THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE MONTPELLIER SUPAGRO

En Génétique et génomique

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau
Portée par

Unité de recherche AGAP

Diversité des populations de maïs du Sud-Ouest de la France : étude de leur origine et de leur adaptation par une approche génomique resituée dans un contexte d'enquêtes ethnobotaniques.

Présentée par Yacine DIAW
Le 05 Février 2021

Sous la direction de Jacques DAVID, Joëlle RONFORT
Sous l'encadrement de Brigitte GOUESNARD

Devant le jury composé de

Mme Maud TENAILLON, Directrice de recherches, CNRS	Président de jury et rapporteur
M. Thierry ROBERT, Maitre de conférence, Université Sorbonne	Rapporteur
M. Diaga DIOUF, Professeur, UCAD du Sénégal	Examinateur
M. Christian LECLERC, Chercheur, CIRAD	Examinateur
M. Jacques DAVID, Professeur, Institut Agro de Montpellier	Directeur de thèse
Mme Brigitte GOUESNARD, Chargé de recherches, INRAE	Encadrante de thèse

UNIVERSITÉ DE MONTPELLIER

Montpellier SupAgro

REMERCIEMENTS :

Au temps où j'étais encore une petite fille, mes parents m'ont demandé ce que je compte faire comme profession dans le futur. Ma réponse a été : « Médecin militaire ». En ce temps, je n'étais qu'une petite fille et je m'étais tellement investie auprès des personnes exerçant des métiers dans le domaine de la médecine et des forces de l'ordre pour savoir comment m'y prendre. D'ailleurs, on me le rappelle encore en me disant « quand est-ce que tu seras colonel DIAW, médecin de l'armée ?! ». Aujourd'hui, je suis fière me présenter en tant que « docteur DIAW » non pas parce que je suis médecin, mais parce que j'ai obtenu mon diplôme de doctorat de Montpellier Sup Agro en génétique et génomique.

Je commencerai tout d'abord par remercier ma famille de m'avoir soutenu, accompagné et aidé tout au long de mon cursus scolaire. Un merci particulier à mon mari Mamadou Lamine DIATTA qui s'est armé de patience, qui a partagé cette aventure avec moi, m'a octroyé tout le soutien nécessaire et a supporté mes mauvaises humeurs durant ma formation. Un merci particulier aussi à mes parents Amadou Lamine DIAW et Mariéme SIDIBE qui ont bien pris soin de ma fille pour que je puisse atteindre mon objectif, à savoir réussir ma thèse.

Ma thèse est pour moi une des plus belles aventures que j'ai vécues durant ces 4 dernières années avec la contribution de nombreuses personnes chercheurs, ingénieurs, techniciens, doctorant et amis, à qui je compte montrer ma gratitude de n'avoir ménager aucun effort pour ma réussite.

Je tiens donc à remercier les personnes qui m'ont permis de réaliser ce projet, mes encadrants Brigitte Gouesnard, Joëlle Ronfort et Jacques David. Nous avons déjà eu l'occasion de travailler ensemble lors de mon stage de Master 2, durant lequel vous m'avez appris comment exploiter les bases de données génétiques, identifier les outils et interpréter les résultats d'analyse. Vous avez augmenté la confiance en moi lorsque vous m'avez annoncé le désir de me soutenir pour entamer mes activités de recherches qui aboutiront à l'obtention de mon diplôme de doctorat. A Brigitte, j'ai essayé de te suivre pour être à la fois rigoureuse, respectueuse de l'ordre et sage ; mais tu es imbattable sur ce domaine. Peut-être que s'est dû au regard que je te porte. En effet, j'ai l'impression d'être avec ma marraine ; tellement je suis à l'aise avec toi. A Jacques, je te voyais uniquement en tant que professeur mais je me suis rendu compte que tu es aussi un éducateur et psychologue. Tes capacités à pousser la réflexion au-delà de l'évidence, à rester positive et à baisser les tensions ne cesseront jamais de m'épater. Joëlle, bien que tu partages ces 3 aptitudes avec Jacques, tu en plus le contrôle sur le temps, une maitrise des techniques de

communication et une bonne organisation. Vous trois êtes la combinaison parfaite qu'il me fallait pour trouver le courage de persévérer et finir en beauté cette belle aventure.

Apprendre la recherche par la recherche n'est pas une chose simple. La réussir demande beaucoup d'énergie, de concentration et surtout de motivation. Eh oui, je ne compte plus le nombre de fois où j'étais découragé. Heureusement que j'ai eu la chance d'évoluer au sein de l'équipe Génomique évolutive et gestion des populations (GE²pop). J'ai vraiment la chance d'être à coté de personnes qui ont le sens de l'écoute, de l'entre-aide et du partage de connaissances. Ce qui est impressionnant avec cette équipe est que je n'ai nullement besoin d'attendre la programmation d'une présentation orale pour parler de mes résultats et recevoir des orientations sur les activités de recherches à venir. En effet, les débats sur les sujets de recherches ont toujours été les bienvenus au moment des pauses cafés. Ils se poursuivent par l'accompagnement de tes interlocuteurs qui t'orientent vers les personnes ressources et te suivent jusqu'à l'obtention du résultat final escompté. Je me rappelle encore les moments de convivialité en salle cafeteria durant laquelle j'oubliai mon stress lié au travail. Merci à l'ensemble des collaborateurs de l'équipe GE^{2}pop pour leur accueil chaleureux et leurs contributions permettant de mener à bien cette étude.

Pour m'avoir aidé à recadrer mon sujet de recherche, à tenir le cap et à garder une organisation efficace sur le long terme, je remercie les membres de mon comité de thèse : Yves Vigouroux, Isabelle Goldringer, Stéphane D. Nicolas, Xavier Perrier et Daniel Fonceka. A travers eux, je remercie tous les professeurs, chercheurs et agents administratifs de l'UMR Génétique Quantitative et Évolution -Le Moulon, de l'UMR AGAP (Amélioration Génétique et Adaptation des Plantes méditerranéennes et tropicales) et de l'UMR DiADE qui ont beaucoup contribuer à la réussite de ce projet.

Je compte aussi remercier l'Institut Sénégalaise de Recherches Agricoles (ISRA) du Sénégal et en particulier le Centre d'Etude Régional pour l'Amélioration de l'Adaptation à la Sécheresse (CERAAS/ISRA) d'avoir financé ma formation avec le projet WAAPP. En parlant de l'ISRA, je n'oublierai pas de remercier Dr Ndiaga CISSE, Ex-Directeur du CERAAS, qui en plus de ces responsabilités m'a encadré, conseillé, orienté de ma formation en Master 2 jusqu'à l'obtention du financement de ma thèse ; et continue de me suivre malgré son départ à la retraite. A travers lui, je remercie aussi tous les agents de l'ISRA du Sénégal.

Je finirai par remercier tous mes amis pour m'avoir facilité le séjour en France, partagé mes peines et soutenu tout au long de mon séjour.

Enfin, je remercie tous les doctorants, surtout ceux du bâtiment ARCAD avec qui j'ai partagé mon bureau : Germain, Pauline, Margaux, Josselin, Céline. Merci pour vos différentes des scripts d'analyse, vos aides à la préparation des réunions avec mes encadrants, les présentations lors des colloques et pour l'intégration. Courage à ceux qui sont encore là en attendant de trinquer avec vous lors de votre soutenance !

Un grand merci à vous tous pour votre grande disponibilité, vos sens de la pédagogie et vos conseils avisés.

Table des matières

1. INTRODUCTION GENERALE 1

1.1. La biodiversité et les ressources génétiques 2

1.1.1. Les ressources génétiques et leurs importances 2

1.1.2. La conservation et la gestion des ressources phytogénétiques 3

1.1.3. Des centres de ressources biologiques pour préserver les ressources génétiques végétales 4

1.2. Le cas du maïs 5

1.2.1. L'utilisation du maïs 5

1.2.2. La collection européenne 6

1.2.3. Les populations locales et leur gestion par les agriculteurs 7

1.3. Apports des marqueurs moléculaires sur les études de la diversité : application au maïs 9

1.3.1. Analyse de l'ADN nucléaire : passage des RFLP à la puce 50K SNP 9

1.3.2. Utilisations des microsatellites chloroplastiques (cpSSR) chez les plantes 10

1.4. Domestication et diffusion de la culture du maïs dans le monde 13

1.5. Objectifs de la thèse 16

2. CHAPITRE I : 20

Genetic diversity of maize landraces from the South-West of France 20

2.1. Présentation générale 21

2.1.1. Contexte 21

2.1.2. Méthodologie 22

2.1.3. Résultats 23

2.1.4. Discussion 24

2.2. Article 1 25

SUPPLEMENTARY FIGURES 47

CHAPITRE II 52

Diversité cytoplasmique des populations de maïs du Sud-Ouest de la France. 52

3.1. Introduction 53

3.2. Matériels et méthodes 55

3.2.1. Matériel génétique 55

3.2.2. Amplification par PCR 55

3.2.3. Analyse de la diversité des cpSSR 57

3.2.4. Relations génétiques entre les haplotypes 57

3.2.5. Analyse cpSSR basée sur les groupes génétiques de maïs 58

3.3. Résultats 59

3.3.1. Variabilité des marqueurs SSR chloroplastiques 59
3.3.2. Le réseau d'haplotypes 61
3.3.3. Diversité chloroplastique des groupes génétiques représentatifs de l'histoire du maïs cultivé 63
3.3.4. Différentiation génétique entre les groupes génétiques basée sur l'analyse cpSSR 66
3.4. Discussion 68
3.4.1. Diversité chloroplastique et histoire ancienne du maïs 68
3.4.2. Diversité chloroplastique et origine des populations du Sud-Ouest de la France 70
3.4.3. Critique sur la méthode d'analyse 71
3.5. Conclusion 72
4. CHAPITRE III 74
Enquête auprès des anciens agriculteurs et pratiques de la gestion des semences des populations de maïs du Sud-Ouest de la France 74
4.1. Introduction 75
4.2. Etude ethnobotanique et application au maïs 76
4.3. Les enquêtes dans les Pyrénées 77
4.3.1. Description de l'enquête 78
4.3.2. Le test de sélection des épis effectué lors de l'enquête 81
4.4. Dispositif expérimental de production d'épi et de mesure des caractères agro morphologique des variétés 81
4.5. Ce que l'enquête nous a appris sur les pratiques de culture du maïs dans le Sud-Ouest de la France jusque dans les années 1980 : 86
4.6. Mesure des pratiques des anciens agriculteurs des Pyrénées sur l'évolution des populations du SWF 88
4.6.1. Présentation générale 88
4.6.2. Article 2: Ethnobotanical and genetic investigations of old maize landraces from the Pyrenees: seed management practices and expected response to selection. 91
MAIN FIGURES 114
MAIN TABLES 121
SUPPLEMENTARY INFORMATION 125
5. DISCUSSION GENERALE ET PERSPECTIVES 130
5.1. Rappel sur les questions de recherche et les objectifs de la thèse 131
5.2. Origine des populations de maïs du Sud-Ouest de la France 132
5.3. Analyse critique des approches sur « bulk » d'ADN et à partir d'une plante par population dans les études de diversité 135
5.4. Conclusion et perspectives 137
BIBLIOGRAPHIE 140
Bibliographie 141

ANNEXES 150
Annexe 1 : tableaux supplémentaires du chapitre 1. 151
Annexe2 : tableaux supplémentaires de l'analyse cpSSR du chapitre 2 167
Annexe 3: tableaux supplémentaires chapitre 3. 181

INTRODUCTION GENERALE

1.1. La biodiversité et les ressources génétiques

1.1.1. Les ressources génétiques et leurs importances

Une ressource génétique est, selon la Convention sur la diversité biologique (CDB, https://www.cbd.int/), un matériel génétique ayant une valeur effective ou potentielle. Elle est liée à la diversité génétique et fonctionnelle des patrimoines génétiques des sous-populations, populations et métapopulations de chaque taxon. Fenzi and Bonneuil (2016) montrent trois grandes périodes qui ont conduit aux changements dans les représentations et la gouvernance des ressources génétiques. La première période comprise entre les années 1920 à 1960 coïncide avec l'émergence du concept de ressources génétiques. La notion de ressources génétiques n'était pas couramment utilisée par les scientifiques à cette période. Cependant, les premières études effectuées par le scientifique russe Nicolaï Vavilov sur les collections d'espèces animales et végétales ont attiré l'attention des hommes sur l'importance de la gestion et de la conservation de la diversité des formes cultivées et sauvages des plantes cultivées. En effet, Vavilov a développé une compréhension de la répartition géographique et de la diversité des formes cultivées localement et a inventé le concept de "centre d'origine", défini comme une zone où l'on trouve une grande diversité de formes cultivées d'une espèce (Bonneuil, 2019; Fenzi and Bonneuil, 2016). Plus tard, entre 1960 et 1992, les risques liés à la faible diversité génétique utilisée dans les exploitations agricoles ont catalysé les mouvements politiques et scientifiques mondiaux en faveur de l'émergence du terme de ressources génétiques et de leur nécessaire conservation pour faire face aux enjeux d'amélioration et de réponse à des catastrophes telle que l'émergence de nouvelles races pour des maladies importantes voire de nouvelles maladies dans le futur. Le remplacement des variétés locales qui constituaient le véritable socle de diversité des espèces par des variétés performantes certes mais de plus en plus homogènes et apparentées sur de très vastes surfaces a amené à une perte de diversité génétique dans de nombreuses plantes cultivées. En réaction, l'identification de cette érosion a entraîné des efforts de collectes des ressources phytogénétiques (RPG), dans le but de soutenir les sélectionneurs avec du matériel génétique pour étendre la variabilité génétique, comme base pour créer de nouvelles variétés. La dernière période, c'est-à-dire de 1992 à nos jours, voit l'évolution du terme « ressources génétiques » vers des termes de biodiversité cultivée ou agro-diversité. On est passé d'une problématique d'amélioration des plantes pour entrer dans un cadrage plus écologique de la question avec des apports de l'écologie des paysages, biologie de la conservation nécessitant des savoirs en anthropologie, gestion dynamique, génétique évolutive et des populations (Bonneuil and Fenzi, 2011). La génétique des populations est

l'étude de la distribution et des changements de la fréquence des versions d'un gène (allèles) dans les populations d'êtres vivants, sous l'influence des forces évolutives (sélection naturelle, dérive génétique, mutation, et migration). Les changements de fréquences des allèles sont un aspect majeur de l'évolution. La fixation de certains allèles conduit à une modification génétique de la population, et l'accumulation de tels changements dans différentes populations peut conduire au processus de spéciation. C'est une discipline qui est née de la synthèse des théories de Mendel, de Darwin et des biométriciens du début du XXème siècle (Provine 1971, cité par Vitalis (2001)).

Lors du Sommet de la Terre à Rio de juin 1992, l'organisation des Nations Unies (ONU, 1992) a reconnu l'importance de ces ressources comme bien commun et comme fondamentale pour le développement durable. Le rôle crucial des RPG pour la sécurité alimentaire et le développement économique au niveau mondial (FAO, 2020a) afin de faire face aux changements environnementaux ou socio-économiques est aujourd'hui largement reconnu. La diversité phytogénétique peut aussi offrir des caractères susceptibles d'aider à relever des défis futurs, tels que la nécessité d'adapter les cultures à l'évolution des conditions climatiques ou à l'apparition de foyers de maladies. Le traité international sur les ressources phytogénétiques (RPG) pour l'alimentation et l'agriculture (TIRPAA), entré en vigueur en 2004, assure la conservation de la biodiversité, les échanges et l'utilisation durable des RPG en garantissant le partage des avantages (FAO, 2020b).

1.1.2. La conservation et la gestion des ressources phytogénétiques

Comme indiqué dans le dossier du GNIS (2020) sur les RPG, la gestion des ressources génétiques consiste à les inventorier, caractériser et évaluer, conserver, régénérer et diffuser. Pour ce faire, deux stratégies globales de gestion des ressources génétiques sont utilisées : la conservation *in situ* et la conservation *ex situ*.

Il y a une différence fondamentale évidente entre ces deux stratégies. La conservation *ex situ* implique l'échantillonnage, le transfert et le stockage d'une population d'une certaine espèce loin de l'emplacement d'origine où il a été trouvé (c'est-à-dire en dehors de son habitat naturel) ; alors que dans la conservation *in situ* (c'est-à-dire la conservation de la diversité dans son habitat) implique la désignation, la gestion et le suivi de la population à l'endroit où il se trouve actuellement et au sein de la communauté à laquelle il appartient (Maxted et al., 1997). La conservation *in situ* a été proposée comme étant une meilleure conservation car, contrairement à la conservation *ex situ,* elle permet la préservation des processus évolutifs.

La conservation *in situ* comprend la conservation des populations sauvages naturelles dans des zones définies et la conservation à la ferme. Cette dernière est la gestion durable de la diversité génétique de variétés de cultures traditionnelles développées localement ou des formes sauvages et adventices associées par les agriculteurs au sein de l'agriculture traditionnelle, de l'horticulture ou systèmes de culture agro-sylvicole.

La gestion *ex situ* est une alternative qui permet de prévoir la restauration des biodiversités et le rétablissement des écosystèmes dans le futur. Il implique l'échantillonnage, le transfert et le stockage d'une population d'une certaine espèce loin de l'emplacement d'origine où il a été trouvé. Les principales infrastructures de stockage de cette stratégie de conservation sont les banques de gènes. Les différentes méthodes de conservations *ex situ* décrites par Negri et al. (2009) sont :

- Le stockage des semences graines issues des collections d'espèces végétales ;
- La conservation par culture *in-vitro* avec la maintenance des explants de tissus dans un environnement stérile et exempt d'agents pathogènes ;
- La banque de gènes en champ avec le transfert d'une collection dans un autre site que son milieu d'origine ;
- Les jardins et conservatoires botaniques ;
- Le stockage des ADN et des pollens en condition réfrigérée.

1.1.3. Des centres de ressources biologiques pour préserver les ressources génétiques végétales

Dans les années 60, l'intérêt premier des ressources génétiques est qu'elles sont des réservoirs de diversités indispensables pour créer de nouveaux cultivars. Cet intérêt des ressources génétiques pour l'amélioration des plantes a conduit à la création de centres internationaux de collections nationales publiques (ou associant privé et publique) où les ressources sont conservées *ex situ*. Ces centres internationaux ont vu le jour dans les années 60-70 sous l'égide du CGIAR (Groupe consultatif pour la recherche agricole internationale) officiellement établi en 1971 (Figure 1). Il existe de nombreuses banques de gènes dans le monde. Le « Svalbard Global Seed Vault », situé en Norvège, est le plus grand réservoir du monde où les RPG sont conservé *ex situ* à -18 degrés dans des chambres fortes en vue de restaurer les espèces menacées d'extinction dans le futur. Contrairement à l'objectif de préservation attribué au « Svalbard Global Seed Vault », les autres banques de gènes ont pour objectif de préserver les RPG et de rendre leurs collections accessibles au monde scientifique pour répondre aux besoin actuels de

recherche et de développement. C'est le cas du CIMMYT (Centre international pour l'amélioration du blé et du maïs) basé à Mexico ou encore l'ICARDA (Centre international pour la recherche agronomique dans les zones arides) qui a débuté ses recherches sur l'orge.

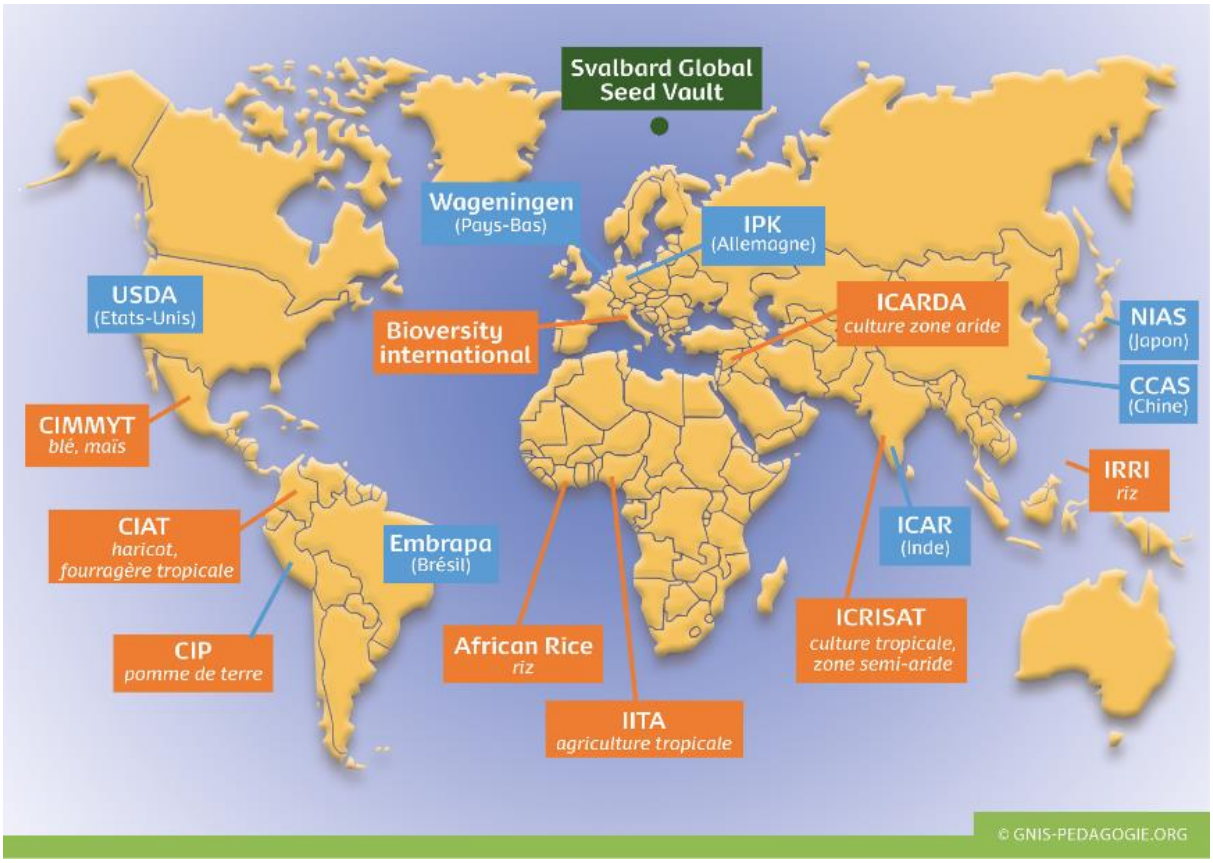


Figure 1 : Carte des principales banques de gènes ex situ dans le monde. Source : https://www.gnis-pedagogie.org/sujet/gestion-des-ressources-genetiques/

1.2. Le cas du maïs

1.2.1. L'utilisation du maïs

Le maïs est actuellement la céréale la plus productive dans le monde (http://www.fao.org/home/en/). Il s'est diffusé à travers le monde et s'est adapté aux différentes conditions pédoclimatiques (Bouchet et al., 2013; Camus-Kulandaivelu et al., 2006; Tenaillon and Charcosset, 2011). Des milliers de variétés sont dénombrées faisant du maïs la première céréale la plus produite aujourd'hui dans le monde (http://www.fao.org/home/en/) et la deuxième céréale la plus produite en France (https://www.franceagrimer.fr). La culture de maïs connait de nombreux types d'utilisations qui permettent de satisfaire des besoins d'ordre alimentaire et industriel comme détaillé dans la fiche sur le maïs du GNIS (Figure 2). Cette diversité d'utilisations conduit aux besoins de diversifier les variétés cultivées en vue de

répondre aux attentes de productions d'où l'intérêt de connaitre la diversité de l'espèce. Cependant, la production mondiale de maïs est toujours influencée par la demande pour l'alimentation animale qui consomme à lui seul 56% des récoltes (OCDE/FAO, 2018). Cette forte demande pour l'alimentation animale est plus importante pour les pays développés en raison de l'expansion et de l'intensification du secteur de l'élevage. Par contre, la production de maïs pour la consommation humaine augmentera plus dans les pays en voie de développement où la croissance démographique est le plus rapide. La fabrication de biocarburants a largement contribué à l'augmentation de la production de maïs ; la principale matière première qui fournit 64% de la production de bioéthanol. Cette demande devrait continuer de croître dans les pays développés et absorbera 14% de la production mondiale de maïs d'ici 2029 (OCDE/FAO, 2020).

1.2.2. La collection européenne

En Europe, l'ECPGR (https://www.ecpgr.cgiar.org/) est un programme de collaboration entre la plupart des pays européens, qui vise à assurer la conservation à long terme et à faciliter l'utilisation des RPG. Le comité consultatif EURISCO mis en place par ce programme européen dispose d'un catalogue de recherche sur les ressources génétiques qui fournit des informations sur plus de 2 millions d'accessions de plantes cultivées et leurs parents sauvages conservés *ex-situ*. Pour le maïs qui est l'objet de cette thèse, plus de 50 000 accessions (populations et lignées) ont été collectées dans 28 pays Européens (https://eurisco.ipk-gatersleben.de/apex/f?p=103:1:0:::::). Le programme ECPGR a aussi démarré la mise en place d'une base de données qui contient les données de passeport de 11865 accessions représentant les collections de maïs des pays européens de l'Ouest et de l'Est. Cette base de données est gérée par l'institut de recherche sur le maïs "Zemun Polje" (MRIZP, https://www.mrizp.rs/emdb/default-en.htm).

En France, la gestion *ex situ* est faite à travers les centres de ressources biologiques (CRB) et les réseaux de conservation des ressources génétiques qui associent la recherche publique et les sélectionneurs privés, de même que les associations d'amateurs. Les ressources phytogénétiques étudiées et préservées par l'INRAE, le CIRAD et l'IRD sont le reflet de l'histoire de l'agriculture. Elles illustrent les différentes valeurs que l'on peut attribuer à la biodiversité, aussi bien en termes de recherches ou de sélections qu'en intérêt patrimonial ou historique.

1.2.3. Les populations locales et leur gestion par les agriculteurs

Environ 27000 accessions de maïs originaires d'Amérique sont conservées *ex-situ* dans les banques de semences (GCDT, 2007; Ortiz et al., 2010). Ces accessions sont en majorité des populations locales, appelées « landraces » en anglais, dont la gestion a été confiée au CIMMYT. Zeven (1998) définit les populations locales comme suit : «une race locale autochtone est une variété avec une grande capacité à tolérer le stress biotique et abiotique, ce qui se traduit par une grande stabilité du rendement et un niveau de rendement intermédiaire dans un système agricole à faibles intrants». Cette définition repose davantage sur les critères agro-morphologique de la population que sur les événements (ou les méthodes) ayant conduit au développement de la population locale ou de son histoire (Casanas et al., 2017). Par la suite, la notion de population locale a évolué comme celle des ressources génétiques et prend de plus en plus en compte le rôle de l'humain dans la création et l'évolution des variétés locales (Negri et al., 2009; Villa et al., 2005). L'action de l'homme sur la gestion des populations conduit ainsi à un niveau de diversité importante sur le phénotype des populations cultivées qui reflète principalement les objectifs de productions des différents exploitants. Tel est le cas pour le maïs chez qui de nombreuses études ethnobotaniques ont été effectuées sur les populations du Mexique en Amérique (Bellon and Brush, 1994; Bellon and Risopoulos, 2001; Louette et al., 1997; Louette and Smale, 1998, 2000; Perales et al., 2003). Les agriculteurs ont adopté les nouvelles variétés (élite) de maïs dans leur exploitation, mais ils n'ont pas pour autant abandonné la culture de leurs populations locales. Les pratiques agricoles sur des nouvelles variétés introduites sont identiques à celles pratiquées sur les populations locales de l'exploitation (Louette and Smale, 1998, 2000). Par ailleurs, Bellon and Brush (1994) ont montré que l'isolement entre ces parcelles de culture des variétés nouvelles introduites dans le système et les populations locales n'est pas suffisante pour empêcher les flux géniques. Pourtant, lorsque les populations gérées à la ferme sont soumises à un flux de gènes important et récurrent par la pollinisation, tous les caractéristiques morphologiques de l'épi sont maintenues (Louette and Smale, 2000). Louette et al. (1997) confirment que la diversité génétique observée chez les populations de maïs du Mexique est due à l'introduction multiple des nouveaux matériels génétiques dans les systèmes de culture, et non pas par l'effet d'un isolement par la distance sur les populations cultivées.

En se référant aux études ethnobotaniques effectuées au Mexique, nous avons remarqué que les chercheurs axaient leurs analyses sur trois principaux facteurs clés qui sont considérés comme

nécessaires d'être pris en compte pour comprendre l'impact des pratiques des agriculteurs sur l'évolution des populations dans les agroécosystèmes. Ces trois facteurs sont :

- Le nombre de variétés ou de races présentes dans le secteur. Ceci dépend fortement des aspects agronomiques propres à la culture, mais aussi de l'utilisation du produit de culture. Le nombre de variétés est aussi l'unité principale de mesure de diversité chez les agriculteurs. Il permet aussi, sur la base du nom vernaculaire que lui donne l'agriculteur, d'identifier l'origine proposable de la population et les liens génétiques entre elles.
- Les échanges de semences entre les acteurs. Généralement, les agriculteurs ont tendance à échanger des semences de leurs récoltes avec des voisins du même terroir. Les échanges en dehors du territoire de culture de la variété sont importants à considérer pour l'évolution de la variété, ils peuvent avoir lieu notamment dans le cas de curiosité de l'agriculteur pour une nouveauté, ou dans le cas de renouvellement des variétés dans le cas de disparition de la variété cultivée.
- Et la gestion de la sélection des semences. Elle permet d'avoir des informations sur la manière des agriculteurs contrôlent l'identité et le maintien d'une variété, sa variabilité génétique, comment ils adaptent ou maintiennent les nouveaux types de variétés pour les besoins futurs. Notons que, aussi bien en Mésoamérique qu'en Europe, les agriculteurs sélectionnent ou sélectionnaient les semences pour la génération suivante uniquement sur épis (Louette, 2000).

Contrairement à la Méso-Amérique (centre de diversité du maïs), peu d'études ont analysé l'impact des pratiques des agriculteurs sur les populations locales de maïs en Europe. A l'heure actuelle, nous pouvons citer l'article de Bitocchi et al. (2009) qui a étudié l'introgression venant des variétés hybrides sur les populations locales encore cultivées en Italie ; et les travaux de sélection participative de l'équipe portugaise de Vaz Patto (Alves et al., 2018; Vaz Patto et al., 2008). En France, les populations locales de maïs sont rarement cultivées sauf exception par les agriculteurs refusant la modernisation (agriculteurs en agriculture biologique notamment). Un regain d'intérêt apparait pour les populations locales. Beigbeder and Carraretto (2018) ont rédigé un livre sur les savoirs et savoir-faire des anciens agriculteurs de maïs des Pyrénées. Cet ouvrage apporte ainsi des éléments sur l'identité des populations de maïs des Pyrénées, indispensables pour la conservation et l'utilisation durable des ressources génétiques.

1.3. Apports des marqueurs moléculaires sur les études de la diversité : application au maïs

Les marqueurs moléculaires sont largement utilisés pour étudier la diversité présente au sein des collections de ressources génétiques ou les populations cultivées dans les agro-systèmes traditionnels. La variabilité qu'ils révèlent ne dépend pas de l'environnement dans lequel se trouve la plante (de Vienne, 1998), et permet de faire des inférences sur l'histoire démographique des populations (changement de taille des populations, flux de gènes notamment). Parmi les marqueurs moléculaires, les RFLP (Restriction Fragment Length polymorphism ; (Gauthier et al., 2002; Reyes-Valdes et al., 2013) et les marqueurs microsatellites ou SSR (Mir et al., 2013; Orozco-Ramirez et al., 2016; Warburton et al., 2011) ont été les plus utilisés pour étudier l'histoire et la diversité du maïs.

1.3.1. Analyse de l'ADN nucléaire : passage des RFLP à la puce 50K SNP

Après une période importante d'analyses basées sur les marqueurs RFLP (Powell et al., 1996), la mise au point de marqueurs microsatellites, plus polymorphes a permis le développement de nouvelles études sur l'histoire du maïs et la définition de grands groupes génétiques (Dubreuil et al., 2006; Matsuoka et al., 2002; Reverter et al., 2014; Smith et al., 1997; Vigouroux et al., 2002). Le maïs étant allogame, ces différentes études ont généralement été acquises sur plusieurs plantes individuelles par population, nécessitant des efforts expérimentaux longs et coûteux. Pour y palier, Dubreuil et al. (1999) ont évalué l'application d'une nouvelle stratégie de génotypage dans l'étude de la diversité appelé analyse en « bulk » d'ADN. L'idée est de mélanger l'ADN de plusieurs plantes (15 plantes en général mais parfois plus) pour représenter une population, de faire l'analyse génétique sur ce mélange et d'estimer, sur cette base, la fréquence des différents allèles présents au sein de la population étudiée (Camus-Kulandaivelu et al., 2006; Rebourg et al., 2003; Reyes-Valdes et al., 2013). A plusieurs reprises, cette approche a été validée en comparant les résultats d'analyses par bulk et ceux obtenus sur des plantes individuelles (Dubreuil et al., 1999; Eschholz et al., 2008; Reif et al., 2005; Sham et al., 2002). Toutefois, l'analyse en « bulk » ne permet pas de détecter les allèles rares ; par exemple, les allèles dont la fréquence est inférieure à 0.05 en analyse RFLPs (Dubreuil et al., 1999) et inférieure à 0.2 en analyse SSRs (Reif et al., 2005) ne sont pas détectés. De plus, cette analyse se base sur des fréquences alléliques et ne permet pas d'avoir l'information sur les génotypes

des individus qui composent une population (Reyes-Valdes et al., 2013) : taux d' hétérozygotes, déséquilibres de liaison entre locus.

Plus récemment, les marqueurs SNPs (Single Nucléotide Polymorphism) révélant le polymorphisme d'un seul nucléotide génomique ont été développés. Ils sont très utilisés car ils permettent de faire des génotypages à faible coût avec des marqueurs répartis sur l'ensemble du génome (Ozerov et al., 2013). Généralement développés à partir d'informations de séquence, les SNP sont abondants et uniformément répartis dans les génomes de la plupart des espèces végétales (Yan et al., 2009b). Ils permettent des analyses de diversité à grande échelle, la cartographie de liaison de haute densité, la cartographie des locus quantitatifs (QTL) et des études d'association à l'échelle du génome ou GWAS (Wang et al., 2012; Yan et al., 2009a). Dès 2011, Ganal et al. (2011) ont développé une puce contenant 50K SNP. La technique de la puce consiste à mesurer les valeurs de deux intensités de fluorescences normalisées « x » et « y » pour chaque allèle de SNP appelé allèle A et B pour des raisons de simplicité (Arca et al., 2021; LaFramboise, 2009; Macgregor et al., 2006). Il s'agit de deux couleurs de fluorescence Cy3 et Cy5 associées chacune à la détection des deux allèles spécifiques de SNP (allèle A et allèle B). Les fréquences alléliques de chaque individu sont estimées en utilisant le ratio des intensités de fluorescences (FIR) mesuré avec le logiciel GenomeStudio. On s'attend à ce que le développement de ce type de marqueur soit associé à un biais systématique notamment si les SNP sont définis sur un panel d'individus réduit (Yan et al., 2009a). Les 50K SNP de la puce maïs ont été sélectionnés à partir des populations NAM qui ont été développées en croisant 25 lignées fondatrices de maïs avec la lignée de référence B73 (http://www.panzea.org/current-project), ce qui représente une diversité relativement importante. Une étude pilote d'estimation des fréquences alléliques de SNP de populations à partir d'un « bulk » d'ADN de 15 plantes a été réalisée par Arca et al. (2021). Ils ont par la suite appliqué cette technique sur un ensemble de 166 populations américaines et européennes ; et ils ont montré que leur structuration génétique était globalement cohérente avec ce qui est connue des populations de maïs en analyse SSR.

1.3.2. Utilisations des microsatellites chloroplastiques (cpSSR) chez les plantes

Durant la fin du 20ème siècle, l'analyse de l'ADN des mitochondries a permis l'acquisition des connaissances puissantes sur l'écologie, l'histoire et l'évolution des populations animales (Provan et al., 2001). Chez les animaux, le génome mitochondrial évolue de manière très différente. Il est caractérisé par une grande taille, des taux de substitution de nucléotides lents

et des niveaux étendus de recombinaisons intramoléculaires. A l'inverse, chez les plantes, l'ADN chloroplastique possède des caractéristiques similaires au génome mitochondrial des animaux. Ceci fait du génome chloroplastique un outil attractif pour les analyses phylogénétiques. Chez le maïs, la première étude sur le génome chloroplastique a été effectué par Maier et al. (1995). Ces auteurs ont montré que le génome du chloroplaste chez le maïs est un ADN circulaire à double brin qui se compose de 140387 paires de bases. Ce génome contient une paire de régions de répétitions inversées (IRA et IRB) avec 22 748 paires de bases chacune. Ces deux régions de répétitions inversées sont séparées par une petite et une grande région de copie simple (SSC et LSC) de 12 536 et 82 355 paires de bases, respectivement. En comparant les génomes chloroplastiques du maïs et du riz (*Oryza sativa*), ils ont montré que le contenu du gène et les positions relatives d'un total de 104 gènes du chloroplaste sont identiques chez ces deux espèces.

Par la suite, Provan et al. (1999) ont réalisé une première étude utilisant des marqueurs microsatellites définis sur le génome chloroplastique pour analyser la diversité génétique chez des accessions des espèces et sous-espèces du genre *Zea*. Ils ont trouvé un niveau de diversité génétique élevé chez les marqueurs cpSSR qui a permis de bien différentier les accessions des espèces et sous-espèces du genre *Zea*. Cette étude a aussi apporté des informations précieuses sur la diversité génétique du maïs cultivée et a permis de faire des inférences sur l'histoire du maïs sur le continent américain (Figure 3).

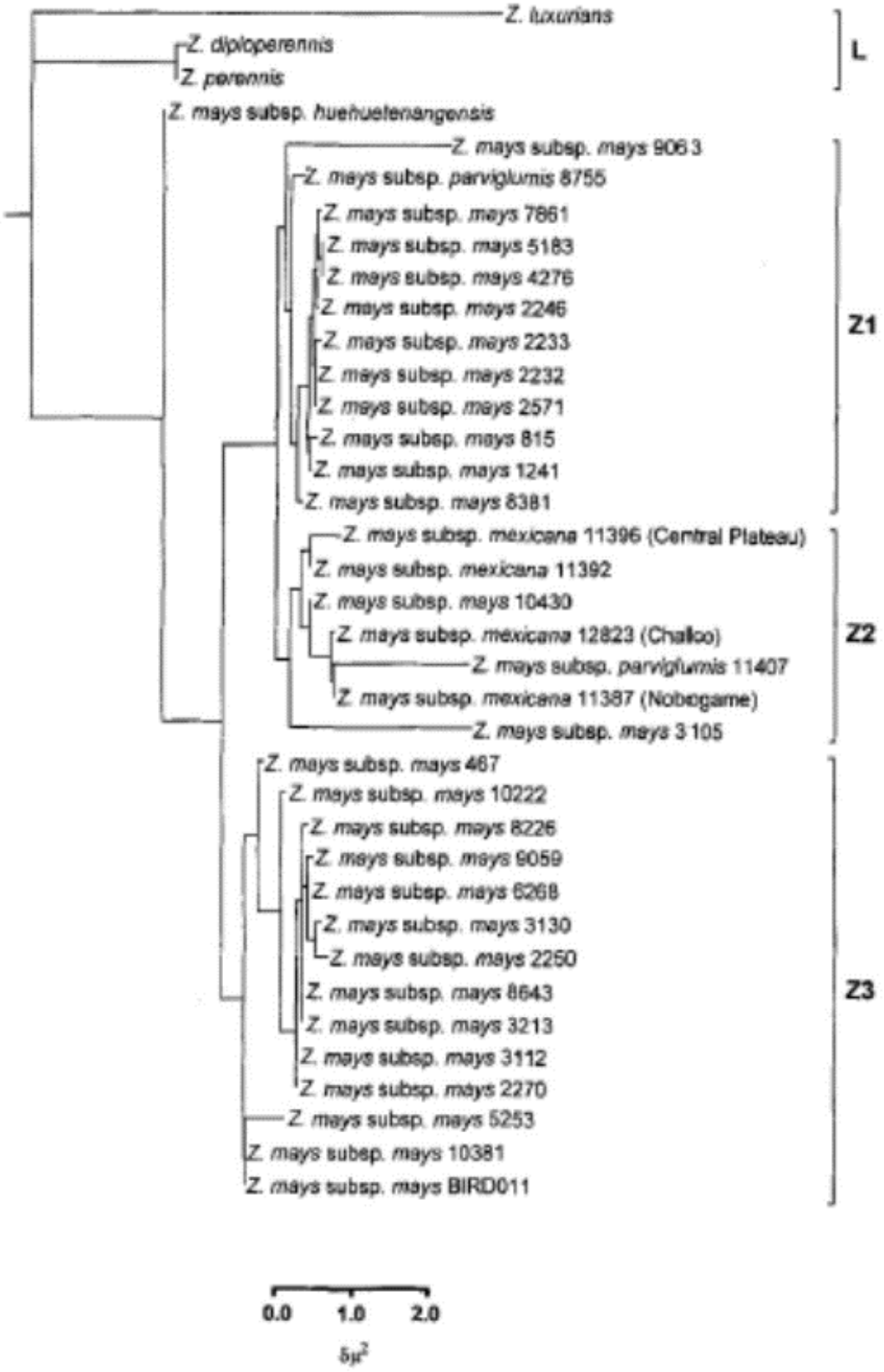


Figure 3 : arbre phylogénétique des liens génétiques entre les 37 accessions du genre Zea basé sur matrice de distance génétique de Goldstein et al. (1995) calculée à partir de 14 marqueurs cpSSR du chloroplaste de maïs. *source : Provan et al. (1999).*

1.4. Domestication et diffusion de la culture du maïs dans le monde

Le maïs est une espèce végétale qui a été domestiquée il y a environ 9000 ans dans la vallée d'Oaxaca au Mexique (Piperno and Flannery, 2001; Ranere et al., 2009), à partir d'une espèce sauvage connue sous le nom de téosinte, identifiée comme étant la sous-espèce *Zea mays ssp parviglumis* (Doebley et al., 1988; Matsuoka et al., 2002; van Heerwaarden et al., 2011). La culture du maïs a d'abord diffusé au sein du continent américain. Vers le nord de l'Amérique elle a débuté au sud-ouest du Mexique pour aller au nord du Mexique jusqu'au nord du continent. Vers le sud, elle a eu lieu du centre d'origine au Nord de l'Amérique du sud (U.S.A, Figure 4) De ces diffusions sont issues les populations des iles Caraïbes et d'autre part les populations dites andéennes venant de la Colombie. De tels résultats sont appuyés par plusieurs études (Dubreuil et al., 2006; Matsuoka et al., 2002; Tenaillon and Charcosset, 2011). Plus tard, la culture du maïs a été introduite en Europe grâce aux échanges socio-économiques qui existaient entre les différents pays des deux continents. Deux voies d'introduction indépendantes ont été identifiées en Europe. La première et la plus soutenue par les études est l'introduction de populations dite « Caribbean » par Christophe Colomb des Caraïbes vers l'Espagne en 1493. La deuxième a eu lieu vers l'an 1539 dans le nord de l'Europe et concerne l'arrivée des populations de maïs cornées appelées « Northern flint ». Depuis le Nord des USA et le Canada. L'étude de Tenaillon and Charcosset (2011) a bien résumé les relations génétiques qui existe entre les populations de maïs de l'Amérique et de l'Europe, comme décrit dans la Figure4 ci-dessous.

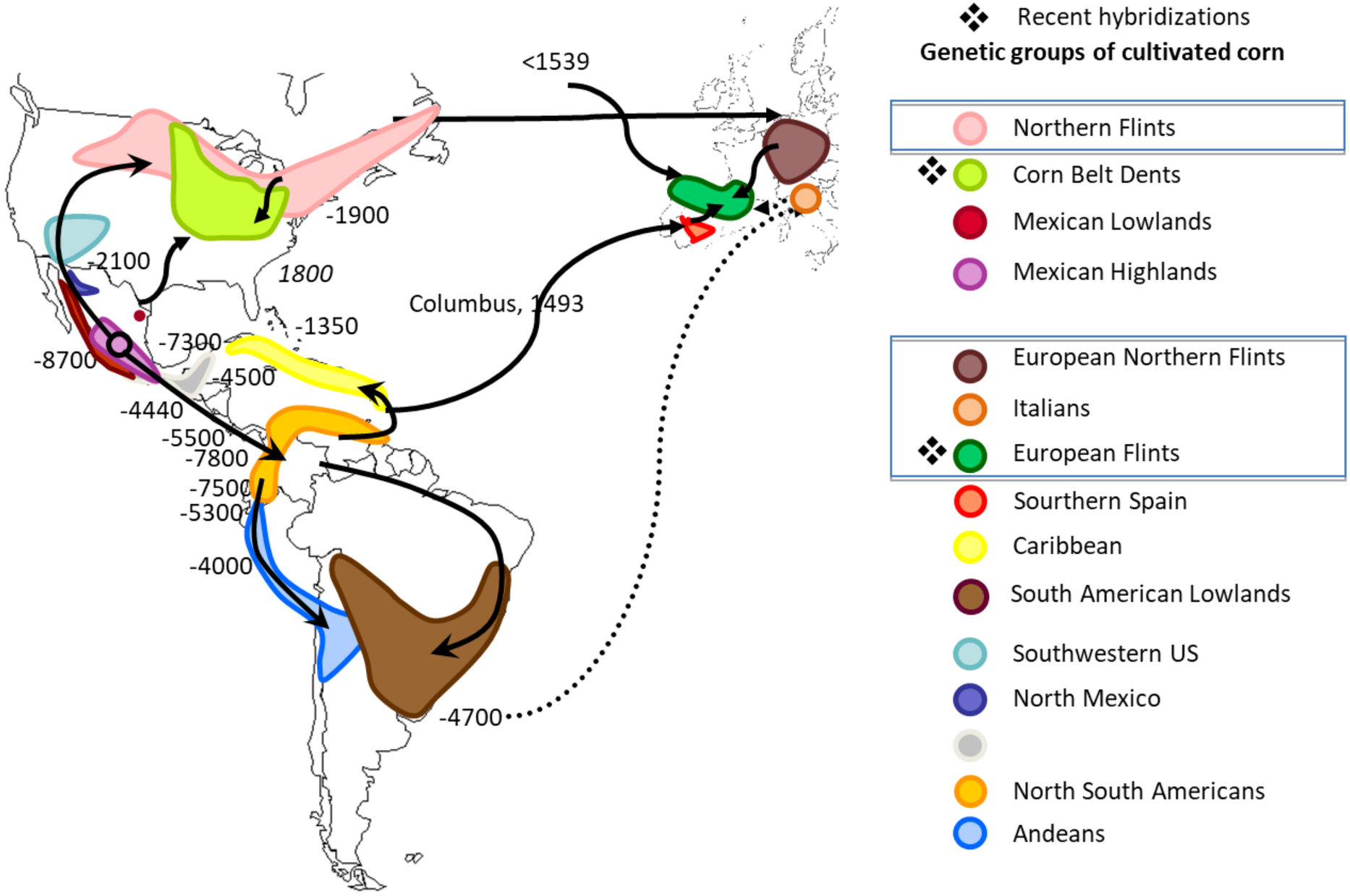


Figure 4 : Centre de domestication et hypothèses de diffusion des populations de maïs à travers l'Amérique et l'Europe. *source Tenaillon and Charcosset (2011).*

Vigouroux et al. (2008) a effectué une étude approfondie sur les populations de maïs collectées en Amérique et ils ont montré que (1) le maïs a eu le temps de se diversifier au Mexique avant sa propagation au Nord et au Sud de l'Amérique, où les populations de maïs ont évolué dans des environnement contrastés ; (2) les populations du Nord et du Sud de l'Amérique sont génétiquement différenciées et (3) des populations « Northern Flint » du Nord de l'Amérique ont été introduites par mouvements démographiques à partir de 1493 vers les régions sur la partie Sud de l'Amérique du Sud. Un autre groupe génétique nommé « Corn Belt Dent » (CBD) a été aussi identifié chez le maïs de l'Amérique. Les populatins de ce groupe CBD sont originaires, vers 1800, d'une hybridation entre des populations Northern Flint et des populatins de maïs appelées « U.S Southern Dent » ; ces dernières sont des populations de maïs provenant du Mexique nouvellement introduites au Sud-Est des U.S.A vers 1500 par les conquistadores espagnols (Doebley et al., 1988; Labate et al., 2003).

Pour l'Europe, les études morphologiques et moléculaires suggèrent que les populations cultivées dans le Nord, le Sud-Ouest et le Sud-Est de l'Europe résultent d'introductions de matériels en provenance de différentes régions du continent américain (Camus-Kulandaivelu et al., 2006; Dubreuil et al., 2006; Gauthier et al., 2002; Mir et al., 2013; Rebourg et al., 2003; Rebourg et al., 2001). Les proximités observées avec les marqueurs moléculaires entre les populations de maïs du Nord de l'Amérique et celles du Nord de l'Europe, ainsi qu'entre les populations des Caraïbes et celles du sud de l'Espagne sont en accord avec les hypothèses d''introduction du maïs en Europe, à partir de matériel de type « Caribbean » (CAR) en 1493 dans le Sud de l'Espagne (Rebourg et al., 2003; Revilla Temiño et al., 1998) et de maïs de type « Northern Flint » (NF) dans le Nord de l'Europe dans la première moitié du XVIe siècle.

Camus-Kulandaivelu et al. (2006) ont analysé, avec des marqueurs SSRs, une collection de 275 populations locales de maïs représentatives de la diversité américaine et européenne. Les données moléculaires ainsi acquises ont fait l'objet d'une approche dite de « clustering » dont le but est de grouper les individus ou les populations analysées en fonction de leur ressemblance génétique. L'étude a permis d'identifié 7 groupes génétiques: un groupe appelé « Caribbean » (CAR) qui regroupe les populations des îles Caraïbes et des populations du Sud de l'Espagne, le groupe dit « Northern Flint » (NF) qui réunit les populations du Nord de l'Amérique et du Nord de l'Europe, le groupe « Mexican » (MEX) regroupant les populations du Mexique, le groupe « Andean » (AND) qui est formé principalement par les populations du Sud de l'Amérique, le groupe « Corn Belt Dent » (CBD) dans les grandes plaines du Nord de l'Amérique, le groupe des « Italian flint » (ITA) contenant le matériel italien et le groupe appelé

« Pyrenees-Galicia » qui regroupe des populations des Pyrénées (française et espagnole) et de Galice en Espagne. Ces mêmes groupes génétiques ont été identifiés par Arca et al. (2020) en analyse SNP avec la puce 50K SNP et l'approche « bulk » d'ADN. Dans toutes ces études, le groupe « Pyrenees-Galicia » apparaît intermédiaire entre les populations de maïs localisées au Sud de l'Espagne et celles localisées au Nord de l'Europe (Brandenburg et al., 2017; Dubreuil et al., 2006). Néanmoins, Mir et al. (2013) ont observé que les populations des Pyrénées sont intermédiaires entre des populations NF et des populations localisées au Nord de l'Amérique du Sud, avec une prédominance génomique issue du Nord de l'Amérique du Sud. Toutes ces études soulignent l'importance des populations NF dans l'adaptation climatique du maïs aux conditions environnementales présentes dans les Pyrénées.

1.5. Objectifs de la thèse

Les études décrites ci-dessus étaient réalisées à très grande échelle et des études plus spécifiques sont nécessaires sur les populations de maïs européens pour préciser les relations génétiques identifiées dans les études décrites ci-dessus. Pour les populations appelées « Pyrénées-Galicia » seules 30 populations en provenance du Sud-Ouest de la France ont été utilisées.

Jusque dans les années 60 au moins, la culture du maïs dans le Sud-Ouest de la France reposait sur des variétés populations locales ou « landraces » évoluant sous l'action simultanée de l'environnement et de la gestion humaine. Entre 1950 et 1980, ces populations locales ont été collectées par l'INRA et ses partenaires puis conservées en banques de graines (Gestion *ex situ*, Centre de Ressource Biologiques INRA Montpellier).

Le Sud-Ouest de la France (SWF) est situé entre l'Espagne et l'Europe du Nord, qui, comme décrit plus haut, sont les deux portes d'introduction du maïs en Europe, à partir du groupe génétique Caraïbes (CAR) pour l'introduction en Espagne et du groupe Northern Flint (NF) pour l'introduction au Nord de l'Europe. Par conséquent, les populations SWF pourraient résulter de la diffusion d'un de ces deux types de matériels ou d'une hybridation entre les populations de maïs du groupe CAR et celles du groupe NF. Le croisement entre ces deux groupes génétiques aurait conduit à l'obtention d'un pool génétique original. En accord avec cette prédiction, Tenaillon and Charcosset (2011) ont indiqué dans leur étude que les populations de maïs cultivées dans les Pyrénées (françaises et espagnoles) et en Galice (Espagne) ne montrent pas de relations de similarité directes avec les groupes génétiques d'Amérique. Dubreuil et al. (2006) ont observé que le génome moyen d'une population cultivée dans les Pyrénées (françaises et espagnoles) ou en Galice est composée d'une fraction de l'ordre

de 17% d'information génétique issue du type « Northern Flint » européen et à 27% de populations d'Espagne. Ces différents résultats suggèrent que les populations de maïs cultivées dans le SWF pourraient résulter aussi d'une hybridation entre les NF du nord de l'Europe et les CAR de l'Espagne. Concernant les caractéristiques morphologiques, on peut noter que ce groupe « Pyrenees-Galicia » est formé par des populations qui présentent des traits morphologiques intermédiaires entre les groupes CAR et NF (Rebourg et al., 2001): un cycle précoce qu'elles ont surement hérité des populations du Nord de l'Europe (NF très précoces mais épis fins et cylindriques) ; et des épis courts et coniques qu'elles auraient hérité des populations du Sud de l'Europe (CAR tardifs mais avec des gros épis coniques).

Par ailleurs, les populations cultivées au sud de l'Europe (Espagne, Italie, Portugal, Grèce et Pyrénées) se montrent plus polymorphes que les populations cultivées à l'Est de l'Europe (Pologne, Tchécoslovaquie, Australie, Allemagne, Suisse et France-Alsace) et dans les Balkans (Eschholz et al., 2008; Gauthier et al., 2002; Rebourg et al., 2001; Revilla Temiño et al., 2003). Revilla Temiño et al. (2003) ont montré que le sud de l'Europe aurait connu plusieurs introductions de populations de maïs des divers groupes génétiques de l'Amériques. Ceci nous permet de poser une deuxième hypothèse sur l'origine des populations SWF à savoir « le maïs cultivé dans le Sud-ouest de la France est issu d'une introduction de populations issues des différents groupes génétiques connus chez le maïs » comme le cas du maïs d'Espagne (Revilla Temiño et al., 2003). Cette deuxième hypothèse pourrait expliquer que des populations des Pyrénées seraient originaires du maïs cultivé dans les régions de la partie Nord de l'Amérique du Sud, comme observé par Mir et al. (2013).

De nombreuses questions concernant la diversité, la structure génétique et l'origine de ce matériel restent donc ouvertes. Dans cette thèse nous avons abordé les questions suivantes :

- Quelle est l'origine des populations de maïs cultivées jusque dans les années 1960-1980 dans le Sud-Ouest de la France ?
- Quelle est la diversité et la structure génétique de ce matériel ? et peut-on expliquer cette structure génétique à partir de critères géographiques ou climatiques ?
- Comment les pratiques paysannes ont-elles façonné ces populations ?

Pour répondre à ces questions, nous avons étudié la diversité génétique des populations de maïs cultivées dans le Sud-Ouest de la France jusque dans les années 60 en étudiant différents descripteurs de la diversité. Ainsi : le premier chapitre (article 1 publié dans la revue Plos One) est consacré à l'analyse de la diversité génétique moléculaire des populations de maïs du Sud-Ouest de la France à partir de la puce 50K SNP (Ganal et al., 2011) et en se basant sur l'approche en bulk d'ADN décrite dans Dubreuil et al. (1999) et Arca et al. (2021). Les données

SNP sont utilisées pour étudier finement la structure génétique de ces populations et comparées à la variabilité SNP observée dans les groupes génétiques retraçant l'histoire du maïs cultivé. L'ensemble de ces analyses est utilisé pour comprendre l'origine des populations du Sud-Ouest de la France.

Le second chapitre est consacré à une étude de ces mêmes populations à l'aide de marqueurs chloroplastiques. Les chloroplastes étant transmis par voie maternelle et la propagation du maïs se faisant à longue distance via les semences, la répartition géographique actuelle des différentes variants de l'ADN chloroplastique devait nous permettre d'identifier plus précisément les voies de migration empruntées par les graines depuis la domestication.

Le chapitre 3 est consacré à l'étude des pratiques locales des agriculteurs dans les Pyrénées françaises. Pour cela nous combinerons l'analyse des caractéristiques agro-morphologiques de ce matériel avec une étude-enquête auprès d'un ensemble d'anciens agriculteurs. L'enquête vise à mieux comprendre les pratiques et critères de sélection qui étaient utilisées par les anciens agriculteurs des Pyrénées françaises.

CHAPITRE I

Genetic diversity of maize landraces from the South-West of France

Yacine Diaw [1,2], Christine Tollon-Cordet [2], Alain Charcosset [3], Stéphane Nicolas [3], Delphine Madur [3], Joëlle Ronfort [2], Jacques David [2], Brigitte Gouesnard [2*]

Authors' affiliations

[1] Institut Sénégalais de Recherches Agricoles, ISRA-CNRA de Bambey, Dakar, Sénégal

[2] AGAP, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France

[3] Génétique Quantitative et Evolution – Le Moulon, Université Paris-Saclay, INRAE, CNRS, AgroParisTech, Gif-sur-Yvette, France

*Corresponding author

E-mail: brigitte.gouesnard@inrae.fr

Article publié sous Plos One • February 1, 2021 • https://doi.org/10.1371/journal.pone.0238334

2.1. Présentation générale

L'article 1 intitulé « Genetic diversity of maize landraces from the South-West of France » a pour objectifs (1) d'évaluer la structure et la diversité génétique des populations locales de maïs qui étaient cultivées dans le sud-ouest de la France (SWF) jusqu'en 1960-1987, et (2) de déterminer les relations génétiques entre ces populations et les autres populations de maïs originaires d'Amérique et d'Europe. Le résumé de l'article 1 ci-dessous sera structuré en 4 parties décrivant respectivement le contexte de l'étude, les stratégies utilisées pour effectuer l'étude, les principaux résultats obtenus et la discussion simplifiée qui en découle. Les tableaux supplémentaires de cet article 1 sont présentés dans l'Annexe 1 de ce manuscrit.

2.1.1. Contexte

Domestiqué il y a environ 9000 ans dans la vallée centrale d'Oaxaca au Mexique (Matsuoka et al., 2002), le maïs a évolué dans des environnements contrastés, conduisant à la création de groupes génétiques distincts adaptés aux différentes conditions climatiques. Des études basées sur des données morphologiques et/ou moléculaires ont montré que les populations cultivées dans le Nord, le Sud-Ouest et le Sud-Est de l'Europe résultent de l'introduction de différentes populations de maïs originaires de différentes régions de l'Amérique (Camus-Kulandaivelu et al., 2006; Dubreuil et al., 2006; Gauthier et al., 2002; Mir et al., 2013; Rebourg et al., 2003; Rebourg et al., 2001). Ces auteurs ont observé des similitudes génétiques entre les populations du Nord de l'Amérique et du Nord de l'Europe, ainsi qu'entre les populations des iles des Caraïbes et du sud de l'Espagne. Ces résultats corroborent avec l'histoire connue sur l'introduction du maïs en Europe, avec l'introduction du maïs dite « Caribbean » (CAR) en 1493 dans le Sud de l'Espagne (Rebourg et al., 2003; Revilla Temiño et al., 1998) et l'introduction de maïs dit « Northern Flint » (NF) au Nord de l'Europe dans la première moitié du XVIe siècle. En analysant les populations américaines et européennes à l'aide de marqueurs SSR, Camus-Kulandaivelu et al. (2006) ont montré que les populations des Pyrénées (France et Espagne) et les populations de la Galice en Espagne forment un groupe de populations original qu'ils nomment « Pyrenees-Galicia ». Le fond génétique de ces populations semble être intermédiaire entre les populations de maïs localisées au Sud de l'Espagne et celles localisées au Nord de l'Europe (Brandenburg et al., 2017; Dubreuil et al., 2006). Néanmoins, Mir et al. (2013) ont soutenu que les populations des Pyrénées présentaient également une origine hybride entre des populations NF et des populations localisées au Nord de l'Amérique du Sud, avec une prédominance génomique issue du Nord de l'Amérique du Sud. Toutes ces études soulignent l'importance des populations NF dans l'adaptation climatique du maïs des

Pyrénées et de la Galice. Cependant, ils sont en désaccord sur l'origine génétique du deuxième ancêtre potentiel des populations SWF.

2.1.2. Méthodologie

Le SWF est l'une des principales zones traditionnelles de culture du maïs en Europe. D'après les archives historiques, la culture du maïs dans cette zone géographique a été signalée pour la première fois en 1626 dans la région du Béarn, et en 1628 et 1637 respectivement dans les villes de Bayonne et de Castelnaudary (Renoux, 2014). Cependant, l'origine et la diversité génétique de ces populations locales sont mal connues.

Pour mener à bien cette étude, nous avons étudié 194 populations de maïs collectées entre 1949 et 1987 (Dallard et al., 2000; INRA, 2017) dans le Sud-Ouest de la France (voir S1 Table dans Annexe 1) puis conservées au Centre de Ressources Biologiques du Maïs (CRB, Mauguio, France, https://urgi.versailles.inra.fr/siregal/siregal/grc.do). Nous avons extrait l'ADN de chacune de ces populations en utilisant la méthode des « bulk » (Dubreuil et al., 1999); puis nous avons fait leur génotypage avec la puce 50K SNP développée par Ganal et al. (2011) et nous avons prédit les fréquences des allèles des SNPs en utilisant la méthode d'estimation des fréquences alléliques développée par Arca et al. (2021).

Pour répondre aux questions sur l'origine de ces populations et pour comparer ces populations avec le matériel américain et européen, nous avons utilisé un panel de référence mondial composé de 64 populations de maïs européennes et 73 populations de maïs américaines (voir tableau S2 de l'article1) préalablement analysées à l'aide de marqueurs SSR par Camus-Kulandaivelu et al. (2006). Pour compléter cet ensemble de référence, nous avons ajouté 2 populations du Portugal et 9 populations de l'Amérique du Sud précédemment analysées par Mir et al. (2013). Ces populations utilisées comme panel de référence mondial ont été étudiées par Arca et al. (2021) avec la puce 50K SNPs en analyse « bulk » d'ADN. Ainsi, en combinant la base de données de Arca et al. (2021), la base de données sur les 194 populations du SWF, nous avons obtenu une base de données de fréquences d'allèles pour 23 412 SNP sur 342 populations de maïs. Sur cette base de données génétiques, nous avons estimé les paramètres de diversité génétique (la richesse allélique, le taux d'hétérozygotie et la distance génétique) ; analysé la structure génétique des populations SWF en utilisant le logiciel ADMIXTURE v1.07 (Alexander et al., 2009). A partir de la structuration mise en évidence, nous avons recherché les facteurs d'évolution de la diversité (analyse de gradient génétique, isolement par la distance, impact des données climatiques). Enfin, pour déterminer les relations génétiques entre les

populations SWF et les populations de diversité mondiale, nous avons effectué une analyse en composantes principales (ACP) sur les 342 populations, des test TreeMix 3-populations (Reich et al., 2009), une analyse « Neighbor joining » sur matrice de distance génétique entre 342 populations et une analyse de distances génétiques estimées uniquement entre 82 populations de l'Amérique et 194 populations du SWF.

