watanabe, S. Japanese intake of flavonoids and isoflavonoids from foods. *J. Epidemiol.* **1998**, *8*, 168–175. [CrossRef] [PubMed]
76. Arai, Y.; Watanabe, S.; Kimira, M.; Shimoi, K.; Mochizuki, R.; Kinae, N. Dietary intakes of flavonols, flavones and isoflavones by Japanese women and the inverse correlation between quercetin intake and plasma LDL cholesterol concentration. *J. Nutr.* **2000**, *130*, 2243–2250. [CrossRef]
77. Yamamoto, S.; Sobue, T.; Sasaki, S.; Kobayashi, M.; Tsugane, S.; Arai, Y.; Uehara, M.; Adlercreutz, H.; Watanabe, S.; Takahashi, T.; et al. Validity and reproducibility of a self-administered food-frequency questionnaire to assess isoflavone intake in a japanese population in comparison with dietary records and blood and urine isoflavones. *J. Nutr.* **2001**, *131*, 2741–2747. [CrossRef]
78. Verkasalo, P.K.; Appleby, P.N.; Allen, N.E.; Davey, G.; Adlercreutz, H.; Key, T.J. Soya intake and plasma concentrations of daidzein and genistein: Validity of dietary assessment among eighty British women (Oxford arm of the European Prospective Investigation into Cancer and Nutrition). *Br. J. Nutr.* **2001**, *86*, 415–421. [CrossRef]
79. Iwasaki, M.; Inoue, M.; Otani, T.; Sasazuki, S.; Kurahashi, N.; Miura, T.; Yamamoto, S.; Tsugane, S. Plasma isoflavone level and subsequent risk of breast cancer among Japanese women: A nested case-control study from the Japan Public Health Center-based prospective study group. *J. Clin. Oncol.* **2008**, *26*, 1677–1683. [CrossRef] [PubMed]
80. van der Velpen, V.; Hollman, P.C.; van Nielen, M.; Schouten, E.G.; Mensink, M.; Veer, P.V.; Geelen, A. Large inter-individual variation in isoflavone plasma concentration limits use of isoflavone intake data for risk assessment. *Eur. J. Clin. Nutr.* **2014**, *68*, 1141–1147. [CrossRef] [PubMed]
81. Gardner, C.D.; Chatterjee, L.M.; Franke, A.A. Effects of isoflavone supplements vs. soy foods on blood concentrations of genistein and daidzein in adults. *J. Nutr. Biochem.* **2009**, *20*, 227–234. [CrossRef]
82. Chan, S.G.; Murphy, P.A.; Ho, S.C.; Kreiger, N.; Darlington, G.; So, E.K.F.; Chong, P.Y.Y. Isoflavonoid content of Hong Kong soy foods. *J. Agric. Food Chem.* **2009**, *57*, 5386–5390. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

IV.3. Etude cellulaire du GPER

IV.3.1. *Docking*

IV.3.1.1. Construction du modèle de récepteur

Dans la littérature, plusieurs protéines ont été utilisées comme références pour la construction du GPER par homologie, notamment la rhodopsine bovine, la protéine ADRB2 et la chémokine CXCR4 [249]. Cependant, chacune de ces références présente des limitations. La rhodopsine bovine, bien qu'ayant une homologie de 44% et seulement 4 gaps (régions d'alignement absentes dans l'une des séquences comparées, le plus souvent correspondant à une insertion ou une délétion) avec le GPER, présente une conformation bovine très différente et une identité de seulement 20,1 % [226].

La recherche BLAST 3D sur la base *Uniprot* a identifié 244 hits, dont 235 protéines revues. En limitant la recherche à l'humain, 13 résultats ont été obtenus et classés selon 3 critères : identité (%), score et *e-value*. Les trois premières protéines présentant aussi des structures 3D ont été analysées.

CXCR-1 : 4 structures RMN sont disponibles, dont une seule contient la quasi-totalité de la séquence. Cependant cette structure présente une dizaine de conformères « tordues » et a été écartée.

CX3CRl : 2 structures obtenues par microscopie électronique (7XBX et 7XBW) sont disponibles. Malgré les scores élevés de 7XBW (49% d'homologie, 31,4% d'identité et 16 gaps), le *protein_report* indique que les deux structures présentent de nombreuses chaines manquantes ou invalides.

CXCR-2 : 29% d'identité, 48% d'homologie et 17 gaps. Trois structures 3D disponibles, dont 2 homologues identiques : 6LFM (dimère du ligand endogène) et 6LFO (monomère du ligand endogène), obtenues par microscopie électronique, et 6LFL, obtenue par cristallographie avec un inhibiteur allostérique et un lysozyme T4. Le lyzozyme T4 est souvent utilisé comme marqueur et aide à la cristallisation notamment des GPCR. Le ligand endogène, le CXCL8 (interleukine-8), existe de manière réversible sous forme de monomères ou de dimères *in vivo*, et les deux formes sont fonctionnelles sur le CXCR2 [250].

L'analyse des structures 3D des protéines homologues au GPER est décrite dans le tableau 13.

Conformère	Résultat de l'alignement
7XBX	La structure correspond au FASTA à partir du résidu Y124 (7XBX) qui équivaut à Y22 dans la protéine humaine (CX3C1), aucun gap n'est présent à partir de là. Le *Protein_report* de Discovery Studio® révèle de nombreux résidus incomplets ou invalides dans les chaines latérales Les deux premiers tours d'hélice de la structure, correspondent à 9 résidus, ne figurent pas dans l'homologie avec CX3C1. Structure Non retenu
7XBW	La structure correspond à la séquence FASTA d'origine (CX3C1). Aucun gap n'est présent et seulement 3 mutations sont observées. Le *Protein_report* indisue de nombreux résidus incomplets à partir du résidu 23 Structure Non retenu
6LFL	La structure présente seulement 3 gaps correspondant à la boucle extracellulaire par rapport à la séquence FASTA du CXCR-2. Le *Protein_report* indique peu de résidus invalides. Cependant, ce conformère diffère du reste en raison de la présence d'un inhibiteur intracellulaire et du lysozyme T4, qui empêche l'activation du récepteur. Cette structure correspond à la forme inactive du récepteur.
6LFO (avec un monomère de CXCL8) et 6LFM (avec un dimère de CXCL8)	Les structures présentent un alignement parfait avec la séquence FASTA d'origine (CXCR-2). Le *Protein_report* indique seulement deux résidus invalides. Les structures correspondent à la forme active en raison de la présence de l'interleukine 8 (CXCL8), un agoniste du récepteur, dans la structure.

Tableau 13. Résultats de l'analyse des structures 3D des protéines homologues au GPER

Les structures 3D 6LFL et 6LFO ont été retenues pour représenter respectivement les formes inactive et active du récepteur GPER. Ce choix s'explique par les caractéristiques suivantes.

- Complétude des structures : faible nombre de résidus invalides et numérotation conforme à la séquence FASTA d'origine.
- Simplicité du ligand : le conformère 6LFO est préféré à 6LFM parce que le ligand associé est sous sa forme monomérique, bien que les deux structures soient équivalentes en termes de conformation protéique.

Les structures de 6LFL et 6LFO ont été minimisées et les gaps de 6LFL complétés.

Génération et classement des modèles

20 modèles de la forme active et 20 modèles de la forme inactive du GPER ont été générés. Ces modèles ont ensuite été classés en fonction de leur RMSD et de leur niveau énergétique.

Pour la forme active, les récepteurs n°17 et n°19 présentent les meilleures énergies avec des RMSD très proches. Le récepteur n°13 présente l'énergie la moins favorable.

Pour la forme inactive, les meilleures énergies sont obtenues avec les récepteurs n°1 et n°8. Le récepteur n°6 présente le niveau énergétique le moins favorable.

L'ensemble des résultats est disponible en Annexe 3.

Analyse des résidus conservés

L'analyse des résidus conservés entre le GPER et le CXCR2 révèle la présence de sept résidus cystéines communs (voir tableau 14).

Cystéine de CXCR2	Cystéine correspondante dans GPER
119	130
166	178
196	205
230	233
263	271
286	294
308	318

Tableau 14. Résidus cystéine conservés entre CXCR2 et GPER

Une analyse comparative des structures 3D du CXCR2 et du GPER révèle également des différences notables, notamment au niveau de l'hélice 3. Le résidu aspartate 332 du CXCR2, situé dans l'hélice 3, joue un rôle crucial dans l'ancrage des ligands et est donc considéré comme un élément clé de la fonction du récepteur [251]. Ce résidu est absent du GPER, suggérant un mécanisme de liaison des ligands distinct.

Néanmoins, le GPER présente un autre résidu aspartate dans l'hélice 2 (Asp 111), un glutamate (Glu 275) et une glutamine (Gln138) dans l'hélice 3. Ces résidus pourraient jouer un rôle similaire à celui de l'aspartate 332 du CXCR2 dans la liaison des ligands au GPER.

IV.3.1.2. Validation du modèle

L'analyse des résultats des courbes ROC du *rescoring* sur le *training-set* est présentée en annexe 4. Elle a permis de sélectionner les isoformes suivantes du GPER :

- 20_actif : cette isoforme présente la meilleure somme d'aires sous la courbe et obtient un score VINA satisfaisant (3e rang).
- 14_inactif : cette isoforme se distingue par une très bonne somme d'aires sous la courbe et un excellent score VINA

La Figure 17 représente une superposition des deux formes du récepteur pour comparaison structurelle.

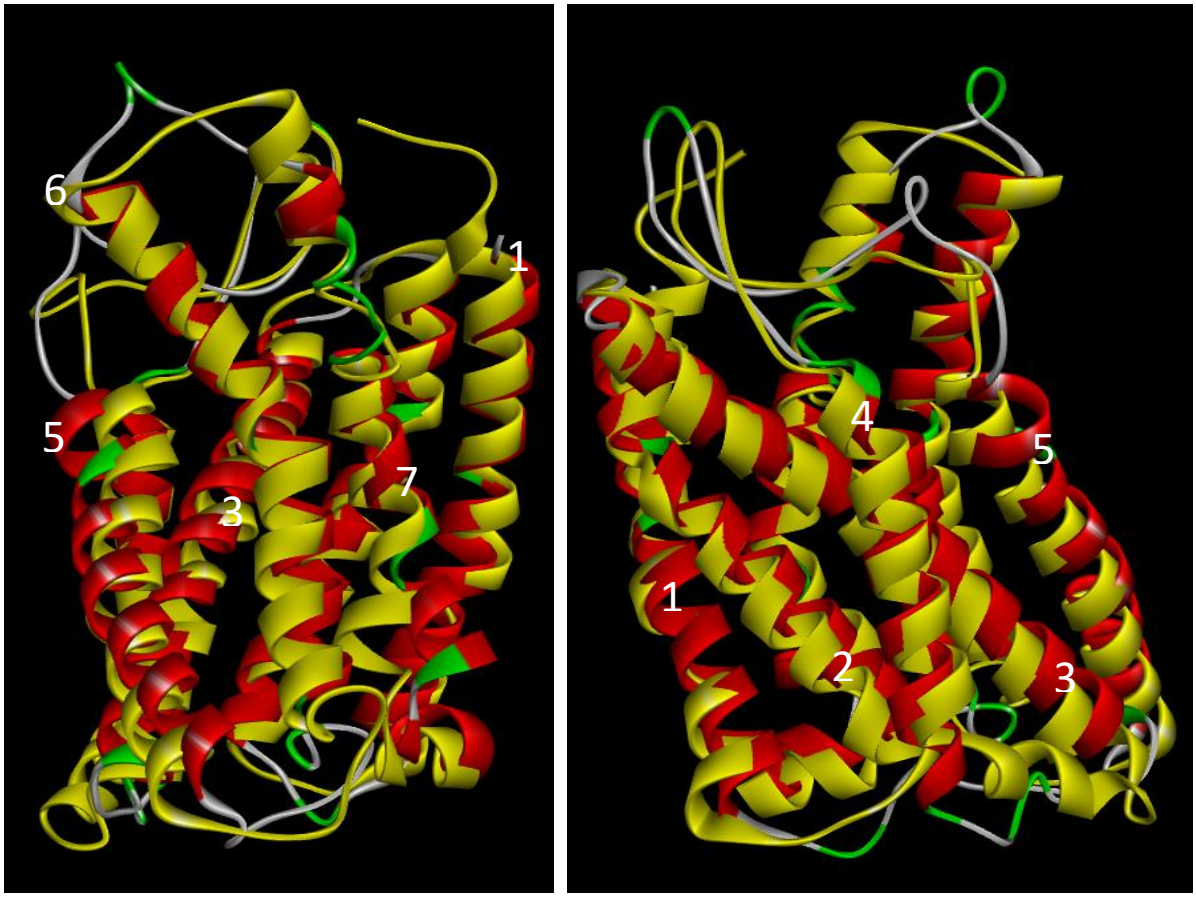


Figure 17. Comparaison des structure 3D des formes construites active (jaune) et inactive (rouge) du GPER vues sous 2 angles différents. Les numéros représentent l'ordre des hélices, la n°7 correspondant à celle liée à la protéine G

Modifications structurelles pour les formes active et inactive du GPER

L'analyse des structures 3D des formes active et inactive du GPER révèle un déplacement notable de l'hélice 5, d'environ 3 à 4 Ångströms vers l'intérieur, dans la forme active. Ce mouvement compacte la partie extra-cellulaire du récepteur. Liu *et al.* [250] ont observé un réarrangement similaire de 6 Ångströms et ont suggéré qu'il pourrait être un marqueur d'activation du récepteur.

Orientation du GPER dans la membrane plasmique

Les GPCR sont des protéines transmembranaires orientés dans la membrane plasmique avec l'extrémité amino-terminale dirigée vers l'extérieur de la cellule et l'extrémité carboxy-terminale située à l'intérieur de la cellule [166].

Zones de liaison du GPER

L'observation des poses 3D des ligands dans les deux structures du GPER sélectionnées révèle la présence de 2 zones de liaison distinctes (Figure 18). Ces zones pourraient jouer un rôle crucial dans la reconnaissance et la liaison des ligands au récepteur.

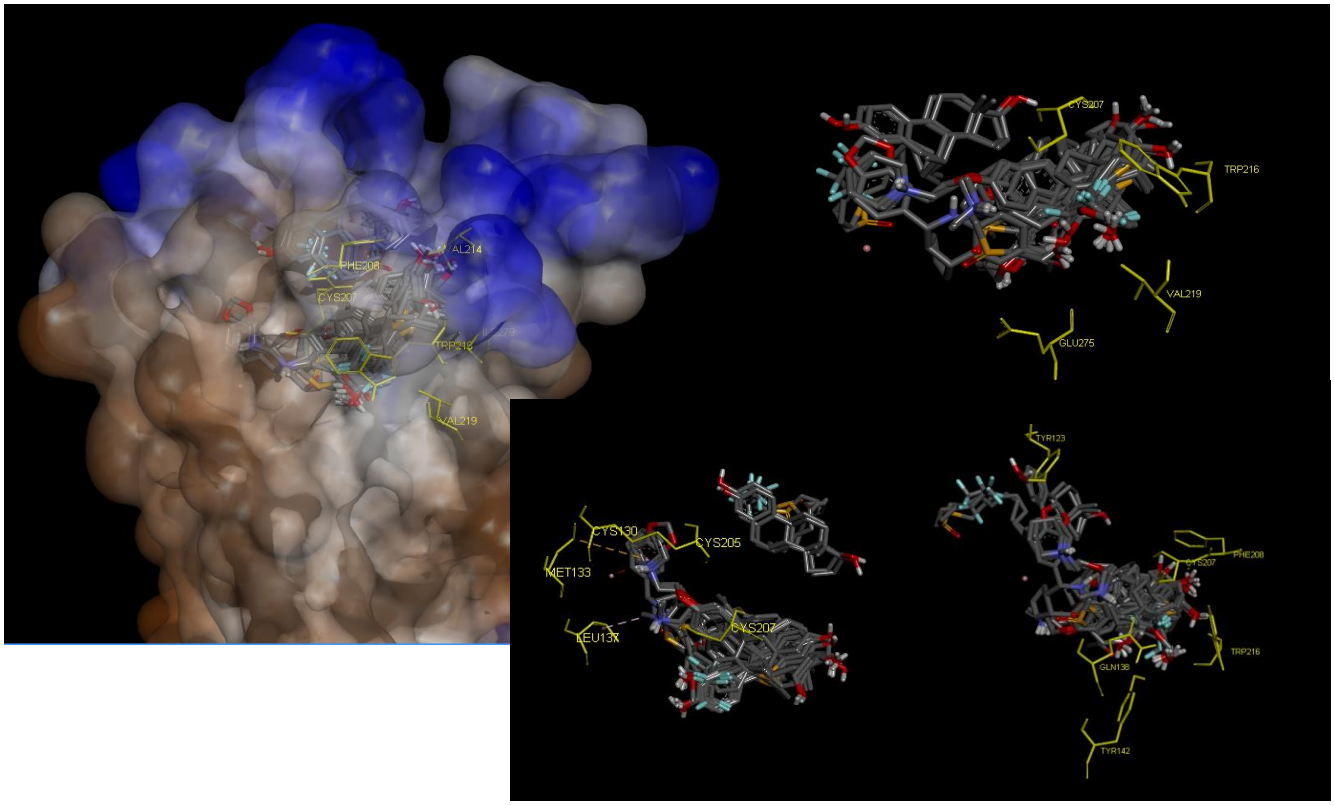


Figure 18. Résidus du GPER_activé impliqués dans l'interaction avec les agonistes du training-set

Le *docking* des agonistes a révélé que les résidus Cys207, Phe208, Trp216, Val214 et Val219 sont les plus récurrents dans le site de liaison. Ces résidus pourraient jouer un rôle crucial dans l'activation du récepteur. De plus, Cys130, Met133, Leu137, Pro192 et Cys 205, semblent également impliqués dans les interactions agonistes-récepteur. Le résidu Gln138 pourrait également être impliqué, dans une moindre mesure.

La deuxième poche de liaison a été occupée uniquement par le fulvestrant et ne possède aucun résidu en commun avec le premier cluster d'agonistes.

Le *docking* des antagonistes et molécules inactives a mis en évidence les résidus Gln82, Cys151, Arg230 comme étant essentiels pour les interactions de la forme inactive du récepteur. Glu162, semble également impliqué dans ces interactions, mais dans une moindre mesure.

Rôle de Cys207 dans l'activation du récepteur

D'après Martinez-Muñoz *et al.* [183], Cys207 pourrait jouer un rôle clé dans le réarrangement conformationnel des hélices 3 et 5, menant à l'activation du récepteur. De plus, ce résidu et Gly306 étaient présents dans le site de reconnaissance de l'agoniste G1 mais absents pour l'antagoniste G15, suggérant leur importance pour la signalisation. Ces mêmes auteur·es ont suggéré que les molécules antagonistes de GPER, comme G15 et G36 n'ont pas d'atome d'acétyle ou d'oxygène, ce qui les empêche d'interagir avec le groupe aminé présent dans le site de liaison des agonistes. Cette absence d'interaction pourrait expliquer leur activité antagoniste [183].

Identification des fonctions de score pour le *docking*

Les résultats des ACP sont rapportés dans le tableau en annexe 5. Cette analyse a permis d'identifier les fonctions de score les plus pertinentes pour effectuer le

consensus pour le *docking* des molécules d'intérêt. Elles sont présentées dans le tableau 15.

Isoforme du GPER	Fonctions de score empiriques	Fonctions de score par potentiel statistique
20_actif	Xscore HPS	DC-pair-surf
14_inactif	Xscore AV	DC-pair-surf

Tableau 15. Fonctions de score retenues pour le consensus

IV.3.1.3. *Docking* des isoflavones et conjugués et *Rescoring*

Le *docking* des substances d'intérêt a été restreint à la *small-box* aux dimensions suivantes : 28, 24 et 36 Ångströms (selon les axes x, y et z), représentée dans la Figure 19. Elle comprend une partie extracellulaire et une partie intramembranaire.

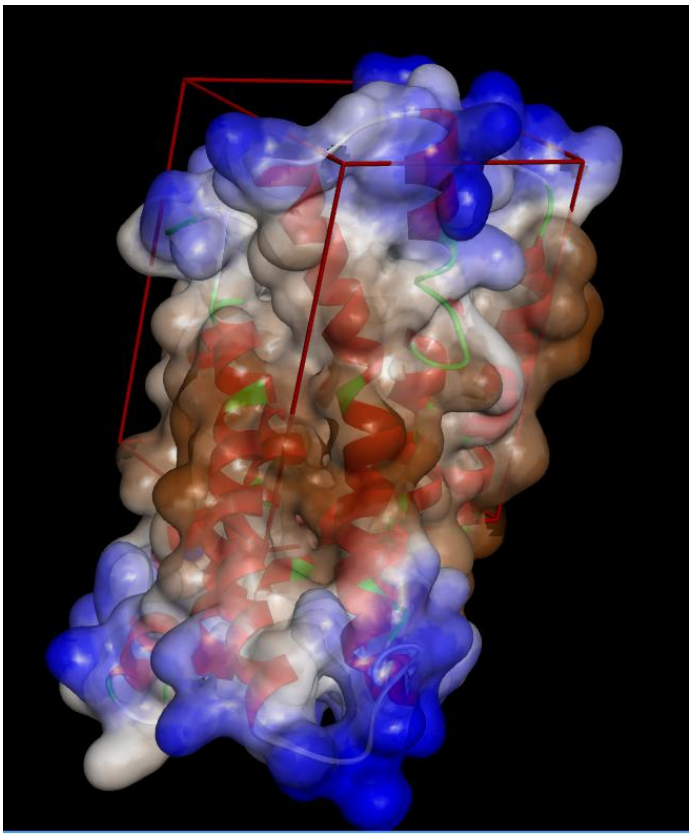


Figure 19. Construction de la structure 3D du GPER avec représentation de sa surface avec en bleu les parties hydrophiles (compartiments extra- et intra-cellulaires) et en marron la partie hydrophobe (intramembrannaire); small-box = grille rouge

L'affinité des IFs et de leurs conjugués glucuronides pour le récepteur GPER a été évaluée par *docking* moléculaire sur une base de données de 5340 composés naturels proches de la GEN. Les résultats semblent indiquer que les IFs et leurs conjugués ne présentent pas une affinité élevée pour le GPER.

En effet, d'après les résultats globaux du *docking*, les IFs et leurs conjugués n'apparaissent pas dans le top 100 des composés ayant la meilleure affinité pour le GPER. Plus précisément, pour dc-pair-surf, les positions les plus élevées pour les IFs et leurs conjugués glucuronides sont 115 pour GEN 4'-O-glucuronide, 136 pour Gly 4-O-glucuronide et 140 pour équol 7-O-glucuronide. Pour la fonction de score Vina, GEN 4'-S-7-glucuronide se classe en 630e position, suivi de DAI-7-O-glucuronide en

811e position. Pour la fonction de score HPS, Gly 7-O-glucuronide se classe seulement en 1492e place.

Il faut cependant noter que le modèle construit *in silico* peut être un peu différent de la protéine elle-même. En effet, les tests d'affinité *in vitro* publiés avec les composés aglycones contredisent ces données [40,252]. Ils tendent à montrer que les IFs aglycones ne sont que 10 fois moins affins pour GPER que E_2. De plus, concernant les composés conjugués, leurs concentrations dans les fluides biologiques peuvent être de 1 000 à 100 000 fois supérieures à celle de l'E_2.

Analyse des interactions ligand-GPER et comparaison avec la littérature

L'analyse des poses 3D et de la poche de liaison du GPER a permis d'identifier les résidus clés potentiellement impliqués dans les interactions avec les IFs et leurs conjugués. Les résidus le plus souvent impliqués sont : His282, Ile279, Cys207, Trp216, Glu213, Phe208, Tyr142 et Gln138. Les conjugués glucuronidés semblent impliquer des résidus supplémentaires tels que : Thr305, Phe206, Cys205, Ala204, Asp202, Met141, Leu137, Ser134, Met133, Cys130, Leu119, Asn118, Asp111 et Gln53. En revanche, les conjugués sulfatés semblent se lier moins profondément dans le récepteur, en impliquant les résidus Arg286 et Gln215.

Ces résultats sont en accord avec les travaux de Mendez-Luna *et al.* [253], qui ont décrit une cavité importante dans le GPER, divisées en trois parties.

- Une première située à la face externe de la membrane : composée, notamment, des résidus Gly58, Leu59, Leu119 (domaine transmembranaire), Ala204, Cys205, Phe206, His282 et Arg286.
- Une partie médiane formée par les résidus Ser62, Glu115, Leu137, Phe208, Gln215 et Ile279.
- Une partie orientée vers la face intérieure de la membrane : composée de Asp111, Leu137, Met141, Gln138, Tyr142, Glu218 et Val219.

L'analyse des interactions a révélé que Phe208 et, dans une moindre mesure, Phe206, sont des résidus essentiels pour les interactions ligands-GPER. Ces deux résidus ont également été identifiés par Mendez-Luna *et al.* [253], qui ont proposé que la proximité de Phe208 avec Cys207 joue un rôle crucial dans le mécanisme d'activation/inactivation du récepteur GPER. Ce mécanisme impliquerait la liaison d'un ligand agoniste contenant des cycles aromatiques dans sa structure au site de liaison du GPER. Phe208 interagirait avec le ligand via des interactions π-π, affectant directement la liaison disulfure formée par Cys207-Cys130. Cette interaction induirait des remaniements conformationnels qui influencent directement les hélices α 3 et 5, initiant le processus d'activation du récepteur [253]. Ce modèle est en accord avec le modèle de « commutateur à bascule global » observé dans d'autres récepteurs de la famille des récepteurs de la rhodopsine.

Analyse des scores de docking

Le tableau 16 présente les scores de *docking* des IFs et de leurs conjugués pour les 3 fonctions de score préalablement identifiées comme les plus pertinentes. Les résultats indiquent que la GEN 4'-O-glucuronide (G4G) et la DAI-7-O-glucuronide (D7G) sont les seules molécules communes dans le top 5 des 3 fonctions de score. Ces deux molécules présentent donc une plus grande probabilité d'agir comme des agonistes du GPER et ont été sélectionnées pour les premiers tests *in vitro*. Il est à noter que l'Equol 7-O-glucuronide se retrouve également parmi les molécules les mieux notées selon

ces 3 fonctions de scores. Cependant, la présence d'Equol étant dépendante de la nature du microbiote intestinal, cette molécule n'a pas été retenue pour les premiers tests.

L'analyse globale du tableau révèle une tendance des molécules les plus lourdes à obtenir de meilleurs scores de *docking*. Cette observation met en lumière l'une des limites de cette approche *in silico*, où la taille et la masse moléculaire d'un composé peuvent influencer artificiellement le nombre d'interactions possibles avec le récepteur et, par conséquent, augmenter la probabilité d'obtenir un bon score d'amarrage. Ce phénomène est particulièrement notable pour les protéines avec de grandes cavités, comme le GPER

xscore_hps		dc_PAIRSURF			vina1
Gly 7-O-glucuronide	6,32	GEN 4'-O-glucuronide	267892	GEN 4'-S-7-glucuronide	8,72
GEN 4'-O-glucuronide	6,274	Gly 4-O-glucuronide	265881	DAI 7-O-glucuronide	8,6
Equol 7-O-glucuronide	6,24	Equol 7-O-glucuronide	265645	GEN 7-glucuronide	8,52
GEN 4'-S-7-glucuronide	6,236	DAI 7-O-glucuronide	256957	GEN 4'-O-glucuronide	8,5
DAI 7-O-glucuronide	6,202	GEN 7-glucuronide	251642	Equol 7-O-glucuronide	8,48
GEN 7-glucuronide	6,192	GEN 4'-S-7-glucuronide	247681	Gly 7-O-glucuronide	8,4
Gly 4-O-glucuronide	6,158	Gly 7-O-glucuronide	237429	GEN 7-sulfate	8,2
GEN 7-sulfate	5,98	4 OH-methyl-équol	186342	Gly 4-O-glucuronide	8,2
DAI 7-sulfate	5,972	Equol	183979	GEN	7,9
GEN 4'-sulfate	5,936	GEN	179441	DAI 7-sulfate	7,86
DAI 4'-sulfate	5,908	GEN 7-sulfate	177859	Eq 4'-sulfate	7,7
Equol 4'-sulfate	5,89	GEN 4'-sulfate	176958	DAI 4'-sulfate	7,7
GEN	5,876	Equol 4'-sulfate	176391	Equol	7,6
Equol	5,874	GEN 4',7-disulfate	173211	DAI	7,54
DAI	5,828	DAI 7-sulfate	172974	GEN 4',7-disulfate	7,52
4 OH-methyl-équol	5,782	Gly	172888	GEN 4'-sulfate	7,4
GEN 4',7-disulfate	5,76	DAI	169817	Gly	7,4
Gly	5,758	DAI 4'-sulfate	165233	4 OH-methyl-Eq	7,28
5-OH-équol	5,714	5-OH-équol	161059	5-OH-équol	7

Tableau 16. Classement des IFs et conjugués selon les scores obtenus avec les fonctions de score les plus pertinentes

Une analyse plus fine du *docking* de G4G a révélé deux emplacements de liaison distincts au sein du GPER Le premier, impliquant le résidu Gln53, se caractérise par une position plus superficielle par rapport au second, qui implique la Met141 (Figue 20).

