ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

AgroParisTech
INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT
PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

Ecole Doctorale ABIES
Agriculture
Agriculture
Alimentation
Food
Biologie
Biology
Environnement
Environment
Santé
Health

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

L'Institut des Sciences et Industries du Vivant et de l'Environnement

(AgroParisTech)

Spécialité : biomathématiques

Présentée et soutenue publiquement par

Mohamed KILEH WAIS

Le 6 Septembre 2012

Méthodes statistiques pour la détection de QTL : nouveaux développements et application chez le canard mulard.

Directeur de thèse : **Jean-Michel ELSEN**
Co-encadrement de thèse : **Christel MARIE-ETANCELIN**

Jury

M. Etienne VERRIER, Professeur, AgroParisTech, Paris	Président
M. Juan José ARRANZ SANTOS, Professeur, Faculté Vétérinaire-Université de Leon	Rapporteur
Mme Elisabeth LE BIHAN-DUVAL, Directrice de Recherche, INRA-URA	Rapporteur
Mme Catherine LARZUL, Chargée de Recherche, INRA-GABI	Examinatrice
Mme Christel MARIE-ETANCELIN, Ingénieur de recherche, INRA-SAGA	Examinatrice
M. Jean-Michel ELSEN, Directeur de recherche, INRA-SAGA	Examinateur

AgroParisTech
INRA-SAGA
Auzeville B.P. 52627 31326 CASTANET TOLOSAN CEDEX

“Nin la sugayow maxad sugi”

“Sirwal ad lahayn dabadada ma qariyo”

“Malintuba wa cimri”

Remerciements

Je tiens à saluer ici et à remercier très vivement les personnes qui, de près ou de loin, ont contribué à la concrétisation de ce travail de thèse de doctorat.

Merci à l'université de Djibouti et au département GA de l'INRA pour le financement.

Merci à **Edouardo Manfredi, Alain Ducos et Christèle Robert-Granié**, directeurs successifs pendant mes quatre années passés à la SAGA, pour les moyens mis à ma disposition.

Je tiens à remercier les membres du jury, en particulier **Etienne VERRIER** qui m'a fait l'honneur d'être le Président du Jury, **Elisabeth LeBihan-Duval** et **Juan José Arranz** pour avoir accepté d'être rapporteurs. Je remercie **Cathérine Larzul** pour avoir accepté d'être examinatrice. Merci pour la relecture de la thèse et pour les améliorations que vous aviez apporté à celle-ci.

Je tiens à remercier très chaleureusement mon directeur de cette thèse, **Jean-Michel Elsen**, pour m'avoir fait confiance, puis pour m'avoir guidé, encouragé, conseillé tout en me laissant une grande liberté. Merci pour le sens de l'humour, toujours une blague pour détendre l'atmosphère à chaque fois où c'était utile. Merci pour le soutien morale à chaque fois que j'en avais besoin et pour la disponibilité sans faille.

Mes remerciements vont également à **Christel Marie-Etancelin**, ma co-encadrante. Merci pour ta gentillesse et ta patience à mon égard durant cette thèse, ainsi que pour tous les conseils et encouragements.

Je tiens à souligner que vos relectures finales méticuleuses de chacun des chapitres m'ont permis de préciser mon propos. Je ne sais comment exprimer ma gratitude à vous deux, autrement qu'en vous promettant d'agir comme vous avec des étudiants dans ma situation, si un jour l'occasion m'est donnée.

Merci à **Hélène Gilbert**, pour m'avoir aidé à comprendre la méthodologie de détection de QTL.

Un grand merci au président de l'université de Djibouti **Dr Djama Mohamed Hassan** pour son soutien indéfectible aux doctorants. Un petit mot pour ma doyenne et néanmoins amie **Madina Daher**, à qui je lui souhaite une carrière à la mesure de ses compétences et ses valeurs.

Je remercie tout le personnel de la SAGA pour leur accueil chaleureux. Je pense particulièrement à mes voisins de bureau **Hervé** (gare aux gorilles) et **Benjamin** (canardo), les meilleurs que je pouvais avoir. Je pense à **Gilbert** pour les dépannages informatiques, à **Valérie**, **Line**, **Dounia, Nancy** et **Carine** pour les démarches administratives, à **Olivier** pour les dépannages de QTLMap, à **Andrès** pour sa disponibilité, à **Zulma** pour les données « Weibull » et à la bande du

« barbec » et autres sorties nocturnes, **Carine, Llibertat, Cecile, Aurélie, Cyrielle, Félicien, Chloe, Yoannah, Guillaume**, **Charlotte**, …

Merci à tous les pensionnaires de la place « **AREZKI** », à mon ami **David** qui a supporté mes humeurs dépressives tout au long de ces années, quelle plaisir de t'avoir connu. Merci à **Abde** pour les sorties à Cadet et les longues après-midi de rigolade. Merci à **Charles-Elie** pour son rire atypique et l'hymne américain. Merci à **Patinet**, pour la relecture et la préparation du pot… la lutte continue.

Merci à ma « dream-team » préférée, **Diouf, Ilyas, Houssein, Mah** et **Said Dabar** pour leur soutien sans faille, les soirées Darfour et tous les moments de détente et fous rires inoubliables passés ensemble. Un clin d'œil amical aux fréquentations occasionnelles de « Darfour » **Zouhour** et **Kanny**.

Merci à mes amis djibouto-toulousains (**Araksan**, **Said**, **Aden** et **Abdillahi**) pour les discussions des Samedis soirs et leur hospitalité, à **Moussa** pour ses appels le matin de bon heure et les bons moments passés à Paris, à **Dr Bouh** pour ces encouragements et ses conseils, à **Greg** et **Haidar** pour leur soutien amical.

Mes dernières pensées vont à **mes parents** qui ont su me donner une éducation de qualité et ont mis à ma disposition les maigres moyens financiers en leur possession pour que j'aille le plus loin possible dans mes études. Je pense à eux aujourd'hui, qu'ils trouvent ici toute ma gratitude et ma reconnaissance pour tous leurs efforts. Merci aussi à mes frères et sœurs qui m'ont apporté leurs soutiens à leur manière, **Said** et son affection sans faille, **Mahamoud** qui a toujours les mots justes, **Souleiman, Wais, Nafissa, Nima, Kadir, Hamza** et **Kafia** ma petite princesse capricieuse. Et bien sure un grand merci à **Oubah** pour son tendre soutien ô combien important pour finir cette thèse.

Enfin, merci à celles et ceux dont j'ai omis les noms par oubli et qui se reconnaîtront dans ces quelques lignes.

Résumé :

La recherche de QTL par régression des phénotypes sur les probabilités de transmission (modèle Haley-Knott) est une méthode très largement utilisée quand on dispose de grandes familles phénotypées par des caractères gaussiens. L'objectif de cette thèse d'un point de vue méthodologique, est de proposer une méthode de détection de QTL qui prend en compte des effectifs de familles petits d'une part, et l'existence de caractères discrets d'autre part. Ainsi, nous proposons, pour répondre à la première question, une approche de détection de QTL intégrant dans le calcul du mérite génétique des individus marqués, les performances calculées sur n générations de descendants. L'obtention d'un mérite génétique dérégressé comme substitut de phénotypes, proposé notamment par Weller *et al* (1990) et Tribout *et al* (2008), est donc généralisée. Ensuite, sont présentés les résultats de comparaisons d'un modèle supposant la normalité des données à un modèle à seuils faisant l'hypothèse d'une distribution continue sous jacente à la distribution observée dans la détection de QTL des caractères discrets. Nous démontrons ici que le modèle discret est plus précis et plus puissant quand le caractère étudié possède trois modalités distribuées de façon déséquilibrée dans la population.
Dans la deuxième partie de la thèse, l'analyse des données du protocole GENECAN a été réalisée. Il s'agit d'identifier les régions du génome ou locus à caractère quantitatif (QTL), associées à des caractères d'intérêt mesurés sur des canards mulards gavés. Le canard mulard est un hybride interspécifique obtenu par croisement d'une cane commune (Anas platyrhynchos) et d'un canard de Barbarie (Cairina moschata). Trois cents quarante deux canes communes conçues en back-cross (BC) ont été générées par croisement d'une lignée de canard Kaiya et d'une lignée de canard Pékin lourd. Ces femelles BC ont été accouplées avec des canards de Barbarie pour produire 1600 canards mulards sur lesquels sont effectuées des mesures de croissance, de métabolisme au cours de la période de croissance et du gavage, d'aptitude au gavage et de qualités du magret et du foie gras. La valeur phénotypique des femelles BC marquées a été estimée, pour chaque caractère, comme étant la valeur moyenne des phénotypes de sa progéniture et pondérée par un coefficient de détermination (CD) fonction du nombre de descendants et de l'héritabilité du caractère étudié. Une carte génétique de 91 marqueurs microsatellites réparties sur 16 groupes de liaison (GL) et couvrant un total de 778 cM a été utilisée. Dans le cadre de l'analyse uni-caractère, vingt-deux QTL significatifs à 1% au niveau du chromosome ont été cartographiés. Ces QTLs sont pour la plupart impliqués dans la variabilité de la qualité du magret et du foie gras. Les zones chromosomiques d'intérêt, identifiées dans le cadre de cette étude devront dans le futur, être densifiées en marqueurs pour faire l'objet d'une cartographie fine.

Mots clés : QTL, mérite génétique, protocole petite fille, caractères discrets, multi-caractères, pléiotropie, canard mulard, qualité du foie, qualité du magret, caractère métabolique.

Abstract:

QTL detection using the regression of phenotypes on transmission probability is largely used when large families phenotyped for Gaussian trait are available. The aim of this thesis from a methodological point of view, is to propose a method for detection of QTL that takes into account the small number of families on the one hand, and the existence of discrete traits on the other. Thus, we propose to answer the first question, an QTL detection approach, integrating in the calculation of genetic merit of genotyped individuals, the performances calculated over n generations of descendants. The use of a 'de-regressed proof' as a phenotype to be analysed, proposed by Weller et al. (1990) and Tribout et al. (2008) is generalized. Next, we present the results of comparisons of a model assuming normality of the data to a thresholds model assuming a continuous distribution underlying the observed distribution in the QTL detection of discrete traits. Here we demonstrate that the discrete model is more accurate and more powerful when the studied trait has three modalities distributed unevenly in the population.

In the second part of the thesis, the data analysis of GENECAN protocol was performed. This is to identify genomic regions or quantitative trait locus (QTL) associated with interest traits measured on over-feed mule ducks. The mule duck is an hybrid duck from a female Common duck (*Anas Platyrhynchos*) and a Muscovy drake (*Cairina moschata*). Three hundred forty two common ducks designed by back-cross (BC) were generated by crossing a line of Kaiya duck and a heavy line of Pekin duck. These BC females were mated with Muscovy ducks to produce 1600 mules ducks which undergo measures of growth, metabolism during the growth and over-feeding periods, over-feeding, of breast muscle and fatty liver qualities. The phenotypic value of genotyped BC females was estimated for each trait as the average phenotypes of their offspring and weighted by a coefficient of determination (CD) function on the number of offspring and heritability of the studied trait. The genetic map comprised 91 microsatellite markers aggregated into 16 linkage groups (LG) and representing 778 cM. For the uni-trait analysis, twenty-two QTL significant at 1% threshold in chromosome-wide have been mapped. These QTLs are mostly involved in the variability of the breast muscle and fatty liver qualities. Chromosomal regions of interest identified in the framework of this study should be in the future be densified to markers to do the fine mapping.

Keys words: QTL, genetic merit, granddaughter designs, threshold traits, multi-traits, pléiotropy, mule duck, liver quality, meat quality, metabolic trait.

Sommaire

Introduction générale 6
Chapitre I. Méthodologie de détection de QTL par analyse de liaison : revue bibliographique 9
1 Notions élémentaires 10
1.1 Introduction 10
1.2 Les marqueurs génétiques 11
1.3 La recombinaison 11
1.4 Les distances génétiques 11
1.5 Les cartes génétiques 12
2 Principes généraux de la détection de QTL 14
2.1 Introduction 14
2.2 Type de populations utilisées dans la détection de QTL chez les animaux 15
2.2.1 Introduction 15
2.2.2 Lignées aux allèles fixés ou population « inbred » 16
2.2.3 Lignées avec des génotypes variés ou population « outbred » 19
2.2.3.1 Protocoles « filles » 20
2.2.3.2 Protocoles « petites-filles » 21
2.3 Les outils statistiques utilisés 22
2.3.1 Introduction 22
2.3.2 Détection marqueur par marqueur 22
2.3.2.1 Approche par analyse de variance 22
2.3.2.2 Approche par maximum de vraisemblance 23
2.3.3 La cartographie d'intervalle 24
2.3.3.1 Principe général (Lander et Botstein 1989) 24
2.3.3.2 Méthode de régression linéaire : le cas d'une population F2 (Haley et Knott, 1992) 26
2.3.3.3 La cartographie d'intervalle appliquée à une population « outbred » 28
2.3.4 Le modèle à effet aléatoire pour les pedigrees complexes 29
2.4 Statistique de test et précision de la position du QTL 31
2.4.1 Introduction 31
2.4.2 Détection avec un seul marqueur 31
2.4.3 Détection avec plusieurs marqueurs : le cas de la cartographie d'intervalle 31
2.4.3.1 Le seuil de significativité 32
2.4.3.2 L'intervalle de confiance 32
3 Améliorations des méthodes de détection de QTL 34
3.1 Autres formes de pénétrance 34
3.1.1 L'approche multi-caractères 34
3.1.2 Etude des caractères discrets ou caractères « à seuil » 35
3.2 Détection multi-QTLs 37
3.2.1 Cartographie d'intervalle composite (CIM) et cartographie de QTL multiple (MQM) 38
3.2.2 La cartographie d'intervalle multiple (MIM) 38
3.3 Amélioration de la puissance : le dispositif petites-filles (Weller et al., 1990) et la prise en compte du « mérite génétique » 39
4 Outils de calcul 42

5 Conclusions 43
Chapitre II. Apports méthodologiques : Amélioration de l'estimation du mérite génétique et prise en compte, en détection de QTL des caractères à seuils 44
1 Prise en compte des performances mesurées sur les générations postérieures à la troisième génération en détection de QTL avec analyse de liaison 45
1.1 Introduction 45
1.2 Article 1 46
1.3 Conclusion et perspective 54
2 Etude comparative des deux modèles dans la détection de QTL des caractères à seuils : considération de la sous-jacente normale ou utilisation du modèle de mélange de distributions normales 55
2.1 Introduction 55
2.2 Méthode 56
2.3 Simulations 58
2.4 Résultats 59
2.5 Discussion et conclusion 62
Chapitre III. Le canard Mulard : zootechnie, paramètres génétiques et QTLs connus pour les caractères d'intérêt 65
1 Génétique et zootechnie du canard mulard 66
1.1 Introduction 66
1.2 Les souches parentales 66
1.3 Zootechnie du mulard 67
1.3.1 Phase de croissance du canard mulard destiné au gavage 67
1.3.2 Phase de gavage du canard mulard 68
1.4 Les caractères zootechniques d'intérêt du canard mulard 69
1.4.1 La croissance 69
1.4.2 L'efficacité alimentaire 70
1.4.3 La composition corporelle et aptitude au gavage 71
1.4.4 La qualité des produits 73
1.4.5 Le comportement 75
1.5 Les déterminismes génétiques des caractères d'intérêt chez le canard mulard 76
1.5.1 Caractères liés à l'élevage des canards mulards 76
1.5.2 Caractères liés au gavage des canards mulards 78
2 Les QTLs identifiés pour les caractères d'intérêt chez le canard 80
2.1 Introduction 80
2.2 Les QTLs détectés pour les caractères d'intérêt 80
2.2.1 Introduction 80
2.2.2 Les caractères de croissance 81
2.2.3 Les caractères liés au métabolisme 82
2.2.3.1 Les caractères liés au stress 82
2.2.3.2 Taux de cholestérol, de glucose et de triglycérides dans le sang 82
2.2.4 Les caractères de qualité des produits 83
Chapitre IV. Détection de QTL dans le protocole de canard gras 84
1 Introduction 85
2 Le dispositif détaillé 86
3 Les mesures des caractères 87
4 Résultats et discussion 94
4.1 Article 2 94
4.2 QTLs significatifs à 5% au niveau du chromosome 125
4.2.1 Les caractères de croissance 125

4.2.2 Les caractères liés au métabolisme durant le gavage et les caractères liés au stress (taux de corticostérone avant et après stress) 125
4.2.3 Les caractères d'aptitude au gavage 126
4.2.4 Les caractères de qualité de produits 126
4.3 Les analyses multi-caractères 127
5 Conclusions et perspectives 130
Conclusions générales 131
Références bibliographiques 135
Table des figures 149
Table des tableaux 150
Annexes 151

Introduction générale

Avec 75% de la production mondiale en 2009, la France est de loin le premier producteur mondial du foie gras (source : CIFOG, 2009) mais aussi le premier consommateur. Ainsi, en 2009, 23 000 tonnes de foie gras cru ont été proposées à la vente en France (source : CIFOG, 2009) dont 18% était importé. La filière « palmipèdes gras » française pèse 1,7 milliards d'euros de chiffre d'affaires. Une grande part de la production française du foie gras (4 684 tonnes dont 43% de foie gras cru en 2008) est destinée à l'exportation. Le commerce extérieur pour les échanges de foie gras (cru et transformé) présente un solde positif de 36 millions d'euros.

Le foie gras, produit traditionnel français, fait partie depuis 2006 du patrimoine culturel et gastronomique (Loi n° 2006-11 du 5 janvier 2006 d'orientation agricole ajoutée à l'article L654-27-1 au code rural français). L'espèce la plus utilisée pour la production du foie gras est le canard mulard (canard issu du croisement d'un mâle de Barbarie et d'une cane commune, généralement une cane Pekin). La filière canard gras concurrence la filière maigre, puisqu'elle génère une part majoritaire de la production de viande de canard en France (58% de la production totale en 2008).

Le canard mulard étant un hybride stérile, le progrès génétique est créé dans les populations parentales. Un grand nombre de caractères (de croissance, de reproduction et d'aptitude au gavage) ont fait l'objet de sélection dans cette espèce. Quelques études analysant la variabilité génétique des caractères d'intérêt chez le mulard ont été publiées (Marie-Etancelin *et al*., 2008 ; Marie-Etancelin et al., 2011) mais aucune de ces études n'a investigué les zones chromosomiques impliquées dans la variabilité de ces caractères.

D'où l'intérêt de mettre en place un programme de détection de QTL (Quantitative Trait Locus dont le polymorphisme a un effet sur la variabilité des caractères quantitatifs) qui permet d'identifier les zones chromosomiques contenant un ou des gènes dont la variation modifie significativement la performance des animaux pour les caractères considérés. Cela permettra aux sélectionneurs d'utiliser cette information pour choisir les animaux (au niveau des populations parentales) porteurs des meilleurs allèles au(x) QTL(s). Les programmes QTL fournissant généralement une information imprécise (le QTL est contenu dans une région chromosomique assez large), il faudra à terme réaliser des travaux complémentaires à nos analyses en procédant à une cartographie fine du génome afin d'aboutir à l'identification d'une éventuelle mutation causale au locus d'un gène.

Les résultats du premier dispositif de détection de QTL pour des canards gavés sont présentés dans cette thèse. Différentes questions méthodologiques posées par le dispositif sont abordées et de nouveaux développements pour y répondre sont présentés. Tout d'abord, comment prendre en compte au mieux les performances des mulards dans une recherche de QTL basée sur les informations marqueurs de leurs mères canes communes, issus d'un croisement (Pékin lourd x commun léger) ? De plus, le caractère non gaussien de certains caractères étudiés (importance des taches sur les viscères, facilité de gavage, …) a apporté un second niveau de complexité méthodologique : en effet, la question d'éventuelles améliorations (en termes de puissance de détection de QTL, de précision d'estimations de la position du QTL, …) apportées par l'utilisation d'un modèle à seuil dans la détection de QTL des caractères discrets est posée.

Dans un premier temps (**chapitre I**) les méthodes de détection de QTL par analyse de liaison sont passées en revue et dans un deuxième temps, sont présentés les développements nouveaux apportés dans le cadre de cette thèse (**chapitre II**). En effet : (i) une approche de détection de QTL prenant en compte, dans le calcul du mérite génétique des individus marqués, les performances calculées sur n générations de descendants est présentée et (ii) un modèle supposant la normalité des données et un modèle à seuil faisant l'hypothèse d'une distribution continue sous jacente à la distribution observée dans la détection de QTL des caractères discrets sont comparés. La puissance de détection du QTL et la précision de l'estimation de sa position sont les deux principaux critères considérés. Ensuite, un tour d'horizon des connaissances zootechniques et génétiques du canard mulard est réalisé (**chapitre III**). Sont notamment présentés les paramètres génétiques estimés et les QTLs cartographiés (tant chez la poule et que chez le canard maigre) pour des caractères d'intérêt chez le canard gavé. Enfin les nouvelles zones chromosomiques identifiées dans le cadre de cette thèse comme porteuses de QTL sont présentées et discutées (**chapitre IV**). Un intérêt particulier est porté aux QTLs significatifs au moins à 1% au niveau du chromosome.

Chapitre I. Méthodologie de détection de QTL par analyse de liaison : revue bibliographique

1 Notions élémentaires

1.1 Introduction

Le génome correspond à l'intégralité du matériel génétique d'un individu, organisé en chromosomes. Les chromosomes, molécules d'Acide DésoxyriboNucléique (ADN), sont associés chez les espèces diploïdes et à reproduction sexuée, par paires de chromosomes homologues (l'un provenant du père et l'autre de la mère). En absence d'anomalie, on compte 40 paires chez les canards, 32 chez les chevaux, 30 chez les bovins et caprins, 27 chez les ovins et 23 paires chez l'Homme (encyclopédie *universalis*).

Les chromosomes sont composés de séquences codantes (transcrites en ARN messager puis traduites en protéines) et de séquences non-codantes (non transcrites ou non traduites). Le gène appartient aux séquences codantes : c'est une unité d'information génétique transmissible, codée par une succession de bases nucléotidiques ayant 4 formes possibles (adénine, guanine, cytosine et thymine). La région du chromosome où se situe un gène particulier est appelée son locus. Chaque gène possède différentes versions appelées allèles.

Si tous les individus d'une même espèce ont les mêmes gènes, chacun a un ensemble spécifique d'allèles. Un individu donné possède, pour chaque gène, un allèle reçu de son père et un allèle reçu de sa mère : c'est son génotype au locus. La séquence ordonnée des allèles de plusieurs locus adjacents forme un haplotype. Chez les espèces diploïdes, chaque individu a deux séries d'haplotypes, l'une héritée de son père et l'autre héritée de sa mère. L'analyse de liaison visant à mettre en évidence les portions de génome (QTL pour Quantitatif Trait Loci ou locus à caractère quantitatif) impliquées dans le déterminisme des caractères d'intérêt a été rendu possible par le développement des cartes génétiques de marqueurs. Il s'agit d'identifier les locus dont la variation allélique est associée à une variation d'un caractère d'intérêt.

Cette détection de QTL nécessite la mise en place de dispositifs animaux spécifiques. Deux types de populations peuvent être utilisés : les populations « inbred », lignées consanguines aux allèles fixés, et les populations « outbred » chez lesquelles les génotypes sont variés. Ces dispositifs font intervenir plusieurs générations d'individus (aux moins deux) composés de plusieurs familles de pères et/ou mères. La cartographie de QTL est d'autant plus précise qu'il y a de parents informatifs, c'est-à-dire hétérozygotes aux marqueurs et aux éventuels QTLs associés. Les croisements entre lignées « inbred » sont de ce point de vue très efficaces. Puis il est nécessaire de réaliser i) le marquage moléculaire de la génération des descendants et, quand c'est possible et utile, de générations antérieures. Il s'agit du génotypage des individus pour un réseau de marqueurs génétiques polymorphes régulièrement espacés sur le génome et repérables dans une carte génétique, et ii) le phénotypage des individus de la dernière génération c'est-à-dire la mesure des caractères d'intérêt.