2.1.3. Résultats

L'analyse d'ADMIXTURE avec les 194 populations SWF a permis d'identifier deux groupes génétiques qui différencient principalement les populations situées dans la partie Est du SWF des populations situées dans la partie Ouest (voir Figure2 et Figure S1 de l'article1, Table S1 dans Annexe 1). Le premier groupe, appelé « East South-West France » (E-SWF), comprend 65 populations. Tandis que le second groupe, appelé « West South-West France » (W-SWF), comprend 126 populations. L'analyse d'ADMIXTURE montre aussi que la plupart des populations collectées dans les départements Haute-Garonne et Ariège sont originaires d'une hybridation entre les groupes W-SWF et E-SWF. La PCoA sur les fréquences des allèles a révélé le même schéma de structuration chez les populations SWF. La richesse allélique (A_r dans l'article) du W-SWF est significativement plus élevée que celle du E-SWF : A_r =1,72 et A_r = 1,68 respectivement (p-valu < 10^{-15} ; voir Tableau 1 de l'article1). D'autre part, le groupe E-SWF présente un niveau de diversité génétique (H_e) plus élevé que le groupe W-SWF : H_e =0,29 et H_e =0,27 respectivement (p-value < 10^{-15}). La différenciation W-SWF/E-SWF explique une proportion faible mais significative de la diversité génétique globale : F_{ST} =0,026 (p-value < 10^{-15}). Nous avons mis en évidence l'existence de deux gradients génétiques : un longitudinal et un latitudinal. Nous avons également identifié une corrélation significative entre le rapport F_{ST} / (1- F_{ST}) (Rousset, 1997) et les distances géographiques (r=0,38, p-value < 10-15, Figure 3d de l'article1). Ce résultat marque clairement un schéma d'isolement par la distance. L'analyse des données climatiques montre des différences de températures et de précipitations mensuelles entre les sites de collecte des populations des groupes E-SWF et W-SWF (tableau S3 de l'article1), mais la régression linéaire multiple entre les matrices de données géographiques et climatiques avec la matrice de données génétiques montre une relation significative uniquement pour la matrice géographique (p-value=0,01 pour les valeurs géographiques VS p-value= 0,67 pour les valeurs climatiques), ce qui suggère que l'histoire démographique (au lieu de la sélection) de ces populations est le principal facteur de différenciation des populations SWF.

Nous avons utilisé le test TreeMix 3-population pour déterminer si les deux groupes du SWF sont originaires d'événements d'hybridation entre les groupes génétiques de maïs identifiés par Camus-Kulandaivelu et al. (2006). Les résultats de cette analyse montrent que les populations du groupe E-SWF sont originaires d'événements d'hybridations principalement entre les populations des groupes CAR et NF (c'est-à-dire ayant des valeurs f3-stat<0 associées avec des Z-score= -10). Par contre, nous n'avons pas observé de signaux d'hybridations pour le groupe W-SWF. L'analyse en composantes principales (ACP) et l'analyse « Neighbor joining » effectuée sur les données SNP des 342 populations de maïs (Fig.4 et Fig.5 de l'article1 et S2 Table dans Annexe 1) montre que la plupart des populations de l'Europe du Nord sont proches des populations de l'Amérique du Nord, et que les populations du sud de l'Espagne sont proches de celles des Caraïbes. Les populations E-SWF sont autant plus proches du groupe NF que les populations W-SWF le sont avec les populations tropicales (c'est-à-dire des groupes AND, CAR et MEX). En recherchant les relations entre les populations européennes de maïs, nous avons trouvé que les populations des groupes E-SWF et W-SWF sont plus proches des populations du groupe « Pyrenees-Galicia » que des populations du groupe ITA.

Les résultats sur l'analyse de la matrice de distance génétique, estimée entre les 82 populations de l'Amérique et les 194 populations SWF, appuient non seulement les résultats observés dans les analyses ci-dessus ; mais suggèrent aussi l'introduction de populations du groupe CBD dans le pool génétique du SWF, ainsi que des échanges avec les populations de maïs du Chili.

2.1.4. Discussion

En utilisant cette base de données génétiques de 23412 marqueurs SNP pour analyser la diversité génétique des populations de maïs collectées dans le Sud-Ouest de la France il y a environ 50 ans, nous avons pu identifier deux groupes, l'un situé à l'Est et l'autre à l'Ouest du Sud-Ouest de la France. Cependant, bien que ces deux groupes de populations aient été identifiés par des analyses différentes (ADMIXTURE, PCoA), la différenciation génétique entre eux est relativement faible (FST = 0,03) et indiquait que chaque groupe contient une variabilité génétique importante. Nous avons ainsi retenu que les populations SWF sont issues du même groupe ancestral de maïs et qu'elles se sont ensuite différenciées en 2 groupes sous l'effet de la dérive génétique et du flux de gènes, comme le montre les analyses de gradients génétiques et d'IBD. Toutefois, les différences génétiques observées chez les populations SWF pourraient également s'expliquer par l'impact des pratiques des anciens agriculteurs. La gestion des semences et les échanges entre agriculteurs dans différentes vallées limitées par des

montagnes accidentées auraient probablement contribué à différencier les populations en deux groupes distincts. Ceci a été étudié dans l'article2 présenté dans ce document (Annexe 3).

En ce qui concerne l'origine des populations SWF, nous avons trouvé une proximité génétique entre les populations SWF et les populations du groupe « Pyrenees-Galicia » localisées en Espagne et au Portugal. Ceci nous a conduit à affirmer que les populations SWF sont initialement issues d'une hybridation entre les populations des groupes CAR et NF. Notre analyse TreeMix 3-population nous a permis de confirmer que les populations NF ont joué un rôle important dans l'évolution des populations SWF. Les analyses dans cette étude ont permis aussi de montrer que les populations SWF ont connu des flux de gènes issues de l'introduction de maïs provenant de différentes régions de l'Amérique et de l'Europe, rendant ainsi difficile la détermination de leur origine génétique. Enfin, notre analyse ACP a montré une proximité génétique entre 10 populations SWF et les populations du groupe CBD ; ce dernier résultant d'hybridation entre des matériaux NF et « U.S. Southern Dent » (Doebley et al., 1988; Labate et al., 2003). Cette relation génétique peut s'expliquer par l'introduction, dès 1948 avec le plan Marshall, des premiers hybrides américains qui ont été testés en France (Renoux, 2014) et adoptées assez rapidement par les agriculteurs (Carraretto, 2005).

Tous ces résultats nous ont permis de décrire au moins deux scénarios d'origine qui tiennent compte de l'influence des populations NF du Nord de l'Europe, principalement dans la partie orientale, et de la proximité génétique des populations SWF avec les races tropicales (voir article1).

2.2. Article 1

PLOS ONE

RESEARCH ARTICLE

Genetic diversity of maize landraces from the South-West of France

Yacine Diaw[1,2], Christine Tollon-Cordet[2], Alain Charcosset[3], Stéphane D. Nicolas[3], Delphine Madur[3], Joëlle Ronfort[2], Jacques David[2], Brigitte Gouesnard[2]*

1 Institut Sénégalais de Recherches Agricoles, ISRA-CNRA de Bambey, Dakar, Sénégal, **2** AGAP, CIRAD, INRAE, Institut Agro, Univ Montpellier, Montpellier, France, **3** INRAE, CNRS, AgroParisTech, GQE—Le Moulon, Université Paris-Saclay, Gif-sur-Yvette, France

* brigitte.gouesnard@inrae.fr

Check for updates

OPEN ACCESS

Citation: Diaw Y, Tollon-Cordet C, Charcosset A, Nicolas SD, Madur D, Ronfort J, et al. (2021) Genetic diversity of maize landraces from the South-West of France. PLoS ONE 16(2): e0238334. https://doi.org/10.1371/journal.pone.0238334

Editor: Tzen-Yuh Chiang, National Cheng Kung University, TAIWAN

Received: August 11, 2020

Accepted: January 15, 2021

Published: February 1, 2021

Copyright: © 2021 Diaw et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: YD received a grant from West African Agricultural Productivity Program (WAAPP, http://www.waapp-ppaao.org/en) given by Institut Sénégalais de Recherches Agricoles (ISRA, https://www.isra.sn/). AC and BG received a grant from project Amaizing ANR-10-BTBR_01 (https://amaizing.fr/en/) of "Agence Nationale de la

Abstract

From the 17th century until the arrival of hybrids in 1960s, maize landraces were cultivated in the South-West of France (SWF), a traditional region for maize cultivation. A set of landraces were collected in this area between the 1950s and 1980s and were then conserved *ex situ* in a germplam collection. Previous studies using molecular markers on approx. twenty landraces from this region suggested that they belonged to a Pyrenees-Galicia Flint genetic group and originated from hybridizations between Caribbean and Northern Flint germplasms introduced to Europe. In this study, we assessed the structure and genetic diversity of 194 SWF maize landraces to better elucidate their origin, using a 50K SNP array and a bulk DNA approach. We identified two weakly differentiated genetic groups, one in the Western part and the other in the Eastern part of the studied region. We highlighted the existence of a longitudinal gradient along the SWF area that was probably maintained through the interplay between genetic drifts and restricted gene flows. The contact zone between the two groups observed near the Garonne valley may be the result of these evolutionnary forces. We found in landraces from the East part of the region significant cases of admixture between landraces from the Northern Flint group and landraces from either the Caribbean, Andean or Italian groups. We then assumed that SWF landraces had a multiple origin with a predonderance of Northern Flint germplasm for the two SWF groups, notably for the East part.

Introduction

Maize was domesticated from teosinte, *Z. mays* L. subsp. *parviglumis*, about 9000 years ago in Southern Mexico [1]. Thereafter, it spread from this domestication center over North and South America [1–3]. During these expansions, maize evolved in contrasted environments, leading to distinct genetic groups adapted to new climates such as cold temperatures and long days in the north and warmer temperatures in coastal Central America. After the discovery of America by Columbus in 1492, maize was introduced to Europe for the first time in 1493 from the Caribbeans to Spain [4–6]. In the first half of the 16th century, a second introduction of

Recherche, https://anr.fr/". The both funders played a role in decision to publish.

Competing interests: The authors have declared that no competing interests exist.

maize is documented in North-Eastern Europe from Northern Flint material originating from North America as suggested by an illustration of Leonard Fuchs, a German herbalist, in 1542 [4, 7]. Genetic and morphological differences between European maize landraces were found in numerous studies: studies based on morphological data separated European maize landraces principally according to their flowering time earliness and related morphological traits [8, 9], and maize landraces from North-Eastern Europe were shown to flower earlier than those from Southern Europe. Similarly, molecular studies confirmed a major differentiation between Northern and Southern Europe maize [4, 7–11]. Gauthier *et al.* [9] also detected genetic differentiation between South-Western (Northern Spain, Portugal and the Pyrenees) and South-Eastern (Greece and Italy) European landraces. These results suggested that the landraces cultivated in the North, South-West and South-East of Europe each resulted from introductions of different maize landraces belonging to different American regions.

Other molecular studies showed that maize landraces cultivated at intermediate latitudes in Europe originated from hybridizations between maize landraces from Caribbean, South American and North American landraces [4, 5, 7, 10–13]. These studies also revealed genetic similarities between American Northern Flint landraces and landraces from Northern Europe, and also between Caribbean and Southern Spain landraces, showing that the first maize landraces introduced from America still have direct representatives in Europe. By analysing both American and European landraces with SSR markers, Camus-Kulandaivelu *et al.* [10] were able to identify 7 genetic groups referred to as Mexican (MEX), Caribbean (CAR), Corn Belt Dent (CBD), Andean (AND), Northern Flint (NF), Italian Flint (ITA) and Pyrenees-Galicia Flint groups. Dubreuil *et al.* [7] found that maize landraces cultivated in Pyrenees and Galicia do not display close similarity with any of the American genetic group. Nevertheless, based on structure analyses, these landraces were shown to be intermediate between landraces from the South of Spain and from the North of Europe. Recently, Brandenburg *et al.* [12] analysed by resequencing first-cycle inbred lines directly derived from American and European landraces and showed that the European Flint group was close to the Northern Flint group. They also detected gene flow from ancestors belonging to Southern European landraces. Mir *et al.* [11] argued that landraces from Pyrenees displayed a hybrid origin between Northern U.S Flint and Northern South-American landraces, with a predominance of the latter. Thus, all these studies highlighted the importance of Northern Flint in the origin of the Pyrenees-Galicia group. The geographic origin of this NF material remains an open question, *i.e.* did it arrive directly from America or indirectly through NF landraces from the North of Europe. These two patterns of NF introductions were proposed by Tenaillon and Charcosset [13] on their map of maize introduction to Europe.

All the above information highlights the need for genetic markers analyses to better understand the evolution of maize landraces located in intermediate regions of Europe. Since maize is an outcrossing species, landraces are expected to display relatively high levels of genetic diversity. To assess this diversity, molecular studies must thus include several individuals. However, for large sets of landraces, molecular characterization based on several individuals per landrace is limited by laborious and costly experimental processes. To overcome this problem, several studies on maize landraces were performed on bulks of several plants per landrace, for RFLP analysis [4, 9, 14, 15] and SSR analyses [7, 10, 11, 16–19]. First, Dubreuil *et al.* [14] evaluated a method based on the RFLP analysis of balanced DNA bulks from several individuals. They found that allelic frequencies could be estimated with a high precision in bulks of 15 individuals. The major disadvantage of bulk DNA analysis was the loss of information on individual genotypes, preventing access to variation at individual level and thus to genetic parameters such as Wright's fixation indices (F_{IS} and F_{IT}). Nevertheless, bulking is highly efficient to estimate most parameters of interest regarding population genetic structure and to make

inference about the evolutionary history of populations. Recently, this bulk approach was investigated on SNP markers in maize [20], leading to the development of a new method for allelic frequency estimations [21].

In this report, we focus on the South-West of France (SWF), one of the main traditional maize cultivation areas in Europe. Maize cultivation in SWF was first reported in 1626 in the 'Béarn' region, and in 1628 and 1637 respectively in the towns of 'Bayonne' and 'Castelnaudary' [22]. Based on historical records, maize spread from its introduction in the South-West of France at the end of the 16th century and was already largely cultivated in this area (from the Pyrenees to the Garonne Valley) at the end of the 17th century [22]. American hybrids were introduced from 1948 onwards in France and maize landraces in SWF were progressively replaced by commercial American x European hybrids. The origin and genetic diversity of these landraces before the introduction of commercial hybrids are poorly known. The aims of this paper are (1) to assess the structure and genetic diversity of traditional maize landraces cultivated in the South-West of France until the 1970s, and (2) to determine genetic relationships between SWF landraces andAmerican and European maize landraces. To do so, we used the 50 K SNPs array [20, 21, 23] to analyse the diversity of 342 maize landraces, including 194 landraces collected in the South-West of France and conserved by INRA since the 1960s; as well as 148 European and American maize landraces already analysed by Camus-Kulandaivelu *et al.* [10] and Mir *et al.* [11] using SSR markers and by Arca *et al.* [23] using SNP analysis.

Materials and methods

Plant material

We studied 194 landraces that had been collected between 1949 and 1987 [24, 25] in the South-West of France (SWF, S1 Table) in the two French administrative regions "Nouvelle Aquitaine" (in 5 districts) and "Occitanie Pyrénées-Mediterranée" (in 9 districts). Since their collection, these landraces have been preserved at the French Maize Biological Resource Center (CRB, Mauguio, France, https://urgi.versailles.inra.fr/siregal/siregal/grc.do) and seed lots regenerated through four successive generations of multiplication using between 100 and 200 full-sib ears for each landrace. Passport data including information about the area of collection (continent, country and district) and geographical coordinates (latitude and longitude) are available for each landrace although geographic coordinates are lacking for 36 of these 194 landraces (S1 Table).

To address questions about the origin of these landraces and to compare landraces from SWF with original American material, we used a worldwide reference panel consisting of 64 European and 73 American maize landraces (S2 Table) previously analysed using SSR markers by Camus-Kulandaivelu *et al.* [10] and using SNP analysis by Arca *et al.* [23]. These authors identified 7 genetic groups for these maize landraces that they termed the Andean (AND), Caribbean (CAR), Mexican (MEX), Northern Flint (NF), Corn Belt Dent (CBD), Italian (ITA) and the Pyrenees-Galicia (PG) groups. The NF, ITA, CBD and PG groups included respectively 2, 1, 2 and 17 landraces from the French Pyrenees. These 22 landraces originating from the region focused in this paper, we chose to remove them from their reference groups and put them in the landraces of the SWF region to assess *a posteriori* their genetic proximity with the other groups. Thus, the size of each reference groups was finaly as follows: 12 for AND landraces, 25 for CAR, 22 for MEX, 37 for NF, 15 for CBD, 17 for ITA and 9 for Pyrenees_Galicia _2 (PG_2) groups (S2 Table). The PG_2 reference group comprised 9 landraces from the PG genetic group identified by Camus-Kulandaivelu *et al.* [10] that originated from outside SWF. To complement this reference set, we added 2 landraces from Portugal and 9 landraces from South America previously analysed by Mir *et al.* [11].

DNA extraction and genotyping with 50K SNP array

We assessed nucleotide diversity in the 194 maize landraces collected in South-West France using the Maize 50K SNP array developed by Ganal *et al.* [20] and DNA bulks representing each landrace. To this aim, DNA was extracted from leaf disks collected on 15 individuals per landrace. To assess population diversity, we selected the 30,068 Panzea markers (PZE-prefix SNPs) proven suitable for diversity analyses [20] and we predicted SNPs allele frequencies using the method described in Arca *et al.* [21] briefly presented below. The approach consists in a two-steps analysis of the relative fluorescence ratio for each allele. The first step consists in determining, for each SNP, whether a landrace is fixed for the A (or for the B) allele, by comparison with the distribution of ratio of A *vs*. B fluorescence within a set of 327 inbred lines. If the SNP is declared polymorphic within the landrace, the second step consists in estimating the allelic frequencies at this locus using a calibrated average curve established from1000 SNPs on polymorphic DNA pools controlled for their composition. Arca *et al.* [21] observed that the mean absolute error (MAE) for allelic frequencies estimation at both steps was on average 7%. For the 82 American and the 66 European maize landraces used as a worldwide reference panel, we used Arca *et al.* [21]'s SNPs database obtained using the same SNP array and following the same procedure. Combining these different datasets, we finally obtained an allele frequency database for 23,412 SNP in which diversity was assessed on 342 landraces.

Diversity and genetic structure analysis of SWF landraces

For each landrace, we estimated the allelic richness (A_r) and gene diversity (H_e, [26]) using a *ad hoc* script in the R language v3.0.3 [27].

To assess the population genetic structure underlying our panel of SWF maize landraces, we used the software ADMIXTURE v1.07 [28]. This method assumes the existence of a predetermined number (K) of clusters and estimates the fraction of ancestry of each accession in each of the K clusters (Q). It also infers the SNP allele frequencies of the ancestral landraces (P). The software requires individual genotypic data, while our SNP analysis on DNA bulks provides allelic frequencies only. To obtain individual genotypic data, we thus simulated 5 haploid genotypes per landrace from allele frequencies under the hypothesis of Hardy-Weinberg equilibrium and no linkage disequilibrium [10, 21, 23]. To limit linkage disequilibrium between SNPs, we divided the maize genetic map into 2500 non-overlapping windows and randomly selected a single SNP in each windows [21, 23]. We explored K values ranging from K = 1 to K = 13. The likely number of genetic groups was estimated using the *DeltaK* parameter following the method proposed by Evanno *et al.* [29]. Each landrace was assigned to the cluster where its fraction of ancestry was highest.

Relationship between landrace genetic structure, geographic and climatic variables

To describe the genetic structure of SWF landraces, we first looked for links between the genetic structure and the geographic localization of the landraces. To this aim, we calculated pairwise genetic distances over the 194 maize landraces using the modified Rogers's distance [30, 31]. On the corresponding genetic distance matrix, we performed a principal coordinate analysis (PCoA) using R ade4 package and estimated correlations between landrace coordinates on the first axis of the PCoA and either (1) the longitude or (2) the latitude of their site of collection. This analysis was based only on those 158 SWF landraces with available geographic coordinates. We also looked for patterns of isolation by distance. To this aim, we performed a regression analysis between the "genetic divergence matrix" which we estimated using the

linearized F_{ST} /(1- F_{ST}) values following Rousset [32], and the pairwise geographic distance matrix calculated using latitude and longitude coordinates of the site of colection of each landraces with spherical law of cosines formula [33]. Pairwise F_{ST} values were calculated over SWF landraces using the Gst estimate proposed by Nei [34]. The statistical significance of the correlation between these two matrices was evaluated using a Mantel test [35], R vegan package.

We also explored relationships between the genetic structure and climatic variables. To this aim, we created a climatic data base containing values of monthly precipitation and monthly mean temperatures of prospection sites for the 194 landraces from the SWF. We first retrieved data for monthly precipitation and monthly minimum and maximum average temperatures from the WORLDCLIM database [36]. Then, we estimated the monthly mean temperatures as mean of the sum of monthly minimum and maximum average temperatures. Considering that maize was cultivated in Europe from May to October, we retained the mean monthly temperature and mean monthly precipitation for only these 6 months. These lead us to obtain a climatic database of 12 variables of maize growing season. When we performed a principal component analysis on these 12 variables with R ade4 package, we observed that the first axis of this PCA explained 82.5% of the climatic variability observed for temperatures and precipitations covered by the locations of our set of 158 SWF landraces. The climatic matrix distance between our 158 landraces was computed as the Euclidian distance on this first PCA axis (Ecodist R package, Goslee and Urban [37]). Thereafter, we used the MRM function of the Ecodist R package [37] to perform a multiple regression of the genetic distance matrix on the geographical and the climatic matrixes. We also compared, through an ANOVA analysis, the climatic characteristics of the different genetic groups identified using ADMIXTURE.

Comparative analyses between SWF landraces and worldwide maize genetic groups

Genetic diversity. To compare allelic richness and gene diversity between well-known maize genetic groups and our set of SWF landraces, we used results obtained using ADMIXTURE on our panel of 194 maize landraces for K = 2 and considered the seven genetic groups identified by Camus-Kulandaivelu *et al.* [10]. Allelic richness and genetic diversity (H_e) were estimated in each of these groups and, to compare allelic richness (A_r) among groups of different sizes, we used the rarefaction method proposed by Petit *et al.* [38] with $n = 1000$ resampling. To estimate the expected heterozygosity (H_e) for each group, we calculated the averaged values of allelic frequencies at the 23K SNP over the landraces owing to each group. Pairwise comparisons of A_r and *He* values between groups were carried out using Wilcoxon signed-rank tests across all the 23 K SNPs. Statistical analyses were computed using a R Core Team [27] script (S. Nicolas, personal communication).

Looking for footprints of hybridization between main maize genetic groups in the origin of SWF landraces. To determine whether SWF landraces originated from hybridizations between ancestral landrace groups, we used the 3-populations test [39] implemented in the TreeMix software version 1.12 [40] This test compares a focus population (or group) X to two reference populations (or groups) Y and W, and calculates an *f3 statistic*, f3(X; Y, W) defined as the product of the difference in allele frequencies between populations X and Y, and the difference in allele frequencies between populations X and W. If the focal population X can be considered as resulting from an admixture or hybridization event between populations Y and W, the value of the *f3-stat* can be negative. A *Z-score* value $<$-2 indicates a significant mixture [39].

In our analysis, we tested two scenarios for the origin of SWF landraces. In the first one, we made the hypothesis that SWF landraces originated from hybridizations between maize

landraces from Caribbean (CAR) and Northern Flint (NF) genetic groups as proposed in previous articles. To this aim, we applied 3 population s tests on triplets composed of a focal SWF group previously identified using ADMIXTURE analyses and the two genetic groups representing CAR and NF. In order to test the distinct involvement of NF materials from North America and NF from Northern Europe in the origin of SWF landraces, we subdivided NF landraces into American Northern Flint (NFA) and European Northern Flint (NFE) groups. Then, we analysed two hypotheses: SWF landraces originate from (1) an hybridization between Caribbean landraces and Northern Flint landraces previously introduced in Europe (NFE), or (2) an hybridization between Caribbean landraces and American Northern Flint landraces (NFA) potentially independently introduced in the South-West of France.

In the second scenario, we assumed that SWF landraces originated from hybridizations but without any *a priori* about the parental groups. For example, to determine whether SWF landraces originated from hybridization between landraces belonging from NF and MEX groups, we performed TreeMix 3-population tests using each of the genetic groups identified in the SWF as the focus group and NF and MEX groups as the parental groups. All these groups were defined on the basis of structure analysis using the arbitrary cut off of upper than 0.7.

To vizualize genetic relationships between SWF landraces and worldwide reference genetic groups, we also performed a Neighbor joining tree with ape R package and a principal component analysis (PCA) with ade4 R package. For both analyses, we used the modified Rogers' genetic distance matrice [30] estimated on SNP data of 342 maize landraces. To perform the PCA analysis, we only used the 72 American landraces representing the 4 main historical maize genetic groups (*i.e.* MEX, AND, CAR and NF groups). The European maize landraces (including SWF landraces) were added as supplementary data. Due to their recent origins, the landraces of CBD group [41] were also added as supplementary data. We used the modified Rogers' distance between each SWF landrace and each American landrace in order to identify the closest American maize genetic group.

Results

Among the 23412 SNPs analysed, 378 exhibited missing data for at least one landrace and discarded, leaving 23034 SNPs. Among them, 15 were monomorphic in all the landraces considered in this study (n = 342). At landrace level, the proportion of polymorphic loci varied from 13.5% for IPK60, a German Northern Flint landrace, to 92.8% for EPZMV23, a landrace from Spain. On average, the proportion of polymorphic loci within landrace was 68.7%.

Genetic diversity and population structure

The proportion of polymorphic SNPs in landraces from the South-West of France varied between 34% and 86%, with an average value of 70%. Averaged over all SNPs, the mean allelic richness (Ar) per landrace was 1.70, and varied from 1.33 to 1.86. Mean gene diversity (H_e) was 0.22 per landrace (min = 0.12; max = 0.27, Fig 1, S1 Table).

Using the ADMIXTURE software on the whole set of SWF landraces, we observed that the maximal deltaK value occurred at K = 2 (S1 Fig). The next two large peaks of deltaK value were observed at K = 3 and K = 5 (S1 Fig). As shown in Fig 2, we identified a stratification in two major groups at K = 2 distinguishing landraces located in the Eastern part of the South-West of France from those located in the Western part. The first group, which was named 'East South-West France' (E-SWF), included 65 landraces. Most of these landraces were collected in Eastern districts such as Ariège, Tarn and Haute-Garonne, but ten of them were collected in Western districts (*i.e.* districts such as Pyrénées Atlantiques, Hautes-Pyrénées, Landes and Gironde). The second group, referred to hereafter as 'West South-West France' (W-SWF),

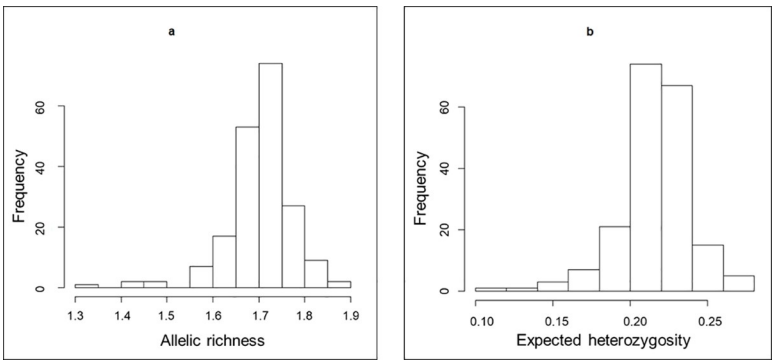


Fig 1. Genetic diversity analysis of SWF landraces. Histogram of (a) allelic richness and (b) expected heterozygosity values estimated for each of the 194 landraces.

https://doi.org/10.1371/journal.pone.0238334.g001

included 126 landraces mostly originating from the Hautes-Pyrénées, Pyrénées-Atlantiques and Landes districts. Twenty-three landraces from this group were nevertheless collected in the Eastern part of the area. Interestingly, most of the landraces collected in the Haute-Garonne and Ariège districts exhibited patterns of admixture between W-SWF and E-SWF groups (Fig 2). The contact zone between E-SWF and W-SWF groups is located between these two districts, near Garonne valley (Fig 2). At K = 3, we observed a third group mostly composed of the 10 landraces that were previously assigned to E-SWF group at K = 2 but located in the Western part of the region (S2 Fig). At K = 5, we observed two groups, E-SWF and W-SWF, a new group shaped by landraces from Gironde and Lande districts (12 and 4 respectively), and two other groups mostly deriving from the third group identified at K = 3.

PCoA on allele frequencies revealed the same population structure pattern. The first axis (representing 13% of the total inertia) separated landraces located in the Western part (W-SWF cyan colour in Fig 3A) from those located in the Eastern part (E-SWF dark blue colour in Fig 3A). The second PCoA axis of Fig 3A (representing 6.3% of total inertia) highlighted

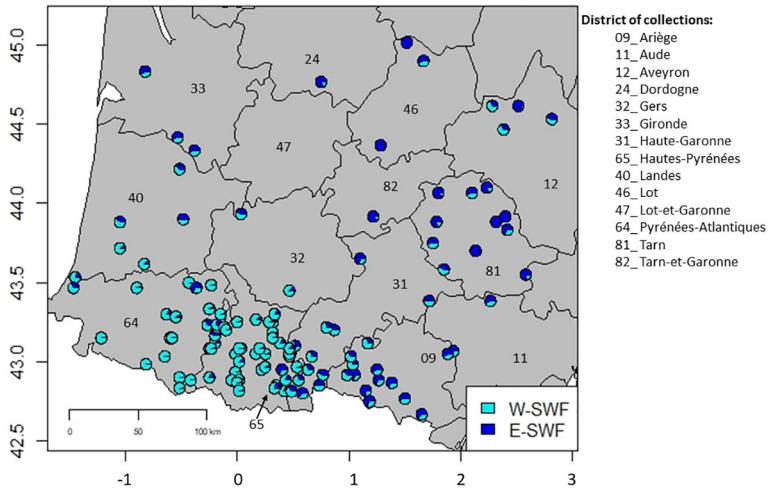


Fig 2. Geographical representation of the 158 maize landraces from the South-West of France (36 landraces having no geographical coordinates were not plotted). The districts of collection sites are represented with administrative numbers. Landraces for which pie-diagrams had both blue and cyan colors exhibited mixed genetic origins between E-SWF and W-SWF groups. The map was created using R software (libraries maps, plotrix, mapproj, devEMF).

https://doi.org/10.1371/journal.pone.0238334.g002

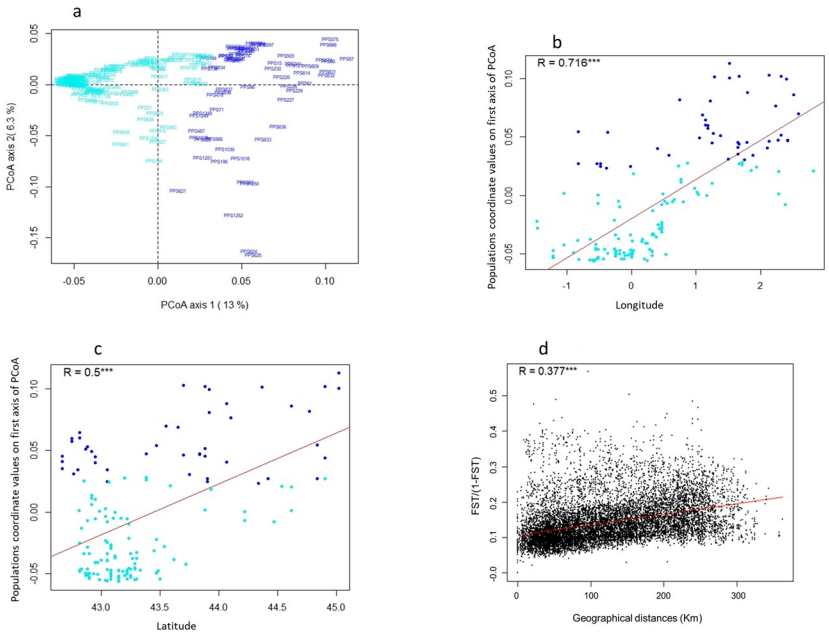


Fig 3. Spatial genetic structuration analysis of SWF landraces. (a) Principal coordinate analysis on SNP data using Rogers 'genetic distance matrix estimated among the 194 landraces. (b) Longitudinal and (c) latitudinal gradient analyses on SNP data of 158 landraces from South-West France using coordinates values on the first axis of PCoA analysis. SWF landraces were colored according to their assignation to E-SWF and W-SWF groups identified with admixture analysis at K = 2. (d) Isolation by distance (IBD) analysis of 158 landraces from South-West France using F_{ST} / (1- F_{ST}) matrix for genetic differentiation estimated following Rousset [32] and geographical distance matrix estimated for each pair of landraces. R = coefficient of correlation. *** = p-value<0.001.

https://doi.org/10.1371/journal.pone.0238334.g003

variation among E-SWF landraces. Six landraces, assigned to the third group identified with ADMIXTURE analysis on SNP data of 194 SWF landraces at K = 3 (S2 Fig), were indeed strongly differentiated from the majority of E-SWF landraces on the PCoA plan 1–2. Fig 3B and 3C highlighted the existence of a longitudinal and a latitudinal gradient (respectively) along the SWF. Significant correlation was found between genetic differenciation based on F_{ST} and geographical distance (Fig 3D).

We analysed the genetic diversity of both groups. The allelic richness of W-SWF was significantly higher than that of E-SWF: A_r = 1.718 and A_r = 1.679 respectively (Wilcoxon's signed-rank test, $p\text{-}value < 10^{-15}$; Table 1). On the other hand, the E-SWF group exhibited a higher level of gene diversity compared to the W-SWF group: H_e = 0.290 and H_e = 0.273 respectively ($p\text{-}value < 10^{-15}$ using Wilcoxon's signed-rank test). The fact that A_r is lower and H_e is higher in E-SWF group compared to W-SWF group indicates that allelic frequencies are more balanced in E-SWF group than in W-SWF group. The W-SWF/E-SWF differentiation explained a low but significant proportion of the overall genetic diversity: F_{ST} = 0.026 ($p\text{-}value < 10^{-15}$ based on permutation test over landraces). The mean F_{ST} between pairs of landraces was 0.104 and 0.168 for the W-SWF and E-SWF groups, respectively. This shows that the E-SWF landraces are, on average, more differentiated from each other than W-SWF landraces.

Relationships between geographic distribution, climate variables and population structure

To determine the main factors underlying the genetic structure observed in SWF land races, we searched for a relationship between genetic variation and either the geographical origin of the landraces or the climatic characteristics of their site of origin. First, we looked for

Table 1. Allelic richness (A_r) and gene diversity (expected heterozygosity, H_e) for the two SWF groups identified with admixture at K = 2 and the 7 genetic groups previously identified in Camus-Kulandaivelu *et al.* [10].

Group.Name	Population Size	Allelic richness (A_r)	Expected heterozygosity (H_e)
Andean	12	1.75	0.31
Caribbean	25	1.72	0.32
Corn Belt Dent	15	1.75	0.33
Pyrenees_Galicia_2	9	1.76	0.30
Italian Flint	17	1.66	0.29
Mexican	22	1.71	0.33
Northern Flint	37	1.52	0.27
W-SWF	129	1.72	0.27
E-SWF	65	1.68	0.29

A_r values were calculated using a rarefaction method as in [38], with 1000 resampling of 9 landraces without replacement in each group, except for the Andes group for which we obtained only 220 possible resampling of 9 landraces. We used a Wilcoxon signed-rank test to analyze statistical differences between peers of H_e (or pair of A_r) values estimated for the 9 groups. All comparisons for pair of H_e (and pair of A_r) were significant (*p-value*< 0.009); except for comparison between He of ITA and SWEF groups (*p-value* = 0.45) and comparison between A_r of Andean and Corn Belt Dent groups. E-SWF = East South-West France; W-SWF = West South-West France.

https://doi.org/10.1371/journal.pone.0238334.t001

associations between (i) landrace coordinates on the first axis of the PCoA (Fig 3A) performed on SNPs and (ii) either the latitude or longitude coordinates of their prospection sites. As shown on Fig 3B and 3C, linear regression analyses evidenced significant correlations with both longitudinal ($r = 0.72$; *p-value* $< 10^{-15}$) and latitudinal coordinates ($r = 0.5$; *p-value* = 10^{-10}), highlighting the existence of both longitudinal and latitudinal genetic gradients. We also identified a significant correlation between the pairwise F_{ST} / (1- F_{ST}) ratio and pairwise geographic distances ($r = 0.38$, *p-value* $< 10^{-15}$, Fig 3D), suggesting isolation by distance.

Examining climate characteristics from May to October in the collection sites of our SWF landraces, we observed that landraces belonging to the E-SWF group originated from sites associated with higher temperatures and weaker precipitations than for W-SWF landraces (S3 Table). Differences in monthly temperatures and precipitations between E-SWF and W-SWF landraces were low, but significant variations were observed for temperature (from June to August) and precipitations (during September and October; see S3 Table, in bold). To explore how much genetic differentiation between SWF landraces could be explained by geographic distance instead of climatic distance, we performed multiple regression analyses of genetic distance on geographic and climatic distance matrices that showed a significant relationship only for the geographic matrix (*p-value* = 0.01 for geographical values VS, *p-value* = 0.67 for climatic values). This suggests that the demographic history (instead of selection) of these populations is the main factor driving population differentiation.

Genetic diversity of SWF landraces compared to the main maize genetic groups

As shown in Table 1, the two genetic groups identified in the South-West of France using ADMIXTURE exhibited larger allelic richness (A_r) and gene diversity (H_e) than the Northern Flint group, but lower allelic richness and gene diversity than the Corn Belt Dent (CBD), Andean (AND) and Caribbean (CAR) groups.

Do SWF landraces originate from hybridization between known maize genetic groups? We used a treeMix 3-population test to determine whether E-SWF and W-SWF groups result from hybridization events beween historical maize genetic groups. Two different hybridization

Table 2. Negative values of f3-stat and Z-score on TreeMix 3-population test with E-SWF group as focus groups.

Focus group	Reference group 1	Reference group 2	f3-stat	Sd	Z-score
E-SWF	AND	NFA	-0.00314147	0.000766376	-4.09912
E-SWF	***CAR***	***NFA***	***-0.00897185***	***0.0008593***	***-10.4409***
E-SWF	CAR	NFE	-0.00739079	0.00082723	-8.93438
E-SWF	***ITA***	***NFA***	***-0.00874928***	***0.000730175***	***-11.9825***

AND = Andean, CAR = Caribbean, ITA = Italian, NFA = Northern Flint landraces from America, NFE = Northern Flint landraces from European. No negative f3-stat values were observed for the W-SWF group.

https://doi.org/10.1371/journal.pone.0238334.t002

scenarios were considered: (1) SWF landraces originated from hybridizations between landraces from Caribbean (CAR) and Northern Flint (NF) groups, or (2) SWF landraces originated from hybridizations between any of the historical genetic groups other than CAR x NF. Negative *f3-stat* values, suggestive of a hybridization event between two reference groups as ancestors, were detected for scenarios 1 and 2 for the E-SWF group (Table 2) but not for the W-SWF group (S4 Table). Z-scores of 3 populations tests showed that the most significant scenarios were mixed events between (i) landraces from NFA and (ii) landraces from either CAR or ITA groups (bold and italic line in Table 2). No negative f3-stat values were observed for W-SWF group.

We did not observe signals for hybridization events involving CBD or MEX groups as origins of E-SWF and W-SWF groups, suggesting that these were not influenced by introductions from MEX landraces, nor by hybrids originating from CBD landraces after their introduction to Europe.

Genetic relatedness between SWF landraces and maize from Europe and America. Results of the PCA analysis performed on SNPs for the whole set of 342 maize landraces are presented in Fig 4 and S2 Table. Fig 4 shows the position of the different groups on the first and second axes of the PCA. The first axis (19.7% inertia) mainly separated Northern Flint landraces (NFA, NFE) from Tropical landraces (AND, CAR, MEX) while the second axis (6.4% inertia) mainly separated South American from Caribbean landraces. As expected, most European NF landraces were close to American NF landraces, and landraces from Southern Spain were close to those from the Caribbean Islands. The two SWF groups occupied a central position on the first PCA plan (Fig 4) but were somewhat separated on the first two axes. Landraces from the E-SWF group were more dispersed than those from W-SWF. On the first axis, E-SWF landraces appeared closer to the NF group than did W-SWF landraces. On the contrary, W-SWF landraces were closer to tropical landraces (AND, CAR and MEX) than to NF landraces (Fig 4). As to relationships between landraces from Europe, we found that landraces from SWF were closer to Pyrenees_Galicia_2 (PG_2) and European Northern Flint landraces (NFE) than to Italian Flint landraces.

To further analyse the relationship between landraces from the SWF groups and reference groups, we performed a neighbor-joining tree based on Roger's distance matrice on SNPs data of 342 landraces (Fig 5). We observed that landraces are distributed in two large groups distinguishing NF landraces from those from tropical regions of America (AND, MEX and CAR groups). The majority of landraces from SWF assigned to E-SWF and W-SWF groups were close to the NF group, suggesting a major influence of NF landraces in the evolution of SWF landraces. Landraces of E-SWF group were closer to NF group than landraces of W-SWF group. The remaining SWF landraces (36 landraces in total, blue and cyan star symbols in Fig 5) were close to landraces from the ITA, PG_2 and CBD groups, suggesting possible recent gene flow between these maize landraces from SWF and those from ITA and CBD groups. The

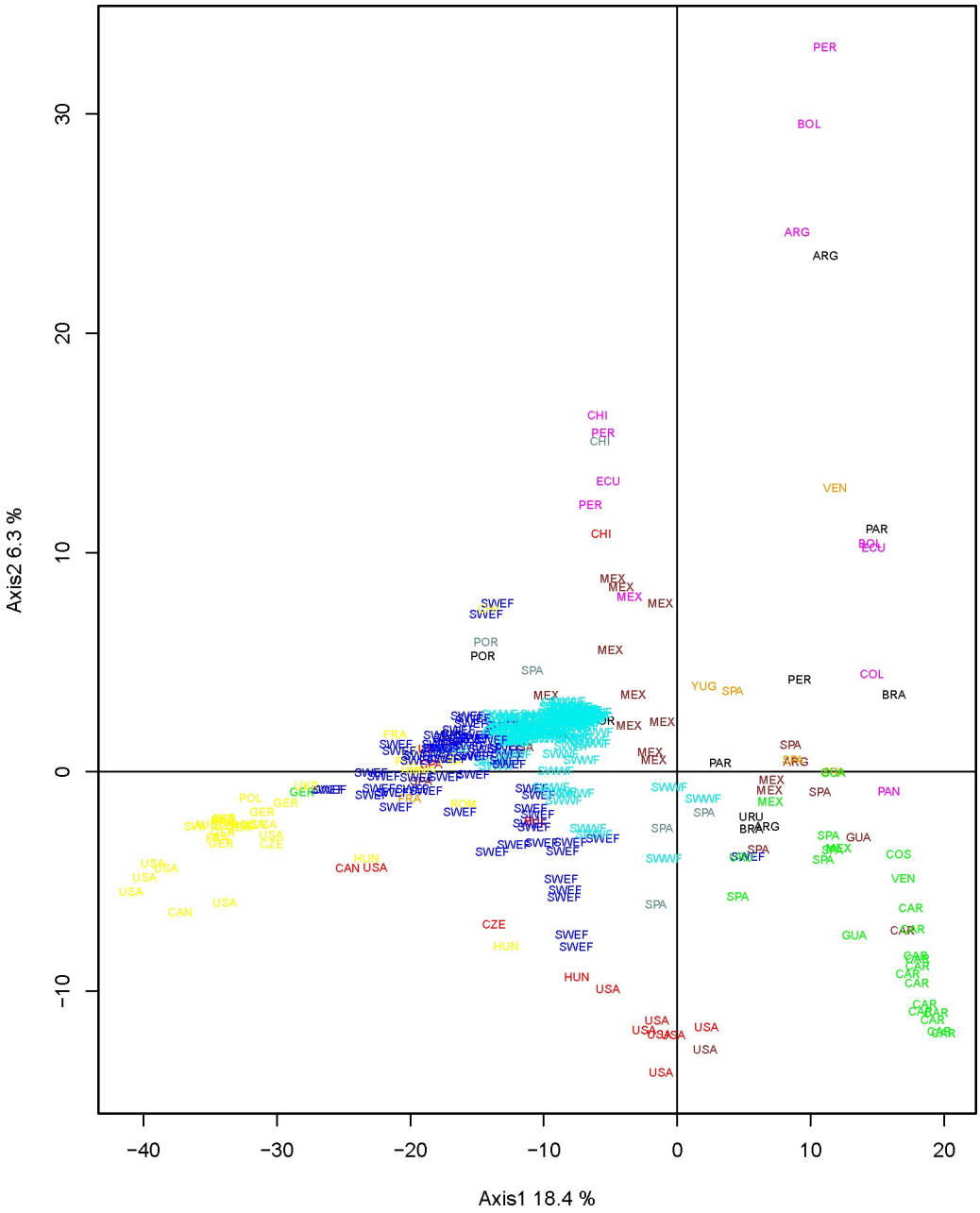


Fig 4. PCA analysis on SNP data of 72 American maize landraces with 260 maize landraces from Europe and those of CBD as supplementary data. Labels indicate landrace country of origins (see S2 Table). South-West France landraces were colored according to their admixture result at K = 2 allowing to distinguish East (in blue) and West South-West France (in cyan) genetic groups. The remaining American and European landraces were colored according to their genetic groups previously identified by Camus-Kulandaivelu et al. [10]: Corn Belt Dent in red, Caribbean in green, Northern Flint in yellow, Mexican in brown, Italian Flint in orange, Andean in magenta, Pyrenees_Galicia_2 in grey. The 9 landraces from South America and 2 landraces from Portugal studied by Mir et al (2017) were colored in black.

https://doi.org/10.1371/journal.pone.0238334.g004

ITA, PG_2 and CBD groups were themselves closer to landraces from Tropical region of America (AND, MEX and CAR groups). We also observed that these 36 landraces from the SWF were assigned to different genetic groups in our ADMIXTURE analysis with SNP data from 194 SWF landraces at K = 5, distinguishing them from E-SWF and W-SWF groups (S2 Fig).

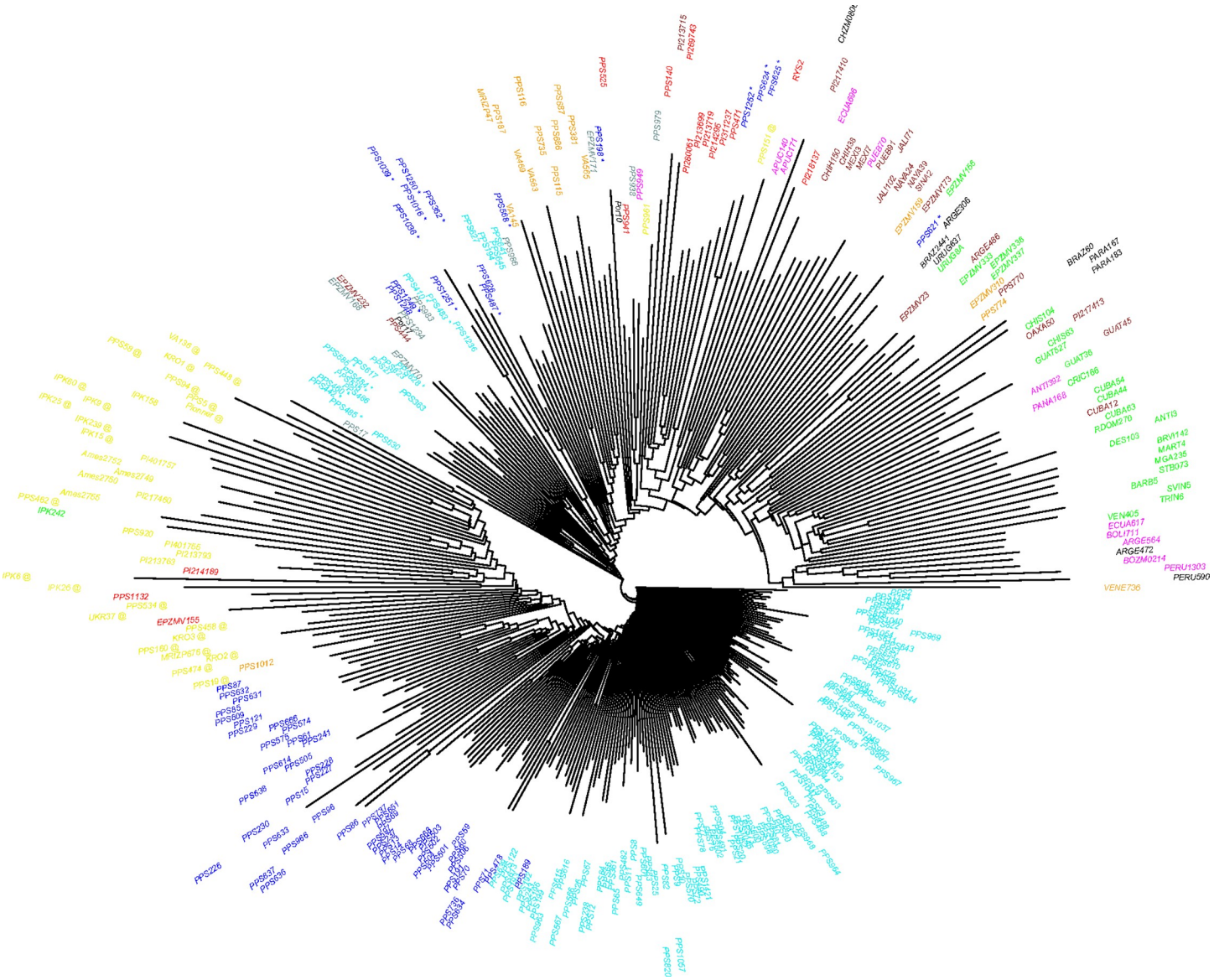


Fig 5. Neighbor joining tree for the 342 populations obtained from Rogers distance matrix. South-West France landraces were colored according to their admixture result at K = 2, allowing to distinguish East (in blue) and West South-West France (in cyan) genetic groups. The remaining American and European landraces were colored according to their genetic groups previously identified by Camus-Kulandaivelu et al. [10]: Corn Belt Dent in red, Caribbean in green, Northern Flint in yellow, Mexican in brown, Italian Flint in orange, Andean in magenta, Pyrenees_Galicia_2 in grey. The 9 landraces from South America and 2 landraces from Portugal studied by Mir et al (2017) were colored in black. SWF landraces assigned to the 3 other groups identified with our ADMIXTURE analysis on SNP data of 194 SWF landraces at K = 5 were represented with star symbols in blue and cyan. The European Northern Flint landraces (NFE) were represented by an @ symbol in yellow.

https://doi.org/10.1371/journal.pone.0238334.g005

To identify genetic distances between landraces from SWF and those from America, we used a heatmap representation with R gg2plot function to visualize the result of genetic distance estimation between SWF landraces and the 82 American landraces (S3 Fig). The heatmap representation of genetic distances showed that the 194 landraces from SWF were closer to 4 Chilean (PPS949 from AND, PPS941 from CBD, PPS961 from NF and PPS938 from MEX groups), 2 CBD (PI214189 and PI280061) and 2 NF (PI213793 and PI401755) landraces from North of America than landraces from others American countries. These results suggested an introduction of Northern Flint and CBD landraces from Northern America into the SWF

genepool, as well as exchanges of maize landraces between Chile and SWF. Comparing the genetic distances of the two SWF groups with the American accessions, we did not evidence any clear difference between the genetic origins of these two groups, except that E-SWF landraces appeared to be close to landraces from the NF group and that W-SWF landraces were closer to Tropical landraces (S3 Fig).

Discussion

Validity and limitations of bulk DNA analysis with the 50 K SNP array

The usefulness of the bulk DNA approach in genetic diversity investigations has been proven in numerous studies [8, 14, 17, 18, 21, 42]. The major constraint observed in the use of this pooled DNA analysis is the loss of information about genetic variation at the individual level [8, 15]. As a result, genetic parameters such as individual's heterozygosity and Wright's F_{IS} fixation indices [31, 34] cannot be estimated directly. Bulk analysis also limits the use of classic genetic tools that require individual genotype information such as the ADMIXTURE [28], GENEPOP [43], HZAR [44] and STRUCTURE [45] softwares. As a result, we chose to estimate genetic parameters and to use software tools that do not require individual genotype information to characterize and analyse these landraces, except for the ADMIXTURE software for which we simulated genetic haplotypes [28].

Despite these constraints, the genetic proximities we revealed between American and European landraces using the 50K SNPs on DNA bulks were in accordance with what is known about maize diffusion in America and the main paths of maize introductions to Europe. Indeed, our PCA and Neighbor-joining analyses allowed distinguishing between the main historical maize genetic groups previously identified by Camus-Kulandaivelu *et al*. [10] using SSR markers. For instance, the genetic similarity between landraces from the South of Spain and Caribbean landraces corroborates the introduction of maize by Columbus from the Caribbean Islands to Spain [4–6]. We also detected strong genetic proximity between Northern Flint and Northern Europe landraces, in accordance with the introduction of maize landraces from Northern America to Northern Europe as suggested by an illustration of Leonard Fuchs, a German herbalist, in 1542 [4, 7].

Landraces of South West of France are differenciated along a longitudinal gradient

Using a genetic database of 23412 SNP markers to analyse the genetic diversity of maize landraces that were collected in the South-West of France circa 50 years ago, we were able to identify two main groups of landraces, located respectively in the Eastern and in the Western side of the South-West of France. This differenciation was consistently found in the different analyses (ADMIXTURE, PCoA, PCA and neighbor-joining tree) even if the genetic differentiation between them was relatively low (F_{ST} = 0.03). These 2 groups displayed a large genetic variability as compared with other groups.

We detected a significant correlation between the pairwise genetic distance and geographical distance matrices, which is indicative of isolation by distance (IBD). We also highlighted a longitudinal gradient along SWF using a multivariate technique (PCoA) based on a genetic distance matrix with SNP data from 158 landraces, which led to a significant correlation coefficient between the first PCoA axis and the longitudinal coordinates of their prospection sites.

The geographical trend observed in the diversity of SWF landraces could have been the result of diversifying selection. The occurrence of a genetical longitudinal gradient along the Pyrenees area as observed in this study has been reported for several species [46–49].

Gradients, such as altitude or longitude, play an important role in the Pyrenees, as they shape the existing proportion of suitable habitats along the mountains [47, 48]. Valbuena-Ureña *et al*. [49] emphasized that longitude is an important gradient to take into account in the Pyrenees, because the influence of the Atlantic ocean provides cooler and wetter climate westwards than in Eastern areas, which are more influenced by the Mediterranean climate.

We found differences for temperature and precipitation between the Eastern and Western parts of SWF, but regression analysis showed that, when considering geographical covariates, climatic variations did not explain genetic variability observed between SWF landraces. Since climate variation are geographically linked with longitude and latitude, we still cannot completely exclude the involvement of climatic variation in the differentiation of SWF landraces into two genetic groups; our results rather suggest that an equilibrium between restricted gene flow and genetic drift is the main driver of population differentiation between landraces in this region. This is supported by the longitudinal and latitudinal gradients that we observed on the first axis of PCA, which is determined by a large number of SNPs. However, the genetic differences observed for SWF landraces could be also explained by farmers' practices: the management of seed exchanges in different valleys limited by hilly mountains probably contributed to differentiate SWF landraces in different districts. However, the ethnobotanical survey realized in 2012–2013 among farmers in the Pyrenean area did not reveal any major differences in farmers' practices between the East and West parts of the Pyrenees [50]. As the differenciation signal can be seen at the genomic level, its suggests that demographic and gene flow forces were at work in SWF evolution.

To elucidate the origin and evolution of maize landraces in the South-West of France, we conducted an investigation of the relationship between SWF and putative ancestors.

The use of genome wide analysis to infer the origin of SWF landraces

A predominant origin from Pyrenees-Galicia. We did not observe any clear differentiation between SWF landraces and those of the Pyrenees-Galicia group, suggesting that SWF landraces originated from a single genetic group close to the ancestors of the Pyrenees-Galicia group (F_{ST} = 0.03 between the Pyrenees_Galicia_2 group and each of the groups from the South-West of France) [4, 7, 8, 10].

Several studies on genetic diversity of European and American maize landraces [4, 5, 7] highlighted a Pyrenees-Galicia Flint group made up of European landraces including a small sample of landraces from the South-West of France and landraces from the North of Spain. This original landrace group was hypothesized to originate from hybridizations between landraces of Caribbean and Northern Flint groups; however less than 30 landraces from Pyrenees-Galicia were used in these previous studies. In our study, using a much larger data set of 194 maize landraces from the SWF, the origin for SWF landraces was more precisely analysed. Our PCA and Neighbor-joining analyses showed genetic proximity between landraces from the Western part of the South-West of France and maize landraces from PG_2 group, supporting genetic relationship mentioned between landraces from Pyrenees (France and Spain) and Galicia in previous studies of maize landraces named as Pyrenees-Galicia Flint group [7, 9, 10]. Based on historical documents, the introduction of maize in the South-West of France was reported by Renoux [22] as originating from Spain. In her study, Vouette [51] reported the name of maize in the Pyrenees valleys as "wheat from Spain", as given by Bonnafous (1836). Similarly, maize is first mentionned in the "Mercuriales of Verfeuil" (near Toulouse) in 1637 as "millet of Spain" [52]. The vernacular names thus only gave indications of where the maize landraces came from before their introduction to a given country. The introduction of SWF maize from Spain in the 17th century is also explained by the coasting trade practiced by

fishermen from French and Spanish fishing harbours on the Atlantic coast [53]. In his work on relations between Spain and the South of France in the 17th century, Mauro [53] described the route of maize that progressed from the West part towards the East part of South-West France. This progression of maize from West to East is consistent with the greater genetic allelic richness observed in the W-SWF group than in the E-SWF group.

A specific influence of Northern Flint landraces in E-SWF. NF landraces played an important role in the evolution of SWF landraces and more generally in the adaptation of maize in Europe [4, 6] and in the development of CBD germplasm [54]. Our TreeMix 3-population analysis allowed us to confirm Northern Flint landraces influence on the evolution of E-SWF group landraces. For all admixture events highlighted in this analysis, E-SWF landraces displayed a mixed origin between NF landraces with either CAR, AND or ITA landrace groups. There are two possible ways to explain the presence of NF marks in the genome of SWF landraces: either directly from a secondary introduction from North America (NFA), or from an introduction from Northern Europe (NFE) as reviewed by Tenaillon and Charcosset [13]. Revilla et al [6] supported the NFA hypothesis by suggesting that NF landraces could have been introduced several times to the North of Spain from the beginning of the 17th century. Later, Brandenburg et al. [12] found that European Flint inbred lines, *i.e.* first-cycle inbred lines directly derived from landraces after few generations of selfing, were issued from an admixture between inbred lines belonging to European Northern Flint landraces and those derived from Southern Spain lines. In our TreeMix 3-population test, we found a predominant involvement of the NFA group in the origin of the E-SWF group. However, our PCA analysis showed a closer genetic proximity between maize landraces from the Eastern part of SWF and Northern Flint landraces located in the North of Europe (NFE), supporting the scenario involving some gene flow from European Northern Flint landraces. These observations are probably due to the genetic similarity between NFA and NFE (F_{ST} = 0.04), which makes it difficult to distinguish between the two sources.

A limited influence fromrecent introduction of CBD varieties. Our PCA analysis showed genetic proximity between 10 SWF landraces and landraces belonging to the Corn Belt Dent genetic group (CBD). CBD is a group of landraces identified in previous studies as resulting from hybridizations between Northern Flint and U.S Southern Dent materials [41, 54]. The presence of CBD genomic traces in SWF can be explained by the recent introduction of American hybrid seeds in Europe, since hybrids belong to the CBD group. Indeed, starting in1948 with the Marshall plan, the first American hybrids were tested in France [22]. Then, from 1957 onwards, the first double hybrids with at least one American parental line were developed in France by INRA. Farmers' adoption of hybrids was fairly rapid [55]. These 10 landraces, showing a genetic proximity with CBD group on the PCA analysis, shaped a specific group with admixture analysis at k = 3. Introgression of CBD germplasm into local landraces has also been observed in Central Italy with very variable effects [56].

Other relationships with American landraces. In our TreeMix 3-population test, we found that the E-SWF group originated from crosses between NF and Andean landraces. Mir *et al.* [11] showed that SWF maize landraces had a Northern South America origin, but our study does not allow us to really assess this path of introduction, as only four maize landraces from such origin were used in this study: VEN405 from CAR, VEN736 from ITA, ANTI392 and PANA168 both from AND groups. Our PCA and Neighbor-joining analyses showed little genetic similarity between W-SWF landraces and those from Northern South-America. It is possible that our study did not capture very well genetic relationships between SWF landraces and these expected Northern South-American parental landraces identified by Mir *et al.* [11] because of a limited sample of South America landraces. In consequence, we had some difficulty to evidence relationships between SWF and Northern South America landraces. Note

that American landraces introduced in Europe evolved in America during about 500 years, which limits the ability of TreeMix 3-population test to detect hybridization events between SWF and American landraces.

Our collection of SWF landraces presented the lowest genetic differentiation with landraces from Chile (except for CHZMO8050) compared to the remaining American landraces. All these Chilean landraces have been shown to originate from different genetic groups [10], with a probable replacement of traditional landraces with relatively recently introduced Northern US materials [3]. Genetic relatedness between Chilean and SWF landraces may be due to the introduction of SWF genepools into Chile or vice-versa. Interestingly, the Spanish (especially Basques and Andalusians) and also other Europeans such as French people (mainly coming from SWF) immigrated mainly to Chile in the second half of the 19th century, although Basque presence in Chile began in the conquistador period. Also exchanges between the Basque country and the South West of France have been taking place for a few centuries (https://en.wikipedia.org/wiki/Immigration_to_Chile, march 2020).

The different scenarios for SWF landrace origin and evolution

More generally, we thought that the South-West of France had been submitted to several introductions of maize with various genetic origins, as already reported for maize landraces in Spain [6] and in Portugal [11, 57]. Evidence for subsequent admixture between the genetic groups involved in these introductions was observed in this study with TreeMix 3-populations tests, PCA analyses and genetic distance analyses. This is particularly true for E-SWF landraces that showed hybridization between landraces from NF and either ITA, or CAR or AND groups in TreeMix 3-populations tests. E-SWF landraces were more scattered than W-SWF landraces, which were more homogenous. The hybridization of E-SWF group built by ITA and NFA groups could be explained by the presence of Italian agricultural workers in the 19^{th} and 20^{th} centuries in SWF [58].

All the results above allowed us to describe at least two scenarios of origin that take into account the influence of Northern Flint landraces from the North of Europe predominently in the Eastern part and the genetic proximity of W-SWF landraces with Tropical ones. For the first scenario, we postulated that, after the introduction of hybridized maize landraces from Spain (Caribbean x Northern Flint landraces, known as Pyrenees-Galicia group), the NFE landraces spread from the North to the Eastern part of SWF where they hybridized together.

In the second scenario, we hypothesized that the first maize landraces introduced to the South-West of France displayed predominantly Northern Flint ancestries as do Galician landraces [11]. Thus, these landraces spread from the Western to the Eastern parts of South-West France. Thereafter, Tropical landraces (from Mexican, Caribbean and Andean genetic groups) would have been introduced secondarily into Western altlantic coast spanning possibily Cantabrian and Basque coasts and hybridized with local landraces forming a new genetic entity which migrated from W-SWF to E-SWF. The gene flows were restricted as supported by IBD indicator, which slowed down the West-East introgression, explaining the current differenciation pattern between W-SWF and E-SWF. Presence of genetic marks from Tropical landraces also resulted from seed exchanges between Mediterranean trademen from Italy, France, Spain and Portugal at the end of the sixteenth century [11, 12]. Revilla Temiño *et al.* [5], Patto *et al.* [57], Brandenburg *et al.* [12] and Mir *et al.* [11] found that the South of Europe has experienced maize landraces introductions from both the Caribbean islands and South America.