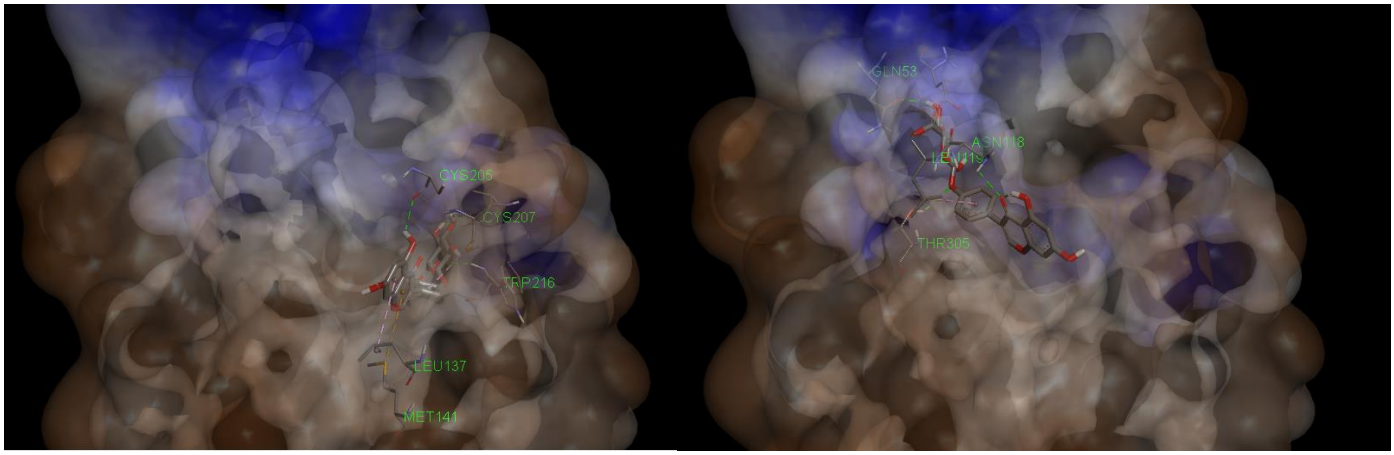


Figure 20. Poses de docking du G4G avec les résidus impliqués dans l'interaction avec le GPER. La surface de la protéine est représentée en terme d'hydrophobicité : partie extracellulaire en bleu et partie intramembrannaire en marron

L'examen détaillé des interactions entre G4G et les résidus du GPER a permis d'identifier plusieurs types d'interactions (Figure 21).

- Interactions π-π : elles impliquent des interactions entre les nuages d'électrons délocalisés des cycles aromatiques de G4G et les résidus aromatiques du GPER.
- Interactions π-sulfure : elles impliquent des interactions entre les nuages d'électrons délocalisés des cycles aromatiques de G4G et les atomes de soufre des résidus cystéine du GPER.
- Interactions conventionnelles : elles incluent des liaisons hydrogène, des liaisons ioniques et des interactions de van der Waals entre les atomes de G4G et les résidus du GPER.

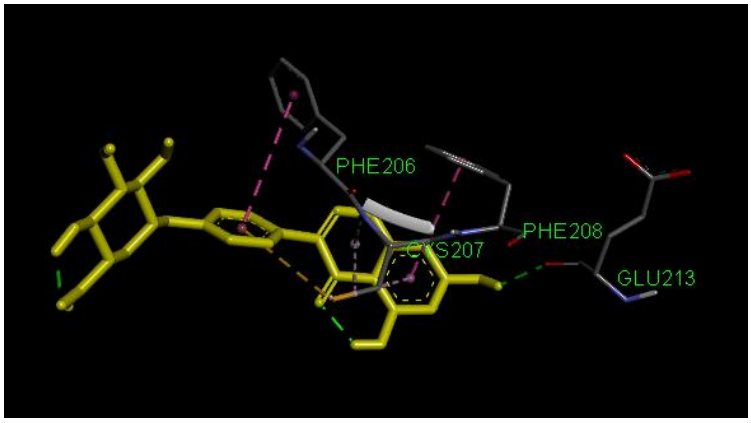


Figure 21. Interactions G4G/GPER :
- *π-π (magenta),*
- *π-sulfure (orange)*
- *conventionnelle (vert).*

Le G4G est représenté en jaune

IV.3.2. Etude de la prolifération

IV.3.2.1. Validation du modèle cellulaire

La validation des modèles cellulaires est une étape cruciale avant toute expérimentation *in vivo*. Dans le cadre de cette étude, le choix des modèles de TNBC s'est basé sur une analyse approfondie de la littérature [168,189,194,200]. Les cellules MDA-MB-436 et MDA-MB-468 ont été sélectionnées en raison de leur absence d'expression des récepteurs ER, PR, HER2 tout en exprimant le GPER.

La présence du GPER dans ces modèles cellulaires a été évaluée par Western blot (Figure 22). Cependant, les résultats obtenus ont été surprenants et en contradiction avec les données de la littérature. En effet, aucune bande correspondant à la taille attendue du GPER (55 kDa) n'a été détectée dans les cellules MDA-MB-436. En revanche, le GPER a été clairement identifié dans les cellules SKBR3 et MCF7, et dans une moindre mesure dans les cellules MDA-MB-468.

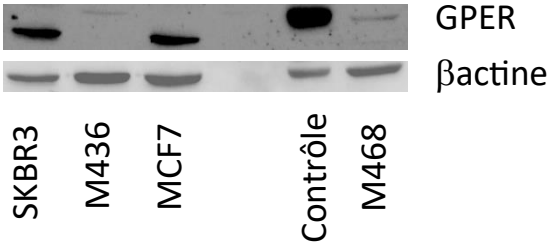


Figure 22. Résultat du western Blot pour le GPER. Anti-GPER 1:600. M436 : MDA-MB-436 et M468 : MDA-MB-468

Les analyses par Western blot ont révélé une divergence dans les tailles des bandes du GPER observées dans les cellules SKBR3 (45 kDa), MCF7 (43,5 kDa) et MDA-MB-468 (49 kDa). Cette variabilité apparente de la taille moléculaire du GPER n'est pas inhabituelle et peut s'expliquer par plusieurs facteurs. En effet, des études antérieures, telles que celles de Chan *et al.* [254] et Jala *et al.* [255],.ont mis en évidence une variabilité de la taille moléculaire du GPER pouvant atteindre 5 kDa. Cette variabilité est attribuée à des modifications post-traductionnelles du GPER, telles que la glycosylation et l'interaction avec d'autres protéines. Ces modifications peuvent augmenter la masse moléculaire apparente du GPER et expliquer les différences de taille observées dans les trois lignées cellulaires étudiées.

L'immunocytochimie présenté dans la Figure 23 a confirmé la présence du GPER dans les cellules SKBR3, MCF7 et MDA-MB-468, bien que son expression soit moins importante dans ces dernières. Il n'a pas été observé dans les MDA-MB-436. De manière surprenante, la majorité du GPER observé dans les trois lignées cellulaires serait localisée dans le cytoplasme, avec une faible proportion seulement à la membrane plasmique. Cette distribution intracellulaire non superficielle du GPER a été observée par une lecture de plusieurs plans intra-cellulaires, mais n'a pas été confirmée par un marquage de colocalisation membranaire.

Cette possible localisation cytoplasmique prédominante du GPER dans les trois lignées cellulaires étudiées, est un résultat préliminaire qui pourrait s'expliquer par la détection de la protéine en cours de synthèse dans le réticulum endoplasmique. Ceci amène à questionner la fiabilité des anticorps utilisés pour reconnaitre la protéine d'intérêt sous sa forme active. En effet, sans information sur la séquence immunogène pour les anticorps, on ne peut pas écarter la possibilité qu'ils reconnaissent des sous-parties du GPER en cours de synthèse.

Si la nature des anticorps n'est pas remise en question, les résultats obtenus pourraient remettre en question la présence constitutive du GPER dans les cellules MDA-MB-436, un modèle cellulaire couramment utilisé pour étudier le TNBC. Dans nos études, cette lignée cellulaire ne devrait donc pas répondre à l'estradiol. Elle pourrait également suggérer un rôle alternatif de ce récepteur dans des processus intracellulaires au-delà de la signalisation membranaire classique. Ces observations soulignent l'importance d'une caractérisation précise des anticorps utilisés pour détecter les récepteurs transmembranaires à structure tertiaire complexe. De plus, ces résultats démontrent la nécessité d'une validation rigoureuse des modèles cellulaires

avant toute étude et d'une interprétation prudente des données, en tenant compte de la variabilité des protéines et de leur localisation subcellulaire.

Concernant les xéno-estrogènes la question reste ouverte notamment à la présence de récepteurs ERRs par exemple.

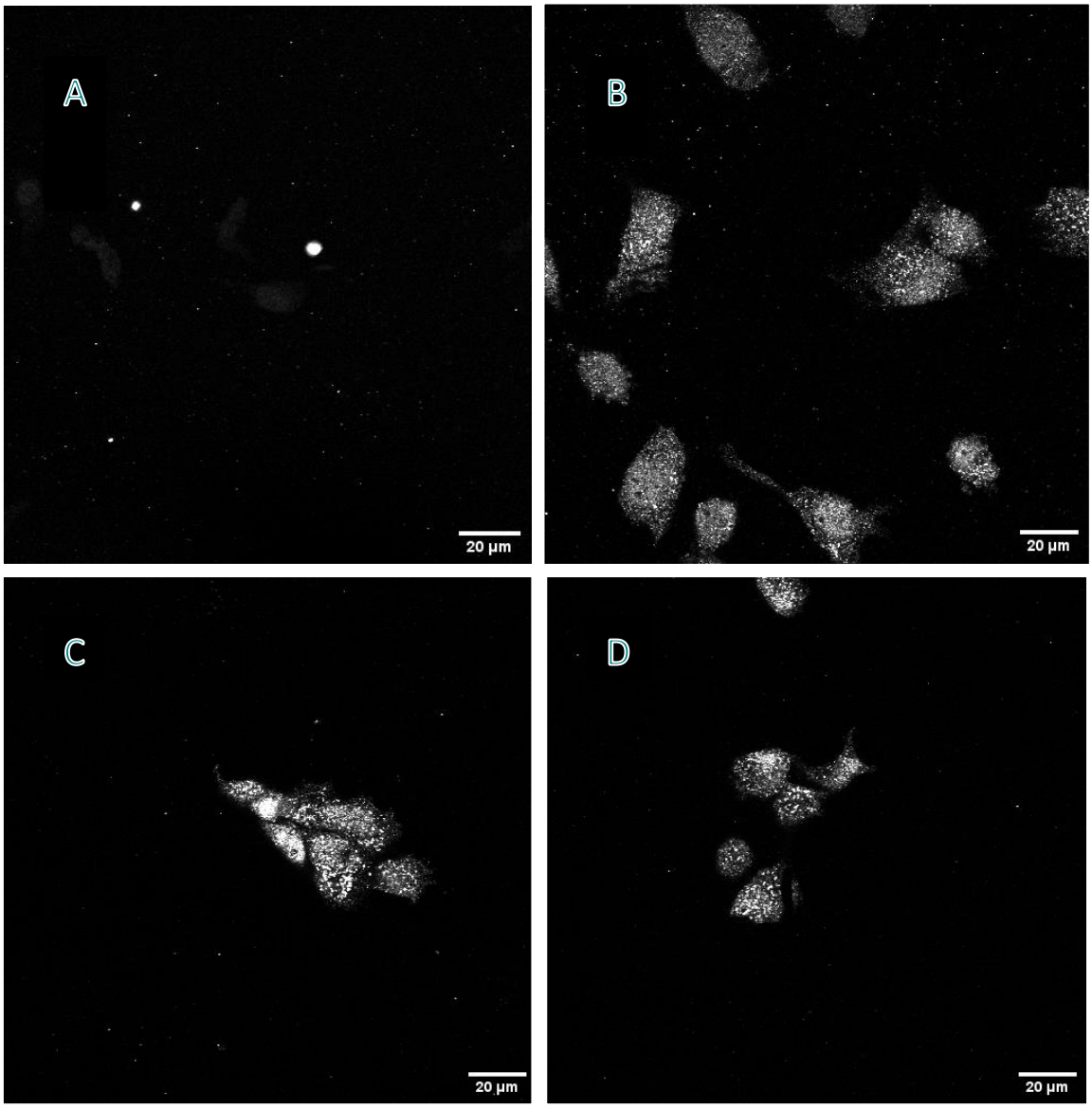


Figure 23.résultats de la détection du GPER par immunocytochimie (anti-GPER 1:200) dans (A) MDA-MB-436, (B) MDA-MB-468, (C) SKBR3 et (D) MCF7

L'expression d'ER a été vérifiée par Western blot dans les différentes lignées cellulaires utilisées. Comme illustré par la Figure 24, ERα n'a été détecté que dans les cellules MCF7. Aucune expression n'a été observée dans les modèles TNBC (MDA-MB-468, MDA-MB-436 et MDA-MB231), ni dans les cellules SKBR3. Ces résultats corroborent les données de la littérature, qui indiquent une

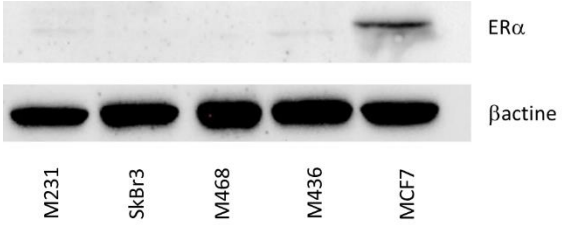


Figure 24. Résultats du western blot pour ER. Anti-ER 1:200. M231 : MDA-MB-231, M436 : MDA-MB-436 et M468 : MDA-MB-468

expression préférentielle d'ERα dans les cellules MCF7 par rapport aux autres modèles cellulaires étudiés.

IV.3.2.2. Effet des isoflavones et implication du GPER

La Figure 25 présente des résultats préliminaires de l'étude de l'effet des conjugués glucuronides sur la prolifération des cellules MDA-MB-468, qui n'expriment pas les ERs. Ces résultats indiquent une augmentation significative de la prolifération cellulaire après traitement avec la D7G à des concentrations de 10 et 100 nM. Ces doses sont couramment mesurées dans les plasmas de consommateur·ices de soja. En revanche, le traitement avec le conjugué G4G n'a induit qu'une légère augmentation de la prolifération qui n'a pas atteint la signification statistique. Les résultats sont exprimés en pourcentage de prolifération par rapport au milieu de culture, considéré comme contrôle.

L'implication du récepteur GPER dans ces effets prolifératifs a été évaluée en bloquant le récepteur avec son antagoniste spécifique G15 à une concentration de 1 µM. Le blocage du GPER a réduit significativement la prolifération induite par l'E_2 et les IFs conjuguées testées, suggérant que le GPER joue un rôle clé dans la médiation de ces effets dans les conditions expérimentales étudiées.

Il est important de noter que ces résultats préliminaires nécessitent des confirmations par des tests complémentaires afin d'évaluer la reproductibilité des données et la validité des conclusions. Des analyses statistiques plus approfondies et des expériences répétées avec des concentrations et des durées de traitement différentes seront nécessaires pour tirer des conclusions définitives sur les effets des IFs et de leurs conjugués glucuronidés et sulfatés sur la prolifération cellulaire et l'implication du GPER dans ces effets. Il serait également intéressant de tester les autres composés conjugués, notamment sulfatés.

La faible activité proliférative de G7G observée dans cette étude est en accord avec les travaux de Yuan *et al.* [86]. Ces auteur·es ont émis l'hypothèse que les effets prolifératifs de la G7G sur les cellules MCF-7, bien que modestes, pourraient résulter de son hydrolyse en GEN au sein de ces cellules. De plus, Islam *et al.* [44] ont démontré dans leurs travaux que les glucuronides des IFs, notamment D7G, n'ont pas d'activité estrogénique intrinsèque via les ERs, et que le faible pourcentage de déconjugaison observé dans les cellules (entre 0,2 et 1,6 %) suffit à expliquer la faible bioactivité observée pour les glucuronides d'IFs. Mais qu'en est-il du GPER ?

Si ce résultat ne peut être directement transposé au modèle cellulaire TNBC utilisé dans la présente étude, il vient néanmoins appuyer partiellement nos observations. En effet, la faible prolifération induite par G7G suggère que sa capacité à activer le récepteur GPER pourrait être limitée, ce qui est en accord avec les résultats de *docking* et d'analyses d'interactions moléculaires présentés précédemment.

Les formes conjuguées d'IFs représentent la majorité des IFs circulantes dans le sang des consommateur·ices de soja, atteignant généralement des concentrations de l'ordre de 1 µM. Nos résultats préliminaires montrent que ces conjugués d'IFs, ne sont pas dépourvus d'activité biologique, même à des doses aussi faibles que 10 nM. Cette donnée remet en question l'idée communément admise au sein de la communauté scientifique, selon laquelle les conjugués d'IFs seraient inactifs. Ces premiers résultats invitent à une approche plus prudente quant à l'impact des IFs sur les cancers du sein.

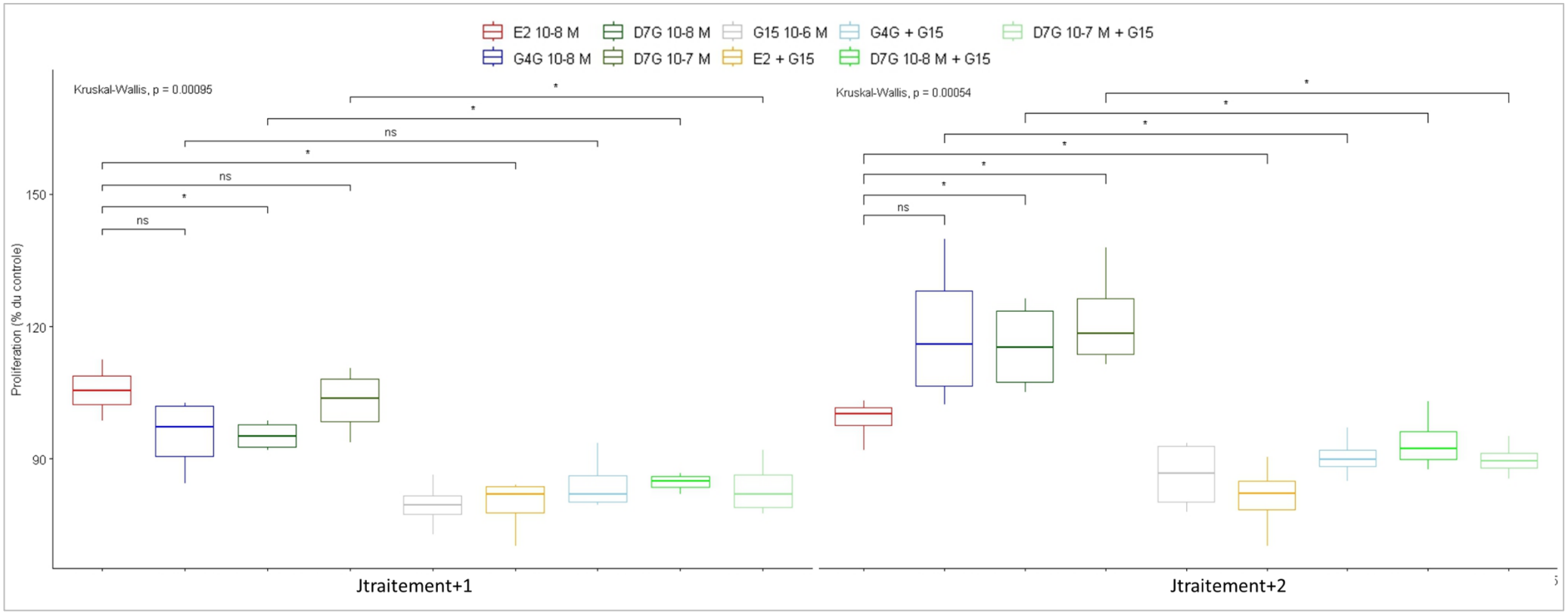


Figure 25. Résultats de prolifération des cellules MDA-MB-468, 24 et 48 heures après traitement à la GEN-4-Glucuronide et DAI-7-Glucuronide.

V. Discussion générale

Ce travail de recherche, initié dès le master, a adopté une approche multidisciplinaire en combinant des études cliniques, des analyses cellulaires, de la modélisation moléculaire et des investigations préindustrielles. L'objectif était d'explorer l'impact des IFs, et plus largement des phytoE, sur la santé humaine sous différents angles.

Comment peut-on effectuer une évaluation précise de l'exposition aux phytoE ?

Des études antérieures [8] estimaient l'exposition aux IFs en France comme relativement faible, autour de 50 µg/jour, bien en deçà des chiffres rapportés dans d'autres pays occidentaux (2 872 µg/jour aux USA [256]; 2 462 µg/jour en Angleterre [114]). Cependant, les résultats de cette recherche révèlent une réalité bien différente. En 2020, en France, l'exposition aux IFs s'avère nettement supérieure aux estimations antérieures, atteignant près de 2 000 µg/jour pour les consommateur·ices involontaires de soja et près de 7 000 µg/jour pour les consommateur·ices intentionnel·les. Chez ces dernier·es, la consommation déclarée varie de moins d'un produit par mois à plus de 9 par semaine. Cette exposition, caractérisée par des composés à la fois estrogéniques, anti-LH, anti-FSH et antithyroïdiens, dépasse largement celle des pesticides PEs. Cette découverte met en lumière l'importance de considérer les IFs comme une source significative d'exposition aux PEs dans la population française et invite à une réévaluation des risques potentiels associés à ces composés.

Les évaluations antérieures de l'exposition aux IFs reposaient souvent uniquement sur la consommation de soja autodéclarée. Cependant, cette approche négligeait la contribution significative du « soja caché » présent dans de nombreux aliments transformés [8]. En s'appuyant sur les travaux antérieurs de Lee *et al.* [117], cette recherche souligne l'importance de prendre en compte ces sources souvent ignorées d'IFs dans l'alimentation en France. La présence d'IFs dans divers produits alimentaires, y compris ceux auxquels les consommateur·ices ne s'attendent pas, conduit à une sous-estimation importante de l'exposition réelle à ces composés. Des études récentes rejoignent les résultats de cette thèse, notamment en montrant que les IFs de soja ont été retrouvées, en 2020, dans 100% des échantillons d'urine d'adolescent·es autrichien·nes [257].

Les erreurs commises dans l'évaluation de l'exposition aux IFs peuvent avoir des conséquences sur les résultats des études cliniques qui comparent différents niveaux d'exposition, qu'elles soient rétrospectives ou prospectives. Elles affectent la répartition des groupes en quartiles ou quintiles et affaiblissent la robustesse des études. Nous pouvons citer comme exemple l'étude d'Iwasaki *et al.* [258] qui évaluait l'association entre les concentrations plasmatiques de GEN et DAI, l'apport en IFs estimé par questionnaire alimentaire et le risque de cancer du sein chez des femmes japonaises. Leur analyse a mis en lumière des divergences notables entre les résultats obtenus à partir des mesures plasmatiques d'IFs et de l'estimation de l'apport par questionnaire. En effet, une corrélation négative significative a été mise en évidence entre la concentration plasmatique de GEN et le risque de cancer du sein, suggérant un effet protecteur potentiel des IFs. En revanche, aucune association significative n'a été observée entre le risque de cancer du sein et l'apport en IFs estimé par questionnaire alimentaire. Ces résultats soulignent les limites de cette dernière approche. Cette étude n'est pas un cas isolé. De nombreux autres travaux évaluant les

associations entre les IFs et la santé n'ont pas confronté les estimations alimentaires avec des dosages biologiques d'IFs [5,6,63].

Concernant les PEs d'origine anthropique, l'équipe de la Pr. Bennetau-Pelissero a proposé une estimation récente en travaillant sur les données des agences sanitaires et des sites professionnels. L'étude, fondée sur les pratiques mises en œuvre à l'Anses a intégré les limites maximales de résidus (LMR) d'une soixantaine de fruits et légumes consommés en France et d'une soixantaine de pesticides PEs pris au sens large. L'étude a révélé une exposition aux PEs de type cocktail d'environ 560 µg/jour. Cependant, cette estimation, basée sur des données chiffrées, peut être considérée comme surévaluée car elle fixe les concentrations de tous les pesticides au niveau des LMR. De plus, elle a pris en compte des substances dont l'effet PE n'est pas complètement établi. A contrario, l'étude a été restreinte aux pesticides autorisés en France, sans inclure les produits autorisés ailleurs mais qui peuvent pourtant contaminer les aliments des Français·es.

Ces résultats invitent à une reconsidération des sources d'exposition aux PEs en France, en intégrant pleinement les IFs dans les évaluations de risque. Des études plus approfondies sont requises pour mieux comprendre les effets potentiels sur la santé d'une exposition élevée aux IFs, surtout dans le contexte d'une exposition simultanée à d'autres PEs [117]. La prise en compte du « soja caché » dans les outils d'évaluation de l'exposition est cruciale pour obtenir une estimation plus précise du risque au sein de la population française.

Cette considération a d'ailleurs conduit au développement et à la validation de nouveaux outils d'évaluation de l'exposition aux IFs. Parmi eux, des questionnaires alimentaires spécifiques de la consommation française ont été élaborés. Les questionnaires ont été bâtis en se basant sur une étude préalable qui a réévalué l'exposition des Français·es aux IFs en analysant des aliments et plus de 12 000 étiquettes de produits alimentaires disponibles en France. Ils ont essentiellement rendu compte de l'ingestion récente d'IFs et de lignanes. Les corrélations obtenues avec les concentrations de phytoE dans les fluides biologiques correspondants (0,83 et 0,6 respectivement pour les IFs et l'ENL urinaires) étaient nettement plus significatives que celles des travaux existants ([25] et références incluses). Cette amélioration des corrélations s'explique par la prise en compte, d'un côté, du « soja caché » dans les habitudes alimentaires des individus, et de l'autre, de la pharmacocinétique de ces composés.

En outre, dans le cadre d'ISOLED, l'introduction d'une méthode d'analyse capillaire inédite des IFs a permis de valider son utilisation pour évaluer l'exposition habituelle aux IFs à long terme. Associée aux questionnaires alimentaires, cette méthode offre un moyen plus précis de surmonter les limites des fluctuations à court terme des taux sanguins et urinaires d'IFs. Elle a permis de déterminer sur un petit échantillon de femmes Françaises que 80% d'entre-elles étaient suffisamment exposées pour que les IFs soient détectables dans leurs cheveux. Ces résultats renforcent l'idée d'une sous-estimation de l'exposition aux IFs de soja, en France. Ces niveaux d'exposition sont conformes à ce qui est observé depuis quelques années dans les pays européens. Ils corroborent également des données obtenues en Allemagne sur des enfants et des adolescents qui présentaient tous des IFs détectables dans leurs urines sur une période de 24 heures [259].

Par ailleurs, les données capillaires, plasmatiques et urinaires ont été collectées pour la première fois sur une femme grande consommatrice de soja. Bien que ce type de profil soit régulièrement signalé dans la présentation de cas pathologiques [20], les concentrations des fluides biologiques ne sont généralement pas communiquées. L'analyse de ce cas a permis d'objectiver la grande variabilité qui existe dans l'exposition « naturelle » aux IFs, c'est-à-dire sans intervention clinique particulière. Cette grande variabilité individuelle avait déjà été soulignée dans l'étude de Lee *et al.* en 2019 [117].

Malgré ces avancées, l'évaluation précise de l'exposition aux IFs reste un défi. La variabilité individuelle du métabolisme et de la composition du microbiote intestinal influence la biodisponibilité et la biotransformation des PEs, ce qui complique davantage l'estimation des niveaux d'exposition intrinsèques [31]. Pour améliorer la précision des outils d'évaluation de l'exposition, des recherches futures devront mieux comprendre et modéliser ces facteurs de variabilité individuelle. Néanmoins, nos résultats suggèrent fortement que l'exposition aux IFs en France est nettement plus élevée que les estimations officielles.

Quels sont les risques véritables associés aux phytoE ?

Les travaux de cette thèse ont porté sur l'évaluation de l'effet des phytoE sur le LED et le TNBC.

a. PhytoE et LED

Les variabilités interindividuelles rendent d'autant plus complexe l'évaluation de l'exposition pour les lignanes, autre famille de phytoE, qui ne sont estrogéniques qu'après métabolisation par la flore intestinale compétente [29]. Cette complexité métabolique rend l'évaluation de leur rôle dans le LED particulièrement difficile. En effet, le lien entre les critères physiopathologiques et les entérolignanes ne peut se faire que sur la base d'analyses plasmatiques. En effet, l'estimation de la consommation des précurseurs d'entérolignanes ne suffit pas à prédire l'exposition aux phytoE et d'en déduire leur effet sur la santé.