Les analyses statistiques nécessaires pour analyser ces données reposent sur diverses techniques telles que la méthode du maximum de vraisemblance ou la

régression linéaire. Dans cette première partie bibliographique nous définirons, très brièvement, les différentes notions biologiques et statistiques indispensables pour comprendre les méthodes de détection de QTL proposées dans la littérature avant de détailler lesdites méthodes.

1.2 Les marqueurs génétiques

Par définition, un marqueur génétique est une séquence polymorphe d'ADN aisément détectable, située dans un locus unique et qui se transmet selon les lois mendéliennes. Il est utilisé pour "baliser" le génome. Comme cela sera décrit au paragraphe 3, l'ordonnancement de ces marqueurs et l'estimation de leurs distances génétiques aboutissent à la construction des cartes génétiques. Les marqueurs génétiques permettent de suivre la transmission d'un segment de chromosome d'une génération à l'autre. En effet, si un allèle X porté par un individu existe chez son père mais pas chez sa mère, l'individu l'a reçu de son père.

Les marqueurs les plus utilisés sont :

- les marqueurs microsatellites. Il s'agit de la répétition simple ou multiple d'un motif court composé de 2 à 10 nucléotides. C'est le nombre de répétitions qui définit les allèles. Pour que ce marqueur puisse être un repère non ambigu, il doit être entouré à droite et à gauche de séquences uniques.
- les marqueurs SNP (Single Nucleotide Polymorphim). Ils se caractérisent par la variation (polymorphisme) d'une seule paire de bases du génome. Les SNP sont très généralement bi-alléliques.

Quel que soit le type de marqueur, lorsque, pour un individu donné, à un locus donné, les 2 chromosomes portent des allèles différents, on dit que l'individu est hétérozygote au marqueur.

1.3 La recombinaison

Les loci sont donc disposés linéairement sur le génome et on considère que deux loci situés sur deux chromosomes différents sont indépendants. Lors de la méiose (processus se déroulant durant la gamétogénèse chez les animaux diploïdes et produisant des cellules haploïdes), un brassage génétique (mélange de gènes) a lieu à la fois au sein des chromosomes et entre les chromosomes.

La probabilité pour que deux loci différents situés sur un même chromosome soient séparés par une recombinaison réciproque (crossing-over) entre ce chromosome et son homologue au cours de la méiose, est d'autant plus grande que ces loci sont plus éloignés l'un de l'autre. Si les loci sont proches ou s'il y a un nombre pair de recombinaisons entre eux, ils restent groupés sur le même chromosome dans le gamète produit par cette méiose. De tels loci, qui tendent à être hérités ensemble, sont dits liés. S'il y a un nombre impair de crossing-over entre eux, une recombinaison est observable.

1.4 Les distances génétiques

La proportion de recombinaison entre deux locus est une mesure de leur éloignement. Des gènes sont considérés comme liés si la probabilité de recombinaison est inférieure à 0,5. Pour des gènes non liés, la probabilité de recombinaison est égale à 0,5. Cependant, cette mesure ne possède pas les propriétés mathématiques des distances, mais, l'espérance du nombre de recombinaisons possède ces attributs.

Pour calculer la distance entre deux gènes sachant le taux de recombinaison r (proportion de gamètes recombinés sur l'ensemble des gamètes transmis) de ces derniers, plusieurs fonctions de calcul sont disponibles :

- La distance de Morgan : d = r

- La fonction de Haldane (Haldane, 1919) qui prend en compte les crossing-over multiples possibles sur un chromosome et suppose l'indépendance entre eux :

$$d = -\frac{1}{2}\ln(1-2r)$$

- La fonction de Kosambi (Kosambi, 1944) qui, plus générale, prend en compte l'interférence entre les différents crossing-over. C'est-à-dire l'observation que la présence d'un crossing-over diminue la probabilité d'un autre crossing-over à proximité (Précis de génomique par Greg Gibson et Spencer V. Muse) :

$$d = \begin{cases} \frac{1}{4}\ln\left(\frac{1+2r}{1-2r}\right) si\ 0 \leq r < 1/2 \\ +\infty\ sinon \end{cases}$$

Les distances ci-dessus sont toutes exprimées en centimorgan (cM). Utilisée de façon inverse, elles permettent le calcul du taux de recombinaison entre deux gènes si la distance qui les sépare est disponible.

1.5 Les cartes génétiques

Une carte génétique est un inventaire de gènes ou de marqueurs génétiques d'une espèce ou d'un groupe d'individus. Elle définit l'ordre relatif des marqueurs génétiques et/ou des gènes, et les distances les séparant les uns des autres sur chaque chromosome. Concrètement c'est une représentation du génome d'une espèce.

La construction d'une carte génétique se fait en deux grandes étapes. On constitue d'abord des groupes de liaison, c'est-à-dire des ensembles de marqueurs liés entre eux. Pour qu'un marqueur donné soit affecté à un groupe de liaison, il doit être lié à au moins un des marqueurs du groupe. Un groupe de liaisons correspond à un chromosome ou à un morceau de chromosome quand les marqueurs qui le composent sont, au moins en partie, ancrés sur ledit chromosome. Pour constituer ces groupes, on utilise la probabilité de recombinaison (r) entre les marqueurs deux à deux, en testant, grâce aux méthodes de maximum de vraisemblance, l'hypothèse H_0: r = 0,5 c'est-à-dire qu'« il y a indépendance des marqueurs » contre l'hypothèse H_1: 0 ≤ r < 0,5 c'est-à-dire que « les deux marqueurs sont liés ». Ensuite les marqueurs sont ordonnés à l'intérieur de chaque groupe de liaison, c'est-à-dire qu'on cherche l'ordre le plus vraisemblable de ces marqueurs. Des logiciels utilisant le maximum de

vraisemblance sont disponibles pour réaliser ce travail, par exemple CRIMAP (Green *et al.,* 1990) ou MapMaker (Lander *et al.*, 1987).

2 Principes généraux de la détection de QTL

2.1 Introduction

En génétique quantitative on considère le plus souvent le modèle infinitésimal pour étudier les caractères d'intérêt. Ce modèle génétique suppose qu'il existe une infinité de gènes qui agissent sur une expression phénotypique donnée avec chacun un effet individuel faible. Le modèle polygénique prend en compte différents types d'effets sur le caractère étudié. En effet, les performances des individus y sont décrites comme la somme d'effets génétiques et d'effets environnementaux, ces derniers pouvant se décomposer à leur tour en plusieurs « sous effets » tels que les effets d'élevages ou des lieux géographiquement différents, des alimentations différentes, des interactions entre les effets, etc…

Sous l'hypothèse d'indépendance des effets, le modèle s'écrit :

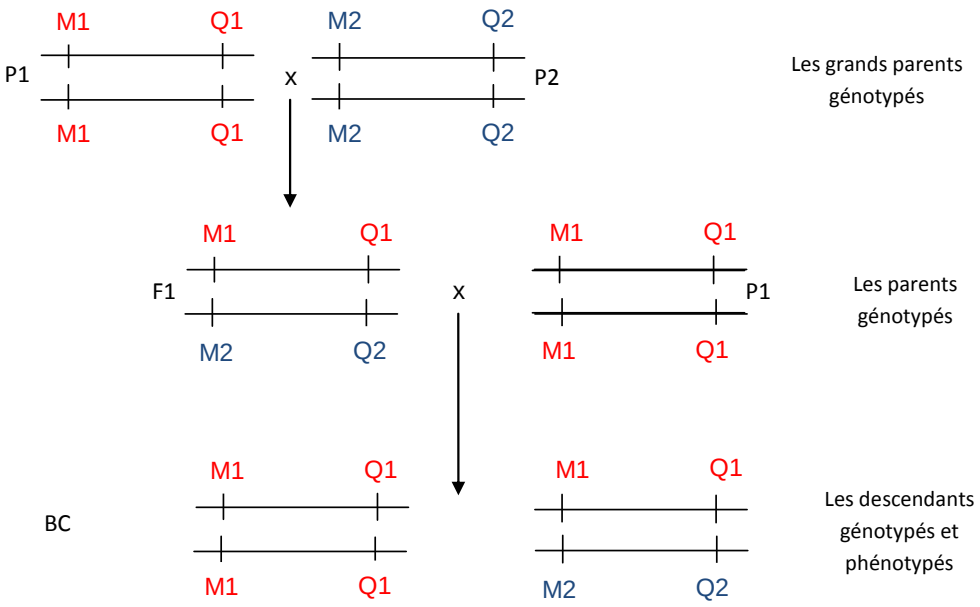
$$Y = \mu + G + E + e$$

où Y est la valeur phénotypique, μ = E(Y) la moyenne des phénotypes Y, G la valeur génotypique, E les effets identifiables du milieu et e est l'erreur résiduelle du modèle.

La partie héritable d'un phénotype réside donc dans sa valeur génétique. La variabilité des valeurs génétiques entre les individus, dans une population donnée, résulte des effets des gènes polymorphes. Parmi tous les gènes qui agissent sur le caractère, certains ont des effets plus importants que les autres. Quand la variabilité phénotypique du caractère d'intérêt étudié est due à l'effet d'un seul gène, on parle de gène majeur. Mais dans la plupart des cas, la variabilité des caractères quantitatifs est contrôlée par plusieurs gènes, chacun ayant un faible effet sur le caractère étudié. De tels gènes, polymorphes, sont localisés en divers points du génome. Un QTL est une portion du génome plus ou moins large où sont localisés un ou plusieurs gènes responsables, au moins en partie, de la variation phénotypique d'un caractère quantitatif d'intérêt. Les marqueurs génétiques permettent de repérer et de localiser ces QTLs.

Le principe général de la détection de QTL par analyse de liaison est le test de différences phénotypiques entre classes de descendants selon l'origine grand-parentale de segments chromosomiques hérités d'un parent. La détection de QTL consiste donc pour un marqueur donné M, à observer dans la descendance d'un parent hétérozygote M1/M2 (M1 et M2 étant les allèles au marqueur M) s'il existe une différence de performance moyenne selon l'allèle du marqueur, M1 ou M2, transmis. L'hypothèse sous-jacente est que, si cette différence existe, elle s'explique par la ségrégation des allèles Q1 et Q2, en un QTL génétiquement lié au marqueur. Par exemple, dans le cas d'un backcross (figure 1.1) entre deux populations P1 (homozygote M1Q1/M1Q1) et P2 (homozygote M2Q2/M2Q2), un père F1 (M1Q1/M2Q2) accouplé à une conjointe de P1 donne un descendant Q1/Q1 lorsqu'il transmet M1 et un descendant Q2/Q1 lorsqu'il transmet M2, en supposant l'absence de recombinaison entre M et Q (Le Roy et Elsen 1995), (Soller et Genizi, 1978), (Niemann-Sorensen et Robertson 1961).

Figure 1.1 : Transmission des allèles au marqueur M et au QTL dans une population BC sans recombinaison



Sous l'hypothèse nulle (absence de QTL), la valeur phénotypique moyenne est indépendante du génotype au marqueur. En conséquence, l'observation d'une différence significative entre les valeurs moyennes du caractère quantitatif des génotypes au marqueur indique la possibilité d'une liaison du marqueur avec un QTL. Des tests statistiques appropriés permettent de tester cette hypothèse. Quand les comparaisons ne concernent que deux classes (comme avec les protocoles backcross ou demi-frères), celles-ci peuvent être comparées avec un simple test-*t*. Dans le cas de populations « outbred », un modèle linéaire d'analyse de variance (Knott *et al.,* 1996) peut être utilisé. La performance d'un individu *i* s'écrit :

$$Y_{ij} = s_i + (2p_{ij} - 1)a_i + e_i$$

Y_{ij} est le phénotype du *j*^{ème} *descendant du père i*, s_i est l'effet père i, p_{ij} est la probabilité que le descendant j ait reçu l'allèle 1 du père i conditionnellement à l'information aux marqueurs. a_i est l'effet de l'allèle 1 du père i au QTL supposé et e_i est la résiduelle du modèle supposée de moyenne nulle et de variance σ^2.

2.2 Type de populations utilisées dans la détection de QTL chez les animaux

2.2.1 Introduction

Mettre en place des méthodes statistiques robustes pour permettre la détection de QTLs (et à terme identifier les gènes sous-jacents) est l'un des plus grands enjeux de la génétique quantitative de ces dernières décennies. En effet, contrairement aux gènes majeurs qui eux, ont des effets importants sur les caractères, les QTLs ne sont souvent responsables que d'une très faible part de la variabilité du caractère et sont

donc plus difficiles à identifier. Des protocoles adaptés doivent être mis en place. Après le choix des parents, fait de façon à maximiser les chances d'obtenir du polymorphisme moléculaire dans la descendance, la cartographie de QTL implique la création d'une descendance en ségrégation.

2.2.2 Lignées aux allèles fixés ou population « inbred »

Une population « inbred » est constituée d'individus hautement consanguins, dont on espère qu'ils sont homozygotes pour la plus grande partie de leurs locus. De telles populations n'existent pas chez les animaux de ferme, mais uniquement chez les plantes et les animaux de laboratoire (souris, …). Cependant beaucoup de protocoles et de méthodes d'analyse dérivent d'études reposant sur ces populations.

Avec ces populations « inbred » les protocoles les plus utilisés consistent :

1. à croiser deux lignées « inbred » pour ainsi obtenir une génération de descendants F_1 dans laquelle tous les individus sont hybrides (supposés hétérozygotes à tous les locus)

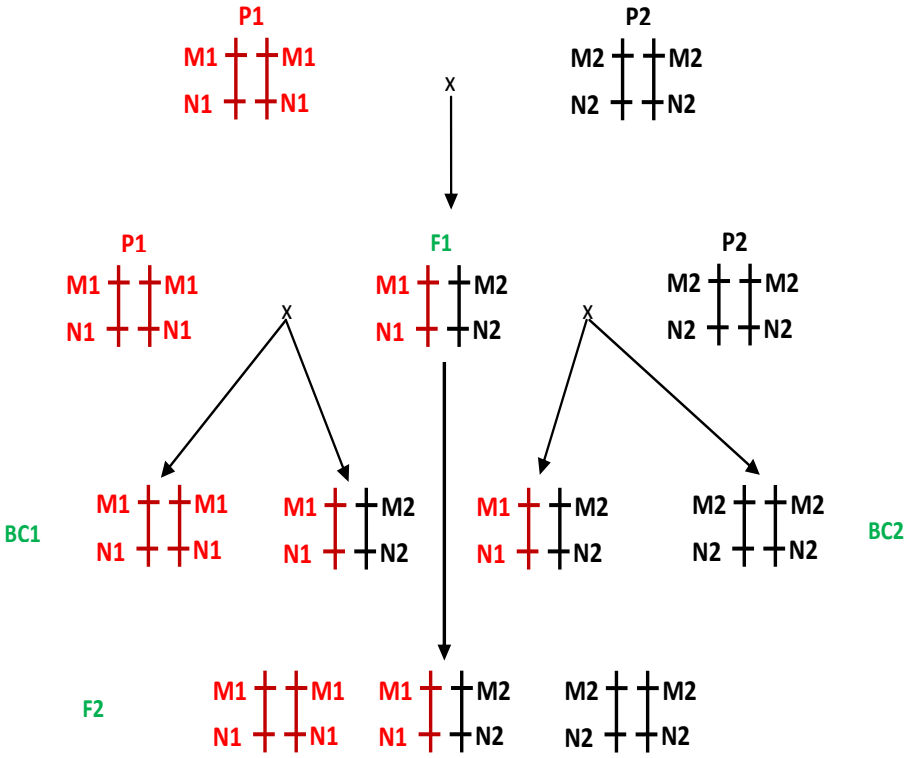
2a. et à croiser cette génération avec l'une des lignées parentales. Le résultat constitue la population rétro-croisée ou population backcross (figure 1.2) : c'est cette population qui sera analysée.

2b. ou à croiser entre eux des individus F1, pour donner une population F2 dont on mesurera les phénotypes pour réaliser la détection de QTLs.

A partir de ces lignées homozygotes fixées pour des allèles différents aux marqueurs, on dispose de parents F1 doubles hétérozygotes au marqueur et au QTL. La cartographie des QTLs s'en trouve simplifiée puisqu'en seconde génération de croisement, la coségrégation des allèles au marqueur et au QTL peut être observée aisément.

Dans le cas d'une population de backcross, on observe qu'une moitié des individus produits sont homozygotes et une moitié hétérozygotes (figure 1.2). Les performances de ces deux sous groupes sont comparées pour cartographier les QTLs. Dans le cas d'une population F2 on observe un quart d'individus homozygotes pour le premier allèle du marqueur, un quart d'homozygotes au deuxième allèle du marqueur et la moitié de la population qui est hétérozygote (figure 1.2). Les performances des deux sous groupes homozygotes peuvent être comparées pour cartographier les QTLs à l'aide d'un simple test de comparaison de moyennes. On peut aussi inclure les hétérozygotes dans l'évaluation et tester la significativité d'un effet génotype sur l'expression des performances des F2.

Figure 1.2 : Protocoles de détection de QTL chez les populations « inbred » sans prise en compte de recombinaison

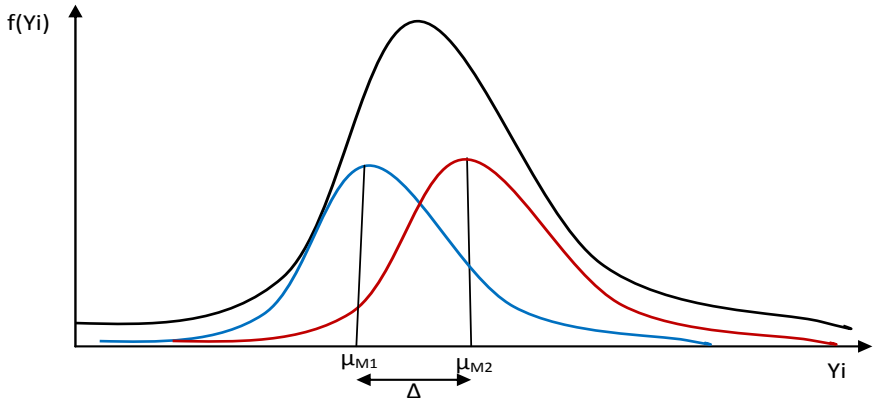


Description : M1 et M2 correspondent aux allèles du premier locus et N1 et N2 ceux du second.

Quel que soit le protocole de croisement choisi, dans le cadre d'une détection de QTL classique, avec des populations « inbred », il est toujours nécessaire i) de réaliser le marquage moléculaire (génotypage) des descendants des parents hétérozygotes au marqueur, ce qui permettra de former des groupes de descendants en fonction des allèles reçus de leurs pères. Les phases (répartition des allèles aux marqueurs sur les chromosomes homologues) de ces derniers se déduisent du fait de l'homozygotie des grands parents. Puis ii) de mesurer les performances de la dernière génération qui peut être la génération génotypée ou les descendants de cette dernière dans le cas d'un testage sur descendance. Ces deux informations combinées permettront de tester la coségrégation entre allèles aux marqueurs transmis par le père et les valeurs phénotypiques des descendants.

Lorsque le marqueur et le QTL sont complètement liés (figure 1.3), les distributions des phénotypes selon les allèles hérités au marqueur sont strictement les mêmes que les distributions des phénotypes selon les allèles au QTL.

Figure 1.3 : Courbes de distributions des phénotypes en fonction des allèles au marqueur reçu



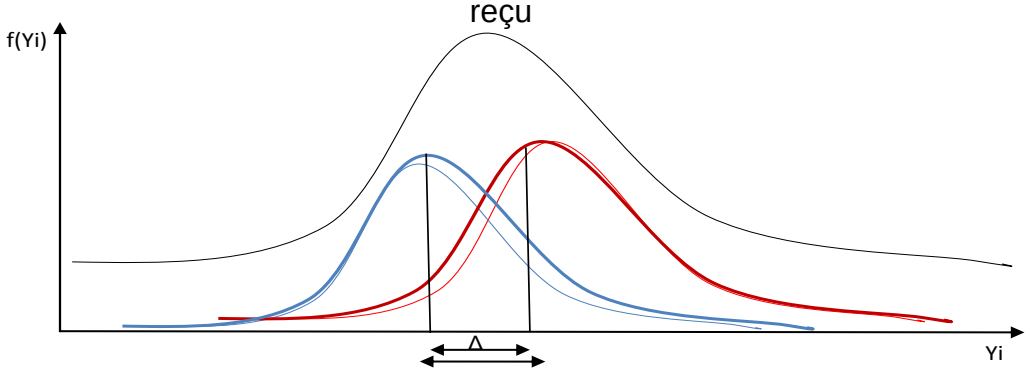
Description : La courbe noire représente la distribution générale des individus sur lesquels les mesures phénotypiques ont été faites, la courbe bleue correspondant à la distribution des descendants qui ont reçu l'allèle M1 au marqueur et la courbe rouge, celle des descendants ayant reçu l'allèle M2. La moyenne générale des performances est μ, celle des descendants qui ont hérité l'allèle M1 au marqueur est μ_{M1} et les individus qui ont reçu l'allèle M2 au marqueur ont μ_{M2} comme moyenne. L'effet de substitution $\Delta = \mu_{M2} - \mu_{M1}$.

Si le marqueur et le QTL ne sont pas complètement liés (ce qui est généralement le cas), on n'a pas les mêmes distributions des performances (figure 1.4) selon les allèles aux marqueurs et au QTL. L'écart des espérances des distributions selon les allèles aux marqueurs est notamment réduit (par rapport à l'écart des distributions selon les allèles au QTL) proportionnellement à la distance entre le marqueur et le QTL, mesurée par le taux de recombinaison r.

$$\Delta = \mu_{M2} - \mu_{M1} = (1-r)\mu_{Q2} + r\mu_{Q1} - \left[(1-r)\mu_{Q1} + r\mu_{Q2}\right]$$
$$= (1-2r)\left(\mu_{Q2} - \mu_{Q1}\right) = (\mathbf{1-2r})\boldsymbol{\alpha}$$

Q_1 et Q_2 sont les allèles au QTL, respectivement associés à M1 et M2, α est l'effet réel du QTL et Δ l'effet de substitution estimé.

Figure 1.4 : Courbes de distributions des phénotypes en fonction des allèles au QTL reçu



Description : La courbe bleu foncé correspond à la distribution des descendants qui ont reçu l'allèle M1 au marqueur et la courbe bleue claire correspond à la distribution des descendants qui ont reçu l'allèle Q1 au QTL. La courbe rouge foncé correspond à la distribution des descendants qui ont reçu l'allèle M2 au marqueur et la courbe rouge claire correspond à la distribution des descendants qui ont reçu l'allèle Q2 au QTL

2.2.3 Lignées avec des génotypes variés ou population « outbred »

Dans les populations « outbred », la prise en compte des structures familiales et de l'informativité (l'hétérozygotie) partielle des marqueurs est importante. En effet les allèles aux marqueurs et au QTL, ainsi que leur association, peuvent être complètement différents d'un père à l'autre. Toutefois les mêmes méthodes de détection de QTL que celles décrites pour les populations « inbred » peuvent être appliquées à ces populations en considérant chaque famille comme une sous-population expérimentale du type F2 ou BC selon les génotypes des parents de la famille. Si le parent est homozygote au marqueur, sa famille sera non informative pour le marqueur en question.

La détection d'un QTL dans les populations « outbred » nécessite des effectifs beaucoup plus importants que dans les populations « inbred ». Les deux tableaux suivants comparent les effectifs nécessaires pour détecter un QTL avec une puissance de 90% et pour un risque **α** de 5% dans une population de lignées « inbred » (BC ou F2) (tableau 1.1) et une population « outbred » menée en F2 ou BC (tableau 1.2) (Soller *et al.,* 1976 ; Soller et Genizi 1978). **a** est l'effet additif du QTL (a=0,125; 0,25 ou 0,5) et, pour simplifier les tableaux, on supposera que le degré de dominance au QTL est nul. Les fréquences des allèles au marqueur (M1 et M2) et au QTL (Q1 et Q2) sont égales à ½.