The contribution of maize landraces from various genetic groups found in this study makes it difficult to determine which one of the two plausible scenarios mentioned above best explains the origin of SWF landraces. Additional historical data on maize introduction to Northern Spain and SWF would certainly help to shed more like light on this aspect.

Conclusion

Assessing the population genetic structure and diversity of a collection of crop accessions is the best way to develop an efficient management strategy and to improve breeding programs. Here we used DNA bulk analyses with a 50K SNP array to investigate the diversity and population structure for a panel of 194 maize landraces collected 30 to 60 years ago in the South-West of France. In this study, we assumed that the maize landraces from this area mostly originated from hybridizations between landraces from the Caribbean and Northern Flint groups. Introduced around 1626 through Atlantic harbours such as Bayonne, these SWF landraces have probably been posteriously submitted to gene flows from landraces belonging to various countries of America and Europe, as expected by Revilla Temiño *et al*. [5] and Patto *et al*. [57] for Spanish and Portuegese landraces respectively. A single mixture of NF and CAR landraces is not sufficient to explain the genetic origin of SWF maize.

The new results obtained for SWF in our study are encouraging to expand the approach to other regions of the world, to get a better understanding of diversity and history of local varieties.

Supporting information

S1 Fig. Plot of DeltaK of the log likelihood from the admixture analysis on SWF landraces SNP data. Group numbers varied from K = 1 to K = 13.
(TIF)

S2 Fig. Bar-plot for K = 2 to K = 5 for admixture analysis with the 194 SWF maize landraces. At K = 2, we differentiated W-SWF (in cyan) and E-SWF (in blue) genetic groups. At K = 3, 15 landraces (in dark-grey) previously assigned to E-SWF group at K = 2 constituted a group distinguished from E-SWF and W-SWF groups. At K = 4, we observed a fourth group consisting of about 10 landraces located principally in the "Lot" and in "Lot et Garonne" districts. At K = 5, landraces from Gironde (in turquoise) differ from W-SWF groups; Landraces from the "Lot" and "Lot et Garonne" districts were integrated again in E-SWF group and we observed differentiation between landraces from the third group at K = 3.
(TIF)

S3 Fig. Heat-map of Rogers 'genetic distance values between the 194 landraces from SWF and the 82 landraces from America. American landraces were sorted according to the latitude of their collection sites and colored as per their genetic groups previously identified by Camus-Kulandaivelu et al. [10]. We also sorted SWF landraces using their ancestries values on W-SWF group obtained with admixture analysis at K = 2; thus SWF landrace numbers from 0 to 65 represent the E-SWF landrace group and SWF landrace numbers from 66 to 194 represent the W-SWF landrace group. Corn Belt Dent in red, Caribbean in green, Northern Flint in yellow, Mexican in brown, Italian Flint in orange, Andean in magenta and the 9 landraces from South America studied by Mir et al (2017) in black.
(TIF)

S1 Table. Combination of passport data (column 1 to 5), genetic diversity analysis (column 6 and 7) and admixture analysis at K = 2 (column 8 to 10) performing on SNP data of the 194 maize landraces from the South-West of France. W-SWF = "West South-West France»; E-SWF = " East South-West France".
(XLSX)

S2 Table. Passport information and PCA analysis result on SNP data of 342 maize landraces.
(XLSX)

S3 Table. Climatic variations between E-SWF and W-SWF groups from May to October. This table contained means of temperature and precipitation for W-SWF and E-SWF groups from May to October. We shaded in bold months for which differences for mean temperatures and precipitations between the two groups were significant (*p-value*<5%).
(XLSX)

S4 Table. TreeMix 3-population test results. This table represents a summary of TreeMix 3-population test results performed using all combinations of three groups shaped by each of the two groups from SWF (E-SWF and W-SWF) considered as focal group, and two other groups (Reference groups 1 and 2) considered as the two potential ancestors of the focal group. MEX = Mexican, AND = Andean, CAR = Caribbean, NFA = Northern Flint material from America, NFE = Northern Flint material from Europe, CBD = Corn belt dent, ITA = Italian, E-SWF = East South-West France, W-SWF = West South-West France.
(XLSX)

Acknowledgments

We are grateful to all colleagues of GE_2POP staff (UMR AGAP-INRAE, France) for reviewing and orientation, Anne Zanetto (French Maize Biological Ressource Center, Montpellier, France) for providing landraces. The computing task was realized on the CIRAD—UMR AGAP HPC Data Center of the South Green Bioinformatics platform (http://www.southgreen.fr/). We thank the two anonymous reviewers for the high quality of their reviews which lead to substantial clarifications in the interpretation of our results.

Author Contributions

Conceptualization: Alain Charcosset, Joëlle Ronfort, Jacques David, Brigitte Gouesnard.

Formal analysis: Yacine Diaw, Brigitte Gouesnard.

Funding acquisition: Yacine Diaw, Alain Charcosset, Brigitte Gouesnard.

Methodology: Christine Tollon-Cordet, Stéphane D. Nicolas, Delphine Madur.

Resources: Christine Tollon-Cordet.

Supervision: Joëlle Ronfort, Jacques David, Brigitte Gouesnard.

Visualization: Yacine Diaw.

Writing – original draft: Yacine Diaw.

Writing – review & editing: Yacine Diaw, Alain Charcosset, Stéphane D. Nicolas, Joëlle Ronfort, Jacques David, Brigitte Gouesnard.

References

1. Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez GJ, Buckler E, Doebley J. A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci U S A. 2002; 99(9):6080–4. https://doi.org/10.1073/pnas.052125199 PMID: 11983901
2. Freitas FO, Bendel G, Allaby RG, Brown TA. DNA from primitive maize landraces and archaeological remains: implications for the domestication of maize and its expansion into South America. Journal of Archaeological Science. 2003; 30(7):901–8.
3. Vigouroux Y, Glaubitz JC, Matsuoka Y, Goodman MM, Jesus Sanchez G, Doebley J. Population Structure and Genetic Diversity of New World Maize Races Assessed by DNA Microsatellites. American Journal of Botany. 2008; 95(10):1240–53. https://doi.org/10.3732/ajb.0800097 PMID: 21632329

4. Rebourg C, Chastanet M, Gouesnard B, Welcker C, Dubreuil P, Charcosset A. Maize introduction into Europe: the history reviewed in the light of molecular data. Theor Appl Genet. 2003; 106(5):895–903. https://doi.org/10.1007/s00122-002-1140-9 PMID: 12647065

5. Revilla Temiño P, Soengas Fernández MdP, Cartea González ME, Malvar Pintos RA, Ordás Pérez A. Isozyme variability among European maize populations and the introduction of maize in Europe. Maydica. 2003; 48:141–52.

6. Revilla Temiño P, Soengas Fernández MdP, Malvar Pintos RA, Cartea González ME, Ordás Pérez A. Isozyme variation and historical relationships among the maize races of Spain. Maydica. 1998; 43:175–82.

7. Dubreuil P, Warburton M, Chastanet M, Hoisington D, Charcosset A. More on the introduction of temperate maize into Europe: Large-scale bulk SSR genotyping and new historical elements. Maydica. 2006; 51(2):281–91.

8. Rebourg C, Gouesnard B, Charcosset A. Large scale molecular analysis of traditional European maize populations. Relationships with morphological variation. Heredity. 2001; 86:574–87. https://doi.org/10.1046/j.1365-2540.2001.00869.x PMID: 11554974

9. Gauthier P, Gouesnard B, Dallard J, Redaelli R, Rebourg C, Charcosset A, et al. RFLP diversity and relationships among traditional European maize populations. Theor Appl Genet. 2002; 105(1):91–9. https://doi.org/10.1007/s00122-002-0903-7 PMID: 12582566

10. Camus-Kulandaivelu L, Veyrieras JB, Madur D, Combes V, Fourmann M, Barraud S, et al. Maize adaptation to temperate climate: Relationship between population structure and polymorphism in the Dwarf8 gene. Genetics. 2006; 172(4):2449–63. https://doi.org/10.1534/genetics.105.048603 PMID: 16415370

11. Mir C, Zerjal T, Combes V, Dumas F, Madur D, Bedoya C, et al. Out of America: tracing the genetic footprints of the global diffusion of maize. Theor Appl Genet. 2013; 126(11):2671–82. https://doi.org/10.1007/s00122-013-2164-z PMID: 23921956

12. Brandenburg JT, Mary-Huard T, Rigaill G, Hearne SJ, Corti H, Joets J, et al. Independent introductions and admixtures have contributed to adaptation of European maize and its American counterparts. PLoS Genet. 2017; 13(3):e1006666. https://doi.org/10.1371/journal.pgen.1006666 PMID: 28301472

13. Tenaillon MI, Charcosset A. A European perspective on maize history. Comptes Rendus Biologies. 2011; 334(3):221–8. https://doi.org/10.1016/j.crvi.2010.12.015 PMID: 21377617

14. Dubreuil P, Rebourg C, Merlino M, Charcosset A. Evaluation of a DNA pooled-sampling strategy for estimating the RFLP diversity of maize populations. Plant Molecular Biology Reporter. 1999; 17 (2):123–38.

15. Rebourg C, Dubreuil P, Charcosset A. Genetic diversity among maize populations: Bulk RFLP analysis of 65 accessions. Maydica. 1999; 44(3):237–49.

16. Bedoya CA, Dreisigacker S, Hearne S, Franco J, Mir C, Prasanna BM, et al. Genetic diversity and population structure of native maize populations in Latin America and the Caribbean. Plos One. 2017; 12(4). https://doi.org/10.1371/journal.pone.0173488 PMID: 28403177

17. Eschholz TW, Peter R, Stamp P, Hund A. Genetic diversity of Swiss maize (Zea mays L. ssp. mays) assessed with individuals and bulks on agarose gels. Genet Resour Crop Evol. 2008; 55(7):971–83.

18. Reif JC, Hamrit S, Heckenberger M, Schipprack W, Maurer HP, Bohn M, et al. Genetic structure and diversity of European flint maize populations determined with SSR analyses of individuals and bulks. Theor Appl Genet. 2005; 111(5):906–13. https://doi.org/10.1007/s00122-005-0016-1 PMID: 16059732

19. Sharma L, Prasanna BM, Ramesh B. Analysis of phenotypic and microsatellite-based diversity of maize landraces in India, especially from the North East Himalayan region. Genetica. 2010; 138 (6):619–31. https://doi.org/10.1007/s10709-010-9436-1 PMID: 20107870

20. Ganal MW, Durstewitz G, Polley A, Berard A, Buckler ES, Charcosset A, et al. A large maize (Zea mays L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS One. 2011; 6(12):e28334. https://doi.org/10.1371/journal.pone.0028334 PMID: 22174790

21. Arca M, Mary-Huard T, Gouesnard B, Bérard A, Bauland C, Combes V, et al. Deciphering the genetic diversity of landraces with high-throughput SNP genotyping of DNA bulks: methodology and application to the maize 50k array. Frontiers in Plant Science. 2021; 11.

22. Renoux J. Le maïs, une plante pour l'intensification écologique. Agricole F, editor. Paris: France Agricole; 2014. 427 p.

23. Arca M, Gouesnard B, Mary-Huard T, Le Paslier M-C, Bauland C, Combes V, et al. Genome-wide SNP genotyping of DNA pools identifies untapped landraces and genomic regions that could enrich the maize breeding pool. bioRxiv. 2020: 2020.09.30.321018.

24. Dallard J, Noel P, Gouesnard B, Boyat A. A network for the management of genetic resources of maize populations in France. Plant Genetic Resources Newsletter. 2000(123):35–40.

25. INRA, Maize database [Internet]. 2017 [cited 5 march 2020]. Available from: http://xcrop.inra.fr/index.php.

26. Nei M. Estimation of Average Heterozygosity and Genetic Distance from a Small Number of Individuals. Genetics. 1978; 89(3):583–90. PMID: 17248844

27. R Core Team: A language and environment for statistical computing [Internet]. R Foundation for Statistical Computing. 2018. Available from: https://www.R-project.org/.

28. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome research. 2009; 19(9):1655–64. https://doi.org/10.1101/gr.094052.109 PMID: 19648217

29. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology. 2005; 14(8):2611–20. https://doi.org/10.1111/j.1365-294X.2005.02553.x PMID: 15969739

30. Rogers JS. Measures of similarity and genetic distance. Studies in Genetics VII, University of Texas Publication 7213, Austin. 1972:145–53.

31. Wright S. Evolution and the Genetics of Populations. University of Chicago Press, Chicago: Evolutionary Biology; 1978.

32. Rousset F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics. 1997; 145(4):1219–28. PMID: 9093870

33. Weisstein EW. "Spherical Trigonometry." From MathWorld—A Wolfram Web Resource. Feb 2020 [Available from: http://mathworld.wolfram.com/SphericalTrigonometry.html].

34. Nei M. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences of the United States of America. 1973; 70(12):3321–3. https://doi.org/10.1073/pnas.70.12.3321 PMID: 4519626

35. Mantel N. Detection of desease clustering and a generalized regression approach. Cancer Res. 1967; 27(2P1):209–20.

36. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology. 2005; 25(15):1965–78.

37. Goslee S, Urban D. The ECODIST package for dissimilarity-based analysis of ecological data. Journal of Statistical Software. 2007; 22:1–19.

38. Petit R, El Mousadik A, Pons O. Identifying populations for conservation on the basis of genetic markers. Conservation biology. 1998; 12(4):844–55.

39. Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. Nature. 2009; 461(7263):489–94. https://doi.org/10.1038/nature08365 PMID: 19779445

40. Pickrell JK, Pritchard JK. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. Plos Genetics. 2012; 8(11):e1002967. https://doi.org/10.1371/journal.pgen.1002967 PMID: 23166502

41. Labate JA, Lamkey KR, Mitchell SE, Kresovich S, Sullivan H, Smith JSC. Molecular and historical aspects of corn belt dent diversity. Crop Science. 2003; 43(1):80–91.

42. Segovia-Lerma A, Cantrell RG, Conway JM, Ray IM. AFLP-based assessment of genetic diversity among nine alfalfa germplasms using bulk DNA templates. Genome. 2003; 46(1):51–8. https://doi.org/10.1139/g02-100 PMID: 12669796

43. Raymond M, Rousset F. GENEPOP (version 1.2): population genetic software for exact tests and ecumenicism. J Hered. 1995; 86:248–9.

44. Derryberry EP, Derryberry GE, Maley JM, Brumfield RT. hzar: hybrid zone analysis using an R software package. Mol Ecol Resour. 2014; 14:652–63. https://doi.org/10.1111/1755-0998.12209 PMID: 24373504

45. Pritchard JK, Stephens M, Donnelly P. Inference of Population Structure Using Multilocus Genotype Data. Genetics. 2000; 155(2):945–59. PMID: 10835412

46. Charrier O, Dupont P, Pornon A, Escaravage N. Microsatellite marker analysis reveals the complex phylogeographic history of Rhododendron ferrugineum (Ericaceae) in the Pyrenees. PloS one. 2014; 9(3):e92976–e. https://doi.org/10.1371/journal.pone.0092976 PMID: 24667824

47. Conord C, Gurevitch J, Fady B. Large-scale longitudinal gradients of genetic diversity: a meta-analysis across six phyla in the Mediterranean basin. 2012; 2:2600–14. https://doi.org/10.1002/ece3.350 PMID: 23145344

48. Oromi N, Richter-Boix A, Sanuy D, Fibla J. Genetic variability in geographic populations of the natterjack toad (Bufo calamita). Ecology and evolution. 2012; 2(8):2018–26. https://doi.org/10.1002/ece3.323 PMID: 22957202

49. Valbuena-Ureña E, Oromi N, Soler-Membrives A, Carranza S, Amat F, Camarasa S, et al. Jailed in the mountains: Genetic diversity and structure of an endemic newt species across the Pyrenees. PLOS ONE. 2018; 13(8):e0200214. https://doi.org/10.1371/journal.pone.0200214 PMID: 30071027

50. Beigbeder J, Carraretto M. les maïs anciens des Pyrenees: savoir et savoir-faire traditionnels. Marrimpouey ed. Pau2018. 121 p.

51. Vouette I. Millet, Buckwheat, Maize and Sorghum: minor grains in the old country system (France, mid. XVIth century–XIXth century): Université Paris-Diderot—Paris VII; 2007.

52. Frêche G. Une enquête sur les prix des produits agricoles dans la région toulousaine: XVe-XIXe siècle. Annales du Midi. 1969:17–39.

53. Mauro F. L'Espagne et le Midi de la France au XVIIe siècle: Aspects économiques et humains. Caravelle Cahiers du monde hispanique et luso-brésilien. 1968:151–62.

54. Doebley J, Wendel JD, Smith JSC, Stuber CW, Goodman MM. The origin of Cornbelt maize—the isozyme evidence. Economic Botany. 1988; 42(1):120–31.

55. Carraretto M. Histoire du maïs, d'une divinité amérindienne à ses avatars transgéniques. CTHS ed. Paris2005. 259 p.

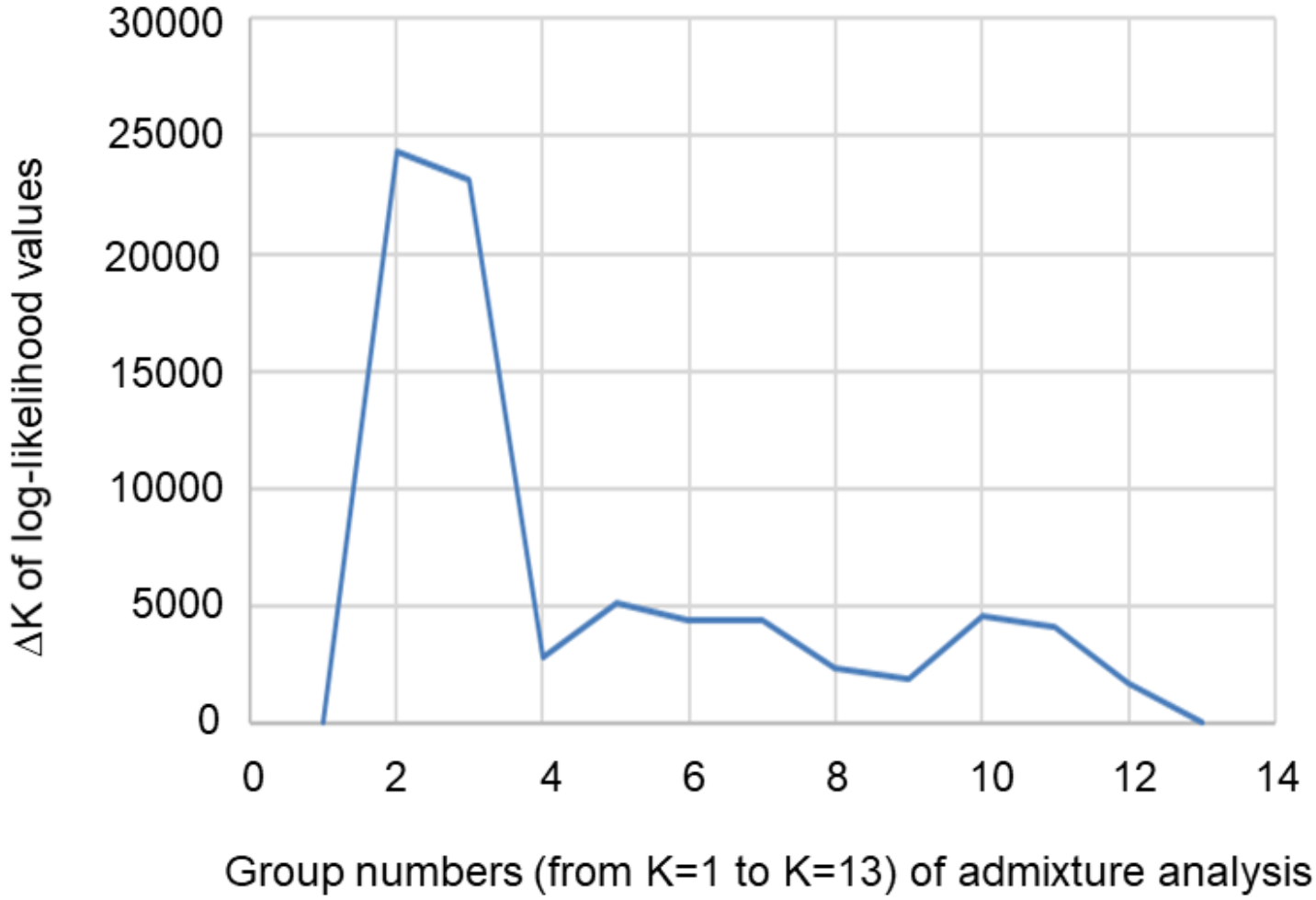
56. Bitocchi E, Nanni L, Rossi M, Rau D, Bellucci E, Giardini A, et al. Introgression from modern hybrid varieties into landrace populations of maize (Zea mays ssp mays L.) in central Italy. Molecular Ecology. 2009; 18(4):603–21. https://doi.org/10.1111/j.1365-294X.2008.04064.x PMID: 19215582

57. Patto MCV, Satovic Z, Pêgo S, Fevereiro P. Assessing the genetic diversity of Portuguese maize germplasm using microsatellite markers. Euphytica. 2004; 137(1):63–72.

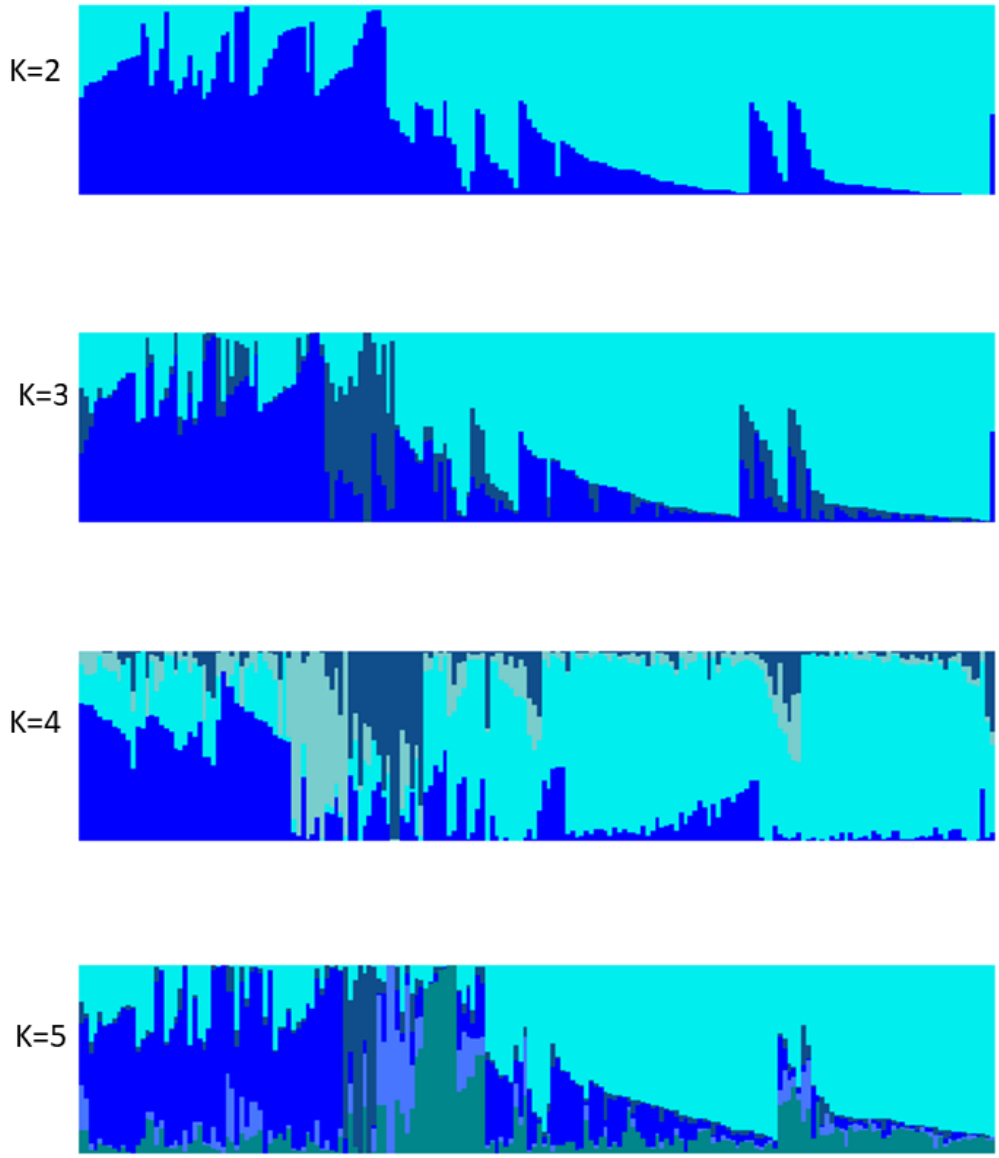
58. Rouch M. L'arrivée et l'implantation des Italiens dans le sud-ouest (1920–1939). École Française de Rome. 1986; 94:693–720.

SUPPLEMENTARY FIGURES

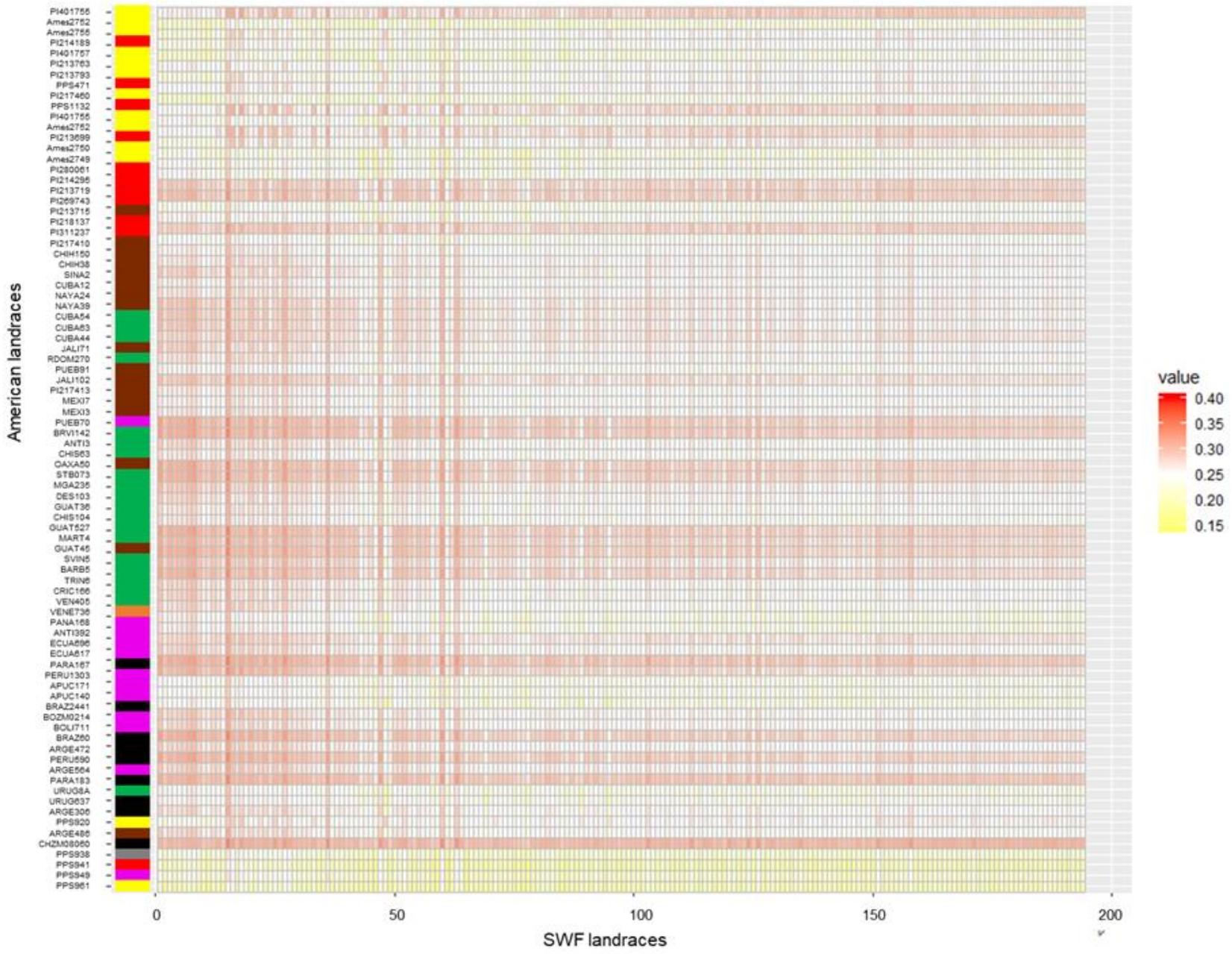
Genetic diversity of maize landraces from the South-West of France.



S1 Fig. Plot of DeltaK of the log likelihood from the admixture analysis on SWF landraces SNP data. *Group numbers varied from K=1 to K=13.*



S2 Fig. Bar-plot for K=2 to K=5 for admixture analysis with the 194 SWF maize landraces. *At K=2, we differentiated W-SWF (in cyan) and E-SWF (in blue) genetic groups. At K=3, 15 landraces (in dark-grey) previously assigned to E-SWF group at K=2 constituted a group distinguished from E-SWF and W-SWF groups. At K=4, we observed a fourth group consisting of about 10 landraces located principally in "Lot" and in "Lot et Garonne" districts. At K=5, landraces from Gironde (in turquoise) differ from W-SWF groups; Landraces from "Lot" and "Lot et Garonne" districts were integrated again in E-SWF group and we observed differentiation between landraces from the third group at K=3.*



S3 Fig. Heat-map of Rogers 'genetic distance values between the 194 landraces from SWF and the 82 landraces from America. *American landraces were sorted by latitude of their collection sites and colored as per their genetic groups previously identified by Camus-Kulandaivelu et al. (2006). We also sorted SWF landraces using their ancestries values on W-SWF group obtained with admixture analysis at K=2; thus SWF landrace numbers from 0 to 65 represent the E-SWF landrace group and SWF landrace numbers from 66 to 194 represent W-SWF landrace group. Corn Belt Dent in red, Caribbean in green, Northern flint in yellow, Mexican in brown, Italian Flint in orange, Andean in magenta and the 9 landraces from South America studied by Mir et al (2017) in black.*

CHAPITRE II

Diversité cytoplasmique des populations de maïs du Sud-Ouest de la France.

3.1. Introduction

Les études sur la diversité et les liens génétiques entre les populations d'espèce animale ou végétale sont le plus souvent basées sur l'analyse du génome nucléaire. Cependant ce type d'ADN n'est pas le seul qui existe dans une cellule végétale ; hors du noyau nucléaire, il existe aussi les génomes des organites qui peuvent être analysés pour connaitre les relations génétiques entre les populations ou les individus d'une espèce animale ou végétale. Contrairement à l'ADN nucléaire qui est transmis par les voies mâle et femelle (pollen et graine) lors de la reproduction, le génome des organites est dans la plupart des cas transmis de manière uni-parentale (Birky, 1995; Conde et al., 1979; Ennos, 1994; Neale et al., 2011; Neale and Sederoff, 1989; Ouborg et al., 1999; Reboud and Zeyl, 1994). La comparaison de la structure génétique de la diversité dans le génome nucléaire et dans celui des organites cytoplasmiques permet donc d'analyser les contributions relatives des flux de pollen et de graines dans les populations (Provan et al., 2001).

Chez les plantes, le séquençage de l'ADN chloroplastique a permis le développement d'amorces universelles pour amplifier des séquences de microsatellites spécifiques des chloroplastes (Powell et al., 1995a; Powell et al., 1995b; Provan et al., 1999b). Ces marqueurs microsatellites chloroplastiques (cpSSR) ont été utilisés pour étudier la diversité et la structure génétique de populations (Ceylan et al., 2014; Ferreira et al., 2015; Minn et al., 2016; Pazouki et al., 2016; Phumichai et al., 2015; Takahashi et al., 2016; Thia et al., 2016; Wang et al., 2011), l'origine génétique et la phylogéographie d'un nombre important d'espèces végétales (Aoki et al., 2016; Chmielewski et al., 2015; Lopez-Caamal et al., 2019; Peng and Chen, 2011), les phénomènes d'hybridation intra-spécifique (Bucci et al., 1998; Cubas et al., 2005; Dzialuk et al., 2009; Semerikova et al., 2011) ou les ratio de flux dus au pollen et aux graines sur la dispersion des espèces (Powell et al., 1995b; Wang et al., 2011; Yang et al., 2008).

L'avantage d'utiliser les cpSSR dans les études de diversité est lié au fait que l'ADN du chloroplaste est supposé être non-recombinant (Marshall et al., 2001) et présente un taux de mutation très faible comparé à l'ADN nucléaire (Provan et al., 1999d). Le taux d'évolution du génome chloroplastique est donc attendu plus faible. Certaines caractéristiques des marqueurs cpSSR peuvent cependant poser problème dans les études de diversité : (1) l'homoplasie définie par la présence à un locus d'allèles de même taille mais issus d'évènements de mutations parallèles mais différents (Estoup et al., 2002; Garza and Freimer, 1996; Navascues and Emerson, 2005), l'hétéroplasmie issue du mélange de génomes chloroplastiques chez les espèces dont les organites sont transmis par les deux parents (Chat et al., 2002; Frey et al., 2005;

Hansen et al., 2007; Scarcelli et al., 2016; Wheeler et al., 2014) et les introgressions cytoplasmiques d'organites chloroplastiques observées suite à certaines hybridations interspécifiques (Provan et al., 1999c; Wheeler et al., 2014). Si on ne tient pas compte de ces phénomènes, l'analyse de données cpSSR peut conduire à un certain nombre d'erreurs dans l'interprétation des liens de parenté entre les populations d'une même espèce (Chmielewski et al., 2015; Hale et al., 2004). Etant donné le faible taux de mutation observé pour les marqueurs de type cpSSR (Provan et al., 1999d), l'impact de l'homoplasie de taille dans les analyses de diversité intra-spécifique est généralement négligé et l'influence des deux autres phénomènes est supposée négligeable pour des études à l'échelle intra-spécifique. Par ailleurs, le génotypage de plusieurs locus SSR permet de définir des haplotypes chloroplastiques combinant la variabilité allélique de plusieurs loci (Garza and Freimer, 1996; Navascues and Emerson, 2005). L'analyse des relations génétiques entre ces haplotypes à travers la construction de réseaux permet de détecter certains phénomènes d'homoplasie et d'en tenir compte dans l'interprétation des résultats (Besnard et al., 2011; Chmielewski et al., 2015; Hatziskakis et al., 2009; Yang et al., 2008; Yu et al., 2010).

Chez le maïs, les chloroplastes sont transmis uniquement via les graines donc par la voie maternelle (Aoki et al., 2016; Conde et al., 1979; Powell et al., 1995a). La diffusion du maïs cultivé se faisant via les semences, l'analyse de la diversité chloroplastique devrait donc nous renseigner sur les grandes voies de diffusion du maïs et sur les pratiques d'échanges de semences. Pour mieux connaitre l'origine génétique des populations du Sud-Ouest de la France nous proposons d'analyser la diversité et la structure génétique de 334 accessions de maïs du genre *Zea* en combinant l'information obtenue à l'aide de 7 marqueurs cpSSR. Ces accessions sont composées de 15 accessions de maïs sauvage issues des 5 sous-espèces de téosintes et de 319 populations de maïs cultivé d'Amérique et d'Europe. Les objectifs de cette étude sont (1) d'évaluer la structure génétique des populations de maïs collectées dans le Sud-Ouest de la France, (2) d'analyser les liens génétiques entre les deux groupes de maïs identifiés au Sud-Ouest de la France par Diaw et al. (2021) et les 7 groupes de maïs identifiés par Camus et al (2006) et (3) de comprendre l'origine génétique des populations du Sud-Ouest de la France. En dehors de notre étude, à notre connaissance, seuls Provan et al. (1999b) et Pressoir and Berthaud (2004) ont utilisé des marqueurs cpSSR pour étudier la diversité chez le maïs. Ces études concernaient respectivement des accessions des espèces et sous-espèces du genre *Zea* et des populations locales de maïs mexicains.

3.2. Matériel et méthodes

3.2.1. Matériel génétique

Nous avons analysé 334 accessions de maïs représentant les différentes espèces et sous espèces du genre *Zea* à l'aide de 7 marqueurs microsatellites chloroplastiques (cpSSR). Cette base de données est composée d'une accession de chacune des espèces suivantes : Zea *mays subsp. perennis*, Zea *mays subsp. diploperennis*, Zea *mays subsp. luxurians*, et Zea *mays subsp. huehuetenangensis* ; 4 accessions de Zea *mays subsp mexicana*, 7 accessions de Zea *mays subsp. parviglumis et* 319 accessions de *Zea mays subsp. mays* (Tableau S1 dans Annexe 2). Les 15 premières accessions sont des téosintes et les 319 dernières accessions sont des populations de maïs cultivé d'Europe et d'Amérique qui contiennent : 125 populations représentants les grands groupes génétiques de référence du maïs cultivé identifiés par Camus-Kulandaivelu et al. (2006) ; et 194 populations collectées dans le Sud-Ouest de la France représentant deux groupes génétiques identifiés à l'aide de marqueurs SNPs nucléaires par Diaw et al. (2021) (chapitre 1). Pour chaque population, un seul individu a été échantillonné et analysé.

3.2.2. Amplification par PCR

Deux amorces (M13 et T7) ont été dessinées pour amplifier 7 cpSSR sur l'ADN du chloroplaste chez les populations de maïs (Tableau 1). Les 7 loci cpSSR sont composés de 6 loci identifiés sur le chloroplaste du maïs (Provan et al., 1999b) et un locus homologue à RCt5 du riz (Ishii and McCouch, 2000). L'information sur leur séquence est donnée dans le Tableau 1. Les « amorces sens » ont été associées avec une queue fluorescente facilitant la lecture du locus cpSSR lors de l'analyse. La PCR a été effectuée dans un volume final de 20 µl de produit réactif contenant 1 Tampon PCR (10 nM Tris-HCP, 50 mM KCl, ph 8,3), 5-20 ng d'ADN, 200 uM de chaque dNTPs, 2 mM de MgCl2, 0.4 pmol amorce M13, 4 pmol d'amorces non marqués, 4 pmol Queuefluo et 0.5 U de Taq polymérase. La lecture des résultats de PCR a été effectuée en utilisant le logiciel GeneMapper v.5 (de « Applied Biosystems »).

Tableau 1 : information sur loci cpSSR et amorces utilisées pour analyser les 319 populations de maïs.

locus	Localisation	MOTIF	Amorces	Espèce d'origine	Taille allèles de référence (en bp)	Tm (°C)	Référence
RCt5	région intergénique	(T)10 (T)5 AC(T)3	ATTTGGAATTTGGACATTTTCG ACTGATTCGTAGGCGTGGAC	riz	143	55	Ishii and McCouch (2000)
ZMCP3764	intron trnK	(A)12	TTTACCAGAATTAGAGCCAGGC CCACGACTGATCCTCAAAGG	maïs	144	60	Provan et al. (1999b)
ZMCP7430	région intergénique psbK-psbI	(A)11	GCAAGCTGCTGTAAGTTTTCG AAGACTTCTCGGCTCTTATCCA	maïs	80	60	
ZMCP12968	région intergénique trnG-trnfM	(T)18	TTTTGGAAAGGGGGATAGGCT GTCTCCGCACTTAGCAGTCC	maïs	148	60	
ZMCP18704	région intergénique psbM-ORF29	(A)15	AAGGTGCGAGATGCTTCAAT ACAATATATCGCACATTTTCCA	maïs	107	60	
ZMCP20824	région intergénique trnC-rpoB	(A)11	TGGGGAATCGTGCTTTAATT TCGTAATTGCACTTAGGTGGG	maïs	120	60	
ZMCP59488	région intergénique psaI-ORF185	(T)11	GCGTAGCCGACTATAGCGAG CAGTAGCGGTCAATTGAAATTG	maïs	172	60	

3.2.3. Analyse de la diversité des cpSSR

Dans un premier temps, nous avons analysé la variabilité génétique pour chaque marqueur cpSSR en estimant le nombre d'allèles et la diversité moyenne H_e à l'aide de la formule de Nei (1987). Nous utilisons le terme allèles pour décrire les variants observés pour chaque marqueur cpSSR (encore appelé locus). Pour chaque individu analysé, on définit un haplotype en combinant l'information observée aux 7 marqueurs cpSSR. Les haplotypes sont nommés en utilisant des combinaisons de lettres alphabétiques. Une fois les haplotypes définis, le génotype de chacun des individus représentant les 319 populations et les 15 accessions de téosinte est affecté à un des haplotypes identifiés.

3.2.4. Relations génétiques entre les haplotypes

Pour étudier l'origine génétique des populations du Sud-Ouest de la France, nous pouvons nous limiter à l'analyse de la présence et de l'absence des différents haplotypes chez les 319 populations de maïs utilisées dans cette étude. Cependant, une telle analyse ne tient pas compte des proximités génétiques qui existent entre ces haplotypes. Pour représenter leurs proximités génétiques, nous avons utilisé un algorithme de réseau de jonction médiane (ou « MJ » en anglais) disponible sous NETWORK version 5.0.1.1 (Flexus Technology Ltd, http://www.fluxus-engineering.com). Ce logiciel est généralement utilisé pour reconstruire les liens phylogénétiques entre différentes espèces mais aussi pour déduire les types ancestraux et les formes dérivées au sein d'un ensemble d'haplotypes. Les haplotypes ancestraux sont supposés être les haplotypes présents chez les populations ancestrales de maïs. Les haplotypes dérivés sont les haplotypes qui dérivent des haplotypes ancestraux suite à des événements de mutation sur ces haplotypes ancestraux. Pour déterminer les haplotypes ancestraux chez le maïs, nous avons intégré les 37 accessions d'espèces et sous-espèces du genre *Zea* étudiées par Provan et al. (1999b) dans notre réseau d'haplotypes (Tableau S2 dans Annexe 2). L'intégration de cette base de données a nécessité un certain nombre d'ajustements sur les données cpSSR. En effet, les conditions d'analyse en laboratoire ne sont pas identiques dans les deux études. De plus, le marqueur RCt5 n'avait pas été utilisé par Provan et al. (1999b). Pour effectuer ces ajustements sur les données cpSSR, nous avons comparé les résultats du séquençage de 3 accessions de téosinte sur la base de 6 marqueurs cpSSR (sauf la RCt5), présentes à la fois dans notre étude et dans celle de Provan et al. (1999a) : 1 de Zea *mays subsp. perennis*, 1 de Zea *mays subsp. diploperennis* et 1 de Zea *mays subsp. huehuetenangensis* (Tableau S2 dans Annexe 2).

3.2.5. Analyse cpSSR basée sur les groupes génétiques de maïs

Notre analyse porte sur 319 populations de maïs. Pour analyser la diversité haplotypique de ces populations, nous nous sommes basés sur les 7 groupes génétiques identifiés par Camus-Kulandaivelu et al. (2006) à l'aide de marqueurs SSR nucléaires et les 2 groupes génétiques identifiés dans le Sud-Ouest de la France à l'aide de marqueurs SNP dans le Chapitre1. Pour chacun de ces groupes définis a *priori*, nous avons analysé la diversité des haplotypes en utilisant le logiciel HAPLOTYPE ANALYSIS version 1.05 (Eliades and Eliades, 2009) pour estimer le nombre d'haplotypes présents dans chaque groupe (*A*), le nombre d'haplotypes spécifiques à chaque groupe (*P*), leur diversité au travers de l'indice de l'hétérozygotie (H_e) de Nei (1973) :

$$He(g) = \frac{N_{(g)}}{N_{(g)} - 1}\left(1 - \sum_{i=1}^{N_h} \bar{P}_{(i,g)}^2\right)$$

Où $\bar{P}_{(i,g)}$= fréquence de l'haplotype (i) dans le groupe g

N_h = nombre d'haplotypes observés dans le groupe g

$N_{(g)}$ = nombre de populations dans le groupe g;

Pour tenir compte du fait que les locus microsatellites mutent selon un modèle de mutation par pas (« Stepwise Mutation Model », SMM), la distance génétique (D^2sh) de Goldstein et al. (1995b) entre les individus au sein du même groupe génétique g a été calculée comme suit

$$D_{sh}^2(g) = \frac{2}{N_{(g)} * \left(N_{(g)} - 1\right)} * \frac{1}{L} * \sum_{x=1}^{N_{(g)}} \sum_{y=x+1}^{N_{(g)}} d_{(x,y)}^2$$

$$d_{(x,y)} = \sum_{k=1}^{L} \left| a_{(x,k)} - a_{(y,k)} \right|$$

Où $a_{(i,k)}$= taille de l'allèle du locus (k) pour l'individu *i*

$a_{(j,k)}$ = taille de l'allèle du locus (k) pour l'individu *j*

$d_{(i,j)}$ =distance génétique estimé entre les individus *i* et *j*, en additionnant les différences du nombre de bases des deux allèles identifiés sur chaque locus chez les deux individus ;

L = nombre de loci.

Nous avons exclu du calcul de la distance de Goldstein les loci dont les différences de taille des allèles ne répondent pas à la norme de SMM ; c'est-à-dire lorsqu'ils présentent des variations de plusieurs paires de base issues d'insertion ou de délétion (Ohta and Kimura, 1973).

Par la suite, pour analyser les proximités génétiques entre ces 9 groupes génétiques avec les 7 marqueurs cpSSR, nous avons calculé les différentiations génétiques (F_{ST}) en utilisant la fonction pairwise.WCfst du package hierfstat (Goudet, 2005) disponible sous R. Nous avons aussi utilisé la fonction boot.ppfst du même package pour estimer l'intervalle de confiance à 95% associé à chaque valeur de F_{ST}.

3.3. Résultats

3.3.1. Variabilité des marqueurs SSR chloroplastiques

Le Tableau 2 donne le résultat du produit de l'amplification obtenu sur les 15 accessions de téosintes et les 319 populations de maïs utilisées dans cette étude. En utilisant 7 marqueurs cpSSR sur l'ensemble de ces populations, nous avons détecté entre 2 et 8 allèles par locus. La combinaison des différents allèles observés aux 7 marqueurs cpSSR permet de définir 21 haplotypes. Parmi eux, les haplotypes AD (36.68% des 319 populations analysées), AA (25.38%), F (11.62%), I (10.09%) et G (8.25%) sont les haplotypes les plus fréquents dans les populations de maïs (Tableau 2). Dans notre échantillonnage, les haplotypes (ZL01, ZP01 et ZD01) observés chez les accessions de maïs de l'espèce *Zea sect. luxuriantes* sont spécifiques à cette section puisqu'ils n'ont pas été détectés chez les populations de maïs cultivé et dans les accessions de téosinte de la section Zea *sect. Zea*. Par contre, nous avons observé que les haplotypes AD et X présents chez des accessions des sous-espèces Zea mays subsp. *parviglumis* et Zea mays subsp. *mexicana* le sont aussi chez des populations de maïs cultivées (Tableau 2).

La variation entre les allèles au sein des cpSSR est conforme au modèle de mutations par pas (Tableau 2). En effet, les marqueurs présentent en grande partie une distribution uni-modale symétrique avec des longueurs différant les unes des autres d'une ou de deux unités de répétition à la fois. Seul le RCt5, un marqueur homologue du riz, fait exception à la règle du modèle « SMM ». On observe en effet un polymorphisme d'insertion /délétion de 16 bases. Nous avons exclu ce locus pour estimer le paramètre D^2sh de Goldstein et al. (1995a).

Tableau 2 : Variabilité allélique des 7 loci cpSSR étudiés chez le maïs, liste des haplotypes identifiés et nombre d'individus analysés portant chacun de ces haplotypes.

Haplotypes	**cpSSR locus**							**Nombre d'individus**
	RCT5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488	
MP01([a])	190	163	124	165	124	136	200	1
MP04([a])	190	161	125	165	125	136	200	1
MP06([a])	190	162	125	165	124	136	200	2
MH01([b])	190	161	121	165	121	136	200	1
X([acd])	190	162	124	165	124	136	200	7
AD([cd])	190	163	125	165	125	136	200	122
AA([d])	190	163	124	165	125	136	200	83
AB([d])	190	163	125	164	125	136	200	6
AI([d])	174	161	120	168	122	138	201	1
D([d])	174	161	120	166	124	138	201	3
F([d])	174	161	120	167	122	138	201	38
G([d])	174	161	120	167	125	138	201	27
I([d])	174	161	120	168	125	138	201	33
J([d])	174	161	120	168	125	138	202	1
K([d])	174	161	120	169	125	138	201	4
N([d])	174	161	125	166	125	136	201	6
P([d])	174	162	127	167	125	136	201	2
Q([d])	174	162	127	168	125	136	201	1
ZD01([e])	174	160	120	162	120	136	199	1
ZP01([f])	174	160	120	163	120	136	199	1
ZL01([g])	174	160	121	162	120	136	199	1
nombre d'allèles	2	4	5	8	5	2	4	

Haplotypes identifiés chez ([a]) les 7 accessions de Zea mays subsp. parviglumis, ([b]) l'accession de Zea mays subsp. huehuetenangensis, ([c]) les 4 accessions de Zea mays subsp. mexicana, ([d]) les populations de maïs cultivée, ([e]) l'accession de Zea mays subsp. diploperennis, ([f]) l'accession de Zea mays subsp. perennis et ([g]) l'accession de Zea mays subsp. luxurians.

3.3.2. Le réseau d'haplotypes

On observe que la taille des allèles identifiés sur les 3 accessions de téosinte dans l'étude de Provan et al. (1999b) est différente de celle observée dans notre étude. Nous avons considéré que ces différences sont dues au fait que les deux analyses ont été faites dans des conditions d'analyse différentes. En nous basant sur les génotypes communs entre l'étude de Provan et al. (1999b) et la nôtre, nous avons obtenu les correspondances entre les tailles observés dans les deux études pour les 6 cpSSR analysés en commun. En nous basant sur le décalage de taille observé (Tableau S3 dans Annexe 2), nous avons adapté notre base de données cpSSR à celle de Provan et al. (1999b).

La Figure 1 donne les résultats de l'analyse des liens génétiques entre les 42 haplotypes identifiés avec les 6 loci cpSSR chez les 334 accessions du genre Zea et les 37 accessions de Provan et al. (1999b). La construction du réseau a nécessité l'utilisation de 9 vecteurs médians (points rouges nommés mv1 à mv9 sur Figure1). La présence de ces nombreux vecteurs médians sur le réseau ne facilite pas la distinction entre les haplotypes ancestraux et les haplotypes dérivés chez le maïs cultivé. Ces vecteurs médians représentent des haplotypes hypothétiques qui auraient existé transitoirement au sein du genre *Zea* ou bien que nous n'avons pas échantillonné. Cependant, le réseau des haplotypes montre l'existence de 3 grands groupes d'haplotypes que nous appellerons haplogroupes. Le premier haplogroupe est formé par les accessions représentant les espèces *Zea* mays subsp. *luxurians*, *Zea* mays subsp. *diploperennis*, *Zea* mays subsp. *perennis* et *Zea* mays subsp. *huehuetnangensis*. Le deuxième haplogroupe est composé uniquement de populations de maïs cultivées alors que l'haplogroupe 3 réunit des populations de maïs cultivées et les accessions des téosintes Zea mays subsp. *mexicana* et Zea mays subsp. *parviglumis.*

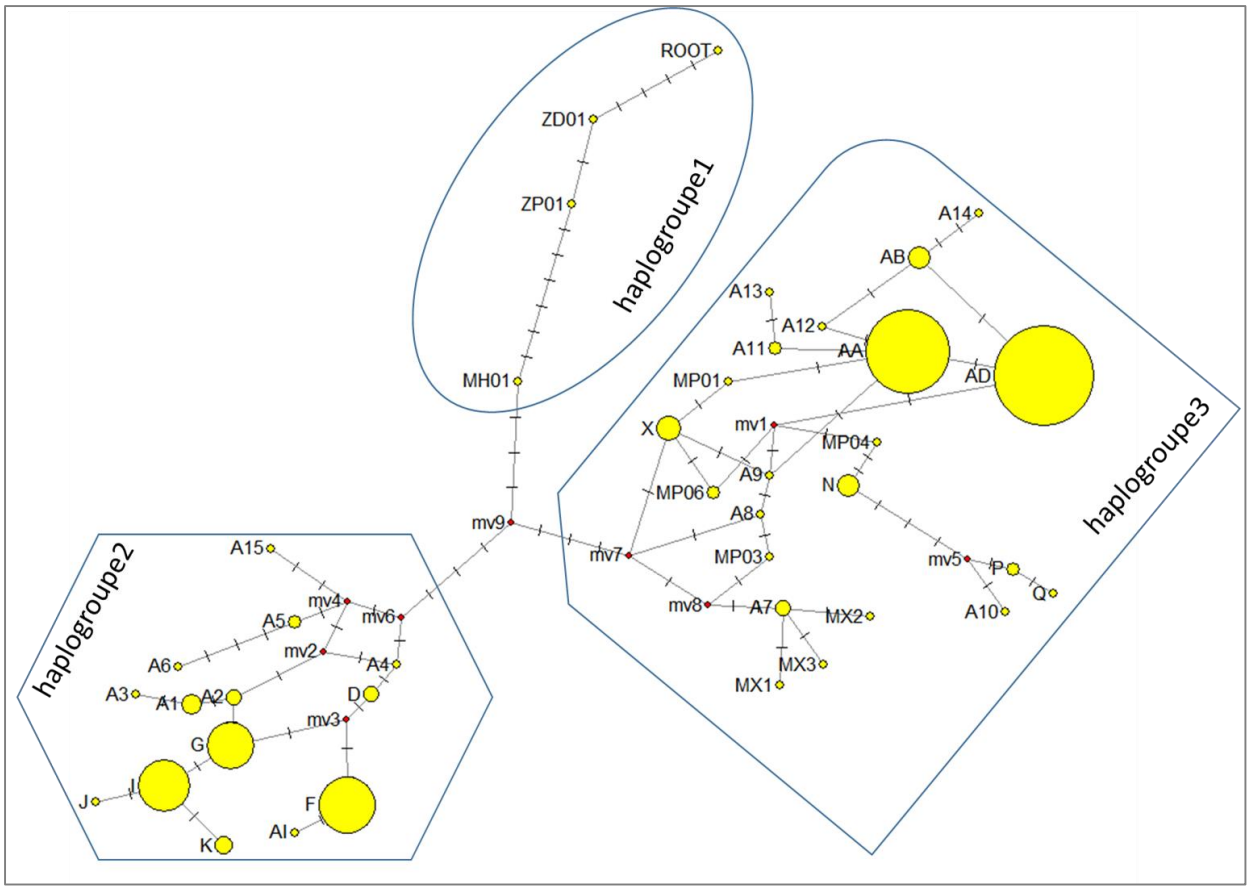


Figure 1 : Réseau d'haplotypes portant sur l'ensemble des haplotypes identifiés au sein des populations de maïs et les accessions de téosinte à partir de la combinaison des variants alléliques observés à l'aide de 6 marqueurs cpSSR. *Les points rouges sur le réseau nommés mv1, mv2, …, mv9 sont des vecteurs médians nécessaire pour mettre des liaisons entre tous les haplotypes. Les haplotypes sont représentés par les cercles jaunes. La taille de chaque cercle est proportionnelle au nombre de populations dont le chloroplaste présente l'haplotype qui lui est associé. Les abréviations à côté des cercles jaunes représentent les haplotypes identifiés chez Zea mays subsp. luxurians (ROOT, cette accession est celle utilisée pour enraciner l'arbre), chez Zea mays subsp. diploperennis (ZD01), chez Zea mays subsp. perennis (ZP01), chez Zea mays subsp. huehuetenangensis (MH01), chez Zea mays subsp. parviglumis (avec le préfixe MP dont le MP02 identifié aussi chez une population de maïs cultivé), chez Zea mays subsp. mexicana (avec le préfixe MX). En dehors de ces abréviations, tous les autres haplotypes sont ceux identifiés uniquement chez les populations de maïs cultivé ; excepté l'haplotype X qui a aussi été identifié chez 3 accessions de Zea mays subsp. parviglumis et 2 Zea mays subsp. mexicana ; et l'haplotype AD identifié chez 2 Zea mays subsp. mexicana. Les informations complémentaires sur ces haplotypes sont disponibles dans les Tableaux S1 et S2 dans Annexe 2.*

3.3.3. Diversité chloroplastique des groupes génétiques représentatifs de l'histoire du maïs cultivé

Les Tableaux 3 et 4 ci-dessous présentent respectivement la répartition des haplotypes chloroplastiques dans les différents groupes génétiques de référence du maïs et les deux groupes génétiques identifiés dans le Sud-Ouest de la France et la diversité génétique observée dans ces différents groupes. Ces groupes sont: le groupe « Mexican » (MEX), « Andean » (AND), « Caribbean » (CAR), « Northern flint » (NF), « Corn Belt Dent » (CBD), « Pyrenees-Galicia2 », « Italian » (ITA), « Eastern South-West France » (E-SWF) et « Western South-West France » (W-SWF). On observe que les haplotypes les plus fréquents (AA, F, I, G et AD) sont présents dans presque tous les groupes génétiques. Les populations de maïs des groupes CAR et AND présentent majoritairement l'haplotype AD (Haplogroupe 2, Tableau 3) alors que les populations du groupe NF présentent majoritairement l'haplotype I (Haplogroupe 3, Tableau 3). Ce dernier est un haplotype qui est absent des populations MEX et faiblement présent dans les populations « Pyrenees-Galicia2 », E-SWF et W-SWF. Dans ces 3 groupes génétiques, les haplotypes AD et AA sont les plus fréquents (Tableau 3). Au sein du maïs cultivé, seulement quatre haplotypes apparaissent spécifiques d'un groupe génétique : l'haplotype X est spécifique du groupe AND, l'haplotype Q du groupe MEX, l'haplotype J du groupe CBD et l'haplotype AI pour NF (Tableau 3). Un résultat notable du tableau 3 est qu'aucun groupe de populations ne contient des haplotypes appartenant à un seul haplogroupe défini ci-dessus, tous les groupes contiennent des haplotypes appartenant aux haplogroupes 2 et 3. En comparant les populations de l'Europe et de l'Amérique, nous observons que les haplotypes AI et D ne sont présents qu'en Europe (Tableau S1 dans Annexe 2). Les haplotypes K (majoritairement présent au Nord de l'Europe) et AI (présent chez une population de Pologne) sont absents chez les populations des groupes E-SWF et W-SWF.

La diversité chloroplastique des groupes génétiques de référence est en moyenne élevée (He > 0.6 ; entre 4 et 7 haplotypes par groupe), seul le groupe ITA présente une diversité vraiment plus faible (Tableau 4). Parmi les groupes de référence du maïs, on note que les plus forts niveaux de diversité sont présents dans les groupes originaires du centre et du sud de l'Amérique, avec notamment le groupe MEX présentant la plus grande valeur de D^2sh. Les deux groupes du Sud-Ouest de la France présentent également des niveaux de diversité importants, la diversité haplotypique étant un peu plus élevée dans le groupe de l'Est, avec des valeurs de He et D^2sh plus forte mais aussi plus d'haplotypes dans ce groupe malgré une taille

d'échantillon deux fois plus faible. Le tableau 4 indique également que ces deux groupes du Sud-Ouest de la France ne présentent pas d'haplotype spécifique, toutes les formes de cpSSR révélées existent dans les populations des groupes de référence du maïs que nous avons étudiées (Tableau 4).

Tableau 3 : répartition des haplotypes au sein des 7 groupes génétiques représentant l'histoire du maïs cultivé (Camus-Kulandaivelu et al., 2006) et des deux groupes génétiques identifiés dans le sud-ouest de la France (chapitre 1).

Groupe génétique	Nombre de populations	Haplogroupe 3							Haplogroupe 2						
		X	AA	AD	AB	N	P	Q	G	I	J	K	D	F	AI
MEX	20		2	9	1			1	4			1		2	
AND	9	1		4	1		1		2						
CAR	21			10	4	1			4	1				1	
NF	36		3	6					1	19		3		3	1
CBD	15			4		1			7	2	1				
Pyrenees_galicia2	9		3	4						1			1		
E-SWF	65		20	23		4			5	7			2	4	
W-SWF	129		55	56						1				17	
ITA	15						1		1	2				11	

AND= « Andean », CAR= « Caribbean », CBD= « Corn Belt Dent », ITA= « Italian », MEX= « Mexican », NF= « Northern flint », E-SWF= « Eastern South-West France », W-SWF= « Western South-West France ».

Tableau 4 : Diversité estimée sur les haplotypes chloroplastiques identifiés à l'aide de 7 marqueurs SSR pour les 7 groupes génétiques représentant l'histoire du maïs cultivé (Camus-Kulandaivelu et al., 2006) et les deux groupes génétiques identifiés dans le sud-ouest de la France (chapitre 1).

Groupe génétique	N	A	P	H_e	D^2sh
AND	9	5	1	0.806	50.246
CAR	21	6	0	0.729	52.279
CBD	15	5	1	0.733	47.121
Pyrenees-Galicia2	9	4	0	0.750	43.889
ITA	15	4	0	0.467	3.611
MEX	20	7	1	0.768	59.224
NF	36	7	1	0.690	46.974
E-SWF	65	7	0	0.766	48.990
W-SWF	129	4	0	0.617	32.010
Valeurs moyennes		**5.444**		**0.703**	**42.705**

N= nombre de populations ; A = nombre d'haplotypes ; P=nombre d'haplotypes spécifiques (ou haplotypes identifiés que chez les populations d'un seul groupe) ; He=hétérozygotie de Nei ; D^2sh= distance génétique estimé sur la base des 6 loci cpSSR (sans tenir compte du locus RCt5 car celui-ci n'est pas compatible avec un modèle de mutation SMM) avec la méthode de Goldstein (Goldstein et al., 1995b). AND= « Andean », CAR= « Caribbean », CBD= « Corn Belt Dent », ITA= « Italian », MEX= « Mexican », NF= « Northern flint », E-SWF= « Eastern South-West France », W-SWF= « Western South-West France ».

3.3.4. Différentiation génétique entre les groupes génétiques basée sur l'analyse cpSSR

Le Tableau 5 renseigne sur la différentiation génétique (F_{ST}) entre les 9 groupes génétiques étudiés. On observe que les fréquences des haplotypes sont proches entre les groupes CAR, AND, MEX et « Pyrenees-Galicia2 », les F_{ST} par paire entre ces 4 groupes étant proche de 0 (Tableau 5). Par contre, la différenciation génétique est forte entre chacun de ces 4 groupes et le groupe NF, les valeurs de F_{ST} étant de l'ordre de 0.3 (Tableau 5). Le groupe CBD est le seul qui n'apparaît pas significativement différencié du groupe NF. Enfin, le groupe ITA apparaît fortement différencié de l'ensemble des groupes génétiques de référence.

La différentiation génétique entre les groupes E-SWF et W-SWF est très faible, égale à 0.06 (significativement différente de 0, Tableau 5). Ces deux groupes apparaissent non différenciés du groupe Pyrénées-Galice. Nous observons que les groupes E-SWF et W-SWF sont plus proches des groupes CAR, MEX et AND que des groupes NF, CBD et ITA (Tableau 5). En particulier, nous avons noté que les populations de l'Est du Sud-Ouest de la France ne sont pas significativement différenciées des groupes AND, CAR et MEX (F_{ST} par paire entre ces groupes et E-SWF est de l'ordre de 0, Tableau 5), tandis que, le groupe de l'Ouest est, pour sa part, très fortement différencié du groupe NF et est plus apparenté, pour son génome chloroplastique, aux populations AND et MEX (Tableau 5).