L'étude ISOLED a révélé une différence significative entre les concentrations plasmatiques d'ENL, un entéro-métabolite des lignanes, chez les femmes atteintes de LED ou de PTI et les volontaires saines. Les groupes de malades présentaient des concentrations plasmatiques d'ENL nettement plus faibles, suggérant un effet potentiellement protecteur de cet entéro-lignane contre le LED. La nature de SERM de l'ENL [28] ainsi que sa faible affinité pour les ER α et β [260], pourraient être impliquées dans cet effet, qui pourrait lui-même être qualifié d'anti-estrogénique.

Malgré l'intérêt de ces résultats, il est important de rester prudent dans leur interprétation. La taille limitée de l'étude ne permet pas de tirer des conclusions définitives sur l'association entre la consommation récente de lignanes et la réduction du risque de poussées lupiques. De plus, l'étude ne permet pas de déterminer si la relation observée entre l'ENL plasmatique et le statut sain est causale (effet de la consommation de lignanes) ou corrélative (simple marqueur d'un microbiote intestinal plus compétent ou d'une plus grande consommation de fruits et légumes). Toutefois, ces résultats renforcent les données indiquant qu'une consommation abondante de fruits, légumes et fibres est bénéfique pour la santé.

Par ailleurs, le profil d'entérolignanes produit par le microbiote intestinal est influencé par la source et le profil des lignanes consommés [27,261]. Cette observation suggère un rôle potentiel du microbiote intestinal dans l'association observée entre l'ENL plasmatique et le LED. Des études futures devraient se concentrer sur l'analyse détaillée de la composition et de la fonction du microbiote intestinal chez les patientes atteintes de LED et les volontaires saines, en corrélation avec leurs niveaux d'ENL et leur consommation de lignanes. Cette approche permettrait de mieux comprendre les interactions complexes entre l'alimentation, le microbiote intestinal et le développement du LED.

De la même manière, l'effet des IFs sur les MAIs pourrait dépendre de la composition de la flore intestinale, notamment dans le cas de l'équol. Cependant, il est important de noter que, contrairement aux lignanes, la DAI et la GEN, principales IFs du soja possèdent une activité estrogénique directe, indépendamment de la flore intestinale. L'absence de distinction claire entre les groupes de l'étude ISOLED en fonction de leur consommation d'IFs souligne la nécessité d'études plus puissantes pour explorer l'association potentielle entre l'exposition aux IFs et le LED. Les données obtenues chez les animaux transgéniques modèles du LED semblent indiquer un effet plutôt aggravant à des doses alimentaires [137]. Si tel s'avérait être le cas chez les humains, cela démontrerait une nouvelle fois que tous les phytoE ne sont pas équivalents et que, si les IFs doivent être considérées avec circonspection, les lignanes issues des fruits et légumes seraient plutôt bénéfiques pour la santé. C'est notamment ce qu'a montré l'étude de cohorte de Touillaud *et al.* sur les cancers du sein [262].

b. IFs et TNBC

Le *docking* moléculaire, en raison des approximations inhérentes aux calculs, ne peut être considéré que comme un outil de présélection de ligands potentiels, nécessitant donc une validation biologique ultérieure. Les résultats des calculs suggèrent une faible affinité des IFs et de leurs conjugués pour le GPER. Cependant, l'absence de détermination des Kd des différents composés dans le cadre de cette étude ne permet pas de tirer des conclusions définitives quant à leur affinité pour ce récepteur. Par ailleurs, l'analyse des poses des ligands a permis d'identifier des résidus clés potentiellement impliqués dans les interactions ligands récepteur, suggérant un mécanisme d'activation du GPER similaire à celui décrit dans la littérature [175]. Ces résultats devraient donc être confirmés par des études *in vitro* plus étoffées que celles réalisées dans le cadre de cette thèse.

A ce titre, les travaux menés sur les cellules TNBC fournissent des résultats préliminaires encourageants concernant l'implication du GPER dans les effets prolifératifs des IFs. Dans un modèle de cellules cancéreuses triple négatives exprimant ce récepteur, deux molécules, correspondant aux formes circulantes principales de conjugués d'IFs, ont été testées à des concentrations physiologiques. Les résultats ont révélé une promotion modeste mais significative de la prolifération cellulaire deux jours après le traitement, réduite en présence d'un antagoniste spécifique du GPER. Ces données soulèvent des questions importantes sur l'effet des autres conjugués d'IFs et la validité des résultats obtenus pour d'autres modèles de TNBC. Des tests complémentaires sont nécessaires pour confirmer ces observations et explorer l'effet d'une gamme plus large de conjugués d'IFs. Pour souligner l'intérêt de ces résultats il faut mentionner qu'actuellement, une controverse existe dans la communauté scientifique autour de l'effet potentiel des IFs sur la prolifération des

cancers du sein *in vivo*. Un des arguments avancés par les chercheur·es qui considèrent que les IFs sont sans effet, est lié aux concentrations et formes circulantes qui seraient insuffisantes et inactives. Néanmoins, l'éventuelle activité des composés conjugués, à des doses circulantes, incite à la prudence, notamment chez les femmes atteintes de cancers hormono-dépendants exprimant des GPER.

Pour autant, les résultats obtenus *in vitro* et *in silico* ne permettent pas aujourd'hui de déterminer si les IFs conjuguées agissent en tant que telles ou si elles sont préalablement déconjuguées par les cellules avant d'agir via le GPER. Si elles sont capables d'agir directement, ce phénomène concernerait toutes les cellules exprimant le GPER. En revanche, si elles doivent être dé-conjuguées pour être actives, leurs actions resteront limitées aux cellules capables de réaliser cette déconjugaison et donc exprimant les enzymes glucuronidases et/ou sulfatases. Au regard de ces informations, et dans le cas où les IFs-glucuronides agiraient seulement après déconjugaison, cela pourrait contribuer à une certaine spécificité tissulaire de l'action des IFs. Compte tenu des variations rapportées dans la littérature de l'action des IFs sur la santé, une telle spécificité tissulaire doit être envisagée.

L'évaluation des activités de conjugaison dans différentes cellules cancéreuses mammaires permet d'illustrer la diversité des profils de métabolisation des PEs. Parmi ces cellules, les cellules MDA-MB-436 (TNBC), présentent un profil de conjugaison de la GEN similaire aux cellules MCF7 [263]. En effet, ces deux lignées cellulaires expriment des activités glucuronosyl- et sulfotransférases, enzymes de phase II impliquées dans la conjugaison des PEs avec des glucuronides ou des sulfates. Cependant, ces cellules présentaient une préférence marquée pour la sulfatation par rapport à la glucuronidation [264] avec quelques différences significatives entre les deux lignées qui pourraient être liées à l'expression d'ERα [263]. Ces observations soulignent l'importance de tester l'effet des conjugués sulfatés, qui pourrait être potentiellement plus prononcés que celui des glucuronides, sur ces modèles cellulaires. Elles confirment également l'importance de prendre en compte la diversité des profils de conjugaison et de déconjuguaison, notamment des PEs, dans les cellules cancéreuses lors de l'évaluation de leurs effets potentiels.

Par ailleurs, les études cellulaires portant sur l'expression et la localisation du GPER ont donné des résultats inattendus. Le GPER n'a en effet pas été détecté dans les cellules MDA-MB-436, un modèle TNBC couramment utilisé dans la littérature [156,180,181]. Par ailleurs, la possible prédominance d'une localisation cytoplasmique du GPER dans les autres lignées cellulaires étudiées soulève des questions quant à la nature des anticorps utilisés et à leur capacité à reconnaître un récepteur transmembranaire enchâssé dans la membrane. Cette observation pourrait suggérer également un rôle potentiel du GPER au-delà de ses fonctions classiques de signalisation membranaire. Il est important de rappeler que la présence d'une protéine dans le cytoplasme d'une cellule est normale, car c'est le lieu de sa synthèse. Cependant, cela ne garantit pas qu'elle y soit sous sa forme active. En effet, les anticorps, en particulier les monoclonaux, peuvent être développés contre de courtes séquences d'acides aminés qui peuvent être plus ou moins masquées lorsque le récepteur est sous sa forme active dans la membrane. Ce point souligne l'importance de développer des outils de détection des protéines transmembranaires spécifiques de leurs formes physiologiques actives. Les informations disponibles sur les anticorps utilisés en immunohistochimie sur les cellules en culture ne permettent pas d'exclure

cette hypothèse. Ces résultats soulignent également la nécessité de valider rigoureusement les modèles cellulaires et de prendre en compte les complexités de l'expression et de la localisation des protéines lors de l'interprétation des données expérimentales.

En outre, la grande majorité des études *in vitro* explorant le rôle du GPER soulève des questions de pertinence physiologique sur son rôle de médiateur des signaux intracellulaires déclenchés par l'E_2. Ces études utilisent souvent des cellules en culture exposées à des concentrations supra-physiologiques d'E_2 [179]. En effet, les concentrations physiologiques d'E_2 circulant dans le plasma sanguin sont de l'ordre de 7,3 à 734,2 pM chez de jeunes femelles rongeurs, et chez les femmes, ces taux varient entre 0,1 et 2,2 nM selon la phase du cycle œstral [265]. Or les concentrations testées, et actives, sont de l'ordre de 10-100 nM. Ces concentrations correspondent aux doses circulantes ne se rencontrant que pendant la grossesse [266]. Ce décalage important entre les concentrations d'E_2 utilisées *in vitro* et les niveaux physiologiques classiques soulève des questions sur le rôle physiologique du GPER dans la médiation des effets de l'E_2 *in vivo* [179]. Alternativement, elles pourraient guider une réflexion sur l'existence de récepteurs potentiellement dédiés à médier des effets estrogéniques seulement au pic du cycle ou pendant la gestation des femelles. Cela impliquerait également que le GPER n'aurait pas de rôle chez les mâles. Par ailleurs, il est important de noter que la fraction d'E_2 libre et biologiquement active ne représente qu'une infime partie de l'E_2 circulante [265], la majorité étant liée aux protéines plasmatiques. Ceci dénote de l'importance de recherches complémentaires utilisant des modèles plus proches de la physiologie humaine pour mieux comprendre l'importance réelle du GPER dans la médiation des effets de l'E_2 *in vivo*.

Quelles recommandations émettre ?

a. A l'échelle des politiques publiques

Sur le registre de la santé publique en France, en 2005, l'AFSSA et l'AFSSAPS ont rendu un rapport commun « Sécurité et bénéfices des phyto-estrogènes apportés par l'alimentation – Recommandations » [79]. Dans cette étude, menée pendant 30 mois par 20 expert·es scientifiques, les agences sanitaires ont recommandé de limiter la consommation de soja chez les nourrissons, les femmes enceintes et celles ayant des antécédents personnels ou familiaux de cancers du sein. Ceci inclue aussi le « soja caché » et est d'autant plus préoccupant que les produits industriels à base de soja contiennent souvent de nombreux autres additifs qui les placent dans la catégories des produits ultra-transformés. Ces additifs, tels que les édulcorants ou agglomérants, pourraient être impliqués dans l'apparition de certains cancers, selon des études épidémiologiques récentes menées au sein de la population française [273].

Une question importante se pose alors : est-il obligatoire de mentionner les taux d'IFs dans les aliments ? En 2005, les expert·es avaient proposé une dose acceptable d'IFs de 1 mg/kg/jour en se basant sur les estimations de consommation en Asie. Ils et elles avaient aussi encouragé les industriels à indiquer les teneurs en phytoE de leurs produits sur les étiquettes [8]. Malheureusement, l'avis des agences sanitaires n'a pas force de loi et cette recommandation n'a pas fait l'objet d'une législation formelle et n'est donc pas appliquée. De plus, bien que l'AFSSA et l'AFSSAPS aient proposé une méthode de calcul des quantités d'IFs, elles n'ont pas validé de technique d'analyse

spécifique pour ces substances. C'est aussi une des raisons pour lesquelles les industriels ne mentionnent toujours pas les teneurs en phytoE de leurs produits.

En Allemagne, l'Institut fédéral d'évaluation des risques (BfR) considère que les informations scientifiques disponibles sur les IFs soulèvent des interrogations quant à leur innocuité. Des perturbations hormonales de la reproduction ont été observées chez des enfants et les effets des IFs sur la thyroïde sont de plus en plus évidents. L'agence estime qu'en Allemagne la consommation de phytoE est sous-estimée, en particulier depuis l'essor récent du végétarisme. Elle recommande une réévaluation de cette exposition afin de formuler des recommandations appropriées [267,268].

En Norvège, le Comité scientifique norvégien pour la sécurité alimentaire (VKM) a évalué la sécurité des IFs consommées sous forme de compléments alimentaires à des doses de 40 ou 80 mg/jour prises pendant une période d'un à trois mois. Les résultats de l'évaluation du VKM suggèrent que ces suppléments d'IFs pourraient présenter un risque d'effets négatifs sur le taux d'hormones chez les femmes pré-ménopausées, les hommes et les adolescent·es. Chez les adolescentes, des perturbations du cycle menstruel sont également observées [269]. Le comité conclu que ces effets négatifs semblaient limités au système hormonal et n'ont entraîné aucune autre conséquence notable pour la santé des populations ciblées par l'étude, sur la durée de celle-ci.

En 2015, l'Autorité européenne de sécurité des aliments (EFSA) a émis un avis sur les effets des IFs présentes dans les compléments alimentaires destinés aux femmes ménopausées. Selon l'agence européenne, il n'y aurait pas de risque même pour des fortes doses allant jusqu'à 120 mg/jour [72]. Cependant, il faut souligner ici que ni les études de toxicologie américaines, ni les effets sur la thyroïde ou les cas cliniques problématiques n'ont été pris en compte.

L'EFSA [72] a examiné les études cliniques portant sur les IFs pour déterminer si elles augmentaient les risques de cancer du sein ou de l'utérus. Cependant, les femmes incluses dans ces études étaient en bonne santé, et le risque de cancer n'était pas l'objectif principal de ces études. Il n'était donc pas systématiquement enregistré. Par conséquent, les résultats sont partiels mais l'agence a conclu que les IFs étaient sans danger. Il est important de noter que dans la vie réelle, les femmes peuvent présenter des pathologies, y compris des cancers qu'elles ignorent. En outre, le risque de cancers du sein est plus élevé en péri-ménopause. Or l'EFSA n'a pas pu formuler d'avis sur cette population en raison du manque de données disponibles.

Dans d'autres pays européens, les agences sanitaires n'ont pas récemment émis d'avis sur les IFs de soja ni sur une dose acceptable. Les avis existants datent d'avant les études de toxicologie américaines. Dans certains pays, les effets estrogéniques bénéfiques des IFs sont admis chez les femmes, mais aucune agence ne s'interroge sur les effets de ces estrogènes sur des populations qui ne devraient pas y être exposées.

Au Canada, Santé Canada, a émis des réserves concernant la consommation de soja et d'IFs. L'agence mentionne que le soja est contre-indiqué chez les hypothyroïdiens et les nourrissons [270]. La Commission Canadienne des Graines publie chaque année un état des lieux de la qualité du soja récolté, incluant la teneur moyenne en IFs. La technique d'analyse retenue par Santé Canada est bien documentée. Tandis qu'aux États-Unis, l'agence de sécurité sanitaire des aliments (FDA) réévalue actuellement ses allégations santé liées au soja [271].

Concernant les phytoE, l'institut étasunien de recherche en santé, le NIH, a financé récemment des études cliniques et des recherches sur des populations humaines pour évaluer les effets délétères des IFs du soja sur la reproduction. Les résultats de ces études sont en accord avec les données de toxicologie obtenues sur les rats, et indiquent une dose délétère d'IFs chez les femmes pré-ménopausées de l'ordre de 50 mg/jour. Cependant, à ce jour, les formules de lait infantile à base de soja ne sont pas interdites aux USA bien que les derniers avis publiés en 2011 [107] ont été discutés au sein des panels d'experts et fait l'objet de critiques de la part de scientifiques étasunien·nes après leur publication. Actuellement, les dernières données scientifiques obtenues par le NIH pourraient conduire à une réévaluation des risques associés à l'utilisation de ces formules chez les nourrissons [272]. Mais pour l'instant, la date précise de cette réévaluation reste inconnue.

b. A l'échelle individuelle et de l'engagement industriel pour adapter les méthodes de préparation

Compte tenu des risques potentiels pour la santé associés à une exposition élevée aux IFs et des préoccupations exprimées par l'AFSSA et l'AFSSAPS en 2005 [8], cette thèse, menée en partenariat avec des industriels de l'agroalimentaire, explore des méthodes visant à réduire leur teneur dans les aliments à base de soja. Cette observation illustre l'implication de certaines entreprises engagées, en particulier Biopress/Berkem, qui soutient financièrement cette thèse CIFRE, témoignant ainsi de leur préoccupation pour la santé des consommateur·ices. A cet effet, les projets PHYTOSECUR 1 et 2 ont démontré l'efficacité des méthodes proposées de transformation du soja pour réduire la teneur en IFs des aliments. A l'échelle préindustrielle, les rinçages des graines de soja dépelliculées ont permis de réduire de 50% les teneurs en IFs. Un autre bénéfice concerne la valorisation des eaux de lavages, à travers la récupération des IFs et de l'eau de traitement, qui semble être une pratique envisageable, réduisant ainsi l'impact écologique et économique. De plus, la reproduction en laboratoire des méthodes traditionnelles de préparation du tempeh, du miso et du tofu a montré une réduction significative des IFs par rapport aux aliments industriels. Cette réduction explique pourquoi la consommation du soja n'a pas eu en Asie, dans le passé, l'impact attendu d'une légumineuse riche en phytoE. En effet, ces populations consommatrices de soja, représentant plus d'un milliard d'individus en 1998 (soit plus d'un tiers de la population mondiale [273]), ont continué, au moins en partie, à cuisiner le soja selon des recettes familiales traditionnelles.

Ces dernières observations soulèvent de nouvelles questions concernant l'estimation de l'exposition actuelle des Asiatiques. Si ces populations continuent à consommer du soja de manière traditionnelle, les corrélations observées en Asie entre la consommation de soja et la réduction des risques de cancers du sein sont-elles réellement liées aux IFs ? Ne devrions-nous pas plutôt envisager des biais liés à d'autres facteurs environnementaux et éventuellement alimentaires ?

Bien que les études PHYTOSECUR 1 et 2 aient apporté des solutions concrètes pour réduire l'exposition aux IFs, tant à l'échelle (pré)industrielle que domestique, des recherches et une collaboration plus poussée avec l'industrie agroalimentaire sont recommandées pour faciliter la mise en œuvre de ces méthodes à plus grande échelle. L'engagement de l'industrie alimentaire est crucial pour développer des produits à base de soja à faible teneur en IFs, tout en préservant leurs qualités organoleptiques et nutritionnelles. Cette collaboration pourrait également contribuer à la mise en place

de stratégies de communication efficaces pour informer les consommateur·ices sur les choix et les pratiques alimentaires éclairés concernant les IFs.

Finalement, les résultats de cette recherche pourraient avoir des implications importantes pour la mise en place de stratégies visant à limiter l'exposition aux IFs dans la population tant au niveau des politiques publiques que des pratiques individuelles. Il est essentiel d'établir des directives claires et fondées sur des preuves concernant la consommation d'IFs, en particulier pour les populations vulnérables. En ce qui concerne l'ensemble des consommateur·ices, le rapport de l'Anses prévu pour juillet 2024, est attendu.

VI. Conclusion générale et perspectives

Ce travail de recherche s'est articulé autour de trois axes principaux : (1) l'évaluation de l'exposition aux IFs en France, (2) l'exploration d'une corrélation potentielle entre l'exposition aux IFs et les poussées de LED ainsi que l'étude de l'implication du GPER dans les effets prolifératifs des IFs sur des modèles cellulaires de TNBC, (3) et enfin la réduction de l'exposition aux IFs via l'alimentation en contribuant au développement de procédures industrielles adéquates.

Les résultats de nos travaux ont révélé une exposition significative aux IFs en France, estimée à environ 7 mg/jour/habitant·e. Cette valeur est nettement supérieure aux estimations antérieures, principalement en raison de la prise en compte du « soja caché » présent dans de nombreux produits ultra-transformés et à l'augmentation récente de la tendance végétarienne dans la population française. Les évaluations antérieures de l'exposition aux IFs se basaient souvent uniquement sur la consommation de soja auto-déclarée. Cette recherche a souligné l'importance de prendre en compte ces sources souvent inaperçues d'IFs dans l'alimentation.

La validation des questionnaires alimentaires a démontré leur pertinence comme outils d'estimation de l'exposition. Notamment, l'intégration de la pharmacocinétique des IFs dans les scores alimentaires a amélioré la précision de l'estimation de l'exposition récente. Par ailleurs, le développement du dosage capillaire des IFs, représentant une avancée méthodologique intéressante, permet de refléter une exposition habituelle et de lisser les variations quotidiennes.

Bien que l'étude ISOLED soit contrainte par la taille limitée de son échantillon, elle laisse entrevoir la possibilité d'un effet protecteur non anticipé de l'ENL contre le LED. Cependant, en raison de ces limitations, il reste incertain de conclure de manière irréfutable quant à l'impact d'une exposition aux IFs.

L'exploration de l'implication du GPER dans les effets des IFs sur les TNBC a permis de mettre en lumière la complexité de la signalisation cellulaire et des interactions ligand-récepteur. Les résultats préliminaires indiquent une implication potentielle du GPER dans les effets prolifératifs des IFs conjuguées.

Les études PHYTOSECUR 1 et 2, en établissant l'efficacité des traitements à l'eau pour réduire la teneur en IFs des aliments à base de soja, représentent une reconnaissance de la valeur des savoirs traditionnels. Les résultats démontrent la possibilité de développer des méthodes de production industrielles et domestiques alternatives pour réduire l'exposition aux IFs. La mise en place de procédés de déflavonoïdisation à l'échelle industrielle et la promotion de méthodes de préparation

traditionnelles pourraient contribuer à limiter l'exposition aux IFs et à réduire les risques potentiels pour la santé.

De manière générale, cette thèse a mis en lumière les complexités de la recherche sur les IFs et a souligné la nécessité de poursuivre les investigations sur leur impact sur la santé humaine. Elle souligne les mérites d'une approche multidisciplinaire et l'importance de prendre en compte les IFs dans l'évaluation des PEs ainsi qu'une vigilance accrue pour les groupes à risque. En comblant quelques lacunes entre l'évaluation de l'exposition, la compréhension mécanistique et les observations cliniques, cette recherche ouvre la voie à l'élaboration de diverses stratégies, aussi bien à l'échelle industrielle que domestique, pour assurer une consommation sûre et durable des aliments à base de soja en réduisant l'exposition de la population.

Ainsi, les résultats obtenus ouvrent des perspectives importantes, à différents niveaux, pour la recherche future et la mise en place de mesures de santé publique visant à protéger la population des effets potentiellement néfastes des IFs.

Au niveau méthodologique

Il serait intéressant d'envisager des méthodes d'estimation de l'exposition des IFs à grande échelle, dans des contextes plus larges que le contexte français. De plus, et comme évoqué précédemment, la mise en place de procédés de déflavonoïdisation à l'échelle industrielle et la promotion de méthodes de préparation traditionnelles pourraient contribuer à limiter l'exposition aux IFs et à réduire les risques potentiels pour la santé, en particulier pour les populations vulnérables. Il parait donc essentiel de poursuivre le développement de méthodes de production d'aliments à base de soja ayant une teneur réduite en IFs.

Au niveau fondamental

Il est important de poursuivre les investigations sur l'impact des IFs sur la santé humaine, notamment sur les MAIs et les cancers hormono-dépendants, en tenant compte des variations individuelles, des facteurs de confusion et des effets cocktails. Des études multicentriques avec des effectifs plus importants sont nécessaires pour confirmer ou infirmer les résultats préliminaires de l'étude ISOLED. Ces études peuvent s'appuyer sur les outils d'estimation de l'exposition développé dans ce travail. Une meilleure compréhension des mécanismes d'action des IFs et leur interaction avec le GPER et d'autres récepteurs aux estrogènes nécessite également la mise en œuvre de nouvelles études. L'exploration de la localisation cellulaire du GPER, des voies de signalisation intracellulaires ainsi que le développement de nouveaux outils de détection plus performants pourraient y contribuer.

VII. Références bibliographiques

1. Gier, K.; Preininger, C.; Sauer, U. A Chip for Estrogen Receptor Action: Detection of Biomarkers Released by MCF-7 Cells through Estrogenic and Anti-Estrogenic Effects. *Sensors* **2017**, *17*, 1760, doi:10.3390/s17081760.

2. Bennetau-Pelissero, C. Positive or Negative Effects of Isoflavones: Toward the End of a Controversy: Response to the Letter from Dr Messina and Dr Badger Following the Publication of the Paper by Fernandez-Lopez A, Lamothe V, Delample M, Denayrolles M and Bennetau-Pelissero C. Entitled: Removing Isoflavones from Modern Soyfood: Why and How? *Food Chem* **2017**, *225*, 293–301, doi:10.1016/j.foodchem.2017.01.033.

3. Canivenc-Lavier, M.C.; Bennetau-Pelissero, C. Phytoestrogens and Health Effects. *Nutrients* **2023**, *15*, 317, doi:10.3390/nu15020317.

4. Unfer, V.; Casini, M.L.; Costabile, L.; Mignosa, M.; Gerli, S.; Di Renzo, G.C. Endometrial Effects of Long-Term Treatment with Phytoestrogens: A Randomized, Double-Blind, Placebo-Controlled Study. *Fertil Steril* **2004**, *82*, 145–148, doi:10.1016/j.fertnstert.2003.11.041.

5. Chavarro, J.E.; Toth, T.L.; Sadio, S.M.; Hauser, R. Soy Food and Isoflavone Intake in Relation to Semen Quality Parameters among Men from an Infertility Clinic. *Human Reproduction* **2008**, *23*, 2584–2590, doi:10.1093/humrep/den243.

6. Touillaud, M.; Gelot, A.; Mesrine, S.; Bennetau-Pelissero, C.; Clavel-Chapelon, F.; Arveux, P.; Bonnet, F.; Gunter, M.; Boutron-Ruault, M.C.; Fournier, A. Use of Dietary Supplements Containing Soy Isoflavones and Breast Cancer Risk among Women Aged >50 y: A Prospective Study. *American Journal of Clinical Nutrition* **2019**, *109*, 606–614, doi:10.1093/ajcn/nqy313.

7. Ariyani, W.; Miyazaki, W.; Amano, I.; Hanamura, K.; Shirao, T.; Koibuchi, N. Soy Isoflavones Accelerate Glial Cell Migration via GPER-Mediated Signal Transduction Pathway. *Front Endocrinol (Lausanne)* **2020**, *11*, doi:10.3389/fendo.2020.554941.

8. AFSSA; AFSSAPS *Sécurité et Bénéfices Des Phyto-Estrogènes Apportés Par l'alimentation-Recommandations*; **2005**;

9. Périan, S.; Vanacker, J.M. GPER as a Receptor for Endocrine-Disrupting Chemicals (EDCs). *Front Endocrinol (Lausanne)* **2020**, *11*, doi:10.3389/fendo.2020.00545.

10. Bergman, Å.; Heindel, J.J.; Jobling, S.; Kidd, K.A.; Zoeller, R.T. *Situation Du Savoir Scientifique Relatif Aux Perturbateurs Endocriniens*; **2012**;

11. Colborn, T. Commentary: Setting Aside Tradition When Dealing with Endocrine Disruptors. *ILAR J* **2004**, *45*.

12. Lagarde, F.; Beausoleil, C.; Belcher, S.M.; Belzunces, L.P.; Emond, C.; Guerbet, M.; Rousselle, C. Non-Monotonic Dose-Response Relationships and Endocrine Disruptors: A Qualitative Method of Assessment. *Environmental Health* **2015**, *14*, doi:10.1186/1476-069X-14-13.

13. National Toxicology Program Multigenerational Reproductive Study of Genistein (CAS No. 446-72-0) in Sprague-Dawley Rats (Feed Study). *Natl Toxicol Program Tech Rep Ser* **2008**, 1–266.

14. Delfosse, V.; Huet, T.; Harrus, D.; Granell, M.; Bourguet, M.; Gardia-Parège, C.; Chiavarina, B.; Grimaldi, M.; Le Mével, S.; Blanc, P.; et al. Mechanistic Insights into the Synergistic Activation of the RXR–PXR Heterodimer by Endocrine Disruptor Mixtures. *Proc Natl Acad Sci U S A* **2021**, *118*, doi:10.1073/PNAS.2020551118/VIDEO-1.

15. Yilmaz, B.; Terekeci, H.; Sandal, S.; Kelestimur, F. Endocrine Disrupting Chemicals: Exposure, Effects on Human Health, Mechanism of Action, Models for Testing and Strategies for Prevention. *Rev Endocr Metab Disord* **2020**, *21*, 127–147, doi:10.1007/s11154-019-09521-z.

16. Basak, S.; Das, M.K.; Duttaroy, A.K. Plastics Derived Endocrine-Disrupting Compounds and Their Effects on Early Development. *Birth Defects Res* **2020**, *112*, 1308–1325, doi:10.1002/bdr2.1741.