Tableau 1.1 : Nombre de descendants nécessaires pour la détection d'un QTL dans un croisement entre lignées homozygotes (« inbred »)

a	0,125	0,25	0,5
BC	2672	668	168
F2	1336	334	84

Tableau 1.2 : Nombre de descendants nécessaires pour la détection d'un QTL dans le cadre d'une population « outbred » (cas d'un ensemble de familles de pleins-frères de type « BC» ou « F2 »)

	a=0,125		a=0,25		a=0,50	
nombre de familles	BC	F2	BC	F2	BC	F2
4	**13200**	14356	3300	3590	824	898
5	**12390**	13620	3098	3404	774	852
6	**12134**	12332	3034	3334	758	834
10	**12518**	13362	3130	3340	782	834
20	**14540**	15154	3634	3790	910	948
100	**26060**	26060	6514	6514	1630	1630

Ces effectifs (tableau 1.2) sont difficilement atteignables pour les espèces d'élevage. Pour se rapprocher des espèces « inbred » et augmenter la puissance du dispositif de détection, on peut utiliser des croisements de type BC ou F2 entre deux races très différentes (exemples : bovins Holstein et Normand, porcs Large White et Meishan) ou deux lignées sélectionnées de façon divergente (exemples : poulets gras et maigre, poulets sélectionnés sur la consommation alimentaire résiduelle élevée (poulets R+) ou faible (poulets R-)), la recherche de QTL étant alors ciblée sur le caractère soumis à sélection.

Parmi les protocoles utilisés pour la cartographie de QTLs dans ces types de populations, figurent des protocoles expérimentaux impliquant deux générations d'individus (protocoles « filles » ou « père ») et des protocoles nécessitants trois générations d'individus (protocoles « petites-filles » ou « grand-père »).

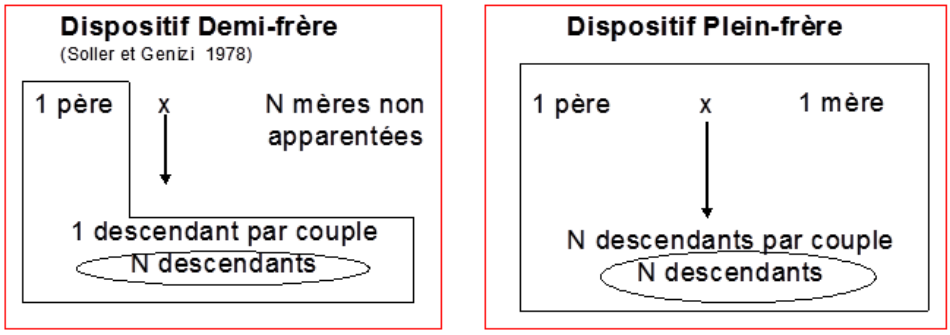
2.2.3.1 Protocoles « filles »

Essentiellement pour des raisons de biologie de l'espèce (monotoque/polytoque), on distingue, dans les protocoles « filles », les dispositifs « demi-frères » et les dispositifs « plein-frères ». Ainsi le dispositif le plus utilisé chez les ruminants est le dispositif « demi-frères » (Soller et Genizi 1978). Dans ce type de dispositif les phases des parents sont reconstituées grâce à l'observation des génotypes des descendants.

Dans le dispositif « demi-frères », les descendants sont l'objet de marquage moléculaire mais aussi des mesures phénotypiques des caractères étudiés. L'analyse de ségrégation se fait donc, sur les descendants phénotypés. La moyenne de performances des individus ayant reçu un allèle parental au marqueur est comparée à celle des descendants ayant hérité l'autre ; cette analyse se faisant intra famille de père.

Pour le dispositif « plein-frères », on prend en compte les allèles transmis par les deux parents à leurs descendants. La figure 1.5 représente ces dispositifs expérimentaux de détection de QTL nécessitant deux générations d'individus (Van der Beek *et al.,* 1995). Les animaux dont le génotype est à déterminer sont encadrés, ceux dont le phénotype est à déterminer sont encerclés.

Figure 1.5 : Dispositifs expérimentaux impliquant deux générations

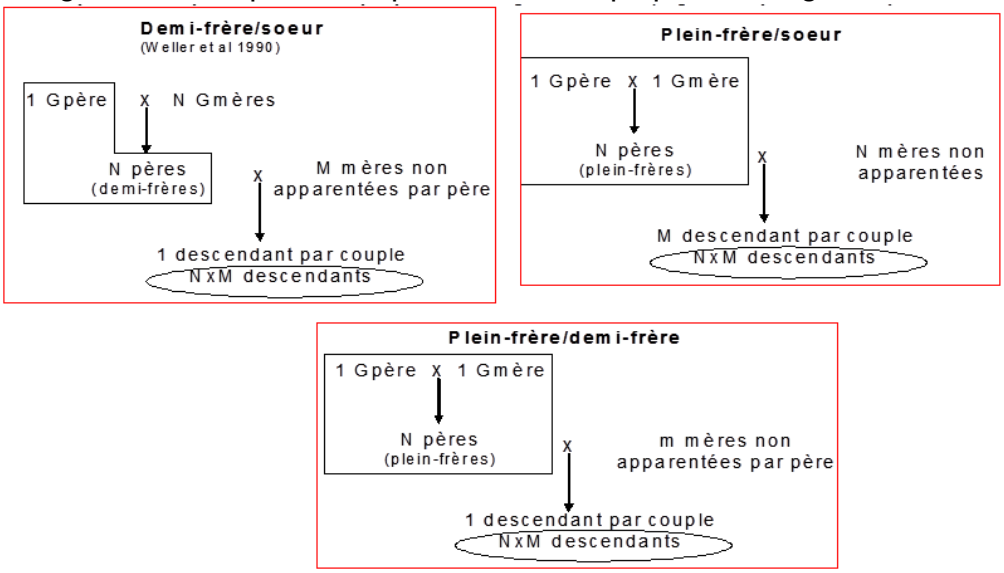


2.2.3.2 Protocoles « petites-filles »

Le principe du protocole « petites-filles » ou « grand père » a été proposé par Weller *et al.* (1990). Le principe est toujours d'observer un écart entre groupes de descendants ayant reçu de leur père un allèle ou l'autre en chaque marqueur, mais cette fois en terme de différence de valeur génétique estimée et non plus de différence de phénotype. Une famille est alors constituée sur 3 générations : un parent hétérozygote (en général un grand père) produisant des descendants (en général des fils) eux-mêmes évalués à partir des phénotypes de leurs descendants (en général les petites filles du grand père). La coségrégation entre allèles au marqueur du grand père et valeurs génétiques des fils est alors testée. L'avantage est que la valeur génétique ayant une variabilité inférieure à celle du phénotype, les différences recherchées sont plus faciles à détecter. Par ailleurs, le nombre de typages moléculaires à réaliser est très diminué par rapport à un protocole « filles » car ceux-ci ne concernent que les grands pères et leurs fils.

La figure 1.6 donne les différents dispositifs expérimentaux de détection de QTL avec trois générations d'individus (Van der Beek *et al.,* 1995). Les animaux dont le génotype est à déterminer sont encadrés, ceux dont le phénotype est à déterminer sont encerclés.

Figure 1.6 : Dispositifs expérimentaux impliquant trois générations



2.3 Les outils statistiques utilisés

2.3.1 Introduction

L'approche statistique de la détection de QTL prend des formes très variées comme l'atteste la littérature. Les premières méthodes consistaient à effectuer une analyse de variance testant l'effet d'un marqueur sur les performances (Niemann-Sorensen et Robertson, 1961). En 1989, Lander et Botstein, s'intéressant à des backcross entre lignées homozygotes, ont proposé une méthode de maximum de vraisemblance pour la recherche de QTL dans un groupe de liaison.

Dans cette partie, nous présenterons une revue bibliographique des différentes méthodes classiques de détection de QTL. Le plus souvent ces méthodes supposent l'homogénéité des variances résiduelles et la normalité des distributions. Dans un premier temps sont abordées les méthodes simples permettant de tester la présence d'un QTL putatif à proximité d'un marqueur. Ces méthodes seront décrites pour des croisements entre lignées consanguines «inbred» et seront ensuite généralisées à l'analyse intra familiale dans des populations «outbred».

2.3.2 Détection marqueur par marqueur

On cherche si un QTL localisé au voisinage du marqueur considéré, serait impliqué dans la variabilité du caractère quantitatif étudié. Deux approches sont envisagées : l'approche par analyse de variance (Soller *et al.,* 1976) et l'approche par maximum de vraisemblance (Luo et Kearsey 1989 ; Rebaï *et al.,* 1995).

2.3.2.1 Approche par analyse de variance

Si on considère que le marqueur et le QTL sont confondus au sens strict du terme, et que donc le taux de recombinaison r est nul, l'approche par analyse de variance consiste à appliquer un modèle linéaire d'analyse de variance à un seul facteur qui est le génotype au marqueur.

Dans le cas d'une lignée « inbred » menée en backcross, le test de l'analyse de variance revient à faire un test de comparaison de moyenne de performances entre les individus homozygotes et les individus hétérozygotes au marqueur.

Pour une population F_2 avec trois phénotypes au marqueur différent (homozygotes 1 / hétérozygotes / homozygotes 2), le modèle de l'analyse de variance s'écrit de la façon suivante :

$$Y = \mu + a + e \quad ; \quad Y = \mu - a + e \quad \text{ou} \quad Y = \mu + d + e$$

selon que l'individu a reçu respectivement les allèles M1M1, M2M2 ou M1M2 au génotype du marqueur (M1 et M2 sont les allèles au marqueur). Y est la performance de l'individu, a est l'effet additif du marqueur, d est l'effet de dominance entre les deux allèles et e est la résiduelle du modèle.

L'hypothèse $H_0 = \{a = d = 0\}$ qui correspond à l'absence de QTL est donc testée grâce à la statistique F du modèle. L'hypothèse alternative est, $H_1 = \{a \neq 0$ et/ou $d \neq 0\}$ est donc définit par « le QTL a un effet additif ou bien un effet dominant ou bien les deux ».

Si l'on suppose maintenant que le taux de recombinaison r entre M et Q n'est pas nul, un parent portant sur le même chromosome les allèles M1 au marqueur et Q1 au QTL associé (se trouvant au voisinage de ce dernier) et qui transmet l'allèle M1 à un descendant ne transmet pas automatiquement l'allèle Q1 au QTL à ce même individu. Les probabilités conditionnelles des génotypes au QTL, sachant le génotype au marqueur, sont des fonctions du taux de recombinaison r. A taille d'échantillon constant, la puissance du test chute considérablement quand ce taux r augmente.

Pour les animaux d'élevage (ou lignées « outbred »), il faut prendre en compte la notion de famille. L'effet du marqueur est donc mesuré intra famille de père, pour les protocoles parentaux (Soller et genizi, 1978) ou intra famille de grand-père, pour les protocoles grand-parentaux (Weller *et al.,* 1990). Pour ces derniers, est pris en compte, en plus de l'effet grand-père, l'effet fils.

Le modèle linéaire s'écrit pour un protocole père :

$$Y_{ijk} = \mu + s_i + a_{ij} + e_{ijk}$$

où Y_{ijk} est la performance du k^{ème} descendant du i^{ème} père possédant le génotype j au marqueur. μ est la moyenne générale des performances, s_i l'effet du père i, a_{ij} est l'effet du marqueur j intra famille de père i (j=1 ou 2, selon que l'allèle reçu du père est le premier ou le second) et e_{ijk} la résiduelle intra-marqueur intra-famille.

2.3.2.2 Approche par maximum de vraisemblance

Dans l'approche par maximum de vraisemblance, on suppose que le caractère quantitatif est distribué normalement dans chaque classe de génotype. Nous développerons ici le cas particulier de croisement entre lignées homozygotes. Dans ce cas, il y a trois classes génotypiques dans le cas d'une F_2 et deux dans le cas d'un Backcross. La vraisemblance des observations, en utilisant les informations apportées par le marqueur s'écrit (Luo et Kearsey 1989) :

$$L = \prod_i f_{M_1M_2}(y_i) \prod_j f_{M_1M_1}(y_j) \prod_k f_{M_2M_2}(y_k)$$

Avec $f_{M\alpha M\beta}$ (y_i) étant la vraisemblance de l'observation faite sur le i^{ème} individu de phénotype au marqueur MαMβ.

Du fait de la recombinaison, chaque classe de génotype au marqueur est un mélange de classes de génotypes au QTL. On peut donc écrire chaque fonction $f_{M\alpha M\beta}$ comme une combinaison linéaire de densités gaussiennes correspondant aux différents génotypes possibles au QTL.

Les coefficients de cette combinaison sont les probabilités conditionnelles P(QiQj/MαMβ) (tableau 1.3) avec, par exemple, pour une population F2, (i,j) $\in$ {(1,1) ;(1,2) ;(2,2)}. Ces probabilités conditionnelles sont fonction de la recombinaison r. On peut alors estimer tous les paramètres du modèle.

Tableau 1.3 : Génotype au QTL conditionnellement à l'information au marqueur (Mackinnon et Georges, 1992)

Génotype au QTL	Génotype au marqueur		
	M1M1	**M1M2**	**M2M2**
Q1Q1	½(1-r)ps	½rps+½(1-r)p(1-s)	½rp(1-s)
Q1Q2	½(1-r)(1-p)s+½rps	½s[(1-r)p+r(1-p)]+½(1-s)[(1-r)(1-p)+rp]	½(1-r)p(1-s)+½r(1-p)(1-s)
Q2Q2	½r(1-p)s	½(1-r)(1-p)s+½r(1-p)(1-s)	½(1-r)(1-p)(1-s)

Il s'agit d'une population descendant d'un père portant l'haplotype M1Q1/M2Q2, accouplé avec des femelles en équilibre de liaison entre marqueur et QTL. p et s sont respectivement la fréquence de l'allèle au QTL Q1 et celle de l'allèle au marqueur M1 dans la population mère.

Cette méthode a pour avantage de ne pas nécessiter l'existence d'une carte génétique complète. Son principal défaut est de ne pas permettre de différentier entre un QTL à effet relativement fort mais loin du marqueur d'un QTL très proche du marqueur avec un effet faible.

2.3.3 La cartographie d'intervalle

2.3.3.1 Principe général (Lander et Botstein 1989)

Lorsqu'une carte génétique est disponible, il est possible d'envisager une démarche de recherche des QTLs sur l'ensemble du génome. L'analyse se fait par groupe de liaison. L'idée générale est de tester l'hypothèse d'absence de QTL en chaque position x du génome. Lander et Botstein (1989) ont proposé la cartographie d'intervalle (ou Interval Mapping) dans le cas d'une population de backcross : cette méthode considère les marqueurs deux à deux définissant ainsi des intervalles successifs sur l'ensemble du génome. Deux marqueurs adjacents étant liés avec un taux de recombinaison r, on teste si un QTL est situé entre ces deux marqueurs. Il s'agit à la fois d'en trouver la position la plus probable, d'en estimer l'effet et de tester sa significativité. Cette technique qui prend en compte deux marqueurs simultanément est plus puissante et robuste que l'analyse uni marqueur (Haley *et al.,* 1994).

Dans ce modèle, la normalité des distributions des performances est supposée. Les génotypes au QTL sont inconnus et le maximum de vraisemblance est exprimé conditionnellement aux génotypes des marqueurs flanquants. Il s'agit donc d'un problème de mélange de distributions. En chaque point de l'intervalle, les proportions du mélange sont connues et dépendent de la distance de ce point à ses marqueurs flanquants.

La cartographie d'intervalle peut être décrite simplement dans le cas d'un BC entre lignées consanguines, tel qu'envisagé par Lander et Botstein (1989). Soient M et N deux marqueurs flanquants une position QTL à tester. Soient M1N1 et Q1 les allèles aux marqueurs flanquants et au QTL dans la première lignée, M2N2 et Q2 dans la deuxième. Le père F1 est M1N1M2N2 et la mère M1N1M1N1. On souhaite tester l'existence d'un QTL à la position x du génome situé entre les marqueurs M et N à d_M centimorgan du premier et d_N centimorgan du second. Une fonction de distance étant choisie, aux valeurs d_M et d_N correspondent des taux de recombinaison r_M et r_N entre le QTL et les marqueurs M et N respectivement. Connaissant le génotype aux

marqueurs d'un descendant, on peut calculer les probabilités selon lesquelles ce dernier a reçu Q1 ou Q2 au QTL putatif à la position testée en considérant les recombinaisons éventuelles.

Par exemple :

$$P(Q1Q2/M1N1M1N1) = \frac{P(Q1Q2\ et\ M1N1M1N1)}{P(M1N1M1N1)}$$

$$= \frac{\frac{1}{2} r_M r_N}{\frac{1}{2}(1-r)} = \frac{r_M r_N}{1-r}$$

r étant le taux de recombinaison entre M et N.

r = P (recombinaison entre M et Q ∩ pas de recombinaison entre Q et N) + P (pas de recombinaison entre M et Q ∩ recombinaison entre Q et N)

En l'absence d'interférence,

$r = r_M (1 - r_N) + r_N (1 - r_M)$ donc $r = r_M + r_N - 2 r_M r_N$.

En effet, dans ce cas, le descendant qui a reçu la portion chromosomique M1N1 de ses deux parents appartient à la classe génotypique Q1Q2 au QTL si son parent F1 (disons le père), triple hétérozygote M1Q1N1/M2Q2N2, lui a transmis Q2 : ceci nous permet de dire qu'il y a eu chez ce parent un double crossing-over, l'un entre M et Q et l'autre entre Q et N. De plus un descendant donné a une chance sur deux de recevoir l'allèle M1 au marqueur M de son père ce qui explique la valeur du numérateur de la probabilité ci-dessus. De la même manière pour que l'individu soit M1N1/M1N1 aux marqueurs, il ne faut pas qu'il y ait de recombinaison visible entre les deux marqueurs d'où le (1-r) et le ½ vient du fait que chaque descendant a une chance sur deux de recevoir l'allèle M1 de son père.

On calcule de la même manière les autres probabilités. Ainsi

$$P(Q1Q1/M1N1M1N1) = 1 - P(Q1Q2/M1N1M1N1) = 1 - \frac{r_M r_N}{1-r}$$

Pour tester l'existence du QTL, on utilise une méthode basée sur la fonction de vraisemblance. Dans l'exemple d'une population de backcross, la vraisemblance, fonction de la position du QTL x, de la variance σ^2 et de moyennes phénotypiques conditionnelles μ et μ+Δ, Δ étant l'effet de substitution au QTL, s'écrit :

$$L(\mu, \sigma^2, \Delta) = \prod \left\{ \left\{ P(d_i^x = 1/M_i N_j).f(y_i/d_i^x = 1) \right\} + \left\{ P(d_i^x = 2/M_i N_j).f(y_i/d_i^x = 2) \right\} \right\}$$

L'événement d^x_i=q (q=1 ou 2 dans le cas d'un backcross) veut dire que l'individu i a reçu de son père F1, à la position x, un segment chromosomique de la population grand-parentale q et, aux marqueurs flanquants Mi et Nj, f (y_i/ d^x_i=1) et f (y_i/ d^x_i=2) sont les fonctions de densité supposées normales de la variable aléatoire y_i, dans les populations de génotype Q1Q1 et Q1Q2 au QTL respectivement.

f(y_i/ d^x_i=1) est donc une normale de moyenne μ et de variance σ^2 et f(y_i/ d^x_i=2) une normale de moyenne μ+Δ et de variance σ^2.

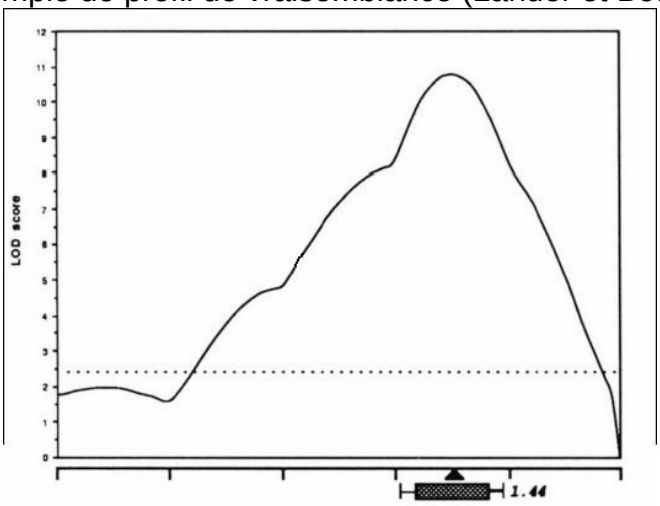
On teste l'hypothèse H1 : « existence de QTL ou Δ≠0 » à la position x contre H0 : « absence de QTL c'est-à-dire Δ=0 ». Pour cela on peut faire un rapport de vraisemblance (LRT) ou utiliser le test de Lod Score (LOD) :

$$LRT = 2ln\left(\frac{L(\hat{\mu},\hat{\Delta},\hat{\sigma}^2)}{L(\hat{\mu},0,\hat{\sigma}^2)}\right) \text{ et } LOD = \log\left(\frac{L(\hat{\mu},\hat{\Delta},\hat{\sigma}^2)}{L(\hat{\mu},0,\hat{\sigma}^2)}\right)$$

Le profil de la vraisemblance décrivant la valeur du rapport de vraisemblance en toutes les positions est ainsi obtenu. Par ailleurs, un seuil de rejet est prédéterminé (détails dans le paragraphe 2.4 de ce chapitre). Au dessus de ce seuil, l'hypothèse H0 (absence de QTL) est rejetée et la localisation la plus probable dans le groupe de liaison étudié est estimée à la position du maximum du rapport de vraisemblance (figure 1.7).

La densité du marquage n'est pas tellement importante pour la détection des QTLs. Darvasi *et al*. (1993) ont montré que la puissance de détection et la précision de la localisation du QTL dépendait principalement de l'effectif de la population.

Figure 1.7 : Exemple de profil de vraisemblance (Lander et Botstein, 1989)



2.3.3.2 Méthode de régression linéaire : le cas d'une population F2 (Haley et Knott, 1992)

Il y a 20 ans les méthodes basées sur la vraisemblance étaient relativement coûteuses en temps de calcul, ce qui constituait un obstacle pour des analyses systématiques. Pour remédier à ce problème, une méthode de régression linéaire multiple pour détecter les QTLs fut proposée par Haley et Knott (1992) et Martinez et Curnow (1992). Cette méthode est asymptotiquement équivalente à celle de Lander et Botstein (1989) (Rebaï *et al.*, 1995) sauf dans le cas de QTLs à effets très faibles (Haley et Knott en 1992). Cette approche correspond à une linéarisation de la vraisemblance (Knapp *et al.*, 1990). Les performances moyennes intra classes génotypiques au QTL sont estimées.

Pour illustrer cette approche, considérons une population F2, et deux marqueurs flanquants M et N, avec (**μ-a**) la moyenne des performances des individus Q1Q1 au QTL, (**μ+d**) la moyenne des performances des individus Q1Q2 et (**μ+a**) la moyenne des performances des individus Q2Q2 ; **μ** étant la moyenne des performances des

individus homozygotes aux QTL, **a** la déviation due à l'effet additif du QTL et **d** la déviation due à l'effet de dominance. La distance entre N et M est connue grâce à la carte génétique et donc le taux de recombinaison **r** entre ces deux marqueurs l'est aussi grâce aux formules de Haldane ou de Kosambi inversées. Si r_M et r_N sont respectivement le taux de recombinaison entre M et Q (QTL situé entre les deux marqueurs) et le taux de recombinaison entre N et Q, on établit donc sous l'hypothèse de non interférence, la relation $r = r_M + r_N - 2r_M r_N$.