Tableau 5 : Matrice des F_{ST} estimés à l'aide de 7 marqueurs cpSSR pour les 7 groupes génétiques représentant l'histoire du maïs cultivé (Camus-Kulandaivelu et al., 2006) et les deux groupes génétiques identifiés dans le sud-ouest de la France (chapitre 1).

	AND	CAR	CBD	Pyrenees_galicia2	ITA	MEX	NF	E-SWF	W-SWF
AND		[-0.08 ; -0.03]	[0.10;0.22]	[-0.10 ; 0]	[0.51 ; 0.66]	[-0.07 ; -0.05]	[0.22 ; 0.32]	[-0.05 ; 0]	[0.03 ; 0.14]
CAR	-0.05		[0.12 ; 0.22]	[-0.06 ; 0.05]	[0.49 ; 0.62]	[-0.04 ; -0.01]	[0.22 ; 0.32]	[-0.03 ; 0.06]	[0.07 ; 0.17]
CBD	**0.16**	**0.18**		[0.21 ; 0.32]	[0.15 ; 0.51]	[0.07 ; 0.14]	[-0.04 ; 0.08]	[0.15 ; 0.22]	[0.32 ; 0.54]
Pyrenees_galicia2	-0.04	0	**0.27**		[0.60 ; 0.75]	[-0.03 ; 0.02]	[0.26 ; 0.38]	[-0.05 ; -0.03]	[-0.04 ; 0]
ITA	**0.59**	**0.56**	**0.31**	**0.68**		[0.43 ; 0.53]	[0.09 ; 0.42]	[0.45 ; 0.56]	[0.59 ; 0.75]
MEX	-0.06	-0.03	**0.1**	-0.010	**0.48**		[0.16 ; 0.22]	[-0.03 ; 0]	[0.08 ; 0.17]
NF	**0.27**	**0.27**	0.02	**0.33**	**0.29**	**0.2**		[0.21 ; 0.29]	[0.36 ; 0.56]
E-SWF	-0.03	0.01	**0.19**	-0.04	**0.51**	-0.01	**0.25**		[0.03 ; 0.09]
W-SWF	**0.08**	**0.13**	**0.44**	-0.03	**0.68**	**0.12**	**0.48**	**0.06**	

La diagonale inférieure contient les valeurs de F_{ST}; les valeurs en gras correspondent à des valeurs significativement différentes de 0 selon l'intervalle de confiance à 95%. La diagonale supérieure contient les intervalles de confiances à 95% estimées par bootstrap sur les populations et associés à chaque F_{ST}. AND= « Andean », CAR= « Caribbean », CBD= « Corn Belt Dent », ITA= « Italian », MEX= « Mexican », NF= « Northern flint », E-SWF= « Eastern South-West France », W-SWF= « Western South-West France ».

3.4. Discussion

Dans cette étude, nous avons analysé la variabilité du génome chloroplastique du maïs, pour les grands groupes génétiques de référence de la forme cultivée, un ensemble de populations de pays collectées dans le sud-ouest de la France et différentes accessions représentants les formes sauvages apparentées au maïs. En analysant la variabilité pour 6 marqueurs microsatellites chloroplastiques, nous avons identifié 21 haplotypes différents. Ces haplotypes ont pu être regroupés en 3 grands groupes sur la base de leur proximité génétique. Ces haplogroupes mettent en évidence une nette distinction entre les 2 téosintes Zea *parviglumis* et Zea *mexicana* (haplogroupe 3) et les autres espèces du genre *Zea* (haplogroupe 1). Le deuxième haplogroupe est composé uniquement de populations de maïs cultivées et ne contient aucune sous-espèce de maïs sauvage. Les haplotypes les plus fréquents des haplogroupes 2 et 3 sont présents dans presque tous les groupes génétiques. Les E-SWF et W-SWF du Sud-Ouest de la France présentent toutes les formes de cpSSR observées chez chacun des 7 groupes génétiques identifiés par Camus-Kulandaivelu et al. (2006). Néanmoins, la diversité chloroplastique indique que les populations de ces deux groupes sont maternellement plus proches des populations issues des régions tropicales et sub-tropicales d'Amérique que des formes Northern Flint.

3.4.1. Diversité chloroplastique et histoire ancienne du maïs

Les proximités génétiques dans notre étude sont en accord avec les résultats obtenus par Provan et al. (1999b) sur les accessions de maïs des espèces et sous-espèces du genre *Zea*. Notre étude montre que le maïs cultivé partage un haplotype (haplotype X) avec la téosinte *Zea mays* subsp. *parviglumis*. Ce résultat est en accord avec les résultats de Matsuoka et al. (2002) qui ont confirmé dans leur étude basée sur des marqueurs SSR nucléaires (99 loci) que les populations de maïs ont été domestiquées à partir la téosinte *Zea mays subsp. parviglumis*, leur seul et unique parent sauvage (Matsuoka et al., 2002). Dès les années 1985, cette origine avait été mise en évidence sur la base de marqueurs enzymatiques (Doebley et al ,1987). Nous avons également retrouvé l'haplotype X au sein des accessions de *Zea mays* subsp. *mexicana*. Ce qui peut s'expliquer par le fait que ce cytoplasme est ancestral à ces deux espèces. Une autre

explication pourrait être que des populations de maïs cultivé seraient originaires d'une domestication à partir d'accessions de *Zea mays* subsp. *mexicana* ou que cet sous-espèce *mexicana* a été impliquée dans des flux de gènes (*via* les graines) depuis les formes sauvages vers les formes cultivées depuis la domestication. Nous avons observé qu'un autre cytoplasme (AD) est partagé entre le maïs cultivé et *Zea mays* subsp. *mexicana ;* mais cette fois-ci nous n'avons pas trouvé ce cytoplasme au sein de la forme *Zea mays* subsp. *parviglumis*. Bien que ce résultat puisse être dû à un simple effet d'échantillonnage, il renforce l'hypothèse selon laquelle *Zea mays* subsp. *mexicana* pourrait avoir joué un rôle dans la domestication ou l'histoire ultérieure du maïs cultivé (via des hybridations notamment). Baltazar et al. (2005) ont montré que le maïs Zea *mays* subsp. *mexicana* sont inter-fertiles, cette hypothèse est donc plausible. Des flux de gènes modérés de la téosinte Zea *mays* subsp. *mexicana* vers les populations de maïs cultivées ont en effet été mis en évidence dans les études de Ellstrand et al. (2007), Matsuoka et al. (2002) et van Heerwaarden et al. (2011). Ellstrand et al. (2007) ont montré que vu les différences sur leurs cycles de floraison, les Zea *mays* subsp. *mexicana* sont plus prompts à recevoir du pollen des populations de maïs que de féconder les épis de ces derniers. Si des flux de gènes ont eu lieu, notre étude confirme bien que c'est la forme *Zea mays* subsp. *mexicana* qui a transmis son cytoplasme à l'occasion de ces flux. Par ailleurs, Provan et al. (1999a) ont montré par analyse cpSSR que certaines populations de maïs cultivées se montraient aussi proches des *Zea mays* subsp. *parviglumis* que des *Zea mays* subsp. *huehuetenengensis*. Ceci suggère que des formes de maïs sauvages autre que les *Zea mays* subsp. *parviglumis* auraient contribué à l'évolution du maïs en Amérique.

L'analyse de la diversité chloroplastique sur les grands groupes génétiques représentants l'histoire évolutive du maïs cultivé a permis de révéler une importante diversité, tant à l'échelle globale qu'à l'échelle des groupes génétiques. Ce résultat, ainsi que les patrons de différenciation observés entre groupes sont globalement en accord avec les données issues d'analyses basées sur des marqueurs microsatellites nucléaires (Matsuoka et al., 2002; Vigouroux et al., 2008). Bien que la différenciation entre les groupes génétiques soient globalement moindre que celle détectée avec les marqueurs nucléaires en général, la subdivision des haplotypes détectés au sein du compartiment cultivé en deux grands haplogroupes est en accord avec l'histoire de l'expansion de la culture du maïs à travers le continent Américain (Rebourg et al., 2003; Tenaillon and Charcosset, 2011; Vigouroux et al., 2008). En effet, même si tous les haplotypes existent dans les différents groupes génétiques, l'haplogroupe 2 est essentiellement formé par le groupe NF dont la majorité des populations

possèdent le l'haplotype I ; tandis que l'haplogroupe 3 est majoritairement composé d'haplotypes appartenant aux groupes génétique représentatif des Caraïbes, de l'Amérique centrale et du sud. Comme attendu étant donné son statut d'espèce sauvage ancêtre du maïs, nous avons pu montrer que le *Zea mays* subsp. *parviglumis* présente les haplotypes appartenant à l'haplogroupe 3, caractéristique des groupes génétiques de la zone d'origine du maïs (MEX) et des populations des Caraïbes, d'Amérique centrale et du sud-ouest de l'Amérique (AND et CAR). Par contre nous n'avons pas observé chez la téosinte de forme haplotypique appartenant à l'haplogroupe 2, plutôt caractéristique des formes cultivées du nord de l'Amérique. L'absence d'haplotypes identifiés chez des formes sauvages de *parviglumis* au sein de l'haplogroupe 2 dans notre jeu de données peut s'expliquer de deux manières : soit il n'a pas été échantillonné parmi nos populations de maïs cultivé ; soit d'autres accessions de téosinte (ou d'autres sous-espèces du genre *Zea*) ont contribué au compartiment cultivé mais n'ont pas encore été explorées ou ont disparu. L'étude de l'origine génétique des populations de maïs portant sur l'analyse de l'ADN nucléaire a conduit à l'identification de groupes génétiques qui réunissent, en leur sein, des populations localisées dans diverses régions géographiques (Camus-Kulandaivelu et al., 2006; Vigouroux et al., 2008). Sur les données cytoplasmiques, nous avons observé un F_{ST} nul entre les groupes AND, MEX et CAR, ce qui montre que la divergence entre ces trois groupes est récente ou que des flux de gènes récents ou réguliers ont lieu entre eux. Ces trois populations apparaissent par contre très éloignées du groupe NF, ce qui suggère une origine un peu différente ou un important goulot d'étranglement lors de la diffusion du maïs cultivé vers le nord de l'Amérique (peu de plantes mères auraient participé à cette diffusion). Un résultat intéressant est que NF et CDB ne sont pas significativement différenciés sur la base des marqueurs cytoplasmiques.

3.4.2. Diversité chloroplastique et origine des populations du Sud-Ouest de la France

L'analyse d'une portion du génome chloroplastique montre une différenciation faible entre les groupes E-SWF et W-SWF. La comparaison des 2 groupes au SWF fait apparaitre un plus grand nombre d'haplotypes pour E-SWF que pour W-SWF. Ceci est cohérent avec les données obtenues sur le génome nucléaire, notamment la plus grande dispersion des populations E-SWF observé sur l'analyse ACP portant sur les données SNP (Diaw et al., 2021). Les F_{ST} entre ces deux groupes et les groupes de référence définis par Camus-Kulandaivelu et al. (2006)

indiquent une forte proximité entre ces populations et les groupes AND, MEX et CAR. Les données obtenues dans les études de Dubreuil et al. (2006) sur le génome nucléaire de populations de maïs montraient une position plus intermédiaire de ces populations entre les populations CAR et les populations NF, et celle de Mir et al. (2013) ont montré une position intermédiaire des populations SWF entre les populations du Nord de l'Amérique du Sud et les NF. Or, l'introduction du maïs dans les Caraïbes s'est effectué à partir des populations de maïs du Nord de l'Amérique du Sud (Vigouroux et al., 2008). Ces deux résultats indiquent que les populations du Sud-Ouest de la France pourraient être issues d'hybridation entre des populations des deux groupes NF et CAR. L'étude chloroplastique suggère donc que cette hybridation aurait eu lieu entre des plantes mères ancêtres des populations SWF issues de semences de type CAR recevant du pollen de populations de type NF. Ces formes CAR peuvent provenir soit d'une introduction de maïs venant d'Amérique directement dans le SWF (Tenaillon and Charcosset, 2011), soit de l'introduction des populations de maïs provenant de l'Espagne dans le SWF (Renoux, 2014). La présence de l'haplotype D de l'haplogroupe2 chez les populations « Pyrenees-Galicia2 » et E-SWF montre bien que des populations de l'Espagne ont été introduites dans le Sud-Ouest de la France.

Le fait que les populations du Sud-Ouest de la France soient également très différenciées du groupe CBD peut s'expliquer de la même manière que celle proposée pour la différenciation avec le groupe NF. La forte proximité observée avec le groupe « Pyrenees-Galicia » montre que les populations du Sud-Ouest de la France et des Pyrénées espagnoles ont une même origine maternelle (cytoplasmique). Par contre la différenciation avec les populations italiennes est très forte, les deux groupes ont donc des origines très différentes.

3.4.3. Critique sur la méthode d'analyse

Contrairement à l'analyse du chapitre 1, l'analyse des marqueurs microsatellites chloroplastiques (cpSSR) porte sur un génotypage d'une plante individuelle par population. Nous avons négligé la variabilité intra-population. La prise en compte de la variabilité chloroplastique intra et inter population aurait nécessité l'analyse de plusieurs individus par population, ce qui est à la fois lourd et très couteux ou une analyse en bulk. Cette dernière a été réalisée au laboratoire avant mes travaux et 77% des populations présentaient un cytotype

unique (Gouesnard, comm. perso.) Nous discuterons de cette méthode dans la partie discussion générale de cette thèse.

3.5. Conclusion

L'analyse de la diversité génétique de populations de maïs avec des microsatellites chloroplastiques est une étude complémentaire à l'étude de l'origine des populations de maïs effectuée sur les données de génotypage avec la puce 50K SNP. Les données obtenues sont particulièrement riches et apportent des informations très intéressantes sur l'origine maternelle des populations du Sud-Ouest de la France, mais aussi sur l'histoire du maïs cultivé. En effet, nous avons pu montrer que les populations de maïs tropicales (AND, CAR et MEX) et notamment, d'après les données nucléaires, les formes Caraïbes sont les formes qui ont le plus contribué aux populations de maïs du sud-ouest de la France par la voie maternelle (via les semences). Nous avons aussi montré la forte proximité entre les groupes CAR, AND et MEX et une différenciation cytoplasmique beaucoup plus nette associée à des haplogroupes différents entre ces populations et le groupe des Northern Flint. Enfin les données cytoplasmiques ouvrent de nouvelles hypothèses sur l'origine maternelle du maïs cultivé, en mettant en évidence l'implication possible d'autres espèces du genre Zea. Notre étude est complémentaire de celle de Provan et al. (1999a), qui est la seule que nous connaissons dans la bibliographie comme analysant la variabilité chloroplastique sur un grand nombre de populations de maïs.

CHAPITRE III

Enquête auprès des anciens agriculteurs et pratiques de la gestion des semences des populations de maïs du Sud-Ouest de la France

4.1. Introduction

L'ethnobotanique est une discipline scientifique qui a été bien décrite par Brousse (2011) dans son journal intitulé « Une analyse historique et ethnobotanique des relations entre les activités humaines et la végétation prairial ». Cet auteur déclare que le terme « *ethno-botany* » est né sous la plume du professeur américain J.W. HARSHBERGER, botaniste et agro-botaniste, dans un article publié par le « Philadelphia Evening Telegram » le 5 décembre 1895. Ce professeur définit l'ethnobotanique comme étant « la science de l'usage des plantes par les peuples aborigènes ». Plus tard en 1978, Richard Ford considère que « l'ethnobotanique est l'étude des interrelations directes entre les humains et les plantes », étendant l'ethnobotanique à l'ensemble des sociétés humaines. Ainsi, Richard Ford initia l'intégration des différentes méthodes et théories s'appuyant souvent sur plusieurs disciplines (Histoire, Géographie botanique, Géographie humaine, Botanique, Génétique, Sociologie, etc), selon les questions de recherches posées par l'étude à mener. L'Ethnobotanique utilise de nombreux sources et moyens d'étude parmi lesquels nous pouvons citer les enquêtes ethnobotanique au sein des ethnies (ou peuple) en place dans les terroirs étudiés (Portères, 1961).

Ce type d'étude est devenu un sujet de recherche important pour les conservateurs/gestionnaires de la diversité des cultures (Badstue et al., 2007; Pautasso et al., 2012). Certaines études ethnobotaniques sont utilisées pour évaluer le maintien des populations locales cultivées dans les systèmes de culture intensive (Ardenghi et al., 2018; Bellon et al., 2011a). Ainsi, par exemple, dans son étude sur les populations de maïs du Nord de l'Italie, Ardenghi et al. (2018) ont effectué l'inventaire des variétés locales de maïs actuellement cultivées dans la zone, évalué leur état de conservation *ex situ* ainsi que les mesures mises en place par les agriculteurs pour empêcher ou limiter les événements d'introgression avec les cultivars élites. Bellon et al. (2011a) déclarent qu'au Mexique, ce sont les maïs des agricultures des hautes terres qui sont les plus vulnérables et que les agriculteurs devront probablement acquérir des semences en dehors de leurs aires géographiques traditionnelles dans le futur.

4.2. Etude ethnobotanique et application au maïs

Le rôle des agriculteurs dans le maintien de la diversité des variétés locales du maïs a été bien documenté (Bellon and Brush, 1994; Orozco-Ramirez et al., 2016). Les connaissances traditionnelles sur la performance des variétés locales ainsi que les pratiques de conservation et de partage des semences sont transmises au sein des agriculteurs d'un même village. Ainsi, l'homme devient un acteur important contribuant à l'évolution de la diversité des variétés locales (Bellon et al., 2017; Doebley, 2004; Perales et al., 2003b). Les sélections effectuées par l'homme sur la plante diffèrent d'une région à une autre mais aussi diffèrent selon les objectifs de production attendus sur la culture. Après chaque récolte, la plupart des producteurs gardent une partie de la récolte qui sera utilisée comme semences pour les semis de l'année suivante. Ils se basent pour cela sur un certain nombre de critères agro-morphologiques de l'épi de maïs. Cette sélection, répétée d'années en années, ainsi que les pratiques de gestion façonnent la diversité génétique et peut entraîner des divergences entre variétés cultivés par différents agriculteurs appartenant au même village ou groupes d'agriculteurs de différents villages voisins. Ces forces évolutives s'ajoutent aux contraintes environnementales des zones de production. En s'appuyant sur ces attendus et afin de comprendre l'origine de la diversité agro-morphologique des populations cultivées jusque dans les années 1960-1980 dans le Sud-Ouest de la France (SWF), nous avons effectué une étude qui associe :

- des résultats d'enquêtes ethnobotaniques pour documenter l'utilisation des populations de maïs par les anciens agriculteurs du SWF et leurs critères de sélection.
- une analyse des différentiels de sélection obtenus sur un test du choix d'épis proposé aux personnes enquêtées avec une estimation de l'héritabilité de différents caractères réalisée en station. L'objectif était de comprendre l'impact des pratiques de sélection des anciens agriculteurs des Pyrénées sur l'évolution des populations de maïs dans le SWF

En 2012-2013, une première enquête dans les Pyrénées françaises a été entreprise pour recueillir des données sur les connaissances et le savoir-faire des anciens agriculteurs (Beigbeder and Carraretto, 2018) et http://pro-mais.org/les-populations-de-mais-pyreneennes). Cependant, les entretiens n'ont pas permis de quantifier les critères de sélection, ni de prédire le changement génétique attendu de cette sélection massale sur les épis. Par conséquent, notre enquête effectué en 2018 dans les Pyrénées françaises, nous a permis à la fois (1) de compléter l'enquête

précédente sur le savoir-faire des anciens agriculteurs de maïs des Pyrénées, (2) de quantifier les différents critères relatifs à la morphologie de l'épi de maïs grâce au test du choix des épis et d'estimer la réponse attendue de cette sélection massale en estimant l'héritabilité des caractères, et (3) de tester si la répartition géographique (Est vs Ouest) des populations de maïs serait issue d'une différence dans la sélection effectuée par les anciens agriculteurs.
Dans ce chapitre, je vous présente de manière détaillée (1) la préparation de l'Enquête effectuée au niveau du SWF et les tests de sélection effectués auprès des personnes enquêtées, (2) la mise en place des différentes expérimentations dans la station de l'INRA à Mauguio et les traits morphologiques mesurés sur les épis de maïs, et (3) un résumé rapide des informations obtenues au cours de l'enquête et des résultats de l'analyse génétique quantitative. Cette dernière partie de ce chapitre sera valorisée sous la forme d'un article scientifique en cours de rédaction intitulé « Ethnobotanical and genetic investigations of old maize landraces from the Pyrenees : seed management practices and expected response to selection » (voir ci-dessous).

4.3. Les enquêtes dans les Pyrénées

Nous avons réalisé une enquête auprès d'anciens agriculteurs ayant contribué à la culture traditionnelle du maïs dans le Sud-Ouest de la France afin de mieux comprendre les critères de sélection (et méthodes de sélection), les pratiques d'échanges et de conservation de semences. Cette étude faisait suite aux enquêtes ethnobotaniques conduites par Maryse Carraretto en 2012-2013, auprès des anciens producteurs de maïs dans la zone des Pyrénées et ayant connu la culture du maïs avant l'arrivée des hybrides (http://pro-mais.org/les-populations-de-mais-pyreneennes). En 2018, elle a rédigé en collaboration avec Jean Beigbeder, vice-président de l'association Promaïs (http://pro-mais.org), un livre intitulé « Les maïs anciens des Pyrénées: savoir et savoir-faire traditionnels » (Beigbeder and Carraretto, 2018).

Pour construire l'enquête et identifier les agriculteurs qui pourront participer à notre enquête en 2018, nous avons travaillé en collaboration avec Jean Beigbeder et Maryse Carraretto qui nous ont fourni les fichiers des anciens agriculteurs de maïs qui avaient participé aux enquêtes effectuées dans les Pyrénées françaises en 2012-2013. La majorité de ces personnes étant inaptes à participer à notre enquête de 2018, nous avons pris contact avec les mairies et les chambres d'agriculture des départements des Pyrénées françaises (Ariège, Hautes Pyrénées, Pyrénées Atlantiques). Ceux-ci nous ont fourni les contacts d'anciens agriculteurs qui auraient connu la culture du maïs jusque dans les années 60. Ainsi, nous avons obtenu une liste de 18

personnes, incluant d'anciens agriculteurs mais aussi des fils d'anciens agriculteurs, qui ont participé à notre enquête en 2018 : 8 en Ariège, 9 en Hautes Pyrénées, 1 en Pyrénées Atlantique (Figure 1).

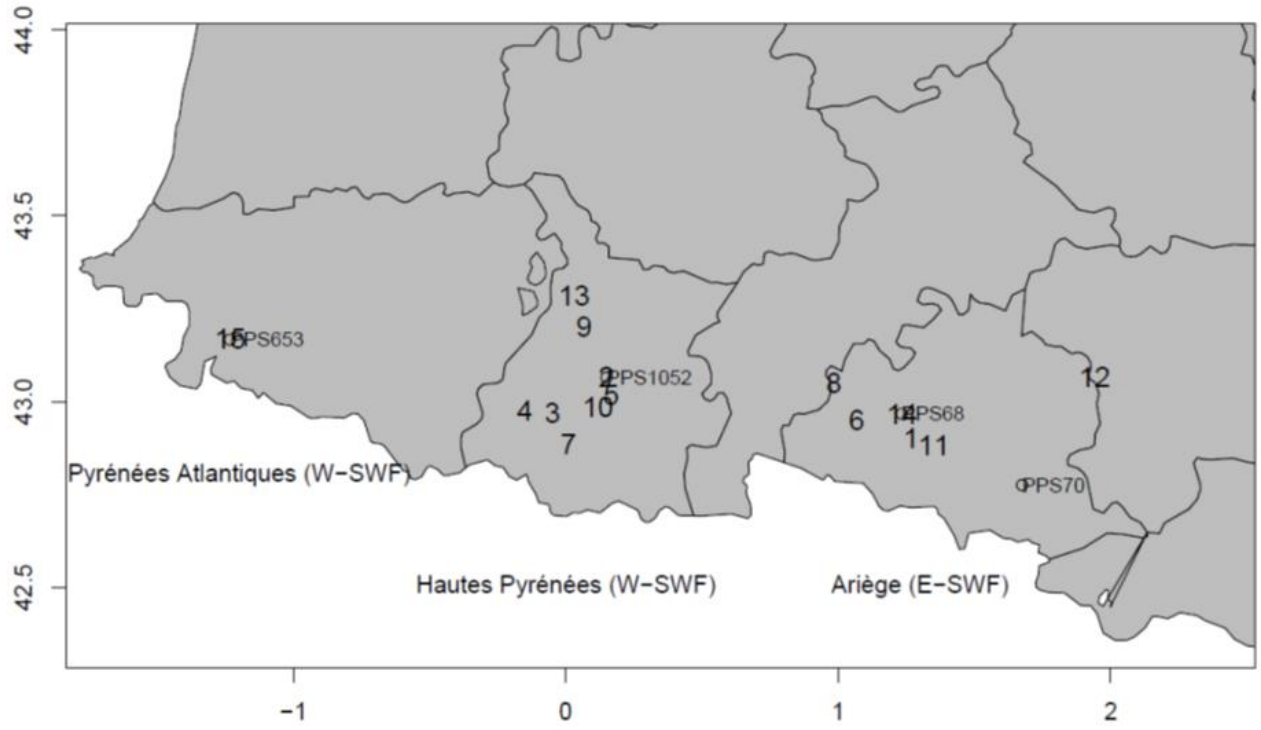


Figure 1 : Répartition géographique des personnes ayant participé à l'enquête de 2018 dans les Pyrénées françaises et des zones de régions de collectes des 4 populations de maïs utilisées dans cette étude. *PPS653 est une population à grains blancs collectée en Pyrénées Atlantiques et PPS1052 est une population à grains jaunes collectée dans les Hautes Pyrénées ; ces deux populations appartiennent au groupe génétique W-SWF. Tandis que PPS68 et PPS70 sont deux populations à grains jaunes et blancs respectivement, appartenant au groupe E-SWF et collectées dans le département d'Ariège. Les anciens agriculteurs et fils d'anciens agriculteurs sont représenté par les numéros 1,2, …,18 sur la carte.*

4.3.1. Description de l'enquête

La forme d'enquête effectuée en 2018 était un entretien ouvert qui durait environ une heure de temps en présence des personnes interviewées seules ou en compagnie des membres de leurs familles (surtout pour les personnes les plus âgées), de moi et de l'encadrante de ma thèse. Nous avons préparé un questionnaire (Figure 2) qui n'a pas été utilisé de façon rigoureuse lors des enquêtes mais elle nous avait servi de repère pour la bonne conduite de la discussion avec les anciens agriculteurs. Sur les recommandations de nos collaborateurs, Jean Beigbeder et Maryse

Carraretto, nous leur avons présenté 4 lots d'épis de maïs pour raviver leurs souvenirs sur la culture du maïs. Ces lots d'épis sont issus des 4 populations de maïs qui ont été choisies pour effectuer l'analyse génétique quantitative associée à cette étude. La présentation des épis de ces populations en début d'enquête a été une astuce clé qui a facilité la discussion et surtout incitait les personnes les plus âgées à s'ouvrir à nous. Elles réagissaient positivement en parlant de la beauté de nos épis de maïs. Ainsi, nous entamons la discussion en leur demandant laquelle de ces populations ressemblait le plus à celle qu'elles cultivaient jusque dans les année 60. Puis nous les interrogions sur le nombre de variétés cultivées dans leurs parcelles, la superficie des parcelles, leurs objectifs de culture, la sélection des semences, les échanges, etc.… A la fin de l'entretien, nous leur proposons de participer à un test de choix des épis afin de quantifier les critères de sélection de l'épi (voir la partie ci-dessous).

Questionnaire d'enquêtes pour les anciens agriculteurs des Pyrénées (juin 2018)

Questions d'ordre générale

- Depuis quand cultiviez-vous du maïs ?
- Comment avez-vous connu la culture du maïs ?
- Pourquoi cultiviez-vous du maïs ?
- Pouvez-vous nous décrire la plante au champ ?
- Comment été organisé la récolte ?
- Vous n'avez cultivé que de(s) variété(s) locale ?
- Avez-vous entendu parler des variétés hybrides, entre 1950-1960 ?
- Parlez-nous du passage des maïs de PAYS au maïs hybride ?

Quelle variété parmi les 4 proposées ressemble le plus à celle que vous cultiviez ?

Questions importantes de l'enquête

- Combien cultiviez de variétés en même temps ?
- Quelles variétés utilisez-vous ? Comment les appeliez-vous ?
- Couleur de la variété ?
- D'où provenaient vos semences ?
- Les choisissez-vous au champ, directement sur les plantes ?
- Comment reconnaissiez-vous les épis à conserver pour les semences de l'année prochaine ?

  le choix des épis était-il fait en groupe ou tout seul ?
- Comment conserviez les épis (avec spathe ou sans spathe) ?
- Utilisiez-vous tous les grains de l'épi ? quelle partie privilégiez-vous ? pourquoi ?
- Comment renouvelez-vous la ou les variétés cultivées ?
- Vous arrive-t-il de partager des semences avec le voisin ? les agriculteurs des autres vallées ?

  Pour quelle(s) raison ?
- Aviez-vous acheter des semences sur le marché ? pour quelle(s) raison(s)

Figure 2 : questionnaire utilisé lors des enquêtes de 2018, auprès des anciens agriculteurs et fils d'anciens agriculteurs de maïs des Pyrénées.

4.3.2. Le test de sélection des épis effectué lors de l'enquête

Pour identifier les critères de sélection les plus importants pour les agriculteurs lors de notre enquête, nous avons amené 4 lots de 150 épis de maïs issus des 4 populations présentées aux anciens agriculteurs et fils d'anciens agriculteurs de maïs des Pyrénées françaises. Nous avons demandé, à chacune des personnes qui ont participé à l'enquête, de choisir 15 épis parmi 150 épis présentés issues d'une même population, le test étant réalisé sur de 1 à 3 populations. Pour choisir ces 4 populations, nous avons défini deux zones de cultures en prenant en compte les groupes génétiques obtenus avec l'analyse d'admixture à k=2 portant sur les données SNP des 194 populations du SWF ; c'est-à-dire les groupes E-SWF et W-SWF. Dans chaque zone de culture, nous avons choisi 2 populations locales de maïs parmi les populations les plus représentatives du groupe pour les variables morphologiques (d'après les données de Gouesnard et al. (1997)) et ayant un taux d'assignation supérieure ou égale à 80% selon le logiciel ADMIXTURE (Alexander et al., 2009). Ces 4 variétés ont été mises en culture en 2017 à la station expérimentale de Mauguio (UE Diascope), pour produire les épis dans des conditions homogènes et construire le dispositif d'analyse de l'héritabilité de différents caractères agro-morphologiques de l'épi (voir paragraphe suivant).

4.4. Dispositif expérimental de production d'épi et de mesure des caractères agro-morphologique des variétés

Les parcelles expérimentales mises en place en 2017 sur le site expérimental de l'UE Diascope à Mauguio ont servi de parcelles pour effectuer des croisements demi-frères et plein-frères afin d'estimer les variances génétiques pour différents caractères agro-morphologiques. Pour cela, dans chacune des 4 populations devant être utilisées pour les enquêtes, nous avons choisi 20 plantes de maïs comme plante mâle et nous avons croisé chacune d'entre elles avec 2 plantes choisies comme plante femelle.

Nous avons récolté 150 épis dans chaque parcelle, et mesuré pour chaque épi la longueur de l'épi complet (c'est-à-dire de la partie basale à la pointe, « E_length1 »), la longueur de l'épi fécondé (c'est-à-dire la partie où les grains de maïs sont présents, « E_length2 »), les diamètres à 1/3 (DIAM2) et à 2/3 (DIAM1) de l'épi, le nombre de rang par épi (« N_row ») et le nombre moyen de grains sur le rang (« NKR »). A partir de ces caractères morphologiques mesurés,

nous avons calculé d'autres caractères agro-morphologiques : le volume de l'épi (« volume », qui donne un aperçu sur la grosseur de l'épi), le diamètre moyen de l'épi (« E_diam »), la conicité, hauteur de l'épi non fécondé (« EAR_TOP »), le nombre moyen de grains par épi (« K_number ») et le poids des grains d'un épi (« K_weight »). Les formules utilisées pour l'estimation de ces caractères sont fournies dans l'encadré 1. En se basant sur les épis choisis par les anciens agriculteurs, nous avons estimé les gradients de sélection (S) pour chacun des traits agro-morphologiques de l'épi avec la formule ci-dessous :

$$S = \mu_{15\ épis} - \mu_{150\ épis}$$

Où $\mu_{15\ épis}$ est la moyenne du trait morphologique étudié sur les 15 épis sélectionnés par l'ancien agriculteur, et $\mu_{150\ épis}$ est la moyenne du trait morphologique étudié sur les 150 épis.

Encadré 1:

Les formules utilisées dans l'article 2 pour l'estimation de :

- « Volume » :

$$Volume = 0.001 * 3.1416 * E_length1 * \frac{E_diam}{4}$$

- « E_diam »

$$E_diam = \frac{DIAM1 + DIAM2}{2}$$

- « Conicity » de Ordas and Deron (1988)

$$Conicity = \left(\frac{(DIAM1 - DIAM2) * 3}{2 * E_lenght1}\right) * 100$$

- « Ear_top »

$$Ear_{top} = E_length1 - E_length2$$

- « K_number »

$$K_number = \mu + a_1 * E_lenght + b_1 * E_diam + c_1 * N_row + d_1 * NKR$$

Avec a_1, b_1, c_1 et d_1 déterminés à partir d'une droite de régression sur un échantillon de 80 épis.

- « K-weight »

$$K_weight = \mu + a_2 * E_lenght + b_2 * E_diam + c_2 * N_row + d_2 * NKR$$

Avec a_2, b_2, c_2 et d_2 déterminés à partir d'une droite de régression sur un échantillon de 80 épis.

Au cours de l'été 2018 à Mauguio, nous avons semé les semences issues des 4 populations en utilisant un dispositif en blocs randomisés avec 2 répétitions par population. Pour chacun des 8 parcelles de population, nous avons semé en ligne les grains des épis issus des 40 croisements plein-frères réalisés durant l'été 2017. Les mesures réalisées sur ce dispositif ont permis d'estimer les variances génétiques des caractères morphologiques avec le modèle animal de Wilson et al. (2010) ci-dessous :

$$y = X\beta + Z1 * a + Z2 * d + e$$

Où *y* est le vecteur du phénotype de l'individu ; *β* est le vecteur des effets fixes ; *a* est le vecteur de la variance génétique additive de l'individu et *d* celui de la variance de dominance ; *X, ZI* et *Z2* sont les matrices de dispositif associées aux effet fixes et aux deux effets aléatoires ; *e* est le vecteur de la variance résiduelle.

L'héritabilité au sens strict ainsi que la réponse attendue par cette sélection massale sur un sexe sont les suivants (Gallais, 1990):

$$R = \frac{1}{2} * h^2 * S$$

Avec

$$h^2 = \frac{{\sigma^2}_A}{\sigma_A^2 + \sigma_D^2 + \sigma_E^2}$$

Où *S* est le gradient de sélection , σ_A^2 est la variance génétique additive, σ_D^2 est la variance génétique de dominance et σ_E^2 est la variance génétique résiduelle.

B136	B50	B39	B84	B154	B118	B52	B34
B127	B41	B1	B108	B135	B98	B53	B30
B134	B55	B3	B81	B140	B115	B59	B22
B122	B58	B31	B96	B129	B88	B70	B9
B159	B70	B29	B118	B151	B104	B79	B23
B151	B73	B26	B88	B148	B106	B51	B19
B138	B49	B7	B86	B133	B119	B46	B2
B145	B59	B8	B119	B142	B97	B57	B1
B143	B56	B5	B115	B145	B89	B77	B20
B147	B54	B18	B91	B155	B91	B78	B37
B140	B46	B30	B92	B143	B117	B47	B29
B125	B72	B9	B116	B152	B94	B43	B25
B150	B43	B37	B112	B160	B111	B50	B5
B124	B68	B36	B85	B132	B81	B44	B15
B144	B69	B24	B87	B147	B114	B48	B16
B130	B60	B38	B103	B130	B82	B80	B39
B160	B47	B10	B120	B158	B96	B54	B13
B152	B78	B22	B83	B123	B112	B74	B17
B133	B71	B20	B82	B128	B110	B60	B28
B141	B64	B23	B110	B121	B83	B68	B33
B129	B74	B16	B102	B156	B90	B45	B35
B153	B57	B15	B101	B139	B87	B41	B24
B158	B52	B40	B109	B159	B105	B75	B40
B132	B65	B14	B104	B124	B107	B62	B4
B128	B75	B2	B89	B137	B95	B69	B21
B146	B53	B35	B105	B134	B92	B64	B14
B135	B66	B11	B97	B157	B84	B55	B32
B149	B80	B32	B99	B144	B100	B66	B31
B148	B79	B4	B106	B127	B101	B42	B18
B137	B42	B19	B111	B126	B102	B63	B10
B155	B76	B21	B98	B141	B86	B73	B7
B139	B63	B28	B95	B122	B120	B72	B26
B123	B62	B13	B113	B150	B85	B61	B27
B157	B48	B33	B90	B153	B113	B56	B38
B142	B45	B25	B114	B125	B93	B49	B3
B156	B51	B12	B107	B149	B108	B71	B36
B154	B44	B17	B100	B131	B116	B58	B12
B121	B77	B6	B94	B146	B109	B67	B6
B131	B67	B34	B117	B136	B99	B76	B11
B126	B61	B27	B93	B138	B103	B65	B8

Figure 3 *: **Plan expérimental de 2018 à Mauguio pour l'estimation de la variance génétique sur les traits agro-morphologique de l'épi chez les 4 populations de maïs.** Dans cette parcelle de culture, les 4 populations de maïs ont été semées en utilisant un dispositif en bloc randomisé avec 2 répétitions par population. La couleur verte représente les bordures de la parcelle de culture. Les couleurs jaune, bleu, orange et violet représentent respectivement les planches de cultures des populations PPS70, PPS68, PPS653 et PPS1052. Chaque planche de cultures comprend les semis en ligne des 40 épis obtenus lors des croisements plein-frères et demi-frères réalisés en 2017 à Mauguio.*

4.5. Ce que l'enquête nous a appris sur les pratiques de culture du maïs dans le Sud-Ouest de la France jusque dans les années 1980 :

Les réponses à notre enquête réalisée auprès de 18 personnes ont permis de confirmer que les agriculteurs cultivaient le maïs pour la consommation humaine et l'alimentation des animaux. Les agriculteurs cultivaient en moyenne 1.4 variétés avec un même usage pour l'alimentation humaine et animale (à l'exception du gavage des oies). La taille des parcelles de culture du maïs était variable (de 0.25 ha à 1.5 ha, moyenne 0.6 ha), les techniques de culture étaient similaires à Est et à l'Ouest, à l'exception du semis pratiqué en ligne à l'Est ou en carré à l'Ouest. Comme déjà décrit dans Beigbeder and Carraretto (2018), la sélection des semences pour la génération suivante se faisait uniquement sur la morphologie des épis, et non sur les plantes au champ. Concernant les échanges de graines, les personnes interrogées nous ont écrit différentes situations. Trois personnes nous ont dit qu'elles n'échangeaient pas les graines ; quatre d'entre elles nous ont dit qu'elles échangeaient des graines à l'intérieur du village, la même variété étant utilisée au sein du village et deux d'entre elles échangeaient des semences avec les villages voisins ou avec un village étranger en cas de voyage (Tableau 1). Au cours de cette enquête, seuls 9 anciens agriculteurs ont accepté de faire le test de sélection pour un total de 17 tests d'épis ; 1 ancien agriculteur ayant accepté de faire le test sur 3 populations, deux sur seulement une population et le reste des agriculteurs sur 2 populations. Des 4 populations montrées aux anciens agriculteurs, 2 étaient de couleur jaunes et 2 de couleurs blanches. Ils ont désigné, en majorité, la population PPS1052 (une population de Bagnère de Bigorre des Hautes Pyrénées) comme étant proche morphologiquement de la variété qu'ils cultivaient (Tableau 1).

Tableau 1 : Récapitulatif sur l'identité, la localisation géographique des personnes ayant participé à l'enquête de 2018, les populations utilisées par les agriculteurs, les échange de semences.

Localisation		Personnes enquêtées		Populations cultivées			Elevage d'oie
Département	commune	Id. (voir Figure1)	âge	population proche de la variété	nombre et couleur des variétés cultivées	échange de semences	
09	Moulin-Neuf	12	ancien agriculteur	PPS1052	2 blanc	?	oui
09	Aleu	1	ancien agriculteur		2 jaune	voisins ou aux alentours	?
09	Massat	11	ancien agriculteur		1 jaune	?	?
09	Engomer	6	ancien agriculteur		2 blanc	?	oui
09	Riverenert 3	14	fils d'ancien agriculteur		?	?	oui
09	Riverenert 2	14	ancien agriculteur		?	voisins	?
09	Rivernevert 1	14	fils d'ancien agriculteur		?	?	?
09	Lacave	8	fils d'ancien agriculteur	PPS68	2 blanc et 1 jaune	?	oui
65	Oursbellile	13	fils d'ancien agriculteur	PPS70	1 blanc	?	non
65	Laloubère	9	fils d'ancien agriculteur		?	?	non
65	Bagnères de Bigorre	2	ancien agriculteur	PPS68	1 jaune et un peu de blanc	?	non
65	Campan 1	5	ancien agriculteur		1 jaune	voisins	non
65	Campan 2	5	ancien agriculteur		1 blonde	non	non
65	Lesponne	10	ancien agriculteur		1 jaune	voisins	non
65	Bun	4	ancien agriculteur	PPS1052	1 jaune	village aux alentours	non
65	Beaucens	3	ancien agriculteur	PPS1052	1 jaune	non	non
65	Esquieze Serre	7	ancien agriculteur		1 jaune	non	non
64	St Jean Pied de Port	15	fils d'ancien agriculteur		?	voisins	non

? = pas de données ; 09 =Ariège ; 65 =Hautes Pyrénées ; 64= Pyrénées Atlantiques. PPS653 est une population à grains blancs collectée au Pyrénées Atlantiques et PPS1052 est une population à grains jaunes collectées dans les Hautes Pyrénées ; ces deux populations appartiennent au groupe génétique W-SWF. Tandis que PPS68 et PPS70 sont deux populations à grains jaunes et blancs respectivement, appartenant au groupe E-SWF et collectées dans le département d'Ariège.

4.6. Mesure des pratiques des anciens agriculteurs des Pyrénées sur l'évolution des populations du SWF

4.6.1. Présentation générale

L'article 2 intitulé «Ethnobotanical and genetic investigations of old maize landraces from the Pyrenees: seed management practices and expected response to selection. » a pour objectifs d'évaluer l'impact des pratiques des anciens agriculteurs des Pyrénées françaises sur l'évolutions des populations de maïs. Le contenu de cet article en cours d'amélioration est résumé dans les paragraphes ci-dessous.

Pour mesurer l'impact des pratiques des anciens agriculteurs des Pyrénées françaises dans l'évolution des populations de maïs du Sud-ouest de la France, nous avons estimé la réponse attendue à la sélection pour des traits morphologiques de l'épi. Les traits morphologiques de l'épi pris en compte dans cette étude sont les traits morphologiques mesurés lors des expérimentations en 2017 et 2018 à Mauguio et définis en se basant sur les informations issues du livre de Beigbeder and Carraretto (2018). L'estimation de la réponse attendue à la sélection qui est un paramètre de génétique quantitative requiert des données sur les valeurs de la différentielle de sélection et de l'héritabilité au sens strict pour chacun des traits morphologiques pris en compte.

Dans notre étude, la différentielle de sélection que les agriculteurs exerçaient sur la variété est estimée en calculant la différence entre la moyenne de chacun des traits morphologiques mesurés sur les 15 épis choisis par les agriculteurs et les 150 épisprésentés. Les résultats de cette analyse ont montré que les agriculteurs sélectionnaient les épis longs et de gros volume. Ces caractères étant corrélés aux caractères de productivité, nous pouvons affirmer que les agriculteurs par leur choix visuel sélectionnaient pour une meilleure productivité. Les résultats montrent aussi que les agriculteurs sélectionnaient des épis bien remplis de la base jusqu'à la pointe, et sains.

Grâce aux expérimentations effectuées en 2017 et 2018 à Mauguio, nous avons estimé les variances génétiques additive, de dominance et environnementale. Nos analyses ont montré que les traits morphologiques de l'épi ayant une faible variance génétique additive sont le nombre de rang de grain par épi (NRG), le nombre de grain par épi (K_number) et le bout non rempli de l'épis (nommé « ear_top »). Nous avons aussi observé que la variance génétique la plus faible était celle de la population PPS68. Nous avons par la suite estimé les valeurs des

héritabilités au sens strict en divisant la variance génétique additive par la variance totale du phénotype (variances additive + dominance + environnementale). Moyenné sur les 4 populations, nous observons que les valeurs d'héritabilité sont de 0.47 pour la longueur de l'épi, 0.57 pour le diamètre moyen de l'épi, 0.33 pour le nombre de rangs de grains sur l'épi et 0.26 pour le nombre de grains par rang (ces traits sont nommés respectivement « E_length» « E_diam », « N_row » et NKR dans l'article2). L'héritabilité du volume de l'épi est égale à 0.53, celle du rendement en grains de 0.28 (« N_kernel ») et celle du poids de grain de 0.36 (« K_weight »).

En combinant ces valeurs et les différentielles de sélection, nous pouvons estimer la réponse à la sélection prévue (R) sur les différents traits morphologiques. L'analyse montre que la réponse attendue en pourcentage de la valeur initiale varie de -3.3% à 13.7%. Pour les principaux caractères (E_length, E_diam, NKR, NRG), les R étaient positifs, en moyenne 1.2 % pour le nombre de rang de grain par épi (NRG), 1.6 % pour le nombre moyen de grains pas rang (NKR), 1.6 % pour E_diam et 3.3 % pour E_length. Les agriculteurs ont tendance à sélectionner des épis productifs, longs, gros et à petits grains. Les valeurs de R sont aussi légèrement plus élevées pour la population PPS1052, qui présentait une part de variance génétique additive élevée par rapport aux autres populations.

Pour mesurer l'impact des forces évolutives, nous avons voulu savoir si les effectifs des épis utilisés pour le semis de la génération suivante n'induisaient une perte d'allèles par dérive génétique. Sur la base des informations recueillies au cours de l'enquête de 2018 sur la densité de semis et la superficie des parcelles, nous avons estimé le nombre des épis nécessaires pour semer la prochaine génération chez un cultivateur qui possède une parcelle de 0,25 ha et qui sème les grains de maïs à une densité de 50 000 plants/ha. Le résultat montre que 53 épis sont nécessaires. Cette taille est suffisante pour conserver les allèles rares à une fréquence de 1% avec une probabilité de 0,99 (Crossa, 1989). De plus, la réduction de l'hétérozygotie a été estimée faible avec une taille de population effective (Ne) autour de 211 avec Nfem=53 et Nmale=53*220. La force de la sélection semble plus importante que celle de la dérive génétique.

Un des objectifs de cette analyse était de mieux comprendre l'évolution des populations de maïs collectées dans le Sud-Ouest de la France. Nous avons donc comparé, via une analyse de variance (ANOVA), les différentielles de sélection standardisées entre les populations issues de la zone Est et celles de la zone Ouest de notre zone d'étude. Les différences inter zone observées ne sont pas significatives. Les variétés issues de la sélection n'étaient pas spécifiques d'une zone, une grande variabilité existait intra zone.

Nous avons vu avec les différentielles de sélection significatives que les agriculteurs savent sélectionner les épis productifs. Avec les héritabilités relativement élevées que nous avons trouvées, nous pourrions nous attendre à une réponse à la sélection aussi élevée. Notons que nos héritabilités ont été obtenues en un seul lieu en station expérimentale, elles sont vraisemblablement surestimées du fait des interactions génotype * environnement non pris en compte et des conditions favorables en station. Les réponses à la sélection réellement réalisées étaient certainement inférieures à ceux que nous avons estimées. La sélection massale a été reconnue comme efficace que pour les caractères héritables, et le rendement est peu héritable. La sélection massale est moins efficace que les sélections sur descendance pratiquées dans la sélection moderne. Avant l'arrivée des hybrides, dans les années 60, le progrès génétique observé sur les populations était faible. L'absence de sélection sur les caractères végétatifs, la faible réponse à la sélection massale sur un caractère peu héritable comme le rendement et le fardeau génétique des populations peuvent expliquer ce faible progrès génétique.

4.6.2. Article 2:

Ethnobotanical and genetic investigations of old maize landraces from the Pyrenees: seed management practices and expected response to selection.

Brigitte Gouesnard [1], Yacine Diaw [1,2], Laurène Gay[1], Joëlle Ronfort [1], Jacques David [1]

Author's affiliations

1 UMR Amélioration Génétique et Adaptation des Plantes méditerranéennes et Tropicales, Univ Montpellier, CIRAD, INRAE, Institut Agro, 10 rue Arthur Young, 34000 Montpellier, France

2 Institut Sénégalais de Recherches Agricoles, ISRA-CNRA de Bambey, Route des Hydrocarbures, Dakar, Sénégal

Corresponding author: Brigitte Gouesnard
Email: brigitte.gouesnard@inrae.fr
ORCID: https://orcid.org/0000-0002-9288-9745

Article2 en cours de révision.

Abstract

Landraces contribute to a large amount to the genetic diversity of maize (*Zea mays* L.). In France, maize landraces are no longer cultivated, with exceptions. Landraces analyzed in this study have been collected in the 1960s and are thus only conserved *ex situ*. The Pyrenees used to be a corn-growing region owing to subsistence farming systems for human consumption and animal feeding. In 2018, a survey was carried out in the Pyrenean region among former farmers or their children in order to document seed management practices and know-how 60 years ago. We asked farmers to select the seeds they would choose if they were to grow maize landraces the following year, and we realized 17 ear tests. Their selection was based on ears only. Old farmers selected very well healthy and productive ears; ears' length and volume were the two first criteria of selection. Both were highly correlated with the kernel weight of the ear. Heritabilities of ear traits were estimated in one trial for four landraces, and were found variable according the trait and the landrace (average 0.36 between 0 and 0.76). We calculated the expected genetic change after one generation of mass selection, following farmer selection criteria. For ear length, the genetic change was expected to be of about 3.4% with variable values (from 1% to 7.5% over the 17 ear tests). Farmers practices were similar in the East and West part of The Pyrenees, and looked like those of Native American maize farmers.

Key words: *Zea mays* L., survey, landraces, seed management, heritability, mass selection

Funding: This research was supported by the project Amaizing ANR-10-BTBR-01 and by West African Agricultural Productivity Program (WAAPP/ISRA) for the grant of Yacine Diaw.

Conflicts of interest/Competing interests: The authors declared no conflict of interest

Consent to participate and consent for publication: All authors consent to participate and agree to publication

Availability of data and material: Data can be asked to the corresponding author

Authors' contributions: BG prepared and investigated the survey, interpreted the data, wrote the manuscript; YD investigated the survey and corrected the manuscript, LG helped in the interpretation of data and corrected the manuscript, JR and JD corrected the manuscript and supervised the study.

Introduction

In 2005, Villa et al. (2005) defined a landrace as "a dynamic population of a cultivated plant that has historical origin, distinct identity and lacks formal crop improvement, as well as often being genetically diverse, locally adapted and associated with traditional farming systems'. Landraces from small-scale farmers contribute to a large amount to the genetic diversity of corn (*Zea mays* L.). In developing countries, landraces are still cultivated, often with some modern cultivars. It is true in Mexico, the center of domestication and diversity of maize. In this country, some studies were as objectives to understand why and how farmers keep cultivating local varieties while improved varieties have been introduced ((Bellon and Brush, 1994);(Bellon and Risopoulos, 2001); (Louette and Smale, 2000); (Rice et al., 1998)). They analyzed the components of farmers' management of genetic resources through the choice of variety, seed flows and seed selection management. Despite significant gene flow through cross-pollination, ear characteristics are maintained through farmers' selection. In developed countries, landraces are no longer cultivated with exception, and maize landraces are conserved ex situ in genebanks (Ortiz et al., 2010). Landraces diversity is searched for their high capacity to tolerate biotic and abiotic stresses as given in definition of a native landrace by Zeven (1998). However, maize landraces exhibit a low yield compared to elite hybrids (52% in Gouesnard et al. (2016)) due to their genetic load. In USA between 1900s and 1930s, their improvement by mass selection of open-pollinated populations did not lead to high genetic progress as found after with hybrids (Duvick, 2005).

Since 1960 landraces are no longer grown in France; they have been progressively replaced by modern hybrid varieties. Only farmers who reject industrialized agriculture and who often practice organic farming still cultivate landraces (Bocci and Chable, 2008). In the 1960s, researchers from INRA (National Institute for Agronomic Research) collected landraces in France in order to build germplasm collections and to preserve the genetic diversity available in these varieties. After 1983, the corresponding collections were managed through an association of public research laboratories and private maize breeders (Dallard et al., 2000). The genetic diversity of these landraces has been studied through many studies supported by INRA and the PROMAIS breeder association (Gallais et al., 2000). A large genetic variability among landraces has been shown. Genetic pools built according their combining abilities for grain and silage production were created in order to be used in plant breeding schemes. The diversity within landraces were evaluated with both enzymatic and RFLP markers: 75% of diversity is within population. This collection included 194 landraces from the South-West of

France, whose 119 were originated from the Pyrenees collected for their cold tolerance and flowering earliness. SNP analyses of the South-Western populations (SWF) allowed to detect two distinct genetic groups in these populations originated from Pyrenees-Galicia genetic group, one located in the eastern part (E-SWF) and the other one located in the western part of SWF (W-SWF) (Diaw et al., 2021). A hybrid zone between these two groups was located in districts of Haute-Garonne and Ariège, near Garonne valley (Diaw et al., 2021). This genetic structure could result from the demographic factors (random drift, restricted gene flow) or/and selection forces (climatic adaptation and selection due to farmer's practices). The effect of demographic factors was demonstrated by the highlighting of isolation by distance. Climatic adaptation was not proposed as main factor despite a slightly warmer and drier climate in the eastern part (Mediterranean climate) than in the western part (Atlantic climate). It remains to analyze the effect of drift and selection by farmers. Here, we aim to analyze these evolutionary forces to better understand the genetic structuration of Pyrenean maize landraces by aggregating an ethnobotanist approach to existing genetic data.

Seed management practices are an important component of the dynamics of genetic diversity within populations. It has become an important research topic for crop diversity curators/managers (Badstue et al., 2007; Pautasso et al., 2012). Thomas et al. (2011) advocated for a multidisciplinary approach for the study of seed exchanges in traditional farming systems, combining analyses of social, biological and environmental factors. Here, as maize landraces are no longer cultivated in the Pyrenees, we collected information about seed management practices through a survey of old farmers and their children, having memories about maize cultivation around 40 to 60 years ago. In 2012-2013, a first survey in the Pyrenees was undertaken to collect data on knowledge and know-how of small-scale farmers (Beigbeder and Carraretto, 2018) and http://pro-mais.org/les-populations-de-mais-pyreneennes) to document the technics of sowing and agricultural practices, the seed management, the man and animal uses of maize, the associated culture with beans, the transition with the arrival of hybrids. The interviews were carried out with around thirty people in 5 Pyrenees valleys. Concerning seed management, the book indicates that the selection of seeds from one generation to the next was only done on ears morphology. Farmers choose the most beautiful ears after harvest. The book reports the verbal description of the ears that were chosen by farmers. A chapter is written on seed renewal and seed exchange. Two situations existed: no exchange at all, and exchange with neighbors in the same valley or on the occasion of a trip such as a wedding. Farmers were said to cultivate one variety, the same for human and animal uses. The smallest maize cultivation areas were reported as to be 0.2 to 0.3 ha. This small surface can lead to drift and inbreeding

effects. However, the interviews could not quantify the selection criteria, nor predict the genetic change expected from this mass selection on ears.

In order to address the selection practices of the farmers on the Pyrenean landraces, we conducted a second survey in 2018 with a test of the choice of the ears in order to quantify the selection criteria. During the survey, information was collected to know the cultivated varieties, their use and their seed management. The approach was narrative and qualitative, and the survey restricted to a low number of persons chosen among informants still alive from the previous survey added with some new old farmers. The first objective of our study was to complement the previous survey on farmers' know-how, their practices were compared to those of the Mesoamerican natives cultivated maize landraces in traditional agrosystems to know their convergence or divergence, and information was used to analyze the effect of random drift as an evolutionary force. The second objective was to determine whether traditional farming practices and landraces were favorable for in farm genetic improvement. Expected response to selection requires high heritability for traits related to the yield and high selection differential exerted by the farmers on their landraces. Heritability estimates were evaluated in a field trial by analyzing full sib and half sib progenies of four Pyrenean maize landraces, and selection differential were estimated with the test of the choice of the ears proposed to old farmers. The third objective was to test whether the east-west genetic structure of the Pyrenean landraces could have been shaped by selection exerted by farmers, or whether genetic drift is expected to be sufficiently strong to create this differentiation.

Material & methods

The survey

The survey was carried out in July 2018 in the Pyrenees. People involved in the survey were issued from the list of informants given by Maryse Carreretto who conducted the survey in 2012-2013 (Beigbeder and Carraretto, 2018), and complemented using information given by town hall secretaries. In total, 18 people participated in the survey: 12 were old farmers (between 85 and 96 years old), and 6 were children of farmers (Table 1). Eight informants were located in the Eastern part of the Pyrenees, and 10 in the Western part (Figure 1 The survey was individual. We first asked the number of varieties, and their uses for human consummation and animal feeding. Second, we obtained information on the size of the plot cultivated with maize

in order to be able to predict the strength of genetic drift on these populations. Third, we asked their know-how on seed management methods such as the selection methods, seed exchanges, seed conservation practices. Farmers gave us also other information about agricultural technics or recipes.

The test of choice of ears and the estimation of selection differential

In order to gather knowledge on seed management and selection criteria, we multiplicated ears from four landraces collected in the 1960s. These landraces were chosen to be representative of the genetic and phenotypic diversity in Pyrenean landraces. We used data on phenotypic traits as described in Gouesnard et al. (1997) and the admixture coefficients for two genetic groups (Diaw et al., 2021). Suppl Figure 1 (see Annexe 3) gives the distribution of flowering date and ear row number for the Pyrenean landraces with positions of the 4 studying landraces. We chose two landraces per genetic zone. We kept one with yellow kernels and the other with white kernels as it was the two main colors found in Pyrenean landraces. Originated from Ariège (E-SWF), there was Riverenert 1 (PPS68 or FRA0410068), with yellow kernels and Aston A (PPS70 or FRA041070), with white kernels. Originated from Pyrénées Atlantique (W-SWF), Saint Jean Pied de Port 3 (PPS 653 or FRA0410653) has white kernels and Bagnères de Bigorre 2 (PPS 1052 or FRA041052), originated from Hautes Pyrénées (W-SWF), has yellow kernels. Seeds from these four landraces were obtained from the French Maize Biological Resources Center located in Montpellier (http://florilege.arcad-project.org/fr/collections/collection-mais) and grown in Mauguio, southern France (43°37'02''N, 3°51'18''E) in 2017 to produce 150 ears by random mating for each landrace, in the same field. During the survey in 2018, ears from one to three populations were presented to the informants with asking them to choose fifteen ears among the 150 ears as they did the selection in the time (Table 2). At the beginning of the survey, we presented uni-colored ears from full-sib crosses of these four landraces in order to ask them which landrace was more like the landrace they were cultivating and to remind them of memories.

We estimated the differential of selection S for each trait as the difference between the average phenotypic value of the 15 ears selected by a farmer and the average phenotypic value of the 150 ears presented ($S = \mu_{15ears} - \mu_{150ears}$). We analyzed the variability of standardized selection differentials by performing an ANOVA analysis (R Development Core Team, 2011) with following factors: zone (E-SWF vs W-SWF), informant inside the zone,

and landraces (PPS68, PPS70, PPS653, PPS1052). Thus, we tested the effect of the zone where the farmers lived on the E-W differentiation.

Ear morphological characterization of landraces

Before the survey, all ears (150 ears per landrace x 4 landraces) were measured for the following traits: total ear length (E_length in mm), ear length with kernels, ear diameter at the 1/3 down of the ear (DIAM1), ear diameter at the 1/3 up of the ear (DIAM2), number of rows (N_row), number of kernels per rows (NKR). From the measured traits, we calculated the diameter of the ear (E_diam in mm estimated by the mean of DIAM1 and DIAM2), the conicity of the ear (((DIAM1-DIAM2)*3)/(2*E_lentgh))*100 following (Ordas and Deron, 1988), the length of the top of ear without kernel (EAR_TOP in mm), the volume of the ear estimated as a cylinder (0.001*3.1416*E_length*E_diam/4). In order to get an estimation of the yield, we estimated the number of kernels per ear (K_number) and the weight of kernels of an ear (K_weight) using a linear regression based on E_length, E_diam, N_row, and NKR. The regression was based on 80 ears per landrace produced in 2018. To validate the estimation, we calculated the coefficient of correlation between estimated values obtained for 30 ears based on a regression using 50 ears, and the observed values on 30 ears. This procedure of cross validation was repeated ten times and showed that K_number and K_weight can be predicted with a good accuracy for each of the four landraces. The coefficients of correlation were upper than 0.7 over the 10 replicates for both K_number and K_weight on the 4 landraces, except for PPS653 for weight of kernels for which 8 replicates coefficients were between 0.4 and 0.7, all significant at 0.05. Following these results, K_number and K_weight were then predicted using the equation of the linear regressions based on E_length, E_diam, N_row and NKR on 80 ears per landrace. To compare the variability of the 4 landraces for ear morphology traits measured on 150 ears each, we performed boxplots on each trait and a principal component analysis (PCA) for a multivariable analysis (R Development Core Team, 2011). On PCA analysis, we added the mean of 150 ears (T0), and the mean of the 15 selected ears (T1 to T9, see Table 2 for the legend).

Estimation of narrow sense heritability for the 10 morphological traits considered in order to predict the expected genetic change after one generation of selection

For each of the four landraces, we built full-sib and half-sib families to estimate the additive and dominance components of genetic variance for a set of quantitative traits. In 2017, twenty random plants, used as male, were crossed to two female plants within each landrace. The plants were randomly sampled from the balanced bulked stock of the gene bank. A total of 40 full-sib families were obtained for each landrace.

In 2018 in Mauguio (near Montpellier, France), 40 full-sib families of each landrace (i.e. 160 FS families in total) were grown in an experimental trial with two replicates. The total trial was therefore organized in 320 rows, with 28 plants belonging to the same full-sib family in each row. The space between two rows was 0.80 m and the density was 70 000 plants / ha. The position of each family within a replicate was randomized. At flowering time, we measured the number of days to pollen shedding and silking. At harvest, a maximum of 15 ears were collected in each plot (12.3 ears per row on average). We measured the same set of quantitative traits as described I, section 3) for the 150 ears. We assumed no maternal effects.

The additive and dominance genetic variance components, as well as heritability values were estimated with a restricted maximum likelihood procedure, implemented in ASReml v1.1 (Gilmour et al., 2009). We fitted an ‘‘animal model’’ (Wilson et al., 2010) with a fixed effect for the replicate and two random effects for the additive genetic and dominance genetic effects:
$y = X\beta + Z_1 a + Z_2 d + e$,

where y is the vector of individual phenotypic values, β is the vector of fixed effects (intercept and replicate effect), a is the vector of individual additive genetic values and d is the vector of individual dominance genetic values. X, Z_1 and Z_2 are the design matrices associated with the fixed effect and the two random effects. e is the vector of residual errors. ***a, d*** and e are assumed to be normally distributed with means of zero, and variances of σ^2_A, σ^2_D and σ^2_E, respectively.