17. Molina, L.; Bustamante, F.A.; Bhoola, K.D.; Figueroa, C.D.; Ehrenfeld, P. Possible Role of Phytoestrogens in Breast Cancer via GPER-1/GPR30 Signaling. *Clin Sci* **2018**, *132*, 2583–2598, doi:10.1042/CS20180885.

18. Barnes, S. The Biochemistry, Chemistry and Physiology of the Isoflavones in Soybeans and Their Food Products. *Lymphat Res Biol* **2010**, *8*, 89–98, doi:10.1089/lrb.2009.0030.

19. Cassidy, A.; Bingham, S.; Setchell, K. Biological Effects of Isoflavones in Young Women: Importance of the Chemical Composition of Soyabean Products. *British Journal of Nutrition* **1995**, *74*, 587–601, doi:10.1079/bjn19950160.

20. Chandrareddy, A.; Muneyyirci-Delale, O.; McFarlane, S.I.; Murad, O.M. Adverse Effects of Phytoestrogens on Reproductive Health: A Report of Three Cases. *Complement Ther Clin Pract* **2008**, *14*, 132–135, doi:10.1016/j.ctcp.2008.01.002.

21. Mierziak, J.; Kostyn, K.; Kulma, A. Flavonoids as Important Molecules of Plant Interactions with the Environment. *Molecules* **2014**, *19*, 16240–16265.

22. Subramanian, S.; Stacey, G.; Yu, O. Distinct, Crucial Roles of Flavonoids during Legume Nodulation. *Trends Plant Sci* **2007**, doi:10.1016/j.tplants.2007.06.006.

23. D'Archivio, M.; Filesi, C.; Di Benedetto, R.; Gargiulo, R.; Giovannini, C.; Masella, R. Polyphenols, Dietary Sources and Bioavailability. *Ann Ist Super Sanità* **2007**, *43*, 348–361.

24. Konar, N.; Poyrazoğlu, E.S.; Demir, K.; Artik, N. Determination of Conjugated and Free Isoflavones in Some Legumes by LC-MS/MS. *Journal of Food Composition and Analysis* **2012**, *25*, 173–178, doi:10.1016/j.jfca.2011.11.004.

25. Bensaada, S.; Raymond, I.; Pellegrin, I.; Viallard, J.-F.; Bennetau-Pelissero, C. Validation of ELISAs for Isoflavones and Enterolactone for Phytoestrogen Intake Assessment in the French Population. *Nutrients* **2023**, *15*, 967, doi:10.3390/nu15040967.

26. Anses *Étude de l'alimentation Totale Française 2 (EAT 2)*; **2011**;

27. Senizza, A.; Rocchetti, G.; Mosele, J.I.; Patrone, V.; Callegari, M.L.; Morelli, L.; Lucini, L.; Jurgó, A.; Fotschki, B.; Sawicki, T. Lignans and Gut Microbiota: An Interplay Revealing Potential Health Implications. *Molecules* **2020**, *25*, 5709, doi:10.3390/molecules25235709.

28. Carreau, C.; Flouriot, G.; Bennetau-Pelissero, C.; Potier, M. Enterodiol and Enterolactone, Two Major Diet-Derived Polyphenol Metabolites Have Different Impact on ERα Transcriptional Activation in Human Breast Cancer Cells. *Journal of Steroid Biochemistry and Molecular Biology* **2008**, *110*, 176–185, doi:10.1016/j.jsbmb.2008.03.032.

29. Kuijsten, A.; Arts, I.C.W.; Vree, T.B.; Hollman, P.C.H. Human Nutrition and Metabolism Pharmacokinetics of Enterolignans in Healthy Men and Women Consuming a Single Dose of Secoisolariciresinol Diglucoside. *J Nutr* **2005**, *135*, 795–801, doi:10.1093/jn/135.4.795.

30. Messina, M.J. Legumes and Soybeans: Overview of Their Nutritional Profiles and Health Effects. *Am J Clin Nutr* **1999**, *70*, 439S–450S.

31. Mathey, J.; Lamothe, V.; Coxam, V.; Potier, M.; Sauvant, P.; -Pelissero, C.B. Concentrations of Isoflavones in Plasma and Urine of Post-Menopausal Women Chronically Ingesting High Quantities of Soy Isoflavones. *J Pharm Biomed Anal* **2006**, *41*, 957–965, doi:10.1016/j.jpba.2006.01.051.

32. Cassidy, A.; Bingham, S.; Setchell, K. Biological Effects of Isoflavones in Young Women: Importance of the Chemical Composition of Soyabean Products. *British Journal of Nutrition* **1995**, *74*, 587–601, doi:10.1079/bjn19950160.

33. Wu, A.H.; Stanczyk, F.Z.; Hendrich, S.; Murphy, P.A.; Zhang, C.; Wan, P.; Pike, M.C. Effects of Soy Foods on Ovarian Function in Premenopausal Women. *Br J Cancer* **2000**, *82*, 1879–1886, doi:10.1054/bjoc.1999.1218.

34. Setchell, K.D.R.; Brown, N.M.; Desai, P.; Zimmer-Nechemias, L.; Wolfe, B.E.; Brashear, W.T.; Kirschner, A.S.; Cassidy, A.; Heubi, J.E. *Bioavailability of Nutrients and Other Bioactive Components from Dietary Supplements Bioavailability of Pure Isoflavones in Healthy Humans and Analysis of Commercial Soy Isoflavone Supplements 1,2*; **2001**; Vol. 131;.

35. Zubik, L.; Meydani, M. Bioavailability of Soybean Isoflavones from Aglycone and Glucoside Forms in American Women. *Am J Clin Nutr* **2003**, *77*, 1459–1465.

36. Setchell, K.D.; Brown, N.M.; Zimmer-Nechemias, L.; Brashear, W.T.; Wolfe, B.E.; Kirschner, A.S.; Heubi, J.E. Evidence for Lack of Absorption of Soy Isoflavone Glycosides in Humans, Supporting the Crucial Role of Intestinal Metabolism for Bioavailability. *Am J Clin Nutr* **2002**, *76*, 447–453.

37. Manach, C.; Scalbert, A.; Morand, C.; Rémésy, C.; Jiménez, L. Polyphenols: Food Sources and Bioavailability. *Am J Clin Nutr* **2004**, *79*, 727–774.

38. Rowland, I.; Faughnan, M.; Hoey, L.; Wähälä, K.; Williamson, G.; Cassidy, A. Bioavailability of Phyto-Oestrogens. *British Journal of Nutrition* **2003**, *89*, S45–S58, doi:10.1079/BJN2002796.

39. Setchell, K.D.R.; Brown, N.M.; Desai, P.; Zimmer-Nechemias, L.; Wolfe, B.E.; Brashear, W.T.; Kirschner, A.S.; Cassidy, A.; Heubi, J.E. Bioavailability of Nutrients and Other Bioactive Components from Dietary Supplements Bioavailability of Pure Isoflavones in Healthy Humans and Analysis of Commercial Soy Isoflavone Supplements. *J. Nutr* **2001**, *131*, 1362–1375, doi:10.1093/jn/131.4.1362S.

40. Ariyani, W.; Miyazaki, W.; Koibuchi, N. A Novel Mechanism of S-Equol Action in Neurons and Astrocytes: The Possible Involvement of Gpr30/Gper1. *Int J Mol Sci* **2019**, *20*, doi:10.3390/ijms20205178.

41. Soukup, S.T.; Al-Maharik, N.; Botting, N.; Kulling, S.E. Quantification of Soy Isoflavones and Their Conjugative Metabolites in Plasma and Urine: An Automated and Validated UHPLC-MS/MS Method for Use in Large-Scale Studies. *Anal Bioanal Chem* **2014**, *406*, 6007–6020, doi:10.1007/s00216-014-8034-y.

42. Hall, K.A.; Filardo, E.J. The G Protein-Coupled Estrogen Receptor (GPER): A Critical Therapeutic Target for Cancer. *Cells* **2023**, *12*, doi:10.3390/CELLS12202460.

43. Shinkaruk, S.; Durand, M.; Lamothe, V.; Carpaye, A.; Martinet, A.; Chantre, P.; Vergne, S.; Nogues, X.; Moore, N.; Bennetau-Pelissero, C. Bioavailability of Glycitein Relatively to Other Soy Isoflavones in Healthy Young Caucasian Men. *Food Chem* **2012**, *135*, 1104–1111, doi:10.1016/j.foodchem.2012.03.135.

44. Islam, M.A.; Bekele, R.; Vanden Berg, J.H.J.; Kuswanti, Y.; Thapa, O.; Soltani, S.; Van Leeuwen, F.X.R.; Rietjens, I.M.C.M.; Murk, A.J. Deconjugation of Soy Isoflavone Glucuronides Needed for Estrogenic Activity. *Toxicol In Vitro* **2015**, *29*, 706–7015, doi:10.1016/j.tiv.2015.01.013.

45. Vergne, S.; Titier, K.; Bernard, V.; Asselineau, J.; Durand, M.; Lamothe, V.; Potier, M.; Perez, P.; Demotes-Mainard, J.; Chantre, P.; et al. Bioavailability and Urinary Excretion of Isoflavones in Humans: Effects of Soy-Based Supplements Formulation and Equol Production. *J Pharm Biomed Anal* **2007**, *43*, 1488–1494, doi:10.1016/j.jpba.2006.10.006.

46. Setchell, K.D. Phytoestrogens: The Biochemistry, Physiology, and Implications for Human Health of Soy Isoflavones. *Am J Clin Nutr* **1998**, *68*, 1333S-1346S.

47. Bennetts, H.W.; Uuderwood, E.J.; Shier, F.L. A Specific Breeding Problem of Sheep on Subterranean Clover Pastures in Western Australia. *Aust Vet J* **1946**, *22*, 2–12, doi:10.1111/j.1751-0813.1946.tb15473.x.

48. National Toxicology Program NTP Technical Report on the Toxicology and Carcinogenesis Study of Genistein (CAS No. 446-72-0) in Sprague-Dawley Rats (Feed Study). *Natl Toxicol Program Tech Rep Ser* **2008**, 1–242.

49. Bensaada, S.; Chabrier, F.; Ginisty, P.; Ferrand, C.; Peruzzi, G.; Valat, M.; Bennetau-Pelissero, C. Improved Food-Processing Techniques to Reduce Isoflavones in Soy-Based Foodstuffs. *Foods 2023, Vol. 12, Page 1540* **2023**, *12*, 1540, doi:10.3390/FOODS12071540.

50. Kodell, R.L.; Gaylor, D.W. Combining Uncertainty Factors in Deriving Human Exposure Levels of Noncarcinogenic Toxicants. *Ann N Y Acad Sci* **1999**, *895*, 188–195, doi:10.1111/j.1749-6632.1999.tb08085.x.

51. Laddha, A.P.; Murugesan, S.; Kulkarni, Y.A. In-Vivo and in-Silico Toxicity Studies of Daidzein: An Isoflavone from Soy. *Drug Chem Toxicol* **2022**, *45*, 1408–1416, doi:10.1080/01480545.2020.1833906.

52. Chandrareddy, A.; Muneyyirci-Delale, O.; McFarlane, S.I.; Murad, O.M. Adverse Effects of Phytoestrogens on Reproductive Health: A Report of Three Cases. *Complement Ther Clin Pract* **2008**, *14*, 132–135, doi:10.1016/j.ctcp.2008.01.002.

53. Chang, E.C.; Charn, T.H.; Park, S.H.; Helferich, W.G.; Komm, B.; Katzenellenbogen, J.A.; Katzenellenbogen, B.S. Estrogen Receptors α and β as Determinants of Gene Expression: Influence of Ligand, Dose, and Chromatin Binding. *Molecular Endocrinology* **2008**, *22*, 1032–1043, doi:10.1210/me.2007-0356.

54. Russo, M.; Russo, G.L.; Daglia, M.; Kasi, P.D.; Ravi, S.; Nabavi, S.F.; Nabavi, S.M. Understanding Genistein in Cancer: The “Good” and the “Bad” Effects: A Review. *Food Chem* **2016**, *196*, 589–600.

55. Hüser, S.; Guth, S.; Joost, H.G.; Soukup, S.T.; Köhrle, J.; Kreienbrock, L.; Diel, P.; Lachenmeier, D.W.; Eisenbrand, G.; Vollmer, G.; et al. Effects of Isoflavones on Breast Tissue and the Thyroid Hormone System in Humans: A Comprehensive Safety Evaluation. *Arch Toxicol* **2018**, *92*, 2703–2748.

56. Obiorah, I.E.; Fan, P.; Jordan, V.C. Breast Cancer Cell Apoptosis with Phytoestrogens Is Dependent on an Estrogen-Deprived State. *Cancer Prevention Research* **2014**, *7*, 939–949, doi:10.1158/1940-6207.CAPR-14-0061.

57. Chen, J.; Duan, Y.; Zhang, X.; Ye, Y.; Ge, B.; Chen, J. Genistein Induces Apoptosis by the Inactivation of the IGF-1R/p-Akt Signaling Pathway in MCF-7 Human Breast Cancer Cells. *Food Funct* **2015**, *6*, 995, doi:10.1039/c4fo01141d.

58. Hatono, M.; Ikeda, H.; Suzuki, Y.; Kajiwara, Y.; Kawada, K.; Tsukioki, T.; Kochi, M.; Suzawa, K.; Iwamoto, T.; Yamamoto, H.; et al. Effect of Isoflavones on Breast Cancer Cell Development and Their Impact on Breast Cancer Treatments. *Breast Cancer Res Treat* **2021**, *185*, 307–316, doi:10.1007/s10549-020-05957-z.

59. Shike, M.; Doane, A.S.; Russo, L.; Cabal, R.; Reis-Filho, J.S.; Gerald, W.; Cody, H.; Khanin, R.; Bromberg, J.; Norton, L. The Effects of Soy Supplementation on Gene Expression in Breast Cancer: A Randomized Placebo-Controlled Study. *J Natl Cancer Inst* **2014**, doi:10.1093/jnci/dju189.

60. Chi, F.; Wu, R.; Zeng, Y.C.; Xing, R.; Liu, Y.; Xu, Z.G. Post-Diagnosis Soy Food Intake and Breast Cancer Survival: A Meta-Analysis of Cohort Studies. *Asian Pacific Journal of Cancer Prevention* **2013**, *14*, 2407–2412, doi:10.7314/APJCP.2013.14.4.2407.

61. Liu, X.O.; Huang, Y.B.; Gao, Y.; Chen, C.; Yan, Y.; Dai, H.J.; Song, F.J.; Wang, Y.G.; Wang, P.S.; Chen, K.X. Association between Dietary Factors and Breast Cancer Risk among Chinese Females: Systematic Review and Meta-Analysis. *Asian Pacific Journal of Cancer Prevention* **2014**, *15*, 1291–1298, doi:10.7314/APJCP.2014.15.3.1291.

62. Ward, H.A.; Kuhnle, G.G.C.; Mulligan, A.A.; Lentjes, M.A.H.; Luben, R.N.; Khaw, K.T. Breast, Colorectal, and Prostate Cancer Risk in the European Prospective Investigation into Cancer and Nutrition-Norfolk in Relation to Phytoestrogen Intake Derived from an Improved Database. *American Journal of Clinical Nutrition* **2010**, *91*, 440–448, doi:10.3945/ajcn.2009.28282.

63. Cassidy, A.; Bingham, S.; Setchell, K. Biological Effects of Isoflavones in Young Women: Importance of the Chemical Composition of Soyabean Products. *British Journal of Nutrition* **1995**, *74*, 587–601, doi:10.1079/BJN19950160.

64. Wu, A.H.; Stanczyk, F.Z.; Hendrich, S.; Murphy, P.A.; Zhang, C.; Wan, P.; Pike, M.C. Effects of Soy Foods on Ovarian Function in Premenopausal Women. *Br J Cancer* **2000**, *82*, 1879–1886, doi:10.1054/bjoc.1999.1218.

65. Cassidy, A.; Bingham, S.; Dr Setchell, K. Biological Effects of a Diet of Soy Protein Rich in Isoflavones on the Menstrual Cycle of Premenopausal Women. *Am J Cliii Nutr* **1994**, *60*, 333–373.

66. Upson, K.; Harmon, Q.E.; Baird, D.D. Soy-Based Infant Formula Feeding and Ultrasound-Detected Uterine Fibroids among Young African-American Women with No Prior Clinical Diagnosis of Fibroids. *Environ Health Perspect* **2016**, *124*, 769–775, doi:10.1289/ehp.1510082.

67. Strom, B.L.; Schinnar, R.; Ziegler, E.E.; Barnhart, K.T.; Sammel, M.D.; Macones, G.A.; Stallings, V.A.; Drulis, J.M.; Nelson, S.E.; Hanson, S.A.; et al. Exposure to Soy-Based Formula in Infancy and Endocrinological and Reproductive Outcomes in Young Adulthood. *J Am Med Assoc* **2001**, *286*, 807–814, doi:10.1001/jama.286.7.807.

68. Otun, J.; Sahebkar, A.; Östlundh, L.; Atkin, S.L.; Sathyapalan, T. Systematic Review and Meta-Analysis on the Effect of Soy on Thyroid Function. *Sci Rep* **2019**, *9*, doi:10.1038/s41598-019-40647-x.

69. Persky, V.W.; Turyk, M.E.; Wang, L.; Freels, S.; Chatterton, R.J.; Barnes, S.; Erdman, J.J.; Sepkovic, D.W.; Bradlow, H.L.; Potter, S. Effect of Soy Protein on Endogenous Hormones in Postmenopausal Women. *Am J Clin Nutr* **2002**, *75*, 145–153.

70. Bell, D.S.H.; Ovalle, F.M. Use of Soy Protein Supplement And Resultant Need for Increased Dose of Levothyroxine. *Endocrine Practice* **2001**, *7*, 193.

71. Li, J.; Teng, X.; Wang, W.; Chen, Y.; Yu, X.; Wang, S.; Li, J.; Zhu, L.; Li, C.; Fan, C.; et al. Effects of Dietary Soy Intake on Maternal Thyroid Functions and Serum Anti-Thyroperoxidase Antibody Level During Early Pregnancy. *J Med Food* **2011**, *14*, 543–550.

72. EFSA Risk Assessment for Peri- and Post-Menopausal Women Taking Food Supplements Containing Isolated Isoflavones. *EFSA Journal* **2015**, *13*, doi:10.2903/j.efsa.2015.4246.

73. Bennetau-Pelissero, C. Positive or Negative Effects of Isoflavones: Toward the End of a Controversy: Response to the Letter from Dr Messina and Dr Badger Following the Publication of the Paper by Fernandez-Lopez A, Lamothe V, Delample M, Denayrolles M and Bennetau-Pelissero C. *Food Chem* **2017**, *225*, 293–301, doi:10.1016/j.foodchem.2017.01.033.

74. Doerge, D.R.; Sheehan, D.M. Goitrogenic and Estrogenic Activity of Soy Isoflavones. *Environ Health Perspect* **2002**, *110*, 349–353.

75. Fernandez-Lopez, A.; Lamothe, V.; Delample, M.; Denayrolles, M.; Bennetau-Pelissero, C. Removing Isoflavones from Modern Soyfood: Why and How? *Food Chem* **2016**, *210*, 286–294, doi:10.1016/j.foodchem.2016.04.126.

76. Uifălean, A.; Schneider, S.; Gierok, P.; Ionescu, C.; Iuga, C.A.; Lalk, M. The Impact of Soy Isoflavones on MCF-7 and MDA-MB-231 Breast Cancer Cells Using a Global Metabolomic Approach. *Int J Mol Sci* **2016**, *17*, doi:10.3390/ijms17091443.

77. Dolinoy, D.C.; Weidman, J.R.; Waterland, R.A.; Jirtle, R.L. Maternal Genistein Alters Coat Color and Protects Avy Mouse Offspring from Obesity by Modifying the Fetal Epigenome. *Environ Health Perspect* **2006**, *114*, 567–572, doi:10.1289/ehp.8700.

78. Harlid, S.; Adgent, M.; Jefferson, W.N.; Panduri, V.; Umbach, D.M.; Xu, Z.; Stallings, V.A.; Williams, C.J.; Rogan, W.J.; Taylor, J.A. Soy Formula and Epigenetic Modifications: Analysis of Vaginal Epithelial Cells from Infant Girls in the IFED Study. *Environ Health Perspect* **2017**, *125*, 447–452, doi:10.1289/EHP428.

79. Lamartiniere, C.A.; Moore, J.B.; Brown, N.M.; Thompson, R.; Hardin, M.J.; Barnes, S. Genistein Suppresses Mammary Cancer in Rats. *Carcinogenesis* **1995**, *16*, 2833–2840.

80. Carreau, C.; Flouriot, G.; Bennetau-Pelissero, C.; Potier, M. Respective Contribution Exerted by AF-1 and AF-2 Transactivation Functions in Estrogen Receptor α Induced Transcriptional Activity by Isoflavones and Equol: Consequence on Breast Cancer Cell Proliferation. *Mol Nutr Food Res* **2009**, *53*, 652–658, doi:10.1002/mnfr.200800061.

81. Seo, H.-S.; Denardo, D.G.; Jacquot, Y.; Laıös, I.; Vidal, D.S.; Rojas Zambrana, C.; Leclercq, G.; Brown, P.H. Stimulatory Effect of Genistein and Apigenin on the Growth of Breast Cancer Cells

Correlates with Their Ability to Activate ER Alpha. *Breast Cancer Res Treat* **2006**, *99*, 121–134, doi:10.1007/s10549-006-9191-2.

82. McMichael-Phillips, D.F.; Harding, C.; Morton, M.; Roberts, S.A.; Howell, A.; Potten, C.S.; Bundred, N.J. Effects of Soy-Protein Supplementation on Epithelial Proliferation Inthe Histologically Normal Human Breast. *Am J Clin Nutr* **1998**, *68*, 1431S-6S.

83. Gann, P.H.; Geiger, A.S.; Helenowski, I.B.; Vonesh, E.F.; Chatterton, R.T. Estrogen and Progesterone Levels in Nipple Aspirate Fluid of Healthy Premenopausal Women: Relationship to Steroid Precursors and Response Proteins. *Cancer Epidemiology Biomarkers and Prevention* **2006**, *15*, 39–44, doi:10.1158/1055-9965.EPI-05-0470.

84. Hargreaves, D.F.; Potten, C.S.; Harding, C.; Shaw, L.E.; Morton, M.S.; Roberts, S.A.; Howell, A.; Bundred, N.J. Two-Week Dietary Soy Supplementation Has an Estrogenic Effect on Normal Premenopausal Breast. *J Clin Endocrinol Metab* **1999**, *84*, 4017–4024.

85. Shike, M.; Doane, A.S.; Russo, L.; Cabal, R.; Reis-Filho, J.S.; Gerald, W.; Cody, H.; Khanin, R.; Bromberg, J.; Norton, L. The Effects of Soy Supplementation on Gene Expression in Breast Cancer: A Randomized Placebo-Controlled Study. *J Natl Cancer Inst* **2014**, *106*, doi:10.1093/jnci/dju189.

86. Yuan, B.; Wang, L.; Jin, Y.; Zhen, H.; Xu, P.; Xu, Y.; Li, C.; Xu, H. Role of Metabolism in the Effects of Genistein and Its Phase II Conjugates on the Growth of Human Breast Cell Lines. *AAPS Journal* **2012**, *14*, 329–344, doi:10.1208/s12248-012-9338-5.

87. Canivenc-Lavier, M.-C.; Bennetau-Pelissero, C. Phytoestrogens and Health Effects. *Nutrients* **2023**, *15*, doi:https://doi.org/10.3390/nu15020317.

88. Liu, J.; Yuan, F.; Gao, J.; Shan, B.; Ren, Y.; Wang, H.; Gao, Y. Oral Isoflavone Supplementation on Endometrial Thickness: A Meta-Analysis of Randomized Placebo-Controlled Trials. *Oncotarget* **2016**, *7*, 17369–17379.

89. Rossi, M.; Edefonti, V.; Parpinel, M.; Lagiou, P.; Franchi, M.; Ferraroni, M.; Decarli, A.; Zucchetto, A.; Serraino, D.; Dal Maso, L.; et al. Proanthocyanidins and Other Flavonoids in Relation to Endometrial Cancer Risk: A Case–Control Study in Italy. *British Journal of Cancer 2013 109:7* **2013**, *109*, 1914–1920, doi:10.1038/bjc.2013.447.

90. Zhong, X. shan; Ge, J.; Chen, S. wei; Xiong, Y. quan; Ma, S. juan; Chen, Q. Association between Dietary Isoflavones in Soy and Legumes and Endometrial Cancer: A Systematic Review and Meta-Analysis. *J Acad Nutr Diet* **2018**, *118*, 637–651, doi:10.1016/J.JAND.2016.09.036.

91. Gómez-Zorita, S.; González-Arceo, M.; Fernández-Quintela, A.; Eseberri, I.; Trepiana, J.; Portillo, M.P. Scientific Evidence Supporting the Beneficial Effects of Isoflavones on Human Health. *Nutrients* **2020**, *12*, doi:10.3390/nu12123853.

92. Pejčic, T.; Zekovic, M.; Bumbaširevic, U.; Kalaba, M.; Vovk, I.; Bensa, M.; Popovic, L.; Tešic, Ž.T. The Role of Isoflavones in the Prevention of Breast Cancer and Prostate Cancer. *Antioxidants* **2023**, *12*, 386, doi:10.3390/antiox12020368.

93. Zhang, H.Y.; Cui, J.; Zhang, Y.; Wang, Z.L.; Chong, T.; Wang, Z.M. Isoflavones and Prostate Cancer: A Review of Some Critical Issues. *Chin Med J (Engl)* **2016**, *129*, 341, doi:10.4103/0366-6999.174488.

94. Mahmoud, A.M.; Yang, W.; Bosland, M.C. Soy Isoflavones and Prostate Cancer: A Review of Molecular Mechanisms. *J Steroid Biochem Mol Biol* **2014**, *140*, 116, doi:10.1016/J.JSBMB.2013.12.010.

95. Hofsjö, A.; Bohm-Starke, N.; Bergmark, K.; Masironi, B.; Sahlin, L. Sex Steroid Hormone Receptor Expression in the Vaginal Wall in Cervical Cancer Survivors after Radiotherapy. *Acta Oncol* **2019**, *58*, 1107–1115, doi:10.1080/0284186X.2019.1598574.

96. De Lima Toccafondo Vieira, M.; Duarte, R.F.; Campos, L.M.M.; Nunan, E. de A. Comparison of the Estrogenic Potencies of Standardized Soy Extracts by Immature Rat Uterotrophic Bioassay. *Phytomedicine* **2008**, *15*, 31–37, doi:10.1016/j.phymed.2007.06.006.

97. Qin, H.; Lin, Z.; Vásquez, E.; Luan, X.; Guo, F.; Xu, L. High Soy Isoflavone or Soy-Based Food Intake during Infancy and in Adulthood Is Associated with an Increased Risk of Uterine Fibroids in Premenopausal Women: A Meta-Analysis ScienceDirect. *Nutrition Research* **2019**, *71*, 30–42, doi:10.1016/j.nutres.2019.06.002.

98. Gao, M.; Wang, H. Frequent Milk and Soybean Consumption Are High Risks for Uterine Leiomyoma: A Prospective Cohort Study. *Medicine* **2018**, *97*, doi:10.1097/MD.0000000000012009.

99. Jacobsen, B.K.; Jaceldo-Siegl, K.; Knutsen, S.F.; Fan, J.; Oda, K.; Fraser, G.E. Soy Isoflavone Intake and the Likelihood of Ever Becoming a Mother: The Adventist Health Study-2. *Int J Womens Health* **2014**, *6*, 377, doi:10.2147/IJWH.S57137.

100. Unfer, V.; Casini, M.L.; Gerli, S.; Costabile, L.; Mignosa, M.; Carlo, G.; Renzo, D. Phytoestrogens May Improve the Pregnancy Rate in in Vitro Fertilization-Embryo Transfer Cycles: A Prospective, Controlled, Randomized Trial. *Fertil Steril* **2004**, *82*, doi:10.1016/j.fertnstert.2004.

101. Gonzalez-Martin, R.; Palomar, A.; Quiñonero, A.; Pellicer, N.; Zuckerman, C.; Whitehead, C.; Scott, R.T.; Dominguez, F. Phytoestrogens Present in Follicular Fluid and Urine Are Positively Associated with IVF Outcomes Following Single Euploid Embryo Transfer. *Int J Mol Sci* **2023**, *24*, doi:10.3390/ijms241310852.

102. Chavarro, J.E.; Mínguez-Alarcón, L.; Chiu, Y.-H.; Gaskins, A.J.; Souter, I.; Williams, P.L.; Calafat, A.M.; Hauser, R. Soy Intake Modifies the Relation Between UrinaryBisphenol A Concentrations and Pregnancy OutcomesAmong Women Undergoing Assisted Reproduction. *J Clin Endocrinol Metab* **2016**, *101*, 1082–1090, doi:10.1210/jc.2015-3473.