En considérant que l'espérance du phénotype d'un individu est linéairement dépendante de l'effet du QTL sachant l'information aux marqueurs :

$$E(Y_i) = P(d_i^x = 1/M_i N_j).(\mu + a) + P(d_i^x = 2/M_i N_j).(\mu + d) + P(d_i^x = 3/M_i N_j).(\mu - a)$$
$$E(Y_i) = \mu + P(d_i^x = 1/M_i N_j).a + (d_i^x = 2/M_i N_j).d - P(d_i^x = 3/M_i N_j).a$$

L'événement d^x_i=q (q=1, 2 ou 3 dans le cas d'une F2) signifie que l'individu i a reçu à la position x un segment chromosomique de la population grand-parentale q et MiNj correspond à l'information aux marqueurs flanquants. Ces probabilités sont des fonctions des taux de recombinaison r, r_M et r_N, facilement calculables, par exemple si l'individu en question est M1M1N1N1 aux marqueurs :

$$P\,(d_i^x = 1\,/\,M1M1N1N1) \;=\; P(Q1Q1\,/\,M1M1N1N1)$$
$$= \frac{P(Q1Q1\;et\;M1M1N1N1)}{P(M1M1N1N1)} \;= \frac{\boldsymbol{(1 - r_M)^2(1 - r_N)^2}}{\boldsymbol{(1 - r)^2}}$$

et

$$P\,(d_i^x = 3\,/\,M1M1N1N1) \;=\; P(Q2Q2\,/\,M1M1N1N1)$$
$$= \frac{P(Q2Q2\;et\;M1M1N1N1)}{P(M1M1N1N1)} \;= \frac{\boldsymbol{r_M^2 r_N^2}}{\boldsymbol{(1 - r)^2}}$$

Donc le coefficient **a** dans la régression multiple est :

$$\left[\frac{(1 - r_M)^2(1 - r_N)^2}{(1 - r)^2}\right] - \left[\frac{r_M^2 r_N^2}{(1 - r)^2}\right] = \frac{[(1 - r_M)^2(1 - r_N)^2 - r_M^2 r_N^2]}{(1 - r)^2}$$

Les différents coefficients **a** et **d** pour les neuf groupes de génotypes aux marqueurs flanquants possibles dans une population de F_2 sont retranscrits sur le tableau 1.4, extrait de Haley et Knott (1992).

Il ne reste donc plus qu'à estimer, grâce à la méthode de régression linéaire multiple, les différents paramètres μ, a et d en chaque position de l'intervalle défini par les deux marqueurs flanquants.

Tableau 1.4 : Espérance des coefficients des composantes de a et d pour toutes les combinaisons de génotypes possibles dans une population F2. D'après Haley et Knott, 1992.

Génotype aux marqueurs	Coefficient de a (déviation due à l'effet additif du QTL)	Coefficient de d (déviation due à l'effet dominant du QTL)
M1M1N1N1	$\frac{(1-r_M)^2(1-r_N)^2 - r_M^2 r_N^2}{(1-r)^2}$	$\frac{r_M(1-r_N)r_N(1-r_M)}{(1-r)^2}$
M1M1N1N2	$\frac{(1-r_M)^2 r_N(1-r_N) - r_M^2 r_N(1-r_N)}{r(1-r)}$	$\frac{r_M(1-r_M)(1-r_N)^2 + r_M(1-r_M)r_N^2}{r(1-r)}$
M1M1N2N2	$\frac{(1-r_M)^2 r_N^2 - (1-r_N)^2 r_M^2}{r^2}$	$\frac{2r_M(1-r_M)r_N(1-r_N)}{r^2}$
M1M2N1N1	$\frac{r_M(1-r_M)(1-r_N)^2 - r_M(1-r_M)r_N^2}{r(1-r)}$	$\frac{(1-r_M)^2 r_N(1-r_N) + r_M^2 r_N(1-r_N)}{r(1-r)}$
M1M2N1N2	0	$\frac{r_M^2 r_N^2 + r_M^2(1-r_N)^2 + (1-r_M)^2 r_N^2 + (1-r_M)^2(1-r_M)^2}{r^2 + (1-r)^2}$
M1M2N2N2	$\frac{r_M(1-r_M)r_N^2 - r_M(1-r_M)(1-r_N)^2}{r(1-r)}$	$\frac{(1-r_M)^2 r_N(1-r_N) + r_M^2 r_N(1-r_N)}{r(1-r)}$
M2M2N1N1	$\frac{r_M^2(1-r_N)^2 - r_N^2(1-r_M)^2}{r^2}$	$\frac{2r_M(1-r_M)r_N(1-r_N)}{r^2}$
M2M2N1N2	$\frac{r_M^2 r_N(1-r_N) - (1-r_M)^2 r_N(1-r_N)}{r(1-r)}$	$\frac{r_M(1-r_M)(1-r_N)^2 - r_M(1-r_M)r_N^2}{r(1-r)}$
M2M2N2N2	$\frac{r_M^2 r_N^2 - (1-r_M)^2(1-r_N)^2}{(1-r)^2}$	$\frac{2r_M(1-r_M)r_N(1-r_N)}{(1-r)^2}$

2.3.3.3 La cartographie d'intervalle appliquée à une population « outbred »

La cartographie d'intervalle peut s'étendre à des populations « outbred ». Les expressions de la vraisemblance du modèle y deviennent plus complexes que dans le cadre d'un protocole idéal (croisement de lignées « inbred »), du fait de la nécessité de prendre en compte la structure familiale et de la non-informativité de marqueurs chez certaines familles.

Le modèle proposé par Le Roy *et al.* (1998) dans le cadre d'un mélange de plein-frères et de demi-frères s'appuie sur le test de rapport de vraisemblance tel que décrit par Lander et Botstein (1989).

On se place dans une population de n familles de père (i=1,…,n) avec n_i accouplements par père i (j=1,…,n_i) et n_{ij} descendants par mère ij. Le nombre total d'animaux composant la population est donc de $N = \sum_{i=1}^{n} \sum_{j=1}^{n_i} n_{ij}$. Les pères et mères sont supposés non apparentés et une mère est accouplée avec un seul père. Si on considère, pour réduire le temps de calcul nécessaire, la phase la plus probable du père, après reconstitution des phases des parents (Elsen *et al.*, 1999 ; Goffinet *et al.*, 1999 ; Mangin *et al.*, 1999), la vraisemblance linéarisée intra-famille de plein-frères, pour une position x donnée, s'écrit (selon Le Roy *et al.*, 1998) :

$$V^x = \prod_i \prod_j \sum_{hd_{ij}} P(hd_{ij}/M_i, \hat{hs}_i) \prod_k f(yp_{ijk}/\hat{hs}_i, hd_{ij}, M_i)$$

où M_i est l'information marqueur pour la famille du père i, $hs_i = \{hs_{ip}, hs_{im}\}$, est la phase du père i, avec hs_{ip} étant l'haplotype du segment chromosomique reçu par le

père i par la voie paternelle et hs_{im} l'haplotype du segment reçu par voie maternelle. L'information marqueur provenant des grands-parents peut aider à déterminer la probabilité des différentes phases possibles. Cette dernière n'étant pas toujours disponible, l'information venant des descendants et des accouplements est aussi utilisée pour déterminer ces probabilités. $\hat{hs}_i$ correspond à la phase des marqueurs génétiques la plus probable pour le père i. hd_{ij} est la phase des marqueurs génétiques pour la mère ij. Dans la pratique, la reconstitution des phases des parents est souvent préliminaire à la mise en œuvre des statistiques de test. Enfin, yp_{ijk} est l'observation phénotypique du descendant ijk (k=1,…,n) pour le caractère quantitatif étudié.

La fonction de pénétrance f est une fonction normale qui s'écrit :

$$f(yp_{ijk}/\hat{h}s_i, hd_{ij}, M_i) = \frac{1}{\sqrt{2\pi}\sigma_i} exp\left(-\frac{1}{2}\left(\frac{yp_{ijk} - \mu_{ijk}}{\sigma_i}\right)^2\right)$$

σ_i étant la variance intra-famille de père i, pour le caractère étudié (Goffinet *et al.*, 1997) et $\mu_{ijk} = \sum_{qs}\sum_{qd} P(d^x_{ijk} = (q_s, q_d)/\hat{h}s_i, hd_{ij}, M_i)(\mu_{iqs} + \mu_{iqd})$.

$P(d^x_{ijk} = (q_s, q_d)/\hat{h}s_i, hd_{ij}, M_i)$ étant la probabilité que le descendant ijk ait reçu de son père i le segment chromosomique q_s à la position x testée (q_s=1 venant du grand-père paternel, q_s=2 venant de la grand-mère paternelle) et qu'il ait reçu de sa mère ij le segment chromosomique q_d à la position testée x (q_d=1 venant du grand-père maternel, q_d=2 venant de la grand-mère maternelle). μ_{iqs} est l'effet du segment q_s reçu du père sur le caractère quantitatif, et μ_{iqd} l'effet pour le segment q_d reçu de la mère ij.

Si μ_i est la moyenne des performances des descendants du père i et α_i l'effet de substitution moyen de QTL estimé intra-famille de père i pour le caractère quantitatif étudié, $\mu_{iqs} = \begin{cases} \mu_i + \frac{1}{2}\alpha_i \; si \; q_s = 1 \\ \mu_i - \frac{1}{2}\alpha_i \; si \; q_s = 2 \end{cases}$.

Le même raisonnement est valable pour le côté maternel.

2.3.4 Le modèle à effet aléatoire pour les pedigrees complexes

Pour ce type de population, le premier modèle à effet aléatoire proposé, est celui de Fernando et Grossman en 1989. Dans ce modèle, la valeur génétique additive d'un animal est décomposée en la somme d'un effet polygénique et de la somme des effets aux QTL supposés additifs (effet de l'allèle au QTL d'origine paternelle + effet de l'allèle d'origine maternelle).

Pour un animal i le modèle s'écrit :

$$y_i = x_i\beta + u_i + v_i + e_i = x_i\beta + u_i + v_{ip} + v_{im} + e_i$$

y_i est la performance de l'individu i,
x_i est la $i^{ème}$ ligne de la matrice d'incidence reliant chaque performance à un individu,

β est le vecteur des effets fixes,
u_i est la valeur polygénique additive,
v_{ip} est l'effet additif de l'allèle paternel du QTL,
v_{im} est l'effet additif de l'allèle maternel du QTL
et e_i est la résiduelle du modèle.

On suppose que : $\mathrm{var(u)} = \mathbf{A}\sigma_u^2$, $\mathrm{var(v)} = \mathbf{G}\sigma_v^2$, $\mathrm{var(e)} = \mathbf{R}\sigma_e^2$, où σ_u^2, σ_v^2 et σ_e^2 sont, respectivement, la variance polygénique, la variance additive du QTL et la variance résiduelle. A est la matrice des relations de parenté, $G\sigma_v^2$ la matrice des variances-covariances des effets additifs du QTL conditionnellement à l'information des marqueurs et R sera prise ici comme la matrice identité.

La cartographie de QTL est réalisée par une estimation des composantes de variance en deux étapes (Georges *et al.,* 2000). Dans un premier temps la matrice G est construite grâce aux probabilités IBD (Identity By Descent ou identité par descendance) des segments chromosomiques considérés. En effet, analogiquement aux relations de parenté entre individus, on définit la parenté entre segments chromosomiques, comme la probabilité qu'ils soient identiques par descendance, c'est-à-dire qu'ils dérivent d'un même segment ancêtre. Différentes méthodes, pour estimer les probabilités IBD ont été proposées (Van Arendonk *et al*., 1994 ; Wang *et al*., 1995 ; Almasy et Blangero, 1998). Ils en existent trois types :

- ***les méthodes utilisant des algorithmes récursifs***

Ces algorithmes permettent de calculer les probabilités IBD de façon récursive en suivant un ordre chronologique (van Arendonk, 1994 ; Wang *et al.,* 1995). Leur principal défaut est leur incapacité à gérer des grandes quantités de données manquantes (Pong-Wong et *al.,* 2001). En effet, ces algorithmes suivant une stratégie de haut en bas, les données manquantes chez des individus en haut du pedigree impliquent des erreurs d'estimation dans tout le reste du pedigree.

- ***les méthodes utilisant des algorithmes basés sur les corrélations***

Ces algorithmes (Almasy et Blangero, 1998) utilisent la matrice de corrélations (Amos, 1994) entre proportion d'allèles IBD partagée entre marqueurs complètement génotypés et QTL. Pour ces algorithmes, les coefficients IBD sont difficiles à estimer pour les pédigrées complexes et avec des données manquantes.

- ***les méthodes utilisant des algorithmes basés sur des simulations***

Les deux types d'algorithmes précédents ne sont pas directement applicables pour des pédigrées où les génotypes aux marqueurs ne sont pas complètement renseignés. Une solution est d'utiliser l'approximation MCMC (Markov Chain Monte Carlo) pour calculer l'espérance de la matrice IBD. Les méthodes MCMC sont adaptées aux pédigrées complexes avec des données manquantes mais sont pénalisées par la lourdeur de leurs calculs.

La deuxième étape est l'estimation des paramètres QTL grâce au modèle linéaire mixte.

2.4 Statistique de test et précision de la position du QTL

2.4.1 Introduction

Quelle que soit la méthode utilisée, pour la cartographie de QTL, les deux hypothèses testées sont les suivantes : -H0 : absence de QTL à la position testée contre H1 : Présence d'un QTL en ségrégation à la position testée. Une valeur de la statistique de test qui correspond par exemple au rapport des vraisemblances sous les deux hypothèses est donc obtenue pour chaque position testée. La position correspondant à la valeur maximale de la statistique de test indique la position d'un éventuel QTL. Dans un premier temps il faut décider si oui ou non cette valeur est significativement élevée conditionnellement à un risque de se tromper α (souvent fixé à 5 ou 1% - ce risque pouvant être fixé au niveau chromosomique ou au niveau du génome), pour considérer la présence ou non d'un QTL à cette position. Deux sortes d'erreurs sont possibles : 1- conclure à la présence d'un QTL en ségrégation alors qu'il n'y en a pas, i.e. rejeter à tort l'hypothèse nulle, donc déclarer un « faux positif », c'est le risque de première espèce ou 2- conclure qu'il n' y a pas de QTL alors qu'il y en a un, i.e. accepter à tort l'hypothèse nulle, donc déclarer un « faux négatif », c'est le risque de deuxième espèce.

Si la valeur maximale du test est « assez » élevée pour que la position testée soit considérée comme la position la plus probable du QTL, il faudra encadrer cette position par un intervalle de confiance.

Les méthodes qui permettent de déterminer le seuil de significativité et la précision de la position du QTL seront présentées dans ce paragraphe.

2.4.2 Détection avec un seul marqueur

Pour la détection de QTL avec un seul marqueur, le test utilisé est celui de Student. En effet, il s'agit de comparer la moyenne de deux sous-populations ayant reçu l'un ou l'autre des allèles aux marqueurs par leur parent hétérozygote. On utilise donc le seuil classique du test T pour déterminer s'il est significatif au niveau voulu (souvent 5%). Pour des populations plus compliquées, un test de Fisher permet de tester l'égalité des moyennes de plusieurs sous-populations (dans le cas d'une population d'élevage avec un protocole plein-frères par exemple). Cela revient à effectuer une analyse de variance avec comme caractère à expliquer les performances des individus et comme caractère explicatif l'allèle au marqueur reçu.

2.4.3 Détection avec plusieurs marqueurs : le cas de la cartographie d'intervalle

Comme précédemment détaillé, différents tests statistiques (test du Log du rapport de vraisemblances, régression, …) peuvent être mis en œuvre pour réaliser une cartographie d'intervalle. Les questions qui se posent sont de savoir comment déterminer le seuil de ce test ainsi que la région de confiance du QTL détecté.

2.4.3.1 Le seuil de significativité

Le test étant calculé sur tout le chromosome, à intervalle régulier (souvent tous les 1 ou 2 cM), nous disposons d'une courbe représentant le test en chaque point du chromosome. Il faut déterminer la valeur du seuil s pour un risque α de première espèce, telle que la probabilité de dépasser ce seuil s en un point donné du chromosome soit de α lorsqu'il n'y a pas de QTL sur ce chromosome. Il faut aussi considérer que le test sera mis en œuvre sur l'ensemble du génome : afin de garantir un niveau d'erreur global de α pour détecter un QTL sur l'ensemble du génome constitué de n chromosomes alors qu'il n'y en a pas, on peut prendre un niveau d'erreur de α/n (test de Bonferroni) pour chaque chromosome.

Dans le choix d'un seuil qui garantisse un niveau d'erreur de α (ou α/n) pour chaque chromosome, la difficulté vient de la non indépendance des statistiques de test aux différentes positions. Les statistiques suivent un processus qui possède un extremum dont la loi de distribution, complexe, n'est pas celle du test en un point. Le seuil de rejet s doit être tel que la probabilité que le maximum des tests lui soit supérieur soit égale à α. Des approximations sont proposées par Lander et Botstein (1989), correspondant à des situations asymptotiques.

D'autres moyens de déterminer ce seuil de test existent tels que l'utilisation de simulations de Monte Carlo ou la méthode de rééchantillonnage. Dans la simulation de Monte Carlo, on simule des variables aléatoires qui représentent les performances des individus sous l'hypothèse nulle d'absence de QTL, en conservant les informations relatives aux marqueurs pour chaque individu, et on calcule la valeur de test pour ces observations simulées. On calcule alors un test T le long du chromosome, dont on déduit la plus grande valeur sup(T). En faisant un grand nombre de fois ce calcul, on obtient la distribution de sup(T) et il suffit de prendre le quantile empirique de cette statistique pour avoir le seuil.

La stratégie de rééchantillonage (Churchill et Doerge. 1994) est similaire à celle de Monte Carlo. Au lieu de tirer des performances selon la distribution liée au modèle, on affecte, pour chacun des individus (défini par sa généalogie et ses génotypes marqueurs), une performance tirée au hasard parmi les observations réellement obtenues. On se place ainsi sous l'hypothèse d'absence de QTL puisque les performances ne sont plus associées aux allèles des marqueurs. L'intérêt de cette méthode est qu'on a une distribution très proche de la distribution réelle des données.

2.4.3.2 L'intervalle de confiance

Une fois le QTL (lorsqu'on considère l'hypothèse qu'il n'y a qu'un seul QTL) détecté et sa position la plus probable sur le génome déterminée, il est nécessaire de compléter ces informations par la construction d'un intervalle de confiance autour de la position x du QTL identifié.

Le principe général de la construction d'un intervalle de confiance I pour la position du QTL x est de déterminer les bornes inférieure et supérieure de I telles que :

$P(x \in I) = 1-\alpha$ pour α fixé entre 0 et 1. 1- α est le niveau de confiance.

En pratique, grâce à la théorie asymptotique classique, pour tout point x (position du QTL putatif) on considère $R = \sup(T)-T(x)$ qui suit un χ^2 à un degré de liberté. L'intervalle de confiance retenu est celui qui contient la position x et dans lequel tous

les autres points sont tels que R soit inférieur au α quantile d'un χ^2 à un degré de liberté. Habituellement, l'intervalle de confiance est construit en retranchant la valeur du χ^2 (3,84 pour un niveau de risque de 5%) à la valeur du maximum du test T et en mettant dans l'intervalle toutes les positions voisines du maximum telles que la valeur du test T est supérieure à cette différence sup (T) - valeur du χ^2.

A noter que Vischer *et al*. (1996) proposent pour calculer l'intervalle de confiance, une méthode non paramétrique (bootstrapping) qui consiste dans un premier temps à faire n tirages avec remise des individus (avec toute leur information). On effectue ensuite autant d'analyses de cartographie de QTL avec les données ainsi obtenues en gardant la généalogie et l'information aux marqueurs. A chaque fois, on retient la valeur du maximum du test statistique. L'intervalle de confiance à 100(1-α)% est obtenu en triant ces derniers et gardant les positions associées aux percentiles extrêmes de la distribution.

3 Améliorations des méthodes de détection de QTL

Dans ce paragraphe nous présentons les différentes améliorations ou raffinements apportés aux méthodes de détection de QTL classiques. Dans un premier temps, sont présentées les méthodes basées sur des fonctions de pénétrances particulières (autre que la fonction de la loi normale) comme la prise en compte de plusieurs caractères à la fois. Dans un second temps, sont abordées les méthodes qui permettent de considérer plusieurs QTLs qui agissent en même temps sur un caractère donné, avant de revenir sur le dispositif petites-filles qui permet de renforcer la puissance de détection de QTL.

3.1 Autres formes de pénétrance

3.1.1 L'approche multi-caractères

Les méthodes de détection de QTL multi-caractères ont été proposées pour répondre à un certain nombre de questions posées par l'utilisation, en détection de QTL, des méthodes dites uni-caractères. En effet, la détection de QTL est souvent réalisée caractère par caractère sans tenir compte des éventuelles corrélations entre ces derniers. Ne serait-il pas plus efficace de regrouper les informations de tous ces caractères d'intérêt dans un même test ? Ce test, pour une position donnée, consisterait à comparer l'hypothèse nulle : "il n'y a pas de QTL agissant sur les caractères étudiés" à l'hypothèse alternative : "il y a un QTL agissant sur au moins un caractère". Ces méthodes permettraient d'améliorer la précision de la détection de QTL par la prise en compte (en plus de la minimisation de la variance résiduelle de chaque caractère) des covariances entre les caractères du modèle.

Pour cette approche multi-caractères, Ronin *et al.* (1995) ont défini une méthode de rapport de vraisemblance dérivée de la méthode uni-caractère, uni-marqueur.

Le principe de cette méthode développée pour des populations « inbred » est de modéliser la coségrégation des performances de deux caractères x et y corrélés, en fonction des génotypes au QTL. Les performances sont supposées suivre intra-classe génotypique des distributions binormales. Si on note S_i la distribution espérée des performances pour chaque classe génotypique au marqueur génétique i testé, on a :

$$S_i(x,y) = \sum_{j=1}^{3} \pi_{ij}(r)\, f_j(x,y)$$

avec $\pi_{ij}(r)$ étant les proportions d'individus de classe génotypique j (j=1 si QQ au QTL ; j=2 si Qq ; j=3 si qq) au QTL dans la classe i au marqueur, r est le taux de recombinaison entre les deux locus et :

$$f_j(x,y) = \frac{1}{\sqrt{2\pi\sigma_{jx}\sigma_{jy}(1-R_j^2)}} \exp\left(-\frac{1}{2(1-R_j^2)} \left[\frac{(x-\mu_{jx})^2}{\sigma_{jx}^2} - 2R_j \frac{(x-\mu_{jx})(y-\mu_{jy})}{\sigma_{jx}\sigma_{jx}} + \frac{(x-\mu_{jy})^2}{\sigma_{jy}^2} \right] \right)$$

la distribution conjointe de x et y pour le génotype j au QTL, avec μ_{jx}, μ_{jy}, σ_{jx}, σ_{jy} les moyennes et les écart-types des caractères x et y pour le génotype j et Rj est le coefficient de corrélation intra classe génotypique j.

Korol *et al.* (1995) ont développé, à partir des travaux de Ronin *et al*. (1995), la méthode de cartographie d'intervalle appliquée aux analyses multi-caractères. Ces auteurs montrent que les analyses multi-caractères permettent d'améliorer la précision d'estimation des paramètres, et ce d'autant plus que la corrélation résiduelle entre les caractères est élevée. Les puissances de détection sont donc augmentées par la prise en compte des distributions conjointes des caractères.

La généralisation de ces méthodes, d'une part pour des populations « outbred » où la prise en compte des structures familles est nécessaire, d'autre part pour un nombre quelconque de caractères (en considérant une fonction multi-normale) entraine une augmentation conséquente des paramètres à estimer. Pour remédier à ce problème, Weller *et al. (*1996) ont proposé de réaliser une analyse en composantes principales (ACP) sur les données phénotypiques pour synthétiser l'information sur les liaisons entre les caractères dans une combinaison linéaire des performances, puis de réaliser une analyse de type uni-caractère marqueur par marqueur, en considérant la variable synthétique ainsi trouvée. Cette approche a été étendue pour la cartographie d'intervalle par Mangin *et al. (*1998).

Plus récemment, Gilbert et Le Roy (2003 ; 2007), en s'inspirant de l'idée de Weller *et al. (*1996), ont utilisé l'analyse discriminante pour construire une combinaison linéaire de caractères qui sera ensuite analysée dans une approche uni-QTL ou multi-QTL. Ces méthodes sont développées pour des populations « outbred » (mélange de demi-frères et de plein frères).

Quand un grand nombre de caractères sont enregistrés, il convient de restreindre la multi dimensionnalité en ne pratiquant les analyses multi-caractères que sur des groupes à choisir. Ces groupements peuvent correspondre à des caractères pour lesquels des QTLs ont été détectés au même endroit ou des caractères corrélés phénotypiquement.