We tested the significance of the random additive effect using a likelihood ratio test to compare this full model with a model where the random additive effect was excluded.

For each trait, we then used the estimates of additive genetic variance (σ^2_A), dominance genetic variance (σ^2_D) and residual variance (σ^2_E) to quantify the narrow sense heritability ($h2$) as σ^2_A / ($\sigma^2_A + \sigma^2_D + \sigma^2_E$). The ration $a^2 = \sigma^2_A / (\sigma^2_A + \sigma^2_D)$ was calculated for each trait in each landrace to measure the weight of additive variance in the genetic variance of the trait.

We used the breeder's equation to predict the expected genetic response of a given quantitative trait to selection with selection only on females, R, using the estimate of the trait's narrow sense heritability, h2, and the selection differential measured for this trait in the experiment with the farmers, S.

$R = ½ * h^2 * S$ (Gallais, 1990)

Estimation of effective size to analyze the effect of random drift

We used the data about field description collected during our survey to estimate the number of ears needed to sow the next generation. As these values were variable among farmers, we simulated the worst scenario, by using the values of the smallest surface, lowest density of plants, and average number of kernels collected on a ear (on 330 kernels in total). Based on the number of female (Nf) and male (Nm) plants necessary for the next generation of maize, we calculated the effective size (Ne) as:

$1/Ne = 1/(4Nf) + 1/(4Nm)$ (Caballero, 1994)

Results

Survey

The small-scale farmers cultivated maize for human consumption and feeding of animals. The agro system was a subsistence agriculture. We found no difference between East and West part of Pyrenees for the use of maize. The name of the dish prepared for human consumption changed according to the region but the recipes were similar: the "milhas" (in East part) or the "pastet" (in West part) were made of maize flour cooked in water or milk, generally after the farmer killed a pig.

Maize surface was variable according to the informants, from 0.25 ha to 2 ha (mean 0.60 ha (sd=0.61)). The technics of cultivation were similar in East and West part except for the sowing:

in row in Ariège (East part), in square (West part). From 2 to 4 kernels were sowed together in the hole. The density was variable too, between 50 000 plants/ha and 100 000 plants per hectare before thinning. A few farmers said that they sowed beans (*Phaseolus vulgaris*) at the same time as maize (4 occurrences). After harvest, braids with husked leaves were made, and ears were dried in front of the house or inside the house.

Table 1 summarizes the information collected during our survey. Most farmers cultivated a single variety (9 occurrences), used for both human consumption and animal feeding. In the East part of the Pyrenees, about half of the farmers cultivated a second variety (4 occurrences with two varieties and one with three), generally a white variety with small kernels, used for force-feeding geese. Farmers cultivated on average 1.4 varieties.

As already described in Beigbeder and Carraretto (2018), the selection of seeds for the next generation was done only on ears' morphology, without any selection on plants in the field. The selection was performed at different times: at the stripping, at the shelling, at the end of the winter. The kernels for the seeds were shelled carefully, often by old people. One informant said us that old people selected on seed size. All informants declared that they eliminated kernels from the top and the bottom of the ear.

Concerning the seed exchange, the informants gave us different cases. Some informants said us to not exchange seeds. Others said us to exchange seeds inside the village, saying that the variety was the same inside the village. Others said us to exchange seeds with neighbors' villages or with a foreign village in case of travels (Table 1).

Morphological and agronomical variation in a panel of 4 landraces

We analyzed the data on the four landraces for ear traits (Figure 2). All the landraces were differentiated for the number of row per ear and the predicted kernel number. Differences of means were observed for at least one landrace for ear length, number of kernels per row, height of the kernel and kernel weight. No difference was observed for conicity and ear diameter (Figure 2). The landraces were differentiated also by kernel color and precocity. Average flowering dates for the 4 landraces were obtained from the trial at Mauguio in 2018 (on full sib progenies). PPS1052 was yellow with 10 kernel rows per ear, and its male flowering date was 726.1 °C (growing degree day base 6) on average. PPS653 was white with kernel rows per ear from 8 to 10, and its male flowering date was 698°C. PPS68 was yellow with a long ear, and a kernel rows per ear from 12 to 14, its male flowering date was 748.5 °C. PPS70 was white with 14 kernel rows per ear on average, its kernels were thin, its male flowering date was 727.5 °C

on average. The two landraces from the eastern part of the south west of France (PPS68 and PPS70) flowered later, had higher row number and smaller kernels than the two populations from the western part for the surveyed region (PPS653 and PPS1052). This differentiation between East and West part of Pyrenees is also found in the Suppl Figure1 (see Annexe 3) for flowering date and ear row number.

We performed a principal component analysis (PCA with R software (R Development Core Team, 2011)) on the 150 ears from the 4 landraces (Figure 3). On the graphic of individuals, the ellipses (70% of variation) around the ears of a landrace overlapped. However, PPS68 and PPS70 (E-SWF) were positioned on the right and upper part of PCA represented by variables correlated with NKR and K_number, whereas PPS653 and PPS1052 (W-SWF) were on the left and below part of PCA represented by variables correlated with K_height. The 2 populations of the East part of the Pyrenees had smaller grains and a better yield than the 2 western landraces. On Figure 3B, correlations between traits showed highly positive correlations between K_weight, K_number, volume, NKR, E_length and E_diam. Whereas, K_height and conicity were negatively correlated with these previous traits. Productive ears appeared with small kernels and rather cyclindrical ears. Based on these results, we expected that if farmers selected ears on their morphology (E_length, volume, N_row, NKR), they would have a positive response on yield.

The additive, dominance and residual components of variance are given in Suppl Table 1 (see Annexe 3). Additive variances were significantly different from 0 for 18 combinations of trait with landraces (out of 40 in total). few found significant dominance variance in only a single trait x landrace combination. The ratio of additive variance on genetic variance (a2) were over 0.5 for 29 combinations of traits and highly variable, even between landraces for the same trait (as illustrated for 4 traits in Figure 4). For example, PPS1052 had very often high a2 values for different traits. Traits with low additive genetic variance were NRG, K_number, K_height, and ear_top, whereas E_diam, E_length and volume of the ear had large additive genetic variance. Narrow sense heritabilities are shown in Suppl Table 1 (see Annexe 3). On average, the value of heritability was 0.358 (ranging from 0 to 0.76). On average over the 4 landraces, we found a heritability of 0.47 for E_length, 0.57 for E_diam, 0.33 for N_row, 0.26 for NKR. For multiplicative traits, the heritability was 0.53 for volume, 0.28 for K_number, 0.36 for K_weight.Variability of heritability for a same trait in different landraces is illustrated in Figure 4 for E_length, NKR, K_number and K_weight. We compared the genetic variances ($\sigma^2_A + \sigma^2_D$)

for each trait by ranking the landraces from the most diverse to the least diverse (data not shown). We observed that the ranking remained the same for all traits, with the lowest genetic variance in landrace PPS68

Choice of the old farmers

We showed some ears of the four populations to the old farmers and asked them to indicate us the ears which look the most to their cultivar. The responses were varied (Table 1)
There is no agreement between the origin of the landraces and the village of the farmers, situated either in Ariège (09, E-SWF) or in Hautes-Pyrénées (65, W-SWF) or Pyrénées Atlantiques (64, W-SWF). Thus, PPS1052 (W-SWF) had been chosen by farmers located in the East part (Moulin Neuf) and in the West part (Bun, Beaucens).

In total with 9 old farmers, 17 tests of ear selection were done (Table 2). Table 3 gives the standardized selection differentials for the 17 tests and for the 10 traits describing the ear morphology. Considering the number of case of significant standardized selection differentials and the level of this parameter, the ear length and the volume of the ear were the first traits selected by the farmers. Considering the phenotypic correlations observed among traits (Figure 3), the differentials of selections were very high and positive for the number of kernels per ear and the weight of kernels per ear. Other traits with high and positive differential of selection were the number of kernels per row, the ear diameter. For ear-top, the differentials of selection were negative, showing that farmers selected full filled ears. Low differentials of selections were found for the number of rows, ear's conicity and the kernel height.
We performed a ANOVA analysis on standardized selection differentials to test if there were significant differences between zones, informants inside a zone and landraces. The results showed no significant differences between zones (Suppl Table 2 in Annexe 3). The only significant difference among informants was observed for ear diameter because the farmer at Oursbellile selected primarily on this criterion. On PCA (Suppl Figure 2), for each landrace, we have included the mean of the 150 ears (suffix T0), and the means of the 15 ears selected by informants (suffix T1 to T9). In all cases, the selection was in the direction of larger ear diameter and volume. Differentials of selection tended to be larger for PPS68 and PPS70 (both E-SWF) than for the two other landraces for which selection differentials were more variable from one farmer to the other.

The farmers have selected sometimes the same ears. For example, 15 ears were selected three times or more for landrace PPS1052 out of seven selections. For the farmer of Campan1 and the white landrace PPS653, no significant differential selection was observed for most traits. This old farmer has selected only the full filling of the ear and their volume.

Expected response to selection

We combined our estimates of narrow sense heritability and the differentials of selection (set to 15/150 = 10% by the experimental design) to predict the univariable expected response to selection (R). The expected response in percent of the original value varied from -3.3% to 13.7% (Suppl Table 3 in Annexe 3). These values were negative only for conicity and K_height. The farmers tended to select for more cylindrical ears with smaller kernels. For the main traits (E_length, E_diam, NKR, NRG), R were positive, on average 1.1% for NRG, 1.7% for NKR, 1.6% for E_diam and 3.4% for E_length. (Suppl Table 3 in Annexe 3). For predicted traits (K_number and K_weight), R were also positive and higher than for previous traits (2.8% and 4.6% on average, respectively). For the multiplicative trait, ear's volume, R was on average equal to 7.3%. Variations were observed among landraces and farmers as illustrated in Figure 5. R values were slightly higher for PPS1052 landrace, which exhibited high additive part in genetic variance. Variability was observed among farmers as illustrated for PPS1052 in Figure 5.

Estimation of the effect of drift

We estimated as 53 ears, the number of ears needed to sow 0.25 ha of maize at density of 50 000 plants/ha, considering 220 kernels taken on the ear (on 330 kernels in total on average over the 4 landraces). These figures are those of the worst situation (small surface, small density). This size is sufficient to conserve rare allele at 1% frequency at the probability of 0.99 Crossa (1989). Also, the reduction of heterozygosity was estimated to be low with an effective population size (Ne) around 211 with Nf=53 and Nm=53*220.

Discussion

We have studied maize landraces from the south west of France that were cultivated 60 years ago and their seed management, by investigating old farmers or their children. Now, there is no longer maize landraces in the Pyrenees (with exceptions). The agriculture of this region is devoted essentially to animal breeding, and maize hybrids are only cultivated in lowlands. From 17th century, SWF farmers have developed know-how on maize cultivation, that we compared to those of native Mesoamerican farmers described in bibliography.

Farmer's management of landraces with comparison with Mesoamerican agrosystems

We analyzed 3 components of farmer's management of genetic resources: choice of the variety, seed flows, and seed selection and management (Bellon et al., 1997).

Choice of the variety

Consistent with Beigbeder and Carraretto (2018), our study indicated that the average number of varieties that was cultivated by each farmer is 1.4 (Table 1), approximately one variety per farmer. In the survey, some farmers who were feeding geese said they cultivated a white variety with small kernels (Millette). With this exception, the same variety was used for human consumption and animal feeding. There was no particularity because of maize use. In Mexico, which is the center of origin of maize, the farmers cultivated several varieties (more than 2 varieties per season (Louette and Smale, 2000); 2.5 varieties in average (Moreno et al., 2006); more than one variety (Rice et al., 1998)). They identified the variety by the color of the kernels, as also described by the old farmers from Pyrenees.

Seed flows

On the informants who remembered the exchanges, 3 said there was no exchange, 4 exchanged sometimes seed with neighbors and 2 with other villages or far away. The seed exchanges were rare and were done inside the village or the family. The exchanges seem quite less important than those reported by Louette et al. (1997) in Mexico. In Cuzalapa, farmers' own harvests represented 45% of the maize area in the study zone, whereas 40% was planted to seed from

other Cuzalapa farmers and 15% was planted to foreign introductions. (Bellon et al., 2011b) reported that 90% of all the seed lots were obtained within 10km of the community in Central Mexico. In our study, none of the informants mentioned crop losses forcing them to get seeds outside their farm. The exchanges were rather than the fact of productivity research or curiosity.

Seed selection and management

All the informants said us that the selection criteria were only based on ears, and never based on the whole plant morphology in the field. Beigbeder and Carraretto (2018) reported the same practice in their book. In Mexico, the studies on seed management by the farmers reported also a selection done on ears (Louette and Smale, 1998; Moreno et al., 2006; Rice et al., 1998). In our survey, two informants (Moulin Neuf and St Jean Pied de Port) said that the selection was done by the head of the farm. Rice et al. (1998), in Mexico (Sierra de la Marta) reported that the choice of the ears was done by the head of the farm or was supervised by him.

All the informants said us that they eliminated the top and the bottom of the ear when they collected seeds for the next generation as it was mentioned in Beigbeder and Carrarretto (2018). A picture was given in their book with two fingers delimitated the portion of the ear from which kernels were obtained. This practice was also mentioned in Mexico and Guatemala (Johannessen, 1982; Louette and Smale, 2000; Perales et al., 2003a). It is a common practice. The kernels of the top of the ear are smaller and attacked by the insects; the kernels from the bottom of the ear are more humid and crooked.

The old farmer of Moulin Neuf cultivated white varieties. When he did the test on the white landrace Aston, he carefully eliminated the multicolored ears. Pyrenean farmers did not cultivate multicolored varieties (Beigbeder and Carraretto, 2018). This was also observed by Louette and Smale (1998) in a test of comparison of random and selected samples, with a black variety which was contaminating by a white one.

Selection of ears and expected response to selection

To get genetic progress, two genetic parameters are needed: a high differential of selection which implies variability of the trait, and a high narrow sense heritability.

Heritability of the traits

On average over the 4 landraces, the estimated heritability varied from 0.26 for NKR to 0.53 for volume. . Thus in our study, the heritabilities were slightly higher for ear diameter and volume than for the other morphological traits. The ratio of additive variance on genetic variance was over 0.5 for 29 cases of 40. These high additive variance values and low dominance variance values led to get high narrow sense heritabilities and an effective selection. We compared these heritability values obtained with mating design I (Full-Sib Design and Half-Sib Design) with those of literature. Hallauer and Miranda (1981) summarized the estimate of additive variance, dominance variance and heritabilty from many studies from the literature. They indicated three classes of traits according to their heritability: h2 < 30 for yield, 30 < h2 < 50 for ear length and ear diameter, 50 < h2 < 70 for kernel-row number, with a2 values varied from 0.6 (yield) to 0.93 (ear row number) Data on intrapopulation variance estimates of ear traits are scarce. We mention the study by Malvar et al. (1996) on 2 Spanish landraces. They found lower heritability than us for ear length (0.36 and 0.21 with a2=0.38 and 0.15, respectively for the two landraces) and h2=0.63 and 0.80 for ear row number with a totally additive variance. Sandoya et al. (2009) in a synthetic population (cycle 0) found h2=41% for ear length and h2=49% for ear row number. The genetic diversity of SWF landraces was quite similar to the materials analyzed by Hallauer and Miranda (1981) and on Spanish populations. However, we can assume that the heritabilities estimated in the field at Mauguio in 2018 are overestimated compared to heritabilities that would have been calculated in the farmers' field. The cultivation conditions are supposed to be better in the experimental station than in the farmer, with therefore higher heritability estimates.

Traits under selection and variability within landrace

All informants eliminated unhealthy ears in their ear choice, as mentioned by Soleri et al. (2000). In our test, the two first traits under selection were the ear length and the ear volume. Johannessen (1982), in Guatemala, mentioned that 2/3 of the farmers selected on ear length.

For Soleri (1999), in Mexico, it was also a main criterion. Louette and Smale (1998) cited as main criteria: ear weight, ear length, length of the ear presenting kernels, total number of kernels, and kernel filling. (Serpolay-Besson et al., 2014) studied the evolution on 2 OPVs in farm in Italy and France during 2 generations, they reported that farmers applied a strong and diverse selection pressure on ear quantitative traits and always selected significantly longer ears. We found a negative differential of the top of the ear (ear_top) showing that farmers selected well filled ears. We found positive responses on volume, number of kernels of an ear, and weight of kernels of an ear. Old farmers integrated the different morphological visual characteristics of the ear to choose ears with a large number of grains and having a high grain weight. As reported by all studies, the choice of the ears was variable according the farmers. In our ANOVA analysis of selection differentials did not show significant differences among farmers except, one for ear diameter. Even if farmers have chosen the longest or largest ears, there is greater variability within farmers than between farmers, except for one farmer who selected on ear diameter.

Univariable expected response to selection

The selection method was the mass selection only on female plants. Its efficiency depended from the heritability and selection differential. After one generation, the univariable expected of selection were positive for the main traits, and even high enough for the target trait, the kernel weight of an ear. The selection differential was calculated with a selection rate of 10%. If we considered a selection rate of 1%, the selection intensity varied from 1.755 to 2.665 (Gallais, 1990). In this case (1%), the expected response to selection would be multiplied by around 1.5 compared with the case of 10%. A selection rate of 1% was more realistic. Thus, we can calculate the number of kernels obtained with one ha sowing at the density of 50 000 plants/ha. The yield is assumed to be of 1500Kg/ha, the weight of 1000 kernels of 0.373 Kg, we obtained around 4 millions of kernels/ha. With these figures, the selection rate is around 1.25%. The yield of 1500 Kg/ha was given by Derieux et al. (1987) as the average yield in France in 1950. (Louette and Smale, 1998) did a test of the choice of ears with this intensity, which was according them the usual selection pressure.

We estimated high expected change after one generation in our study. On average on the 17 test, R was 3.47% for E_length, and 4.62% for K_weight. In one part, selection rate would be upper than 10% and so R would be higher than estimated. In the other part, heritabilities were overestimated in our experimental design under good agronomical conditions, and R was

calculated on one generation without taken into account genotype * environment interaction. In literature reviewed by (Hallauer and Miranda, 1981), mass selection has been shown to be effective in modifying highly heritable traits in maize. For yield (p 214), the average gain per cycle varied from -1% to 19.1% with an average at 3.25% for 18 experiments with various selection intensities and number of cycles. We report, below, two experiences made with farmers.

A low expected response to selection was observed by Soleri (1999). She did a selection exercise in which farming households made a 10% selection of the best ears for planting seed from a 100 ears sample from their own maize population or another population from their community. Response to selection was assessed in a field trial in which three successive generations of farmer selections and two corresponding successive generations of random selections all from the same maize population were grown from each of six households, three from each community. For 8 morphological and phenological traits and 3 comparisons, the contrast was significant at 0.05 for 3 cases on the 6 households. The same trend was observed for post-harvest traits. The author concluded that the additive genetic variance for farmers' selection criteria was very low relative to environmental variance, resulting in no statistically significant response. In the partipatory maize breeding program conducted by stratified mass selection in Portugal (Alves et al., 2018), responses to selection for grain yield were significant for one population and no significant for the other population. Genotype * environment interactions reduce the expected genetic progress from one year to the next. Considering the evolution of grain yield over years from 1900 to 1930 in USA, Duvick (2005) reported no high response to selection of open-pollinated varieties with a yield under 20q/ha. Thus, mass selection is recognized as less efficient than the methods developed for the creation of hybrids, for yield.

The evolutionary forces explained the genetic structuration in SWF

Diaw et al. (2021) highlighted the existence of two genetic groups in SWF (E-SWF and W-SWF). Demographic forces (random drift, restricted gene flow with significant isolation by distance) were presented as having led to this structuration. However, adaptation by the farmers' practices could not be excluded. We tested this selective hypothesis in our study, and we found no significant difference in the selection differentials between the two areas (E-SWF vs W-SWF). Yet, the number of informants was very limited and the investigation relates to memories of a time 60 years ago. The number of informants who lived this period and still in

good health is low. Our informants remembered well the cultivation of the local maize varieties, the memories focused on the most significant events like the festivities when they were killing the pig. In addition to detailing the selection practices, our survey was also useful to document effective population sizes and thereby the strength of genetic drift. We estimated that the number of ears necessary for sowing the next generation was 53 for an area of 0.25ha. This can be considered large enough to avoid significant drift effects. Besides, the estimated response to selection in our study is high. A consequence of a high selection response could have been a reduction in variability if gene flows were low. However, the expected heterozygoty of the 4 landraces were from 0.200 to 0.225 (Diaw et al., 2021). So, the real response to selection would be lower than that we estimated and/or gene flows were high. Expected heterozygoty of a landrace measured its diversity. In our case, it is linked with morphological variability as PPS68 having the lowest variability has also the lowest expected heterozygoty.

The transition to hybrids

The transition to hybrids occurred rapidly, between 1950 and 1960 (Beigbeder and Carraretto 2018). Such a rapid abandon of the landraces is surprising and suggests that the benefits of hybrids was obvious in terms of yield, possibilities of genetic improvement, and/or cultivation practices. In general, some farmers stopped the maize cultivation, notably because of the rural exodus, and other farmers began hybrid cultivation. They well observed the differences between hybrids and their local varieties as reported by Beigbeder and Carraretto (2018). In the hybrid varieties, the kernel was less round, had less flour, were later in flowering, their husks easily detachable, coped with high density better with flat roots. The productivity gain of hybrids was clear. However, some of the farmers, such as old farmers or those who fed geese, kept local varieties longer. The current situation is the abandonment of the culture of local maize populations in the Pyrenees. Very few people still cultivate landraces in their garden or some organic farmers keep on their cultivation. Landraces are conserved ex situ, they have been researched notably for their precocity and digestibility in silage. Others have extracted doubled haploids in order to use their diversity immediately for modern breeding (Strigens et al., 2013).

Conclusion

Our objectives were (1) to know the know-how on the maize cultivation and in particular the management of seeds, (2) to assess the selection differentials practiced by farmers on the ears and to estimate the heritability in the narrow sense of the traits under selection in order to determine whether traditional farming practices and landraces are favorable for in farm genetic improvement, (3) to compare the practices in the East and West of Pyrenees to test whether they are at the origin of the genetic and morphological differentiation observed between the 2 zones.

In just over two centuries, the farmers of the Pyrenees developed a know-how for maize cultivation that look like native Mesoamericans. There was resemblance in the selection done only on ear, and in the way to collect seeds on the ear. In Pyrenees, maize was importance for human consumption and animal feeding. Although, the Pyrenees was a secondary diversification zone for maize, the farmers have integrated maize farming into their culture. The book of Beigbeder and Carraretto (2018) gives a list of words in local language describing the different parts of maize and actions.

The selection criteria were related with productivity of the ear: its length, the number of kernels per row, its volume. These traits are visually appreciable. The health of the ears was also important. Finally, for white varieties, a selection was done on the kernel color. Never Pyrenean farmers cultivated multicolored ears (Beigbeder and Carraretto 2018). The farmers practiced a high selection differential on the morphological ear traits and within populations the heritability in the narrow sense of these traits were high which could contribute to a high mass selection response which was diminished with the genotype * environment interaction effects and the lack of selection on vegetative traits such as lodging tolerance, and tolerance to high densities.

Although we observed slightly differences on ear morphological traits between West and East of SWF (later flowering date, greater row numbers in the East), the practices and selection criteria were the same in the West as East, according to our study on a small number of informants. After 60 years, it is difficult to collect know-how on old practices because of the age of the informants. We investigated the last survivors knowing the 60s maize landraces.

Acknowledgements

This research was supported by the project Amaizing ANR-10-BTBR-01 and by West African Agricultural Productivity Program (WAAPP/ISRA). We are grateful to Dr Ndiaga Cissé (Centre d'étude régional pour l'amélioration de l'adaptation à la sécheresse/ISRA, Thies, Senegal) for the grant, to Jean Beigbeder (Promais) for discussion, to Maryse Carreretto for her

aid for the survey, to Anne Zanetto and Christophe Raynaud (maize biological resources center of Montpellier) for the supply of landraces seeds, to Romain Chapuis, Serge Malavieille and their colleagues for their help of the experiment in INRA experimental unit at Mauguio. This work was supported by the CIRAD - UMR AGAP HPC Data Center of the South Green Bioinformatics platform (http://www.southgreen.fr/).

Bibliography

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome research* **19**, 1655-1664.

Alves, M. L., Belo, M., Carbas, B., Brites, C., Paulo, M., Mendes-Moreira, P., Brites, C., Bronze, M. D., Satovic, Z., and Patto, M. C. V. (2018). Long-term on-farm participatory maize breeding by stratified mass selection retains molecular diversity while improving agronomic performance. *Evolutionary Applications* **11**, 254-270.

Ardenghi, N. M. G., Rossi, G., and Guzzon, F. (2018). Back to beaked: Zea mays subsp. mays Rostrata Group in northern Italy, refugia and revival of open-pollinated maize landraces in an intensive cropping system. *PeerJ* **6**, e5123-e5123.

Badstue, L. B., Bellon, M. R., Berthaud, J., Ramirez, A., Flores, D., and Juarez, X. (2007). The dynamics of farmers' maize seed supply practices in the Central Valleys of Oaxaca, Mexico. *World Development* **35**, 1579-1593.

Beigbeder, J., and Carraretto, M. (2018). "les maïs anciens des Pyrenees : savoir et savoir-faire traditionnels," Marrimpouey/Ed., Pau.

Bellon, M. R., and Brush, S. B. (1994). KEEPERS OF MAIZE IN CHIAPAS, MEXICO. *Economic Botany* **48**, 196-209.

Bellon, M. R., Dulloo, E., Sardos, J., Thormann, I., and Burdon, J. J. (2017). In situ conservation-harnessing natural and human-derived evolutionary forces to ensure future crop adaptation. *Evolutionary applications* **10**, 965-977.

Bellon, M. R., Hodson, D., and Hellin, J. (2011a). Assessing the vulnerability of traditional maize seed systems in Mexico to climate change. *Proceedings of the National Academy of Sciences* **108**, 13432-13437.

Bellon, M. R., Hodson, D., and Hellin, J. (2011b). Assessing the vulnerability of traditional maize seed systems in Mexico to climate change. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 13432-13437.

Bellon, M. R., Pham, J. L., and Jackson, M. T. (1997). Genetic conservation: a role for rice farmers. *In* "Plant genetic conservation, the in situ approach" (N. Maxted, B. Ford-Lloyd and J. G. Hawkes, eds.), pp. 263-289. Chapman & Hall, London.

Bellon, M. R., and Risopoulos, J. (2001). Small-scale farmers expand the benefits of improved maize germplasm: A case study from Chiapas, Mexico. *World Development* **29**, 799-811.

Bocci, R., and Chable, V. (2008). Peasant seeds in Europe: Stakes and prospects. *Cahiers Agricultures* **17**, 216-221.

Brousse, C. (2011). Une analyse historique et ethnobotanique des relations entre les activités humaines et la végétation prairiale *Fourrages* **208**, 245-251

Caballero, A. (1994). Developments in the prediction of effective population size. *Heredity* **73**, 657-679.

Crossa, J. (1989). Methodologies for estimating the sample-size required for genetic conservation of outbreeding crops. *Theoretical and Applied Genetics* **77**, 153-161.

Dallard, J., Noel, P., Gouesnard, B., and Boyat, A. (2000). A network for the management of genetic resources of maize populations in France. *Plant Genetic Resources Newsletter*, 35-40.

Derieux, M., Darrigrand, M., Gallais, A., Barriere, Y., Bloc, D., and Montalant, Y. (1987). Evaluation of the genetic gain in grain maize during the last 30 years in France. *Agronomie* **7**, 1-11.

Diaw, Y., Tollon-Cordet, C., Charcosset, A., Nicolas, S. D., Madur, D., Ronfort, J., David, J., and Gouesnard, B. (2021). Genetic diversity of maize landraces from the South-West of France. *PLOS ONE* **16**, e0238334.

Doebley, J. (2004). The genetics of maize evolution. *Annu Rev Genet* **38**, 37-59.

Duvick, D. N. (2005). Genetic progress in yield of United States maize (Zea mays L.). *Maydica* **50**, 193-2002.

Gallais, A. (1990). "Théorie de la sélection en amélioration des plantes," Masson, Paris.

Gallais, A., Barriere, A., Boyat, A., Charcosset, A., Dallard, J., Derieux, M., Desselle, J. L., Dubreuil, P., Duval, H., Garnier, P., Gouesnard, B., Lavergne, V., Lefort, M., Miclo, P., Montalant, Y., Panouille, A., Pollacsek, M., Rebourg, C., Rimieri, P., Monod, J. P., Baratin, A., Baron, M., Berthe, G., Cambolive, M., Carolo, P., and Devaux, F. (2000). A French cooperative programme for management and utilization of maize genetic resources. *In* "Broadening the genetic base of crop production" (C. S. H.D. Cooper, and T. Hodgkin., ed.), pp. 331-340. Cabi, Wallingford UK.

Gilmour, A., Gogel, B., Cullis, B., and Thomson, R. (2009). ASReml User Guide release 3.0. (V.-I. Ltd., ed.), Hemel Hempstead, UK.

Gouesnard, B., Dallard, J., Panouille, A., and Boyat, A. (1997). Classification of French maize populations based on morphological traits. *Agronomie* **17**, 491-498.

Gouesnard, B., Zanetto, A., and Welcker, C. (2016). Identification of adaptation traits to drought in collections of maize landraces from southern Europe and temperate regions. *Euphytica* **209**, 565-584.

Hallauer, A., and Miranda, J. B. (1981). "Quantitative genetics in maize breeding," Iowa State University, Ames, IA.

Johannessen, C. L. (1982). Domestication process of maize continues in Guatemala. *Economic Botany* **36**, 84-99.

Louette, D., Charrier, A., and Berthaud, J. (1997). In situ conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community. *Economic Botany* **51**, 20-38.

Louette, D., and Smale, M. (1998). "Farmers' seed selection practices and maize variety characteristics in a traditionally-based Mexican community," CIMMYT.

Louette, D., and Smale, M. (2000). Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico. *Euphytica* **113**, 25-41.

Malvar, R. A., Ordas, A., Revilla, P., and Cartea, M. E. (1996). Estimates of genetic variances in two Spanish populations of maize. *Crop Science* **36**, 291-295.

Moreno, L. L., Tuxill, J., Moo, E. Y., Reyes, L. A., Alejo, J. C., and Jarvis, D. I. (2006). Traditional maize storage methods of mayan farmers in Yucatan, Mexico: Implications for seed selection and crop diversity. *Biodiversity and Conservation* **15**, 1771-1795.

Ordas, A., and Deron, A. M. (1988). A method to measure conicalness in maize. *Maydica* **33**, 261-267.

Orozco-Ramirez, Q., Ross-Ibarra, J., Santacruz-Varela, A., and Brush, S. (2016). Maize diversity associated with social origin and environmental variation in Southern Mexico. *Heredity* **116**, 477-484.

Ortiz, R., Taba, S., Tovar, V. H. C., Mezzalama, M., Xu, Y., Yan, J., and Crouch, J. H. (2010). Conserving and Enhancing Maize Genetic Resources as Global Public Goods–A Perspective from CIMMYT. *Crop Science* **50**, 13.

Pautasso, M., Aistara, G., Barnaud, A., Caillon, S., Clouvel, P., Coomes, O. T., Delêtre, M., Demeulenaere, E., Santis, P., Döring, T., Eloy, L., Emperaire, L., Garine, E., Goldringer, I., Jarvis, D., Joly, H. I., Leclerc, C., Louafi, S., Martin, P., Massol, F., McGuire, S., McKey, D., Padoch, C., Soler, C., Thomas, M., and Tramontini, S. (2012). Seed exchange networks for agrobiodiversity conservation. A review. *Agronomy for Sustainable Development* **33**, 151-175.

Perales, H., Brush, S. B., and Qualset, C. O. (2003a). Dynamic management of maize landraces in Central Mexico. *Economic Botany* **57**, 21-34.

Perales, R. H., Brush, S. B., and Qualset, C. O. (2003b). Landraces of maize in Central Mexico: an altitudinal transect. *Economic Botany* **57**, 7-20.

Portères, R. (1961). L'ethnobotanique : Place - Objet - Méthode - Philosophie. *Journal d'agriculture traditionnelle et de botanique appliquée*, 102-109.

R Development Core Team (2011). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rice, E., Smale, M., and Blanco, J. L. (1998). Farmers' use of improved seed selection practices in Mexican maize: Evidence and issues from the Sierra de Santa Marta. *World Development* **26**, 1625-1640.

Sandoya, G., Malvar, R. A., Revilla, P., and Butron, A. (2009). Effects of selection for maize resistance to Sesamia nonagrioides on the additive and dominant components of genetic variance. *Plant Breeding* **128**, 244-248.

Serpolay-Besson, E., Guiliano, S., Schermann, N., and Chable, V. (2014). Evaluation of evolution and diversity of maize open-pollinated varieties cultivaed under contrasted environmental and farmers'selection pressures : a phenotypical approach. *Open Journal of Genetics* **4**, 125-145.

Soleri, D. (1999). Developing methodologies to understand farmer-managed maize folk varieties and farmer seed selection in the Central Valleys of Oaxaca, Mexico., University of Arizona, Tucson.

Soleri, D., Smith, S. E., and Cleveland, D. A. (2000). Evaluating the potential for farmer and plant breeder collaboration: A case study of farmer maize selection in Oaxaca, Mexico. *Euphytica* **116**, 41-57.

Strigens, A., Schipprack, W., Reif, J. C., and Melchinger, A. E. (2013). Unlocking the genetic diversity of maize landraces with doubled haploids opens new avenues for breeding. *PLoS One* **8**, e57234.

Thomas, M., Dawson, J. C., Goldringer, I., and Bonneuil, C. (2011). Seed exchanges, a key to analyze crop diversity dynamics in farmer-led on-farm conservation. *Genetic Resources and Crop Evolution* **58**, 321-338.

Villa, T., Maxted, N., Scholten, M., and Ford-Lloyd, B. (2005). Defining and identifying crop landraces. *Plant Genetic Resources* **3**, 373-384.

Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B., and Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology* **79**, 13-26.

Zeven, A. C. (1998). Landraces : a review of definitions and classifications. *Euphytica* **104**, 127-139.

MAIN FIGURES

Ethnobotanical and genetic investigations of old maize landraces from the Pyrenees : seed management practices and expected response to selection.

Figure 1: Localization of the villages where the informants live numbered according to Table 1, with the localization of the collection site of the four studied maize landraces.

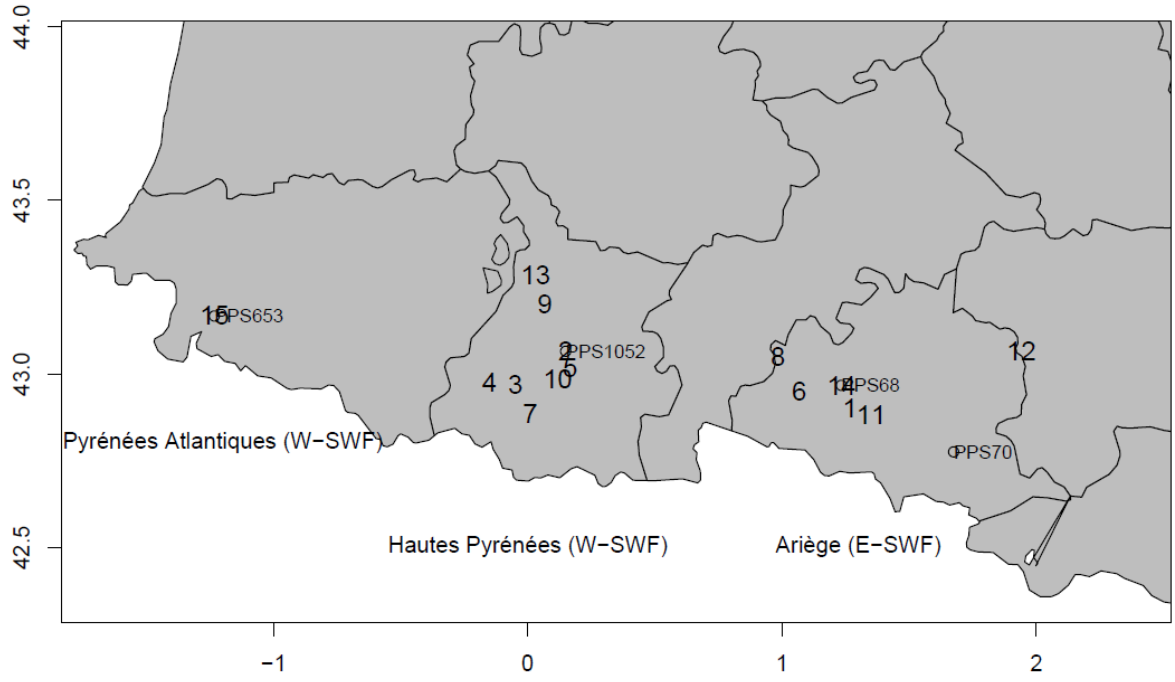


Figure 2: Boxplot for the four landraces (PPS68, PPS70, PPS653, and PPS1052) for 9 morphological traits of ear obtained from the 150 ears per landrace. *E_length or length of the ear, E_diam or diameter of the ear, conicity, N_row or number of row, NKR or number of kernel per row, K_height or height of a kernel, volume, K_number or predicted number of kernels on the ear, K_weigth or predicted weight of kernel on the ear. A same letter on the top of the boxplot indicated two populations with no significant different means.*

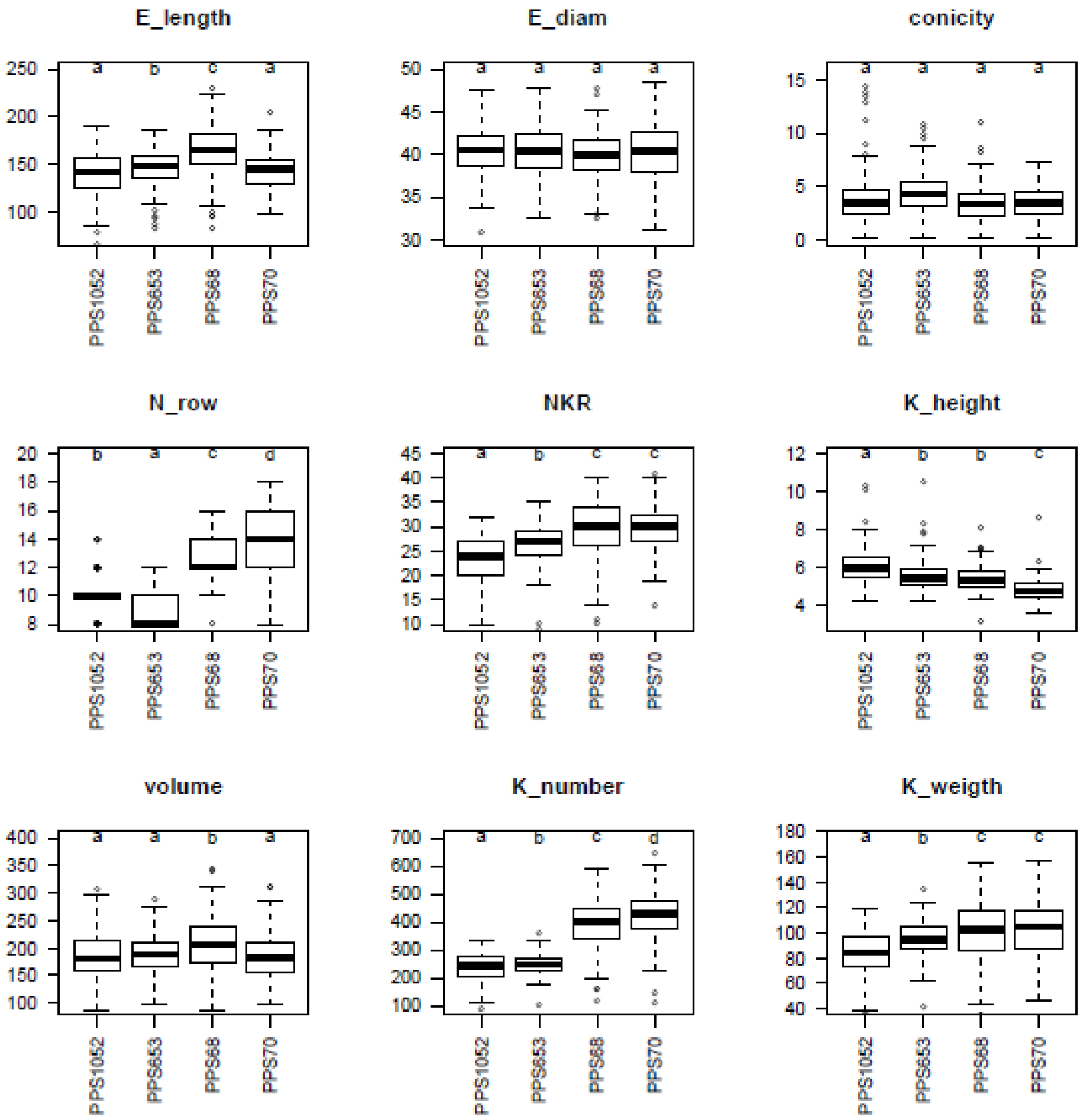
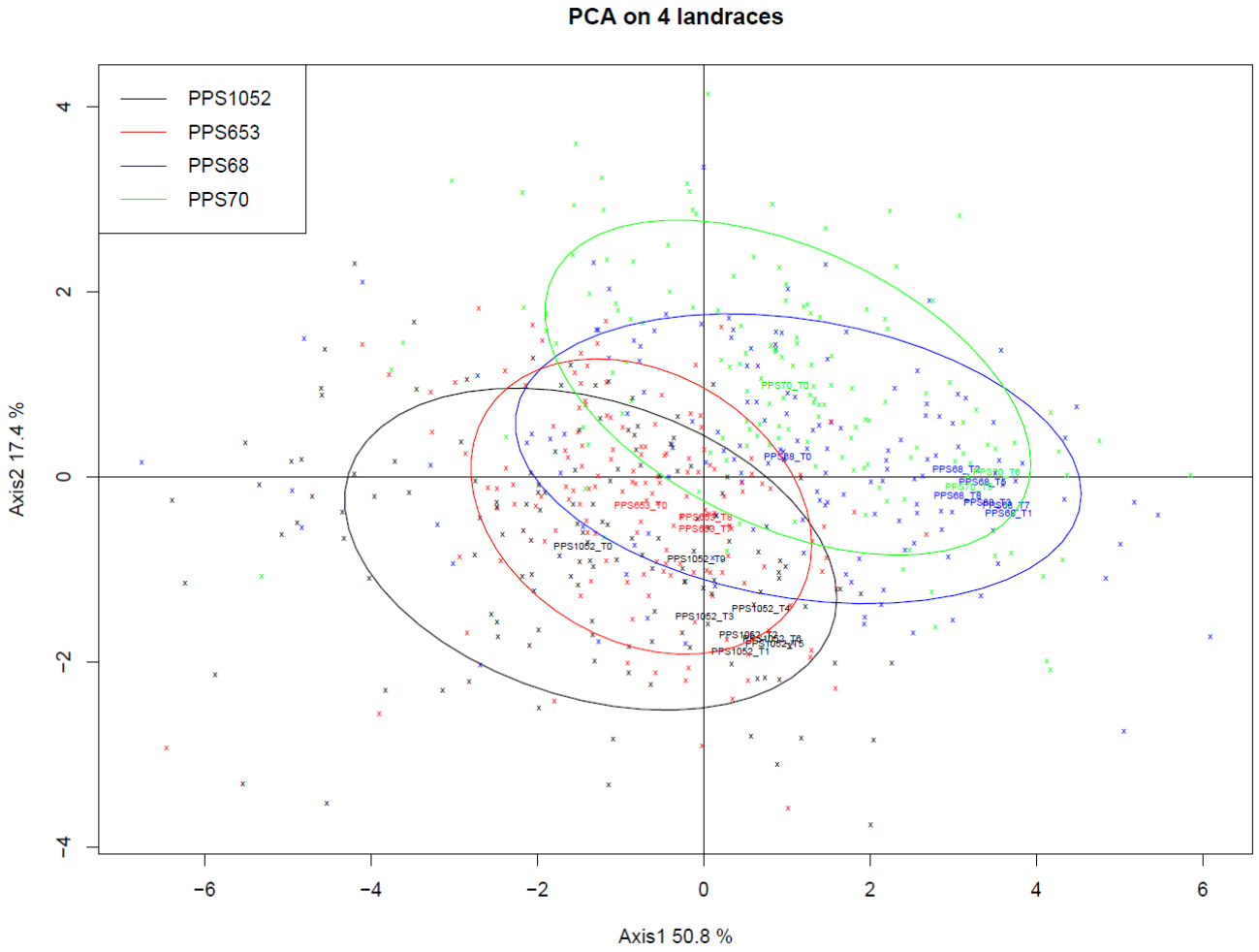


Figure 3: Principal component analysis performed on the 150 ears * 4 landraces with position of the mean of the 150 ears (T0), and the means of the 15 selected ears (from T1 to T9). *A: axis 1 and 2 for the landraces, B: axis 1 and 2 for the variables. In A: the ellipses represent 70% of the variation within the landrace.*

A



B

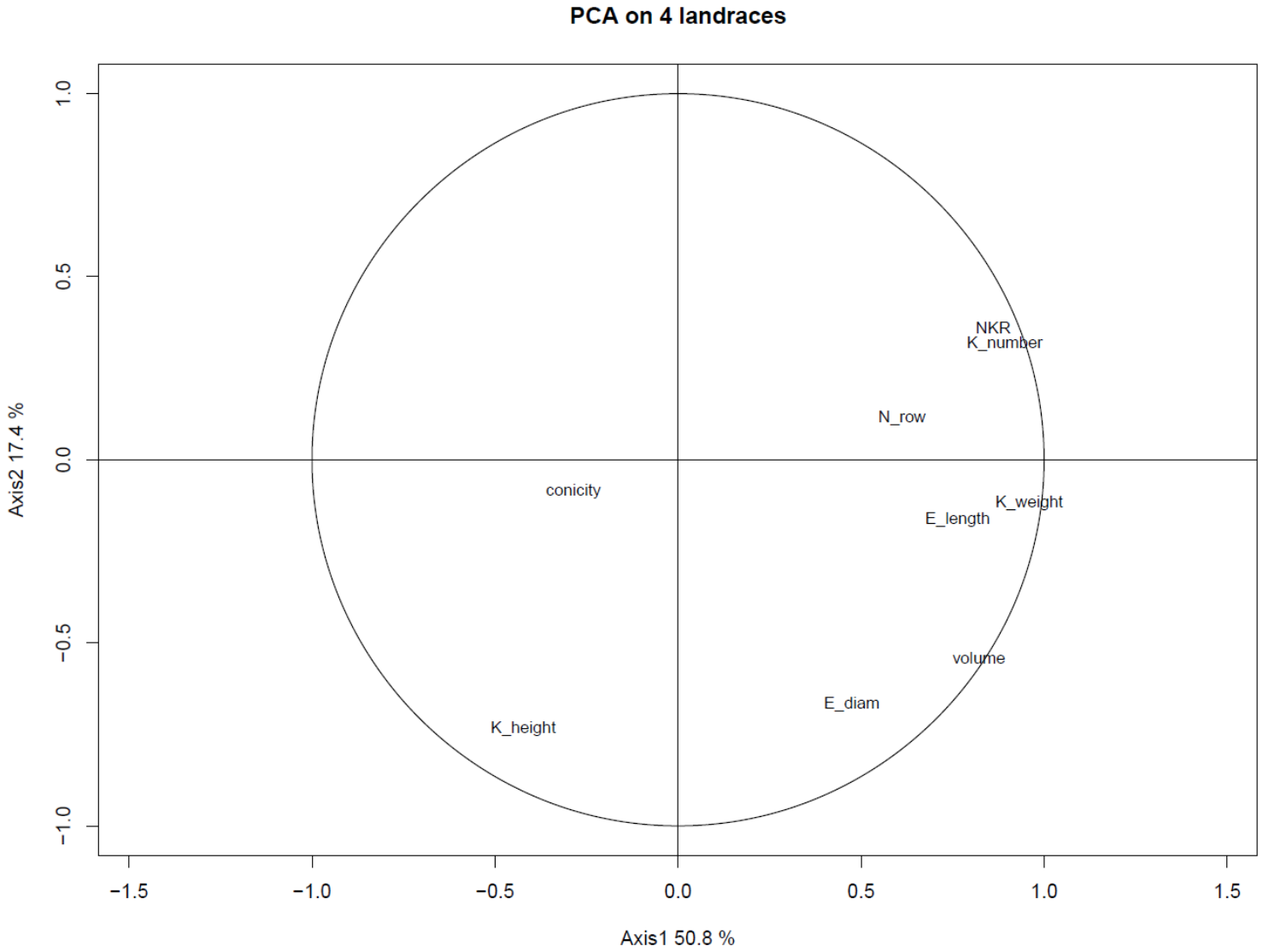

PCA on 4 landraces
Axis2 17.4 %
Axis1 50.8 %
1.0
0.5
0.0
−0.5
−1.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
NKR
K_number
N_row
conicity
K_weight
E_length
volume
E_diam
K_height

***Figure 4 : Heritabilities for 4 ear traits for the 4 landraces**. PPS68: blue, PPS70: green, PPS653: red and PPS1052:grey. The values of the part additive variance on genetic variance (a2) is given with horizontal lines.*

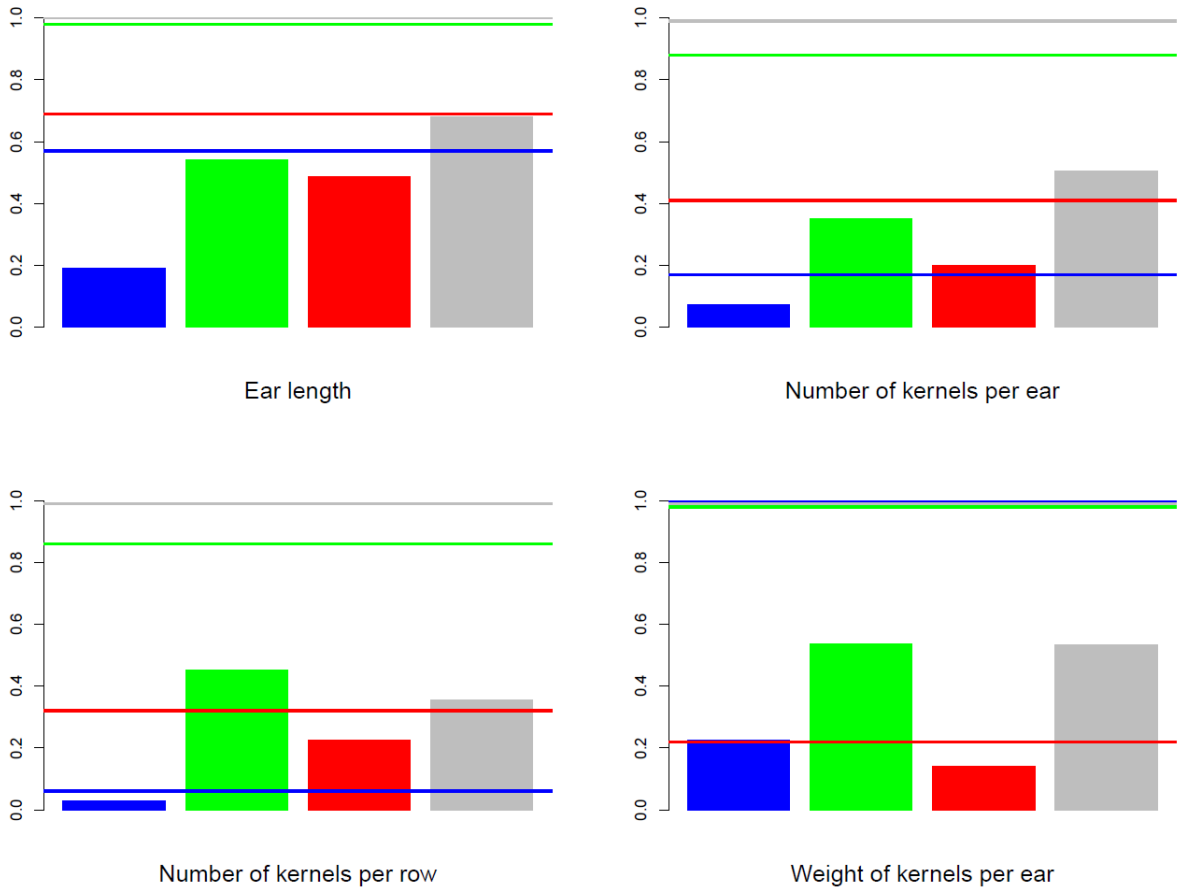
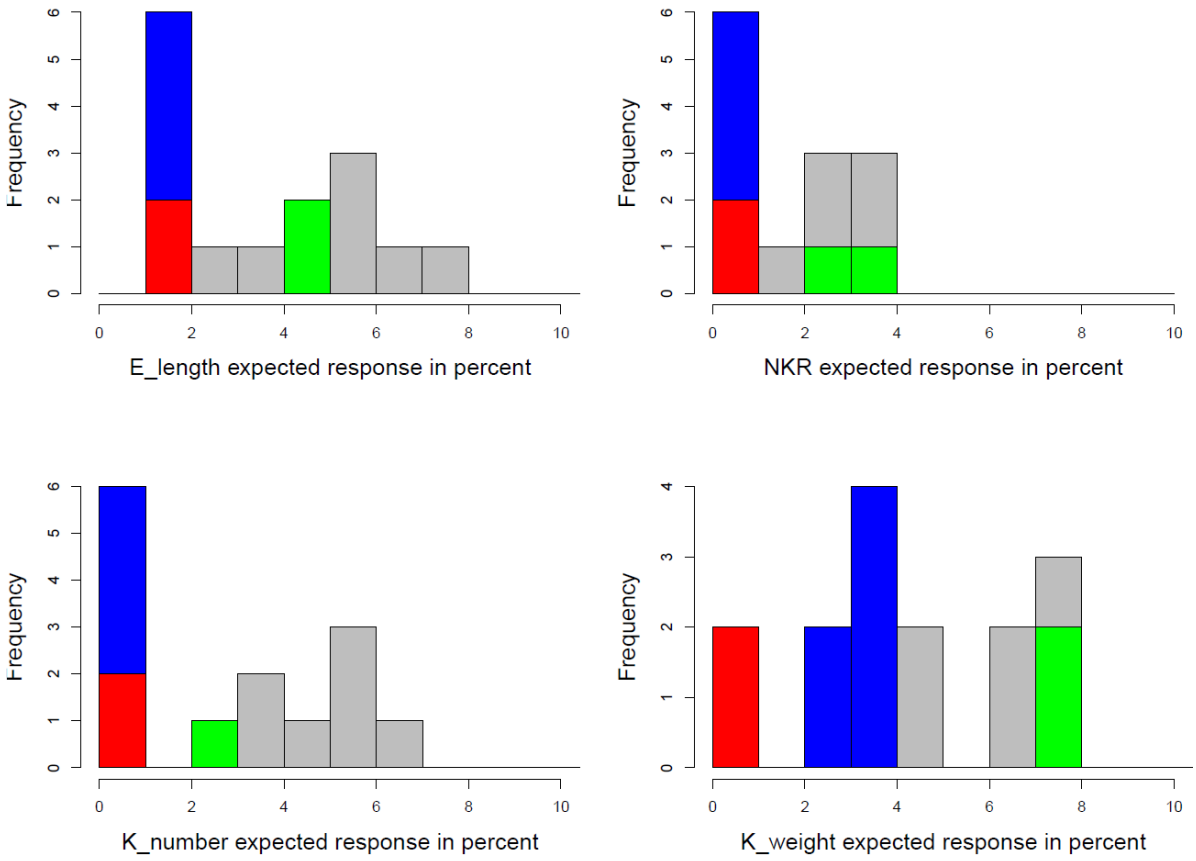


Figure 5 : Histogramme of the expected selection response (in percent of initial values) for 4 ear traits and the 4 landraces. *PPS68: blue, PPS70: green, PPS653: red; PPS1052:grey.*



MAIN TABLES

Ethnobotanical and genetic investigations of old maize landraces from the Pyrenees : seed management practices and expected response to selection.