103. Beaton, L.K.; McVeigh, B.L.; Dillingham, B.L.; Lampe, J.W.; Duncan, A.M. Soy Protein Isolates of Varying Isoflavone Content Do Not Adversely Affect Semen Quality in Healthy Young Men. *Fertil Steril* **2010**, *94*, 1717–1722, doi:10.1016/j.fertnstert.2009.08.055.

104. Chavarro, J.E.; Toth, T.L.; Sadio, S.M.; Hauser, R. Soy Food and Isoflavone Intake in Relation to Semen Quality Parameters among Men from an Infertility Clinic. *Human Reproduction* **2008**, *23*, 2584–2590, doi:10.1093/humrep/den243.

105. Mumford, S.L.; Kim, S.; Chen, Z.; Barr, D.B.; Louis, G.M.B. Urinary Phytoestrogens Are Associated with Subtle Indicators of Semen Quality among Male Partners of Couples Desiring Pregnancy. *Journal of Nutrition* **2015**, *145*, 2535–2541, doi:10.3945/jn.115.214973.

106. Watanabe, S.; Terashima, K.; Sato, Y.; Arai, S.; Eboshida, A. Effects of Isoflavone Supplement on Healthy Women. *BioFactors* **2000**, *12*, 233–241, doi:10.1002/biof.5520120136.

107. McCarver, G.; Bhatia, J.; Chambers, C.; Clarke, R.; Etzel, R.; Foster, W.; Hoyer, P.; Leeder, J.S.; Peters, J.M.; Rissman, E.; et al. NTP-CERHR Expert Panel Report on the Developmental Toxicity of Soy Infant Formula. *Birth Defects Res B Dev Reprod Toxicol* **2011**, *92*, 421–468, doi:10.1002/bdrb.20314.

108. Becker, M.; Hesse, V. Minipuberty: Why Does It Happen? *Horm Res Paediatr* **2020**, *93*, 76–84, doi:10.1159/000508329.

109. Blanco Mejia, S.; Messina, M.; Li, S.S.; Viguiliouk, E.; Chiavaroli, L.; Khan, T.A.; Srichaikul, K.; Mirrahimi, A.; Sievenpiper, J.L.; Kris-Etherton, P.; et al. A Meta-Analysis of 46 Studies Identified by the FDA Demonstrates That Soy Protein Decreases Circulating LDL and Total Cholesterol Concentrations in Adults. *Journal of Nutrition* **2019**, *149*, 968–981, doi:10.1093/jn/nxz020.

110. Yamagata, K.; Yamori, Y. Potential Effects of Soy Isoflavones on the Prevention of Metabolic Syndrome. *Molecules* **2021**, *26*, doi:10.3390/MOLECULES26195863.

111. Irace, C.; Marini, H.; Bitto, A.; Altavilla, D.; Polito, F.; Adamo, E.B.; Arcoraci, V.; Minutoli, L.; Di Benedetto, A.; Di Vieste, G.; et al. Genistein and Endothelial Function in Postmenopausal Women with Metabolic Syndrome. *Eur J Clin Invest* **2013**, *43*, 1025–1031, doi:10.1111/eci.12139.

112. Ganai, A.A.; Farooqi, H. Bioactivity of Genistein: A Review of in Vitro and in Vivo Studies. *Biomedicine & Pharmacotherapy* **2015**, *76*, 30–38, doi:10.1016/j.biopha.2015.10.026.

113. Simental-Mendia, L.E.; Gotto, A.M.; Atkin, S.L.; Banach, M.; Pirro, M.; Sahebkar, A. Effect of Soy Isoflavone Supplementation on Plasma Lipoprotein(a) Concentrations: A Meta-Analysis. *J Clin Lipidol* **2018**, *12*, 16–24, doi:10.1016/j.jacl.2017.10.004.

114. Zamora-Ros, R.; Knaze, V.; Luján-Barroso, L.; Kuhnle, G.G.C.; Mulligan, A.A.; Touillaud, M.; Slimani, N.; Romieu, I.; Powell, N.; Tumino, R.; et al. Dietary Intakes and Food Sources of Phytoestrogens in the European Prospective Investigation into Cancer and Nutrition (EPIC) 24-Hour Dietary Recall Cohort. *Eur J Clin Nutr* **2012**, *66*, 932–941, doi:10.1038/EJCN.2012.36.

115. Desmawati, D.; Sulastri, D. Phytoestrogens and Their Health Effect. *Open Access Maced J Med Sci* **2019**, *7*, 495–499, doi:10.3889/oamjms.2019.086.

116. Clarke, D.B.; Lloyd, A.S. Dietary Exposure Estimates of Isoflavones from the 1998 UK Total Diet Study. *Food Addit Contam* **2004**, *21*, 305–316, doi:10.1080/02652030410001668781.

117. Lee, A.; Beaubernard, L.; Lamothe, V.; Bennetau-Pelissero, C. New Evaluation of Isoflavone Exposure in the French Population. *Nutrients* **2019**, *11*, doi:10.3390/nu11102308.

118. Rizzati, V.; Briand, O.; Guillou, H.; Gamet-Payrastre, L. Effects of Pesticide Mixtures in Human and Animal Models: An Update of the Recent Literature. *Chem Biol Interact* **2016**, *254*, 231–246, doi:10.1016/J.CBI.2016.06.003.

119. Thongprakaisang, S.; Thiantanawat, A.; Rangkadilok, N.; Suriyo, T.; Satayavivad, J. Glyphosate Induces Human Breast Cancer Cells Growth via Estrogen Receptors. *Food and Chemical Toxicology* **2013**, *59*, 129–136, doi:10.1016/j.fct.2013.05.057.

120. McKinlay, R.; Plant, J.A.; Bell, J.N.B.; Voulvoulis, N. Calculating Human Exposure to Endocrine Disrupting Pesticides via Agricultural and Non-Agricultural Exposure Routes. *Science of the Total Environment* **2008**, *398*, 1–12.

121. Wang, H.-J.; Murphy, P.A. Mass Balance Study of Isoflavones during Soybean Processing. *J. Agric. Food Chem* **1996**, *44*, 2377–2383, doi:10.1021/jf950535p.

122. Anses INCA 2 : Les Résultats d'une Grande Étude, Étude Individuelle Nationale Sur Les Consommations Alimentaires 2006-2007.

123. Anses INCA 3 : Evolution Des Habitudes et Modes de Consommation, de Nouveaux Enjeux En Matière de Sécurité Sanitaire et de Nutrition.

124. Brouwer-Brolsma, E.M.; Brandl, B.; Buso, M.E.C.; Skurk, T.; Manach, C. Food Intake Biomarkers for Green Leafy Vegetables, Bulb Vegetables, and Stem Vegetables: A Review. *Genes Nutr* **2020**, *15*, doi:10.1186/s12263-020-00667-z.

125. Pujos-Guillot, E.; Hubert, J.; Martin, J.F.; Lyan, B.; Quintana, M.; Claude, S.; Chabanas, B.; Rothwell, J.A.; Bennetau-Pelissero, C.; Scalbert, A.; et al. Mass Spectrometry-Based Metabolomics for the Discovery of Biomarkers of Fruit and Vegetable Intake: Citrus Fruit as a Case Study. *J Proteome Res* **2013**, *12*, 1645–1659, doi:10.1021/pr300997c.

126. Grace, P.B.; Taylor, J.I.; Low, Y.-L.; Luben, R.N.; Mulligan, A.A.; Botting, N.P.; Dowsett, M.; Welch, A.A.; Khaw, K.-T.; Wareham, N.J.; et al. Phytoestrogen Concentrations in Serum and Spot Urine as Biomarkers for Dietary Phytoestrogen Intake and Their Relation to Breast Cancer Risk in European Prospective Investigation of Cancer and Nutrition-Norfolk. *Cancer epidemiology, biomarkers & prevention* **2004**, *13*, 698–708.

127. Heald, C.L.; Bolton-Smith, C.; Ritchie, M.R.; Morton, M.S.; Alexander, F.E. Phyto-Oestrogen Intake in Scottish Men: Use of Serum to Validate a Self-Administered Food-Frequency Questionnaire in Older Men. *Eur J Clin Nutr* **2006**, *60*, 129–135, doi:10.1038/sj.ejcn.1602277.

128. Inserm Maladies Auto-Immunes · Inserm, La Science Pour La Santé Available online: https://www.inserm.fr/dossier/maladies-auto-immunes/ (accessed on 12 April 2024).

129. Borchers, A.T.; Gershwin, M.E. Sociological Differences between Women and Men: Implications for Autoimmunity. *Autoimmun Rev* **2012**, *11*, A413–A421, doi:10.1016/J.AUTREV.2011.11.016.

130. Calixto, O.-J.; Anaya, J.-M. Socioeconomic Status. The Relationship with Health and Autoimmune Diseases. **2014**, doi:10.1016/j.autrev.2013.12.002.

131. Moroni, L.; Bianchi, I.; Lleo, A. Geoepidemiology, Gender and Autoimmune Disease. *Autoimmun Rev* **2012**, *11*, A386–A392, doi:10.1016/J.AUTREV.2011.11.012.

132. Yoon, K.; Kim, N. Roles of Sex Hormones and Gender in the Gut Microbiota. *J Neurogastroenterol Motil* **2021**, *27*, 314–325, doi:10.5056/JNM20208.

133. Kim, Y.S.; Unno, T.; Kim, B.Y.; Park, M.S. Sex Differences in Gut Microbiota. *World Journal of Men's Health* **2020**, *38*, 48–60, doi:10.5534/wjmh.190009.

134. Mayor, E. Gender Roles and Traits in Stress and Health. *Front Psychol* **2015**, *6*, doi:10.3389/fpsyg.2015.00779.

135. Lleo, A. Geoepidemiology and the Key Role of Sex Chromosomes on Autoimmune Diseases. In *Principles of Gender-Specific Medicine*; Elsevier, **2023**; pp. 331–346.

136. Zhao, J.-H.; Sun, S.-J.; Horiguchi, H.; Arao, Y.; Kanamori, N.; Kikuchi, A.; Oguma, E.; Kayama, F. A Soy Diet Accelerates Renal Damage in Autoimmune MRL/Mp-Lpr/Lpr Mice. *Int Immunopharmacol* **2005**, *5*, 1601–1610, doi:10.1016/j.intimp.2005.04.011.

137. Edwards, M.R.; Dai, R.; Heid, B.; Cecere, T.E.; Khan, D.; Mu, Q.; Cowan, C.; Luo, X.M.; Ahmed, S.A. Commercial Rodent Diets Differentially Regulate Autoimmune Glomerulonephritis, Epigenetics and Microbiota in MRL/Lpr Mice. *Int Immunol* **2017**, *29*, 263–276, doi:10.1093/intimm/dxx033.

138. Ameer, M.A.; Chaudhry, H.; Mushtaq, J.; Khan, O.S.; Babar, M.; Hashim, T.; Zeb, S.; Tariq, M.A.; Patlolla, S.R.; Ali, J.; et al. An Overview of Systemic Lupus Erythematosus (SLE) Pathogenesis, Classification, and Management. *Cureus* **2022**, *14*, doi:10.7759/cureus.30330.

139. Arnaud, L.; Gavand, P.E.; Voll, R.; Schwarting, A.; Ois Maurier, F.; Blaison, G.; Magy-Bertrand, N.; Pennaforte, J.-L.; Peter, H.-H.; Kieffer, P.; et al. Predictors of Fatigue and Severe Fatigue in a Large International Cohort of Patients with Systemic Lupus Erythematosus and a Systematic Review of the Literature. *Rheumatology* **2019**, *58*, 987–996, doi:10.1093/rheumatology/key398.

140. Aparicio-Soto, M.; Sánchez-Hidalgo, M.; Alarcón-De-La-Lastra, C. An Update on Diet and Nutritional Factors in Systemic Lupus Erythematosus Management. *Nutr Res Rev* **2017**, *30*, 118–137, doi:10.1017/S0954422417000026.

141. Connelly, K.; Morand, E.F. Systemic Lupus Erythematosus: A Clinical Update. *Intern Med J* **2021**, *51*, 1219–1228, doi:10.1111/imj.15448.

142. Etchegaray-Morales, I.; Mendoza-Pinto, C.; Munguía-Realpozo, P.; Osorio-Peña, Á.D.; Ibañez-Ovando, S.; Pineda, C.; García-Carrasco, M. Systemic Lupus Erythematosus, a Leading Cause of Death in Young Mexican Females: A Nationwide Population-Based Study, 2000-2020. *Rheumatol Int* **2022**, *42*, 1715–1720, doi:10.1007/S00296-022-05154-9.

143. Yen, E.Y.; Singh, R.R. Brief Report: Lupus—An Unrecognized Leading Cause of Death in Young Females: A Population-Based Study Using Nationwide Death Certificates, 2000–2015. *Arthritis and Rheumatology* **2018**, *70*, 1251–1255, doi:10.1002/art.40512.

144. Gulati, G.; Brunner, H.I. Environmental Triggers in Systemic Lupus Erythematosus. *Semin Arthritis Rheum* **2018**, *47*, 710–717, doi:10.1016/J.SEMARTHRIT.2017.10.001.

145. Somers, E.C.; Richardson, B.C. Environmental Exposures, Epigenetic Changes and the Risk of Lupus. *Lupus* **2014**, *23*, 568–576, doi:10.1177/0961203313499419.

146. Huang, R.G.; Li, X.B.; Wang, Y.Y.; Wu, H.; Li, K. Di; Jin, X.; Du, Y.J.; Wang, H.; Qian, F.Y.; Li, B.Z. Endocrine-Disrupting Chemicals and Autoimmune Diseases. *Environ Res* **2023**, *231*, 116222, doi:10.1016/J.ENVRES.2023.116222.

147. Borchers, A.T.; Naguwa, S.M.; Shoenfeld, Y.; Gershwin, M.E. The Geoepidemiology of Systemic Lupus Erythematosus. **2009**, doi:10.1016/j.autrev.2009.12.008.

148. Csaba, G. Effect of Endocrine Disruptor Phytoestrogens on the Immune System: Present and Future. *Acta Microbiol Immunol Hung* **2018**, *65*, 1–14, doi:10.1556/030.65.2018.018.

149. Hong, Y.H.; Wang, T.C.; Huang, C.J.; Cheng, W.Y.; Lin, B.F. Soy Isoflavones Supplementation Alleviates Disease Severity in Autoimmune-Prone MRL-Lpr/Lpr Mice. *Lupus* **2008**, *17*, 814–821, doi:10.1177/0961203308091287.

150. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **2021**, *71*, 209–249, doi:10.3322/caac.21660.

151. Gospodarowicz, M.; Benedet, L.; Hutter, R. V; Fleming, I.; Henson, D.E.; Sobin, L.H. History and International Developments in Cancer Staging. *Cancer Prev Control* **1998**, *2*, 262–268.

152. Grosclaude, P.; Bouvier, A.-M.; Trétarre, B.; Delafosse, P. Collecting and Using Data on Cancer Stages: Problems Posed in Population Studies. *Bulletin Epidémiologique Hebdomadaire* **2013**, *43–45*, 574–580.

153. Gaffney, D.; Matias-Guiu, X.; Mutch, D.; Scambia, G.; Creutzberg, C.; Fotopoulou, C.; Berek, J.S.; Concin, N. 2023 FIGO Staging System for Endometrial Cancer: The Evolution of the Revolution. *Gynecol Oncol* **2024**, *184*, 245–253, doi:10.1016/J.YGYNO.2024.02.002.

154. Caldon, C.E. Estrogen Signaling and the Dna Damage Response in Hormone Dependent Breast Cancers. *Front Oncol* **2014**, *4 MAY*, doi:10.3389/fonc.2014.00106.

155. Emons, G. Hormone-Dependent Cancers: Molecular Mechanisms and Therapeutical Implications. *Cells* **2023**, *12*.

156. Wang, C.; Li, J.; Ye, S.; Zhang, Y.; Li, P.; Wang, L.; Wang, T.H. Oestrogen Inhibits VEGF Expression and Angiogenesis in Triple-Negative Breast Cancer by Activating GPER-1. *Journal of Industrial and Management Optimization* **2017**, *13*, 3802–3811, doi:10.7150/jca.29233.

157. Koboldt, D.C.; Fulton, R.S.; McLellan, M.D.; Schmidt, H.; Kalicki-Veizer, J.; McMichael, J.F.; Fulton, L.L.; Dooling, D.J.; Ding, L.; Mardis, E.R.; et al. Comprehensive Molecular Portraits of Human Breast Tumours. *Nature* **2012**, *490*, 61–70, doi:10.1038/nature11412.

158. Rouhimoghadam, M.; Lu, A.S.; Salem, A.K.; Filardo, E.J. Therapeutic Perspectives on the Modulation of G-Protein Coupled Estrogen Receptor, GPER, Function. *Front Endocrinol (Lausanne)* **2020**, *11*, doi:10.3389/fendo.2020.591217.

159. Lappano, R.; Pisano, A.; Maggiolini, M. GPER Function in Breast Cancer: An Overview. *Front Endocrinol (Lausanne)* **2014**, *5*, 66, doi:10.3389/FENDO.2014.00066/BIBTEX.

160. Łukasiewicz, S.; Czeczelewski, M.; Forma, A.; Baj, J.; Sitarz, R.; Stanisławek, A. Breast Cancer—Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies—An Updated Review. *Cancers (Basel)* **2021**, *13*.

161. Daly, A.A.; Rolph, R.; Cutress, R.I.; Copson, E.R. A Review of Modifiable Risk Factors in Young Women for the Prevention of Breast Cancer. *Breast Cancer: Targets and Therapy* **2021**, *13*, 241–257.

162. Biancolella, M.; Ouédraogo, N.L.M.; Zongo, N.; Zohoncon, T.M.; Testa, B.; Rizzacasa, B.; Latini, A.; Conte, C.; Compaore, T.R.; Ouedraogo, C.M.R.N.; et al. Breast Cancer in West Africa: Molecular Analysis of BRCA Genes in Early-Onset Breast Cancer Patients in Burkina Faso. *Hum Genomics* **2021**, *15*, doi:10.1186/s40246-021-00365-w.

163. Beasley, J.M.; Coronado, G.D.; Livaudais, J.; Angeles-Llerenas, A.; Ortega-Olvera, C.; Romieu, I.; Lazcano-Ponce, E.; Torres-Mejía, G. Alcohol and Risk of Breast Cancer in Mexican Women. *Cancer Causes and Control* **2010**, *21*, 863–870, doi:10.1007/s10552-010-9513-x.

164. Arem, H.; Loftfield, E. Cancer Epidemiology: A Survey of Modifiable Risk Factors for Prevention and Survivorship. *Am J Lifestyle Med* **2018**, *12*, 200–210, doi:10.1177/1559827617700600.

165. Lasnier, A. Le Cancer Du Sein : Physiopathologie, Prédispositions Génétiques et Stratégies Thérapeutiques, Clermont-Ferrand, **2023**.

166. Petrie, W.K.; Dennis, M.K.; Hu, C.; Dai, D.; Arterburn, J.B.; Smith, H.O.; Hathaway, H.J.; Prossnitz, E.R. G Protein-Coupled Estrogen Receptor-Selective Ligands Modulate Endometrial Tumor Growth. *Obstet Gynecol Int* **2013**, *2013*, 1–17, doi:10.1155/2013/472720.

167. Pepermans, R.A.; Sharma, G.; Prossnitz, E.R. G Protein-Coupled Estrogen Receptor in Cancer and Stromal Cells: Functions and Novel Therapeutic Perspectives. *Cells* **2021**, *10*, 1–16, doi:10.3390/cells10030672.

168. Ignatov, A.; Ignatov, T.; Roessner, A.; Costa, S.D.; Kalinski, T. Role of GPR30 in the Mechanisms of Tamoxifen Resistance in Breast Cancer MCF-7 Cells. *Breast Cancer Res Treat* **2010**, *123*, 87–96, doi:10.1007/S10549-009-0624-6/FIGURES/6.

169. Molina, L.; Bustamante, F.; Ortloff, A.; Ramos, I.; Ehrenfeld, P.; Figueroa, C.D. Continuous Exposure of Breast Cancer Cells to Tamoxifen Upregulates GPER-1 and Increases Cell Proliferation. *Front Endocrinol (Lausanne)* **2020**, *11*, doi:10.3389/fendo.2020.563165.

170. Yao, J.; Deng, K.; Huang, J.; Zeng, R.; Zuo, J. Progress in the Understanding of the Mechanism of Tamoxifen Resistance in Breast Cancer. *Front Pharmacol* **2020**, *11*, doi:10.3389/fphar.2020.592912.

171. Hsu, L.H.; Chu, N.M.; Lin, Y.F.; Kao, S.H. G-Protein Coupled Estrogen Receptor in Breast Cancer. *Int J Mol Sci* **2019**, *20*, doi:10.3390/ijms20020306.

172. Girgert, R.; Emons, G.; Gründker, C. Estrogen Signaling in ERα-Negative Breast Cancer: ERβ and GPER. *Front Endocrinol (Lausanne)* **2019**, *10*, 781, doi:10.3389/FENDO.2018.00781/BIBTEX.

173. Pupo, M.; Bodmer, A.; Berto, M.; Maggiolini, M.; Dietrich, P.-Y.; Picard, D. A Genetic Polymorphism Repurposes the G-Protein Coupled and Membrane-Associated Estrogen Receptor GPER to a Transcription Factor-like Molecule Promoting Paracrine Signaling between Stroma and Breast Carcinoma Cells. *Oncotarget* **2017**, *8*, 46728–46744.

174. Chrisostome Bopassa, J.; Eghbali, M.; Toro, L.; Stefani, E. A Novel Estrogen Receptor GPER Inhibits Mitochondria Permeability Transition Pore Opening and Protects the Heart against Ischemia-Reperfusion Injury. *Am J Physiol Heart Circ Physiol* **2010**, *298*, 16–23, doi:10.1152/ajpheart.00588.2009.-Several.

175. Méndez-Luna, D.; Martínez-Archundia, M.; Maroun, R.C.; Ceballos-Reyes, G.; Fragoso-Vázquez, M.J.; González-Juárez, D.E.; Correa-Basurto, J. Deciphering the GPER/GPR30-Agonist and Antagonists Interactions Using Molecular Modeling Studies, Molecular Dynamics, and Docking Simulations. *J Biomol Struct Dyn* **2015**, *33*, 2161–2172, doi:10.1080/07391102.2014.994102.

176. Filardo, E.J.; Thomas, P. Minireview: G Protein-Coupled Estrogen Receptor-1, GPER-1: Its Mechanism of Action and Role in Female Reproductive Cancer, Renal and Vascular Physiology. *Endocrinology* **2012**, *153*, 2953, doi:10.1210/EN.2012-1061.

177. Prossnitz, E.R.; Barton, M. The G Protein-Coupled Oestrogen Receptor GPER in Health and Disease: An Update. *Nat Rev Endocrinol* **2023**, *19*, 407–424, doi:10.1038/s41574-023-00822-7.

178. Dennis, M.K.; Burai, R.; Ramesh, C.; Petrie, W.K.; Alcon, S.N.; Nayak, T.K.; Bologa, C.G.; Leitao, A.; Brailoiu, E.; Deliu, E.; et al. In Vivo Effects of a GPR30 Antagonist. *Nat Chem Biol* **2009**, *5*, 421–427, doi:10.1038/nchembio.168.

179. Luo, J.; Liu, D. Does GPER Really Function as a G Protein-Coupled Estrogen Receptor in Vivo? *Front Endocrinol (Lausanne)* **2020**, *11*, doi:10.3389/fendo.2020.00148.

180. Xu, S.; Yu, S.; Dong, D.; Lee, L.T.O. G Protein-Coupled Estrogen Receptor: A Potential Therapeutic Target in Cancer. *Front Endocrinol (Lausanne)* **2019**, *10*, doi:10.3389/fendo.2019.00725.

181. Yu, T.; Liu, M.; Luo, H.; Wu, C.; Tang, X.; Tang, S.; Hu, P.; Yan, Y.; Wang, Z.; Tu, G. GPER Mediates Enhanced Cell Viability and Motility via Non-Genomic Signaling Induced by 17β-Estradiol in Triple-Negative Breast Cancer Cells. *Journal of Steroid Biochemistry and Molecular Biology* **2014**, *143*, 392–403, doi:10.1016/j.jsbmb.2014.05.003.

182. Zhang, D.; Chen, H.; Wang, J.; Ji, J.; Imam, M.; Zhang, Z.; Yan, S. Current Progress and Prospects for G Protein-Coupled Estrogen Receptor in Triple-Negative Breast Cancer. *Front Cell Dev Biol* **2024**, *12*, doi:10.3389/fcell.2024.1338448.

183. Martínez-Muñoz, A.; Prestegui-Martel, B.; Méndez-Luna, D.; Fragoso-Vázquez, M.J.; García-Sánchez, J.R.; Bello, M.; Martínez-Archundia, M.; Chávez-Blanco, A.; Dueñas-González, A.; Mendoza-Lujambio, I.; et al. Selection of a GPER1 Ligand via Ligand-Based Virtual Screening Coupled to Molecular Dynamics Simulations and Its Anti-Proliferative Effects on Breast Cancer Cells. *Anticancer Agents Med Chem* **2018**, *18*, 1629–1638, doi:10.2174/1871520618666180510121431.

184. Ariazi, E.A.; Brailoiu, E.; Yerrum, S.; Shupp, H.A.; Slifker, M.J.; Cunliffe, H.E.; Black, M.A.; Donato, A.L.; Arterburn, J.B.; Oprea, T.I.; et al. The G Protein-Coupled Receptor GPR30 Inhibits Proliferation of Estrogen Receptor-Positive Breast Cancer Cells. *Cancer Res* **2010**, *70*, 1184–1194, doi:10.1158/0008-5472.CAN-09-3068.

185. Filardo, E.J.; Quinn, J.A.; Raymond Frackelton, A.; Bland, K.I. Estrogen Action Via the G Protein-Coupled Receptor, GPR30: Stimulation of Adenylyl Cyclase and CAMP-Mediated Attenuation of the Epidermal Growth Factor Receptor-to-MAPK Signaling Axis. *Molecular Endocrinology* **2002**, *16*, 70–84, doi:10.1210/MEND.16.1.0758.

186. Chan, Q.K.Y.; Lam, H.M.; Ng, C.F.; Lee, A.Y.Y.; Chan, E.S.Y.; Ng, H.K.; Ho, S.M.; Lau, K.M. Activation of GPR30 Inhibits the Growth of Prostate Cancer Cells through Sustained Activation of Erk1/2, c-Jun/c-Fos-Dependent Upregulation of P21, and Induction of G2 Cell-Cycle Arrest. *Cell Death Differ* **2010**, *17*, 1511–1523, doi:10.1038/cdd.2010.20.

187. Wang, C.; Lv, X.; He, C.; Hua, G.; Tsai, M.Y.; Davis, J.S. The G-Protein-Coupled Estrogen Receptor Agonist G-1 Suppresses Proliferation of Ovarian Cancer Cells by Blocking Tubulin Polymerization. *Cell Death Dis* **2013**, *4*, doi:10.1038/cddis.2013.397.

188. Weißenborn, C.; Ignatov, T.; Poehlmann, A.; Wege, A.K.; Costa, S.D.; Ana, ·; Zenclussen, C.; Ignatov, A. GPER Functions as a Tumor Suppressor in MCF-7 and SK-BR-3 Breast Cancer Cells. *J Cancer Res Clin Oncol* **2014**, *140*, 663–671, doi:10.1007/s00432-014-1598-2.

189. Ignatov, A.; Ignatov, T.; Weißenborn, C.; Eggemann, H.; Bischoff, J.; Semczuk, A.; Roessner, A.; Serban, •; Costa, D.; Kalinski, T. G-Protein-Coupled Estrogen Receptor GPR30 and Tamoxifen Resistance in Breast Cancer. *Breast Cancer Res Treat* **2011**, *128*, 457–466, doi:10.1007/s10549-011-1584-1.

190. Chen, Z.-J.; Wei, W.; Jiang, G.-M.; Liu, H.; Wei, W.-D.; Yang, X.; Wu, Y.-M.; Liu, H.; Wong, C.K.C.; Du, J.; et al. Activation of GPER Suppresses Epithelial Mesenchymal Transition of Triple Negative Breast Cancer Cells via NF-KB Signals. *Mol Oncol* **2016**, *10*, 775–788, doi:10.1016/j.molonc.2016.01.002.

191. Liang, S.; Chen, Z.; Jiang, G.; Zhou, Y.; Liu, Q.; Su, Q.; Wei, W.; Du, J.; Wang, H. Activation of GPER Suppresses Migration and Angiogenesis of Triple Negative Breast Cancer via Inhibition of NF-KB/IL-6 Signals. *Cancer Lett* **2016**, *386*, 12–23, doi:10.1016/j.canlet.2016.11.003.

192. Luo, H.; Liu, M.; Luo, S.; Yu, T.; Wu, C.; Yang, G.; Tu, G. Dynamic Monitoring of GPER-Mediated Estrogenic Effects in Breast Cancer Associated Fibroblasts: An Alternative Role of Estrogen in Mammary Carcinoma Development. *Steroids* **2016**, *112*, 1–11, doi:10.1016/j.steroids.2016.03.013.