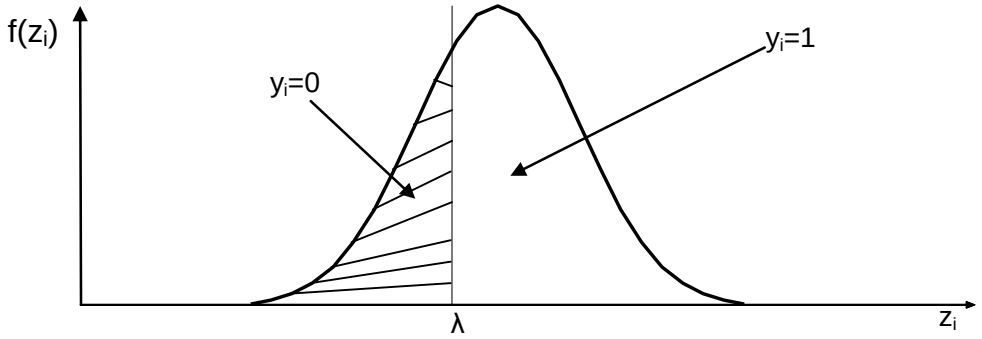
3.1.2 Etude des caractères discrets ou caractères « à seuil »

Les méthodes de détection de QTL décrites jusqu'ici supposent la normalité de la distribution des caractères étudiés. Cette hypothèse est restrictive dans beaucoup de cas (données discrètes, données de survie, …) par exemple la qualité du foie gras exprimée par une note allant de 1 à 5.

Nous développons ici le cas des caractères « à seuil » ou discrets qui ne présentent qu'un nombre limité de valeurs ordonnées.

Le premier modèle prenant en comptent des caractères « à seuil » a été proposé par Wright (1934), pour les caractères binaires. Dans ce modèle, Wright considérait une distribution normale (z_i) de moyenne μ et de variance σ^2 sous-jacente à la distribution binaire observée. Il définit un seuil λ tel que l'observation y_i est 0 si la sous-jacente z_i est au dessous de λ et y_i=1 si elle est au dessus de λ (figure 1.8).

Figure 1.8 : Relation entre variable discrète et la variable sous-jacente



Dans ce cas, les probabilités qu'un individu tiré au hasard dans la population présente un phénotype $y_i=0$ ou $y_i=1$ sont données respectivement par :

$$P_0 = \int_{-\infty}^{\lambda} \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_i - \mu}{\sigma} \right)^2 \right) \right] dz_i$$

ou

$$P_1 = \int_{\lambda}^{-\infty} \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_i - \mu}{\sigma} \right)^2 \right) \right] dz_i$$

La vraisemblance à maximiser pour estimer les paramètres est donc la suivante :

$$V = \prod_i \left[\int_{-\infty}^{\lambda} \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_i - \mu}{\sigma} \right)^2 \right) \right] dz_i \right]^{1-y_i} x \left[\int_{\lambda}^{+\infty} \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_i - \mu}{\sigma} \right)^2 \right) \right] dz_i \right]^{y_i}.$$

Ce modèle s'étend très aisément au cas où le caractère discret possède plus de deux modalités, ces modalités étant ordonnées. Si n est le nombre des modalités du caractère, on aura n-1 seuils à définir.

La détection de QTL, pour ce type de caractère, est réalisée de la même manière que pour les caractères distribués normalement (voir le paragraphe 2 du chapitre II), en utilisant la fonction de pénétrance détaillée ci-dessus. Ainsi sous l'hypothèse H1 d'existence de QTL, la fonction de pénétrance pour l'individu i, pour un caractère binaire, s'écrit (Hackett et Weller, 1995 ; Xu et Atchley, 1996):

$$P_i = \delta\phi(z_i - \mu + x_{ij}\alpha_j) + (1-\delta)(1 - \phi(z_i - \mu + x_{ij}\alpha_j)).$$

Φ étant la fonction de répartition de la loi normale de moyenne μ et d'écart type σ, et δ est l'indicatrice, prenant comme valeur 1 ou 0 selon que la performance de l'individu i est égale à 0 ou 1. x_{ij} est la probabilité que l'individu i ait reçu l'allèle j au QTL sachant l'information au marqueur et α_j est l'effet QTL.

Sous l'hypothèse H0 d'absence de QTL, cette pénétrance s'écrit :

$$P_i = \delta\phi(z_i - \mu) + (1-\delta)(1 - \phi(z_i - \mu)).$$

La manipulation de ce modèle étant difficile, du fait de la prise en compte de l'intégration numérique, le modèle logistique peut être utilisé pour estimer une

approximation de $\Psi(\epsilon)$ (Bonney, 1986 ; Xu et Atchley, 1996). Le modèle logistique est exprimé comme suit :

$$\Psi(\epsilon) = \frac{e^{\epsilon}}{1+e^{\epsilon}} .$$

La relation approximative entre la fonction ϕ et un modèle logistique est :

$$\phi(\epsilon) \approx \Psi(\frac{\pi}{3}\epsilon).$$

$$\text{Donc } \phi\left(z_i - \mu + x_j\alpha_j\right) \approx \frac{e^{\frac{\pi}{3}(z_i-\mu+x_j\alpha_j)}}{1+e^{\frac{\pi}{3}(z_i-\mu+x_j\alpha_j)}}.$$

3.2 Détection multi-QTLs

Ce point a fait l'objet de plusieurs études bibliographiques très complètes (Elsen, 2001 ; Gilbert, 2003) : le texte qui suit s'en inspire largement.

L'analyse uni-QTL, qui compare l'hypothèse de la ségrégation d'un QTL en une position du génome (H1) à celle de l'absence totale de QTL (H0) dans le groupe de liaison correspondant, comporte des limites assez importantes.

En effet, la distribution de la statistique de test sous H0 peut être modifiée par la présence d'une variabilité génétique due à d'autres QTLs que celui testé. Cela peut conduire dans certaines situations à sous estimer considérablement le seuil de rejet de H0 et à conclure à tort à la présence d'un QTL.

Par ailleurs, dans le cas où deux QTLs en phase sont situés dans le même groupe de liaison, la méthode d'analyse uni-QTL fait souvent apparaître un QTL, qualifié de QTL fantôme, qui sera localisé entre les deux vrais QTL. Si au contraire, les deux QTLs localisés dans le même groupe de liaison n'agissent pas dans le même sens, la co-transmission, fréquente du fait de la liaison génétique, de deux allèles à effets contradictoires diminuera fortement la probabilité de détection de chacun des QTLs dans le modèle uni-QTL. La présence de plusieurs QTLs dans le même groupe de liaison a donc des effets très importants sur la qualité des estimations des positions et des effets de QTL (Haley et Knott 1992, Martinez et Curnow 1992).

Pour résoudre ces problèmes, plusieurs méthodes d'analyse qui prennent en compte la présence de plusieurs QTLs dans un groupe de liaison, ont été proposées.

Haley et Knott (1992), Martinez et Curnow (1992) ont proposé de faire la détection simultanée de deux QTLs situés dans deux intervalles différents en considérant non plus les deux possibilités de transmission par un parent d'un segment chromosomique selon l'origine grand-parentale, mais les quatre possibilités de co-transmission des deux segments chromosomiques correspondant aux deux positions testées. Cette méthode a été proposée dans le cadre de la régression linéaire multiple (du phénotype sur les marqueurs). Les coefficients de la régression s'interprètent en termes de localisations et d'effets de QTL. Les effets d'interaction entre les deux QTLs peuvent être pris en compte (Haley et Knott 1996) dans cette analyse, ce qui augmente fortement le nombre de paramètres à estimer. L'analyse de détection se fait

alors dans un espace bidimensionnel (avec un profil de vraisemblance à deux dimensions), la quantité des calculs nécessaires augmentant considérablement. A noter que cette méthode a été initiée par Lander et Botstein en 1989. Ces derniers proposaient d'effectuer l'analyse en maintenant l'unidimensionnalité de la cartographie d'intervalle : ils suggéraient de tester la présence d'un second QTL dans le groupe de liaison en intégrant en covariable la position et l'effet du premier QTL dans le nouveau modèle d'analyse. La généralisation dans le cas de plusieurs QTLs, proprement dit, est abordée ci-dessous.

3.2.1 Cartographie d'intervalle composite (CIM) et cartographie de QTL multiple (MQM)

Les méthodes CIM pour Combined Interval Mapping de Zeng (1994) et MQM pour Multiple QTL Mapping de Jansen *et al.* (1993, 1994a, 1994b et 1996), sont des méthodes similaires, indépendamment développées par ces auteurs. Elles permettent de prendre en compte, hormis la position étudiée, l'impact d'autres régions chromosomiques sur la variabilité du caractère. Ces méthodes ont été développées pour des croisements entre des lignées « inbred ». La CIM consiste à ajouter en cofacteurs, dans le modèle de cartographie d'intervalle, les effets marqueurs. Ainsi, les effets d'autres régions chromosomiques sont intégrés dans l'analyse en utilisant les marqueurs flanquant ces régions, comme cofacteurs dans une régression linéaire multiple. Le modèle des performances pour un croisement entre lignées « inbred » est le suivant (Zeng, 1994):

$$y_j = \mu + \sum_{i=1}^{t} b_i x_{ji} + \varepsilon_j$$

Pour tout individu j=1,…,n et t marqueurs génétiques, avec x_{ji}, la variable indicatrice du génotype au $i^{ième}$ marqueur pour l'individu j,

μ la moyenne du modèle,

b_i le coefficient de la régression partielle du phénotype y sur le $i^{ième}$ marqueur conditionné aux autres marqueurs,

et ε_j est l'erreur aléatoire normalement distribuée avec une moyenne nulle.

3.2.2 La cartographie d'intervalle multiple (MIM)

Zeng *et al.* (1999) et Kao *et al.* (1999) ont proposé la méthode MIM (Multiple Interval Mapping), qui est une généralisation de la CIM (Combined Interval Mapping) pour un nombre quelconque de QTL.

En s'inspirant du modèle décrit précédemment pour la CIM, celui du MIM est défini de la façon suivante (Kao *et al*., 1999):

$$y_j = \mu + \sum_{j=1}^{m} a_j \, x_{ij}^{*} + \sum_{j \neq k}^{m} \delta_{jk} \left(w_{ik} x_{ij}^{*} x_{ik}^{*} \right) + \varepsilon_i$$

Pour tout individu i=1,…,n et m QTL, avec x_{ik}^* la variable indicatrice du génotype au $j^{ième}$ QTL pour l'individu i,

a_j l'effet principal du QTL j,

δ_{jk} la variable indicatrice de l'épistasie entre les QTLs Q_j et Q_k,

w_{jk} l'effet d'épistasie entre les QTLs Q_j et Q_k,

et ε_i est l'erreur aléatoire normalement distribuée avec une moyenne nulle.

Il s'agit d'une méthode itérative : on part d'un modèle avec absence totale de QTL, et à chaque itération le modèle peut être augmenté ou diminué d'un ou de plusieurs QTLs à la fois.

Le principal défaut de cette méthode est le nombre relativement élevé de paramètres à estimer et donc le temps de calcul. En effet, il faut estimer, à chaque itération, la position et l'effet pour chaque QTL, ainsi que les effets d'épistasie entre ces QTLs. Un autre problème est celui du choix des QTLs et de l'effet à considérer dans le modèle. Il existe deux procédures qui permettent de faire ce choix :

- la procédure empirique de Kao *et al.* (1999), qui consiste à rechercher les QTLs, en faisant évoluer les modèles. On suppose d'abord l'hypothèse d'absence complète de QTL et on fixe les seuils significatifs qui conditionneront l'entrée et la sortie des QTLs. Si à une position donnée le test est supérieur au seuil, l'existence d'un QTL est supposée. Une fois ce test effectué sur l'ensemble du génome, on réalise une analyse multidimensionnelle de détection de QTL.

- la procédure systématique développée indépendamment par Carlborg *et al*. (2000) et Nakamichi *et al*. (2001) qui consiste à chercher simultanément plusieurs QTLs sur l'ensemble du génome. Elle s'appuie sur l'Algorithme Génétique en situation de deux QTLs pour les premiers et de plusieurs QTLs pour les seconds. Elle permet, non seulement de réduire considérablement le temps de calcul, mais aussi et surtout de réaliser une analyse de détection plus complète et plus puissante.

Les méthodes ci-dessus décrites ont été développées pour des croisements de lignées « inbred » et sont donc difficilement extrapolables pour des populations « outbred ». En effet elles perdent beaucoup de leur puissance du fait du nombre élevé de paramètres à estimer. Cependant, certains auteurs ont proposé des extensions des méthodes présentées (Jansen *et al.*, 1996 et 1998 ; Kao *et al.*, 1999), pour l'application à des populations « outbred ».

Ces méthodes ne sont pas sans rappeler celles de la sélection génomique. En effet celles-ci reviennent à tester, en même temps, les effets de tous les marqueurs sur le caractère étudié, pour éventuellement ne garder que les marqueurs qui ont les plus grands effets. Les modèles utilisés sont très semblables à ceux décrits dans ce paragraphe.

3.3 Amélioration de la puissance : le dispositif petites-filles (Weller *et al*., 1990) et la prise en compte du « mérite génétique »

Une façon d'améliorer la puissance de détection de QTL est d'utiliser le protocole dit « petites-filles ».

L'idée générale du dispositif petites-filles (ou protocole grand-père) proposé par Weller *et al.* (1990), est toujours d'observer un écart de performances entre groupes de descendants ayant reçu de leur parent l'un ou l'autre des allèles grand-parentaux aux marqueurs testés. Dans ce dispositif, les descendants ne sont toutefois pas phénotypés : les phénotypes sont mesurés sur leurs propres descendants, par exemple dans le cadre d'un testage sur descendances. Les mesures phénotypiques réalisées sur les petites-filles sont alors affectées à leurs parents marqués. Une famille est donc constituée de trois générations : un parent dont on espère qu'il est hétérozygote à la fois au marqueur et au QTL (un grand père) produisant des descendants (des fils) eux-mêmes évalués à partir des phénotypes de leurs descendants (les petites filles du grand père). La coségrégation entre allèles au marqueur du grand père et valeurs génétiques des fils est alors testée.

Le modèle statistique est le même que celui précédemment décrit :

$$Y_j = s_i + (2p_{ij} - 1)a_i + e_{ij}$$

Mais ici Y_i est la performance moyenne des petits descendants corrigée par les effets environnementaux : en pratique, on utilise le DYD ou Daughter Yield Deviation (VanRaden et Wiggans, 1991). Dans les cas où les DYD ne sont pas disponibles, ils sont remplacés par des index « dérégressés », qui sont en principe équivalents (Thomsen et *al.*, 2001).

Dans le modèle ci-dessus, s_i est l'effet père i, p_{ij} est la probabilité que le descendant j ait reçu l'allèle 1 du père i conditionnellement à l'information aux marqueurs, a_i est l'effet de l'allèle 1 du père i et e_{ij} la résiduelle du modèle qui suit une normale de moyenne nulle et de variance σ^2/CD. Le CD (ou coefficient de détermination) est fonction de nombre de descendants et permet de pondérer le DYD : plus le nombre de descendants est faible, plus la variabilité résiduelle est élevée.

L'utilisation d'un dispositif « petites-filles » permet d'augmenter la puissance de détection de QTL car la différence de performances moyennes standardisées entre groupes (selon l'allèle reçu au marqueur) augmente considérablement. En effet, le DYD assimilé au phénotype du fils (individus de la deuxième génération) est affecté d'une variance qui est, par rapport à la variance phénotypique du caractère, « divisée » par le nombre de petites-filles (σ^2/n). Pour un nombre raisonnable de descendants par père, le gain de puissance est réel malgré la division par deux de l'effet QTL ($\alpha/2$) due au fait que l'allèle grand-paternel du fils n'étant transmis qu'à une petite-fille sur deux.

L'inconvénient de ce dispositif est qu'il implique l'existence d'une structure de population particulière, qui est celle rencontrée chez les bovins laitiers, qui bénéficient de plus d'un système de contrôle de performances efficace, permettant l'acquisition des phénotypes. Imaginé à l'origine pour les bovins laitiers, le dispositif « petites-filles » a depuis été mis en œuvre dans différentes espèces. On peut citer par exemple Tribout *et al*. (2008) avec un dispositif porcin. Tribout *et al*. (2008) combinent, pour la prédiction du mérite génétique des reproducteurs, les performances propres des animaux génotypés et celles de leurs descendants. Cela permet de renforcer la précision de la prédiction du mérite génétique pour les parents n'ayant pas suffisamment de descendants. La valeur du mérite génétique est

calculée, dans cette étude, en généralisant le DYD (VanRaden et Wiggans 1991), et en prenant en compte les éventuelles performances propres des parents. Pour les caractères de production (une seule performance par individu) le mérite génétique d'un père i est calculé selon la formule suivante :

$$MG_i = \frac{y_i + \frac{\lambda}{2}\sum_{j=1}^{n}\left[\frac{q_j(y_j - \frac{\hat{u}_{mj}}{2})}{1+\lambda q_j}\right]}{W}, \text{ avec } W = 1 + \frac{\lambda}{4}\sum_{j=1}^{n}\left(\frac{q_j}{1+q_j\lambda}\right).$$

$$\lambda = \frac{\sigma_e^2}{\sigma_a^2} = \frac{variance\ résiduelle}{variance\ génétique\ additive},$$

n est le nombre de descendants du parent i ayant une performance,
y_i est la performance propre du parent i et y_j celle de son $j^{ième}$ descendant,
$\hat{u}_{mj}$ est la valeur génétique additive estimée du second parent du $j^{ième}$ descendant de i, q_j est égale à 2 si le second parent du $j^{ième}$ descendant de i est connu et 4/3 sinon.

Ces formules sont généralisées, par ces mêmes auteurs, pour les caractères répétés.

4 Outils de calcul

Différents logiciels ont été développés pour la mise en œuvre de détections de QTL, les plus connus en génétique animale étant QTLMap (Filangi *et al.,* 2010) et QTLexpress récemment repris par GridQTL (http://www.gridqtl.org.uk/) (Seaton *et al*., 2002). Dans le cadre de cette thèse, le logiciel QTLMap (Elsen *et al*., 1999) a été utilisé. Ce logiciel, disponible à l'adresse https://qgp.jouy.inra.fr/, est développé par les chercheurs de l'INRA. Il est conçu pour la détection de QTL (analyse de liaison, déséquilibre de liaison, …) dans des populations « outbred ». Il permet d'analyser des dispositifs de plein-frères, de demi-frères ou des populations de mélanges de plein et de demi-frères. Le modèle implémenté repose sur une fonction de vraisemblance approchée. Il est possible, grâce à ce logiciel, de mener différentes méthodes de détection : analyse uni-QTL uni-caractère, analyse multi-caractères ou recherche de deux QTLs liées influençant le même caractère. Les différents seuils de significativité, au niveau du groupe de liaison (ou chromosome) et du génome sont calculés par simulation. QTLMap donne aussi la position la plus probable pour le QTL mais pas l'intervalle de confiance associé à cette position.

Pour le besoin de cette thèse, de nouveaux développements ont été effectués dans ce logiciel. En effet, une version de QTLMap qui prend en compte des phénotypes mesurés sur les descendants des individus marqués a été conçue. Cette version permet aussi de calculer et d'intégrer dans l'analyse le coefficient de détermination (CD) à la fois pour les approches uni-caractère comme multi-caractères.

Des développements ont aussi été réalisés pour simuler dans quelle mesure (en termes de nombre de modalités et de la proportion de ces derniers) les caractères discrets pouvaient être traités comme des caractères normaux.

5 Conclusions

Les méthodes conçues pour la cartographie des QTLs par l'analyse de liaison sont variées et diverses, elles utilisent les techniques statistiques telles que l'analyse de variance et le maximum de vraisemblance. Toutefois le principe reste le même, et nécessite que la population d'étude présente une variabilité génétique. La population parentale peut être choisie pour garantir cette diversité, par exemple en accouplant des races très différentes maximisant ainsi l'hétérozygotie chez les parents.

Le choix de la méthode statistique utilisée dépend entre autres du type de protocole expérimental, de la distribution des caractères étudiés et des effectifs d'animaux.

Une fois la position la plus vraisemblable d'un QTL trouvée, un seuil est fixé grâce aux différentes méthodes détaillées ci-dessus. La décision de considérer cette position comme étant effectivement celle d'un QTL est prise au vu de ce seuil. Enfin un intervalle de confiance est calculé autour de cette position.

Les QTLs détectés dans le cadre des analyses uni-caractère-uni-QTL sont généralement complétées par des analyses multi-caractères et/ou multi-QTLs. L'analyse multi-caractères permet la détection de nouveaux QTLs car les informations fournies au modèle sont plus importantes que dans le cadre de l'analyse uni-caractère. Elle permet aussi de mettre en évidence les QTLs pléiotropiques si les caractères contrôlés par des QTLs positionnés au même endroit sont regroupés. Une analyse multi-QTLs permet, quant à elle, de voir si plusieurs QTLs avec des petits effets situés à des positions très proches ne sont pas à l'origine du QTL détecté.

Chapitre II. Apports méthodologiques : Amélioration de l’estimation du mérite génétique et prise en compte, en détection de QTL des caractères à seuils

1 Prise en compte des performances mesurées sur les générations postérieures à la troisième génération en détection de QTL avec analyse de liaison

1.1 Introduction

La détection de QTL par l'analyse de liaison est aujourd'hui une technique largement documentée dans la littérature (Soller et Genizi, 1978 ; Lander et Botstein, 1989 ; Weller *et al.,* 1990). Le principe et les modèles statistiques utilisés sont décrits dans le chapitre I. Elle consiste à observer, intra-famille de père (ou mère, génération G1), si les performances des descendants (génération G2) sont significativement différentes entre groupes définis par la combinaison allélique reçue du père aux marqueurs informatifs flanquant la position testée. Dans l'approche par régression (Haley et Knott, 1992), les performances des descendants sont régressées par les probabilités de transmission des allèles parentales au QTL sachant les informations aux marqueurs flanquants. Avec ce même modèle, il est possible de regrouper le mérite génétique des descendants, calculé à partir des performances mesurées sur les petits-descendants (G3), à la place de leurs performances propres si celles-ci ne sont pas mesurables. Cette modélisation correspond au protocole « petites-filles » (Weller *et al.,* 1990). L'intérêt de ce protocole est multiple. Il permet notamment d'augmenter la puissance de détection de QTL (paragraphe 3.3 du chapitre I). Tribout *et al.* (2008) ont étendu ce dispositif au cas où les descendants possèdent des performances propres.

Dans le calcul du mérite génétique des descendants génotypés, il peut être envisagé de prendre en compte, en plus de leurs performances et celles des petits-descendants, les performances des générations suivantes (G4, G5,…), quand elles sont disponibles.

Nous proposons, dans un article accepté par *Journal of Animal Breeding and Genetics*, une méthode permettant de prendre en compte ces « nouvelles » performances dans la détection de QTL. Cette méthode renforce la puissance des dispositifs de détection de QTL. Cela est particulièrement vrai pour les protocoles où les effectifs par familles sont réduits, comme c'est souvent le cas en génétique humaine ou pour des espèces où l'insémination artificielle n'est pas totalement maitrisée. Toutefois, les intervalles entre générations doivent être suffisamment courts pour que la mise en place d'un tel dispositif soit pratique.

Nous montrons comment obtenir la meilleure combinaison linéaire de performances de la descendance (G2, G3, …) qu'il convient d'analyser dans un modèle de régression de type Haley/Knott comme « performance » étendue des individus de la deuxième génération (G2) génotypés. Nous analysons les relations entre ces combinaisons linéaires et l'estimation BLUP des valeurs génétiques de ces individus G2.