Table 1: List of the informants given by village, with information on number and color of the varieties, exchange, force-feeding geese, type of informant. *09 Ariège for East part of the Pyrénées, 65 Hautes Pyrénées and 64 Pyrénées Atlantiques for the West part of the Pyrénées.*

village	district	Farmer code (Legend figure1)	number of landraces	exchange	force-feeding geese	Variety type	investigated person
Moulin-Neuf	09	12	2 white	?	yes	PPS1052	old farmer
Aleu	09	1	2 yellow	neighbors or further	?		old farmer
Massat	09	11	1 yellow	?	?		old farmer
Engomer	09	6	2 white	?	yes		old farmer
Riverenert 3	09	14	?	?	yes		children
Riverenert 2	09	14	?	neighbors	?		old farmer
Rivernevert 1	09	14	?	?	?		children
Lacave	09	8	2 white and 1 yellow	?	yes	PPS68	children
Oursbellile	65	13	1 white	?	no	PPS70	children
Laloubère	65	9	?	?	no		children
Bagnères de Bigorre	65	2	1 yellow and few ear of white	?	no	PPS68	old farmer
Campan 1	65	5	1 yellow	neighbors	no		old farmer
Campan 2	65	5	1 blonde and 1 copper color	no	no		old farmer
Lesponne	65	10	1 yellow	neighbors	no		old farmer
Bun	65	4	1 yellow	village around	no	PPS1052	old farmer
Beaucens	65	3	1 yellow	no	no	PPS1052	old farmer
Esquieze Serre	65	7	1 yellow	no	no		old farmer
St Jean Pied de Port	64	15	?	neighbors	no		children

Table 2: Distribution of the landraces presented to the informant for the test of the ears. *PPS1052 is a yellow landrace from the west part of the Pyrénées, PPS653 is a white landrace from the west part of the Pyrénées, PPS68 is a yellow landrace from the east part of the Pyrénées, and PPS70 is a white landrace from the east part of the Pyrénées. The column named "Test" gives the legend of Figure 3.*

Informant	district	Test	PPS1052	PPS653	PPS68	PPS70
Aleu	09	T3	x		x	
Beaucens	65	T9	x			
Bun	65	T2	x		x	
Campan1	65	T7		x	x	
Campan2	65	T8		x	x	
Engomer	09	T5	x		x	x
Esquiezes_serre	65	T1	x		x	
Moulin_Neuf	09	T4	x			
Oursbellile	65	T6	x			x

Table 3: Standardized selection differential for ten ear morphological traits for the 17 tests of ears identified by informant and landrace.
*significant test values: * at 5%, ** at 1%.*

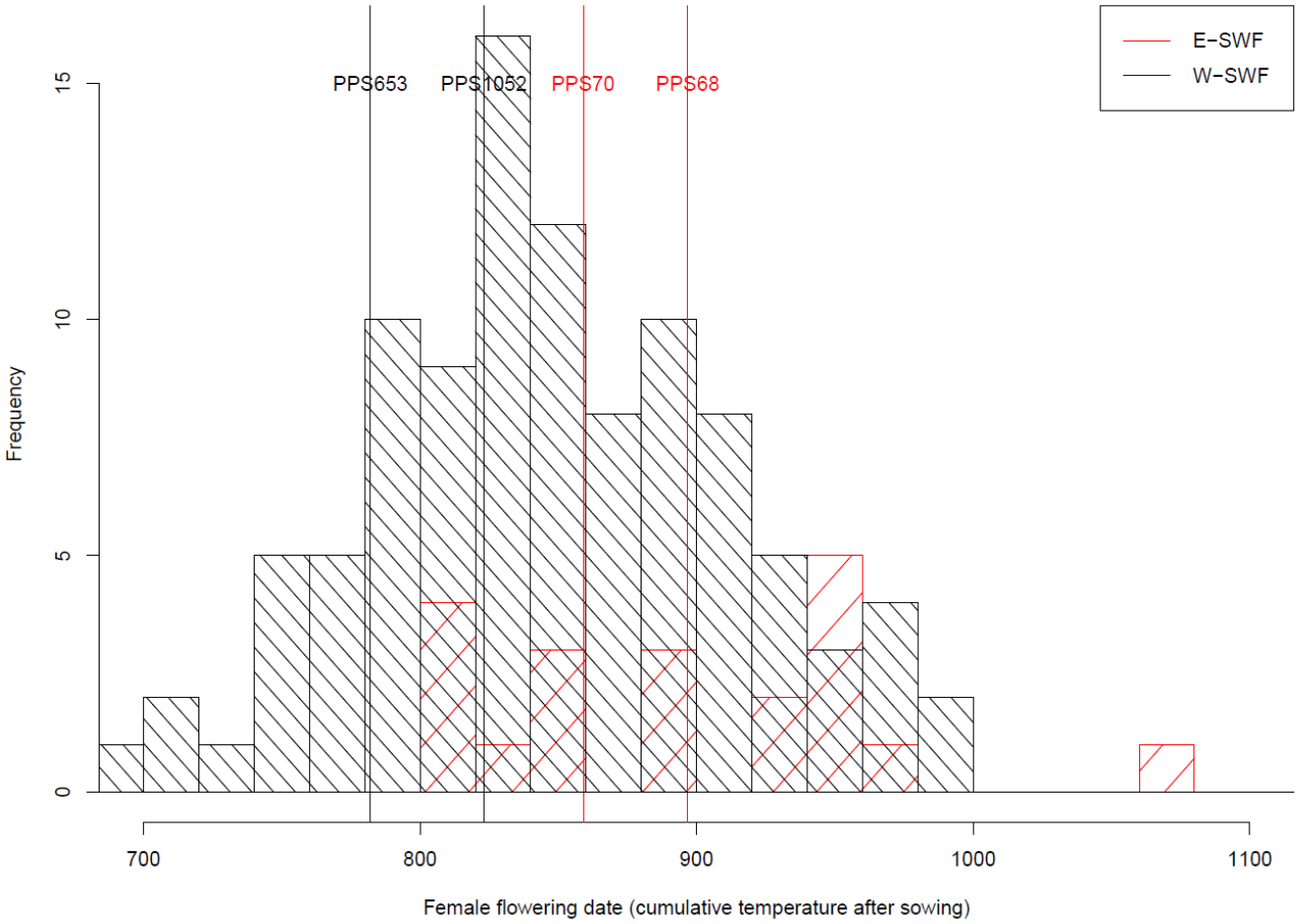
Landrace	Informant	E_length	E_diam	conicity	ear_top	N_row	NKR	K_height	volume	K_number	K_weight
PPS1052	Aleu	0.702**	0.762**	-0.251	-0.345**	0.493*	0.447**	0.081	0.908**	0.600**	0.757**
PPS68	Aleu	1.097**	0.465	-0.389	-0.296*	0.440	0.909**	0.064	1.111**	1.015**	1.066**
PPS1052	Beaucens	0.513*	0.544**	-0.150	-0.345**	0.208	0.684**	-0.41*	0.629**	0.691**	0.705**
PPS1052	Bun	0.992**	0.920**	-0.172	-0.345**	0.777**	0.644**	0.119	1.222**	0.879**	1.042**
PPS68	Bun	0.816**	0.408*	-0.368*	-0.103	0.028	0.931**	-0.42**	0.848**	0.827**	0.916**
PPS653	Campan 1	0.503	0.225	0.0083	-0.349**	0.289	0.408	0.001	0.550**	0.548	0.608
PPS68	Campan 1	0.856**	0.879**	-0.246	-0.020	0.604	0.920**	-0.34	1.233**	1.124**	1.190**
PPS653	Campan 2	0.624**	0.057	-0.324	-0.349**	0.093	0.503*	0.028	0.472	0.536	0.577
PPS68	Campan 2	0.664**	0.748**	0.0592	-0.280	0.111	0.831**	-0.38*	0.978**	0.815**	0.983**
PPS1052	Engomer	1.358**	0.864**	-0.172	-0.263*	0.588*	0.855**	0.159	1.441**	0.983**	1.187**
PPS68	Engomer	0.856**	0.599*	-0.255	-0.290	0.111	1.031**	-0.44**	1.022**	0.967**	1.089**
PPS70	Engomer	1.191**	0.878**	-0.400	-0.371**	0.238	0.803**	0.329	1.349**	0.802**	1.264**
PPS1052	Esquieze Serre	0.995**	1.010**	0.1804	-0.263*	0.683*	0.578**	0.177	1.292**	0.798**	1.011**
PPS68	Esquieze Serre	1.127**	0.615*	-0.508**	-0.449**	0.604	0.942**	0.092	1.242**	1.134**	1.191**
PPS1052	Moulin-Neuf	1.092**	0.813**	-0.178	-0.345**	0.398	0.934**	-0.16	1.200**	0.987**	1.108**
PPS1052	Oursbellile	1.025**	1.096**	-0.109	-0.307**	0.872*	0.776**	-0.01	1.382**	1.043**	1.190**
PPS70	Oursbellile	1.191**	0.953**	-0.122	-0.031	0.494	1.025**	0.005	1.429**	1.111**	1.335**

SUPPLEMENTARY FIGURES

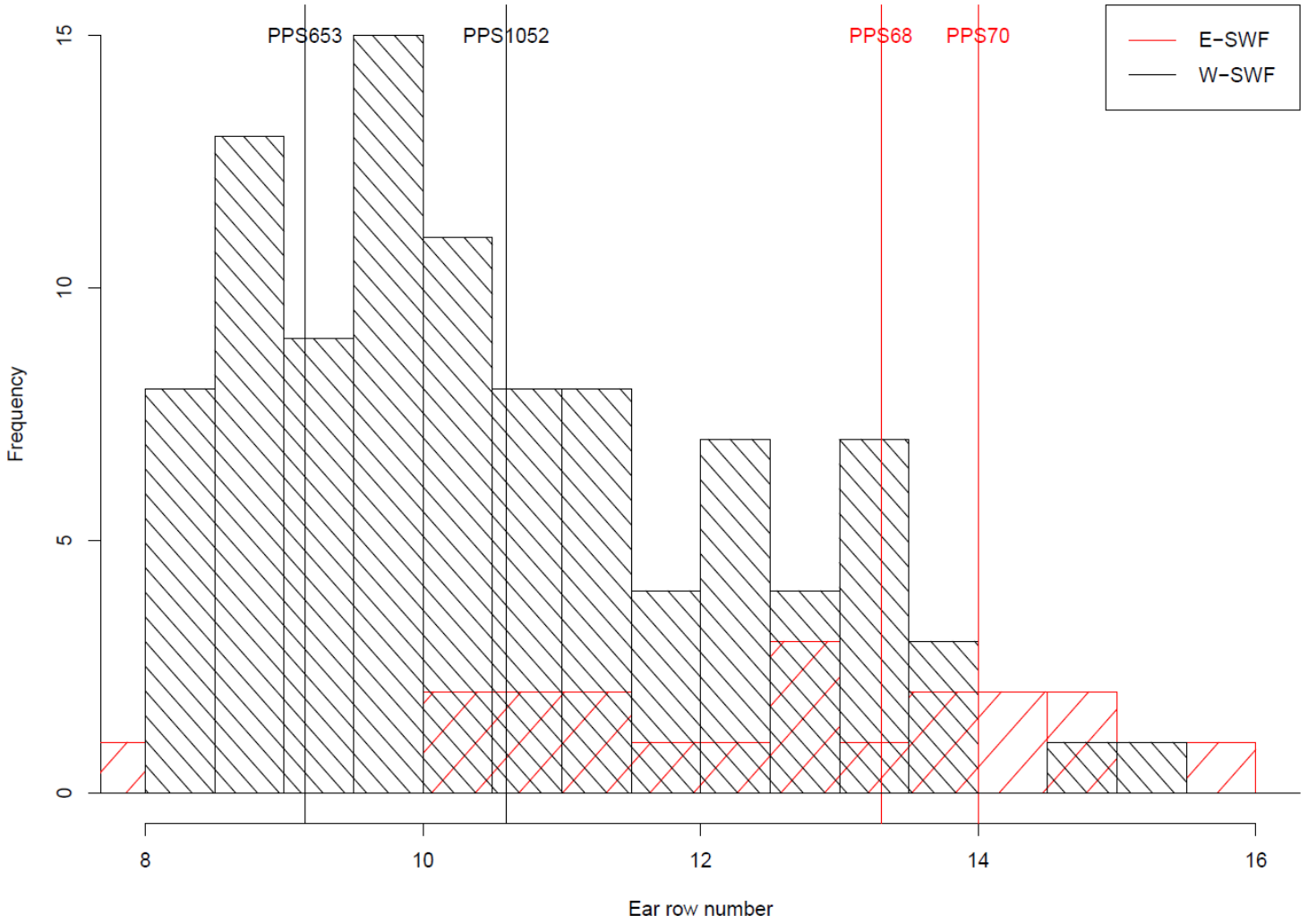
Ethnobotanical and genetic investigations of old maize landraces from the Pyrenees : seed management practices and expected response to selection

Suppl Figure 1 : Distribution of the values of Pyrenean landraces (districts 09, 31, 64, 65) for female flowering date (A) and for number of rows (B), separated in two groups (E-SWF in red and W-SWF in black). *The position of the 4 studying landraces was indicated with a line.*

A



B


PPS653
PPS1052
PPS68
PPS70
E-SWF
W-SWF
Frequency
0
5
10
15
8
10
12
14
16
Ear row number

DISCUSSION GENERALE ET PERSPECTIVES

5.1. Rappel sur les questions de recherche et les objectifs de la thèse

Comme le Sud- Ouest de la France inclut la région pyrénéenne française, nous pourrions nous attendre à ce que les populations collectées dans cette région présentent des caractéristiques similaires à celles des populations des Pyrénées (françaises et espagnoles) et de la Galice. Ces derniers forment un groupe original de maïs nommé « Pyrenees_Galicia » (Camus-Kulandaivelu et al., 2006). Cependant, un débat persiste sur les ancêtres potentiels de ce groupe comme en attestent les études suivantes :

- Dubreuil et al. (2006) ont émis l'hypothèse que les populations de maïs des Pyrénées (françaises et espagnoles) et de la Galice sont originaires d'une hybridation entre les populations de maïs issues des groupes génétiques « Northern flint » (NF) et « Caribbean » (CAR). Ils ont identifié la présence d'allèles spécifiques à ces deux groupes génétiques ; mais sur la base d'analyses de structure, les populations des Pyrénées et de la Galice se sont révélées intermédiaires entre le maïs du sud de l'Espagne et du nord de l'Europe (respectivement 27% et 17 % de leur génome).
- Mir et al. (2013) ont déclaré que les populations des Pyrénées et de la Galice sont issues d'une hybridation entre les populations de maïs caractéristiques du Nord de l'Amérique du sud et des NF, les populations des Pyrénées étant plus proches des populations du Nord de l'Amérique du Sud, et les populations de Galice montrant une plus forte proximité avec les NF.
- Brandenburg et al. (2017) ont montré que ces populations sont originaires d'une hybridation entre les NF et les populations du Sud de l'Europe (Espagne et Italie). Cette étude est basée sur l'analyse de 67 lignées consanguines de premier cycle incluant 14 lignées directement dérivées de populations « Corn Belt Dent » (CBD), 5 des NF de l'Amérique, 8 des NF de l'Europe, 13 lignées « European flint » dérivées des populations du groupe « Pyrenees_Galicia », 3 du groupe « Italian », 7 du l'Amérique du Sud, 7 des « Mexican » et 5 des « Caribbean » (CAR).

Ce qui est intéressant à connaitre sur l'histoire du maïs en Europe est que deux principales portes d'introductions du maïs ont été l'Espagne en 1493 via des populations originaires des Caraïbes donc de type CAR (Rebourg et al., 2003; Revilla Temiño et al., 2003; Revilla Temiño et al., 1998) et le Nord de l 'Europe vers le 16ème siècle avec l'introduction des formes NF depuis le nord du continent américain (Dubreuil et al., 2006; Rebourg et al., 2003). Donc ladite

zone hybride, les Pyrénées (françaises et espagnoles) et la Galice en Espagne, correspond à une zone de contact entre les populations de ces deux groupes génétiques sur l'espace européen. Cependant, entre 1493 (année d'introduction des premiers maïs en Europe) et 1960 (année de collecte des populations de maïs au SWF), l'Europe a connu d'innombrables nouvelles introductions de maïs provenant des diverses régions géographiques de l'Amérique. Mir et al. (2013) ont montré que les populations NF ont été introduites au Portugal au début du 16éme siècle. A partir 17ème siècle, des populations de maïs du Nord de l'Amérique (incluant des CBD) ont été introduites dans le Nord et Nord-Est de l'Espagne (Revilla Temiño et al., 1998).

Prenant en compte tous ces éléments, nous nous sommes intéressés à déterminer l'origine des populations locales de maïs cultivées dans le Sud-Ouest de la France (SWF) avant l'introduction des hybrides commerciaux vers 1960. Pour y arriver, nous avons proposé d'analyser la diversité génétique de ces populations à l'aide de marqueurs moléculaires et d'interroger les agriculteurs ayant cultivé ou connu ces variétés pour qu'ils nous renseignent sur les pratiques culturales et les critères de sélection mis en œuvre sur ce matériel. Nous avons utilisé des marqueurs SNP pour (1) évaluer la structure et la diversité génétique des populations traditionnelles de maïs cultivées dans le Sud-Ouest de la France jusqu'en 1960-1980 et (2) déterminer les relations génétiques entre ces populations et les populations de maïs d'Amérique et d'Europe. Nous avons par la suite utilisé des marqueurs chloroplastiques cpSSR pour suivre de manière plus directe les relations génétiques liant ces populations aux groupes génétiques de référence du maïs et approfondir ainsi notre compréhension de leur origine génétique.

Dans ce qui suit, nous rappelons comment ces différentes études nous ont permis de documenter l'histoire des populations cultivées dans le Sud-Ouest de la France. Puis nous parlerons de l'impact des pratiques des anciens agriculteurs de maïs des Pyrénées sur l'évolution de ces populations, mais aussi des contraintes que nous avons rencontrées au cours des analyses de ces données.

5.2. Origine des populations de maïs du Sud-Ouest de la France

Sur la base des études antérieures sur la diversité génétique du maïs, nous avons émis l'hypothèse que les populations du Sud-Ouest de la France sont originaires d'une hybridation entre des populations des groupes NF et CAR (Brandenburg et al., 2017; Dubreuil et al., 2006;

Rebourg et al., 2003). Les résultats des chapitre 1 et 2 sont en accord avec cette hypothèse d'hybridation mais indique également des origines plus complexes impliquant notamment des flux de gènes entre d'autres matériels tropicaux et les populations NF ; le matériel tropical étant le parent maternel (chapitre 2). Des introductions plus récentes ont été trouvées avec du matériel CBD (chapitre 1). L'implication trouvée du groupe NF dans l'origine des populations du SWF confirme leur importance dans les populations européennes (Dubreuil et al., 2006; Rebourg et al., 2003).

Deux groupes génétiques ont pu être identifiés au sein des populations SWF. Les F_{ST} obtenus entre ces groupes sur les données SNP et cpSSR sont respectivement égaux à 0.03 et 0.06, ce qui signifie que la structure des populations SWF en deux groupes (chapitre1) n'explique qu'une faible proportion de la variabilité génétique globale de ce matériel. La différentiation entre les deux groupes E-SWF et W-SWF est plus importante pour les données cpSSR que pour les données SNP. La diversité des haplotypes montre que seuls 3 haplotypes (présents chez 5.7% des populations SWF) permettent de différentier les deux groupes du SWF ; en dehors de quoi les populations SWF semblent avoir les mêmes parents maternels. Par conséquent, nous pouvons dire que les populations SWF sont issues d'un même groupe ancestral de maïs. Nos résultats ont également montré une forte proximité entre ces deux groupes et les populations du groupe « Pyrenees-Galicia ». Dans le chapitre2, nous avons observé que le fond génétique des populations « Pyrenees-Galicia2 » est identique à celui des groupes E-SWF et W-SWF (F_{ST}=0), suggérant que les premières populations de maïs pourraient avoir été introduites dans le SWF à partir de l'Espagne (Renoux, 2014).

Dans le chapitre3, nous avons étudié l'impact des pratiques des anciens agriculteurs sur l'évolution des populations de maïs dans le SWF. Une de nos hypothèses était que la structure des populations SWF en deux groupes génétiques identifiée par analyse d'ADMIXTURE (chapitre 1) pourrait être due à des méthodes de gestion des semences et de sélection différentes entre l'Est et l'Ouest de cette région géographique. En effet, la sélection par l'homme est en général basée sur des critères agro-morphologiques liés au type de consommation ou à des usages spécifiques. Elle conduit généralement à des différentiations morphologiques entre les populations cultivées. Le changement des valeurs des traits morphologiques provoqué par la sélection humaine dépend à la fois de l'héritabilité du caractère pris en compte et d'un paramètre nommé différentiel de sélection (Falconer, 1975). A travers nos enquêtes ethnobotaniques et nos analyses de génétique quantitative sur quelques populations de maïs cultivées jusque dans les années 1960 dans le SWF, nous n'avons pas détecté de différence significative sur les

critères de sélection appliqués par les anciens agriculteurs à l'Est et à l'Ouest du SWF. Nous avons observé que la variabilité des différentiels de sélection est plus variable intra zone qu'inter zone (Annexe 8). Notre enquête était cependant limitée en nombre d'agriculteurs. Par ailleurs ces populations avaient des usages particuliers que nous n'avons pas bien identifiés durant l'enquête comme le gavage des oies qui se fait dans E-SWF par la culture de populations à petits grains et grand nombre de rangs (Millette localisée dans le Lauragais, E-SWF).

Grâce à ces études, nous avons également appris que le choix des graines servant de semences pour la génération suivante se fait uniquement sur les critères morphologiques de l'épi : longueur de l'épi, le nombre de rang de grains par épi, le volume de l'épi et le nombre de grain sur le rang. Par contre, il semble que les agriculteurs ne prenaient pas en compte la précocité des plantes, qui est le principal caractère différenciant les populations de maïs NF et CAR.

La plupart des agriculteurs nous ont dit qu'ils cultivaient une seule variété dans des parcelles de cultures dont la taille variait entre 0.25ha et 1.5ha (Beigbeder and Carraretto, 2018). En considérant la plus petite surface et les densités de semis pratiquées, nous avons calculé le nombre d'épis utilisés pour le semis. Ce nombre permet de conserver les allèles rares qui ont une fréquence de 1% avec une probabilité de 0,99 (Annexe 8). Par conséquent, nous pouvons conclure que comparée à l'effet de la sélection, la dérive génétique a dû impacter faiblement la variabilité génétique des populations SWF. Compte tenu des sélections réalisées par les agriculteurs sur les différents lots, nous nous attendions à ce que les populations de maïs soient morphologiquement identiques sous l'effet de la sélection humaine. Pourtant, l'analyse de la base de données morphologiques obtenue par Gouesnard et al. (1997) montre une différenciation morphologique significative entre les populations des groupe E-SWF et W-SWF . Les populations du groupe W-SWF sont plus précoces, ont un plus petit nombre de rangs de grains par épi et des grains plus gros que les populations E-SWF. Les tests de choix des épis ont montré un fort différentiel de sélection : les agriculteurs savent parfaitement repérer les épis les plus productifs. Nous pourrions nous attendre à une baisse de la diversité du fait de la sélection, pourtant les analyses SNP ont montré une forte diversité au sein des variétés (chapitre 1). Les enquêtes ont montré que certains agriculteurs pouvaient échanger des semences, le plus souvent au sein du village mais parfois plus loin.

5.3. Analyse critique des approches sur « bulk » d'ADN et à partir d'une plante par population dans les études de diversité

Les marqueurs génétiques représentent des outils inestimables pour étudier la diversité génétique des populations d'espèces animales et végétales. Les analyses moléculaires avec des marqueurs SNP et plusieurs individus analysés par population chez le maïs sont assez rares. On peut citer les travaux des équipes allemandes : Mayer et al. (2017) avec plus de 20 individus sur 35 populations à l'aide de la puce 600K SNP; Unterseer et al. (2016) avec entre 22 et 24 individus pour 38 populations avec la puce 600K SNP; et Hölker et al. (2019) avec 48 individus pour 3 populations avec la même puce SNP. Les résultats de ces études montrent que l'indice de fixation (F_{IS}) de Wright (1978) est faible et non significatif pour la plupart des populations. La diversité moyenne est de l'ordre de 75 % à l'échelle intra-population (73.1 % dans Mayer et al, 2017), en accord avec les résultats obtenus sur des marqueurs RFLP ou à partir de marqueurs iso-enzymatiques (où la part de la variance intra-population représentant environ ¾ de la diversité totale observée ; Rebourg et al. (2003)). Cependant, l'analyse de la variabilité génétique inter-population et intra-population pour un grand nombre de populations nécessite d'importants moyens. Face à cette lourdeur, deux types d'approches de génotypages moins longues et moins onéreuses permettant d'analyser la diversité génétique peuvent être utilisés.

Une première stratégie consiste à réaliser un génotypage sur « bulk » d'ADN. L'ADN est obtenu en mélangeant plusieurs individus par population. L'avantage de cette méthode est qu'elle permet d'estimer la variabilité génétique intra-population sur un grand nombre d'individus pour un coût relativement réduit (i.e. estimation des fréquences alléliques sur 15 individus/30 allèles pour les marqueurs SNPs dans notre étude). Cependant, l'approche « bulk » ne donne pas accès à l'information sur le génotype des individus qui composent chaque population (Dubreuil et al., 1999; Reyes-Valdes et al., 2013). Ceci réduit notre capacité à utiliser certains logiciels d'analyse de génétique des populations (ADMIXTURE d'Alexander et al. (2009), GENEPOP de Raymond and Rousset (1995), HZAR de Derryberry et al. (2014) and STRUCTURE de Pritchard et al. (2000)), et d'estimation des paramètres démographiques et historiques (méthode ABC, Bertorelle et al. (2010)) pour inférer l'origine génétique des populations.

La deuxième méthode de génotypage est l'approche une plante individuelle par population. C'est une méthode d'analyse qui permet d'étudier la diversité génétique uniquement entre populations d'une espèce à grande échelle avec un seul individu par population. On peut citer les travaux de van Heerwaarden et al. (2011) (468 SNPs) sur la diversité des téosintes et des populations de maïs et de Westengen et al. (2012) sur la diversité des maïs africains. Les résultats *Zea mays subsp. parviglumis*de van Heerwaarden et al. (2011) sont cohérents avec l'histoire connue sur la domestication du maïs à partir de *Zea mays subsp. parviglumis* et sa propagation à travers le monde. Sur les nSSR, Vigouroux et al. (2008) ont analysé 945 accessions de maïs représenté par un total de 964 populations pour étudier l'histoire de la domestication et de la diffusion du maïs en Amérique. Ils ont trouvé une faible différenciation entre les 4 groupes de populations identifiés avec le programme structure (environ 8%). L'avantage de cette approche est que les logiciels sur l'analyse de la diversité utilisant des données sur le génotype individuel peuvent être utilisés. Contrairement à l'approche plante individuelle, l'approche « bulk » a l'intérêt de prendre en compte la diversité intra-population au sein d'une espèce allogame tel que le maïs. Ainsi, la population est caractérisée par les fréquences alléliques des différents allèles obtenus sur les marqueurs utilisés. Nous avons choisi la méthode bulk pour les SNP qui prend en compte la variabilité intra-population et permet l'estimation du taux d'hétérozygotie. Et, nous avons dû simuler des haplotypes pour l'analyse d'ADMIXTURE sous l'hypothèse d'équilibre Hardy-Weinberg.

L'approche bulk a fait l'objet de nombreux articles où des auteurs l'ont comparé à l'analyse de plusieurs individus par populations, avec des isoenzymes ou des marqueurs moléculaires tels que les RFLP et les SSR. En 2015, au cours de mon stage de Master 2, j'ai pu réaliser une comparaison de l'approche bulk et l'approche une plante individuelle en analyse SNP sur des données expérimentales acquises sur 119 populations de maïs des Pyrénées et 16 populations témoins représentant les différents groupes génétiques à l'origine des populations européennes (Diaw, comm. perso). Les premiers résultats de cette comparaison ont montré une corrélation significative de 0.71 entre les distances de Rogers par paires de populations calculées selon les deux approches.

L'approche bulk nécessite des calibrations pour pouvoir estimer les fréquences alléliques. Pour les SNP, Arca et al. (2021) ont construit des courbes de calibration avec des mélanges en proportion connue. Pour les SSR, on peut confondre la présence d'un allèle avec un « stutter » (pic supplémentaire) de l'allèle suivant en poids moléculaire. L'analyse bulk a été pratiquée pour les nSSR avec la méthode de correction définie par Pierre Dubreuil (Dubreuil et al., 1999;

Dubreuil et al., 2006). Elle consiste en la sélection de nSSR répétables et l'analyse de témoins pour déterminer le « stutter ratio » afin de corriger les hauteurs de pics. Avec nos 7 marqueurs cpSSR, nous n'avons pas pu faire cette correction. Nous avons choisi les marqueurs de l'étude de Provan et al. (1999) qui se montraient polymorphes et stables : 6 cpSSR auxquels nous avons rajouté un marqueur supplémentaire venant du riz. Notre étude de l'analyse des cpSSR basée sur une plante individuelle par population ne conduit pas à des cas de confusion sur la détermination des haplotypes comme avec une analyse en bulk, qui a été réalisée au laboratoire avant mes travaux.

5.4. Conclusion et perspectives

Les diverses analyses effectuées au cours de cette thèse nous ont permis de documenter la variabilité moléculaire des populations cultivées dans le Sud-Ouest de la France avant l'arrivée des formes hybrides, de mieux comprendre leur histoire et d'identifier les relations multiples entre ces populations de maïs et les grands groupes de référence du maïs d'Amérique et d'Europe. Pour cela nous avons utilisé deux méthodes de génotypage à savoir le « bulk » d'ADN sur les données SNP et l'analyse par plante individuelle sur les données cpSSR. Parmi les études qu'il me semblerait intéressant de développer, il serait utile de poursuivre l'analyse des erreurs induites par les approches bulk (voir ci-dessus, l'étude que j'ai réalisée pendant mon Master). Il serait notamment intéressant d'analyser l'impact de l'échantillonnage des plantes sur la structure de la diversité via des approches analytiques et par simulations. Dans ce dernier cas, il s'agira de construire une distribution des fréquences alléliques imitant la distribution réelle (forme globalement en U).

Dans le chapitre2, nous avons choisi d'analyser un seul individu par population pour des questions de fiabilités sur l'interprétation des résultats. Etant donné les défauts liés aux marqueurs SSR et notamment les risques liés à l'homoplasie, une perspective de notre étude serait de développer des approches de re-séquençage de l'ADN chloroplastique qui est actuellement la méthode de génotypage le plus utilisée dans les études de phylogéographie basées sur les chloroplastes (Morris and Shaw, 2018). Une telle étude chez le maïs nécessitera le génotypage de génomes chloroplastiques. La méthode d'analyse proposée par Mariac et al. (2014) pourrait être une solution pour récupérer des génomes de chloroplaste entier à un coût modeste ; le génome chloroplastique de l'accession de maïs séquencé par Maier et al. (1995) servant alors de base de référence pour cette étude.

Une autre perspective serait d'analyser l'origine génétique des populations de maïs d'Italie. Dans les études de Camus-Kulandaivelu et al. (2006) et de Rebourg et al. (2001) , les populations italiennes utilisées forment un groupe qui se montre différencié des autres populations de maïs d'Amérique et d'Europe. Tenaillon and Charcosset (2011) ont émis l'hypothèse d'une troisième introduction indépendante de maïs d'Amérique qui seraient les potentielles ancêtres du matériel italien en Europe. Tandis que Brandenburg et al. (2017) indiquent plutôt une émergence de lignées italiennes à partir d'un mélange entre les cornés espagnols et cornés européens. Dans l'analyse des populations pyrénéennes en 2015 (comm. pers.), les populations italiennes ne paraissaient pas avoir d'incidence sur la diversité observée. En effet, aucune introgression de génome de populations italiennes n'a été observée au sein des populations des Pyrénées. L'analyse cpSSR a montré une forte différence entre ces populations et les populations du sud-ouest de la France, l'haplotype F, fortement caractéristique des populations NF d'Amérique étant prépondérant dans le groupe italien (chapitre2).

Dans cette étude sur l'origine des populations de maïs du SWF nous n'avons pas mis en œuvre des études de simulations basées sur l'ABC (ou « approximate bayesian computation »). Du fait des origines multiples mises en évidence, les scénarios auraient été difficiles à définir et à tester et nous ne disposions pas du génotypage de plantes individuelles. Ceci reste une perspective à cette étude.

BIBLIOGRAPHIE

Bibliographie

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome research* **19**, 1655-1664.

Alves, M. L., Belo, M., Carbas, B., Brites, C., Paulo, M., Mendes-Moreira, P., Brites, C., Bronze, M. D., Satovic, Z., and Patto, M. C. V. (2018). Long-term on-farm participatory maize breeding by stratified mass selection retains molecular diversity while improving agronomic performance. *Evolutionary Applications* **11**, 254-270.

Aoki, K., Ueno, S., Kamijo, T., Setoguchi, H., Murakami, N., Kato, M., and Tsumura, Y. (2016). Detecting east–west genetic differentiation in Castanopsis (Fagaceae) on the main islands of Japan and north–south on the Ryukyu Islands, based on chloroplast haplotypes. *Plant Systematics and Evolution* **302**, 1093-1107.

Arca, M., Gouesnard, B., Mary-Huard, T., Le Paslier, M.-C., Bauland, C., Combes, V., Madur, D., Charcosset, A., and Nicolas, S. D. (2020). Genome-wide SNP genotyping of DNA pools identifies untapped landraces and genomic regions that could enrich the maize breeding pool. *bioRxiv*, 2020.09.30.321018.

Arca, M., Mary-Huard, T., Gouesnard, B., Bérard, A., Bauland, C., Combes, V., Madur, D., Charcosset, A., and Nicolas, S. D. (2021). Deciphering the Genetic Diversity of Landraces With High-Throughput SNP Genotyping of DNA Bulks: Methodology and Application to the Maize 50k Array. *Frontiers in Plant Science* **11**.

Ardenghi, N. M. G., Rossi, G., and Guzzon, F. (2018). Back to beaked: Zea mays subsp. mays Rostrata Group in northern Italy, refugia and revival of open-pollinated maize landraces in an intensive cropping system. *PeerJ* **6**, e5123-e5123.

Badstue, L. B., Bellon, M. R., Berthaud, J., Ramirez, A., Flores, D., and Juarez, X. (2007). The dynamics of farmers' maize seed supply practices in the Central Valleys of Oaxaca, Mexico. *World Development* **35**, 1579-1593.

Baltazar, B. M., de Jesús Sánchez-Gonzalez, J., de la Cruz-Larios, L., and Schoper, J. B. (2005). Pollination between maize and teosinte: an important determinant of gene flow in Mexico. *Theor Appl Genet* **110**, 519-26.

Beigbeder, J., and Carraretto, M. (2018). "les maïs anciens des Pyrenees : savoir et savoir-faire traditionnels," Marrimpouey/Ed., Pau.

Bellon, M. R., and Brush, S. B. (1994). KEEPERS OF MAIZE IN CHIAPAS, MEXICO. *Economic Botany* **48**, 196-209.

Bellon, M. R., Dulloo, E., Sardos, J., Thormann, I., and Burdon, J. J. (2017). In situ conservation-harnessing natural and human-derived evolutionary forces to ensure future crop adaptation. *Evolutionary applications* **10**, 965-977.

Bellon, M. R., Hodson, D., and Hellin, J. (2011). Assessing the vulnerability of traditional maize seed systems in Mexico to climate change. *Proceedings of the National Academy of Sciences* **108**, 13432-13437.

Bellon, M. R., and Risopoulos, J. (2001). Small-scale farmers expand the benefits of improved maize germplasm: A case study from Chiapas, Mexico. *World Development* **29**, 799-811.

Bertorelle, G., Benazzo, A., and Mona, S. (2010). ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Mol Ecol* **19**, 2609-25.

Besnard, G., Hernandez, P., Khadari, B., Dorado, G., and Savolainen, V. (2011). Genomic profiling of plastid DNA variation in the Mediterranean olive tree. *Bmc Plant Biology* **11**.

Birky, C. W. (1995). Uniparental inheritance of mitochondrial and chloroplast genes-mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 11331-11338.

Bitocchi, E., Nanni, L., Rossi, M., Rau, D., Bellucci, E., Giardini, A., Buonamici, A., Vendramin, G. G., and Papa, R. (2009). Introgression from modern hybrid varieties into landrace populations of maize (Zea mays ssp mays L.) in central Italy. *Molecular Ecology* **18**, 603-621.

Bonneuil, C. (2019). Seeing nature as a 'universal store of genes': How biological diversity became 'genetic resources', 1890–1940. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* **75**, 1-14.

Bonneuil, C., and Fenzi, M. (2011). Des ressources génétiques à la biodiversité cultivée. La carrière d'un problème public mondial. *Revue d'anthropologie des connaissances* **5, 2**, 206-233.

Bouchet, S., Servin, B., Bertin, P., Madur, D., Combes, V., Dumas, F., Brunel, D., Laborde, J., Charcosset, A., and Nicolas, S. (2013). Adaptation of Maize to Temperate Climates: Mid-Density Genome-Wide Association Genetics and Diversity Patterns Reveal Key Genomic Regions, with a Major Contribution of the Vgt2 (ZCN8) Locus. *PLoS One* **8**.

Brandenburg, J. T., Mary-Huard, T., Rigaill, G., Hearne, S. J., Corti, H., Joets, J., Vitte, C., Charcosset, A., Nicolas, S. D., and Tenaillon, M. I. (2017). Independent introductions and admixtures have contributed to adaptation of European maize and its American counterparts. *PLoS Genet* **13**, e1006666.

Brousse, C. (2011). Une analyse historique et ethnobotanique des relations entre les activités humaines et la végétation prairiale *Fourrages* **208**, 245-251

Bucci, G., Anzidei, M., Madaghiele, A., and Vendramin, G. G. (1998). Detection of haplotypic variation and natural hybridization in halepensis-complex pine species using chloroplast simple sequence repeat (SSR) markers. *Molecular Ecology* **7**, 1633-1643.

Camus-Kulandaivelu, L., Veyrieras, J. B., Madur, D., Combes, V., Fourmann, M., Barraud, S., Dubreuil, P., Gouesnard, B., Manicacci, D., and Charcosset, A. (2006). Maize adaptation to temperate climate: Relationship between population structure and polymorphism in the Dwarf8 gene. *Genetics* **172**, 2449-2463.

Carraretto, M. (2005). "Histoire du maïs, d'une divinité amérindienne à ses avatars transgéniques," CTHS/Ed., Paris.

Casanas, F., Simo, J., Casals, J., and Prohens, J. (2017). Toward an Evolved Concept of Landrace. *Frontiers in Plant Science* **8**.

Ceylan, A., Ocal, N., and Akbulut, M. (2014). Genetic diversity among the Turkish common bean cultivars (Phaseolus vulgaris L.) as assessed by SRAP, POGP and cpSSR markers. *Biochemical Systematics and Ecology* **54**, 219-229.

Chat, J., Decroocq, S., Decroocq, V., and Petit, R. J. (2002). A case of chloroplast heteroplasmy in kiwifruit (Actinidia deliciosa) that is not transmitted during sexual reproduction. *Journal of Heredity* **93**, 293-300.

Chmielewski, M., Meyza, K., Chybicki, I. J., Dzialuk, A., Litkowiec, M., and Burczyk, J. (2015). Chloroplast microsatellites as a tool for phylogeographic studies: the case of white oaks in Poland. *Iforest-Biogeosciences and Forestry* **8**, 765-771.

Conde, M. F., Pring, D. R., and Levings, C. S., III (1979). Maternal inheritance of organelle DNA's in Zea mays-Zea perennis reciprocal crosses. *Journal of Heredity* **70**, 2-4.

Crossa, J. (1989). Methodologies for estimating the sample-size required for genetic conservation of outbreeding crops. *Theoretical and Applied Genetics* **77**, 153-161.

Cubas, P., Pardo, C., and Tahiri, H. (2005). Genetic variation and relationships among Ulex (Fabaceae) species in southern Spain and northern Morocco assessed by chloroplast microsatellite (cpSSR) markers. *American Journal of Botany* **92**, 2031-2043.

Dallard, J., Noel, P., Gouesnard, B., and Boyat, A. (2000). A network for the management of genetic resources of maize populations in France. *Plant Genetic Resources Newsletter*, 35-40.

de Vienne, D. (1998). "Les marqueurs moléculaires en génétique et biotechnologies végétales," Quae/Ed. INRA, Paris.

Derryberry, E. P., Derryberry, G. E., Maley, J. M., and Brumfield, R. T. (2014). hzar: hybrid zone analysis using an R software package. *Mol Ecol Resour* **14**, 652-663.

Diaw, Y., Tollon-Cordet, C., Charcosset, A., Nicolas, S. D., Madur, D., Ronfort, J., David, J., and Gouesnard, B. (2021). Genetic diversity of maize landraces from the South-West of France. *PLOS ONE* **16**, e0238334.

Doebley, J. (2004). The genetics of maize evolution. *Annu Rev Genet* **38**, 37-59.

Doebley, J., Wendel, J. D., Smith, J. S. C., Stuber, C. W., and Goodman, M. M. (1988). The origin of Cornbelt maize - the isozyme evidence. *Economic Botany* **42**, 120-131.

Dubreuil, P., Rebourg, C., Merlino, M., and Charcosset, A. (1999). Evaluation of a DNA pooled-sampling strategy for estimating the RFLP diversity of maize populations. *Plant Molecular Biology Reporter* **17**, 123-138.

Dubreuil, P., Warburton, M., Chastanet, M., Hoisington, D., and Charcosset, A. (2006). More on the introduction of temperate maize into Europe: Large-scale bulk SSR genotyping and new historical elements. *Maydica* **51**, 281-291.

Dzialuk, A., Muchewicz, E., Boratynski, A., Montserrat, J. M., Boratynska, K., and Burczyk, J. (2009). Genetic variation of Pinus uncinata (Pinaceae) in the Pyrenees determined with cpSSR markers. *Plant Systematics and Evolution* **277**, 197-205.

Eliades, N.-G., and Eliades, D. (2009). "HAPLOTYPE ANALYSIS: software for analysis of haplotype data."

Ellstrand, N. C., Garner, L. C., Hegde, S., Guadagnuolo, R., and Blancas, L. (2007). Spontaneous hybridization between maize and teosinte. *J Hered* **98**, 183-7.

Ennos, R. A. (1994). Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* **72**, 250-259.

Eschholz, T. W., Peter, R., Stamp, P., and Hund, A. (2008). Genetic diversity of Swiss maize (Zea mays L. ssp. mays) assessed with individuals and bulks on agarose gels. *Genetic Resources and Crop Evolution* **55**, 971-983.

Estoup, A., Jarne, P., and Cornuet, J. M. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* **11**, 1591-1604.

Falconer, D. S. (1975). "Introduction to quantitative genetics," Pearson Education India.

FAO (2020a). Commission des ressources génétiques pour l'alimentation et l'agriculture (CGRFA). Vol. 2020.

FAO (2020b). Traité international sur les ressources phytogénétiques pour l'alimentation et l'agriculture (TIRPAA).

Fenzi, M., and Bonneuil, C. (2016). From “Genetic Resources” to “Ecosystems Services”: A Century of Science and Global Policies for Crop Diversity Conservation. *Culture, Agriculture, Food and Environment* **38**, 72-83.

Ferreira, V., Castro, I., Rocha, J., Crespi, A. L., Pinto-Carnide, O., Amich, F., Almeida, R., and Carnide, V. (2015). Chloroplast and nuclear DNA studies in Iberian Peninsula endemic Silene scabriflora subspecies using cpSSR and ISSR markers: Genetic diversity and phylogenetic relationships. *Biochemical Systematics and Ecology* **61**, 312-318.

Frey, J. E., Frey, B., and Forcioli, D. (2005). Quantitative assessment of heteroplasmy levels in Senecio vulgaris chloroplast DNA. *Genetica* **123**, 255-261.

Ganal, M. W., Durstewitz, G., Polley, A., Berard, A., Buckler, E. S., Charcosset, A., Clarke, J. D., Graner, E. M., Hansen, M., Joets, J., Le Paslier, M. C., McMullen, M. D., Montalent, P., Rose, M., Schon, C. C., Sun, Q., Walter, H., Martin, O. C., and Falque, M. (2011). A large maize (Zea mays L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS One* **6**, e28334.

Garza, J. C., and Freimer, N. B. (1996). Homoplasy for size at microsatellite loci in humans and chimpanzees. *Genome Research* **6**, 211-217.

Gauthier, P., Gouesnard, B., Dallard, J., Redaelli, R., Rebourg, C., Charcosset, A., and Boyat, A. (2002). RFLP diversity and relationships among traditional European maize populations. *Theoretical and Applied Genetics* **105**, 91-99.

GCDT (2007). Global strategy for the ex-situ conservation and utilization of Maize germplasm. pp. 62. Global Crops Diversity Trust, Rome, Italy.

GNIS (2020). Ressources génétiques, Un patrimoine à préserver. Vol. 2020, https://www.gnis-pedagogie.org/sujet/enjeu-des-ressources-genetiques/.
Goldstein, D. B., Ruiz Linares, A., Cavalli-Sforza, L. L., and Feldman, M. W. (1995a). An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**, 463-471.
Goldstein, D. B., Ruiz Linares, A., Cavalli-Sforza, L. L., and Feldman, M. W. (1995b). Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 6723-6727.
Goudet, J. (2005). hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes* **5**, 184-186.
Gouesnard, B., Dallard, J., Panouille, A., and Boyat, A. (1997). Classification of French maize populations based on morphological traits. *Agronomie* **17**, 491-498.
Hale, M. L., Borland, A. M., Gustafsson, M. H. G., and Wolff, K. (2004). Causes of size homoplasy among chloroplast microsatellites in closely related Clusia species. *Journal of Molecular Evolution* **58**, 182-190.
Hansen, A. K., Escobar, L. K., Gilbert, L. E., and Jansen, R. K. (2007). Paternal, maternal, and biparental inheritance of the chloroplast genome in Passiflora (Passifloraceae): Implications for phylogenetic studies. *American Journal of Botany* **94**, 42-46.
Hatziskakis, S., Papageorgiou, A. C., Gailing, O., and Finkeldey, R. (2009). High chloroplast haplotype diversity in Greek populations of beech (Fagus sylvatica L.). *Plant Biology* **11**, 425-433.
Hölker, A. C., Mayer, M., Presterl, T., Bolduan, T., Bauer, E., Ordas, B., Brauner, P. C., Ouzunova, M., Melchinger, A. E., and Schön, C. C. (2019). European maize landraces made accessible for plant breeding and genome-based studies. *Theoretical and Applied Genetics* **132**, 3333-3345.
INRA (2017). INRA, Maize database. http://xcrop.inra.fr/index.php.
Ishii, T., and McCouch, S. R. (2000). Microsatellites and microsynteny in the chloroplast genomes of Oryza and eight other Gramineae species. *Theoretical and Applied Genetics* **100**, 1257-1266.
Labate, J. A., Lamkey, K. R., Mitchell, S. E., Kresovich, S., Sullivan, H., and Smith, J. S. C. (2003). Molecular and historical aspects of corn belt dent diversity. *Crop Science* **43**, 80-91.
LaFramboise, T. (2009). Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Res* **37**, 4181-93.
Lopez-Caamal, A., Ferrufino-Acosta, L. F., Diaz-Maradiaga, R. F., Rodriguez-Delcid, D., Mussali-Galante, P., and Tovar-Sanchez, E. (2019). Species distribution modelling and cpSSR reveal population history of the Neotropical annual herb Tithonia rotundifolia (Asteraceae). *Plant Biology* **21**, 248-258.
Louette, D. (2000). Traditional management of seed and genetic diversity: what is a landrace? *In* "Genes From the Field: On-farm Conservation of Crop Diversity." (S. B. Brush, ed.), pp. 119-142. IPGRI and IDRC, New-York.
Louette, D., Charrier, A., and Berthaud, J. (1997). In situ conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community. *Economic Botany* **51**, 20-38.
Louette, D., and Smale, M. (1998). "Farmers' seed selection practices and maize variety characteristics in a traditionally-based Mexican community," CIMMYT.
Louette, D., and Smale, M. (2000). Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico. *Euphytica* **113**, 25-41.
Macgregor, S., Visscher, P. M., and Montgomery, G. (2006). Analysis of pooled DNA samples on high density arrays without prior knowledge of differential hybridization rates. *Nucleic Acids Res* **34**, e55.
Maier, R. M., Neckermann, K., Igloi, G. L., and Kössel, H. (1995). Complete Sequence of the Maize Chloroplast Genome: Gene Content, Hotspots of Divergence and Fine Tuning of Genetic Information by Transcript Editing. *Journal of Molecular Biology* **251**, 614-628.
Mariac, C., Scarcelli, N., Pouzadou, J., Barnaud, A., Billot, C., Faye, A., Kougbeadjo, A., Maillol, V., Martin, G., Sabot, F., Santoni, S., Vigouroux, Y., and Couvreur, T. L. P. (2014). Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for

population genetics and phylogeography studies. *Molecular Ecology Resources* **14**, 1103-1113.
Marshall, H. D., Newton, C., and Ritland, K. (2001). Sequence-repeat polymorphisms exhibit the signature of recombination in lodgepole pine chloroplast DNA. *Molecular Biology and Evolution* **18**, 2136-2138.
Matsuoka, Y., Vigouroux, Y., Goodman, M. M., Sanchez, G. J., Buckler, E., and Doebley, J. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci U S A* **99**, 6080-4.
Maxted, N., Ford-Lloyd, B. V., and Hawkes, J. G. (1997). Complementary conservation strategies. *In* "Plant Genetic Conservation: The in situ approach" (N. Maxted, B. V. Ford-Lloyd and J. G. Hawkes, eds.), pp. 15-39. Springer Netherlands, Dordrecht.
Mayer, M., Unterseer, S., Bauer, E., de Leon, N., Ordas, B., and Schon, C. C. (2017). Is there an optimum level of diversity in utilization of genetic resources? *Theor Appl Genet* **130**, 2283-2295.
Minn, Y., Gailing, O., and Finkeldey, R. (2016). Genetic diversity and structure of teak (Tectona grandis L. f.) and dahat (Tectona hamiltoniana Wall.) based on chloroplast microsatellites and Amplified Fragment Length Polymorphism markers. *Genetic Resources and Crop Evolution* **63**, 961-974.
Mir, C., Zerjal, T., Combes, V., Dumas, F., Madur, D., Bedoya, C., Dreisigacker, S., Franco, J., Grudloyma, P., Hao, P. X., Hearne, S., Jampatong, C., Laloe, D., Muthamia, Z., Nguyen, T., Prasanna, B. M., Taba, S., Xie, C. X., Yunus, M., Zhang, S., Warburton, M. L., and Charcosset, A. (2013). Out of America: tracing the genetic footprints of the global diffusion of maize. *Theoretical and Applied Genetics* **126**, 2671-2682.
Morris, A. B., and Shaw, J. (2018). Markers in time and space: A review of the last decade of plant phylogeographic approaches. *Mol Ecol* **27**, 2317-2333.
Navascues, M., and Emerson, B. C. (2005). Chloroplast microsatellites: measures of genetic diversity and the effect of homoplasy. *Molecular Ecology* **14**, 1333-1341.
Neale, D., Marshall, K., and Harry, D. (2011). Inheritance of chloroplast and mitochondrial DNA in incense-cedar (Calocedrus decurrens). *Canadian Journal of Forest Research* **21**, 717-720.
Neale, D., and Sederoff, R. (1989). Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **77**, 212-216.
Negri, V., Maxted, N., and Veteläinen, M. (2009). European landraces: on-farm conservation, management and use. (ECPGR, ed.), Vol. 15, pp. 358 p.
Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* **70**, 3321-3323.
Nei, M. (1987). "Molecular evolutionnary genetics," Columbia University Press/Ed., New York.
OCDE/FAO (2018). Perspectives agricoles de l'OCDE et de la FAO 2018-2027. pp. 125-144.
OCDE/FAO (2020). "Perspectives agricoles de l'OCDE et de la FAO 2020-2029."
Ohta, T., and Kimura, M. (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genet Res* **22**, 201-4.
ONU (1992). Conférence des Nations Unies sur l'environnement et le développement, du 3 au 14 juin 1992, Rio de Janeiro. Vol. 2020, New York.
Ordas, A., and Deron, A. M. (1988). A method to measure conicalness in maize. *Maydica* **33**, 261-267.
Orozco-Ramirez, Q., Ross-Ibarra, J., Santacruz-Varela, A., and Brush, S. (2016). Maize diversity associated with social origin and environmental variation in Southern Mexico. *Heredity* **116**, 477-484.
Ortiz, R., Taba, S., Tovar, V. H. C., Mezzalama, M., Xu, Y., Yan, J., and Crouch, J. H. (2010). Conserving and Enhancing Maize Genetic Resources as Global Public Goods–A Perspective from CIMMYT. *Crop Science* **50**, 13.
Ouborg, N. J., Piquot, Y., and Van Groenendael, J. M. (1999). Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology* **87**, 551-568.

Ozerov, M., Vasemägi, A., Wennevik, V., Niemelä, E., Prusov, S., Kent, M., and Vähä, J.-P. (2013). Cost-effective genome-wide estimation of allele frequencies from pooled DNA in Atlantic salmon (Salmo salarL.). *BMC Genomics* **14**, 12.

Pautasso, M., Aistara, G., Barnaud, A., Caillon, S., Clouvel, P., Coomes, O. T., Delêtre, M., Demeulenaere, E., Santis, P., Döring, T., Eloy, L., Emperaire, L., Garine, E., Goldringer, I., Jarvis, D., Joly, H. I., Leclerc, C., Louafi, S., Martin, P., Massol, F., McGuire, S., McKey, D., Padoch, C., Soler, C., Thomas, M., and Tramontini, S. (2012). Seed exchange networks for agrobiodiversity conservation. A review. *Agronomy for Sustainable Development* **33**, 151-175.

Pazouki, L., Shanjani, P. S., Fields, P. D., Martins, K., Suhhorutsenko, M., Viinalass, H., and Niinemets, U. (2016). Large within-population genetic diversity of the widespread conifer Pinus sylvestris at its soil fertility limit characterized by nuclear and chloroplast microsatellite markers. *European Journal of Forest Research* **135**, 161-177.

Peng, Y. H., and Chen, K. (2011). Phylogeographic Pattern of Populus cathayana in the Southeast of Qinghai-Tibetan Plateau of China Revealed by cpSSR Markers. *Silva Fennica* **45**, 583-594.

Perales, R. H., Brush, S. B., and Qualset, C. O. (2003). Landraces of maize in Central Mexico: an altitudinal transect. *Economic Botany* **57**, 7-20.

Phumichai, C., Phumichai, T., and Wongkaew, A. (2015). Novel Chloroplast Microsatellite (cpSSR) Markers for Genetic Diversity Assessment of Cultivated and Wild Hevea Rubber. *Plant Molecular Biology Reporter* **33**, 1486-1498.

Piperno, D. R., and Flannery, K. V. (2001). The earliest archaeological maize (Zea mays L.) from highland Mexico: New accelerator mass spectrometry dates and their implications. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2101-2103.

Portères, R. (1961). L'ethnobotanique : Place - Objet - Méthode - Philosophie. *Journal d'agriculture traditionnelle et de botanique appliquée*, 102-109.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., and Rafalski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* **2**, 225-238.

Powell, W., Morgante, M., Andre, C., McNicol, J. W., Machray, G. C., Doyle, J. J., Tingey, S. V., and Rafalski, J. A. (1995a). Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. *Current Biology* **5**, 1023-1029.

Powell, W., Morgante, M., McDevitt, R., Vendramin, G. G., and Rafalski, J. A. (1995b). Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. *Proceedings of the National Academy of Sciences* **92**, 7759-7763.

Pressoir, G., and Berthaud, J. (2004). Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity* **92**, 88-94.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**, 945-959.

Provan, J., Lawrence, P., Young, G., Wright, F., Bird, R., Paglia, G., Cattonaro, F., Morgante, M., and Powell, W. (1999a). Analysis of the genus Zea (Poaceae) using polymorphic chloroplast simple sequence repeats. *Plant Systematics and Evolution* **218**, 245-256.

Provan, J., Lawrence, P., Young, G., Wright, F., Bird, R., Paglia, G., Cattonaro, F., Morgante, M., and Powell, W. (1999b). Analysis of the genusZea (Poaceae) using polymorphic chloroplast simple sequence repeats. *Plant Systematics and Evolution* **218**, 245-256.

Provan, J., Powell, W., and Hollingsworth, P. M. (2001). Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution* **16**, 142-147.

Provan, J., Russell, J. R., Booth, A., and Powell, W. (1999c). Polymorphic chloroplast simple sequence repeat primers for systematic and population studies in the genus Hordeum. *Molecular Ecology* **8**, 505-511.

Provan, J., Soranzo, N., Wilson, N. J., Goldstein, D. B., and Powell, W. (1999d). A low mutation rate for chloroplast microsatellites. *Genetics* **153**, 943-947.

Ranere, A. J., Piperno, D. R., Holst, I., Dickau, R., and Iriarte, J. (2009). The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5014-5018.

Raymond, M., and Rousset, F. (1995). GENEPOP (version 1.2): population genetic software for exact tests and ecumenicism. *J Hered* **86**, 248–249.

Reboud, X., and Zeyl, C. (1994). Organelle inheritance in plants. *Heredity* **72**, 132-140.

Rebourg, C., Chastanet, M., Gouesnard, B., Welcker, C., Dubreuil, P., and Charcosset, A. (2003). Maize introduction into Europe: the history reviewed in the light of molecular data. *Theoretical and Applied Genetics* **106**, 895-903.

Rebourg, C., Gouesnard, B., and Charcosset, A. (2001). Large scale molecular analysis of traditional European maize populations. Relationships with morphological variation. *Heredity* **86**, 574-587.

Reich, D., Thangaraj, K., Patterson, N., Price, A. L., and Singh, L. (2009). Reconstructing Indian population history. *Nature* **461**, 489-94.

Reif, J. C., Hamrit, S., Heckenberger, M., Schipprack, W., Maurer, H. P., Bohn, M., and Melchinger, A. E. (2005). Genetic structure and diversity of European flint maize populations determined with SSR analyses of individuals and bulks. *Theoretical and Applied Genetics* **111**, 906-913.

Renoux, J. (2014). "Le maïs, une plante pour l'intensification écologique," France Agricole, Paris.

Reverter, A., Henshall, J. M., McCulloch, R., Sasazaki, S., Hawken, R., and Lehnert, S. A. (2014). Numerical analysis of intensity signals resulting from genotyping pooled DNA samples in beef cattle and broiler chicken1. *Journal of Animal Science* **92**, 1874-1885.

Revilla Temiño, P., Soengas Fernández, M. d. P., Cartea González, M. E., Malvar Pintos, R. A., and Ordás Pérez, A. (2003). Isozyme variability among European maize populations and the introduction of maize in Europe. *Maydica* **48**, 141-152.

Revilla Temiño, P., Soengas Fernández, M. d. P., Malvar Pintos, R. A., Cartea González, M. E., and Ordás Pérez, A. (1998). Isozyme variation and historical relationships among the maize races of Spain. *Maydica* **43**, 175-182.

Reyes-Valdes, M. H., Santacruz-Varela, A., Martinez, O., Simpson, J., Hayano-Kanashiro, C., and Cortes-Romero, C. (2013). Analysis and Optimization of Bulk DNA Sampling with Binary Scoring for Germplasm Characterization. *PLoS One* **8**.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-28.

Scarcelli, N., Mariac, C., Couvreur, T. L. P., Faye, A., Richard, D., Sabot, F., Berthouly-Salazar, C., and Vigouroux, Y. (2016). Intra-individual polymorphism in chloroplasts from NGS data: where does it come from and how to handle it? *Molecular Ecology Resources* **16**, 434-445.

Semerikova, S. A., Semerikov, V. L., and Lascoux, M. (2011). Post-glacial history and introgression in Abies (Pinaceae) species of the Russian Far East inferred from both nuclear and cytoplasmic markers. *Journal of Biogeography* **38**, 326-340.

Sham, P., Bader, J. S., Craig, I., O'Donovan, M., and Owen, M. (2002). DNA Pooling: a tool for large-scale association studies. *Nat Rev Genet* **3**, 862-71.

Smith, J. S. C., Chin, E. C. L., Shu, H., Smith, O. S., Wall, S. J., Senior, M. L., Mitchell, S. E., Kresovich, S., and Ziegle, J. (1997). An evaluation of the utility of SSR loci as molecular markers in maize (Zea mays L): comparisons with data from RFLPS and pedigree. *Theoretical and Applied Genetics* **95**, 163-173.

Takahashi, Y., Yokoi, S., and Takahata, Y. (2016). Genetic divergence of turnip (Brassica rapa L. em. Metzg. subsp. rapa) inferred from simple sequence repeats in chloroplast and nuclear genomes and morphology. *Genetic Resources and Crop Evolution* **63**, 869-879.

Tenaillon, M. I., and Charcosset, A. (2011). A European perspective on maize history. *Comptes Rendus Biologies* **334**, 221-228.

Thia, J. A., Hale, M. L., and Chapman, H. M. (2016). Interspecific comparisons with chloroplast SSR loci reveal limited genetic variation in Nigerian montane forests: A study on Cordia millenii (West

African Cordia), Entandrophragma angolense (tiama mahogany), and Lovoa trichilioides (African walnut). *Tropical Conservation Science* **9**, 321-337.

Unterseer, S., Pophaly, S. D., Peis, R., Westermeier, P., Mayer, M., and Seidel, M. A. (2016). A comprehensive study of the genomic differentiation between temperate dent and flint maize. *Genome Biol*.

van Heerwaarden, J., Doebley, J., Briggs, W. H., Glaubitz, J. C., Goodman, M. M., Gonzalez, J. D. S., and Ross-Ibarra, J. (2011). Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 1088-1092.

Vaz Patto, M. C., Moreira, P. M., Almeida, N., Satovic, Z., and Pego, S. (2008). Genetic diversity evolution through participatory maize breeding in Portugal. *Euphytica* **161**, 283-291.

Vigouroux, Y., Glaubitz, J. C., Matsuoka, Y., Goodman, M. M., Jesus Sanchez, G., and Doebley, J. (2008). Population Structure and Genetic Diversity of New World Maize Races Assessed by DNA Microsatellites. *American Journal of Botany* **95**, 1240-1253.

Vigouroux, Y., Jaqueth, J. S., Matsuoka, Y., Smith, O. S., Beavis, W. F., Smith, J. S. C., and Doebley, J. (2002). Rate and pattern of mutation at microsatellite loci in maize. *Molecular Biology and Evolution* **19**, 1251-1260.

Villa, T., Maxted, N., Scholten, M., and Ford-Lloyd, B. (2005). Defining and identifying crop landraces. *Plant Genetic Resources* **3**, 373-384.

Vitalis, R. (2001). Subdivided population genetics: theory and applications, Université Montpellier II - Sciences et Techniques du Languedoc.

Wang, M., Yan, J., Zhao, J., Song, W., Zhang, X., Xiao, Y., and Zheng, Y. (2012). Genome-wide association study (GWAS) of resistance to head smut in maize. *Plant Sci* **196**.

Wang, Y. H., Jiang, H., Peng, S. M., and Korpelainen, H. (2011). Genetic structure in fragmented populations of Hippophae rhamnoides ssp sinensis in China investigated by ISSR and cpSSR markers. *Plant Systematics and Evolution* **295**, 97-107.

Warburton, M. L., Wilkes, G., Taba, S., Charcosset, A., Mir, C., Dumas, F., Madur, D., Dreisigacker, S., Bedoya, C., Prasanna, B. M., Xie, C. X., Hearne, S., and Franco, J. (2011). Gene flow among different teosinte taxa and into the domesticated maize gene pool. *Genetic Resources and Crop Evolution* **58**, 1243-1261.

Westengen, O. T., Berg, P. R., Kent, M. P., and Brysting, A. K. (2012). Spatial structure and climatic adaptation in African maize revealed by surveying SNP diversity in relation to global breeding and landrace panels. *Plos One* **7**.

Wheeler, G. L., Dorman, H. E., Buchanan, A., Challagundla, L., and Wallace, L. E. (2014). A review of the prevalence, utility, and caveats of using chloroplast simple sequence repeats for studies of plant biology. *Applications in Plant Sciences* **2**, 1400059.

Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B., and Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology* **79**, 13-26.

Wright, S. (1978). "Evolution and the Genetics of Populations," Evolutionary Biology, University of Chicago Press, Chicago.

Yan, J., Shah, T., Warburton, M. L., Buckler, E. S., McMullen, M. D., and Crouch, J. (2009a). Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PLoS One* **4**.

Yan, J., Yang, X., Shah, T., Sánchez-Villeda, H., Li, J., Warburton, M., Zhou, Y., Crouch, J. H., and Xu, Y. (2009b). High-throughput SNP genotyping with the GoldenGate assay in maize. *Mol Breed* **25**.

Yang, J., Li, S., Sun, G., Yuan, Y., and Zhao, G. (2008). Population structure and genetic variation in the genus Dipteronia Oliv. (Aceraceae) endemic to China as revealed by cpSSR analysis. *Plant Systematics and Evolution* **272**, 97-106.

Yu, H. U. I., Nason, J. D., Ge, X., and Zeng, J. (2010). Slatkin's Paradox: when direct observation and realized gene flow disagree. A case study in Ficus. *Molecular Ecology* **19**, 4441-4453.

Zeven, A. C. (1998). Landraces : a review of definitions and classifications. *Euphytica* **104**, 127-139.

ANNEXES

Annexe 1 : tableaux supplémentaires du chapitre 1.

S1 Table. Combination of passport data (column 1 to 5), genetic diversity analysis (column 6 and 7) and admixture analysis at K=2 (column 8 to 10) performing on SNP data of the 194 maize landraces from the South-West of France.

W-SWF= "West South-West France » ; E-SWF =" East South-West France".