193. Luo, H.; Yang, G.; Yu, T.; Luo, S.; Wu, C.; Sun, Y.; Liu, M.; Tu, G. GPER-Mediated Proliferation and Estradiol Production in Breast Cancer-Associated Fibroblasts. *Endocr Relat Cancer* **2014**, *21*, 355–369, doi:10.1530/ERC-13-0237.

194. Mo, Z.; Liu, M.; Yang, F.; Luo, H.; Li, Z.; Tu, G.; Yang, G. GPR30 as an Initiator of Tamoxifen Resistance in Hormone-Dependent Breast Cancer. *Breast Cancer Research* **2013**, *15*, 1–15, doi:10.1186/BCR3581/FIGURES/9.

195. Chen, Y.; Li, Z.; He, Y.; Shang, D.; Pan, J.; Wang, H.; Chen, H.; Zhu, Z.; Wan, L.; Wang, X. Estrogen and Pure Antiestrogen Fulvestrant (ICI 182 780) Augment Cell-Matrigel Adhesion of MCF-7 Breast Cancer Cells through a Novel G Protein Coupled Estrogen Receptor (GPR30)-to-Calpain Signaling Axis. *Toxicol Appl Pharmacol* **2014**, *275*, 176–181, doi:10.1016/j.taap.2014.01.005.

196. Catalano, S.; Giordano, C.; Panza, S.; Chemi, F.; Bonofiglio, D.; Lanzino, M.; Rizza, P.; Romeo, F.; Fuqua, S.A.W.; Maggiolini, M.; et al. Tamoxifen through GPER Upregulates Aromatase Expression: A Novel Mechanism Sustaining Tamoxifen-Resistant Breast Cancer Cell Growth. *Breast Cancer Res Treat* **2014**, *146*, 273–285, doi:10.1007/s10549-014-3017-4.

197. Lappano, R.; Rosano, C.; De Marco, P.; De Francesco, E.M.; Pezzi, V.; Maggiolini, M. Estriol Acts as a GPR30 Antagonist in Estrogen Receptor-Negative Breast Cancer Cells. *Mol Cell Endocrinol* **2010**, *320*, 162–170, doi:10.1016/j.mce.2010.02.006.

198. Girgert, R.; Emons, G.; Gründker, C. Inactivation of GPR30 Reduces Growth of Triple-Negative Breast Cancer Cells: Possible Application in Targeted Therapy. *Breast Cancer Res Treat* **2012**, *134*, 199–205, doi:10.1007/S10549-012-1968-X/FIGURES/4.

199. Steiman, J.; Peralta, E.A.; Louis, S.; Kamel, O. Biology of the Estrogen Receptor, GPR30, in Triple Negative Breast Cancer. *The American Journal of Surgery* **2013**, *206*, 698–703, doi:10.1016/j.amjsurg.2013.07.014.

200. Zhou, K.; Sun, P.; Zhang, Y.; You, X.; Li, P.; Wang, T. Estrogen Stimulated Migration and Invasion of Estrogen Receptor-Negative Breast Cancer Cells Involves an Ezrin-Dependent Crosstalk

between G Protein-Coupled Receptor 30 and Estrogen Receptor Beta Signaling. *Steroids* **2016**, *111*, 113–120, doi:10.1016/j.steroids.2016.01.021.

201. Albanito, L.; Sisci, D.; Aquila, S.; Brunelli, E.; Vivacqua, A.; Madeo, A.; Lappano, R.; Pandey, D.P.; Picard, D.; Mauro, L.; et al. Epidermal Growth Factor Induces G Protein-Coupled Receptor 30 Expression in Estrogen Receptor-Negative Breast Cancer Cells. *Endocrinology* **2008**, *149*, 3799–3808, doi:10.1210/EN.2008-0117.

202. Broselid, S.; Cheng, B.; Sjöström, M.; Lövgren, K.; Klug-De Santiago, H.L.P.; Belting, M.; Jirström, K.; Malmström, P.; Olde, B.; Bendahl, P.O.; et al. G Protein-Coupled Estrogen Receptor Is Apoptotic and Correlates with Increased Distant Disease-Free Survival of Estrogen Receptor-Positive Breast Cancer Patients. *Clinical Cancer Research* **2013**, *19*, 1681–1692, doi:10.1158/1078-0432.CCR-12-2376.

203. Poola, I.; Abraham, J.; Liu, A.; Marshalleck, J.J.; Dewitty, R.L. The Cell Surface Estrogen Receptor, G Protein-Coupled Receptor 30 (GPR30), Is Markedly Down Regulated During Breast Tumorigenesis. *Basic and Clinical Research* **2008**, 1–65.

204. Kuo, W.H.; Chang, L.Y.; Liu, D.L.Y.; Hwa, H.L.; Lin, J.J.; Lee, P.H.; Chen, C.N.; Lien, H.C.; Yuan, R.H.; Shun, C.T.; et al. The Interactions Between GPR30 and the Major Biomarkers in Infiltrating Ductal Carcinoma of the Breast in an Asian Population. *Taiwan J Obstet Gynecol* **2007**, *46*, 135–145, doi:10.1016/S1028-4559(07)60007-2.

205. Weißenborn, C.; Ignatov, T.; Ochel, H.J.; Costa, S.D.; Zenclussen, A.C.; Ignatova, Z.; Ignatov, A. GPER Functions as a Tumor Suppressor in Triple-Negative Breast Cancer Cells. *J Cancer Res Clin Oncol* **2014**, *140*, 713–723, doi:10.1007/s00432-014-1620-8.

206. Okamoto, M.; Mizukami, Y. GPER Negatively Regulates TNFα-Induced IL-6 Production in Human Breast Cancer Cells via NF-KB Pathway. *Endocr J* **2016**, *63*, 485–493.

207. Talia, M.; De Francesco, E.M.; Rigiracciolo, D.C.; Muoio, M.G.; Muglia, L.; Belfiore, A.; Maggiolini, M.; Sims, A.H.; Lappano, R. The G Protein-Coupled Estrogen Receptor (GPER) Expression Correlates with pro-Metastatic Pathways in ER-Negative Breast Cancer: A Bioinformatics Analysis. *Cells* **2020**, *9*, doi:10.3390/cells9030622.

208. Liu, Y.; Zou, T.; Wang, S.; Chen, H.; Su, D.; Fu, X.; Zhang, Q.; Kang, X. Genistein-Induced Differentiation of Breast Cancer Stem/Progenitor Cells through a Paracrine Mechanism. *Int J Oncol* **2016**, *48*, 1063–1072, doi:10.3892/ijo.2016.3351.

209. Vivacqua, A.; Bonofiglio, D.; Albanito, L.; Madeo, A.; Rago, V.; Carpino, A.; Musti, A.M.; Picard, D.; Andò, S.; Maggiolini, M. 17β-Estradiol, Genistein, and 4-Hydroxytamoxifen Induce the Proliferation of Thyroid Cancer Cells through the G Protein-Coupled Receptor GPR30. *Mol Pharmacol* **2006**, *70*, 1414–1423, doi:10.1124/MOL.106.026344.

210. Lin, B.C.; Suzawa, M.; Blind, R.D.; Tobias, S.C.; Bulun, S.E.; Scanlan, T.S.; Ingraham, H.A. Stimulating the GPR30 Estrogen Receptor with a Novel Tamoxifen Analogue Activates SF-1 and Promotes Endometrial Cell Proliferation., doi:10.1158/0008-5472.CAN-08-1622.

211. AFSSA; Anses *Les Études de l'alimentation Totale En France-EAT*; **2012**;

212. INRA; AFSSA *Etude de l'alimentation Totale Française Mycotoxines, Minéraux et Éléments Traces (EAT1)*; **2004**;

213. Lee, A.; Bensaada, S.; Lamothe, V.; Lacoste, M.; Bennetau-Pelissero, C. Endocrine Disruptors on and in Fruits and Vegetables: Estimation of the Potential Exposure of the French Population. *Food Chem* **2022**, *373*, doi:10.1016/j.foodchem.2021.131513.

214. Bensaada, S.; Raymond, I.; Breton, M.; Pellegrin, I.; Viallard, J.-F.; Bennetau-Pelissero, C. Development of an Assay for Soy Isoflavones in Women's Hair. *Nutrients* **2022**, *14*, doi:10.3390/nu14173619.

215. Lequin, R.M. Enzyme Immunoassay (EIA)/Enzyme-Linked Immunosorbent Assay (ELISA). *Clin Chem* **2005**, *51*, 2415–2418, doi:10.1373/clinchem.2005.051532.

216. Bennetau-Pelissero, C.; Arnal-Schnebelen, B.; Lamothe, V.; Sauvant, P.; Sagne, J.L.; Verbruggen, M.A.; Mathey, J.; Lavialle, O. ELISA as a New Method to Measure Genistein and Daidzein in Food and Human Fluids. *Food Chem* **2003**, *82*, 645–658, doi:10.1016/S0308-8146(03)00121-3.

217. Le Houerou, C.; Bennetau-Pelissero, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Syntheses of Novel Hapten-Protein Conjugates for Production of Highly Specific Antibodies to Formononetin, Daidzein and Genistein. *Tetrahedron* **2000**, *56*, 295–301, doi:10.1016/S0040-4020(99)00993-X.

218. Bennetau-Pelissero, C.; Le Houérou, C.; Lamothe, V.; Le Menn, F.; Babin, P.; Bennetau, B. Synthesis of Haptens and Conjugates for ELISAs of Phytoestrogens. Development of the Immunological Tests. *J Agric Food Chem* **2000**, *48*, 305–311, doi:10.1021/jf990896v.

219. Shinkaruk, S.; Lamothe, V.; Schmitter, J.M.; Fructus, A.; Sauvant, P.; Vergne, S.; Degueil, M.; Babin, P.; Bennetau, B.; Bennetau-Pelissero, C. Synthesis of Haptens and Conjugates for ELISA of Glycitein: Development and Validation of an Immunological Test. *J Agric Food Chem* **2008**, *56*, 6809–6817, doi:10.1021/jf800663s.

220. Shinkaruk, S.; Pinot, E.; Lamothe, V.; Schmitter, J.M.; Baguenard, L.; Bennetau, B.; Bennetau - Pelissero, C. Design and Validation of a Novel Immunological Test for Enterolactone. *Talanta* **2014**, *119*, 116–124, doi:10.1016/J.TALANTA.2013.10.034.

221. Vergne, S.; Bennetau-pelissero, C.; Lamothe, V.; Chantre, P.; Potier, M.; Asselineau, J.; Perez, P.; Durand, M.; Moore, N.; Sauvant, P. Higher Bioavailability of Isoflavones after a Single Ingestion of a Soya-Based Supplement than a Soya-Based Food in Young Healthy Males. *British Journal of Nutrition* **2008**, *99*, 333–344, doi:10.1017/S0007114507803953.

222. Aydin, S. A Short History, Principles, and Types of ELISA, and Our Laboratory Experience with Peptide/Protein Analyses Using ELISA. *Peptides (N.Y.)* **2015**, *72*, 4–15, doi:10.1016/j.peptides.2015.04.012.

223. Bensaada, S.; Peruzzi, G.; Cubizolles, L.; Denayrolles, M.; Bennetau-Pelissero, C. Traditional and Domestic Cooking Dramatically Reduce Estrogenic Isoflavones in Soy Foods. *Foods* **2024**, *13*, 999, doi:10.3390/foods13070999.

224. ElKaim, J. Drug Design in Silico, Criblage Virtuel de Protéines à Visée Thérapeutique, **2011**.

225. Thomas, P.; Pang, Y.; Filardo, E.J.; Dong, J. Identity of an Estrogen Membrane Receptor Coupled to a G Protein in Human Breast Cancer Cells. *Endocrinology* **2005**, *146*, 624–632, doi:10.1210/en.2004-1064.

226. Revankar, C.M.; Prossnitz, E.R.; Arterburn, J.B.; Cimino, D.F.; Sklar, L.A. A Transmembrane Intracellular Estrogen Receptor Mediates Rapid Cell Signaling. *Science (1979)* **2005**, *307*, 1625–1630, doi:10.1126/science.1105776.

227. O'Dea, A.; Sondergard, C.; Sweeney, P.; Arnatt, C.K. A Series of Indole-Thiazole Derivatives Act as GPER Agonists and Inhibit Breast Cancer Cell Growth. *ACS Med Chem Lett* **2018**, *9*, 901–906, doi:10.1021/acsmedchemlett.8b00212.

228. Bologa, C.G.; Revankar, C.M.; Young, S.M.; Edwards, B.S.; Arterburn, J.B.; Kiselyov, A.S.; Parker, M.A.; Tkachenko, S.E.; Savchuck, N.P.; Sklar, L.A.; et al. Virtual and Biomolecular Screening Converge on a Selective Agonist for GPR30. *Nat Chem Biol* **2006**, *2*, 207–212, doi:10.1038/nchembio775.

229. Dennis, M.K.; Field, A.S.; Burai, R.; Ramesh, C.; Petrie, W.K.; Bologa, C.G.; Oprea, T.I.; Yamaguchi, Y.; Hayashi, S.I.; Sklar, L.A.; et al. Identification of a GPER/GPR30 Antagonist with Improved Estrogen Receptor Counterselectivity. *Journal of Steroid Biochemistry and Molecular Biology* **2011**, *127*, 358–366, doi:10.1016/j.jsbmb.2011.07.002.

230. Koganti, S.; Snyder, R.; Gumaste, U.; Karamyan, V.T.; Thekkumkara, T. 2-Methoxyestradiol Binding of GPR30 down-Regulates Angiotensin AT 1 Receptor. *Eur J Pharmacol* **2014**, *723*, 131–140, doi:10.1016/j.ejphar.2013.10.064.

231. Liu, S.; Fu, R.; Zhou, L.H.; Chen, S.P. Application of Consensus Scoring and Principal Component Analysis for Virtual Screening against β-Secretase (BACE-1). *PLoS One* **2012**, *7*, doi:10.1371/journal.pone.0038086.

232. Trott, O.; Olson, A.J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J Comput Chem* **2010**, *31*, 455–461, doi:10.1002/jcc.21334.

233. Wang, R.; Lai, L.; Wang, S. Further Development and Validation of Empirical Scoring Functions for Structure-Based Binding Affinity Prediction. *J Comput Aided Mol Des* **2002**, *16*, 11–26.

234. Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-Based Scoring Function to Predict Protein-Ligand Interactions. *J. Mol. Bio* **2000**, 337–356.

235. Wang, R.; Wang, S. How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment. *J Chem Inf Comput Sci* **2001**, *41*, 1422–1426, doi:10.1021/ci010025x.

236. Boakye Oppong, F.; Yao Agbedra, S. Assessing Univariate and Multivariate Normality, A Guide For Non-Statisticians. *Mathematical Theory and Modeling* **2016**, *6*.

237. Mohd Razali, N.; Bee Wah, Y. Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *Journal of Statistical Modeling and Analytics* **2011**, *2*, 21–33.

238. Ranstam, J. Multiple P-Values and Bonferroni Correction. *Osteoarthritis Cartilage* **2016**, *24*, 763–764, doi:10.1016/J.JOCA.2016.01.008.

239. Morimoto, Y.; Beckford, F.; Franke, A.A.; Maskarinec, G. Urinary Isoflavonoid Excretion as a Biomarker of Dietary Soy Intake during Two Randomized Soy Trials. *Asia Pac J Clin Nutr* **2014**, *23*, 205–209.

240. Kim, J.; Kim, H.J.; Joung, H.; Park, M.K.; Li, S.; Song, Y.; Franke, A.A.; Paik, H.Y. Overnight Urinary Excretion of Isoflavones as an Indicator for Dietary Isoflavone Intake in Korean Girls of Pubertal Age. *British Journal of Nutrition* **2010**, *104*, 709–715, doi:10.1017/S0007114510000978.

241. French, M.R.; Thompson, L.U.; Hawker, G.A. Validation of a Phytoestrogen Food Frequency Questionnaire with Urinary Concentrations of Isoflavones and Lignan Metabolites in Premenopausal Women. *J Am Coll Nutr* **2007**, *26*, 76–82, doi:10.1080/07315724.2007.10719588.

242. Chávez-Suárez, K.M.; Ortega-Vélez, M.I.; Valenzuela-Quintanar, A.I.; Galván-Portillo, M.; López-Carrillo, L.; Esparza-Romero, J.; Saucedo-Tamayo, M.S.; Robles-Burgueño, M.R.; Palma-Durán, S.A.; Gutiérrez-Coronado, M.L.; et al. Phytoestrogen Concentrations in Human Urine as Biomarkers for Dietary Phytoestrogen Intake in Mexican Women. *Nutrients* **2017**, *9*, doi:10.3390/nu9101078.

243. Horner, N.K.; Kristal, A.R.; Prunty, J.; Skor, H.E.; Potter, J.D.; Lampe, J.W. Dietary Determinants of Plasma Enterolactone. *Cancer Epidemiology, Biomarkers & Prevention* **2002**, *11*, 121–126.

244. Krogholm, K.S.; Bysted, A.; Brantsaeter, A.L.; Jakobsen, J.; Rasmussen, S.E.; Kristoffersen, L.; Toft, U. Evaluation of Flavonoids and Enterolactone in Overnight Urine as Intake Biomarkers of Fruits, Vegetables and Beverages in the Inter99 Cohort Study Using the Method of Triads. *British Journal of Nutrition* **2000**, *108*, 1904–1912, doi:10.1017/S0007114512000104.

245. Katz-Agranov, N.; Zandman-Goddard, G. The Microbiome and Systemic Lupus Erythematosus. *Immunol Res* **2017**, *65*, 432–437, doi:10.1007/S12026-017-8906-2.

246. Mohd, R.; Chin, S.F.; Shaharir, S.S.; Cham, Q.S. Involvement of Gut Microbiota in SLE and Lupus Nephritis. *Biomedicines* **2023**, *11*, doi:10.3390/biomedicines11030653.

247. Toumi, E.; Goutorbe, B.; Plauzolles, A.; Bonnet, M.; Mezouar, S.; Militello, M.; Mege, J.L.; Chiche, L.; Halfon, P. Gut Microbiota in Systemic Lupus Erythematosus Patients and Lupus Mouse Model: A Cross Species Comparative Analysis for Biomarker Discovery. *Front Immunol* **2022**, *13*, doi:10.3389/fimmu.2022.943241.

248. Carreau, C.; Flouriot, G.; Bennetau-Pelissero, C.; Potier, M. Enterodiol and Enterolactone, Two Major Diet-Derived Polyphenol Metabolites Have Different Impact on ERα Transcriptional Activation in Human Breast Cancer Cells. *J Steroid Biochem Mol Biol* **2008**, *110*, 176–185, doi:10.1016/J.JSBMB.2008.03.032.

249. Khan, S.U.; Ahemad, N.; Chuah, L.H.; Naidu, R.; Htar, T.T. G Protein-Coupled Estrogen Receptor-1: Homology Modeling Approaches and Application in Screening New GPER-1 Modulators. *J Biomol Struct Dyn* **2022**, *40*, 3325–3335, doi:10.1080/07391102.2020.1844059.

250. Liu, K.; Wu, L.; Yuan, S.; Wu, M.; Xu, Y.; Sun, Q.; Li, S.; Zhao, S.; Hua, T.; Liu, Z.J. Structural Basis of CXC Chemokine Receptor 2 Activation and Signalling. *Nature* **2020**, *585*, 135–140, doi:10.1038/s41586-020-2492-5.

251. Bartuzi, D.; Kaczor, A.A.; Targowska-Duda, K.M.; Matosiuk, D. Recent Advances and Applications of Molecular Docking to g Protein-Coupled Receptors. *Molecules* **2017**, *22*.

252. Arterburn, J.B.; Prossnitz, E.R. G Protein-Coupled Estrogen Receptor GPER: Molecular Pharmacology and Therapeutic Applications. *Annu Rev Pharmacol Toxicol* **2023**, *63*, 295–320, doi:10.1146/annurev-pharmtox-031122.

253. Méndez-Luna, D.; Guzmán-Velázquez, S.; Padilla-Martínez, I.I.; García-Sánchez, J.R.; Bello, M.; García-Vázquez, J.B.; Mendoza-Figueroa, H.L.; Correa-Basurto, J. GPER Binding Site Detection and Description: A Flavonoid-Based Docking and Molecular Dynamics Simulations Study. *Journal of Steroid Biochemistry and Molecular Biology* **2024**, *239*, doi:10.1016/j.jsbmb.2024.106474.

254. Chan, Y.-T.; C-Y Lai, A.; Lin, R.-J.; Wang, Y.-H.; Wang, Y.-T.; Chang, W.-W.; Wu, H.-Y.; Lin, Y.-J.; Chang, W.-Y.; Wu, J.-C.; et al. GPER-Induced Signaling Is Essential for the Survival of Breast Cancer Stem Cells. *Int. J. Cancer* **2020**, *146*, 1674–1685, doi:10.1002/ijc.32588.

255. Rao Jala, V.; Radde, B.N.; Haribabu, B.; Klinge, C.M. Enhanced Expression of G-Protein Coupled Estrogen Receptor (GPER/GPR30) in Lung Cancer. *BMC Cancer* **2012**, *12*, doi:10.1186/1471-2407-12-624.

256. Horn-Ross, P.L.; Lee, M.; John, E.M.; Koo, J. Sources of Phytoestrogen Exposure among Non-Asian Women in California, USA. *Cancer Causes and Control* **2000**, *11*, 299–302.

257. Ayeni, K.I.; Jamnik, T.; Fareed, Y.; Flasch, M.; Braun, D.; Uhl, M.; Hartmann, C.; Warth, B. The Austrian Children's Biomonitoring Survey 2020 Part B: Mycotoxins, Phytotoxins, Phytoestrogens and Food Processing Contaminants. *Food and Chemical Toxicology* **2023**, *182*, doi:10.1016/j.fct.2023.114173.

258. Iwasaki, M.; Inoue, M.; Otani, T.; Sasazuki, S.; Kurahashi, N.; Miura, T.; Yamamoto, S.; Tsugane, S. Plasma Isoflavone Level and Subsequent Risk of Breast Cancer among Japanese Women: A Nested Case-Control Study from the Japan Public Health Center-Based Prospective Study Group. *Journal of Clinical Oncology* **2008**, *26*, 1677–1683, doi:10.1200/JCO.2007.13.9964.

259. Degen, G.H.; Blaszkewicz, M.; Shi, L.; Buyken, A.E.; Remer, T. Urinary Isoflavone Phytoestrogens in German Children and Adolescents – A Longitudinal Examination in the DONALD Cohort. *Mol Nutr Food Res* **2011**, *55*, 359–367, doi:10.1002/mnfr.201000325.

260. Mueller, S.O.; Simon, S.; Chae, K.; Metzler, M.; Korach, K.S. Phytoestrogens and Their Human Metabolites Show Distinct Agonistic and Antagonistic Properties on Estrogen Receptor α (ERα) and ERβ in Human Cells. *Toxicological Sciences* **2004**, *80*, 14–25, doi:10.1093/toxsci/kfh147.

261. Corona, G.; Kreimes, A.; Barone, M.; Turroni, S.; Brigidi, P.; Keleszade, E.; Costabile, A. Impact of Lignans in Oilseed Mix on Gut Microbiome Composition and Enterolignan Production in Younger Healthy and Premenopausal Women: An in Vitro Pilot Study. *Microb Cell Fact* **2020**, *19*, doi:10.1186/s12934-020-01341-0.

262. Touillaud, M.S.; Thiébaut, A.C.M.; Fournier, A.; Niravong, M.; Boutron-Ruault, M.C.; Clavel-Chapelon, F. Dietary Lignan Intake and Postmenopausal Breast Cancer Risk by Estrogen and Progesterone Receptor Status. *J Natl Cancer Inst* **2007**, *99*, 475–486, doi:10.1093/jnci/djk096.

263. Spink, D.C.; Spink, B.C.; Cao, J.Q.; Depasquale, J.A.; Pentecost, B.T.; Fasco, M.J.; Li, Y.; Sutter, T.R. Differential Expression of CYP1A1 and CYP1B1 in Human Breast Epithelial Cells and Breast Tumor Cells. *Carcinogenesis* **1998**, *19*, 291–298.

264. Bursztyka, J.; Perdu, E.; Pettersson, K.; Pongratz, I.; Fernández-Cabrera, M.; Olea, N.; Debrauwer, L.; Zalko, D.; Cravedi, J.P. Biotransformation of Genistein and Bisphenol A in Cell Lines Used for Screening Endocrine Disruptors. *Toxicology in Vitro* **2008**, *22*, 1595–1604, doi:10.1016/j.tiv.2008.06.013.

265. Nagata, C.; Wada, K.; Yamakawa, M.; Nakashima, Y.; Hayashi, M.; Takeda, N.; Yasuda, K. Sleep-Related Factors and Circulating Levels of Sex Hormones in Premenopausal Japanese Women. *Endocr J* **2023**, *70*, 267–273.

266. Branch, D.W. Physiologic Adaptations of Pregnancy. *American Journal of Reproductive Immunology* **1992**, *28*, 120–122, doi:10.1111/j.1600-0897.1992.tb00771.x.

267. German Federal Institute for Risk Assessment (BfR) *Foods and Food Supplements Containing Soy: Health Aspects*; **2023**;

268. German Federal Institute for Risk Assessment (BfR) *Isolated Isoflavones Are Not without Risk*; **2007**;

269. Norwegian Scientific Committee for Food Safety (VKM) *Risk Assessment of "Other Substances"-Isoflavones from Soy*; **2017**;

270. Santé Canada *Rapport Technique Pour Résumer La Justification Scientifique Des Nouvelles Directives de La Direction Des Produits de Santé Naturels Sur La Réglementation Des Produits à Base d'isoflavones de Soya*; **2009**;

271. Mayne, S. Statement from Susan Mayne, Ph.D., on Proposal to Revoke Health Claim That Soy Protein Reduces Risk of Heart Disease.

272. Suen, A.A.; Kenan, A.C.; Williams, C.J. Developmental Exposure to Phytoestrogens Found in Soy: New Findings and Clinical Implications. *Biochem Pharmacol* **2022**, *195*, 114848, doi:10.1016/J.BCP.2021.114848.