1.2 Article 1

QTL detection from regression analysis of “generalized de-regressed proof” information

M. Kileh Wais[1, 2], J.M. Elsen[1]

[1] INRA, UR 631, Station d’Amélioration Génétique des Animaux, 31326 Castanet Tolosan, France
[2] University of Djibouti, Centre de Recherche Universitaire de Djibouti, Djibouti

Email : Mohamed Kileh Wais – MohamedKileh@gmail.com ; Jean-Michel Elsen - Jean-Michel.Elsen@toulouse.inra.fr

Soumis et accepté par Journal of Animal Breeding and Genetics

Journal of
Animal Breeding and Genetics

J. Anim. Breed. Genet. ISSN 0931-2668

SHORT COMMUNICATION

QTL detection from regression analysis of 'generalized de-regressed proof' information

M. Kileh Wais[1,2] & J.M. Elsen[1]

1 INRA, UR 631, Station d'Amélioration Génétique des Animaux, Castanet Tolosan, France
2 Université de Djibouti, Centre de Recherche Universitaire de Djibouti, Djibouti

Keywords
De-regressed proof; genetic merit; QTL; granddaughter designs.

Correspondence
M.K. Wais, INRA, UR 631, Station d'Amélioration Génétique des Animaux, 31326 Castanet Tolosan, France. Tel: 33 561285376; Fax: 33 561285353; E-mail: mohamedkileh@gmail.com

Received: 4 July 2011;
accepted: 19 November 2011

Summary

QTL detection using the regression of phenotypes on transmission probability is largely used when large families are available. In three generations designs, the use of a 'de-regressed proof' as a phenotype to be analysed was proposed by Weller *et al.* (1990) and Tribout *et al.* (2008). Our work generalizes this approach. A score (that we define as a 'generalized de-regressed proof') is described, which combines performance phenotypes recorded in multigenerational offspring of genotyped individuals. Estimation of the QTL effect on this score with a simple regression is unbiased. The link between this score and the BLUP animal model of the polygenic effect is demonstrated. The theory is developed and two simple examples illustrate how this technique can be implemented.

Introduction

Within-family analysis of the linkage between quantitative performances and the transmissions of marker alleles is a commonly used technique for QTL detection (Soller *et al.* 1976; Soller & Genizi 1978; Lander & Botstein 1989; Weller *et al.* 1990; Haley *et al.* 1994). Even if, because of the availability of large-scale SNP genotyping techniques, linkage disequilibrium analysis is now possible and powerful in many species, the within-family linkage approach is still attractive both for its robustness and easiness. Various models were proposed, including regression techniques (Haley & Knott 1992), maximum likelihood mixture analyses (Elsen *et al.* 1999) and variance components analyses (e.g. Uimari & Hoeschele 1997; George *et al.* 2000).

In regression approaches, the quantitative performances of half-sibs or full-sibs are related to the transmission, from the parents to the progeny, of markers flanking the position at which a segregating QTL is hypothesized: progeny quantitative performances are regressed on the QTL alleles transmission probabilities, given the marker information.

The standard design thus includes two generations: the parents (G1) tested to be heterozygous for a QTL, and their progeny (G2), in which both the quantitative trait and marker genotypes are determined.

This design can be enriched by assessing a third generation (the grand-offspring, G3). In this case, the marker genotypes of G2 individuals are related to the mean performance of their own G3 offspring. This was proposed by Weller *et al.* (1990) and largely implemented, in particular in dairy cattle, in so-called granddaughter designs (e.g. Zhang *et al.* 1998; Farnir *et al.* 2002; Bennewitz *et al.* 2003; Boichard *et al.* 2003).

In these designs, the performances of the third generation are summarized rather than analysed independently. The most routinely employed practical solution is to use, as pseudo-phenotypes for the genotyped G2 individuals their 'daughter yield deviation' (DYD), which are generally available with their 'estimated breeding values' (EBV) (*cf.* the

© 2011 Blackwell Verlag GmbH

doi:10.1111/j.1439-0388.2011.00981.x

granddaughter designs mentioned previously). A few alternative, but less powerful, solutions have been proposed such as direct analysis of EBVs or 'de-regressed proofs' that is EBVs corrected for their precision (Thomsen *et al.* 2001), as also proposed by Lien *et al.* (1995) who used a 'regenerated right-hand side of mixed model equations'.

More recently, Tribout *et al.* (2008) extended the procedure to situations where performances are available in both G2 and G3 generations, with repetitions. The authors defined a 'genetic merit' or 'unregressed summary of own and progeny performances', a generalization of the use of DYD in granddaughter designs, as the phenotype to be analysed.

In practice, the precision of such summarized phenotypes is variable between G2 individuals. Taking account this heterogeneity of residual variances, the regression approach seems obvious, and currently available software (e.g. QTLExpress, Seaton *et al.* 2002; QTLMap, Filangi *et al.* 2010) can be used directly. Typically, precision is quantified by the squared correlation between the estimated and true genetic values and referred to as the determination coefficient or reliability (VanRaden & Wiggans 1991; Georges *et al.* 1995).

Our objective in this paper is to generalize this approach to any type of pedigree descendant from the G2. We define a 'generalized de-regressed proof' (GDRP) to be used as G2 individuals pseudo-phenotypes in simple regression approach. We describe how this GDRP is linked to BLUP animal model estimations.

Methods

Notations

A set of n_1 unrelated sire families ($i = 1 \ldots n_1$) was considered. The progeny ($j = 1 \ldots n_{2i}$) of sire i were born from different dams. The probabilities of i to j transmission of alleles for a putative QTL located at a given position, conditional to marker information, were estimated: $t_{j,2i-1}$(resp.$t_{j,2i}$) is the probability that i transmitted its first (resp. second) QTL allele to j. It should be noted that $t_{j,2i} = 1 - t_{j,2i-1}$. The corresponding sire i QTL allele effects are designated by a'_i and a''_i.

The performance trait of the animal k, say y_{ik}, (belonging to G2 or G3+) depends on QTL allele inherited from the sire i and on its polygenic value u_k and a residual e_k, both two quantities being considered as random normal variables. The effect of QTL alleles that are not inherited from G1 contributes to the residual effect.

One step model

In the simplest situation ($k = j$ is a direct progeny of parent i), the regression model is:

$$y_{ij} = t_{j,2i-1} \cdot a'_i + t_{j,2i} \cdot a''_i + u_{ij} + e_{ij}$$

In the case of G3 progeny ($k = jl$), with only one parent belonging to G2, the performance y_{ijl} of the progeny l ($l = 1 \ldots n_{3ij}$) of the sire j born to grandsire i is:

$$y_{jl} = \frac{1}{2} t_{j,2i-1} \cdot a'_i + \frac{1}{2} t_{j,2i} \cdot a''_i + u_{ij} + e_{ij}$$

More generally, n individuals are considered, from which n_1 belong to the parental generation, in which the transmission of marker alleles is traced, n_2 belong to generation G2 which is genotyped, and n_3 are the progeny of subsequent generations (G3+). $n \geq n_1 + n_2 + n_3$.

Assuming there are no fixed effects, the description model is

$$y = ZPWTa + Zu + e \qquad (1)$$

where:

y is a vector of N performances from individuals belonging to various generations, including G2 and G3+

a a $2\,n_1$-vector of fixed QTL allele effects in G1 parents

u a n-vector of polygenic effects of pedigreed individuals, a random normal multi-variable with a 0 mean and D covariances matrix

e a N-vector of residual effects of pedigreed individuals, a random normal multi-variable with a 0 mean and $R = I \cdot \sigma_e^2$ covariances matrix

Z the $N \times n$ incidence matrix linking the N performances to the n individuals of the pedigree (Z_{ij} is 1 if the i th phenotype ($i = 1 \ldots N$) belongs to the j th individual ($j = 1 \ldots n$), 0 if not)

W the $n \times n_2$ incidence matrix which designates, among the n individuals, the n_2 individuals which belong to G2 (W_{ij} is 1 if the i th individual of the pedigree ($i = 1 \ldots n$) is the j th G2 individual ($j = 1 \ldots n_2$), 0 if not)

T the $n_2 \times 2n_1$ matrix of G1 to G2 transmission probabilities, conditional to marker information

P the $n \times n$ matrix of QTL allele transmission between generations. P is a lower triangular matrix that can be built by sorting the blocs by generation (First G1, G2 then...). It is constructed recursively as: $p_{ii} = 1$ and $p_{ij} = \frac{1}{2} p_{sj} + \frac{1}{2} p_{dj}$ where s and d are the indices of i's parents, with $j < i$ and $p_{ij} = 0$ for $j > i$.

© 2011 Blackwell Verlag GmbH

The y covariance matrix is $var(Zu + e) = V = ZDZ' + R$, where $D = A\sigma^2_u$, A the additive relationship matrix and $R = I\sigma^2_e$.

It should be noted that A and P matrices are linked by $A = PCP'$, where C is a diagonal matrix in which the c_{kk} element is 1, ¾ or ½ (in the non-inbred situation) when none, 1 or 2 of the k parents are known (Quaas 1984).

Following this model [I], the generalized least squares estimation of QTL effects is

$$\hat{a} = (T'W'P'Z'V^{-1}ZPWT)^{-1}T'W'P'Z'V^{-1}y$$

Equivalent two steps model

This $\hat{a}$ estimation may be reformulated as

$$\hat{a} = (T'(QVQ')^{-1}T)^{-1}T'(QVQ')^{-1}z$$

where $Q = (W'P'Z'V^{-1}ZPW)^{-1}W'P'Z'V^{-1}$.

Thus, estimating the QTL effect from model [I] is equivalent to the analysis of a linear combination $z = Qy$ of the y phenotypes (called hereafter the 'score'), using the model:

$$z = Ta + e, with\ e \sim N(0, QVQ') \qquad (2)$$

which is the model currently used in QTL regression approaches (Seaton *et al.* 2002; Filangi *et al.* 2010).

Is this score consistent with the DYD, 'de-regressed proof' or 'genetic merit' used by others?

Meaning of the score

This quantity is linked to the BLUP animal model estimations of the polygenic values u, ignoring the QTL effect a, *that is*, following the model $y = Zu + e$.

Indeed, $z = (W'P'Z'V^{-1}\ ZPW)^{-1}W'(PC)^{-1}PCP'Z'V^{-1}y$

Noting that the BLUP of u is $\hat{u} = DZ'V^{-1}y = \sigma^2_u PCP'Z'V^{-1}y$

We get

$$z = (W'P'Z'V^{-1}ZPW)^{-1}W'(PC)^{-1}\hat{u}/\sigma^2_u$$

$$z = (W'P'Z'V^{-1}ZPW)^{-1}W'C^{-1}P^{-1}\hat{u}/\sigma^2_u$$

The P^{-1} inverse of the P matrix is quite simple: a lower triangular matrix in which diagonal elements are 1 and off-diagonal elements are 0, except for $(P^{-1})_{is}$ and $(P^{-1})_{id}$ which are $-1/2$, where s and d are the indices of i's parents (if they belong to the pedigree list).

Thus, the i th element of $P^{-1}\hat{u}$ is the BLUP $\hat{u}_i$ of u_i in deviation from the halved known parental values: $\hat{u}_i - \frac{1}{2}(\hat{u}_s + \hat{u}_d)$, $\hat{u}_i - \frac{1}{2}\hat{u}_s$, $\hat{u}_i - \frac{1}{2}\hat{u}_d$ or $\hat{u}_i$, when the pedigree list includes both parents, the sire only, the dam only or none of them, respectively.

Finally, $\tilde{u} = W'C^{-1}P^{-1}\hat{u}$ is the subset of standardized deviations $\tilde{u}_i$ of the n_2 G2 individuals.

The variance $V(\hat{u})$ of $\hat{u}$ is

$$V(\hat{u}) = \sigma^4_u PCP'Z'V^{-1}ZPCP$$

The variance $V(\tilde{u})$ of $\tilde{u}$ is $V(\tilde{u}) = W'C^{-1}P^{-1}V(\hat{u})P^{-1}C^{-1}W$, *that is*:

$$V(\tilde{u}) = \sigma^4_u W'P'Z'V^{-1}ZPW \qquad (3)$$

Thus,

$$z = \sigma^2_u V(\tilde{u})^{-1}\tilde{u} \qquad (4)$$

demonstrating that the score z generalizes the genetic merit defined by [2].

The covariances matrix of z is

$$V_z = \sigma^4_u V(\tilde{u})^{-1} \qquad (5)$$

When a QTL regression model ($z = Ta + e$) has to be applied to the GDRP z, the heterogeneity of the residual variances $var(\varepsilon)_{ii} = \sigma^4_u(V(\tilde{u})^{-1})_{ii}$ must be considered. Nothing that $[V(\tilde{u})]_{ii}/\sigma^2_u$ is the reliability coefficient CD_i of the BLUP of $[u_i - \frac{1}{2}(u_s + u_d)]/C_{ii}$, it must be emphasized that this residual variances heterogeneity is not measured by the inverse of this reliability coefficient but by $\sigma^2_u(V(\tilde{u})^{-1})_{ii}$. Nevertheless, when the off-diagonal elements of the $V(\tilde{u})$ matrix are negligible, both expressions are very close.

Complete treatment of the information should include the covariances between z_i which are not considered in QTL regression approaches (e.g. Haley & Knott 1992; Elsen *et al.* 1999).

Illustration: application to the classical situations

To illustrate the method developed previously, it will be applied to two standard situations: daughter and granddaughter designs.

Daughter design

In this basic situation, none of the G3+ information is used, and the method is expected to be reduced to the direct analysis of y information, without variance heterogeneity. No fixed effects are considered here.

Equations are provided for a single family to simplify the demonstration, and the sire i indice is omitted. Let d be the number of G2 daughters of this sire.

The matrices describing the designs follows:

$$Z=\begin{pmatrix}0&1&\dots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\dots&1\end{pmatrix}\quad W=\begin{pmatrix}0&\dots&0\\ 1&\dots&0\\ \vdots&\ddots&\vdots\\ 0&\dots&1\end{pmatrix}\quad P=\begin{pmatrix}1&0&\dots&0\\ 1/2&1&\dots&0\\ \vdots&\vdots&\ddots&\vdots\\ 1/2&0&\dots&1\end{pmatrix}\quad C=\begin{pmatrix}1&0&\dots&0\\ 0&3/4&\dots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\dots&3/4\end{pmatrix}$$

Let σ_u^2 and σ_e^2 be the genetic and residual variances, with $\sigma^2=\sigma_u^2+\sigma_e^2$, the additive relationship matrix and y covariance matrix come to be

$$D=A\sigma_u^2=PCP'\sigma_u^2=\begin{pmatrix}1&1/2&\dots&1/2\\ 1/2&1&\dots&1/4\\ \vdots&\vdots&\ddots&\vdots\\ 1/2&1/4&\dots&1\end{pmatrix}\sigma_u^2$$

$$V=ZDZ'+R=\begin{pmatrix}1&1/4&\dots&1/4\\ 1/4&1&\dots&1/4\\ \vdots&\vdots&\ddots&\vdots\\ 1/4&1/4&\dots&1\end{pmatrix}\sigma_u^2+\begin{pmatrix}1&0&\dots&0\\ 0&1&\dots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\dots&1\end{pmatrix}\sigma_e^2$$

V is a linear combination $aI_d + bJ_d$, where I_d is the $d\times d$ identity matrix and J_d a $d\times d$ matrix in which all the elements are 1. The coefficients are given by $Ma=\sigma^2-1/4\sigma_u^2$ and $b=1/4\sigma_u^2$. Thus, $V^{-1}=\frac{1}{a}I_d-\frac{b}{a[a+db]}J_d$.

As expected we find that $z_j=y_j$ and $var(z_j)=\sigma_u^4[V(\vec{u})^{-1}]^{jj}=\sigma^2$, the elements usually analysed in QTL regression analysis of sire families.

Granddaughter design

In this design, the only recorded phenotypes belong to the G3 generation. To simplify the presentation, we consider a single grandsire in a balanced situation (p sires/grandsire and d daughters/sire, giving $n=1+p+pd$). The matrices describing the designs follow:

$$Z=\begin{pmatrix}0&0&\dots&0&1&\dots&0\\ \vdots&\vdots&\ddots&\vdots&\vdots&\ddots&\vdots\\ 0&0&\dots&0&0&\dots&1\end{pmatrix}\quad W=\begin{pmatrix}0&\dots&0\\ 1&\dots&0\\ \vdots&\ddots&\vdots\\ 0&\dots&1\\ 0&\dots&0\\ \vdots&\ddots&\vdots\\ 0&\dots&0\end{pmatrix}\quad C=\begin{pmatrix}1&0&\dots&0\\ 0&3/4&\dots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\dots&3/4\end{pmatrix}$$

Noting $\lambda=\frac{\sigma_2-1/4\sigma_u^2}{1/4\sigma_u^2}$, we get $V^{-1}=\frac{4}{\sigma_u^2}[\frac{1}{\lambda}I_d-\frac{1}{\lambda(\lambda+d)}J_d]$.

Applying $\hat{u}=DZ'V^{-1}y$ and noting $\bar{y}$ the mean of the daughters' performances, the estimated polygenic values of the sire ($\hat{u}_0$) and its daughters ($\hat{u}_j$) are given by

$\hat{u}_0=2\frac{d}{d+2}\bar{y}$ and $\hat{u}_j=\frac{3}{2}(y_j-\frac{1}{2}\hat{u}_0)+\frac{1}{2}\hat{u}_0$

Thus, $\tilde{u}_j=\frac{4}{3}(\hat{u}_j-\frac{1}{2}\hat{u}_0)$.

On the other hand, $V(\vec{u})=\sigma_u^4\,W'P'Z'V^{-1}ZPW$ gives here $V(\vec{u})=\frac{4\sigma_u^2}{\lambda}[I_d-\frac{1}{\lambda+d}J_d]$ and $V(\vec{u})^{-1}=\frac{1}{\sigma_u^2}V$

$$P=\begin{pmatrix}1&0&\dots&0&0&\dots&0&\dots&0&\dots&0\\ 1/2&1&\dots&0&0&\dots&0&\dots&0&\dots&0\\ \vdots&\vdots&\ddots&\vdots&\vdots&\ddots&\vdots&\ddots&\vdots&\ddots&\vdots\\ 1/2&0&\dots&1&0&\dots&0&\dots&0&\dots&0\\ 1/4&1/2&\dots&0&1&\dots&0&\dots&0&\dots&0\\ \vdots&\vdots&\ddots&\vdots&\vdots&\ddots&\vdots&\ddots&\vdots&\ddots&\vdots\\ 1/4&1/2&\dots&0&0&\dots&1&\dots&0&\dots&0\\ \vdots&\vdots&&\vdots&\vdots&\ddots&\vdots&\ddots&\vdots&\ddots&\vdots\\ 1/4&0&\dots&1/2&0&\dots&0&\dots&1&\dots&0\\ \vdots&\vdots&\ddots&\vdots&\vdots&\ddots&\vdots&\ddots&\vdots&\ddots&\vdots\\ 1/4&0&\dots&1/2&0&\dots&0&\dots&0&\dots&1\end{pmatrix}$$

Using [III], [IV] and [V], we find that $z_j = 2\,y_{ij}$, *that is to say* twice the mean performances of sire j daughters and

$$var(z_j) = \sigma_a^4 (V(\vec{u})^{-1})_{jj} = \sigma^2 \frac{1 + \frac{1}{4}(d-1)h^2}{\frac{1}{4}d}$$

Discussion and Conclusion

We generalized to any number of generations the de-regressed proof to be analysed as a phenotype in QTL regression approaches as proposed for instance by Haley & Knott (1992) (QTL regression). In QTL regression, the performances of G2 progenies of a parent supposed to be heterozygous for a QTL are correlated with the genetic markers transmitted by this parent to its descendants. The GDRP used to replace these performances is based on the BLUP estimation of the polygenic value independently of the QTL. For a given progeny, this estimation combines its own performances and the performances of all its recorded progeny regardless of the generation number. From the simple algebra developed in this paper, it is easily possible to calculate both the GDRP [IV] and its precision [V] which should be considered as a factor of heterogeneity in the QTL regression model.

Extending this procedure to repeated measures is straightforward. In this case, a Zv term, where v is an individual random normal effect of 0 mean and $I \cdot \sigma_v^2$ covariances, is added to the equation (1) describing the trait. The covariance matrix becomes $V = ZDZ' + ZZ'\sigma_v^2 + R$, and the elements $(\vec{u}) = \sigma_a^4 W' P' Z' V^{-1} Z P W$, $z = \sigma_a^2 V(\vec{u})^{-1}\vec{u}$ and $\sigma_a^2 (V(\vec{u})^{-1})_{ii}$ are directly deduced from this new V matrix.

Similarly, considering fixed effects in the model [I] is performed easily by adding a $X\beta$ term. In this case, the inverse V^{-1} is replaced in the transformation matrix Q by $V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}$. Noting that the BLUP estimation u is now $\hat{u} = DZ'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})y$, it is easy to demonstrate that the transformation Qy is the GDRP.

It must be emphasized that, in our presentation, the BLUP estimation of the polygenic effect u (i) did not include the QTL effect (the model used was $y = Zu + e$ and not $y = ZPWTa + Zu + e$) and (ii) was only based on the performances measured in the descendants (belonging to G2 or G3+) of genotyped animals.

In our algebra, the A relationship matrix between the polygenic effects u of the genotyped and/or phenotyped n individuals was only based on the bottom part of the pedigree, *that is*, did not include possible relations between G1 reproducers, which could share ancestors (the top part of the pedigree): the G1 parents were supposed unrelated, an hypothesis made in standard QTL regression approaches. It is not proven that the QTL effects estimations from GDRP obtained in this more general situation (the case often considered, e.g. Lien *et al.* 1995) equal $\hat{a}$, the QTL effect estimation from [I]. Even if including ancestral relationships between parents in QTL regression approach is straightforward in the mixed model framework, extending the GDRP to this situation needs new developments.

Weller *et al.* (1990) clearly demonstrated that the granddaughter design is more powerful than the daughter design, in particular for small QTL effects and moderate heritability. This extra power arises from the more precise estimation of the contrasts of G1 (grandsire) origins (here estimated by $\hat{a}$), which is because of the recording of large numbers of granddaughters, as compared to the estimations based only on the performances of the daughters. In their algebra, the power was estimated considering a simplified chi-squared analysis and depended on a non-central parameter inversely proportional to the contrast variances. In our approach, these variances are 'replaced' by the precision of $\hat{a}$estimations, *that is*, $V(\hat{a}) = (T'V(\vec{u})T)^{-1}$. When including more generations in the 'generalized de-regressed proof', the CD_i reliability coefficient, inversely proportional to $V(\vec{u})$ increases, as does the equivalent non-central parameter and the detection power.

However, this gain in power, which comes from extra information to the grand progeny (G3) information, may be reduced if G3 information is already abundant. In such a situation, the CD_i precision obtained using only the G3 data is high, and adding extra generation information will have a limited impact. Thus, working with generalized de-regressed proofs will be mostly useful for species in which artificial insemination is not used on a large scale, such as small ruminants or birds.

Inversely, in the case of lowly heritable traits, use of our generalized de-regressed proof will be advantageous. This is directly in the line with the conclusions of Weller *et al.* (1990) on the effect of heritability.

Finally, the G1-G2 DNA samples, or marker data from these samples, must still be available when phenotypes from extra generations are recorded, and this constraint may limit the practical use of generalized de-regressed proof to a limited number of generations.

The method has been implemented as one of the options available in QTLMap software (http://dga7.-jouy.inra.fr/qtlmap).

Acknowledgements

MKW was supported by the State of Djibouti and the INRA Animal division. This work was partially financed by the ANR GENECAN project. We thank Christèle Robert Granié for her helpful comments.

Competing interests

The authors declare they have no competing interests.

Authors contributions

JME drafted the manuscript. Both authors participated in the development of the method and read and approved the final manuscript

References

Bennewitz J., Reinsch N., Grohs C., Levéziel H., Malafosse A., Thomsen H., Xu N., Looft C., Kühn C., Brockmann G.A., Schwerin M., Weimann C., Hiendleder S., Erhardt G., Medjugorac I., Russ I., Förster M., Brenig B., Reinhardt F., Reents R., Averdunk G., Blümel J., Boichard D., Kalm E. (2003) Combined analysis of data from two granddaughter designs: A simple strategy for QTL confirmation and experimental power in dairy cattle. *Genet. Sel. Evol.*, **35**, 319–338.