CodeColl	Commonname	District	LatSiteColl	LongSiteColl	genetic diversity		admixture analysis at K=2		
					Allelic richness	Expected heteorzygotity (He)	E-SWF	W-SWF	cluster
PPS597	Seix.Bardot	Ariège	42.81666667	1.15	1.657	0.205	0.74	0.26	E-SWF
PPS1016	Couserans	Ariège	42.91666667	1.05	1.604	0.195	0.73	0.27	E-SWF
PPS737	Couflens.1	Ariège	42.75	1.183333333	1.672	0.211	0.72	0.28	E-SWF
PPS237	Seix	Ariège	42.81666667	1.15	1.665	0.203	0.71	0.29	E-SWF
PPS69	Couflens.2	Ariège	42.75	1.183333333	1.679	0.212	0.70	0.30	E-SWF
PPS13	Aleu	Ariège	42.88333333	1.266666667	1.655	0.200	0.66	0.34	E-SWF
PPS14	Massat	Ariège	42.8666667	1.3833333	1.722	0.202	0.66	0.34	E-SWF
PPS191	Aston D	Ariège	42.66666667	1.65	1.707	0.219	0.65	0.35	E-SWF
PPS71	Aston B	Ariège	42.66666667	1.65	1.840	0.263	0.63	0.37	E-SWF
PPS966	Ariege 3	Ariège	NA	NA	1.551	0.163	0.60	0.40	E-SWF
PPS478	Labastide de Boussignac	Ariège	43.05	1.883333333	1.825	0.257	0.59	0.41	E-SWF
PPS68	Riverenert 1	Ariège	42.95	1.25	1.640	0.201	0.59	0.41	E-SWF
PPS70	Aston A	Ariège	42.66666667	1.65	1.682	0.213	0.58	0.42	E-SWF
PPS736	Ariege 1	Ariège	42.76666667	1.5	1.672	0.203	0.51	0.49	E-SWF
PPS668	Millette Montagne Noire	Aude	43.3833333	2.2666667	1.743	0.215	0.58	0.42	E-SWF
PPS614	Espeyrac	Aveyron	44.61666667	2.516666667	1.701	0.200	0.91	0.09	E-SWF
PPS121	Beaumont	Dordogne	44.76666667	0.75	1.610	0.181	0.84	0.16	E-SWF
PPS621	Pop Gers	Gers	NA	NA	1.762	0.239	0.60	0.40	E-SWF
PPS626	Cugnieres	Gironde	44.41666667	-0.533333333	1.818	0.258	0.56	0.44	E-SWF
PPS487	Cazalis	Gironde	44.33333333	-0.383333333	1.835	0.261	0.53	0.47	E-SWF
PPS230	Sainte Livrade	Haute Garonne	43.65	1.1	1.491	0.152	0.73	0.27	E-SWF
PPS568	Moustajon	Haute Garonne	42.8	0.583333333	1.765	0.245	0.62	0.38	E-SWF
PPS1039	Sarrancolin	Hautes Pyrénées	42.95	0.4	1.409	0.133	0.66	0.34	E-SWF
PPS1036	Saint Creac.A	Hautes Pyrénées	43.05	-0.016666667	1.649	0.185	0.58	0.42	E-SWF
PPS624	Landes 3	Landes	NA	NA	1.677	0.200	0.86	0.14	E-SWF
PPS625	Landes 4	Landes	NA	NA	1.650	0.196	0.85	0.15	E-SWF
PPS1250	Le Muret.1	Landes	44.833	-0.825	1.568	0.184	0.75	0.25	E-SWF
PPS1251	Le Muret.2	Landes	44.833	-0.825	1.855	0.266	0.61	0.39	E-SWF
PPS1248	Pop Cornée.1	Landes	43.9	-0.483333333	1.731	0.227	0.54	0.46	E-SWF
PPS1249	Pop Cornée.2	Landes	43.9	-0.483333333	1.746	0.232	0.51	0.49	E-SWF
PPS87	Cressensac 1	Lot	45.01666667	1.516666667	1.650	0.177	1.00	0.00	E-SWF
PPS631	Saint Pantaleon	Lot	44.3666667	1.2833333	1.687	0.197	0.96	0.04	E-SWF
PPS632	Cressensac 2	Lot	45.01666667	1.516666667	1.634	0.193	0.96	0.04	E-SWF
PPS85	Floirac Camboni A	Lot	44.9	1.666666667	1.591	0.173	0.96	0.04	E-SWF
PPS86	Floirac	Lot	44.9	1.666666667	1.644	0.191	0.59	0.41	E-SWF
PPS229	Marmande.34	Lot et Garonne	NA	NA	1.637	0.200	0.89	0.11	E-SWF
PPS638	Marmande.32	Lot et Garonne	NA	NA	1.692	0.208	0.89	0.11	E-SWF
PPS227	Marmande.21	Lot et Garonne	NA	NA	1.772	0.231	0.88	0.12	E-SWF
PPS226	Marmande.10	Lot et Garonne	NA	NA	1.336	0.118	0.87	0.13	E-SWF
PPS228	Marmande.26	Lot et Garonne	NA	NA	1.740	0.214	0.86	0.14	E-SWF
PPS633	Marmande.12	Lot et Garonne	NA	NA	1.732	0.208	0.84	0.16	E-SWF
PPS362	Marmande.35	Lot et Garonne	NA	NA	1.701	0.223	0.77	0.23	E-SWF
PPS96	Marmande.31	Lot et Garonne	NA	NA	1.670	0.201	0.72	0.28	E-SWF
PPS198	Marmande.36	Lot et Garonne	NA	NA	1.695	0.212	0.68	0.32	E-SWF
PPS634	Marmande.45	Lot et Garonne	NA	NA	1.674	0.200	0.58	0.42	E-SWF
PPS637	Marmande.20	Lot et Garonne	NA	NA	1.443	0.145	0.54	0.46	E-SWF
PPS636	Marmande.19	Lot et Garonne	NA	NA	1.497	0.153	0.52	0.48	E-SWF
PPS609	Grand Roux Basque 3	Pyrenees Atlantiques	NA	NA	1.576	0.175	0.91	0.09	E-SWF
PPS651	Auga 2	Pyrenees Atlantiques	43.46666667	-0.366666667	1.741	0.216	0.65	0.35	E-SWF
PPS574	Fraysse Alban	Tarn	43.91666667	2.4	1.765	0.219	0.97	0.03	E-SWF

CodeColl	Commonname	District	LatSiteColl	LongSiteColl	genetic diversity		admixture analysis at K=2		
					Allelic richness	Expected heteorzygotity (He)	E-SWF	W-SWF	cluster
PPS575	Cabrilles	Tarn	43.7	2.133333333	1.634	0.187	0.97	0.03	E-SWF
PPS666	Villefranche d Albigeois.A	Tarn	43.88333333	2.316666667	1.684	0.199	0.97	0.03	E-SWF
PPS61	Vaour	Tarn	44.06666667	1.8	1.765	0.228	0.90	0.10	E-SWF
PPS505	Pampelone	Tarn	44.1	2.233333333	1.659	0.201	0.82	0.18	E-SWF
PPS1252	Isle sur Tarn	Tarn	43.88333333	1.783333333	1.670	0.214	0.79	0.21	E-SWF
PPS15	Lacaune	Tarn	43.55	2.5833333	1.650	0.182	0.77	0.23	E-SWF
PPS503	Semport	Tarn	NA	NA	1.729	0.217	0.67	0.33	E-SWF
PPS501	Borie Lautrec	Tarn	43.7	2.133333333	1.724	0.219	0.66	0.34	E-SWF
PPS60	Poulinet Alban.B	Tarn	43.83333333	2.416666667	1.726	0.220	0.66	0.34	E-SWF
PPS504	Villefranche d Albigeois.B	Tarn	43.88333333	2.316666667	1.675	0.210	0.64	0.36	E-SWF
PPS506	Poulinet Alban.A	Tarn	43.83333333	2.416666667	1.713	0.227	0.61	0.39	E-SWF
PPS59	Montirat Monesties	Tarn	44.06666667	2.1	1.755	0.226	0.58	0.42	E-SWF
PPS502	Lautrec	Tarn	43.7	2.133333333	1.702	0.211	0.56	0.44	E-SWF
PPS189	Saint Lieux les Lavaur	Tarn	43.75	1.75	1.729	0.225	0.52	0.48	E-SWF
PPS241	Millette de Finham.AGPM	Tarn et Garonne	43.91666667	1.216666667	1.769	0.222	0.89	0.11	E-SWF
PPS196	Moulin Neuf	Ariège	43.06666667	1.933333333	1.722	0.225	0.46	0.54	W-SWF
PPS238	Riverenert 2	Ariège	42.95	1.25	1.673	0.206	0.40	0.60	W-SWF
PPS26	Argein	Ariège	42.91666667	0.983333333	1.692	0.216	0.39	0.61	W-SWF
PPS67	Prat Bonrepaux	Ariège	43.03333333	1.016666667	1.709	0.225	0.33	0.67	W-SWF
PPS66	Cazavet	Ariège	42.98333333	1.033333333	1.740	0.232	0.31	0.69	W-SWF
PPS65	Sainte Croix	Ariège	43.11666667	1.166666667	1.700	0.219	0.28	0.72	W-SWF
PPS639	Millette du Lauragais 2	Aude	43.3833333	1.7166667	1.712	0.223	0.49	0.51	W-SWF
PPS122	Millette du Lauragais 1	Aude	43.383	1.717	1.741	0.228	0.49	0.51	W-SWF
PPS613	Millette La Cassignole	Aude	NA	NA	1.698	0.222	0.47	0.53	W-SWF
PPS615	Saint Come	Aveyron	44.53333333	2.816666667	1.749	0.237	0.46	0.54	W-SWF
PPS616	Saint Parthem.A	Aveyron	44.61666667	2.283333333	1.792	0.243	0.45	0.55	W-SWF
PPS617	Saint Parthem.B	Aveyron	44.61666667	2.283333333	1.638	0.197	0.31	0.69	W-SWF
PPS27	Saint Christophe	Aveyron	44.46666667	2.383333333	1.715	0.217	0.31	0.69	W-SWF
PPS570	Moncassin	Gers	43.45	0.466666667	1.676	0.214	0.13	0.87	W-SWF
PPS483	Gironde 5	Gironde	NA	NA	1.816	0.243	0.45	0.55	W-SWF
PPS410	Gironde 2	Gironde	NA	NA	1.744	0.219	0.42	0.58	W-SWF
PPS486	Gironde 8	Gironde	NA	NA	1.683	0.202	0.21	0.79	W-SWF
PPS442	Gironde 3	Gironde	NA	NA	1.593	0.183	0.17	0.83	W-SWF
PPS485	Gironde 7	Gironde	NA	NA	1.720	0.205	0.17	0.83	W-SWF
PPS35	Gironde 1	Gironde	NA	NA	1.637	0.191	0.14	0.86	W-SWF
PPS460	Gironde 4	Gironde	NA	NA	1.650	0.183	0.13	0.87	W-SWF
PPS484	Gironde 6	Gironde	NA	NA	1.653	0.190	0.09	0.91	W-SWF
PPS585	Gironde 9	Gironde	NA	NA	1.591	0.172	0.04	0.96	W-SWF
PPS963	Val d'Aran	Haute Garonne	42.85	0.733333333	1.707	0.219	0.50	0.50	W-SWF
PPS193	Aurignac	Haute Garonne	43.2	0.866666667	1.699	0.219	0.48	0.52	W-SWF
PPS567	Boutx B	Haute Garonne	42.91666667	0.766666667	1.614	0.194	0.40	0.60	W-SWF
PPS566	Boutx A	Haute Garonne	42.91666667	0.766666667	1.668	0.212	0.35	0.65	W-SWF
PPS12	Portet	Haute Garonne	42.8133	0.4861	1.644	0.206	0.33	0.67	W-SWF
PPS84	Sauveterre de Comminges 1	Haute Garonne	43.03333333	0.666666667	1.705	0.219	0.29	0.71	W-SWF
PPS482	Cuguron	Haute Garonne	43.1	0.516666667	1.733	0.226	0.28	0.72	W-SWF
PPS361	Cassagnabere 1	Haute Garonne	43.21666667	0.8	1.717	0.225	0.27	0.73	W-SWF
PPS1421	Pouyastruc	Haute Pyrénées	43.267	0.167	1.709	0.221	0.10	0.90	W-SWF
PPS8	Salechan	Hautes Pyrénées	42.95	0.633333333	1.759	0.238	0.29	0.71	W-SWF
PPS11	Sost	Hautes Pyrénées	42.8833333	0.55	1.739	0.214	0.27	0.73	W-SWF
PPS601	Nestier 2	Hautes Pyrénées	43.05	0.466666667	1.726	0.230	0.27	0.73	W-SWF
PPS820	Blanche Htes.Pyr.C1	Hautes Pyrénées	NA	NA	1.556	0.173	0.25	0.75	W-SWF
PPS83	Guchan	Hautes Pyrénées	42.83333333	0.333333333	1.698	0.218	0.22	0.78	W-SWF
PPS25	Bize	Hautes Pyrénées	43.03333333	0.466666667	1.687	0.212	0.22	0.78	W-SWF
PPS9	Bareilles	Hautes Pyrénées	42.88333333	0.433333333	1.691	0.217	0.20	0.80	W-SWF
PPS491	Saint Laurent de Neste 3	Hautes Pyrénées	43.08333333	0.466666667	1.722	0.229	0.18	0.82	W-SWF
PPS76	Nestier 1	Hautes Pyrénées	43.05	0.466666667	1.726	0.227	0.18	0.82	W-SWF
PPS1153	Col du Soulor	Hautes Pyrénées	42.9	-0.25	1.775	0.240	0.17	0.83	W-SWF
PPS1037	Saint Creac.B	Hautes Pyrénées	43.05	-0.016666667	1.741	0.229	0.17	0.83	W-SWF
PPS490	Bugard	Hautes Pyrénées	43.25	0.316666667	1.693	0.214	0.15	0.85	W-SWF

CodeColl	Commonname	District	LatSiteColl	LongSiteColl	genetic diversity		admixture analysis at K=2		
					Allelic richness	Expected heteorzygotity (He)	E-SWF	W-SWF	cluster
PPS77	Saint Laurent de Neste 1	Hautes Pyrénées	43.08333333	0.466666667	1.748	0.233	0.15	0.85	W-SWF
PPS10	Estarvielle	Hautes Pyrénées	42.8166667	0.4166667	1.744	0.218	0.15	0.85	W-SWF
PPS82	Bazus Aure	Hautes Pyrénées	42.85	0.35	1.669	0.210	0.14	0.86	W-SWF
PPS78	Saint Laurent de Neste 2	Hautes Pyrénées	43.08333333	0.466666667	1.727	0.228	0.13	0.87	W-SWF
PPS604	Lannemezan	Hautes Pyrénées	43.11666667	0.383333333	1.747	0.236	0.13	0.87	W-SWF
PPS1041	Gazost	Hautes Pyrénées	43	0.016666667	1.812	0.250	0.13	0.87	W-SWF
PPS72	Vidou 1	Hautes Pyrénées	43.3	0.333333333	1.668	0.206	0.13	0.87	W-SWF
PPS602	Saint Laurent de Neste 4	Hautes Pyrénées	43.08333333	0.466666667	1.719	0.221	0.11	0.89	W-SWF
PPS1045	Bagnères de Bigorre 1	Hautes Pyrénées	43.05	0.166666667	1.755	0.231	0.11	0.89	W-SWF
PPS20	Campan Galade	Hautes Pyrénées	42.95	0.216666667	1.671	0.210	0.09	0.91	W-SWF
PPS1043	Argeles Gazost	Hautes Pyrénées	42.9	-0.25	1.729	0.227	0.08	0.92	W-SWF
PPS1038	Saint Creac.C	Hautes Pyrénées	43.05	-0.016666667	1.783	0.240	0.08	0.92	W-SWF
PPS1052	Bagnères de Bigorre 2	Hautes Pyrénées	43.05	0.166666667	1.717	0.226	0.07	0.93	W-SWF
PPS21	Sainte Marie de Campan	Hautes Pyrénées	42.96666667	0.25	1.678	0.216	0.07	0.93	W-SWF
PPS1051	Lourdes	Hautes Pyrénées	43.08333333	0.033333333	1.770	0.227	0.06	0.94	W-SWF
PPS600	Serre Rustaing	Hautes Pyrénées	43.25	0.283333333	1.711	0.221	0.05	0.95	W-SWF
PPS964	Argeles	Hautes Pyrénées	43.08333333	0.2	1.606	0.180	0.05	0.95	W-SWF
PPS240	Saint Sauveur	Hautes Pyrénées	42.81666667	0.016666667	1.706	0.218	0.05	0.95	W-SWF
PPS598	Bugard A	Hautes Pyrénées	43.25	0.316666667	1.679	0.212	0.05	0.95	W-SWF
PPS22	Esterre	Hautes Pyrénées	42.8666667	0	1.716	0.205	0.04	0.96	W-SWF
PPS1044	Pierrefitte	Hautes Pyrénées	42.93333333	-0.016666667	1.758	0.230	0.04	0.96	W-SWF
PPS1050	Tarbes	Hautes Pyrénées	43.25	0	1.672	0.210	0.03	0.97	W-SWF
PPS73	Begole	Hautes Pyrénées	43.15	0.316666667	1.704	0.220	0.03	0.97	W-SWF
PPS24	Viey	Hautes Pyrénées	42.88333333	0.016666667	1.682	0.213	0.03	0.97	W-SWF
PPS81	Luz	Hautes Pyrénées	42.81666667	0.016666667	1.679	0.210	0.03	0.97	W-SWF
PPS1046	Arrens	Hautes Pyrénées	42.9	-0.25	1.731	0.228	0.03	0.97	W-SWF
PPS1053	Ourdes	Hautes Pyrénées	42.96666667	0.533333333	1.740	0.228	0.03	0.97	W-SWF
PPS599	Burg	Hautes Pyrénées	43.18333333	0.316666667	1.717	0.226	0.02	0.98	W-SWF
PPS79	Aspin ez Angles	Hautes Pyrénées	43.08333333	0.016666667	1.692	0.218	0.01	0.99	W-SWF
PPS80	Esquieze Serre	Hautes Pyrénées	42.86666667	0	1.668	0.207	0.01	0.99	W-SWF
PPS603	Grust	Hautes Pyrénées	42.88333333	-0.05	1.674	0.209	0.01	0.99	W-SWF
PPS627	Pessan	Landes	44.21666667	-0.516666667	1.702	0.225	0.48	0.52	W-SWF
PPS199	Parleboscq	Landes	43.93333333	0.033333333	1.707	0.218	0.44	0.56	W-SWF
PPS363	Landes 2	Landes	43.883	-1.05	1.761	0.236	0.39	0.61	W-SWF
PPS628	Dubos	Landes	44.21666667	-0.516666667	1.823	0.242	0.38	0.62	W-SWF
PPS623	CETA des Landes	Landes	NA	NA	1.765	0.238	0.34	0.66	W-SWF
PPS630	Quarantaun	Landes	43.53333333	-1.45	1.778	0.240	0.20	0.80	W-SWF
PPS622	Roux de Chalosse	Landes	43.61666667	-0.833333333	1.753	0.237	0.12	0.88	W-SWF
PPS98	Landes 1	Landes	43.717	-1.05	1.724	0.222	0.08	0.92	W-SWF
PPS243	Floirac Camboni B	Lot	44.9	1.666666667	1.720	0.218	0.50	0.50	W-SWF
PPS194	Grand Roux Basque 1	Pyrenees Atlantiques	NA	NA	1.819	0.260	0.49	0.51	W-SWF
PPS1236	Coarraze	Pyrenees Atlantiques	43.16666667	-0.2	1.856	0.270	0.45	0.55	W-SWF
PPS645	Bordes	Pyrenees Atlantiques	43.23333333	-0.266666667	1.788	0.243	0.33	0.67	W-SWF
PPS641	Grand Roux Basque 5	Pyrenees Atlantiques	43.4666667	-1.4666667	1.771	0.246	0.30	0.70	W-SWF
PPS649	Gomer	Pyrenees Atlantiques	43.23333333	-0.183333333	1.687	0.217	0.24	0.76	W-SWF
PPS818	Lucq de Bearn	Pyrenees Atlantiques	43.3	-0.633333333	1.814	0.252	0.14	0.86	W-SWF
PPS1040	Monein 2	Pyrenees Atlantiques	43.28333333	-0.55	1.798	0.250	0.13	0.87	W-SWF
PPS1057	Lestelle Betharram	Pyrenees Atlantiques	43.11666667	-0.2	1.655	0.182	0.13	0.87	W-SWF
PPS822	Monein C1	Pyrenees Atlantiques	43.28333333	-0.55	1.789	0.241	0.08	0.92	W-SWF
PPS969	Monein 1	Pyrenees Atlantiques	43.28333333	-0.55	1.684	0.211	0.07	0.93	W-SWF

CodeColl	Commonname	District	LatSiteColl	LongSiteColl	genetic diversity		admixture analysis at K=2		
					Allelic richness	Expected heteorzygotity (He)	E-SWF	W-SWF	cluster
PPS1042	Morlaas	Pyrenees Atlantiques	43.33333333	-0.25	1.774	0.238	0.06	0.94	W-SWF
PPS1048	Orthez d'Asson.A	Pyrenees Atlantiques	43.08333333	-0.25	1.783	0.239	0.06	0.94	W-SWF
PPS642	Roux du Pays	Pyrenees Atlantiques	NA	NA	1.753	0.238	0.06	0.94	W-SWF
PPS644	Sainte Suzanne	Pyrenees Atlantiques	NA	NA	1.671	0.211	0.06	0.94	W-SWF
PPS647	Lucgarier 1	Pyrenees Atlantiques	43.21666667	-0.183333333	1.783	0.244	0.06	0.94	W-SWF
PPS648	Lucgarier 2	Pyrenees Atlantiques	43.21666667	-0.183333333	1.777	0.241	0.05	0.95	W-SWF
PPS1031	Arrien	Pyrenees Atlantiques	43.3	-0.15	1.687	0.212	0.05	0.95	W-SWF
PPS2	Saint Jean Pied de Port 1	Pyrenees Atlantiques	43.15	-1.216666667	1.733	0.224	0.05	0.95	W-SWF
PPS607	Doré de Bordes	Pyrenees Atlantiques	43.23333333	-0.266666667	1.708	0.218	0.05	0.95	W-SWF
PPS646	Barzin	Pyrenees Atlantiques	43.21666667	-0.116666667	1.735	0.220	0.04	0.96	W-SWF
PPS652	Mouhoux	Pyrenees Atlantiques	43.48333333	-0.233333333	1.739	0.231	0.04	0.96	W-SWF
PPS643	Salies du Bearn 2	Pyrenees Atlantiques	43.46666667	-0.9	1.702	0.219	0.03	0.97	W-SWF
PPS962	Asson	Pyrenees Atlantiques	43.08333333	-0.233333333	1.698	0.212	0.03	0.97	W-SWF
PPS650	Auga 1	Pyrenees Atlantiques	43.46666667	-0.366666667	1.726	0.226	0.03	0.97	W-SWF
PPS1049	Orthez d'Asson.B	Pyrenees Atlantiques	43.08333333	-0.25	1.714	0.217	0.03	0.97	W-SWF
PPS967	Pontacq	Pyrenees Atlantiques	43.2	-0.1	1.636	0.191	0.03	0.97	W-SWF
PPS23	Laruns	Pyrenees Atlantiques	42.8833333	-0.4166667	1.699	0.217	0.02	0.98	W-SWF
PPS1	Salies de Bearn 1	Pyrenees Atlantiques	43.46666667	-0.9	1.709	0.225	0.02	0.98	W-SWF
PPS608	Doré du Pays	Pyrenees Atlantiques	NA	NA	1.743	0.232	0.02	0.98	W-SWF
PPS653	Saint Jean Pied de Port 3	Pyrenees Atlantiques	43.15	-1.216666667	1.719	0.224	0.01	0.99	W-SWF
PPS1047	Saint Engrace	Pyrenees Atlantiques	42.98333333	-0.816666667	1.730	0.231	0.01	0.99	W-SWF
PPS235	Grand Roux Basque 2	Pyrenees Atlantiques	NA	NA	1.723	0.227	0.01	0.99	W-SWF
PPS821	Val du Gave d'Aspe	Pyrenees Atlantiques	43.05	0.25	1.719	0.223	0.01	0.99	W-SWF
PPS610	Grand Roux Basque 4	Pyrenees Atlantiques	NA	NA	1.715	0.225	0.01	0.99	W-SWF
PPS965	Barzun	Pyrenees Atlantiques	43.21666667	-0.116666667	1.725	0.224	0.01	0.99	W-SWF
PPS1054	Monein 3	Pyrenees Atlantiques	43.28333333	-0.55	1.733	0.233	0.01	0.99	W-SWF
PPS968	Urdos	Pyrenees Atlantiques	42.83333333	-0.516666667	1.689	0.205	0.01	0.99	W-SWF
PPS498	Estaut 2	Pyrenees Atlantiques	42.9	-0.516666667	1.665	0.211	0.01	0.99	W-SWF
PPS823	Lourdios.C1	Pyrenees Atlantiques	43.03333333	-0.65	1.733	0.224	0.01	0.99	W-SWF
PPS611	Saint Jean Pied de Port 2	Pyrenees Atlantiques	43.15	-1.216666667	1.719	0.228	0.00	1.00	W-SWF
PPS640	Doré du Pays Louvrigny	Pyrenees Atlantiques	43.5	-0.433333333	1.730	0.236	0.00	1.00	W-SWF
PPS488	Estaut 1	Pyrenees Atlantiques	42.9	-0.516666667	1.669	0.209	0.00	1.00	W-SWF
PPS1058	Gurmencon	Pyrenees Atlantiques	43.15	-0.6	1.730	0.226	0.00	1.00	W-SWF
PPS1154	Oloron Sainte Marie	Pyrenees Atlantiques	43.15	-0.583333333	1.715	0.222	0.00	1.00	W-SWF
PPS192	Cambon Le Lavaur	Tarn	43.58333333	1.85	1.695	0.217	0.43	0.57	W-SWF

S2 Table. Passport information and PCA analysis result on SNP data of 342 maize landraces.

CAM=Central America and Mexico ; SAM = South America ; NAM = North America ; SW France= South_West of France

CodeColl	Commonname	LatSiteColl	LongSiteColl	ZoneOri	PaysOri	cluster.Name	membership	Axis1	Axis2	Axis3	Axis4	Id_country initial
ANTI392	Amagaceno	6.3333333	-75.5	CAM	Colombia	Andes	0.44	20.21	2.72	-2.46	2.81	COL
ECUA617	Tusilla	-4.0833333	-78.55	CAM	Ecuador	Andes	0.59	20.21	3.80	-6.97	8.50	ECU
ECUA696	Canguil	-0.2166667	-78.25	CAM	Ecuador	Andes	0.65	2.96	-2.70	-17.32	-1.92	ECU
PUEB70	Conico	18.8666667	-97.3666667	CAM	Mexico	Andes	0.52	5.26	-7.49	-16.93	-14.24	MEX
PANA168	Salvadoreno	8.1333333	-81.4166667	CAM	Panama	Andes	0.37	21.50	-1.16	1.35	1.49	PAN
ARGE564	Cateto Amarillo	-26.0333333	-65.5	SAM	Argentina	Andes	0.76	15.31	8.35	-19.22	17.03	ARG
BOLI711	Perola	-21.3	-64.4333333	SAM	Bolivia	Andes	0.80	20.10	1.79	-7.53	12.55	BOL
BOZM0214	Uchuquilla	-17.35	-66.0166667	SAM	Bolivia	Andes	0.93	16.33	10.65	-23.33	18.02	BOL
PPS949	Conaripe Choshuenco	-39.5666667	-72.0166667	SAM	Chile	Andes	0.47	0.78	4.61	-13.79	12.66	CHI
APUC140	Confite_Puneno	-13.35	-72.55	SAM	Peru	Andes	0.90	1.14	-1.79	-14.45	4.43	PER
APUC171	San_Geronimo	-13.35	-73.2166667	SAM	Peru	Andes	0.72	2.31	-0.76	-17.74	3.45	PER
PERU1303	Confite_Morocho	-12.45	-75.6833333	SAM	Peru	Andes	0.94	17.19	13.21	-25.30	20.07	PER
ANTI3	-	17.0833333	-61.7166667	Caribbean	Antigua	Caribbean	0.95	23.80	-8.19	9.13	0.74	CAR
BARB5	Coastal_Tropical_Flint	13	-59.0333333	Caribbean	Barbados	Caribbean	0.77	23.36	-5.92	7.31	-0.78	CAR
BRVI142	Early_Caribbean	18.5	-65.5	Caribbean	British Virgin Islands	Caribbean	0.94	24.21	-7.83	9.79	1.50	CAR
CUBA44	Cuban_Flint	20.3833333	-76.4333333	Caribbean	Cuba	Caribbean	0.76	22.74	-6.34	4.82	0.85	CAR
CUBA54	Chandelle	20.9	-76.25	Caribbean	Cuba	Caribbean	0.75	23.29	-7.67	6.13	-0.15	CAR
CUBA63	Cuban_Flint	20.6666667	-75.6666667	Caribbean	Cuba	Caribbean	0.56	23.08	-6.90	5.48	0.10	CAR
RDOM270	Chandelle	19.8	-70.6833333	Caribbean	Dominican Republic	Caribbean	0.93	23.25	-8.35	7.20	0.42	CAR
DES103	Early_Caribbean	16.3333333	-61.05	Caribbean	Guadeloupe	Caribbean	0.93	23.17	-7.69	9.08	1.17	CAR
MGA235	Early_Caribbean	16.3333333	-63.05	Caribbean	Guadeloupe	Caribbean	0.68	24.18	-8.47	8.66	0.87	CAR
STB073	Chandelle	16.3333333	-63.3833333	Caribbean	Guadeloupe	Caribbean	0.86	23.66	-8.68	7.83	1.35	CAR
MART4	Early_Caribbean	14.75	-61.1833333	Caribbean	Martinique	Caribbean	0.92	24.49	-7.89	9.70	0.91	CAR
SVIN5	Coastal_Tropical_Flint	13.0333333	-61.0166667	Caribbean	Saint Vincent	Caribbean	0.90	23.14	-7.38	6.55	0.25	CAR
TRIN6	Tuson	10.5	-61.1666667	Caribbean	Trinite	Caribbean	0.52	22.44	-7.94	6.92	1.02	CAR
CRIC166	Salvadoreno	9.9833333	-84.0666667	CAM	Costa Rica	Caribbean	0.79	21.96	-3.47	3.28	0.59	COS
GUAT36	San_Marceno	15.4166667	-91.1	CAM	Guatemala	Caribbean	0.57	18.93	-7.34	5.41	1.94	GUA
GUAT527	Quicheno_early	15.1666667	-91.25	CAM	Guatemala	Caribbean	0.46	18.20	-3.59	-2.04	-3.72	GUA
CHIS104	Zapalote_Grande	15.3333333	-92.6666667	CAM	Mexico	Caribbean	0.62	13.82	-4.84	-2.72	-8.76	MEX
CHIS63	Tepecintle	16.7666667	-93.3666667	CAM	Mexico	Caribbean	0.42	18.63	-5.30	0.64	-4.65	MEX
VEN405	Tuson	9.5	-66.3	CAM	Venezuela	Caribbean	0.55	22.55	-4.79	3.86	1.03	VEN
IPK242	Mecklenburger	47.85	12.5833333	Europe	Germany	Caribbean	0.36	-26.48	2.76	2.50	0.36	GER
EPZMV156	Amarillo_de_Utrera	37.1833333	-5.7833333	Europe	Spain	Caribbean	0.44	11.67	-3.27	7.60	-0.47	SPA
EPZMV333	Conil	36.267	-6.083	Europe	Spain	Caribbean	0.80	18.78	0.79	6.61	-3.12	SPA
EPZMV336	Vejer	36.25	-5.883	Europe	Spain	Caribbean	0.81	18.93	1.46	7.03	-4.50	SPA
EPZMV337	Villamartin	38.7	-5.9833333	Europe	Spain	Caribbean	0.41	18.48	0.55	7.68	-4.37	SPA
URUG8A	Cateto_Sulino	-34.15	-56.45	SAM	Uruguay	Caribbean	0.49	10.71	-3.80	3.58	2.23	URU

CodeColl	Commonname	LatSiteColl	LongSiteColl	ZoneOri	PaysOri	cluster.Name	membership	Axis1	Axis2	Axis3	Axis4	Id_country initial
PPS525	Manolovo	42.7	24.6666667	Europe	Bulgaria	Corn_Belt_Dent	0.80	-5.78	3.74	6.45	2.82	BUL
RYS2	Slovenska Zlta	50.0833333	17.0833333	Europe	Czechoslovakia	Corn_Belt_Dent	0.93	-7.94	-10.04	6.35	5.26	CZE
PPS140	Foldeaki_Corne	47.5166667	19.0333333	Europe	Hungary	Corn_Belt_Dent	0.48	-1.13	-10.48	8.65	6.79	HUN
EPZMV155	Amarillo de Maranon	42.6166667	-1.75	Europe	Spain	Corn_Belt_Dent	0.39	-13.34	7.67	2.25	-1.53	SPA
PI214189	Falconer	49.8666667	-97.1666667	NAM	Canada	Corn_Belt_Dent	0.39	-16.70	-8.52	-1.48	-1.27	CAN
PI213699	Krug_Yellow_Dent	42.4	-93.0666667	NAM	Usa	Corn_Belt_Dent	0.73	5.70	-11.77	7.82	5.33	USA
PI213719	Reids_Yellow_Dent	38.4666667	-92.5	NAM	Usa	Corn_Belt_Dent	0.80	5.94	-12.53	8.31	5.58	USA
PI214295	Pride_of_Saline	38.5	-97.65	NAM	Usa	Corn_Belt_Dent	0.85	6.62	-12.42	8.35	5.45	USA
PI218137	Tesuque_Pueblo	35.8	-105.983333	NAM	Usa	Corn_Belt_Dent	0.76	-9.22	-8.17	-9.47	-9.55	USA
PI269743	Shoepeg	38	-93	NAM	Usa	Corn_Belt_Dent	0.94	5.51	-12.66	9.96	5.15	USA
PI280061	Lancaster_Surecrop	40.4833333	-82.55	NAM	Usa	Corn_Belt_Dent	0.89	1.97	-10.69	6.43	3.61	USA
PI311237	Hickory_King	34	-84.7	NAM	Usa	Corn_Belt_Dent	0.63	8.99	-11.74	8.28	4.42	USA
PPS1132	Smut Nose	45	-70	NAM	Usa	Corn_Belt_Dent	0.77	-14.30	-9.88	-2.10	-1.15	USA
PPS471	Minnesota_13	45.5833333	-94.4666667	NAM	Usa	Corn_Belt_Dent	0.86	4.57	-11.79	8.29	5.13	USA
PPS941	Curacalco	-39.5166667	-71.8666667	SAM	Chile	Corn_Belt_Dent	0.32	0.86	3.88	-8.67	9.30	CHI
PPS1016	Couserans	42.9166667	1.05	SWF	SWF	E-SWF	0.73	-4.74	0.35	3.76	1.02	E-SWF
PPS1036	Saint Creac.A	43.05	-0.01666667	SWF	SWF	E-SWF	0.58	-5.32	4.77	3.30	0.27	E-SWF
PPS1039	Sarrancolin	42.95	0.4	SWF	SWF	E-SWF	0.66	-5.85	3.40	3.36	0.63	E-SWF
PPS121	Beaumont	44.7666667	0.75	SWF	SWF	E-SWF	0.84	-17.42	4.84	1.59	-0.38	E-SWF
PPS1248	Pop Cornée.1	43.9	-0.48333333	SWF	SWF	E-SWF	0.54	-8.96	6.12	-5.45	4.47	E-SWF
PPS1249	Pop Cornée.2	43.9	-0.48333333	SWF	SWF	E-SWF	0.51	-7.96	7.01	-5.62	4.18	E-SWF
PPS1250	Le Muret.1	44.833	-0.825	SWF	SWF	E-SWF	0.75	-2.94	-2.40	4.62	2.56	E-SWF
PPS1251	Le Muret.2	44.833	-0.825	SWF	SWF	E-SWF	0.61	-2.06	0.11	3.82	1.78	E-SWF
PPS1252	Isle sur Tarn	43.8833333	1.78333333	SWF	SWF	E-SWF	0.79	-1.41	-7.36	2.30	1.37	E-SWF
PPS13	Aleu	42.8833333	1.26666667	SWF	SWF	E-SWF	0.66	-12.32	7.87	1.04	-1.33	E-SWF
PPS14	Massat	42.8666667	1.3833333	SWF	SWF	E-SWF	0.66	-12.35	6.41	1.69	-0.44	E-SWF
PPS15	Lacaune	43.55	2.5833333	SWF	SWF	E-SWF	0.77	-15.27	6.93	1.84	-0.41	E-SWF
PPS189	Saint Lieux les Lavaur	43.75	1.75	SWF	SWF	E-SWF	0.52	-8.30	9.25	1.49	-0.59	E-SWF
PPS191	Aston D	42.6666667	1.65	SWF	SWF	E-SWF	0.65	-10.57	7.54	2.22	-0.59	E-SWF
PPS198	Marmande.36	NA	NA	SWF	SWF	E-SWF	0.68	0.08	3.55	5.93	1.79	E-SWF
PPS226	Marmande.10	NA	NA	SWF	SWF	E-SWF	0.87	-12.67	3.90	1.77	0.86	E-SWF
PPS227	Marmande.21	NA	NA	SWF	SWF	E-SWF	0.88	-11.06	1.99	2.35	0.38	E-SWF
PPS228	Marmande.26	NA	NA	SWF	SWF	E-SWF	0.86	-14.12	4.85	2.48	-0.07	E-SWF
PPS229	Marmande.34	NA	NA	SWF	SWF	E-SWF	0.89	-16.72	0.05	-0.06	1.07	E-SWF
PPS230	Sainte Livrade	43.65	1.1	SWF	SWF	E-SWF	0.73	-12.98	5.96	1.69	0.15	E-SWF
PPS237	Seix	42.8166667	1.15	SWF	SWF	E-SWF	0.71	-13.49	7.55	1.53	-0.92	E-SWF
PPS241	Millette de Finham.AGPM	43.9166667	1.21666667	SWF	SWF	E-SWF	0.89	-14.72	5.32	1.03	-0.10	E-SWF
PPS362	Marmande.35	NA	NA	SWF	SWF	E-SWF	0.77	-2.42	-3.38	5.10	2.68	E-SWF
PPS478	Labastide de Boussignac	43.05	1.88333333	SWF	SWF	E-SWF	0.59	-5.74	4.27	2.71	0.93	E-SWF
PPS487	Cazalis	44.3333333	-0.38333333	SWF	SWF	E-SWF	0.53	-5.43	1.58	3.33	0.73	E-SWF
PPS501	Borie Lautrec	43.7	2.13333333	SWF	SWF	E-SWF	0.66	-10.82	8.38	1.57	-0.28	E-SWF
PPS502	Lautrec	43.7	2.13333333	SWF	SWF	E-SWF	0.56	-10.95	9.09	1.37	-0.36	E-SWF
PPS503	Semport	NA	NA	SWF	SWF	E-SWF	0.67	-11.23	8.08	1.50	-0.41	E-SWF

CodeColl	Commonname	LatSiteColl	LongSiteColl	ZoneOri	PaysOri	cluster.Name	membership	Axis1	Axis2	Axis3	Axis4	Id_country initial
PPS504	Villefranche d Albigeois.B	43.8833333	2.31666667	SWF	SWF	E-SWF	0.64	-10.65	8.23	1.64	-0.17	E-SWF
PPS505	Pampelone	44.1	2.23333333	SWF	SWF	E-SWF	0.82	-13.01	6.01	0.71	-0.60	E-SWF
PPS506	Poulinet Alban.A	43.8333333	2.41666667	SWF	SWF	E-SWF	0.61	-11.11	9.32	0.95	-0.41	E-SWF
PPS568	Moustajon	42.8	0.58333333	SWF	SWF	E-SWF	0.62	-2.87	-0.40	4.04	1.52	E-SWF
PPS574	Fraysse Alban	43.9166667	2.4	SWF	SWF	E-SWF	0.97	-14.49	4.18	1.32	0.51	E-SWF
PPS575	Cabrilles	43.7	2.13333333	SWF	SWF	E-SWF	0.97	-16.52	6.60	0.79	0.01	E-SWF
PPS59	Montirat Monesties	44.0666667	2.1	SWF	SWF	E-SWF	0.58	-10.92	8.13	1.90	-0.67	E-SWF
PPS597	Seix.Bardot	42.8166667	1.15	SWF	SWF	E-SWF	0.74	-13.66	7.23	1.69	-0.62	E-SWF
PPS60	Poulinet Alban.B	43.8333333	2.41666667	SWF	SWF	E-SWF	0.66	-11.72	8.67	1.16	-0.46	E-SWF
PPS609	Grand Roux Basque 3	NA	NA	SWF	SWF	E-SWF	0.91	-22.09	3.86	1.89	-0.07	E-SWF
PPS61	Vaour	44.0666667	1.8	SWF	SWF	E-SWF	0.90	-15.96	1.32	1.59	0.86	E-SWF
PPS614	Espeyrac	44.6166667	2.51666667	SWF	SWF	E-SWF	0.91	-14.95	3.33	1.74	0.73	E-SWF
PPS621	Pop Gers	NA	NA	SWF	SWF	E-SWF	0.60	11.45	-3.56	3.76	1.96	E-SWF
PPS624	Landes 3	NA	NA	SWF	SWF	E-SWF	0.86	-0.87	-8.72	4.09	1.59	E-SWF
PPS625	Landes 4	NA	NA	SWF	SWF	E-SWF	0.85	-0.56	-9.87	4.22	1.80	E-SWF
PPS626	Cugnieres	44.4166667	-0.53333333	SWF	SWF	E-SWF	0.56	-5.19	1.65	3.68	1.40	E-SWF
PPS631	St_Pantaleon	44.3666667	1.2833333	SWF	SWF	E-SWF	0.96	-18.15	2.67	1.56	0.30	E-SWF
PPS632	Cressensac 2	45.0166667	1.51666667	SWF	SWF	E-SWF	0.96	-16.50	2.83	1.51	0.08	E-SWF
PPS633	Marmande.12	NA	NA	SWF	SWF	E-SWF	0.84	-6.69	-0.22	3.59	1.27	E-SWF
PPS634	Marmande.45	NA	NA	SWF	SWF	E-SWF	0.58	-8.23	7.67	2.16	-0.17	E-SWF
PPS636	Marmande.19	NA	NA	SWF	SWF	E-SWF	0.52	-7.57	4.53	1.29	-0.26	E-SWF
PPS637	Marmande.20	NA	NA	SWF	SWF	E-SWF	0.54	-7.77	4.85	1.14	-0.45	E-SWF
PPS638	Marmande.32	NA	NA	SWF	SWF	E-SWF	0.89	-8.41	-0.77	3.79	1.73	E-SWF
PPS651	Auga 2	43.4666667	-0.36666667	SWF	SWF	E-SWF	0.65	-12.83	7.54	1.44	-1.26	E-SWF
PPS666	Villefranche d Albigeois.A	43.8833333	2.31666667	SWF	SWF	E-SWF	0.97	-16.24	6.03	0.90	0.46	E-SWF
PPS668	Millette_Montagne_Noire	43.3833333	2.2666667	SWF	SWF	E-SWF	0.58	-10.65	9.07	0.96	-0.59	E-SWF
PPS68	Riverenert 1	42.95	1.25	SWF	SWF	E-SWF	0.59	-12.85	8.55	1.81	-0.91	E-SWF
PPS69	Couflens.2	42.75	1.18333333	SWF	SWF	E-SWF	0.70	-13.39	7.73	1.01	-0.69	E-SWF
PPS70	Aston A	42.6666667	1.65	SWF	SWF	E-SWF	0.58	-9.36	9.00	2.09	-0.62	E-SWF
PPS71	Aston B	42.6666667	1.65	SWF	SWF	E-SWF	0.63	-5.15	2.44	3.11	0.71	E-SWF
PPS736	Ariege 1	42.7666667	1.5	SWF	SWF	E-SWF	0.51	-8.11	7.61	2.33	-0.29	E-SWF
PPS737	Couflens.1	42.75	1.18333333	SWF	SWF	E-SWF	0.72	-13.44	7.80	1.38	-1.14	E-SWF
PPS85	Floirac Camboni A	44.9	1.66666667	SWF	SWF	E-SWF	0.96	-18.43	4.01	1.05	0.20	E-SWF
PPS86	Floirac	44.9	1.66666667	SWF	SWF	E-SWF	0.59	-9.92	7.80	2.15	-0.24	E-SWF
PPS87	Cressensac 1	45.0166667	1.51666667	SWF	SWF	E-SWF	1.00	-21.64	2.54	1.14	0.32	E-SWF
PPS96	Marmande.31	NA	NA	SWF	SWF	E-SWF	0.72	-10.44	5.47	2.53	0.51	E-SWF
PPS966	Ariege 3	NA	NA	SWF	SWF	E-SWF	0.60	-10.79	7.64	2.35	-0.58	E-SWF
VENE736	Sabanero	8.4333333	-71.2	CAM	Venezuela	Italian	0.32	17.62	9.71	-8.41	3.81	VEN
PPS1012	Ratenelle	46.5166667	5	Europe	France	Italian	0.51	-15.92	3.70	2.70	0.95	FRA
PPS115	Marano	44.5166667	11.4333333	Europe	Italy	Italian	0.68	0.83	21.07	5.12	-5.55	ITA
PPS116	VA204_Spezia	44.1166667	9.8333333	Europe	Italy	Italian	0.85	7.12	22.62	7.21	-8.47	ITA
PPS187	VA1_Ott_Tortonese	44.7833333	10.3166667	Europe	Italy	Italian	0.94	10.85	23.72	-0.58	-3.70	ITA
PPS381	VA46_Quarantino_Famiglia	45.0333333	10.3166667	Europe	Italy	Italian	0.95	6.62	22.22	6.91	-4.72	ITA

CodeColl	Commonname	LatSiteColl	LongSiteColl	ZoneOri	PaysOri	cluster.Name	membership	Axis1	Axis2	Axis3	Axis4	Id_country initial
PPS686	VA5_Nostrano_dell_Isola	44.616	8.25	Europe	Italy	Italian	0.90	7.18	22.42	5.17	-5.82	ITA
PPS687	VA20_Pignoletto	44.65	7.483	Europe	Italy	Italian	0.81	7.40	22.13	7.47	-7.38	ITA
PPS735	VA84_Cinquantino_Rosso	45.566	11.916	Europe	Italy	Italian	0.53	1.37	21.12	5.74	-5.48	ITA
VA145	Nostrano_Quarantino	46.5	11.0666667	Europe	Italy	Italian	0.49	-8.09	16.20	3.45	-3.17	ITA
VA469	Trecchinese_di_Lagonegro	40.05	15.85	Europe	Italy	Italian	0.87	10.44	22.07	-0.46	-4.41	ITA
VA563	Locale_di_Macerata	43.1666667	10.5666667	Europe	Italy	Italian	0.49	6.55	21.11	2.40	-5.03	ITA
VA565	Cinquantino_di_Montagnana	46.0333333	10.3166667	Europe	Italy	Italian	0.70	6.71	23.83	5.55	-5.97	ITA
EPZMV159	Pabilillo_de_Granada	37.1833333	-3.6	Europe	Spain	Italian	0.61	10.42	13.98	2.57	-4.25	SPA
EPZMV310	Andaluz	38.0333333	-5.9833333	Europe	Spain	Italian	0.66	18.98	5.56	4.52	-3.54	SPA
PPS774	Trimesino_ABC	39.3666667	-5.9833333	Europe	Spain	Italian	0.53	15.56	8.16	4.53	-3.64	SPA
MRIZP47	Zuti_Rani	44.8	20.4666667	Europe	Yugoslavia	Italian	0.85	6.76	18.80	4.74	-4.23	YUG
CUBA12	Cuban_Flint	22.4166667	-83.7333333	Caribbean	Cuba	Mexican	0.82	22.43	-6.66	5.38	-0.11	CAR
GUAT45	Oloton	14.6333333	-90.5166667	CAM	Guatemala	Mexican	0.78	19.24	-5.51	0.64	-1.48	GUA
CHIH150	Palomero_Toluqueno	29.3666667	-111.083333	CAM	Mexico	Mexican	0.76	-1.07	-9.05	-12.06	-10.68	MEX
CHIH38	Apachito	27.8	-111.066667	CAM	Mexico	Mexican	0.62	4.03	-7.62	-14.05	-12.87	MEX
JALI102	Tabloncillo	19.4833333	-103.283333	CAM	Mexico	Mexican	0.63	7.21	-6.13	-8.67	-10.67	MEX
JALI71	Harinoso_de_Ocho_Occidentales	20.1666667	-103.05	CAM	Mexico	Mexican	0.74	5.37	-7.20	-11.54	-13.85	MEX
MEXI3	Conico	19.2833333	-99.65	CAM	Mexico	Mexican	0.66	4.33	-7.14	-17.66	-14.33	MEX
MEXI7	Cacahuacintle	19.2833333	-99.6333333	CAM	Mexico	Mexican	0.86	4.85	-7.65	-17.62	-14.83	MEX
NAYA24	Harinoso_de_Ocho	21.95	-105.3	CAM	Mexico	Mexican	0.46	6.41	-7.45	-8.17	-11.73	MEX
NAYA39	Reventador	21.9333333	-105.3	CAM	Mexico	Mexican	0.85	6.32	-7.34	-7.82	-12.77	MEX
OAXA50	Zapalote_Chico	16.5833333	-94.6	CAM	Mexico	Mexican	0.65	14.13	-4.71	-3.57	-7.51	MEX
PI217413	Zapalote_Chico	19.3833333	-99.15	CAM	Mexico	Mexican	0.54	13.93	-5.42	-3.48	-8.43	MEX
PUEB91	Arrocillo_Amarillo	19.75	-97.55	CAM	Mexico	Mexican	0.76	7.21	-6.47	-15.48	-13.11	MEX
SINA2	Chapalote	24.8	-107.4	CAM	Mexico	Mexican	0.67	4.80	-6.62	-9.42	-12.33	MEX
EPZMV173	Blanco_de_Ricote	38.1333333	-1.3666667	Europe	Spain	Mexican	0.96	15.38	9.50	4.33	-1.21	SPA
EPZMV23	Basto	37.3666667	-5.9833333	Europe	Spain	Mexican	0.40	12.88	-0.43	4.77	-0.90	SPA
EPZMV232	Amarillo_Pequeno_de_Bernedo	42.6166667	-3.1666667	Europe	Spain	Mexican	0.31	-14.79	4.59	1.82	-0.53	SPA
PPS444	Norteno	42.8666667	-8.55	Europe	Spain	Mexican	0.45	-12.26	4.90	-0.43	-1.67	SPA
PPS770	Andaluz_ABC	38.0166667	-4.5166667	Europe	Spain	Mexican	0.67	17.93	3.81	4.66	-3.34	SPA
PI213715	Gourdseed_Dent	37.5	-96	NAM	Usa	Mexican	0.97	9.33	-12.42	8.76	5.79	USA
PI217410	Papago_Flour_Corn	32.25	-112	NAM	Usa	Mexican	0.64	-2.44	-8.27	-10.20	-13.79	USA
ARGE486	Cuarento Cateto	-34.9	-57.95	SAM	Argentina	Mexican	0.31	15.76	3.27	2.80	-1.53	ARG
IPK60	Vorarlberger Weisser	47.4166667	9.75	Europe	Austria	Northern_Flint	0.94	-35.45	-1.68	2.91	2.89	AUS
PPS58	Schindelmeiser	48.25	16.3666667	Europe	Austria	Northern_Flint	0.97	-33.95	-0.04	3.73	4.72	AUS
PPS462	Stupicka_Rana	50.0833333	15.0833333	Europe	Czechoslovakia	Northern_Flint	0.89	-29.83	1.80	5.81	0.40	CZE
PPS19	Liesse	49.6	3.8166667	Europe	France	Northern_Flint	0.56	-13.21	7.37	2.61	-1.04	FRA
PPS474	Ain	46.2	5.2166667	Europe	France	Northern_Flint	0.82	-17.73	8.96	1.66	-0.23	FRA
PPS5	Wantzenau	48.65	7.8166667	Europe	France	Northern_Flint	0.86	-32.74	-2.10	2.50	2.34	FRA
IPK15	Caspersmeyer_II	52.5	12.2166667	Europe	Germany	Northern_Flint	0.84	-29.62	-2.75	1.13	3.16	GER
IPK158	Mahndorfer	52.0666667	11.0666667	Europe	Germany	Northern_Flint	0.96	-33.04	-1.48	2.14	2.78	GER
IPK239	Strenzfelder	51.7833333	11.7166667	Europe	Germany	Northern_Flint	0.96	-31.09	-1.35	2.33	1.61	GER
IPK25	Dr._Delliles_Körnermais	52.5	10.8833333	Europe	Germany	Northern_Flint	0.97	-33.73	-3.02	3.59	4.23	GER

CodeColl	Commonname	LatSiteColl	LongSiteColl	ZoneOri	PaysOri	cluster.Name	membership	Axis1	Axis2	Axis3	Axis4	Id_country initial
IPK26	Rimpaus_Binder	52.1	10.1833333	Europe	Germany	Northern_Flint	0.79	-17.11	9.56	3.12	-3.33	GER
IPK6	Terras_Binder	52.1	10.85	Europe	Germany	Northern_Flint	0.95	-28.48	2.33	2.80	-0.56	GER
IPK9	Pautzfelder	49	11	Europe	Germany	Northern_Flint	0.94	-33.59	-1.78	3.01	3.59	GER
Pionner	Jaune de bade	48.5666667	8.5833333	Europe	Germany	Northern_Flint	0.84	-32.17	-2.20	2.30	2.81	GER
PPS151	Martonvasari	47.5166667	19.7	Europe	Hungary	Northern_Flint	0.66	-6.90	-9.47	6.51	4.07	HUN
PPS458	Iregszemscei_12_hetes	47.5166667	20.3666667	Europe	Hungary	Northern_Flint	0.59	-16.18	7.21	3.11	-0.42	HUN
PPS534	Magyarovari	47.5166667	21.0333333	Europe	Hungary	Northern_Flint	0.45	-19.02	-3.09	3.67	3.44	HUN
VA136	Nostrano_di_Bolzano	46.8	11.9	Europe	Italy	Northern_Flint	0.90	-32.19	2.12	3.68	2.16	ITA
KRO1	Szyldeca	52.4166667	16.8833333	Europe	Poland	Northern_Flint	0.97	-31.45	-0.30	2.93	1.74	POL
KRO2	Oksonska	52.4166667	17.55	Europe	Poland	Northern_Flint	0.74	-16.79	8.33	2.66	-1.54	POL
KRO3	Rozyllo	52.4166667	18.2166667	Europe	Poland	Northern_Flint	0.44	-13.58	7.99	2.65	-0.54	POL
PPS448	Wielkopolenka	51.15	21.8333333	Europe	Poland	Northern_Flint	0.91	-30.73	2.52	2.68	1.49	POL
PPS160	Ariesan	44.4166667	26.1	Europe	Romania	Northern_Flint	0.34	-11.99	3.19	3.50	1.10	ROM
PPS94	Reinthaler	46.9166667	7.45	Europe	Switzerland	Northern_Flint	0.97	-34.83	-2.21	1.94	2.50	SWI
UKR37	Ukraine_1	50.4166667	30.5166667	Europe	Ukraine	Northern_Flint	0.93	-23.80	0.95	-0.23	0.67	UKR
MRIZP676	Zuti_Murt-Optak	43.4666667	20.4666667	Europe	Yugoslavia	Northern_Flint	0.48	-15.84	7.72	3.57	0.06	YUG
Ames2755	Canada_Yellow_Flint	52	-106	NAM	Canada	Northern_Flint	0.58	-29.85	-7.50	1.12	0.42	CAN
Ames2749	Palmer_Rhode_Island_Flint	41.9	-71.9	NAM	Usa	Northern_Flint	0.96	-34.03	-7.07	0.42	1.65	USA
Ames2750	Wakefields Rhode Island Flint	41.9	-72.5666667	NAM	Usa	Northern_Flint	0.96	-30.34	-7.40	-1.69	-0.88	USA
Ames2752	Quicks_Rhode_Island_Flint	42.45	-72.0166667	NAM	Usa	Northern_Flint	0.93	-32.22	-7.00	-0.60	1.60	USA
PI213763	Mandan_Pyrenees_Galicia__Flint_Flour	47.55	-99.5833333	NAM	Usa	Northern_Flint	0.64	-21.89	-8.26	-4.07	-3.10	USA
PI213793	Assiniboine	46.8166667	-100.816667	NAM	Usa	Northern_Flint	0.74	-22.53	-7.71	-4.33	-3.48	USA
PI217460	King Philip	45	-70.6666667	NAM	Usa	Northern_Flint	0.92	-27.23	-7.00	1.47	2.60	USA
PI401755	Sioux Tribe	43.0166667	-102.55	NAM	Usa	Northern_Flint	0.48	-23.47	-7.86	-4.54	-3.88	USA
PI401757	Gaspe_Flint	48.8166667	-64.25	NAM	Usa	Northern_Flint	0.89	-33.18	-6.00	-0.22	0.83	USA
PPS920	Balsa	-34.8	-61.6666667	SAM	Argentina	Northern_Flint	0.75	-25.89	-7.08	-3.99	-2.63	ARG
PPS961	Choshuenco_Aesa	-40.2333333	-72.0166667	SAM	Chile	Northern_Flint	0.28	-7.32	0.09	-7.65	7.41	CHI
Por10	MILHO_BRANCO_PEROLA_VERDIAL	41.467	-7.85	Europe	Portugal	Others	0.00	0.41	-0.79	-3.08	2.92	POR
Por17	MILHO_BRANCO	40.6	-8.083	Europe	Portugal	Others	0.00	-8.94	5.18	-4.32	3.47	POR
ARGE306	NA	-34.5	-68.5	SAM	Argentina	Others	0.00	12.55	-2.37	2.73	2.17	ARG
ARGE472	NA	-23.2	-65.35	SAM	Argentina	Others	0.00	17.37	8.08	-18.30	15.79	ARG
BRAZ2441	NA	-15.78	-47.936	SAM	Brazil	Others	0.00	11.60	-3.81	1.94	2.19	BRA
BRAZ60	NA	-22.217	-54.8	SAM	Brazil	Others	0.00	21.68	-1.61	-2.22	9.38	BRA
CHZM08060	NA	-37.75	-72	SAM	Chile	Others	0.00	-7.66	-10.44	-10.79	-7.06	CHI
PARA183	NA	-27	-57	SAM	Paraguay	Others	0.00	9.47	-4.99	-0.97	10.94	PAR
PERU590	NA	-25.298	-57.637	SAM	Paraguay	Others	0.00	20.47	2.59	-8.24	10.07	PAR
PARA167	NA	-12.083	-76.95	SAM	Peru	Others	0.00	14.70	-2.35	-2.59	12.94	PER
URUG637	NA	-34.337	-57.729	SAM	Uruguay	Others	0.00	11.56	-2.96	1.86	2.83	URU
PPS17	Barisis	49.5666667	3.3166667	Europe	France	Pyrenees_Galicia_2	0.72	-4.55	11.63	3.17	-2.34	FRA
PPS1234	Portugal	38.7166667	-9.1166667	Europe	Portugal	Pyrenees_Galicia_2	0.48	-8.90	6.67	-4.17	3.49	POR
EPZMV168	Aduna	43.2666667	-2.0333333	Europe	Spain	Pyrenees_Galicia_2	0.69	-6.58	8.22	0.26	0.70	SPA
EPZMV171	Rojo_de_Tolosa	43.1166667	-2.0666667	Europe	Spain	Pyrenees_Galicia_2	0.33	4.35	4.21	5.66	-1.55	SPA
EPZMV70	Vasco_1	42.8166667	-1.6166667	Europe	Spain	Pyrenees_Galicia_2	0.75	-8.33	9.42	0.92	-1.82	SPA

CodeColl	Commonname	LatSiteColl	LongSiteColl	ZoneOri	PaysOri	cluster.Name	membership	Axis1	Axis2	Axis3	Axis4	Id_country initial
PPS979	Agraar_la_Coruna_1	43.3666667	-8.3666667	Europe	Spain	Pyrenees_Galicia_2	0.63	4.32	-3.77	6.56	2.16	SPA
PPS983	Orense	42.3166667	-7.8333333	Europe	Spain	Pyrenees_Galicia_2	0.77	-6.17	10.45	-0.84	0.97	SPA
PPS986	Orendain	43.0666667	-2.1166667	Europe	Spain	Pyrenees_Galicia_2	0.34	7.23	2.90	3.69	-1.00	SPA
PPS938	Salta Lomocura	-38.85	-72.3666667	SAM	Chile	Pyrenees_Galicia_2	0.29	0.97	3.82	-12.82	12.58	CHI
PPS1	Salies de Bearn 1	43.4666667	-0.9	SWF	SWF	W-SWF	0.98	-3.43	11.38	2.47	-1.82	W-SWF
PPS10	Estarvielle	42.8166667	0.4166667	SWF	SWF	W-SWF	0.85	-5.70	10.35	1.87	-1.27	W-SWF
PPS1031	Arrien	43.3	-0.15	SWF	SWF	W-SWF	0.95	-2.92	11.90	2.90	-1.29	W-SWF
PPS1037	Saint Creac.B	43.05	-0.01666667	SWF	SWF	W-SWF	0.83	-4.47	9.76	2.35	-1.30	W-SWF
PPS1038	Saint Creac.C	43.05	-0.01666667	SWF	SWF	W-SWF	0.92	-3.50	11.01	2.60	-1.01	W-SWF
PPS1040	Monein 2	43.2833333	-0.55	SWF	SWF	W-SWF	0.87	-1.62	9.66	3.26	-0.54	W-SWF
PPS1041	Gazost	43	0.01666667	SWF	SWF	W-SWF	0.87	-3.22	10.07	2.98	-0.93	W-SWF
PPS1042	Morlaas	43.3333333	-0.25	SWF	SWF	W-SWF	0.94	-2.99	10.11	2.72	-1.35	W-SWF
PPS1043	Argeles Gazost	42.9	-0.25	SWF	SWF	W-SWF	0.92	-3.33	11.41	2.23	-1.40	W-SWF
PPS1044	Pierrefitte	42.9333333	-0.01666667	SWF	SWF	W-SWF	0.96	-3.23	10.37	1.79	-1.61	W-SWF
PPS1045	Bagnères de Bigorre 1	43.05	0.16666667	SWF	SWF	W-SWF	0.89	-4.19	10.33	3.05	-1.26	W-SWF
PPS1046	Arrens	42.9	-0.25	SWF	SWF	W-SWF	0.97	-3.70	11.21	2.69	-1.71	W-SWF
PPS1047	Saint Engrace	42.9833333	-0.81666667	SWF	SWF	W-SWF	0.99	-1.42	11.88	2.22	-1.92	W-SWF
PPS1048	Orthez d'Asson.A	43.0833333	-0.25	SWF	SWF	W-SWF	0.94	-3.18	10.92	2.67	-1.22	W-SWF
PPS1049	Orthez d'Asson.B	43.0833333	-0.25	SWF	SWF	W-SWF	0.97	-3.79	11.77	2.64	-1.31	W-SWF
PPS1050	Tarbes	43.25	0	SWF	SWF	W-SWF	0.97	-3.73	11.73	2.66	-1.30	W-SWF
PPS1051	Lourdes	43.0833333	0.03333333	SWF	SWF	W-SWF	0.94	-3.52	11.27	2.28	-1.70	W-SWF
PPS1052	Bagnères de Bigorre 2	43.05	0.16666667	SWF	SWF	W-SWF	0.93	-5.05	10.54	3.10	-1.25	W-SWF
PPS1053	Ourdes	42.9666667	0.53333333	SWF	SWF	W-SWF	0.97	-2.88	11.37	2.43	-1.85	W-SWF
PPS1054	Monein 3	43.2833333	-0.55	SWF	SWF	W-SWF	0.99	-2.24	11.97	2.44	-1.80	W-SWF
PPS1057	Lestelle Betharram	43.1166667	-0.2	SWF	SWF	W-SWF	0.87	-5.29	9.14	2.37	-0.95	W-SWF
PPS1058	Gurmencon	43.15	-0.6	SWF	SWF	W-SWF	1.00	-2.90	12.57	2.75	-1.64	W-SWF
PPS11	Sost	42.8833333	0.55	SWF	SWF	W-SWF	0.73	-8.25	10.14	1.81	-1.64	W-SWF
PPS1153	Col du Soulor	42.9	-0.25	SWF	SWF	W-SWF	0.83	-3.95	8.21	2.68	-1.27	W-SWF
PPS1154	Oloron Sainte Marie	43.15	-0.58333333	SWF	SWF	W-SWF	1.00	-2.28	12.35	3.16	-1.17	W-SWF
PPS12	Portet	42.8133	0.4861	SWF	SWF	W-SWF	0.67	-7.92	9.24	2.21	-1.46	W-SWF
PPS122	Millette du Lauragais 1	43.383	1.717	SWF	SWF	W-SWF	0.51	-8.58	9.12	1.69	-1.10	W-SWF
PPS1236	Coarraze	43.1666667	-0.2	SWF	SWF	W-SWF	0.55	-1.32	2.86	4.42	0.75	W-SWF
PPS1421	Pouyastruc	43.267	0.167	SWF	SWF	W-SWF	0.90	-1.71	9.29	2.89	-0.43	W-SWF
PPS192	Cambon Le Lavaur	43.5833333	1.85	SWF	SWF	W-SWF	0.57	-7.62	9.39	1.80	-0.28	W-SWF
PPS193	Aurignac	43.2	0.86666667	SWF	SWF	W-SWF	0.52	-8.27	9.48	1.23	-0.78	W-SWF
PPS194	Grand Roux Basque 1	NA	NA	SWF	SWF	W-SWF	0.51	4.69	0.08	4.78	0.53	W-SWF
PPS196	Moulin Neuf	43.0666667	1.93333333	SWF	SWF	W-SWF	0.54	-7.31	8.42	1.88	-0.51	W-SWF
PPS199	Parleboscq	43.9333333	0.03333333	SWF	SWF	W-SWF	0.56	-6.71	8.86	2.05	-1.01	W-SWF
PPS2	Saint Jean Pied de Port 1	43.15	-1.21666667	SWF	SWF	W-SWF	0.95	-2.63	11.35	2.79	-1.78	W-SWF
PPS20	Campan Galade	42.95	0.21666667	SWF	SWF	W-SWF	0.91	-4.80	11.44	2.41	-1.64	W-SWF
PPS21	Sainte Marie de Campan	42.9666667	0.25	SWF	SWF	W-SWF	0.93	-4.49	11.08	1.88	-1.46	W-SWF
PPS22	Esterre	42.8666667	0	SWF	SWF	W-SWF	0.96	-2.93	11.70	2.71	-1.98	W-SWF
PPS23	Laruns	42.8833333	-0.4166667	SWF	SWF	W-SWF	0.98	-2.53	11.76	2.31	-1.21	W-SWF