273. Maddison, A. *The World Economy*; OECD, **2001**; ISBN 9789264186088.

Annexes

Table des matières

Annexe 1 : Code_R® _____ 240

Annexe 2 : ISOLED_RESULTATS DES DOSAGES DE PHYTO-ESTROGENES_____ 251

1. Dosages dans les plasmas _____ 251
2. Dosages dans les urines _____ 253
3. Dosages dans les cheveux _____ 255

Annexe 3 : Classement des modèles de GPER générés par Modeler® _____ 258

Annexe 4 : Résultats de l'analyse ROC du *training set* _____ 259

Annexe 5 : Résultats des ACP_*Docking_training_set* _____ 261

Annexe 1 : Code_R®

a. Observation des données ISOLED et test de normalité

```
#################
# DATA IMPORT
#################
setwd("D:/THESE/ISOLED/Analyse_R")
library(tidyverse)
Data<-read.csv('D:/THESE/ISOLED/Analyse_R/Data_ISOLED.csv')
Data<- Data[,-1]
Data  <- rename(Data,
                Hair_G_D     = Hair.GEN...DAI..ng.mL.,
                Hair_IFs     = Hair.IFs..ng.mg.,
                Urine_G      = GEN.urin..microg.L.,
                Urine_D      = Daidztein.urin..microg.L.,
                Urine_E      = Equol.urin..microg.L.,
                Urine_IFs    = IFs.urin..mg.L.,
                Score_urin_IFs   = Score.urin.IFs,
                Score_urin_G     = Score.urin.GEN,
                Score_urin_D     = Score.urin.DAI,
                Plasma_G     = GEN.plasm..ng.mL.,
                Plasma_D     = DAI.plasm..ng.mL.,
                Plasma_E     = Equol.plasm..ng.mL.,
                Plasma_IFs   = IFs.plasm..ng.mL.,
                Score_plasm_IFs  = Score.plasm,
                Score_plasm_G = Score.plasm.GEN,
                Score_plasm_D = Score..plasm.DAI,
                Urine_ENL =  ENL.urin.microg.L,
                Score_Urin_ENL = Score.urin.ENL,
                Plasma_ENL = ENL.ng.mL,
                Score_plasm_ENL = Score.ENL.plasma
                )

str(Data)
summary(Data)

# Controlling column types
Data<-Data |> mutate(Status = factor(Status))
```

```r
#################
# TIDY DATA
#################
#library(tidyverse)

## DATA OBSERVATION DATA BRUT
## ==========================
Data1 <- na.omit(Data)
print(class(Data1))

Col_status<- c("red","orange","green4")

pdf("1_Data_Observation.pdf", width=9, height=9)
par(mfrow=c(6,4), mar=c(1, 3, 1, 1))
for (i in 2:23) {
  plot(Data1$Status, Data1[,i], ylab="", cex.axis=0.6, col=Col_status)
  nam<-names(Data1)[i]
  mtext(nam,2,2.2,cex=0.7)
}
dev.off()

# Data observation without outliers

Subset <- subset(Data1, Urine_G < 400)
Subset <- subset(Subset, Urine_D < 600)
Subset <- subset(Subset, Urine_ENL < 8000)
#EN plus pour pdf_2
Subset <- subset(Subset, Hair_G_D < 4)
Subset <- subset(Subset, IHS < 12)
#EN plus pour pdf_3
Subset <- subset(Subset, Plasma_G < 50)


pdf("3_Data_Observation_without_outliers.pdf", width=14, height=9)
par(mfrow=c(6,4), mar=c(1, 3, 1, 1))
for (i in 2:23) {
  plot(Subset$Status, Subset[,i], ylab="", cex.axis=0.6, col=Col_status)
  nam<-names(Subset)[i]
  mtext(nam,2,2.2,cex=0.7)
}
dev.off()


## NORMALITE DES DONNEES.
## ============================
#QQPLOT
Donnee<-Data[,2:23]
x <- rnorm(Donnee$Hair_G_D )
qqnorm(x)
qqline(x, col = "red")
```

```r
x <- rnorm(Donnee$Hair_IFs )
qqnorm(x)
qqline(x, col = "red")

x <- rnorm(Donnee$IHS )
qqnorm(x)
qqline(x, col = "red")

x <- rnorm(Donnee$IFS )
qqnorm(x)
qqline(x, col = "red")

x <- rnorm(Donnee$Urine_G)
qqnorm(x)
qqline(x, col = "red")

x <- rnorm(Donnee$Urine_D)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Urine_E)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Urine_IFs)
qqnorm(x)
qqline(x, col = "red")

x <- rnorm(Donnee$Score_urin_IFs)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Score_urin_G)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Score_urin_D)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Plasma_G)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Plasma_D)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Plasma_E)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Plasma_IFs)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Score_plasm)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Score_plasm_G)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Score_plasm_D)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Score_plasm_ENL)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Score_Urin_ENL)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Plasma_ENL)
qqnorm(x)
qqline(x, col = "red")
x <- rnorm(Donnee$Urine_ENL)
qqnorm(x)
qqline(x, col = "red")

# Decrire les variables avec des histogrammes
par(mar=c(3, 5, 1, 1))
layout(matrix(c(1:22),11,2))
for(i in 1:22) {hist(Donnee[,i],main=names(Donnee)[i],xlab="")}
layout(1)

#TEST ANDERSON-DARLING DE NORMALITE
install.packages("nortest")
library(nortest)
#ad.test(mtcars$mpg)
#for(i in 1:22)
ad.test(Donnee$Hair_G_D)
ad.test(Donnee$Hair_IFs)
ad.test(Donnee$Plasma_G)
ad.test(Donnee$Plasma_D)
ad.test(Donnee$Plasma_E)
ad.test(Donnee$Plasma_ENL)
ad.test(Donnee$Urine_G)
ad.test(Donnee$Urine_D)
ad.test(Donnee$Urine_E)
ad.test(Donnee$Urine_IFs)
ad.test(Donnee$Urine_ENL)
ad.test(Donnee$Score_urin_G)
ad.test(Donnee$Score_urin_D)
ad.test(Donnee$Score_Urin_ENL)
ad.test(Donnee$Score_urin_IFs)
ad.test(Donnee$Score_plasm_IFs)
ad.test(Donnee$Score_plasm_G)
ad.test(Donnee$Score_plasm_D)
ad.test(Donnee$Score_plasm_ENL)
ad.test(Donnee$IHS)
ad.test(Donnee$IFS)
ad.test(Donnee$Plasma_IFs)
```

b. Corrélations_ISOLED

```r
library("dplyr")

# CHARGER DONNEES
setwd("D:/THESE/ISOLED/Analyse_R")
Data <- read.csv("Data_ISOLED.csv")

# RENOMMER LES VARIABLES
Hair_IFs <- Data$Hair.IFs..ng.mg.
IHS <- Data$IHS
IFS <- Data$IFS
Hair_G_D  <- Data$Hair.GEN...DAI..ng.mL.
Hair_IFs <- Data$Hair.IFs..ng.mg.
Urine_G <- Data$GEN.urin..microg.L.
Urine_D <- Data$Daidztein.urin..microg.L.
Urine_E <- Data$Equol.urin..microg.L.
Urine_IFs <- Data$IFs.urin..mg.L.
Score_urin_IFs <- Data$Score.urin.IFs
Score_urin_G <- Data$Score.urin.GEN
Score_urin_D <- Data$Score.urin.DAI
Plasma_G <- Data$GEN.plasm..ng.mL.
Plasma_D <- Data$DAI.plasm..ng.mL.
Plasma_E <- Data$Equol.plasm..ng.mL.
Plasma_IFs <- Data$IFs.plasm..ng.mL.
Score_plasm_IFs <- Data$Score.plasm
Score_plasm_G <- Data$Score.plasm.GEN
Score_plasm_D <- Data$Score..plasm.DAI
Urine_ENL <- Data$ENL.urin.microg.L
Score_Urin_ENL <- Data$Score.urin.ENL
Plasma_ENL <- Data$ENL.ng.mL
Score_plasm_ENL <- Data$Score.ENL.plasma

View(Data)

# PERFORM PEARSON CORRELATION ANALYSIS

corr <- cor.test(x = Hair_IFs, y = IHS, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(x = Hair_IFs, y = IHS, method = "spearman", use = "complete.obs")
# SCATTERPLOT
plot(x=Hair_IFs, y=IHS, main = "Correlation de Spearman", xlab = "Concentration d'isoflavones dans les cheveux",
     ylab = "IHS",cex.lab=1.2, pch=19, col="lightblue")
# Regression line
abline(lm(IHS~Hair_IFs), col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 3, y = 20, cex = 3)
```

```r
corr <- cor.test(x = IFS, y = IHS, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(x = IFS, y = IHS, method = "spearman", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x=IFS, y=IHS, main = "Correlation de Spearman", xlab = "IFS",
     ylab = "IHS", cex.lab=1.2, pch=19, col="lightblue")
# Regression line
abline(lm(IHS~IFS), col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 5, y = 20, cex = 3)

corr <- cor.test(x = Hair_G_D, y = Hair_IFs, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(x = Hair_G_D, y = Hair_IFs, method = "spearman", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x = Hair_G_D, y = Hair_IFs, main = "Correlation de Spearman",
     xlab = "Concentrations capillaires de Genisteine et Daidzeine",
     ylab = "Scores capillaires d'isoflavones ", cex.lab=1.2, pch=19, col="lightblue")
# Regression line
abline(lm(Hair_IFs~Hair_G_D), col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 100, y = 15, cex = 3)


corr <- cor.test(x = Urine_ENL, y = Score_Urin_ENL, method = "pearson", exact = FALSE)
corr
spe_cor <- cor(x = Urine_ENL, y = Score_Urin_ENL, method = "pearson", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x = Urine_ENL, y = Score_Urin_ENL, main = "Correlation de Pearson",
     xlab = "Concentrations urinaires d'Enterolactone", ylab = "Scores urinaires d'Enterolactone",
     cex.lab=1.2, pch=19, col="lightblue")
# Regression line
abline(lm(Score_Urin_ENL~Urine_ENL), col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 2000, y = 23, cex = 3)


corr <- cor.test(x = Plasma_ENL, y = Score_plasm_ENL, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(x = Plasma_ENL , y = Score_plasm_ENL, method = "spearman", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x = Plasma_ENL, y = Score_plasm_ENL, main = "Correlation de Spearman",
     xlab = "Concentrations plasmatiques d'Enterolactone", ylab = "Scores plasmatiques d'Enterolactone",
     cex.lab=1.2, pch=19, col="lightblue")
# Regression line
abline(lm(Score_plasm_ENL~Plasma_ENL), col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 20, y = 30, cex = 3)
```

```r
##### transformation ln
Data[Data==0] <- NA #replace 0 with NA
Data
x <- log(Urine_G)
y <- Score_urin_G
corr <- cor.test(x, y, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(log(Urine_G), Score_urin_G, method = "spearman", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x , y, main = "Correlation de Spearman", xlab = "Concentrations urinaires de Genisteine transformees",
     ylab = "Scores urinaires de Genisteine", cex.lab=1.2, pch=19, col="lightblue")
# Regression line
Data[Data$GEN.urin..microg.L. != 0, ]
x[is.infinite(x)] <- NA
Data[!is.data.frame(Data$GEN.urin..microg.L.),]
abline(lm(y~x), Data, col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 6, y = 6, cex = 3)

Data <- read.csv("Data_ISOLED.csv")
Data[Data==0] <- NA
Data
x <- log(Urine_IFs)
y <- Score_urin_IFs
corr <- cor.test(x = Urine_IFs, y = Score_urin_IFs, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(log(Urine_IFs), Score_urin_IFs, method = "spearman", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x , y , main = "Correlation de Spearman", xlab = "Concentrations urinaires d'isoflavones transformees",
     ylab = "Scores urinaires d'isoflavones", cex.lab=1.2, pch=19, col="lightblue")
# Regression line
Data[Data$IFs.urin..mg.L. != 0, ]
x[is.infinite(x)] <- NA
Data[!is.data.frame(Data$IFs.urin..mg.L.),]
abline(lm(y~x), Data,col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = -2, y = 13, cex = 3)
```

```r
Data <- read.csv("Data_ISOLED.csv")
Data[Data==0] <- NA
Data
x <- log(Urine_D)
y <- Score_urin_D
corr <- cor.test(x = Urine_D, y = Score_urin_D, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(x = Urine_D, y = Score_urin_D, method = "spearman", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x , y , main = "Correlation de Spearman",
     xlab = "Concentrations urinaires de Daidzeine transformees",
     ylab = "Scores urinaires de Daidzeine", cex.lab=1.2, pch=19, col="lightblue")
# Regression line
Data[Data$Daidztein.urin..microg.L. != 0, ]
x[is.infinite(x)] <- NA
Data[!is.data.frame(Data$Daidztein.urin..microg.L.),]
abline(lm(y~x), col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 5, y = 6, cex = 3)

Data <- read.csv("Data_ISOLED.csv")
Data[Data==0] <- NA
Data
x <- log(Plasma_D)
y <- Score_plasm_D
corr <- cor.test(x = Plasma_D, y = Score_plasm_D, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(x = Plasma_D, y = Score_plasm_D, method = "spearman", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x , y , main = "Correlation de Spearman",
     xlab = "Concentrations plasmatiques de Daidzeine transformees", ylab = "Scores plasmatiques de Daidzeine",
     cex.lab=1.2, pch=19, col="lightblue")
# Regression line
Data[Data$DAI.plasm..ng.mL. != 0, ]
x[is.infinite(x)] <- NA
Data[!is.data.frame(Data$DAI.plasm..ng.mL.),]
abline(lm(y~x), Data, col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 2, y = 10, cex = 3)
```

```r
Data <- read.csv("Data_ISOLED.csv")
Data[Data==0] <- NA
Data
x <- log(Plasma_IFs)
y <- Score_plasm_IFs
corr <- cor.test(x = Plasma_IFs, y = Score_plasm_IFs, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(x = Plasma_IFs, y = Score_plasm_IFs, method = "spearman", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x , y , main = "Correlation de Spearman",
     xlab = "Concentrations plasmatiques d'isoflavones transformees",
     ylab = "Scores plasmatiques d'isoflavones",
     cex.lab=1.2, pch=19, col="lightblue")
# Regression line
Data[Data$IFs.plasm..ng.mL. != 0, ]
x[is.infinite(x)] <- NA
Data[!is.data.frame(Data$IFs.plasm..ng.mL.),]
abline(lm(y~x), col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 2, y = 13, cex = 3)

Data <- read.csv("Data_ISOLED.csv")
Data[Data==0] <- NA
Data
x <- log(Plasma_G)
y <- Score_plasm_G
corr <- cor.test(x = Plasma_G, y = Score_plasm_G, method = "spearman", exact = FALSE)
corr
spe_cor <- cor(x = Plasma_G, y = Score_plasm_G, method = "spearman", use = "complete.obs")
# SCATTERPLOT  (ajouter (xaxs = "i" , yaxs = "i") si on veut faire commenser les axes à 0
plot(x , y , main = "Correlation de Spearman",
     xlab = "Concentrations plasmatiques de Genisteine transformees",
     ylab = "Scores plasmatiques de Genisteine",
     cex.lab=1.2, pch=19, col="lightblue")
# Regression line
Data[Data$GEN.plasm..ng.mL. != 0, ]
x[is.infinite(x)] <- NA
Data[!is.data.frame(Data$GEN.plasm..ng.mL.),]
abline(lm(y~x), col="red")
# Pearson correlation
text(paste("Correlation :", round(x = spe_cor, 2)), x = 2, y = 13, cex = 3)
```

c. ACP_ISOLED

```
library("FactoMineR")
library("factoextra")
library("dplyr")
library("stats")
setwd("D:/THESE/ISOLED/Analyse_R")
#################################################
####AVEC toutes les variables
Data <- read.csv("Data_ISOLED.csv")
Data  <- rename(Data,
                Hair_G_D     = Hair.GEN...DAI..ng.mL.,
                Hair_IFs     = Hair.IFs..ng.mg.,
                Urine_G      = GEN.urin..microg.L.,
                Urine_D      = Daidztein.urin..microg.L.,
                Urine_E      = Equol.urin..microg.L.,
                Urine_IFs    = IFs.urin..mg.L.,
                Score_urin_IFs    = Score.urin.IFs,
                Score_urin_G      = Score.urin.GEN,
                Score_urin_D      = Score.urin.DAI,
                Plasma_G     = GEN.plasm..ng.mL.,
                Plasma_D     = DAI.plasm..ng.mL.,
                Plasma_E     = Equol.plasm..ng.mL.,
                Plasma_IFs   = IFs.plasm..ng.mL.,
                Score_plasm_IFs  = Score.plasm,
                Score_plasm_G = Score.plasm.GEN,
                Score_plasm_D = Score..plasm.DAI,
                Urine_ENL =  ENL.urin.microg.L,
                Score_Urin_ENL = Score.urin.ENL,
                Plasma_ENL = ENL.ng.mL,
                Score_plasm_ENL = Score.ENL.plasma
)

x = Data[1:63,- c(1,3) ]
str(x)
x <- na.omit(x)
summary(x)
x.active <- x[, 2:22]
######
#CENTRER ET REDUIRE LES DONNES
#####
x_norm <- scale(x.active, center = TRUE, scale = TRUE)
apply(x_norm, MARGIN = 2, FUN = mean)
apply(x_norm, MARGIN = 2, FUN = sd)
```

```
######
#ACP
#####
res.pca <- PCA(x_norm, graph = FALSE)
print(res.pca)
#Extract the eigenvalues/variances of principal components
#An eigenvalue > 1 indicates that PCs account for more variance than accounted
#by one of the original variables in standardized data. This is commonly used as
#a cutoff point for which PCs are retained. This holds true only when the data
#are standardized.
eig.val <- get_eigenvalue(res.pca)
eig.val
#Visualize the eigenvalues
fviz_eig(res.pca, addlabels = TRUE, ylim = c(0, 50))
#Extract the results for individuals and variables :
var <- get_pca_var(res.pca)
var
fviz_pca_var(res.pca, col.var = "black")
var$cos2
#visualize the cos2 of variables on all the dimension
library("corrplot")
corrplot(var$cos2, is.corr=FALSE)
# Total cos2 of variables on Dim.1 and Dim.2
fviz_cos2(res.pca, choice = "var", axes = 1:2)
# Color by cos2 values: quality on the factor map
fviz_pca_var(res.pca, col.var = "cos2",
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE # Avoid text overlapping
)
#Contributions of variables to PCs
head(var$contrib, 18)
corrplot(var$contrib, is.corr=FALSE)
# Contributions of variables to PC1
fviz_contrib(res.pca, choice = "var", axes = 1, top = 10)
# Contributions of variables to PC2
fviz_contrib(res.pca, choice = "var", axes = 2, top = 10)
# Contributions of variables to PC3
fviz_contrib(res.pca, choice = "var", axes = 3, top = 10)
#Dimension description
res.desc <- dimdesc(res.pca, axes = c(1,2,3), proba = 0.05)
# Description of dimension 1
res.desc$Dim.1
# Description of dimension 2
res.desc$Dim.2
# Description of dimension 3
res.desc$Dim.3
```

```r
#Visualize the results individuals and variables, respectively.
ind <- get_pca_ind(res.pca)
ind
fviz_pca_ind(res.pca, col.ind = "cos2",
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE # Avoid text overlapping (slow if many points)
)
fviz_pca_ind(res.pca,
             geom.ind = "point",
             col.ind = x$Status, # color by groups
             palette = c("red","orange","green4"),
             addEllipses = TRUE, # Concentration ellipses
             legend.title = "Groups"
)

# Add confidence ellipses
fviz_pca_ind(res.pca, geom.ind = "point", col.ind = x$Status,
             palette = c("red","orange","green4"),
             addEllipses = TRUE, ellipse.type = "confidence",
             legend.title = "Groups"
)
########3D PLOTTING
library("plotly")
prin_comp <- prcomp(x.active, rank. = 3)

components <- prin_comp[["x"]]
components <- data.frame(components)
components$PC2 <- -components$PC2
components$PC3 <- -components$PC3
components = cbind(components, x$Status)

tot_explained_variance_ratio <- summary(prin_comp)[["importance"]]['Proportion of Variance',]
tot_explained_variance_ratio <- 100 * sum(tot_explained_variance_ratio)

tit = 'VARIANCE TOTALE EXPLIQUUEE = 77.3'

fig <- plot_ly(components, x = ~PC1, y = ~PC2, z = ~PC3, color = ~x$Status,
               colors = c("red","orange","green4") ) %>%
  add_markers(size = 12)


fig <- fig %>%
  layout(
    title = tit,
    scene = list(bgcolor = "#e5ecf6")
  )

fig
```

d. Prolifération cellulaire_GPER_test de significativité

```
##############################################
#PROLIFERATION_SIGNIFICATIVITE
#KRUSKALL-WALLIS

#Data import
setwd("D:/THESE/GPER/Prolif/R")
library(tidyverse)
library(ggpubr)
#library(ggplot2)
#library(dplyr)
library(rstatix)
my_data <- read.csv("Data_436_13_14_15_J6.csv")
#print the head of the file
head(my_data)
# Show the group levels
dataFrame <- read.csv("Data_436_13_14_15_J6.csv", stringsAsFactors=T)
levels(dataFrame$Group)
dataFrame$Group <- fct_relevel(dataFrame$Group,
                               c("Milieu", "G 10-8 M", "G 10-6 M", "D 10-6 M", "G15 10-6M",
                                 "G10-8 M + G15", "G10-6 M + G15", "D10-6 M + G15") )

group_by(dataFrame, Group) %>%
  summarise(
    count = n(),
    mean = mean(Prolif_percent, na.rm = TRUE),
    sd = sd(Prolif_percent, na.rm = TRUE),
    median = median(Prolif_percent, na.rm = TRUE),
    IQR = IQR(Prolif_percent, na.rm = TRUE)
  )
# Box plots
# ++++++++++++++++++++
# Plot Fluo by group and color by group
ggboxplot(dataFrame, x = "Group", y = "Prolif_percent")

#ggboxplot(dataFrame, x = "Group", y = "Fluo",
          # color = "Group", palette = c("#FFFF00", "#99FF33", "#66CC00", "#336600",
          #"#003300", "#CCFFFF", "#99FFFF", "#3333FF", "#000099"),
          # order = c("Medium", "G 10-8 M", "G 5x10-8 M", "G 10-7 M ", "G 10-6 M",
          #"D 10-8 M", "D 5x10-8 M", "D 10-7 M ", "D 10-6 M"),
          # ylab = "Fluorescence", xlab = "Treatment" )
```

```r
# Mean plots
# ++++++++++++++++++++
# Plot weight by group
# Add error bars: mean_se
# (other values include: mean_sd, mean_ci, median_iqr, ....)
#library("ggpubr")
#ggline(dataFrame, x = "Group", y = "Fluo",
       #add = c("mean_se", "jitter"),
       #order = c("Medium", "G 10-8 M", "G 5x10-8 M", "G 10-7 M ", "G 10-6 M",
       #"D 10-8 M", "D 5x10-8 M", "D 10-7 M ", "D 10-6 M", "DMSO"),
       #ylab = "Fluorescence", xlab = "Treatment")
#Compute Kruskal-Wallis test
kruskal.test(Prolif_percent ~ Group, data = dataFrame)
#Multiple pairwise-comparison between groups
pairwise.wilcox.test(dataFrame$Fluo, dataFrame$Group,
                     p.adjust.method = "BH", exact = FALSE)
pairwise.wilcox.test(dataFrame$Fluo, dataFrame$Group,
                     p.adjust.method = "bonferroni", exact = FALSE)


#####representation graph avec significativité
my_comparisons <- list(c("Milieu","G 10-8 M"), c("Milieu","G 10-6 M"),
                       c("Milieu","D 10-6 M"), c("Milieu","G15 10-6M"),
                       c("G 10-8 M","G10-8 M + G15"), c("G 10-6 M","G10-6 M + G15"),
                       c("D 10-6 M","D10-6 M + G15"))
ggboxplot(dataFrame, x = "Group", y = "Prolif_percent") +
  stat_compare_means(method = "kruskal.test", label.y = 500) +
  stat_compare_means(comparisons = my_comparisons, method = "wilcox.test",
                     label.y = c(210, 230, 250, 270, 350, 370, 390), label="p.adj.signif")
ggboxplot(dataFrame, x = "Group", y = "Prolif_percent") +
  stat_compare_means(method = "kruskal.test", label.y = 500) +
  stat_compare_means(comparisons = my_comparisons, method = "wilcox.test",
                     label.y = c(210, 230, 250, 270, 350, 370, 390), label="p.signif",
                     cex.axis = 0.1)
```

Annexe 2 : ISOLED_RESULTATS DES DOSAGES DE PHYTO-ESTROGENES

Deux familles de phyto-estrogènes ont été dosés. Les IFs (GEN, DAI et son métabolite Equol) ainsi que l'ENL (issu de la métabolisation des lignanes par la flore intestinale). Les valeurs manquantes dans les tableaux de résultats indiquent des concentrations non détectables ou non quantifiables dans les échantillons.

1. Dosages dans les plasmas

Plasma

N°	GEN ng/mL	DAI ng/mL	Equol ng/mL	IFs totales ng/mL	ENL ng/mL	Score ENL plasma	Score IFs plasma 24-48h	Score plasma GEN	Score plasma DAI
01			**0,89**	**0,89**	**20,61**	10,25	1,75	1,875	1,625
02			**1,61**	**1,61**	**10,12**	14,1	0,625	0,75	0,5
03	**45,11**	**24,99**	**3,56**	**73,66**	**38,40**	20,4	2	2,375	1,625
04			**0,87**	**0,87**	**9,47**	9,4	5,6125	5,825	5,4
05			**0,45**	**0,45**	**27,43**	19,7	1,5125	1,575	1,45
06	**47,30**	**43,25**	**1,83**	**92,38**	**24,40**	13,65	6,6	7,225	5,975
07	**17,72**	**7,24**	**3,22**	**28,18**	**22,91**	15,6	2,7875	2,975	2,6
08		**15,75**	**1,16**	**16,91**	**40,85**	16,05	5,1625	5,25	5,075
09			**1,57**	**1,57**	**22,47**	12,6	3,9125	4	3,825
10			**0,58**	**0,58**	**32,41**	11,05	1,725	1,875	1,575
11	**24,13**	**26,74**	**3,02**	**53,88**	**39,13**	14,45	3,6	3,625	3,575
12			**1,17**	**1,17**	**10,92**	15,45	2,675	2,725	2,625
13		**2,98**	**0,74**	**3,72**	**15,70**	15,65	1,05	1,1	1
14			**1,23**	**1,23**	**29,62**	10,75	0,6	0,6	0,6
15		**4,11**	**1,23**	**5,33**	**31,12**	16,5	1,45	1,5	1,4
16					**13,76**	6,4	2,8	2,8	2,8
17	**10,31**	**6,74**	**2,86**	**19,90**	**11,36**	6,45	3,1125	3,325	2,9
18		**2,11**	**0,70**	**2,81**	**13,45**	10,75	0,9125	0,975	0,85
19		**3,24**	**2,61**	**5,85**	**21,97**	13,45	2,25	2,275	2,225
20		**4,11**	**2,61**	**6,71**	**27,39**	17,15	2,3125	2,625	2
21			**0,99**	**0,99**	**57,17**	38,35	3,475	3,75	3,2
22	**5,70**	**6,21**		**11,91**	**34,10**	15,15	1,3625	1,45	1,275
23	**1,20**		**1,10**	**2,30**	**13,81**	15,95	1,4625	1,525	1,4
24			**1,89**	**1,89**	**18,30**	8,1	1,45	1,5	1,4
25		**3,73**		**3,73**	**45,63**	19,4	2,5125	2,75	2,275
26					**4,92**	13,35	1,6375	1,7	1,575
27					**26,59**	15,2	1,35	1,375	1,325
28					**8,96**	7,35	0,825	0,825	0,825
29					**11,65**	16,6	2,5625	2,875	2,25
30	**468,15**	**253,17**	**8,28**	**759,60**	**5,66**	11,65	16,5625	16,275	16,85
31					**12,67**	6,95	5,8125	5,875	5,75
32		**22,53**		**22,53**	**10,10**	12,55	3,65	3,4	3,9
33	**10,88**	**68,12**		**79,01**	**9,10**	9,95	1,5	1,5	1,5
34					**21,52**	11,2	0,625	0,625	0,625
35		**16,96**		**16,96**	**13,91**	15,4	5,8375	6,05	5,625
36	**38,70**	**7,13**	**1,98**	**47,81**	**22,52**	11,65	2,35	2,625	2,075

37					**11,93**	9,1	0,9375	1	0,875
38					**11,68**	11,4	1,6875	1,75	1,625
39	**138,15**	**53,25**	**3,40**	**194,80**	**7,31**	6,5	6,75	7	6,5
40					**21,63**	12,1	2,1	2,125	2,075
41					**6,41**	6,1	2,2	2,5	1,9
42					**13,09**	7,4	1,3125	1,375	1,25
43					**13,60**	7,2	0,625	0,625	0,625
44					**6,69**	7,75	0,825	0,825	0,825
45					**6,03**	11,8	0,91875	0,95	0,8875
46					**7,62**	10,24	1,2625	1,375	1,15
47					**6,11**	10,15	1,5375	1,6	1,475
48	**67,82**	**133,84**	**46,57**	**248,22**	**5,61**	4,5	5,625	5,75	5,5
49	**36,38**	**36,68**		**73,06**	**28,28**	20,35	3,45	3,7	3,2
50					**1,94**	8,5	0,6	0,6	0,6
51					**8,19**	5,25	1,1	1,1	1,1
52					**10,22**	6,6	1,5375	1,625	1,45
53	**83,70**	**40,30**	**15,95**	**139,95**	**25,04**	21,45	6,4875	7,525	5,45
54	**39,05**	**22,81**	**4,96**	**66,82**	**21,73**	13,5	4,2	4,475	3,925
55					**12,71**	9,7	1,725	1,85	1,6
56					**6,67**	14,35	1,8875	1,975	1,8
57					**7,61**	3,2	0,8	0,8	0,8
58	**46,17**			**46,17**	**11,87**	18	1,3875	1,575	1,2
59	**37,62**		**15,54**	**53,17**	**0,93**	10,3	0,7875	0,75	0,825
60	**49,22**		**20,57**	**69,79**	**1,50**	8,7	2,9375	3	2,875
61	**86,15**		**7,25**	**93,40**	**0,65**	11,7	2,1875	2,375	2
62	**46,70**		**6,10**	**52,80**	**4,37**	16,4	1,925	2	1,85
63	**13,13**			**13,13**	**0,73**	10,4	5,5625	5,875	5,875
Nombre	21,00	23,00	32,00	40,00	63,00	63,00	63,00	63,00	63,00
Moyenne	62,54	35,04	5,20	57,89	879,36	9,47	2,66	2,79	2,55
Ecart-type	98,30	56,10	8,97	126,22	1665,79	4,29	2,46	2,49	2,45
Médiane	39,05	16,96	1,93	16,94	304,73	9,00	1,89	1,98	1,63
Max	468,15	253,17	46,57	759,60	8219,26	27,75	16,56	16,28	16,85
Min	1,20	2,11	0,45	0,45	6,76	0,00	0,60	0,60	0,50

2. Dosages dans les urines

Les concentrations de Génistéine, Daidzéine et Equol présentées ici sont corrigées par rapport à la créatinine.