Boichard D., Grohs C., Bourgeois F., Cerqueira F., Faugeras R., Neau A., Rupp R., Amigues Y., Boscher M.Y. (2003) Detection of genes influencing economic traits in three French dairy cattle breeds. *Genet. Sel. Evol.*, **35**, 77–101 (available at: http://www.ncbi.nlm.nih.gov/pubmed?term=%22Lev%C3%A9ziel%20H%22%5BAuthor%5D; last accessed 15 December 2011).

Elsen J.M., Mangin B., Goffinet B., Boichard D., Le Roy P. (1999) Alternative models for QTL detection in livestock - I General introduction. *Genet. Sel. Evol.*, **31**, 213–224.

Farnir F., Grisart B., Coppieters W., Riquet J., Berzi P., Cambisano N., Karim L., Mni M., Moisio S., Simon P., Wagenaar D., Vilkki J., Georges M. (2002) Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbred half-sib pedigrees: revisiting the location of a quantitative trait locus with major effect on milk production on bovine chromosome 14. *Genetics*, **161**, 275–287.

Filangi O., Moreno C., Gilbert H., Legarra A., Le Roy P., Elsen J.M. (2010) QTLMap, a Software for QTL Detection in Outbred Populations. Proceedings of the 9th WCGALP, 1–6 August Leipzig, Germany.

George A.W., Visscher P.M., Haley C.S. (2000) Mapping quantitative trait loci in complex pedigrees: a two-step variance component approach. *Genetics*, **156**, 2081–2092.

Georges M., Nielsen D., Mackinnon M., Mishra A., Okimoto R., Pasquino A.T., Sargeant L.S., Sorensen A., Steele M.R., Zhao X., Womack J.E., Hoeschele I. (1995) Mapping Quantitative Trait Loci Controlling Milk Production in Dairy Cattle by Exploiting Progeny Testing. *Genetics*, **139**, 907–920.

Haley C.S., Knott S.A. (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, **69**, 315–324.

Haley C.S., Knott S.A., Elsen J.M. (1994) Mapping Quantitative Trait Loci in Crosses between Outbred Lines Using Least Squares. *Genetics*, **136**, 1195–1207.

Lander E.S., Botstein D. (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121**, 185–199.

Lien S., Gomez-Raya L., Steine T., Fimland E., Rogne S. (1995) Associations Between Casein Haplotypes and Milk Yield Traits. *J. Dairy Sci.*, **78**, 2047–2056.

Quaas R.L (1984) Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics*, **32**, 949–953.

Seaton G., Haley C.S., Knott S.A., Kearsey M., Visscher P.M. (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics*, **18**, 339–340.

Soller M., Genizi A. (1978) The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations. *Biometrics*, **34**, 47–55.

Soller M., Brody T., Genizi A. (1976) On the power of experiments to detect marker-linked quantitative effects in crosses between inbred lines. *Theor. Appl. Genet.*, **47**, 35.

Thomsen H., Reinsch N., Xu N., Looft C., Grupe S., Kühn C., Brockmann G.A., Schwerin M., Leyhe-Horn B., Hiendleder S., Erhardt G., Medjugorac I., Russ I., Förster M., Brenig B., Reinhardt F., Reents R., Blümel J., Averdunk G., Kalm E. (2001) Comparison of estimated breeding values, daughter yield deviations and de-regressed proofs within a whole genome scan for QTL. *J. Anim. Breed. Genet.*, **118**, 357–370.

Tribout T., Iannuccelli N., Druet T., Gilbert H., Riquet J., Gueblez R., Mercat M.J., Bidanel J.P., Milan D., Le Roy P. (2008) Detection of quantitative trait loci for reproduction and production traits in Large White and French Landrace pig populations. *Genet. Sel. Evol.*, **40**, 61–78.

© 2011 Blackwell Verlag GmbH

Uimari P., Hoeschele I. (1997) Mapping-linked quantitative trait loci using Bayesian analysis and Markov chain Monte Carlo algorithms. *Genetics*, **146**, 735–743.

VanRaden P.M., Wiggans G.R. (1991) Derivation, Calculation, and Use of National Animal Model Information. *J. Dairy Sci.*, **71**, 2737–2746.

Weller J.I., Kashi Y., Soller M. (1990) Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Sci.*, **73**, 2525–2537.

Zhang Q., Boichard D., Hoeschele I., Ernst C., Eggen A., Murkve B., Pfister-Genskow M., Witte L.P., Grignola F.E., Uimari P., Thaller G., Bishop M.D. (1998) Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. *Genetics*, **149**, 1959–1973.

1.3 Conclusion et perspective

La méthode décrite dans l'article permet donc d'étendre celle proposée par Weller *et al*. (1990), puis repris par Tribout *et al*. (2008) à un nombre quelconque de générations. L'article présente comment obtenir « le mérite génétique dé-régressé » et le coefficient de détermination associé, pour chacun des individus de la deuxième génération génotypée. Des extensions pour les mesures répétées et la prise en compte des effets fixes dans le modèle sont proposées.

L'apport de nouvelles mesures phénotypiques permettra d'améliorer la puissance de détection. Toutefois, lorsque les tailles des familles sont suffisamment larges (comme chez les bovins laitiers, par exemple) pour garantir un bon niveau d'estimation du mérite génétique pour les individus marqués, l'intérêt peut être limité.

Un certain nombre d'améliorations peuvent être apportées aux développements proposés. Notamment, dans nos développements algébriques, nous considérons que les individus de la première génération sont indépendants, hypothèse assez répandue en détection de QTL par analyse de liaison, ce qui peut ne pas correspondre à la réalité. La prise en compte d'éventuels liens de parenté entre les pères fondateurs peut être une des améliorations à apporter à ce modèle.

2 Etude comparative des deux modèles dans la détection de QTL des caractères à seuils : considération de la sous-jacente normale ou utilisation du modèle de mélange de distributions normales

2.1 Introduction

Comme il est dit ci-dessus (paragraphe 3, chapitre I), certains caractères d'intérêt économique ou biologique pour la production animale ont une distribution discrète, et donc non gaussienne. Comme les caractères continus, ces caractères discrets présentent en général un déterminisme complexe et leur variabilité est contrôlée par de nombreux gènes. Une hypothèse fréquente est que la variabilité de ce type de caractère (disons Y) est sous l'influence d'une variable gaussienne sous-jacente (disons Z), dont l'échelle est jalonnée de seuils (s_1, s_2, …, s_n) : quand Z est dans l'intervalle [s_i, s_{i+1}], la variable Y vaut i+1 : on parle de caractères à seuils (Falconer, 1958). Le premier à aborder le modèle à seuil en génétique fut Wright (1934). Depuis, plusieurs auteurs ont analysé des caractères de ce type, le plus souvent dans le cadre polygénique classique, i.e. sans prendre en compte l'information moléculaire, pour estimer des paramètres génétiques tels que l'héritabilité, ou des effets de facteurs connus tels que le sexe de l'animal (Fouley *et al.,* 1983 ; Gianola et Foulley, 1983 ; Harville et Mee, 1984 ; Weller *et al.,* 1988 ; Weller et Gianola, 1989 ; Jamrozik *et al.,* 1991 ; Weller et Ron, 1992 ; Weller *et al.,* 1992). D'autres auteurs (Hackett et Weller, 1995 ; Xu et Atchley, 1996) intègrent, dans leurs études sur les caractères à seuils, l'information moléculaire, dans le but de cartographier des QTLs. Leurs méthodes sont développées pour des populations « inbred ». Elles ont été étendues pour des dispositifs « outbred » par Yi et Xu (1999). Plus récemment, Xu *et al.* (2005) ont proposé une méthode multi-caractères de détection de QTL pour les caractères binaires. Tous ces auteurs considèrent une distribution normale sous-jacente à la distribution réelle de la variable discontinue étudiée.

Plusieurs études concernant des caractères à seuils (Hackett et Weller, 1995 ; Xu et Atchley, 1996 ; Rao et Xu, 1998 ; Kadarmideen *et al*., 2000) comparent les précisions des estimations de certains paramètres (notamment la position estimée du QTL et le puissance de détection du modèle) obtenues en mettant en œuvre soit un modèle non-linéaire, soit la méthode usuellement utilisée pour les caractères à distributions normales. La différence entre ces deux modèles est donc la fonction de pénétrance considérée. Tous ces auteurs ont considéré des lignées croisées, et concluent que les avantages de l'utilisation d'un modèle discret pour l'analyse des caractères à seuils sont négligeables par rapport à l'utilisation du modèle normal. Hackett et Weller (1995) obtiennent des estimations de positions légèrement meilleures pour le modèle discret alors que Rao et Xu (1998) observent une faible supériorité en termes de puissance de détection pour le modèle discret comparé au modèle normal. Toutefois, certains auteurs (Visscher *et al.,* 1996 ; Rebai, 1997 ; Kadarmideen *et al.,* 2000) rapportent que les deux méthodes d'analyse présentent des caractéristiques équivalentes. Pour les caractères binaires, le modèle discret est même moins robuste qu'un modèle normal (Xu et Atchley, 1996 ; Yi et Xu, 1999).

Cette différence de robustesse entre les deux modèles serait liée, selon ces auteurs, à des pertes d'informations occasionnées par le passage de la variable observée à la variable sous-jacente à celle-ci.

Notre travail vient en complément de ceux des auteurs précédemment cités. En effet, nous proposons d'élargir les cas de figure simulés et de comparer les 2 approches à la fois sur la précision de la position estimée du QTL et sur la puissance de détection. Ainsi, nous simulons 16 cas de figures (certains cas n'ont fait l'objet d'aucune étude), un nombre important comparé à la littérature. Les deux modèles comparés de détection de QTL par analyse de liaison, utilisent la méthode des marqueurs flanquants (Lander et Botstein, 1989). Un protocole simple est simulé et différents cas sont considérés en jouant sur le nombre de modalités du caractère considéré, leurs fréquences dans la population ainsi que l'effet du QTL simulé et sa position.

2.2 Méthode

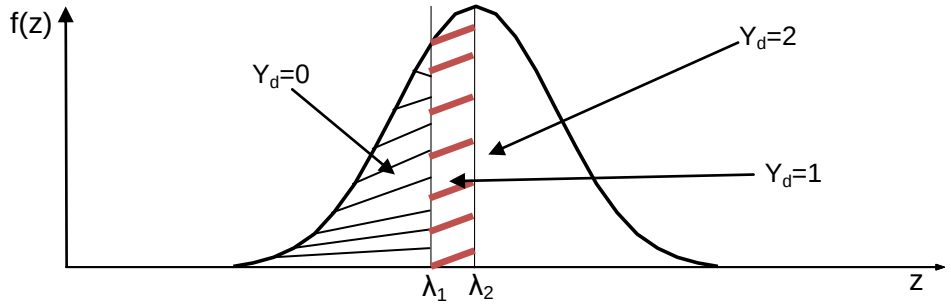
Les modèles

Le modèle normal est utilisé, classiquement, pour les caractères quantitatifs à distribution normale. Son application dans la détection de QTL dans les populations d'élevage est décrite dans le premier chapitre. Dans le modèle discret on fait l'hypothèse d'une variable aléatoire distribuée normalement, sous-jacente au caractère à seuils étudié. Cette variable aléatoire Z se caractérise ainsi :

Si $\lambda_{i-1} < Z < \lambda_i$ alors Y=i avec $i \in \{1, \dots, m\}$ et $\lambda_0 = -\infty$ et $\lambda_m = +\infty$

Les λ_i sont les seuils fixés, définis pour l'échelle sous-jacente grâce aux catégories observées sur la variable discrète Y et son nombre des modalités. La figure 2.1 représente un caractère à seuils présentant trois modalités (0, 1 et 2) avec 30% de la population totale présentant le phénotype 0, 20% le phénotype 1 et 50% le phénotype 2. m=3 est le nombre des modalités du caractère.

Figure 2.1 : Illustration d'un caractère à seuils avec 3 modalités



Dans ce cas particulier, la probabilité qu'un individu d pris au hasard dans la population soit de phénotype exprimé y_d=0 est de 30%.

Dans le cas plus général (avec un nombre quelconque m de modalités), la transformation de Z en Y peut être résumée en termes de probabilité, de la façon suivante :

$$\begin{cases} P(Y=1) = P(-\infty < Z < \lambda_1) \\ P(Y=2) = P(\lambda_1 < Z < \lambda_2) \\ \vdots \\ P(Y=m) = P(\lambda_{m-1} < Z < +\infty) \end{cases}$$

Par ailleurs, la variable aléatoire Z suit une loi normale. Cette hypothèse est notamment compatible avec celle d'un contrôle du caractère par plusieurs gènes, chacun ayant un petit effet. Pour une position donnée du génome, la performance Z_d, d'un individu d, peut s'écrire comme suit :

$$Z_d = \mu + x_{jd}\alpha_j + e_d$$

μ est la moyenne générale, x_{jd} la probabilité que l'individu d ait reçu l'allèle j au QTL sachant l'information aux marqueurs, α_j est l'effet QTL et e_j la résiduelle du modèle d'espérance nulle.

Les probabilités précédentes, pour un individu d, peuvent être calculées de la façon suivante :

$$P(Y_d = 1) = P(-\infty < Z_d < \lambda_1) = \int_{-\infty}^{\lambda_1} f(Z_d)dz$$

$$P(Y_d = 2) = P(\lambda_1 < Z_d < \lambda_2) = \int_{\lambda_1}^{\lambda_2} f(Z_d)dz$$

$$\vdots$$

$$P(Y_d = m) = P(\lambda_{m-1} < Z_d < +\infty) = \int_{\lambda_{m-1}}^{+\infty} f(Z_d)dz = 1 - \int_{-\infty}^{\lambda_{m-1}} f(Z_d)dz$$

$$\Leftrightarrow$$

$$P(Y_d = 1) = \phi(\lambda_1 - E(Z_d)) = \boldsymbol{\phi(\lambda_1 - \mu - x_{jd}\alpha_j)}$$

$$P(Y_d = 2) = \phi(\lambda_2 - E(z_d)) - \phi(\lambda_1 - E(Z_d)) = \boldsymbol{\phi(\lambda_2 - \mu - x_{jd}\alpha_j) - \phi(\lambda_1 - \mu - x_{jd}\alpha_j)}$$

$$\vdots$$

$$P(Y_d = m) = 1 - \phi(\lambda_{m-1} - E(Z_d)) = \boldsymbol{1 - \phi(\lambda_{m-1} - \mu - x_{jd}\alpha_j)}$$

f est la densité de la loi normale standard et ϕ sa fonction de répartition. On suppose ici, que Z_d est une variable aléatoire réduite.

Simulations et détection de QTL

Les simulations des populations et l'estimation des paramètres de détections de QTL (position, effet de substitution,…) ont été effectuées avec le logiciel QTLMap (Filangi *et al.,* 2010). Un module prenant en compte la sous-jacente normale aux caractères à seuils a été implémenté à cette occasion.

La méthode du maximum de vraisemblance détaillé dans le paragraphe 2.3.2.2 du chapitre I a été mise en œuvre. Seule la fonction de pénétrance est modifiée et s'écrit maintenant de la façon suivante :

$$\sum_{i=1}^{m} \delta_i(\phi(\lambda_i - \boldsymbol{\mu - x_j\alpha_j}) - \phi(\lambda_{i-1} - \boldsymbol{\mu - x_j\alpha_j})$$

δ_i est un indicateur prenant la valeur 1 si l'individu d présente un phénotype visible y_d=i, et 0 sinon.

Le rapport des vraisemblances est maximisé pour estimer les paramètres du modèle, il s'agit des seuils λ, des effets des locus α et de la distance entre la position estimée du QTL et les marqueurs flanquants.

2.3 Simulations

Pour des questions de facilité, on simule un protocole de type F2 avec sept pères fondateurs hétérozygotes aux marqueurs, chacun ayant été accouplé avec dix mères. Chacune des mères possède six produits. On a donc au total, 420 descendants. Pour chacun d'entre eux, on considère une carte génétique composée d'un chromosome de 1 Morgan, couvert de 11 marqueurs, régulièrement espacés de 10cM.

Dans toutes les situations étudiées, l'héritabilité est fixée à h^2=0,35. Le tableau 2.1 présente les caractéristiques du QTL (sa position et son effet de substitution) et du caractère (le nombre de modalités et leurs fréquences dans la population) dans les différentes simulations effectuées.

Les positions du QTL

Dans les différents dispositifs simulés, le QTL est tantôt placé à l'extrémité du chromosome (à 20cM), tantôt placé au milieu du chromosome (à 50cM), cela afin de vérifier si la position du QTL a un effet sur la précision de l'estimation des paramètres de détection de QTL.

Tableau 2.1 : Les caractéristiques des différentes populations simulées

N° Cas simulé	Position du QTL simulé (cM)	Effet du QTL simulé	Nombre de modalités[1] (fréquences)
1	0,5	fort (0,8)	2 (50% et 50%)
2	0,2	fort (0,8)	2 (50% et 50%)
3	0,5	faible (0,2)	2 (50% et 50%)
4	0,2	faible (0,2)	2 (50% et 50%)
5	0,5	fort (0,8)	2 (10% et 90%)
6	0,2	fort (0,8)	2 (10% et 90%)
7	0,5	faible (0,2)	2 (10% et 90%)
8	0,2	faible (0,2)	2 (10% et 90%)
9	0,5	fort (0,8)	3 (30%, 30% et 40%)
10	0,2	fort (0,8)	3 (30%, 30% et 40%)
11	0,5	faible (0,2)	3 (30%, 30% et 40%)
12	0,2	faible (0,2)	3 (30%, 30% et 40%)
13	0,5	fort (0,8)	3 (10%, 10% et 80%)
14	0,2	fort (0,8)	3 (10%, 10% et 80%)
15	0,5	faible (0,2)	3 (10%, 10% et 80%)
16	0,2	faible (0,2)	3 (10%, 10% et 80%)

[1]Nombre de modalités du caractère discret simulé

Les effets de substitution du QTL

Dans nos simulations, l'effet de substitution du QTL est fixé à 80% de l'écart phénotypique (donc très fort) ou à 20% (donc faible).

Pour une combinaison donnée de ces paramètres (soit un des 16 cas simulés répertoriés dans le tableau 2.1), 500 simulations permettent d'obtenir les distributions des estimations des modèles. Par ailleurs, les seuils de rejet de l'hypothèse d'absence de QTL sont obtenus par 500 simulations sous l'hypothèse d'absence de QTL, afin d'estimer la puissance de détection de QTL des deux modèles. Concrètement cette puissance est estimée par la proportion de LRT dépassant le seuil empirique de 5%.

Les critères de comparaison

Les critères pris en compte sont la puissance de détection de QTL et la précision de la position estimée du QTL. A cette fin sont comparés les écart-types de ces positions estimées et leurs erreurs moyennes quadratiques par rapport à la position simulée. Par ailleurs, le temps de calcul est aussi considéré comme un des éléments de l'efficacité des modèles.

2.4 Résultats

Nous détaillons ici, les résultats de 4 des 16 cas simulées. Elles correspondent aux quatre types de caractères simulés (cas simulé N°1, 5, 9 et 13). Les résultats des 12 autres cas simulés sont présentés dans l'Annexe A.

Cas simulé N° 1

Dans ce cas, le caractère discret simulé possède deux modalités (50%/50%), le QTL simulé a un effet fort (0,8) et est situé au milieu (50cM) du chromosome. Dans cette situation, le QTL est bien localisé par les deux modèles (tableau 2.2). Le modèle normal présente toutefois une puissance de détection plus élevée, avec un temps de calcul deux fois plus faible que pour le modèle discret. Quand le caractère étudié présente deux modalités, avec une fréquence équilibrée dans la population, le modèle normal est donc plus attractif pour la détection de QTL. Si on diminue l'effet du QTL simulé (cas simulé N°3 – annexe A), la puissance de détection reste toujours meilleure dans le modèle normal, 9,2% contre 8,3% pour le modèle discret. Toujours pour ce type de caractère, la position du QTL (qu'il soit à l'extrémité ou au milieu du chromosome) ne change rien quant à l'efficacité du modèle normal par rapport au modèle discret. Toutefois, quand le QTL est à l'extrémité du chromosome (20cM) et que le QTL présente un effet faible (cas simulé N°4 – annexe A), on note une très mauvaise estimation, par les deux modèles, de la position du QTL.

Tableau 2.2 : paramètres estimés pour les deux modèles comparés grâce aux données simulées du cas 1

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,49	0,49
Ecart-type	-	0,18	0,15
Erreur moyenne quadratique	-	0,03	0,02
puissance (%)	-	68,9	76,0

Cas simulé N°5

Le caractère a deux modalités en fréquences déséquilibrées (10%/90%). Le QTL simulé a un effet fort (0,80) et est situé au milieu (50cM) du chromosome (tableau 2.3). Dans ce cas, les précisions relatives à l'estimation de la position du QTL sont similaires et satisfaisantes (assez proche de la valeur vraie) pour les deux modèles. La puissance de détection du modèle normal est supérieure à celle du modèle discret. Si on place le QTL à l'extrémité du chromosome (20cM), tous les autres paramètres étant inchangés (cas simulé N°6 – annexe A), la puissance reste sensiblement la même, et ce pour les deux modèles. La position du QTL est assez bien estimée pour les deux modèles, tout comme le cas N°5. Le modèle normal reste toujours privilégié pour sa puissance de détection supérieure à celle du modèle discret, ce dernier quant à lui présente une meilleure précision pour l'estimation de la position du QTL. Par ailleurs, si on diminue l'effet de substitution du QTL (0,2 au lieu de 0,8), les autres paramètres étant inchangés (cas simulé N°7 – annexe A), la position du QTL est beaucoup moins bien estimée (erreur moyenne quadratique=0,10 pour les deux modèles). Enfin si le QTL présente un effet faible (0,20) et est positionné à l'extrémité du chromosome (cas simulé N°8 – annexe A), on obtient des précisions assez médiocres de l'estimation de la position du QTL : dans ce cas, le modèle discret présente un léger avantage (7,4% contre 6,3%) en termes de puissance de détection de QTL par rapport au modèle normal.

Tableau 2.3 : paramètres estimés pour les deux modèles comparés grâce aux données simulées du cas 5

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,45	0,48
Ecart-type	-	0,22	0,19
Erreur moyenne quadratique	-	0,05	0,04
puissance (%)	-	46,6	52,6

Cas simulé N°9

Dans ce cas, on simule un caractère en fréquences déséquilibrées à trois modalités (30%/30%/40%). Le QTL simulé présente un effet fort (0,80) et est situé au milieu (50cM) du chromosome (tableau 2.4). La précision de l'estimation de la position du QTL est assez satisfaisante pour les deux modèles. En effet, l'erreur moyenne quadratique est seulement de 0,03 pour le modèle normal et 0,04 pour le modèle discret. Le modèle normal présente une puissance supérieure de 5 points par rapport au modèle discret. Avec le coût en temps de calcul élevé du modèle discret par rapport au modèle normal (quasiment, deux fois plus de temps), le modèle normal peut être préféré au premier pour ce type de caractères. Les puissances de détection des deux modèles sont quasiment similaires (66,2% pour le modèle discret et 65,2% pour le modèle normal) si le QTL est positionné à l'extrémité du chromosome (cas simulé N°10 – annexe A), le reste des paramètres étant inchangé par rapport au cas 9. On remarque même un changement de tendance, dans la mesure où le modèle discret est plus puissant que le modèle normal dans le cas 9, l'écart restant néanmoins assez minime. Comme pour les autres types de caractères précédemment décrit, si le QTL est placé à 20cM et a un effet faible (cas simulé N°12 – annexe A), la précision de l'estimation de la position est assez mauvaise. En conclusion, les deux modèles présentent des précisions d'estimation et des puissances de détection comparables, on note toutefois une légère supériorité pour le modèle normal quand le QTL recherché est suffisamment fort et qu'il n'est pas à l'extrémité du chromosome.