CodeColl	Commonname	LatSiteColl	LongSiteColl	ZoneOri	PaysOri	cluster.Name	membership	Axis1	Axis2	Axis3	Axis4	Id_country initial
PPS235	Grand Roux Basque 2	NA	NA	SWF	SWF	W-SWF	0.99	-2.76	11.86	2.36	-1.48	W-SWF
PPS238	Riverenert 2	42.95	1.25	SWF	SWF	W-SWF	0.60	-8.08	9.27	1.58	-1.15	W-SWF
PPS24	Viey	42.8833333	0.01666667	SWF	SWF	W-SWF	0.97	-3.19	11.99	2.08	-1.85	W-SWF
PPS240	Saint Sauveur	42.8166667	0.01666667	SWF	SWF	W-SWF	0.95	-3.62	11.93	1.72	-1.60	W-SWF
PPS243	Floirac Camboni B	44.9	1.66666667	SWF	SWF	W-SWF	0.50	-7.82	8.40	1.99	-0.74	W-SWF
PPS25	Bize	43.0333333	0.46666667	SWF	SWF	W-SWF	0.78	-6.98	10.38	2.78	-1.04	W-SWF
PPS26	Argein	42.9166667	0.98333333	SWF	SWF	W-SWF	0.61	-8.95	9.36	2.04	-1.52	W-SWF
PPS27	Saint Christophe	44.4666667	2.38333333	SWF	SWF	W-SWF	0.69	-4.21	6.92	3.22	-0.13	W-SWF
PPS35	Gironde 1	NA	NA	SWF	SWF	W-SWF	0.86	-3.95	11.22	1.63	-1.48	W-SWF
PPS361	Cassagnabere 1	43.2166667	0.8	SWF	SWF	W-SWF	0.73	-7.10	10.00	1.75	-1.59	W-SWF
PPS363	Landes 2	43.883	-1.05	SWF	SWF	W-SWF	0.61	-13.15	7.20	2.14	0.32	W-SWF
PPS410	Gironde 2	NA	NA	SWF	SWF	W-SWF	0.58	-2.08	5.04	3.39	-0.20	W-SWF
PPS442	Gironde 3	NA	NA	SWF	SWF	W-SWF	0.83	-4.34	9.74	2.03	-0.72	W-SWF
PPS460	Gironde 4	NA	NA	SWF	SWF	W-SWF	0.87	-5.23	10.84	1.97	-1.40	W-SWF
PPS482	Cuguron	43.1	0.51666667	SWF	SWF	W-SWF	0.72	-7.46	10.23	2.34	-1.33	W-SWF
PPS483	Gironde 5	NA	NA	SWF	SWF	W-SWF	0.55	-3.16	3.50	3.03	0.02	W-SWF
PPS484	Gironde 6	NA	NA	SWF	SWF	W-SWF	0.91	-4.44	11.49	1.75	-1.67	W-SWF
PPS485	Gironde 7	NA	NA	SWF	SWF	W-SWF	0.83	-4.95	9.30	2.29	-1.10	W-SWF
PPS486	Gironde 8	NA	NA	SWF	SWF	W-SWF	0.79	-5.47	9.31	2.13	-1.22	W-SWF
PPS488	Estaut 1	42.9	-0.51666667	SWF	SWF	W-SWF	1.00	-2.50	11.54	2.35	-2.85	W-SWF
PPS490	Bugard	43.25	0.31666667	SWF	SWF	W-SWF	0.85	-4.32	11.36	2.39	-1.03	W-SWF
PPS491	Saint Laurent de Neste 3	43.0833333	0.46666667	SWF	SWF	W-SWF	0.82	-5.50	10.87	2.57	-1.30	W-SWF
PPS498	Estaut 2	42.9	-0.51666667	SWF	SWF	W-SWF	0.99	-2.05	11.58	2.55	-2.42	W-SWF
PPS566	Boutx A	42.9166667	0.76666667	SWF	SWF	W-SWF	0.65	-8.79	9.79	1.17	-0.98	W-SWF
PPS567	Boutx B	42.9166667	0.76666667	SWF	SWF	W-SWF	0.60	-8.31	9.67	1.22	-1.23	W-SWF
PPS570	Moncassin	43.45	0.46666667	SWF	SWF	W-SWF	0.87	-4.18	10.33	2.10	-1.59	W-SWF
PPS585	Gironde 9	NA	NA	SWF	SWF	W-SWF	0.96	-4.22	11.33	1.43	-1.72	W-SWF
PPS598	Bugard A	43.25	0.31666667	SWF	SWF	W-SWF	0.95	-3.22	10.66	2.73	-1.30	W-SWF
PPS599	Burg	43.1833333	0.31666667	SWF	SWF	W-SWF	0.98	-3.81	11.60	2.37	-1.65	W-SWF
PPS600	Serre Rustaing	43.25	0.28333333	SWF	SWF	W-SWF	0.95	-3.56	11.14	2.66	-1.38	W-SWF
PPS601	Nestier 2	43.05	0.46666667	SWF	SWF	W-SWF	0.73	-7.39	10.12	1.49	-1.42	W-SWF
PPS602	Saint Laurent de Neste 4	43.0833333	0.46666667	SWF	SWF	W-SWF	0.89	-5.66	10.81	1.74	-1.27	W-SWF
PPS603	Grust	42.8833333	-0.05	SWF	SWF	W-SWF	0.99	-3.45	11.85	2.33	-1.90	W-SWF
PPS604	Lannemezan	43.1166667	0.38333333	SWF	SWF	W-SWF	0.87	-4.76	10.84	2.97	-1.38	W-SWF
PPS607	Doré de Bordes	43.2333333	-0.26666667	SWF	SWF	W-SWF	0.95	-3.07	11.19	2.52	-1.16	W-SWF
PPS608	Doré du Pays	NA	NA	SWF	SWF	W-SWF	0.98	-3.51	11.50	2.91	-1.74	W-SWF
PPS610	Grand Roux Basque 4	NA	NA	SWF	SWF	W-SWF	0.99	-1.71	12.25	3.02	-1.19	W-SWF
PPS611	Saint Jean Pied de Port 2	43.15	-1.21666667	SWF	SWF	W-SWF	1.00	-2.80	12.02	2.65	-1.42	W-SWF
PPS613	Millette La Cassignole	NA	NA	SWF	SWF	W-SWF	0.53	-8.35	9.42	1.69	-1.37	W-SWF
PPS615	Saint Come	44.5333333	2.81666667	SWF	SWF	W-SWF	0.54	-9.33	7.62	2.80	-0.48	W-SWF
PPS616	Saint Parthem.A	44.6166667	2.28333333	SWF	SWF	W-SWF	0.55	-8.74	6.75	2.68	0.31	W-SWF
PPS617	Saint Parthem.B	44.6166667	2.28333333	SWF	SWF	W-SWF	0.69	-7.61	8.85	2.30	-1.04	W-SWF
PPS622	Roux de Chalosse	43.6166667	-0.83333333	SWF	SWF	W-SWF	0.88	-3.68	10.49	2.44	-0.59	W-SWF

CodeColl	Commonname	LatSiteColl	LongSiteColl	ZoneOri	PaysOri	cluster.Name	membership	Axis1	Axis2	Axis3	Axis4	Id_country initial
PPS623	CETA des Landes	NA	NA	SWF	SWF	W-SWF	0.66	-3.78	5.64	3.58	0.04	W-SWF
PPS627	Pessan	44.2166667	-0.51666667	SWF	SWF	W-SWF	0.52	-0.60	1.45	4.04	0.76	W-SWF
PPS628	Dubos	44.2166667	-0.51666667	SWF	SWF	W-SWF	0.62	-4.29	5.16	3.02	-0.14	W-SWF
PPS630	Quarantaun	43.5333333	-1.45	SWF	SWF	W-SWF	0.80	-2.96	8.51	2.32	-1.40	W-SWF
PPS639	Millette_du_Lauragais	43.3833333	1.7166667	SWF	SWF	W-SWF	0.51	-8.09	8.94	1.16	-1.08	W-SWF
PPS640	Doré du Pays Louvrigny	43.5	-0.43333333	SWF	SWF	W-SWF	1.00	-2.27	12.26	2.89	-1.65	W-SWF
PPS641	Grand Roux Basque	43.4666667	-1.4666667	SWF	SWF	W-SWF	0.70	7.16	4.31	3.64	-1.14	W-SWF
PPS642	Roux du Pays	NA	NA	SWF	SWF	W-SWF	0.94	-3.23	10.99	2.52	-1.33	W-SWF
PPS643	Salies du Bearn 2	43.4666667	-0.9	SWF	SWF	W-SWF	0.97	-3.14	12.16	2.81	-1.65	W-SWF
PPS644	Sainte Suzanne	NA	NA	SWF	SWF	W-SWF	0.94	-3.74	11.73	2.71	-1.98	W-SWF
PPS645	Bordes	43.2333333	-0.26666667	SWF	SWF	W-SWF	0.67	4.49	4.65	3.02	-0.93	W-SWF
PPS646	Barzin	43.2166667	-0.11666667	SWF	SWF	W-SWF	0.96	-3.84	11.09	2.13	-1.63	W-SWF
PPS647	Lucgarier 1	43.2166667	-0.18333333	SWF	SWF	W-SWF	0.94	-3.28	11.04	2.18	-1.12	W-SWF
PPS648	Lucgarier 2	43.2166667	-0.18333333	SWF	SWF	W-SWF	0.95	-3.28	11.33	2.42	-1.67	W-SWF
PPS649	Gomer	43.2333333	-0.18333333	SWF	SWF	W-SWF	0.76	-6.79	10.38	2.50	-0.94	W-SWF
PPS65	Sainte Croix	43.1166667	1.16666667	SWF	SWF	W-SWF	0.72	-7.35	9.03	2.45	-1.11	W-SWF
PPS650	Auga 1	43.4666667	-0.36666667	SWF	SWF	W-SWF	0.97	-3.37	12.24	2.43	-1.43	W-SWF
PPS652	Mouhoux	43.4833333	-0.23333333	SWF	SWF	W-SWF	0.96	-3.17	11.30	2.17	-1.36	W-SWF
PPS653	Saint Jean Pied de Port 3	43.15	-1.21666667	SWF	SWF	W-SWF	0.99	-2.98	11.99	2.71	-1.36	W-SWF
PPS66	Cazavet	42.9833333	1.03333333	SWF	SWF	W-SWF	0.69	-7.26	8.61	2.03	-1.23	W-SWF
PPS67	Prat Bonrepaux	43.0333333	1.01666667	SWF	SWF	W-SWF	0.67	-7.22	9.87	1.85	-1.23	W-SWF
PPS72	Vidou 1	43.3	0.33333333	SWF	SWF	W-SWF	0.87	-5.68	10.28	2.22	-1.53	W-SWF
PPS73	Begole	43.15	0.31666667	SWF	SWF	W-SWF	0.97	-2.94	12.06	2.72	-1.46	W-SWF
PPS76	Nestier 1	43.05	0.46666667	SWF	SWF	W-SWF	0.82	-6.21	10.85	2.42	-1.71	W-SWF
PPS77	Saint Laurent de Neste 1	43.0833333	0.46666667	SWF	SWF	W-SWF	0.85	-5.75	11.07	1.98	-1.82	W-SWF
PPS78	Saint Laurent de Neste 2	43.0833333	0.46666667	SWF	SWF	W-SWF	0.87	-5.32	10.75	2.54	-1.49	W-SWF
PPS79	Aspin ez Angles	43.0833333	0.01666667	SWF	SWF	W-SWF	0.99	-4.18	11.67	2.69	-2.02	W-SWF
PPS8	Salechan	42.95	0.63333333	SWF	SWF	W-SWF	0.71	-6.86	9.48	2.05	-1.13	W-SWF
PPS80	Esquieze Serre	42.8666667	0	SWF	SWF	W-SWF	0.99	-3.66	11.73	2.35	-1.53	W-SWF
PPS81	Luz	42.8166667	0.01666667	SWF	SWF	W-SWF	0.97	-2.98	11.96	2.37	-1.61	W-SWF
PPS818	Lucq de Bearn	43.3	-0.63333333	SWF	SWF	W-SWF	0.86	-2.38	9.09	3.65	-1.15	W-SWF
PPS82	Bazus Aure	42.85	0.35	SWF	SWF	W-SWF	0.86	-4.93	11.00	2.23	-1.92	W-SWF
PPS820	Blanche Htes.Pyr.C1	NA	NA	SWF	SWF	W-SWF	0.75	-6.17	10.61	2.03	-1.90	W-SWF
PPS821	Val du Gave d'Aspe	43.05	0.25	SWF	SWF	W-SWF	0.99	-2.61	12.21	2.90	-1.82	W-SWF
PPS822	Monein C1	43.2833333	-0.55	SWF	SWF	W-SWF	0.92	-1.49	10.64	3.10	-0.99	W-SWF
PPS823	Lourdios.C1	43.0333333	-0.65	SWF	SWF	W-SWF	0.99	-2.68	12.04	2.39	-1.88	W-SWF
PPS83	Guchan	42.8333333	0.33333333	SWF	SWF	W-SWF	0.78	-7.94	10.60	2.57	-1.30	W-SWF
PPS84	Sauveterre de Comminges 1	43.0333333	0.66666667	SWF	SWF	W-SWF	0.71	-8.07	10.03	1.87	-1.73	W-SWF
PPS9	Bareilles	42.8833333	0.43333333	SWF	SWF	W-SWF	0.80	-5.02	10.14	2.21	-2.01	W-SWF
PPS962	Asson	43.0833333	-0.23333333	SWF	SWF	W-SWF	0.97	-3.50	10.90	2.68	-1.02	W-SWF
PPS963	Val d'Aran	42.85	0.73333333	SWF	SWF	W-SWF	0.50	-7.41	6.99	2.14	0.19	W-SWF
PPS964	Argeles	43.0833333	0.2	SWF	SWF	W-SWF	0.95	-2.52	11.87	2.37	-1.65	W-SWF
PPS965	Barzun	43.2166667	-0.11666667	SWF	SWF	W-SWF	0.99	-2.99	12.30	2.58	-1.46	W-SWF

CodeColl	Commonname	LatSiteColl	LongSiteColl	ZoneOri	PaysOri	cluster.Name	membership	Axis1	Axis2	Axis3	Axis4	Id_country initial
PPS967	Pontacq	43.2	-0.1	SWF	SWF	W-SWF	0.97	-3.07	11.90	2.36	-1.72	W-SWF
PPS968	Urdos	42.8333333	-0.51666667	SWF	SWF	W-SWF	0.99	-3.42	11.20	1.84	-1.65	W-SWF
PPS969	Monein 1	43.2833333	-0.55	SWF	SWF	W-SWF	0.93	-2.29	11.39	2.75	-1.36	W-SWF
PPS98	Landes 1	43.717	-1.05	SWF	SWF	W-SWF	0.92	-3.78	11.55	1.97	-1.56	W-SWF

S3 Table. Climatic variations between E-SWF and W-SWF groups from May to October.

This table contained mean of temperature and mean of precipitation for W-SWF and E-SWF groups from May to October. We shaded in bold months for which differencies for mean temperatures between the two groups were significant (p-value<5%). Same thing was doing for precipitation.

admixture groups at K=2	Precipitation (mm)						Temperature (°C)					
	May	June	July	August	**September**	**October**	May	*June*	*July*	*August*	September	October
E-SWF	88,64	79,87	54,45	73,09	**76,11**	**74,32**	12,97	**16,7**	**18,98**	**18,87**	16,36	11,68
W-SWF	92,88	84,43	57,32	75,8	**80,19**	**79,61**	12,06	**15,73**	**17,89**	**17,89**	15,61	11,11
pvalues	0,1	0,1	0,13	0,19	0,02	0,03	0,06	0,03	0,01	0,02	0,08	0,17

S4 Table. TreeMix 3-population test results.

This table represents a summary of TreeMix 3-population test results performed using all combinations of three groups shaped by each of the two groups from SWF (E-SWF and W-SWF) considered as focal group, and two other groups (Reference groups 1 and 2) considered as the two potential ancestors of the focal group. MEX= Mexican, AND= Andean, CAR= Caribbean, NFA= Northern flint material from America, PYR_GAL2= Pyrenees_Galicia2, NFE= Northern flint material from Europe, CBD= Corn belt dent, ITA= Italian, E-SWF= East South-West France, W-SWF= West South-West France.

Focus group	Reference group 1	Reference group 2	f3-stat	sd	Zscore
E-SWF	AND	CAR	0.0497717	0.00205181	24.2575
E-SWF	AND	CBD	0.024694	0.00120503	20.4925
E-SWF	AND	ITA	0.03689	0.00194592	18.9576
E-SWF	AND	MEX	0.0418331	0.00177397	23.5817
E-SWF	AND	NFA	-0.00314147	0.00076638	-4.09912
E-SWF	AND	NFE	-0.00116918	0.00078013	-1.4987
E-SWF	AND	PYR_GAL2	0.0175387	0.00147696	11.8748
E-SWF	CAR	CBD	0.0341757	0.00207075	16.504
E-SWF	CAR	ITA	0.0413742	0.00224242	18.4507
E-SWF	CAR	MEX	0.0512962	0.00213462	24.0305
E-SWF	CAR	NFA	-0.00897185	0.0008593	-10.4409
E-SWF	CAR	NFE	-0.00739079	0.00082723	-8.93438
E-SWF	CAR	PYR_GAL2	0.0168397	0.00153525	10.9687
E-SWF	CBD	ITA	0.0168069	0.00157083	10.6994
E-SWF	CBD	MEX	0.0287615	0.00135882	21.1665
E-SWF	CBD	NFA	0.00682359	0.00099255	6.87478
E-SWF	CBD	NFE	0.00419657	0.00084092	4.99045
E-SWF	CBD	PYR_GAL2	0.00710907	0.00086584	8.21064
E-SWF	ITA	MEX	0.0306017	0.00177774	17.2139
E-SWF	ITA	NFA	-0.00874928	0.00073018	-11.9825
E-SWF	ITA	NFE	-0.00096053	0.00073961	-1.2987
E-SWF	ITA	PYR_GAL2	0.0200593	0.00154623	12.973
E-SWF	MEX	NFA	0.00062584	0.00069	0.907024
E-SWF	MEX	NFE	9.76E-05	0.0008087	0.120663
E-SWF	MEX	PYR_GAL2	0.0138033	0.00126795	10.8863
E-SWF	NFA	NFE	0.026499	0.0012767	20.7559
E-SWF	NFA	PYR_GAL2	0.00091267	0.00065504	1.39331
E-SWF	NFE	PYR_GAL2	0.00435525	0.00071467	6.09406
W-SWF	AND	CAR	0.0432697	0.00283688	15.2526
W-SWF	AND	CBD	0.0303737	0.00177867	17.0766
W-SWF	AND	ITA	0.027425	0.00255536	10.7323

Focus group	Reference group 1	Reference group 2	f3-stat	sd	Zscore
W-SWF	AND	MEX	0.0399197	0.00249446	16.0034
W-SWF	AND	NFA	0.0123474	0.0019643	6.28592
W-SWF	AND	NFE	0.0101058	0.0018471	5.47115
W-SWF	AND	PYR_GAL2	0.00643596	0.00111786	5.7574
W-SWF	CAR	CBD	0.0380955	0.00266752	14.2812
W-SWF	CAR	ITA	0.0301494	0.00301023	10.0156
W-SWF	CAR	MEX	0.047623	0.00287876	16.5429
W-SWF	CAR	NFA	0.00475717	0.0018128	2.62421
W-SWF	CAR	NFE	0.00212429	0.00164714	1.28968
W-SWF	CAR	PYR_GAL2	0.00397715	0.00098807	4.02517
W-SWF	CBD	ITA	0.0177637	0.00188313	9.43308
W-SWF	CBD	MEX	0.03727	0.0019959	18.6733
W-SWF	CBD	NFA	0.0327343	0.00224719	14.5668
W-SWF	CBD	NFE	0.0258933	0.00204318	12.673
W-SWF	CBD	PYR_GAL2	0.00642816	0.0009565	6.72047
W-SWF	ITA	MEX	0.0239655	0.00226464	10.5825
W-SWF	ITA	NFA	0.00201673	0.00148137	1.36139
W-SWF	ITA	NFE	0.00559153	0.00134663	4.15225
W-SWF	ITA	PYR_GAL2	0.0042337	0.00082283	5.14526
W-SWF	MEX	NFA	0.0189435	0.00191014	9.91736
W-SWF	MEX	NFE	0.0142013	0.00175142	8.10844
W-SWF	MEX	PYR_GAL2	0.00552938	0.00087901	6.29047
W-SWF	NFA	NFE	0.058005	0.00308724	18.7886
W-SWF	NFA	PYR_GAL2	0.010041	0.00118796	8.45232
W-SWF	NFE	PYR_GAL2	0.00926961	0.00117668	7.87778

Annexe2 : Tableau supplémentaire de l'analyse cpSSR du chapitre 2

Tableau S 1 : résultat du génotypage avec les marqueurs cpSSR.

Combinaison des données passeport (colonne 1 à 9), des groupes génétiques (colonne 10 et 11) et de l'analyse du génome chloroplastique basée sur 7 marqueurs cpSSR. NF= « Northern flint » ; CAR= « Caribbean » ; AND= « Andean » ; CBD= « Corn Belt Dent » ; ITA= « Italian » ; MEX= « Mexican » ; PG2= « Pyrenees_Galicia2 » ; W-SWF = « West South-West France » ; E-SWF = « East South-West France » ; CAM= « Central America and Mexico » ; NAM= « North America » ; SAM= « South America ».

Code	**ZoneOri**	**PaysOri**	**Accession_label**	**Zea subsp.**	**Provant et al (1999a)**	**Groupes génétiques**	**Haplotypes**	**Marqueurs cpSSR**						
								RCt5	**Z3764**	**Z7430**	**Z12968**	**Z18704**	**Z20824**	**Z59488**
Ames2749	NAM	Usa	Ames2749_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
Ames2750	NAM	Usa	Ames2750_DivCor	cultivé	non	NF	F	174	161	120	167	122	138	201
Ames2752	NAM	Usa	Ames2752_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
Ames2755	NAM	Canada	Ames2755_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
ANTI3	Caribbean	Antigua	ANTI3_DivCor	cultivé	non	CAR	AB	190	163	125	164	125	136	200
APUC140	SAM	Peru	APUC140_DivCor	cultivé	non	AND	AD	190	163	125	165	125	136	200
APUC171	SAM	Peru	APUC171_DivCor	cultivé	non	AND	AD	190	163	125	165	125	136	200
ARGE486	SAM	Argentina	ARGE486_DivCor	cultivé	non	MEX	AD	190	163	125	165	125	136	200
ARGE564	SAM	Argentina	ARGE564_DivCor	cultivé	non	AND	G	174	161	120	167	125	138	201
BARB5	Caribbean	Barbados	BARB5_DivCor	cultivé	non	CAR	G	174	161	120	167	125	138	201
BOLI711	SAM	Bolivia	BOLI711_DivCor	cultivé	non	AND	G	174	161	120	167	125	138	201
BRVI142	Caribbean	British Virgin Islands	BRVI142_DivCor	cultivé	non	CAR	AB	190	163	125	164	125	136	200
CHIH150	CAM	Mexico	CHIH150_DivCor	cultivé	non	MEX	AD	190	163	125	165	125	136	200
CHIS104	CAM	Mexico	CHIS104_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200
CHIS63	CAM	Mexico	CHIS63_DivCor	cultivé	non	CAR	G	174	161	120	167	125	138	201
CRIC166	CAM	Costa Rica	CRIC166_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200
CUBA12	Caribbean	Cuba	CUBA12_DivCor	cultivé	non	MEX	G	174	161	120	167	125	138	201
CUBA44	Caribbean	Cuba	CUBA44_DivCor	cultivé	non	CAR	G	174	161	120	167	125	138	201
CUBA54	Caribbean	Cuba	CUBA54_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200
CUBA63	Caribbean	Cuba	CUBA63_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200
DES103	Caribbean	Guadeloupe	DES103_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200
ECUA617	CAM	Ecuador	ECUA617_DivCor	cultivé	non	AND	AD	190	163	125	165	125	136	200
ECUA696	CAM	Ecuador	ECUA696_DivCor	cultivé	non	AND	AD	190	163	125	165	125	136	200
EPZMV155	Europe	Spain	EPZMV155_DivCor	cultivé	non	CBD	I	174	161	120	168	125	138	201
EPZMV156	Europe	Spain	EPZMV156_DivCor	cultivé	non	CAR	F	174	161	120	167	122	138	201

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
EPZMV159	Europe	Spain	EPZMV159_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
EPZMV168	Europe	Spain	EPZMV168_DivCor	cultivé	non	PG2	AA	190	163	124	165	125	136	200
EPZMV171	Europe	Spain	EPZMV171_DivCor	cultivé	non	PG2	AD	190	163	125	165	125	136	200
EPZMV173	Europe	Spain	EPZMV173_DivCor	cultivé	non	MEX	K	174	161	120	169	125	138	201
EPZMV23	Europe	Spain	EPZMV23_DivCor	cultivé	non	MEX	F	174	161	120	167	122	138	201
EPZMV232	Europe	Spain	EPZMV232_DivCor	cultivé	non	MEX	AA	190	163	124	165	125	136	200
EPZMV310	Europe	Spain	EPZMV310_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
EPZMV337	Europe	Spain	EPZMV337_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200
EPZMV70	Europe	Spain	EPZMV70_DivCor	cultivé	non	PG2	AA	190	163	124	165	125	136	200
GUAT36	CAM	Guatemala	GUAT36_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200
GUAT45	CAM	Guatemala	GUAT45_DivCor	cultivé	non	MEX	AB	190	163	125	164	125	136	200
GUAT527	CAM	Guatemala	GUAT527_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200
IPK15	Europe	Germany	IPK15_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
IPK158	Europe	Germany	IPK158_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
IPK239	Europe	Germany	IPK239_DivCor	cultivé	non	NF	K	174	161	120	169	125	138	201
IPK242	Europe	Germany	IPK242_DivCor	cultivé	non	CAR	I	174	161	120	168	125	138	201
IPK25	Europe	Germany	IPK25_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
IPK26	Europe	Germany	IPK26_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
IPK6	Europe	Germany	IPK6_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
IPK60	Europe	Austria	IPK60_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
IPK9	Europe	Germany	IPK9_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
JALI102	CAM	Mexico	JALI102_DivCor	cultivé	non	MEX	AD	190	163	125	165	125	136	200
JALI71	CAM	Mexico	JALI71_DivCor	cultivé	non	MEX	AD	190	163	125	165	125	136	200
KRO1	Europe	Poland	KRO1_DivCor	cultivé	non	NF	F	174	161	120	167	122	138	201
KRO2	Europe	Poland	KRO2_DivCor	cultivé	non	NF	AD	190	163	125	165	125	136	200
KRO3	Europe	Poland	KRO3_DivCor	cultivé	non	NF	AI	174	161	120	168	122	138	201
MEXI3	CAM	Mexico	MEXI3_DivCor	cultivé	non	MEX	AD	190	163	125	165	125	136	200
MGA235	Caribbean	Guadeloupe	MGA235_DivCor	cultivé	non	CAR	AB	190	163	125	164	125	136	200
MRIZP47	Europe	Yugoslavia	MRIZP47_DivCor	cultivé	non	ITA	G	174	161	120	167	125	138	201
MRIZP676	Europe	Yugoslavia	MRIZP676_DivCor	cultivé	non	NF	AA	190	163	124	165	125	136	200

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
NAYA24	CAM	Mexico	NAYA24_DivCor	cultivé	non	MEX	G	174	161	120	167	125	138	201
NAYA39	CAM	Mexico	NAYA39_DivCor	cultivé	non	MEX	G	174	161	120	167	125	138	201
OAXA50	CAM	Mexico	OAXA50_DivCor	cultivé	non	MEX	AD	190	163	125	165	125	136	200
PERU1303	SAM	Peru	PERU1303_DivCor	cultivé	non	AND	X	190	162	124	165	124	136	200
PI213699	NAM	Usa	PI213699_DivCor	cultivé	non	CBD	G	174	161	120	167	125	138	201
PI213715	NAM	Usa	PI213715_DivCor	cultivé	non	MEX	G	174	161	120	167	125	138	201
PI213719	NAM	Usa	PI213719_DivCor	cultivé	non	CBD	G	174	161	120	167	125	138	201
PI213763	NAM	Usa	PI213763_DivCor	cultivé	non	NF	AA	190	163	124	165	125	136	200
PI213793	NAM	Usa	PI213793_DivCor	cultivé	non	NF	AA	190	163	124	165	125	136	200
PI214189	NAM	Canada	PI214189_DivCor	cultivé	non	CBD	AD	190	163	125	165	125	136	200
PI214295	NAM	Usa	PI214295_DivCor	cultivé	non	CBD	G	174	161	120	167	125	138	201
PI217410	NAM	Usa	PI217410_DivCor	cultivé	non	MEX	Q	174	162	127	168	125	136	201
PI217413	CAM	Mexico	PI217413_DivCor	cultivé	non	MEX	AD	190	163	125	165	125	136	200
PI217460	NAM	Usa	PI217460_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
PI218137	NAM	Usa	PI218137_DivCor	cultivé	non	CBD	AD	190	163	125	165	125	136	200
PI269743	NAM	Usa	PI269743_DivCor	cultivé	non	CBD	G	174	161	120	167	125	138	201
PI280061	NAM	Usa	PI280061_DivCor	cultivé	non	CBD	AD	190	163	125	165	125	136	200
PI311237	NAM	Usa	PI311237_DivCor	cultivé	non	CBD	N	174	161	125	166	125	136	201
PI401755	NAM	Usa	PI401755_DivCor	cultivé	non	NF	AD	190	163	125	165	125	136	200
PI401757	NAM	Usa	PI401757_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
Pionner	Europe	Germany	Pionner_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
PPS1	SW France	Pyrenees Atlantiques	Salies de Bearn 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS10	SW France	France	Estarvielle	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS1012	Europe	France	PPS1012_DivCor	cultivé	non	ITA	I	174	161	120	168	125	138	201
PPS1016	SW France	Ariège	Couserans	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS1031	SW France	Pyrenees Atlantiques	Arrien	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1036	SW France	Hautes Pyrénées	Saint Creac.A	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS1037	SW France	Hautes Pyrénées	Saint Creac.B	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1038	SW France	Hautes Pyrénées	Saint Creac.C	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
PPS1039	SW France	Hautes Pyrénées	Sarrancolin	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS1040	SW France	Pyrenees Atlantiques	Monein 2	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1041	SW France	Hautes Pyrénées	Gazost	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1042	SW France	Pyrenees Atlantiques	Morlaas	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1043	SW France	Hautes Pyrénées	Argeles Gazost	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1044	SW France	Hautes Pyrénées	Pierrefitte	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1045	SW France	Hautes Pyrénées	Bagnères de Bigorre 1	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1046	SW France	Hautes Pyrénées	Arrens	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1047	SW France	Pyrenees Atlantiques	Saint Engrace	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1048	SW France	Pyrenees Atlantiques	Orthez d'Asson.A	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1049	SW France	Pyrenees Atlantiques	Orthez d'Asson.B	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1050	SW France	Hautes Pyrénées	Tarbes	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS1051	SW France	Hautes Pyrénées	Lourdes	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1052	SW France	Hautes Pyrénées	Bagnères de Bigorre 2	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1053	SW France	Hautes Pyrénées	Ourdes	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1054	SW France	Pyrenees Atlantiques	Monein 3	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS1057	SW France	Pyrenees Atlantiques	Lestelle Betharram	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1058	SW France	Pyrenees Atlantiques	Gurmencon	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS11	SW France	France	Sost	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS1132	NAM	Usa	PPS1132_DivCor	cultivé	non	CBD	J	174	161	120	168	125	138	202
PPS115	Europe	Italy	PPS115_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
PPS1153	SW France	Hautes Pyrénées	Col du Soulor	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS1154	SW France	Pyrenees Atlantiques	Oloron Sainte Marie	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS116	Europe	Italy	PPS116_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
PPS12	SW France	Haute Garonne	Portet	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS121	SW France	Dordogne	Beaumont	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS122	SW France	Aude	Millette du Lauragais 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
PPS1234	Europe	Portugal	PPS1234_DivCor	cultivé	non	PG2	I	174	161	120	168	125	138	201
PPS1236	SW France	Pyrenees Atlantiques	Coarraze	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS1248	SW France	Landes	Pop Cornée.1	cultivé	non	E-SWF	D	174	161	120	166	124	138	201
PPS1249	SW France	Landes	Pop Cornée.2	cultivé	non	E-SWF	D	174	161	120	166	124	138	201
PPS1250	SW France	Landes	Le Muret.1	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS1251	SW France	Landes	Le Muret.2	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS1252	SW France	Tarn	Isle sur Tarn	cultivé	non	E-SWF	G	174	161	120	167	125	138	201
PPS13	SW France	Ariège	Aleu	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS14	SW France	France	Massat	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS140	Europe	Hungary	PPS140_DivCor	cultivé	non	CBD	G	174	161	120	167	125	138	201
PPS1421	SW France	Hautes Pyrénées	Pouyastruc	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS15	SW France	France	Lacaune	cultivé	non	E-SWF	I	174	161	120	168	125	138	201
PPS151	Europe	Hungary	PPS151_DivCor	cultivé	non	NF	G	174	161	120	167	125	138	201
PPS160	Europe	Romania	PPS160_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
PPS17	Europe	France	PPS17_DivCor	cultivé	non	PG2	AA	190	163	124	165	125	136	200
PPS187	Europe	Italy	PPS187_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
PPS189	SW France	Tarn	Saint Lieux les Lavaur	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS19	Europe	France	PPS19_DivCor	cultivé	non	NF	AD	190	163	125	165	125	136	200
PPS191	SW France	Ariège	Aston D	cultivé	non	E-SWF	N	174	161	125	166	125	136	201
PPS192	SW France	Tarn	Cambon Le Lavaur	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS193	SW France	Haute Garonne	Aurignac	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS194	SW France	Pyrenees Atlantiques	Grand Roux Basque 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS196	SW France	Ariège	Moulin Neuf	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS198	SW France	Lot et Garonne	Marmande.36	cultivé	non	E-SWF	F	174	161	120	167	122	138	201
PPS199	SW France	Landes	Parleboscq	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS2	SW France	Pyrenees Atlantiques	Saint Jean Pied de Port 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS20	SW France	Hautes Pyrénées	Campan Galade	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS21	SW France	Hautes Pyrénées	Sainte Marie de Campan	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS22	SW France	France	Esterre	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
PPS226	SW France	Lot et Garonne	Marmande.10	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS227	SW France	Lot et Garonne	Marmande.21	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS228	SW France	Lot et Garonne	Marmande.26	cultivé	non	E-SWF	I	174	161	120	168	125	138	201
PPS229	SW France	Lot et Garonne	Marmande.34	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS23	SW France	France	Laruns	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS230	SW France	Haute Garonne	Sainte Livrade	cultivé	non	E-SWF	I	174	161	120	168	125	138	201
PPS235	SW France	Pyrenees Atlantiques	Grand Roux Basque 2	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS237	SW France	Ariège	Seix	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS238	SW France	Ariège	Riverenert 2	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS24	SW France	Hautes Pyrénées	Viey	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS240	SW France	Hautes Pyrénées	Saint Sauveur	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS241	SW France	Tarn et Garonne	Millette de Finham.AGPM	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS243	SW France	Lot	Floirac Camboni B	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS25	SW France	Hautes Pyrénées	Bize	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS26	SW France	Ariège	Argein	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS27	SW France	Aveyron	Saint Christophe	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS35	SW France	Gironde	Gironde 1	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS361	SW France	Haute Garonne	Cassagnabere 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS362	SW France	Lot et Garonne	Marmande.35	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS363	SW France	Landes	Landes 2	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS381	Europe	Italy	PPS381_DivCor	cultivé	non	ITA	I	174	161	120	168	125	138	201
PPS410	SW France	Gironde	Gironde 2	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS442	SW France	Gironde	Gironde 3	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS444	Europe	Spain	PPS444_DivCor	cultivé	non	MEX	AA	190	163	124	165	125	136	200
PPS448	Europe	Poland	PPS448_DivCor	cultivé	non	NF	K	174	161	120	169	125	138	201
PPS458	Europe	Hungary	PPS458_DivCor	cultivé	non	NF	AD	190	163	125	165	125	136	200
PPS460	SW France	Gironde	Gironde 4	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS462	Europe	Czechoslovakia	PPS462_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
PPS471	NAM	Usa	PPS471_DivCor	cultivé	non	CBD	G	174	161	120	167	125	138	201

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
PPS474	Europe	France	PPS474_DivCor	cultivé	non	NF	F	174	161	120	167	122	138	201
PPS478	SW France	Ariège	Labastide de Boussignac	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS482	SW France	Haute Garonne	Cuguron	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS483	SW France	Gironde	Gironde 5	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS484	SW France	Gironde	Gironde 6	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS485	SW France	Gironde	Gironde 7	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS486	SW France	Gironde	Gironde 8	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS487	SW France	Gironde	Cazalis	cultivé	non	E-SWF	G	174	161	120	167	125	138	201
PPS488	SW France	Pyrenees Atlantiques	Estaut 1	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS490	SW France	Hautes Pyrénées	Bugard	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS491	SW France	Hautes Pyrénées	Saint Laurent de Neste 3	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS498	SW France	Pyrenees Atlantiques	Estaut 2	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS5	Europe	France	PPS5_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
PPS501	SW France	Tarn	Borie Lautrec	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS502	SW France	Tarn	Lautrec	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS503	SW France	Tarn	Semport	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS504	SW France	Tarn	Villefranche d Albigeois.B	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS505	SW France	Tarn	Pampelone	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS506	SW France	Tarn	Poulinet Alban.A	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS525	Europe	Bulgaria	PPS525_DivCor	cultivé	non	CBD	AD	190	163	125	165	125	136	200
PPS566	SW France	Haute Garonne	Boutx A	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS567	SW France	Haute Garonne	Boutx B	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS568	SW France	Haute Garonne	Moustajon	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS570	SW France	Gers	Moncassin	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS574	SW France	Tarn	Fraysse Alban	cultivé	non	E-SWF	I	174	161	120	168	125	138	201
PPS575	SW France	Tarn	Cabrilles	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS58	Europe	Austria	PPS58_DivCor	cultivé	non	NF	K	174	161	120	169	125	138	201
PPS585	SW France	Gironde	Gironde 9	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS59	SW France	Tarn	Montirat Monesties	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
PPS597	SW France	Ariège	Seix.Bardot	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS598	SW France	Hautes Pyrénées	Bugard A	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS599	SW France	Hautes Pyrénées	Burg	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS60	SW France	Tarn	Poulinet Alban.B	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS600	SW France	Hautes Pyrénées	Serre Rustaing	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS601	SW France	Hautes Pyrénées	Nestier 2	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS602	SW France	Hautes Pyrénées	Saint Laurent de Neste 4	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS603	SW France	Hautes Pyrénées	Grust	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS604	SW France	Hautes Pyrénées	Lannemezan	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS607	SW France	Pyrenees Atlantiques	Doré de Bordes	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS608	SW France	Pyrenees Atlantiques	Doré du Pays	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS609	SW France	Pyrenees Atlantiques	Grand Roux Basque 3	cultivé	non	E-SWF	I	174	161	120	168	125	138	201
PPS61	SW France	Tarn	Vaour	cultivé	non	E-SWF	G	174	161	120	167	125	138	201
PPS610	SW France	Pyrenees Atlantiques	Grand Roux Basque 4	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS611	SW France	Pyrenees Atlantiques	Saint Jean Pied de Port 2	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS613	SW France	Aude	Millette La Cassignole	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS614	SW France	Aveyron	Espeyrac	cultivé	non	E-SWF	I	174	161	120	168	125	138	201
PPS615	SW France	Aveyron	Saint Come	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS616	SW France	Aveyron	Saint Parthem.A	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS617	SW France	Aveyron	Saint Parthem.B	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS621	SW France	Gers	Pop Gers	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS622	SW France	Landes	Roux de Chalosse	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS623	SW France	Landes	CETA des Landes	cultivé	non	W-SWF	I	174	161	120	168	125	138	201
PPS624	SW France	Landes	Landes 3	cultivé	non	E-SWF	F	174	161	120	167	122	138	201
PPS625	SW France	Landes	Landes 4	cultivé	non	E-SWF	F	174	161	120	167	122	138	201
PPS626	SW France	Gironde	Cugnieres	cultivé	non	E-SWF	G	174	161	120	167	125	138	201
PPS627	SW France	Landes	Pessan	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS628	SW France	Landes	Dubos	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
PPS630	SW France	Landes	Quarantaun	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS631	SW France	France	Saint Pantaleon	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS632	SW France	Lot	Cressensac 2	cultivé	non	E-SWF	F	174	161	120	167	122	138	201
PPS633	SW France	Lot et Garonne	Marmande.12	cultivé	non	E-SWF	N	174	161	125	166	125	136	201
PPS634	SW France	Lot et Garonne	Marmande.45	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS636	SW France	Lot et Garonne	Marmande.19	cultivé	non	E-SWF	N	174	161	125	166	125	136	201
PPS637	SW France	Lot et Garonne	Marmande.20	cultivé	non	E-SWF	N	174	161	125	166	125	136	201
PPS638	SW France	Lot et Garonne	Marmande.32	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS639	SW France	France	Millette du Lauragais 2	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS640	SW France	Pyrenees Atlantiques	Doré du Pays Louvrigny	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS641	SW France	France	Grand Roux Basque 5	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS642	SW France	Pyrenees Atlantiques	Roux du Pays	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS643	SW France	Pyrenees Atlantiques	Salies du Bearn 2	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS644	SW France	Pyrenees Atlantiques	Sainte Suzanne	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS645	SW France	Pyrenees Atlantiques	Bordes	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS646	SW France	Pyrenees Atlantiques	Barzin	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS647	SW France	Pyrenees Atlantiques	Lucgarier 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS648	SW France	Pyrenees Atlantiques	Lucgarier 2	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS649	SW France	Pyrenees Atlantiques	Gomer	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS65	SW France	Ariège	Sainte Croix	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS650	SW France	Pyrenees Atlantiques	Auga 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS651	SW France	Pyrenees Atlantiques	Auga 2	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS652	SW France	Pyrenees Atlantiques	Mouhoux	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS653	SW France	Pyrenees Atlantiques	Saint Jean Pied de Port 3	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS66	SW France	Ariège	Cazavet	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
PPS666	SW France	Tarn	Villefranche d Albigeois.A	cultivé	non	E-SWF	I	174	161	120	168	125	138	201
PPS668	SW France	France	Millette Montagne Noire	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS67	SW France	Ariège	Prat Bonrepaux	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS68	SW France	Ariège	Riverenert 1	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS686	Europe	Italy	PPS686_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
PPS687	Europe	Italy	PPS687_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
PPS69	SW France	Ariège	Couflens.2	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS70	SW France	Ariège	Aston A	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS71	SW France	Ariège	Aston B	cultivé	non	E-SWF	G	174	161	120	167	125	138	201
PPS72	SW France	Hautes Pyrénées	Vidou 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS73	SW France	Hautes Pyrénées	Begole	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS735	Europe	Italy	PPS735_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
PPS736	SW France	Ariège	Ariege 1	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS737	SW France	Ariège	Couflens.1	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS76	SW France	Hautes Pyrénées	Nestier 1	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS77	SW France	Hautes Pyrénées	Saint Laurent de Neste 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS770	Europe	Spain	PPS770_DivCor	cultivé	non	MEX	F	174	161	120	167	122	138	201
PPS78	SW France	Hautes Pyrénées	Saint Laurent de Neste 2	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS79	SW France	Hautes Pyrénées	Aspin ez Angles	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS8	SW France	Hautes Pyrénées	Salechan	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS80	SW France	Hautes Pyrénées	Esquieze Serre	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS81	SW France	Hautes Pyrénées	Luz	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS818	SW France	Pyrenees Atlantiques	Lucq de Bearn	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS82	SW France	Hautes Pyrénées	Bazus Aure	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS820	SW France	Hautes Pyrénées	Blanche Htes.Pyr.C1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS821	SW France	Pyrenees Atlantiques	Val du Gave d’Aspe	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS822	SW France	Pyrenees Atlantiques	Monein C1	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS823	SW France	Pyrenees Atlantiques	Lourdios.C1	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
PPS83	SW France	Hautes Pyrénées	Guchan	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS84	SW France	Haute Garonne	Sauveterre de Comminges 1	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS85	SW France	Lot	Floirac Camboni A	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS86	SW France	Lot	Floirac	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS87	SW France	Lot	Cressensac 1	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS9	SW France	Hautes Pyrénées	Bareilles	cultivé	non	W-SWF	F	174	161	120	167	122	138	201
PPS920	SAM	Argentina	PPS920_DivCor	cultivé	non	NF	AD	190	163	125	165	125	136	200
PPS938	SAM	Chile	PPS938_DivCor	cultivé	non	PG2	AD	190	163	125	165	125	136	200
PPS94	Europe	Switzerland	PPS94_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
PPS941	SAM	Chile	PPS941_DivCor	cultivé	non	CBD	I	174	161	120	168	125	138	201
PPS949	SAM	Chile	PPS949_DivCor	cultivé	non	AND	P	174	162	127	167	125	136	201
PPS96	SW France	Lot et Garonne	Marmande.31	cultivé	non	E-SWF	AD	190	163	125	165	125	136	200
PPS961	SAM	Chile	PPS961_DivCor	cultivé	non	NF	AD	190	163	125	165	125	136	200
PPS962	SW France	Pyrenees Atlantiques	Asson	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS963	SW France	Haute Garonne	Val d'Aran	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS964	SW France	Hautes Pyrénées	Argeles	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS965	SW France	Pyrenees Atlantiques	Barzun	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS966	SW France	Ariège	Ariege 3	cultivé	non	E-SWF	AA	190	163	124	165	125	136	200
PPS967	SW France	Pyrenees Atlantiques	Pontacq	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS968	SW France	Pyrenees Atlantiques	Urdos	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS969	SW France	Pyrenees Atlantiques	Monein 1	cultivé	non	W-SWF	AA	190	163	124	165	125	136	200
PPS979	Europe	Spain	PPS979_DivCor	cultivé	non	PG2	AD	190	163	125	165	125	136	200
PPS98	SW France	Landes	Landes 1	cultivé	non	W-SWF	AD	190	163	125	165	125	136	200
PPS983	Europe	Spain	PPS983_DivCor	cultivé	non	PG2	D	174	161	120	166	124	138	201
PPS986	Europe	Spain	PPS986_DivCor	cultivé	non	PG2	AD	190	163	125	165	125	136	200
PUEB70	CAM	Mexico	PUEB70_DivCor	cultivé	non	AND	AB	190	163	125	164	125	136	200
PUEB91	CAM	Mexico	PUEB91_DivCor	cultivé	non	MEX	AD	190	163	125	165	125	136	200
RDOM270	Caribbean	Dominican Republic	RDOM270_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200

Code	ZoneOri	PaysOri	Accession_label	Zea subsp.	Provant et al (1999a)	Groupes génétiques	Haplotypes	Marqueurs cpSSR						
								RCt5	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488
RYS2	Europe	Czechoslovakia	RYS2_DivCor	cultivé	non	CBD	G	174	161	120	167	125	138	201
SINA2	CAM	Mexico	SINA2_DivCor	cultivé	non	MEX	AD	190	163	125	165	125	136	200
STB073	Caribbean	Guadeloupe	STB073_DivCor	cultivé	non	CAR	AB	190	163	125	164	125	136	200
SVIN5	Caribbean	Saint Vincent	SVIN5_DivCor	cultivé	non	CAR	N	174	161	125	166	125	136	201
TRIN6	Caribbean	Trinite	TRIN6_DivCor	cultivé	non	CAR	G	174	161	120	167	125	138	201
UKR37	Europe	Ukraine	UKR37_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
URUG8A	SAM	Uruguay	URUG8A_DivCor	cultivé	non	CAR	AD	190	163	125	165	125	136	200
VA136	Europe	Italy	VA136_DivCor	cultivé	non	NF	I	174	161	120	168	125	138	201
VA145	Europe	Italy	VA145_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
VA469	Europe	Italy	VA469_DivCor	cultivé	non	ITA	P	174	162	127	167	125	136	201
VA563	Europe	Italy	VA563_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
VA565	Europe	Italy	VA565_DivCor	cultivé	non	ITA	F	174	161	120	167	122	138	201
MX02			69-12	spontané	non		AD	190	163	125	165	125	136	200
MX06			JSG-318	spontané	non		AD	190	163	125	165	125	136	200
MX12			WST92-4	spontané	non		X	190	162	124	165	124	136	200
MX15			WST92-2	spontané	non		X	190	162	124	165	124	136	200
MP01			W71-2	spontané	non		X	190	162	124	165	124	136	200
MP02			K67-7	spontané	non		MP01	190	163	124	165	124	136	200
MP03			K67-6	spontané	non		X	190	162	124	165	124	136	200
MP04			K67-15	spontané	non		MP04	190	161	125	165	125	136	200
MP06			K67-20	spontané	non		MP06	190	162	125	165	124	136	200
MP12			K69-14	spontané	non		MP06	190	162	125	165	124	136	200
MP13			WST85-3	spontané	non		X	190	162	124	165	124	136	200
MH01			Teo hueue	spontané	oui		MH01	190	161	121	165	121	136	200
ZL01			Teo Guatemala	spontané	non		ZL01	174	160	121	162	120	136	199
ZD01			San Miguel	spontané	oui		ZD01	174	160	120	162	120	136	199
ZP01			Piedra ancha	spontané	oui		ZP01	174	160	120	163	120	136	199

Tableau S 2 : Analyse cpSSR sur les 37 accessions des espèce et sous-espèces du genre Zea de Provan et al (1999a) basé sur notre base de données.

Zone d'origine	accession no.	Provan et al (1999a) : données réelles						Provan et al (1999a) : données modifiées en 2020						Type d'Haplotypes
		Z3764	Z7430	Z12968	Z18704	Z20824	Z59488	Z3764	Z7430	Z12968	Z18704	Z20824	Z59488	
Antigua	1241	144	84	147	107	119	172	163	124	164	125	136	200	A12
Austria	2571	144	84	148	107	119	172	163	124	165	125	136	200	AA2
Bolibia	10222	142	80	149	106	120	173	161	120	166	124	137	201	A4
Bolivia	6268	142	80	150	108	120	173	161	120	167	126	137	201	A1
Bolivia	7861	144	84	148	108	120	172	163	124	165	126	137	200	A13
Colombia	3105	143	84	148	107	119	172	162	124	165	125	136	200	A9
Colombia	3112	142	80	150	107	120	173	161	120	167	125	137	201	A2
Colombia	3130	142	80	151	108	120	173	161	120	168	126	137	201	A3
Colombia	3213	142	80	150	108	120	173	161	120	167	126	137	201	A1
Ecuador	8226	142	80	150	108	120	173	161	120	167	126	137	201	A1
Ecuador	8643	142	80	150	108	120	173	161	120	167	126	137	201	A1
Guatemala	5183	144	84	148	108	119	172	163	124	165	126	136	200	A11
Guatemala	5253	142	80	145	108	120	173	161	120	162	126	137	201	A6
Mexico	467	144	80	148	107	120	173	163	124	165	125	136	200	AA2
Mexico	815	144	87	147	107	119	172	163	124	165	125	136	200	AA2
Mexico	2232	144	84	148	107	119	172	163	124	165	125	136	200	AA2
Mexico	2233	144	84	148	107	119	172	161	120	167	125	137	201	A2
Mexico	2246	144	84	148	107	119	172	161	120	167	125	137	201	A2
Mexico	2250	142	80	150	107	120	173	163	120	165	125	137	201	A15
Mexico	2270	142	80	150	107	120	173	163	127	164	125	136	200	A14
Peru	8381	143	83	148	107	119	172	161	120	164	125	137	201	A5
Peru	10381	142	80	147	107	120	173	162	123	162	124	136	200	MP02
Peru	10430	143	83	145	106	119	172	162	123	165	125	136	200	A8
Uruguay	4276	144	84	148	108	119	172	163	124	165	126	136	200	A11
USA	BIRD011	142	80	147	107	120	173	161	120	164	125	137	201	A5
Venezuela	9059	142	80	150	108	120	173	161	120	167	126	137	201	A1
Venezuela	9063	143	86	150	108	119	173	162	126	167	126	136	201	A10
Z.perennis	**9475**	**141**	**80**	**146**	**102**	**119**	**171**	**160**	**120**	**163**	**120**	**136**	**199**	**ZP01**
Z.diploperennis	**9476**	**141**	**80**	**145**	**102**	**119**	**171**	**160**	**120**	**162**	**120**	**136**	**199**	**ZD01**
Z.luxurians	11083	141	80	142	102	119	171	160	120	159	120	136	199	ZL01 ou ROOT
Z.mays_huehueteangenis	**9479**	**143**	**80**	**147**	**107**	**119**	**172**	**161**	**121**	**165**	**121**	**136**	**200**	**MH01**
Z.mays_parviglumis1	11407	143	83	145	106	119	172	162	123	162	124	136	200	MP02
Z.mays_parviglumis2	11387	143	83	145	106	119	172	162	123	162	124	136	200	MP02
Z.mays_parviglumis3	8755	143	83	147	107	119	172	162	123	164	125	136	200	MP03
Z.mays_mexicana1	11392	144	83	145	106	119	172	163	123	162	124	136	200	MX1
Z.mays_mexicana2	11396	143	83	144	106	119	172	162	123	161	124	136	200	MX2
Z.mays_mexicana3	12823	143	83	145	106	119	173	162	123	162	124	136	201	MX3

Tableau S 3 : différences en nombres de bases des tailles d'allèles observés chez 3 téosintes analysées avec sur 6 cpSSR entre séquençage de Provan et al (1999a) et séquençage de 2017_2020.

Loci cpSSR	**taille des alléles (en nombre de base)**		**différence (en nombre de base)**
	séquençage Provan et al (1999a)	séquençage 2017_2020	
Z3764	141	160	19
	142	161	19
Z7430	80	120	40
Z12968	146	163	17
	145	162	17
	147	165	18
Z18704	102	120	18
	104	121	17
Z20824	119	136	17
Z59488	171	199	28
	172	200	28

Annexe 3: tableaux supplémentaires chapitre 3.

Suppl Table 1: Heritability values for morphological traits of ear.

Additive, dominance and residual components of variance with the standard deviations for the ten morphological traits of ear. a2 the ratio of additive variance on genetic variance, h2 the narrow sense heritability. ** significant at 1%, * significant at 5%.

Trait	Landrace	var_add	std_add	var_dom	std_dom	var_res	std_res	a2	h2
E_length	PPS68	97.87	81.19	74.70	136.29	342.41	73.76	0.57	0.190
E_length	PPS70	210.31**	53.97	0.00	0.00	178.70	30.13	1.00	0.541
E_length	PPS653	161.83	108.12	73.34	166.05	98.06	82.89	0.69	0.486
E_length	PPS1052	318.24**	79.29	0.00	0.00	148.63	42.34	1.00	0.682
E_diam	PPS68	3.44*	0.91	0.00	0.00	3.90	0.53	1.00	0.469
E_diam	PPS70	5.76*	3.89	3.36	6.02	1.83	3.00	0.63	0.526
E_diam	PPS653	35.62**	9872.26	30.00	8229.88	0.01	3.74	0.54	0.543
E_diam	PPS1052	7.46	1.84	0.00	0.00	2.77	0.97	1.00	0.729
conicity	PPS68	0.35	0.34	0.52	0.60	0.85	0.32	0.40	0.205
conicity	PPS70	0.69**	0.20	0.00	0.00	1.49	0.13	1.00	0.318
conicity	PPS653	0.83	0.25	0.00	0.00	1.79	0.17	1.00	0.317
conicity	PPS1052	0.85*	0.23	0.00	0.00	1.16	0.14	1.00	0.424
ear_top	PPS68	0.00	0.00	84.54**	23.07	42.23	18.13	0.00	0.000
ear_top	PPS70	5.86	6.37	10.14	11.50	24.01	6.31	0.37	0.147
ear_top	PPS653	2.92	4.33	1.31	8.44	61.25	5.82	0.69	0.045
ear_top	PPS1052	8.20	9.41	13.06	17.18	51.62	9.72	0.39	0.113
N_row	PPS68	0.51	0.69	1.53	1.31	1.32	0.71	0.25	0.152
N_row	PPS70	1.47	1.25	1.76	2.12	1.60	1.11	0.46	0.305
N_row	PPS653	0.67	0.43	0.09	0.65	1.23	0.33	0.89	0.338
N_row	PPS1052	1.44**	0.37	0.00	0.00	1.27	0.21	1.00	0.533
NKR	PPS68	0.58	2.84	8.40	6.05	11.50	3.48	0.06	0.028

Trait	Landrace	var_add	std_add	var_dom	std_dom	var_res	std_res	a2	h2
NKR	PPS70	7.76	4.67	1.26	6.82	8.17	3.32	0.86	0.452
NKR	PPS653	3.42	3.82	7.09	6.88	4.60	3.68	0.33	0.226
NKR	PPS1052	4.99	1.37	0.00	0.00	9.09	0.88	1.00	0.354
K_height	PPS68	0.13	0.04	0.00	0.00	0.28	0.02	1.00	0.316
K_height	PPS70	0.18	0.13	0.13	0.21	0.09	0.10	0.57	0.446
K_height	PPS653	0.29	0.18	0.09	0.28	0.29	0.14	0.77	0.432
K_height	PPS1052	0.25	0.16	0.07	0.25	0.42	0.13	0.77	0.332
volume	PPS68	578.62	165.20	0.00	0.00	1294.45	112.04	1.00	0.309
volume	PPS70	1328.81	748.39	97.29	1056.12	759.85	498.89	0.93	0.608
volume	PPS653	920.19	912.90	1154.78	1061.07	20.99	668.22	0.44	0.439
volume	PPS1052	1939.87	478.62	0.00	0.00	621.87	251.47	1.00	0.757
K_number	PPS68	405.17	832.53	2020.97	1668.95	3097.23	942.95	0.17	0.073
K_number	PPS70	2057.87	1267.64	281.36	1876.28	3537.64	936.27	0.88	0.350
K_number	PPS653	247.77	248.24	357.71	440.85	640.70	238.37	0.41	0.199
K_number	PPS1052	893.31	231.63	0.00	0.00	879.33	132.20	1.00	0.504
K_weigth	PPS68	87.75	27.66	0.00	0.00	299.10	21.61	1.00	0.227
K_weigth	PPS70	210.61	117.09	0.19	163.96	181.64	77.50	1.00	0.537
K_weigth	PPS653	25.75	38.98	89.28	75.29	66.89	41.45	0.22	0.142
K_weigth	PPS1052	134.70	34.75	0.00	0.00	117.75	19.53	1.00	0.534

Suppl Table 2: P values of the ANOVA analysis performed on standardized selection differentials.

Factor	DF	conicity	E_diam	E_length	ear_top	K_heigth	K_number	K_weigth	N_row	NKR	volume
Zone	1	0.353	0.178	0.134	0.442	0.282	0.726	0.165	0.601	0.155	0.141
Informant/Zone	7	0.99	0.003	0.322	0.837	0.398	0.343	0.067	0.248	0.173	0.058
Landrace	3	0.448	0	0.772	0.542	0.062	0.096	0.041	0.141	0.006	0.056

Suppl Table 3 : Univariable expected response to selection.

These values were estimated in percent for the 9 morpholocial ear traits for the 17 tests of ears.

Landrace	Informant	conicity	E_diam	K_height	E_length	K_number	NKR	N_row	K_weight	volume
PPS68	Aleu	-1.9	0.8	0.1	1.7	0.9	0.3	0.4	3.0	3.9
PPS68	Bun	-1.8	0.7	-0.8	1.2	0.7	0.3	0.0	2.6	3.0
PPS68	Campan 1	-1.2	1.5	-0.7	1.3	1.0	0.3	0.6	3.4	4.3
PPS68	Campan 2	0.3	1.3	-0.7	1.0	0.7	0.2	0.1	2.8	3.4
PPS68	Engomer	-1.2	1.0	-0.8	1.3	0.8	0.3	0.1	3.1	3.6
PPS68	Esquieze Serre	-2.4	1.1	0.2	1.7	1.0	0.3	0.6	3.4	4.4
PPS70	Engomer	-2.6	2.0	0.9	4.2	2.9	2.8	0.6	7.1	9.7
PPS70	Oursbellile	-0.8	2.2	0.0	4.2	4.0	3.5	1.1	7.5	10.3
PPS653	Campan 1	0.1	0.4	0.0	1.6	0.8	0.7	0.7	0.6	2.3
PPS653	Campan 2	-2.4	0.1	0.1	2.0	0.8	0.9	0.2	0.6	2.0
PPS1052	Aleu	-3.3	2.0	0.2	3.9	3.5	1.7	1.8	4.8	8.6
PPS1052	Beaucens	-2.0	1.4	-1.0	2.8	4.0	2.7	0.8	4.5	6.0
PPS1052	Bun	-2.3	2.4	0.3	5.5	5.1	2.5	2.9	6.6	11.6
PPS1052	Engomer	-2.3	2.3	0.4	7.5	5.7	3.3	2.2	7.5	13.7
PPS1052	Esquieze Serre	2.4	2.7	0.4	5.5	4.6	2.3	2.5	6.4	12.3
PPS1052	Moulin-Neuf	-2.4	2.2	-0.4	6.0	5.7	3.6	1.5	7.0	11.4
PPS1052	Oursbellile	-1.4	2.9	0.0	5.7	6.0	3.0	3.2	7.6	13.1

Résumé

Les populations locales ou « landraces » sont, par leur diversité, des ressources génétiques fondamentales pour faire face aux besoins de l'agriculture. Dans le Sud-Ouest de la France, zone traditionnelle de culture du maïs, les populations de maïs ont été cultivées jusque dans les années 60. Une collection a été constituée entre 1950 et 1980 et conservée *ex situ*. L'objectif de cette thèse est d'étudier leur diversité génétique. Dans un premier temps, nous avons combiné une étude du génome nucléaire de ces populations avec la puce 50K SNP en approche « bulk » d'ADN et une étude du génome chloroplastique avec des marqueurs microsatellites chloroplastiques en plante individuelle, pour analyser la diversité génétique de ces populations de maïs et de déterminer leurs liens génétiques avec les populations de maïs d'Amérique et d'Europe. Dans un deuxième temps, nous avons combiné une enquête ethnobotanique et une analyse génétique quantitative pour comprendre comment les populations de maïs ont évolué dans le Sud-Ouest de la France sous l'effet des pratiques de gestion des anciens agriculteurs des Pyrénées françaises. Ces approches complémentaires de la thèse montrent que les populations du Sud-Ouest de la France sont un groupe original de maïs originaire d'une hybridation entre populations de maïs des régions tropicales et des populations « Northern Flint » (NF), qui sont respectivement les parents femelle et male. Elles auraient hérité une part importante du génome nucléaire des populations NF comme identifiée avec les données nucléaires, les rendant ainsi plus précoces que leur parent maternel issus du maïs des régions tropicales. Nous avons identifié deux groupes génétiques faiblement différenciés au sein de la collection, un dans le partie Ouest et l'autre dans la partie Est. Nous avons mis en évidence l'existence d'un gradient longitudinal le long de la région qui a été probablement maintenu par l'action combinée de la dérive génétique et de flux géniques restreints plutôt que par l'action du climat. La zone de contact entre les 2 groupes se trouve près de la vallée de la Garonne. Nous avons proposé aux anciens agriculteurs enquêtés de choisir 15 épis parmi 150. Les différentiels de sélection calculés par ce test ne montrent pas de différences significatives entre les zones Ouest et Est des Pyrénées. Les anciens agriculteurs des 2 zones sélectionnaient des épis longs, volumineux, bien remplis et sains. Les populations de maïs du Sud-Ouest de la France sont issues de ces pratiques de sélection.

Mots clés : Zea mays L., populations locales, Sud-Ouest de la France, ethnobotanique, puce 50K SNP, microsatellite chloroplastique.

Abstract

Local populations or "landraces" are, by their diversity, fundamental genetic resources to cope the needs of agriculture. In south-western France, a traditional growing area for maize, landraces were cultivated until the 1960s. A collection was built up between 1950 and 1980 and conserved *ex situ*. The aim of this thesis is to study their genetic diversity. In a first step, we combined a study of the nuclear genome using bulk of DNA with the 50K SNP array and a study of the chloroplastic genome using chloroplastic microsatellite markers on one individual plant per population, to analyze the genetic diversity of these landraces and to determine their relationship with maize landraces from America and Europe. In a second step, we combined an ethnobotanical survey with quantitative genetic analysis to understand how maize landraces have evolved in south-western France under the management of old farmers in the French Pyrenees. These complementary approaches using in this thesis show that landraces of south-western France are an original group of maize originating from a hybridization between maize landraces from tropical regions and "Northern Flint" (NF), which are respectively the female and male parents. They are believed to have inherited a significant part of the nuclear genome of the NF populations as identified with the nuclear data, making them earlier than their maternal parent from tropical landraces. We have identified two weakly differentiated genetic groups within the collection, one in the western part and the other in the eastern part. We highlighted the existence of a longitudinal gradient along the region that was probably maintained by interplay effect between genetic drift and restricted gene flow rather than by impact of climate. The contact zone between the 2 groups is located near the Garonne valley. We proposed to the former farmers surveyed to choose 15 ears among 150. The selection differentials calculated by this test do not show significant differences between the western and eastern part of the Pyrenees. The former farmers in both zones selected long, bulky, well-filled and healthy ears. The maize populations of South-West France are the result of these selection practices.

Keywords : *Zea mays L.*, landraces, South-Western France, ethnobotany, 50K SNP array, chloroplastic microsatellite.