Urine

N° Participante	GEN µg/L	DAI µg/L	Equol µg/L	IF totaux µg/L	ENL µg/L	Score urin, ENL	Score Alimentaire Urine 24-48h	Score Urine GEN	Score urine DAI
01		**27,88**		**27,88**	**283,03**	10,1	1	0,5	0,5
02	**21,51**	**42,24**	**23,11**	**86,86**	**428,59**	10,75	1	0,5	0,5
03	**186,69**	**293,17**	**26,55**	**506,41**	**1221,58**	16,25	2,2	1,1	1,1
04	**97,39**	**342,86**	**47,59**	**487,84**	**852,83**	7,125	8	4	4
05	**30,23**	**110,40**	**21,22**	**161,85**	**4695,25**	19,35	3	1,5	1,5
06	**987,53**	**2338,61**	**303,58**	**3629,71**	**244,07**	13,05	7,4	3,7	3,7
07	**26,80**	**55,33**	**15,23**	**97,35**	**714,50**	12,25	4	2	2
08	**209,96**	**915,21**	**93,32**	**1218,49**	**521,96**	14	6,8	3,35	3,35
09	**70,02**	**40,37**	**35,28**	**145,67**	**198,05**	11,5	3	1,5	1,5
10	**35,06**	**41,88**	**8,88**	**85,82**	**236,84**	8,75	1,6	0,8	0,8
11		**69,42**		**69,42**	**641,94**	12,1	2,4	0,7	0,7
12	**58,84**	**63,50**		**122,34**	**426,38**	10,75	2,6	1,3	1,3
13	**206,38**	**454,14**	**27,82**	**688,34**	**991,27**	13,25	4,6	2,3	2,3
14	**14,69**	**15,43**	**8,56**	**38,68**	**322,35**	6,6	1,2	0,6	0,6
15	**28,91**	**30,02**	**11,85**	**70,78**	**1035,72**	16,35	0,4	0,2	0,2
16	**27,39**	**93,99**	**15,06**	**136,44**	**189,17**	5,35	3	1,5	1,5
17	**159,35**	**173,59**	**14,65**	**347,58**	**105,11**	6,55	4,6	2,05	2,05
18		**41,94**		**41,94**	**323,06**	7	1,5	0,75	0,75
19	**34,56**	**49,65**		**84,21**	**665,40**	10,1	1	0,5	0,5
20	**39,29**	**137,29**		**176,57**	**685,53**	10,5	3,4	1,6	1,6
21	**69,20**	**245,75**	**24,77**	**339,71**	**8076,13**	27,75	5,4	2,7	2,7
22	**23,66**	**56,48**	**9,84**	**89,98**	**954,75**	9,6	2,2	1,1	1,1
23	**42,38**	**62,25**	**26,70**	**131,33**	**1620,81**	11,5	2,6	1,3	1,3
24	**23,70**	**22,88**	**19,41**	**65,99**	**35,09**	5,875	0,4	0,2	0,2
25	**28,69**	**59,47**	**37,01**	**125,17**	**253,50**	14,1	3,2	1,6	1,6
26	**21,43**	**23,14**	**11,15**	**55,71**	**7,72**	9,5	1,1	0,55	0,55
27	**28,59**	**43,01**	**16,90**	**88,50**	**914,93**	11,25	1,2	0,6	0,6
28	**38,74**	**26,02**	**6,29**	**71,05**	**68,11**	6,25	1,4	0,7	0,7
29		**85,15**	**97,86**	**183,01**	**2094,24**	11,75	4,1	1,8	1,8
30	**231158,12**	**222912,86**	**313,53**	**454384,51**	**259,97**	9,875	16,9	7,95	7,95
31	**46,86**	**40,91**	**11,47**	**99,24**	**65,52**	5,375	1,5	0	0,7
32					**12,45**	8,875	NA	NA	NA
33	**30,88**	**86,09**	**35,78**	**152,75**	**304,73**	7,875	3	1,5	1,5
34	**23,54**	**43,93**		**67,46**	**95,68**	7,85	1	0,5	0,5
35	**316,16**	**492,53**	**87,13**	**895,82**	**1277,77**	11,1	5	2,5	2,5
36	**28,11**	**206,80**	**22,05**	**256,95**	**554,38**	9	3	1	1
37	**16,18**	**19,47**	**12,04**	**47,68**	**294,03**	6,375	1,6	0,8	0,8
38	**50,61**	**80,69**	**13,29**	**144,59**	**183,25**	8,875	2,8	1,4	1,4

39	**2188,47**	**3208,22**	**68,26**	**5464,94**	**71,85**	4,75	9,95	4,75	4,75
40					**22,67**	5,25	NA	NA	NA
41	**95,29**	**99,21**	**48,48**	**242,98**	**237,78**	8,5	2,4	1,2	1,2
42	**22,42**	**53,39**	**16,13**	**91,93**	**18,90**	5,625	1,6	0,8	0,8
43	**22,68**	**26,67**		**49,35**	**176,31**	5,75	1	0,5	0,5
44					**104,22**	3,875	NA	NA	NA
45			**9,53**	**9,53**	**112,12**	7,45	0,7	0,35	0,35
46	**57,62**		**34,35**	**91,97**	**23,78**	8,9	1	0,5	0,5
47	**41,40**	**78,70**	**3,37**	**123,47**	**6,76**	7	2,8	1,4	1,4
48	**114,59**	**70,10**	**20,61**	**205,30**	**420,43**	3,375	3	1,5	1,5
49	**45,30**	**2009,67**	**21,07**	**2076,04**	**4988,97**	13,5	5,9	2,7	2,7
50	**55,38**	**47,03**	**48,71**	**151,12**	**37,46**	5	1,2	0,6	0,6
51		**24,49**	**6,50**	**30,99**	**528,26**	9,25	1,2	0,6	0,6
52		**23,36**	**14,40**	**37,76**	**423,14**	5,5	1,8	0,9	0,9
53	**4634,79**	**2252,21**	**1469,81**	**8356,81**	**4360,90**	14,6	7,8	3,4	3,4
54	**148,78**	**523,04**	**120,79**	**792,61**	**1314,81**	10,35	3,4	1,7	1,7
55	**83,07**	**99,95**	**66,22**	**249,24**	**801,02**	7,225	1,8	0,9	0,9
56	**12,13**	**28,52**	**6,85**	**47,50**	**202,28**	8,625	2,2	1,1	1,1
57	**38,68**	**51,49**	**31,56**	**121,74**	**36,97**	2,875	1	0,5	0,5
58	**93,44**	**399,25**	**76,67**	**569,36**	**8219,26**	0	2,2	1,1	1,1
59	**20,82**	**127,93**	**26,27**	**175,02**	**228,70**	12	1,2	0,6	0,6
60	**42,89**	**58,35**	**ND**	**101,25**	**185,99**	10	2	1	1
61	**123,45**	**379,59**	**21,40**	**524,44**	**434,74**	11,25	3	1,5	1,5
62	**43,87**	**77,45**	**33,83**	**155,16**	**521,99**	14,25	2,6	1,3	1,3
63	**36,23**	**140,04**	**ND**	**176,28**	**64,39**	7,25	2,6	1,3	1,3
Nombre	53,00	58,00	50,00	60,00	63,00	63,00	63,00	63,00	63,00
Moyenne	4567,98	4131,43	70,85	8087,81	12,40	2,66	2,79	2,55	12,40
Ecart-type	31730,8	29238,11	210,86	58608,60	5,40	2,46	2,49	2,45	5,40
Médiane	42,4	69,76	22,58	133,89	11,65	1,89	1,98	1,63	11,65
Max	231158,1	222912,9	1469,81	454384,51	38,35	16,56	16,28	16,85	38,35
Min	12,13	15,43	3,37	9,53	3,20	0,60	0,60	0,50	3,20

3. Dosages dans les cheveux

CHEVEUX

N° Participante	**GEN ng/mg**	± ET	**DAI ng/mg**	± ET	**Equol ng/mg**	± ET	**GEN+DAI ng/mg**	± ET	**GEN+DAI+ Equol ng/mg**	± ET	IHS (score capillaire)
01	**0.00**	± 0.00	**0.14**	± 0.00	**0.00**	± 0.00	**0.14**	± 0.00	**0.14**	± 0.00	2.304
02	**0.76**	± 0.00	**0.11**	± 0.00	**0.10**	± 0.00	**0.87**	± 0.00	**0.97**	± 0.00	1.6
03	**0.00**	± 0.00	**0.19**	± 0.02	**0.14**	± 0.02	**0.19**	± 0.02	**0.33**	± 0.04	2.81
04	**0.69**	± 0.00	**1.14**	± 0.20	**0.12**	± 0.06	**1.83**	± 0.20	**1.95**	± 0.25	5.22
05	**0.00**	± 0.00	**0.15**	± 0.01	**0.00**	± 0.00	**0.15**	± 0.01	**0.15**	± 0.01	1.39
06	**0.00**	± 0.00	**0.37**	± 0.00	**0.12**	± 0.03	**0.37**	± 0.00	**0.48**	± 0.03	3.42
07	**0.00**	± 0.00	**0.13**	± 0.00	**0.00**	± 0.00	**0.13**	± 0.00	**0.13**	± 0.00	2.27
08	**2.08**	± 0.17	**2.44**	± 0.46	**0.17**	± 0.01	**4.52**	± 0.63	**4.69**	± 0.64	16.2
09	**0.00**	± 0.00	**0.40**	± 0.00	**0.08**	± 0.00	**0.40**	± 0.00	**0.48**	± 0.00	3.88
10	**NA**										
11	**1.18**	± 0.06	**0.87**	± 0.05	**0.00**	± 0.00	**2.05**	± 0.11	**2.05**	± 0.11	9.79
12	**NA**										
13	**0.29**	± 0.06	**0.27**	± 0.10	**0.00**	± 0.00	**0.55**	± 0.16	**0.55**	± 0.16	3.78
14	**0.00**	± 0.00	**0.25**	± 0.00	**0.00**	± 0.00	**0.25**	± 0.00	**0.25**	± 0.00	2.02
15	**0.95**	± 0.42	**0.40**	± 0.09	**0.00**	± 0.00	**1.35**	± 0.52	**1.35**	± 0.52	5.98
16	**1.21**	± 0.30	**0.89**	± 0.15	**0.08**	± 0.01	**2.10**	± 0.45	**2.10**	± 0.45	5.2
17	**1.23**	± 0.26	**0.63**	± 0.16	**0.00**	± 0.00	**1.87**	± 0.42	**1.87**	± 0.42	7.13
18	**4.25**	± 1.17	**4.97**	± 0.29	**0.00**	± 0.00	**9.22**	± 1.46	**9.22**	± 1.46	22.77
19	**0.34**	± 0.04	**0.71**	± 0.03	**0.00**	± 0.00	**1.04**	± 0.07	**1.04**	± 0.07	4.32
20	**1.08**	± 0.11	**0.82**	± 0.12	**0.00**	± 0.00	**1.89**	± 0.23	**1.89**	± 0.23	6.48
21	**2.44**	± 0.15	**0.71**	± 0.09	**0.16**	± 0.04	**3.15**	± 0.24	**3.30**	± 0.28	10.17
22	**0.43**	± 0.05	**0.66**	± 0.01	**0.14**	± 0.00	**1.10**	± 0.06	**1.24**	± 0.06	5.51
23	**0.26**	± 0.05	**0.30**	± 0.07	**0.04**	± 0.01	**0.56**	± 0.12	**0.59**	± 0.13	7.13
24	**0.61**	± 0.06	**0.34**	± 0.10	**0.20**	± 0.01	**0.95**	± 0.16	**1.15**	± 0.17	8.16
25	**0.21**	± 0.09	**1.16**	± 0.48	**0.06**	± 0.00	**1.38**	± 0.57	**1.43**	± 0.57	4.75

26	**0.62**	±	0.02	**0.60**	±	0.04	**0.14**	±	0.03	**1.22**	±	0.05	**1.36**	±	0.08	4.69
27	**0.36**	±	0.02	**0.35**	±	0.00	**0.07**	±	0.02	**0.71**	±	0.02	**0.78**	±	0.04	4.05
28	**2.28**	±	0.49	**1.67**	±	0.28	**ND**	±	ND	**3.96**	±	0.77	**3.96**	±	0.77	12.18
29	**1.18**	±	0.17	**2.27**	±	0.35	**ND**	±	ND	**3.45**	±	0.52	**3.45**	±	0.52	13
30	**0.36**	±	0.04	**0.57**	±	0.02	**0.05**	±	0.00	**0.93**	±	0.06	**0.98**	±	0.06	4.74
31	**0.95**	±	0.06	**0.56**	±	0.14	**0.04**	±	0.01	**1.51**	±	0.19	**1.55**	±	0.20	5.94
32	**0.57**	±	0.05	**1.62**	±	0.66	**0.18**	±	0.06	**2.18**	±	0.71	**2.36**	±	0.77	9.69
33	**NA**															
34	**0.31**	±	0.12	**0.31**	±	0.05	**0.04**	±	0.01	**0.63**	±	0.17	**0.67**	±	0.18	3.74
35	**0.36**	±	0.02	**0.61**	±	0.03	**0.00**	±	0.00	**0.96**	±	0.05	**0.96**	±	0.05	3.98
36	**0.00**	±	0.00	**1.62**	±	0.66	**0.06**	±	0.02	**1.62**	±	0.66	**1.67**	±	0.68	7.44
37	**0.65**	±	0.02	**0.13**	±	0.07	**0.13**	±	0.01	**0.78**	±	0.08	**0.90**	±	0.09	2.1
38	**0.37**	±	0.07	**1.62**	±	0.66	**0.08**	±	0.01	**1.99**	±	0.73	**2.08**	±	0.74	6.75
39	**0.46**	±	0.04	**0.13**	±	0.07	**0.08**	±	0.00	**0.58**	±	0.11	**0.67**	±	0.11	4.06
40	**NA**															
41	**0.24**	±	0.00	**0.25**	±	0.00	**0.05**	±	0.00	**0.49**	±	0.00	**0.54**	±	0.00	4.96
42	**0.31**	±	0.06	**1.36**	±	0.00	**0.00**	±	0.00	**1.66**	±	0.06	**1.66**	±	0.06	7.41
43	**NA**															
44	**0.00**	±	0.00	**0.00**	±	0.00	**0.00**	±	0.00	**0.00**	±	0.00	**0.00**	±	0.00	1.43
45	**0.00**	±	0.00	**0.36**	±	0.00	**0.12**	±	0.05	**0.36**	±	0.00	**0.48**	±	0.05	2.0352
46	**0.30**	±	0.04	**0.38**	±	0.06	**0.00**	±	0.00	**0.67**	±	0.10	**0.67**	±	0.10	5.4
47	**0.00**	±	0.00	**0.22**	±	0.00	**0.06**	±	0.00	**0.22**	±	0.00	**0.28**	±	0.00	3.6
48	**0.21**	±	0.04	**0.20**	±	0.00	**0.16**	±	0.02	**0.41**	±	0.04	**0.56**	±	0.05	2.75
49	**0.42**	±	0.06	**1.61**	±	0.41	**0.14**	±	0.02	**2.04**	±	0.46	**2.18**	±	0.48	7.66
50	**1.06**	±	0.12	**0.54**	±	0.12	**0.29**	±	0.05	**1.60**	±	0.24	**1.89**	±	0.29	5.32
51	**0.62**	±	0.03	**0.20**	±	0.05	**0.12**	±	0.02	**0.81**	±	0.07	**0.94**	±	0.10	2.6
52	**0.48**	±	0.01	**0.39**	±	0.07	**0.31**	±	0.01	**0.88**	±	0.08	**1.18**	±	0.09	1.75
53	**1.12**	±	0.11	**0.66**	±	0.01	**0.38**	±	0.00	**1.78**	±	0.12	**2.16**	±	0.12	5.52
54	**0.28**	±	0.04	**0.16**	±	0.01	**0.24**	±	0.04	**0.43**	±	0.05	**0.67**	±	0.09	1.56
55	**0.69**	±	0.13	**0.64**	±	0.01	**0.25**	±	0.00	**1.33**	±	0.15	**1.58**	±	0.15	3.52
56	**0.31**	±	0.00	**0.30**	±	0.03	**0.16**	±	0.00	**0.61**	±	0.03	**0.76**	±	0.04	1.79

57	**0.83**	±	0.08	**1.03**	±	0.05	**0.28**	±	0.02	**1.86**	±	0.13	**2.14**	±	0.15	6.91
58	**0.55**	±	0.10	**0.47**	±	0.02	**0.14**	±	0.01	**1.01**	±	0.11	**1.15**	±	0.12	8.8
59	**NA**															
60	**0.81**	±	0.16	**0.71**	±	0.00	**0.36**	±	0.02	**1.52**	±	0.16	**1.88**	±	0.18	NA
61	**1.14**	±	0.01	**0.65**	±	0.03	**0.38**	±	0.01	**1.78**	±	0.04	**2.17**	±	0.05	4.37
62	**2.54**	±	0.26	**1.78**	±	0.14	**0.68**	±	0.10	**4.32**	±	0.41	**5.00**	±	0.50	5.95
63	**1.30**	±	0.03	**1.09**	±	0.09	**0.35**	±		**2.40**	±	0.11	**2.75**	±	0.11	5.74
Nombre	57.00			57.00			57.00			57.00			57.00			57.00
Moyenne	0.70			0.76			0.12			1.45			1.57			5.60
Ecart-type	0.78			0.80			0.13			1.47			1.50			3.85
Médiane	0.46			0.56			0.08			1.04			1.18			4.86
Max	4.25			4.97			0.68			9.22			9.22			22.77
Min	0.00			0.00			0.00			0.00			0.00			1.39

Annexe 3 : Classement des modèles de GPER générés par Modeler®

Conformère	Energie	RMSD	Conformère	Energie	RMSD
GPER_active.B99990017.pdb	1683.4179	1.2221	**GPER_inactive.B99990008.pdb**	1426.9666	1.0835
GPER_active.B99990019.pdb	1811.1548	1.2226	**GPER_inactive.B99990019.pdb**	1607.3263	1.1391
GPER_active.B99990002.pdb	1912.5820	1.2687	**GPER_inactive.B99990009.pdb**	1447.049	1.1814
GPER_active.B99990005.pdb	1870.2982	1.2702	**GPER_inactive.B99990013.pdb**	1520.9132	1.1921
GPER_active.B99990001.pdb	1823.6481	1.2831	**GPER_inactive.B99990001.pdb**	1420.8795	1.1924
GPER_active.B99990015.pdb	1898.7347	1.2872	**GPER_inactive.B99990014.pdb**	1617.1383	1.2053
GPER_active.B99990011.pdb	1862.1276	1.3210	**GPER_inactive.B99990012.pdb**	1592.915	1.2067
GPER_active.B99990016.pdb	1848.7568	1.3315	**GPER_inactive.B99990015.pdb**	1433.61	1.2121
GPER_active.B99990010.pdb	1976.2648	1.3346	**GPER_inactive.B99990016.pdb**	1698.0452	1.2307
GPER_active.B99990013.pdb	2028.7507	1.3387	**GPER_inactive.B99990010.pdb**	1615.5397	1.2339
GPER_active.B99990009.pdb	1944.1920	1.3539	**GPER_inactive.B99990011.pdb**	1688.5306	1.2443
GPER_active.B99990007.pdb	1913.2238	1.3593	**GPER_inactive.B99990007.pdb**	1724.2626	1.2811
GPER_active.B99990004.pdb	1999.9658	1.3883	**GPER_inactive.B99990002.pdb**	1642.1397	1.289
GPER_active.B99990008.pdb	1699.3928	1.4024	**GPER_inactive.B99990020.pdb**	1678.0125	1.2916
GPER_active.B99990018.pdb	1961.7670	1.4060	**GPER_inactive.B99990003.pdb**	1476.5647	1.2928
GPER_active.B99990014.pdb	1959.5353	1.4068	**GPER_inactive.B99990018.pdb**	1610.7395	1.2937
GPER_active.B99990006.pdb	1836.3374	1.4104	**GPER_inactive.B99990017.pdb**	1509.6608	1.3285
GPER_active.B99990012.pdb	1864.8197	1.4382	**GPER_inactive.B99990004.pdb**	1566.6992	1.3301
GPER_active.B99990003.pdb	1917.8219	1.4636	**GPER_inactive.B99990005.pdb**	1641.7177	1.3855
GPER_active.B99990020.pdb	1895.2064	1.6900	**GPER_inactive.B99990006.pdb**	1792.6381	1.4391

Annexe 4 : Résultats de l'analyse ROC du *training set*

Récepteur	Surf	dc-pair	dc-pair-surf	Dp-pair	Dp-pair-surf	vina	xscore_av	xscore_hms	xscore_hps	xscore_hss	Somme
active1	0.747	0.758	0.737	0.717	0.747	0.616	0.758	0.828	0.783	0.758	6.702
active2	0.576	0.465	0.465	0.515	0.535	0.566	0.566	0.545	0.611	0.566	4.834
active3	0.737	0.808	0.808	0.818	0.808	0.51	0.859	0.879	0.788	0.687	6.965
active4	0.818	0.838	0.859	0.869	0.859	0.712	0.828	0.808	0.869	0.808	7.45
active5	0.657	0.848	0.828	0.657	0.657	0.586	0.697	0.788	0.646	0.677	6.384
active6	0.758	0.828	0.828	0.778	0.788	0.566	0.808	0.848	0.727	0.682	6.853
active7	0.778	0.889	0.889	0.768	0.778	0.737	0.869	0.828	0.869	0.707	7.334
active8	0.343	0.626	0.556	0.586	0.586	0.323	0.414	0.434	0.424	0.384	4.333
active9	0.475	0.455	0.465	0.505	0.495	0.475	0.465	0.485	0.586	0.556	4.487
active10	0.788	0.778	0.747	0.859	0.828	0.439	0.727	0.818	0.692	0.51	6.398
active11	0.848	0.899	0.899	0.778	0.818	0.535	0.869	0.848	0.859	0.788	7.293
active12	0.707	0.889	0.879	0.768	0.778	0.566	0.778	0.788	0.727	0.697	6.87
active13	0.677	0.727	0.727	0.677	0.707	0.571	0.732	0.798	0.697	0.586	6.222
active14	0.859	0.697	0.717	0.747	0.747	0.419	0.737	0.788	0.778	0.677	6.307
active15	0.919	0.808	0.828	0.758	0.758	0.51	0.909	0.854	0.889	0.747	7.061
active16	0.616	0.454	0.434	0.566	0.556	0.338	0.51	0.505	0.556	0.465	4.384
active17	0.747	0.828	0.798	0.727	0.717	0.667	0.808	0.768	0.758	0.636	6.707
active18	0.778	0.848	0.788	0.808	0.778	0.606	0.788	0.818	0.798	0.788	7.02
active19	0.566	0.859	0.869	0.788	0.798	0.566	0.859	0.828	0.818	0.758	7.143
active20	0.788	0.879	0.889	0.848	0.859	0.707	0.838	0.788	0.843	0.869	7.52
Somme	14.182	15.181	15.01	14.537	14.597	11.015	14.819	15.044	14.718	13.346	128.267
inactive1	0.889	0.818	0.808	0.838	0.848	0.551	0.889	0.818	0.848	0.798	7.216
inactive2	0.939	0.818	0.848	0.788	0.818	0.611	0.828	0.864	0.758	0.773	7.106
inactive3	0.768	0.848	0.859	0.768	0.758	0.535	0.828	0.838	0.879	0.747	7.06
inactive4	0.758	0.758	0.747	0.737	0.727	0.586	0.843	0.798	0.869	0.72	6.785
inactive5	0.778	0.717	0.758	0.697	0.707	0.535	0.798	0.818	0.798	0.677	6.505
inactive6	0.737	0.768	0.778	0.737	0.737	0.621	0.717	0.788	0.763	0.576	6.485
inactive7	0.859	0.859	0.818	0.848	0.828	0.52	0.758	0.747	0.773	0.727	6.878
inactive8	0.778	0.828	0.828	0.798	0.788	0.662	0.808	0.808	0.859	0.798	7.177

Récepteur	Surf	dc-pair	dc-pair-surf	Dp-pair	Dp-pair-surf	vina	xscore_av	xscore_hms	xscore_hps	xscore_hss	Somme
inactive9	0.677	0.717	0.758	0.727	0.768	0.838	0.838	0.798	0.808	0.707	6.959
inactive10	0.879	0.869	0.919	0.778	0.788	0.485	0.869	0.869	0.854	0.864	7.295
inactive11	0.697	0.838	0.808	0.859	0.859	0.737	0.758	0.758	0.768	0.727	7.112
inactive12	0.848	0.889	0.919	0.889	0.879	0.611	0.828	0.798	0.768	0.808	7.389
inactive13	0.848	0.848	0.859	0.828	0.838	0.566	0.869	0.838	0.768	0.818	7.232
inactive14	0.848	0.859	0.869	0.838	0.838	0.768	0.788	0.768	0.758	0.727	7.213
inactive15	0.889	0.798	0.838	0.727	0.727	0.621	0.798	0.828	0.778	0.727	6.842
inactive16	0.848	0.939	0.929	0.869	0.889	0.566	0.899	0.838	0.899	0.823	7.651
inactive17	0.707	0.838	0.838	0.818	0.808	0.591	0.818	0.874	0.798	0.707	7.09
inactive18	0.798	0.798	0.838	0.838	0.838	0.606	0.828	0.798	0.848	0.778	7.17
inactive19	0.636	0.717	0.737	0.717	0.707	0.535	0.707	0.808	0.717	0.606	6.251
inactive20	0.758	0.788	0.798	0.778	0.778	0.515	0.753	0.848	0.788	0.677	5.935
Somme	15.939	16.312	16.554	15.877	15.928	12.06	16.222	16.302	16.099	14.785	139.351

Annexe 5 : Résultats des ACP_*Docking_training_set*

Cinq meilleures fonctions de score de la composante principale 1

Isoforme	xscore_hps	xscore_hms	xscore_av	dp_PAIR	SURF	dp_PAIRSURF	dc_PAIR	dc_PAIRSURF
1_actif	0.9774293		0.9696185			0.9650741	0.9834981	0.9841396
2_actif	0.9357039		0.9429635	0.9547628		0.9502116		0.9215943
3_actif				0.9549648	0.969506	0.9569419	0.949874	0.948728
4_actif	0.9707869		0.9564876		0.9778452		0.9818277	0.97916
5_actif				0.9554297	0.9414961	0.959152	0.9462554	0.9451094
6_actif			0.9602373	0.9591257		0.9642249	0.9540807	0.9658008
7_actif	0.9763067		0.9600865		0.9551254		0.9830993	0.9858973
8_actif	0.9773446		0.9637647	0.9468324	0.9585725	0.9524886		
9_actif			0.9561022	0.9491518		0.9552737	0.9683336	0.9729568
10_actif				0.9348431	0.924186	0.9358796	0.9746241	0.9656938
11_actif	0.9643369		0.9657274		0.9739706		0.964402	0.9799839
12_actif			0.923008	0.945774		0.9489122	0.9661223	0.9651991
13_actif	0.9719389		0.9481561	0.9423158		0.9483108		0.9467903
14_actif	0.938235		0.9792369	0.9501925	0.9372389	0.9561193		
15_actif			0.9539587		0.9757511	0.9463828	0.9690455	0.9720455
16_actif	0.7360682			0.8040895		0.8109182	0.7782183	0.7820973
17_actif				0.9748285	0.548958	0.9762845	0.9730418	0.9771518
18_actif			0.9183002	0.958379		0.9546995	0.9360248	0.9387591
19_actif			0.9623837	0.9679607		0.9701886	0.9762082	0.9783725
20_actif	0.9754986		0.9663101			0.9568222	0.9815075	0.9866681
1_inactif	0.9441169		0.9432113			0.9439245	0.9699714	0.9713439
2_inactif				0.9556869	0.9241738	0.9609218	0.9704558	0.9821502
3_inactif	0.938045	0.9317409	0.945612				0.9535694	0.9617761
4_inactif			0.9306674	0.9526733		0.9561043	0.9698658	0.9653332
5_inactif			0.9492343	0.9446135		0.9487149	0.9720093	0.9711541
6_inactif			0.951566	0.9448674	0.9706385		0.9501531	0.9706228
7_inactif	0.9419956		0.9318404	0.9556956		0.9575634	0.937357	
8_inactif	0.9664391		0.9581739	0.9383645	0.9623535	0.9431353		
9_inactif			0.9220587	0.9353654		0.9375655	0.9761279	0.9772057
10_inactif			0.9542499	0.9356917	0.9621797		0.963834	0.9738643
11_inactif	0.9700374		0.9378513	0.941369			0.9676842	0.9613435
12_inactif	0.95244		0.9475887	0.9578254		0.9650933		0.9452954
13_inactif	0.9397185			0.9361219		0.9396001	0.9580846	0.9670758
14_inactif			0.9463517	0.9661254		0.9667579	0.9706774	0.9753085
15_inactif			0.9523866	0.953506		0.9573035	0.9629435	0.9557395
16_inactif			0.9546192	0.9567764		0.958477	0.9691338	0.9717367
17_inactif	0.9633825			0.9531914		0.9552719	0.9498264	0.9545173
18_inactif	0.9613467		0.9326008			0.9351636	0.9848319	0.987942
19_inactif	0.974213			0.9616639		0.9640622	0.9678962	0.971521
20_inactif			0.9602952	0.9535465		0.9565901	0.9743724	0.9779489