Tableau 2.4 : paramètres estimés pour les deux modèles comparés grâce aux données simulées du cas 9

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,49	0,50
Ecart-type	-	0,19	0,16
Erreur moyenne quadratique	-	0,04	0,03
puissance (%)	-	59,2	64,8

Cas simulé N°13

Cette fois, un caractère à trois modalités en fréquences très inégales (10%/10%/80%) est simulé. Le QTL simulé présente un effet fort (0,8) et est situé au milieu (50cM) du chromosome (tableau 2.5). Ici encore, les paramètres de détection sont bien estimés par les deux modèles, les erreurs moyennes quadratiques étant négligeables. La puissance de détection du modèle discret est presque deux fois plus forte par rapport au modèle normal. Le modèle discret est donc le plus adapté pour ce type de caractère. Toutefois le temps de calcul (environ 2heures pour les 500 simulations ie autour de 20 secondes par analyse) peut constituer une barrière à l'utilisation de manière systématique de ce modèle. Si le QTL est positionné à l'extrémité du chromosome (cas simulé N°14 – annexe A), les précisions des estimations restent similaires et l'intérêt en termes de puissance de détection du modèle discret s'accentue. En effet, le modèle discret présente dans ce cas, une puissance de 67,1% contre 34,5% pour le modèle normal. Si, en plus d'être situé à

l'extrémité du chromosome, le QTL présente un effet de substitution faible (cas simulé N°16 – annexe A), les estimations de la position ne sont pas précises. L'erreur moyenne quadratique par rapport à la position vraie est de 0,16 pour le modèle discret, elle est de 0,18 pour le modèle normal. Dans ce dernier cas, la puissance de détection du modèle discret est 2,7 fois supérieure à la puissance du modèle normal. En conclusion, pour des caractères de ce type (caractères à trois modalités et de distribution déséquilibrée) le modèle discret semble plus approprié pour la détection de QTL et ce, quelque soit les caractéristiques du QTL simulé.

Tableau 2.5 : paramètres estimés pour les deux modèles comparés grâce aux données simulées du cas 13

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,50	0,51
Ecart-type	-	0,16	0,23
Erreur moyenne quadratique	-	0,03	0,05
puissance (%)	-	67,4	36,7

2.5 Discussion et conclusion

Précisions des paramètres

Dans la plupart des cas, les précisions des estimations des paramètres par les deux modèles sont satisfaisantes. Quand le QTL simulé présente un effet de substitution élevé sa position est toujours bien estimée. En effet, l'erreur moyenne quadratique (EMQ) est inférieure à 0,05 dans 75% des cas. Si le QTL n'est responsable que d'un faible pourcentage de la variabilité du caractère et qu'il est positionné au milieu du chromosome, les simulations montrent que les précisions des estimations restent acceptables (l'EMQ ne dépasse guère 0,06), et ce pour les deux modèles étudiés. Si l'effet de substitution du QTL est faible et qu'il est situé à l'extrémité du chromosome, sa position n'est jamais bien estimée.

Temps de calcul

Le modèle discret est beaucoup plus gourmand en temps de calcul que le modèle normal. Cette observation étant faite pour un code informatique faiblement optimisé, il faut en moyenne deux fois plus de temps, pour l'analyse par rapport au modèle normal. Insistons sur le fait que le modèle normal a été mis en place et programmé dans le logiciel QTLMap (Filangi *et al.,* 2010) bien avant le modèle discret, et a donc bénéficié d'un certain nombre d'améliorations qui optimisent les temps de calcul. A moyen terme, des améliorations de ce genre seront opérées pour le modèle discret. Quoiqu'il en soit, les temps de calcul resteront plus importants par la nature même de ce modèle non linéaire.

Modèle normal ou modèle discret ?

En oubliant le temps de calcul qui est amené à évoluer, la précision de la position du QTL estimée et la puissance de détection sont des éléments solides,

pour comparer les deux modèles étudiés. Pour ce qui est de l'estimation de la position, les deux modèles présentent des précisions intéressantes et assez similaires : dans 88% des cas, l'écart des EMQ entre ces 2 modèles sont inférieurs à deux points. Ces observations sont conformes à la littérature (Visscher *et al.,* 1996 ; Rebai, 1997 ; Kadarmideen *et al.,* 2000). La différence des précisions peut être jugée significative dans deux (cas simulés N°8 et N°14) des 16 cas simulés, et c'est le modèle discret qui s'avère le plus efficace. Il s'agit de cas où les modalités du caractère sont déséquilibrées dans la population et le QTL placé à l'extrémité du chromosome.

La puissance de détection est le critère le plus discriminant entre les deux modèles. Si on ne s'intéresse qu'à ce critère, le modèle normal est plus puissant pour les caractères binaires. Cette conclusion est vraie quelque soit les caractéristiques du QTL simulés. Quand la fréquence des modalités est déséquilibrée, le modèle normal reste néanmoins efficace sauf quand l'effet du QTL est très faible et que le QTL est placé à l'extrémité du chromosome. Il est fait état dans la littérature de ce faible apport du modèle discret, en termes de précision d'estimation et de puissance de détection par rapport au modèle normal, pour des caractères binaires avec des modalités en fréquences déséquilibrées. Yi et Xu (1999) montrent même que l'utilisation d'un modèle à seuils a plus d'apport en termes de précision d'estimation et de puissance de détection pour un caractère binaire avec des modalités en fréquences équilibrées, ces apports étant moindres voire inexistants pour un caractère binaire avec des modalités en fréquences déséquilibrées dans la population. A noter que dans ce dernier cas, les estimations des positions sont assez moyennes (EMQ=0,16 pour le modèle discret et 0,19 pour le modèle normal).

Les caractères à trois modalités se comportent différemment. En effet, si les modalités sont distribuées de façon homogène (30%/30%/40%) dans la population, le modèle normal est toujours le plus puissant, excepté quand le QTL est situé à l'extrémité du chromosome où l'on observe des puissances du même ordre pour les deux modèles. Par contre, si les modalités sont distribuées de façon déséquilibrée le modèle discret présente la puissance de détection la plus élevée, et ce quelque soit les caractéristiques du QTL.

Conclusion

Nous montrons donc que le modèle discret est plus précis et plus puissant quand le caractère étudié possède trois modalités distribuées de façon déséquilibrée dans la population. Ce résultat n'a pour l'heure jamais été publié. Toutefois, Hackett et Weller (1995) montrent l'intérêt du caractère discret en termes de précision des estimations de paramètres pour un caractère à cinq modalités, avec une hétérogénéité de fréquences dans la population. Selon ces mêmes auteurs, si les 5 modalités du caractère sont représentées de façon équilibrée dans la population, les deux modèles étudiés présentent des précisions similaires.

Pour les caractères binaires, le modèle normal présente des précisions et une puissance de détection similaires, voire légèrement meilleures, que le modèle discret, exception faite du cas où le QTL est placé à l'extrémité du chromosome avec un effet de substitution faible. Dans ce cas particulier, les précisions des estimations sont médiocres, indépendamment des modèles, et les puissances très faibles.

Perspectives

Même si nous apportons quelques éclaircissements sur l'intérêt d'utiliser le modèle discret, utilisant la sous-jacente normale aux caractères discrets, nos résultats doivent être confirmés et complétés par des simulations supplémentaires. En effet, l'effectif devra être accru dans ces nouvelles simulations pour mesurer l'influence de la taille du protocole sur l'efficacité de l'un ou l'autre des deux modèles : cela peut être envisagé en doublant le nombre de pères fondateurs ou des descendants par mère.

Une autre piste à explorer serait de quantifier l'effet de l'héritabilité du caractère sur la robustesse et la précision des estimations pour les deux modèles considérés. En effet considérant une héritabilité fixe pour tous nos simulations, nous ne sommes pas en mesure de dire si oui ou non, ce paramètre a une quelconque influence dans l'appréciation des modèles. Cependant, ces nouvelles simulations ne peuvent être réalisées que si le temps de calcul est significativement réduit, notamment pour le modèle discret qui, en l'état, est assez lourd à mettre en œuvre.

Chapitre III. Le canard Mulard : zootechnie, paramètres génétiques et QTLs connus pour les caractères d'intérêt

1 Génétique et zootechnie du canard mulard

1.1 Introduction

Le canard mulard, hybride infertile produit d'un croisement intergénérique entre une cane commune (*Anas platyrhynchos*) et un canard de Barbarie (*Cairina moschata*), fournit à lui seul environ 92% de la production de foie gras en France. Connu en France depuis le XVII siècle (Duhart, 2004), sa rusticité en élevage, sa facilité de gavage, ses capacités d'ingestion et son foie gras de poids élevé et de bonne qualité font de lui le canard le plus gavé en France (Guy *et al.,* 1995). En effet, il bénéficie d'un effet d'hétérosis favorable pour la capacité d'ingestion et la production de foie gras (consommation alimentaire et poids de foie gras supérieurs à la moyenne des espèces parentales) et le taux de fonte lors de la pasteurisation ou de la stérilisation du foie gras est faible (Guy *et al.,* 1995). En outre le canard mulard est valorisé pour ses parties carnées comme le magret : une part majoritaire de la production de viande de canard est donc liée à la commercialisation du foie gras. Ainsi, en 2008, 145 535 tonnes de viande ont été produites par la filière palmipède gras contre 103 540 pour la filière canard maigre (source CIFOG, 2009).

Dans ce paragraphe nous ferons le point sur les connaissances d'ordre génétique concernant le mulard. Après la description des phases d'élevage et de gavage du canard mulard, cet hybride sera comparé aux deux souches parentales afin de présenter les spécificités du premier relativement à ses parents. Puis, une synthèse des paramètres génétiques des différents caractères d'intérêt chez le canard à gaver (caractère de qualité de foie gras et de magret notamment) sera présentée.

1.2 Les souches parentales

Le canard mulard est le produit du croisement entre une cane commune (*Anas platyrhynchos*) et un canard de Barbarie (*Cairina moschata*), deux espèces génétiquement isolées depuis 50 millions d'années (Brun *et al*., 2005).

On s'intéresse naturellement à la variante domestique du canard commun, sa variante sauvage étant connue sous le nom de canard colvert. Domestiqué depuis environ 4000 ans en Asie du sud-est, les canards communs sont notamment utilisés pour la production de la viande ou d'œufs de consommation (Brun *et al.,* 2005). Les races utilisées pour la production des œufs de consommation sont très pondeuses et de petit gabarit alors que celles utilisées pour la production de viande sont plus lourdes (Brun *et al.,* 2005). La race maternelle utilisée pour la production de mulards destinés au gavage en France, est un Pékin lourd au plumage blanc (Rouvier, 1992). Cette cane Pékin, utilisée comme mère du mulard, est généralement le résultat d'un croisement entre 2 lignées sélectionnées, l'une pour le poids corporel et l'autre pour la production d'œufs. Son poids à l'âge adulte est compris entre 4 et 5 kilos (Brun *et al*., 2005). Il existe d'autres types de canards communs, comme la cane Tsaiya qui est réputée très pondeuse et donc utilisé, notamment dans les pays du sud-est asiatique, pour la production d'œufs. On peut citer aussi, le canard de Rouen, élevé

généralement pour la production de viande : ce fut la race utilisée pour produire les mulards destinés au gavage, avant l'introduction du Pékin en France (Brun *et al.,* 2005).

Le canard de Barbarie a, lui, été domestiqué par les sud-amérindiens, puis a été introduit en Europe par les Espagnols au XVIe siècle (Duhart, 2004). Différentes souches de canard de Barbarie présentant des colorations variées, sont utilisées comme père du mulard (Migliore *et al.,* 1988).

L'accouplement naturel entre ces deux espèces de canards est rare et le taux de fertilité des œufs qui en résulte est faible (Watanabe, 1961 ; Rouvier *et al*., 1988). Il peut être favorisé par l'Homme qui peut intervenir afin d'immobiliser la cane. En effet, le comportement de la cane commune, réputée très agile, la rend réticente envers un accouplement avec le canard de Barbarie, qui, lui est un animal calme et silencieux.

L'insémination artificielle d'une cane commune par de la semence de Barbarie est bien maitrisée à Taiwan (Tai Lui J.J. et Tai C., 1984). Cette pratique est mise en œuvre deux fois par semaine à intervalles de temps successifs de 3 et 4 jours -la semence étant utilisée immédiatement après sa collecte- et conduit à des taux de fertilité pouvant atteindre les 80% (Rouvier *et al.,* 1984). Arrivée en France dans la deuxième moitié des années 80 (Rouvier *et al.,* 1988), l'insémination artificielle a grandement contribué à l'essor que connait la production de mulard hexagonale. En effet, la production du foie gras de mulard a été multipliée par 6 en 25 ans (1985-2010) (source : CIFOG, 2010).

Le croisement réciproque entre canard de Barbarie et canard commun est aussi possible, il est désigné par le terme de Hinny. Dans le croisement Hinny (mâle commun x femelle Barbarie), l'insémination artificielle n'est pas très efficace : son taux de réussite est autour de 45,8% contre 72,0% pour le croisement mulard (mâle Barbarie x femelle commun) (Brun *et al.,* 1999). Cela peut expliquer que le Hinny ne soit pas élevé pour le gavage, même si des études montrent sa capacité de produire un foie gras de poids élevé (Larzul *et al.,* 2006) sans toutefois atteindre les poids du foie gras du mulard.

1.3 *Zootechnie* du mulard

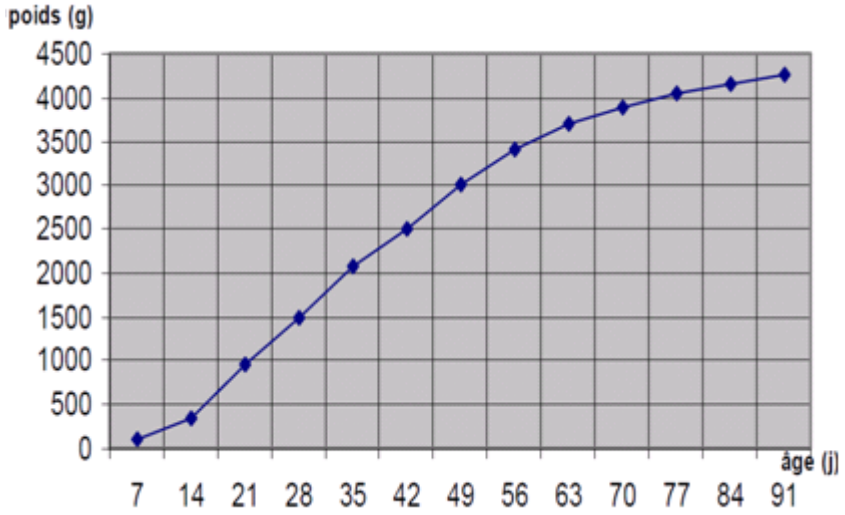
1.3.1 Phase de croissance du canard mulard destiné au gavage

Jusqu'à 12 semaines d'âge, l'animal est en phase de croissance (figure 3.1). Il est dans un premier temps (jusqu'à trois à quatre semaines d'âge) placé en poussinière, dans une ambiance contrôlée (salle ventilée, température basse et régulière, etc…) avant d'être sorti sur parcours (Guémené *et al.,* 2007). Cette première phase de croissance a pour objectif de préparer le canard à la mise en gavage. L'animal doit arriver en gavage avec une bonne musculature des pectoraux, le squelette bien développé et en excellente santé. Tout cela est contrôlé entre autres par la qualité et la quantité de la ration alimentaire administrée, la gestion de la densité d'élevage et les soins vétérinaires. Jusqu'à six semaines d'âge le canard

se nourrit à volonté : c'est un régime *ad libitum.* Au delà de la sixième semaine d'âge, la quantité d'aliment distribuée au lot est contrôlée (Guémené *et al.,* 2007). Réduite dans un premier temps jusqu'à dix semaines d'âge, la ration journalière du canard augmente graduellement de la dixième semaine d'âge jusqu'au début du gavage : c'est la période de pré-gavage. La courbe de croissance de l'animal (figure 3.1) présente un point d'inflexion vers la huitième semaine. Ensuite, principalement les pectoraux et les tissus adipeux continuent à se développer. Dans cette phase de pré-gavage, l'animal apprend à consommer un maximum d'aliment en peu de temps et peu de repas jusqu'à ce qu'il soit capable de prendre en 2 repas, sa ration quotidienne. Cela prépare aussi les hépatocytes à la surcharge lipidique et le tube digestif, au niveau du pseudo-jabot, à la distension (Guémené et Guy, 2004).

Cette première phase d'élevage est en général 6 à 7 fois plus longue que la phase de gavage proprement dite qui n'excède jamais deux semaines (Guémené *et al.,* 2007).

Figure 3.1 : Courbe de croissance théorique du canard mulard pendant la durée d'élevage

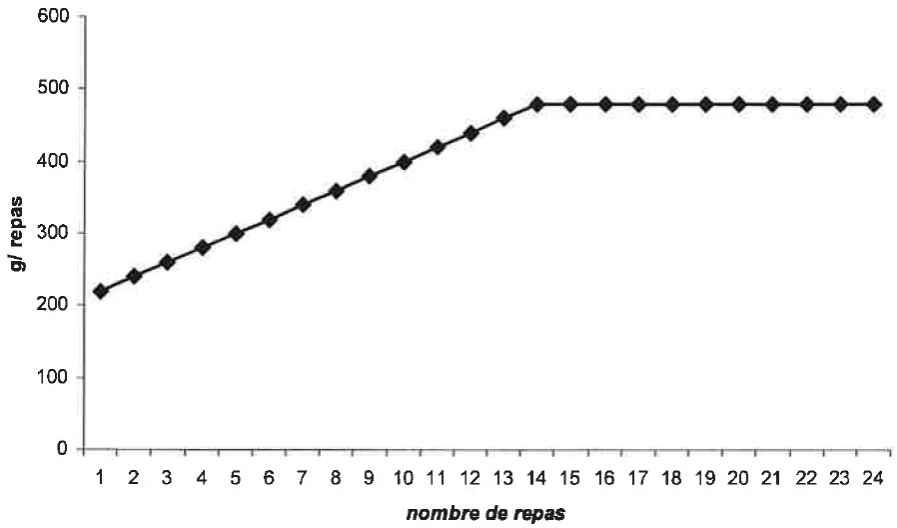


1.3.2 Phase de gavage du canard mulard

Le gavage consiste à faire consommer quotidiennement aux canards une quantité importante d'un aliment riche en amidon (du maïs généralement) très énergétique qui conduit à un engraissement intensif et à la formation de foie gras. L'animal est en général nourri deux fois par jour, avec des quantités d'aliments progressivement de plus en plus importantes. Il existe plusieurs plans de distribution : on peut citer le modèle « palier tardif » qui consiste à distribuer de façon croissante, en augmentant de 20 g d'un repas à l'autre (figure 3.2), pour atteindre un pallier autour de 500g / repas. Les animaux sont logés, pendant cette période, dans des cages collectives de 3 ou 4 canards ou des cages individuelles. Ce dernier type de logement est amené à disparaitre à l'échéance 2015 à cause des nouvelles

réglementations européennes (Directive 98/58/CE) sur le bien-être animal. Pourtant, il présentait plusieurs avantages du point de vue du gaveur : lors du gavage, l'animal se présentait toujours dans la bonne position et il n'existait pas de risque de confondre les animaux c'est-à-dire d'attraper un canard déjà gavé (Guémené et Guy, 2004).

Figure 3.2 : Courbe de gavage théorique – modèle « palier tardif »



1.4 Les caractères zootechniques d'intérêt du canard mulard

Contrairement aux pays du sud-est asiatique, l'élevage du mulard exclusivement pour la production de viande n'existe pas en France : c'est le canard de Barbarie qui est élevé à ces fins. En France, le canard mulard est spécifiquement élevé pour la production de foie gras, néanmoins sa carcasse est aussi bien valorisée. Cela conduit à prendre en compte, en plus de son aptitude au gavage, du poids et de la qualité du foie, la qualité et le poids des éléments qui composent la carcasse du mulard tels que le magret et la cuisse. Une large gamme de caractères couvrant la croissance, le comportement, l'aptitude au gavage, la composition corporelle et l'efficacité alimentaire a donc été étudiée chez le canard mulard.

1.4.1 La croissance

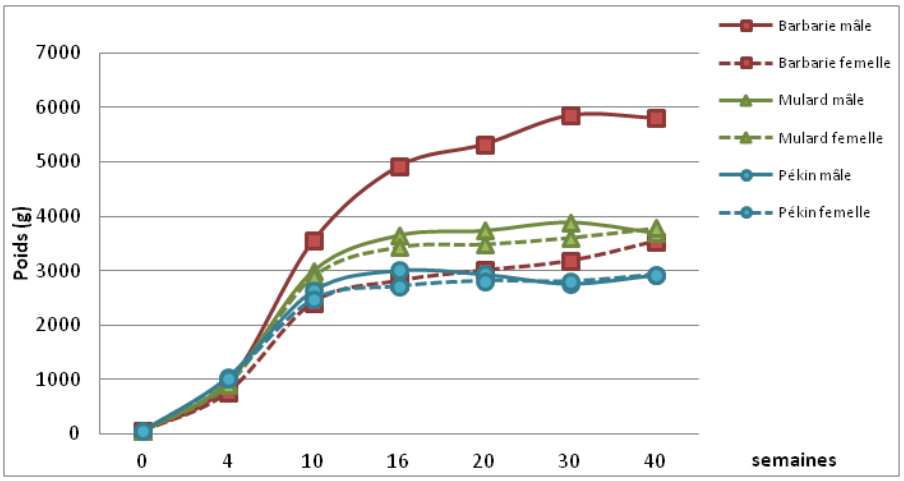
La courbe de croissance du canard a beaucoup été étudiée (Knizetova *et al.,* 1991 ; Gille et Salomon, 1994 ; Knizetova *et al.,* 1995 ; Baéza *et al.,* 1997 ; Maruyama *et al.,* 1999 ; Vitezica *et al.,* 2010). C'est une courbe sigmoïde dont les caractéristiques (l'âge au point d'inflexion et le poids asymptotique) varient d'une

espèce à l'autre. L'âge au point d'inflexion est compris entre 24,1 et 26,8 jours pour les canes Pekin (Knizetova *et al.,* 1991). Il est plus tardif pour les mâles de Barbarie (entre 34,6 et 37,2 jours selon Béaza *et al.* 1997). Pour le canard mulard, la croissance est intermédiaire entre celles des deux espèces parentales : l'âge au point d'inflexion est situé entre 25,0 et 30,5 jours (Gille et Salomon, 1994 ; Setiawan *et al.,* 1994).

La croissance du canard mulard jusqu'à son abattage a souvent été comparée à celles de ses deux souches parentales (Olver *et al.,* 1977), les espèces étant abattues plus ou moins tôt selon leur niveau de précocité. Contrairement au canard mulard et au Pékin, le canard de Barbarie présente un dimorphisme sexuel assez prononcé au niveau du poids vif (Baéza *et al.,* 2001). Cela est vrai aussi pour le Hinny, ce qui laisse supposer le rôle de la femelle Barbarie dans ce dimorphisme sexuel (Marie-Etancelin *et al.,* 2008).

Jusqu'à dix semaines d'âge, les canards mulards et les canards communs sont plus lourds que les canards de Barbarie (Olver *et al.,* 1977). A partir de dix semaines, et en prenant en compte le sexe des animaux, c'est le mâle de Barbarie qui est le plus lourd et le canard commun le plus léger, le mulard ayant un poids intermédiaire (figure 3.3 - Tai et Rouvier, 1998). Pour les femelles, la cane mularde s'avère être la plus lourde (Tai et Rouvier, 1998) par rapport aux canes des deux souches parentales (la cane commune et la cane de Barbarie).

Figure 3.3 : Courbes des moyennes de poids mesurés à différents âges pour chacun des sexes des trois génotypes (Tai et Rouvier, 1998)



1.4.2 L'efficacité alimentaire

L'efficacité alimentaire en croissance correspond au gain en poids vif observé au cours d'une période donnée et exprimé en pourcent de la quantité d'aliment ingérée au cours de cette même période. Cette notion peut être étendue à la phase de gavage en mesurant le poids du foie gras de l'animal par rapport à la quantité de

maïs ingéré. Dans ces 2 cas, l'objectif principal est la réduction des coûts alimentaires.

Durant la période de croissance, l'étude des indices de consommation à différents âges montre que les mulards présentent une efficacité alimentaire intermédiaire à celles des espèces parentales (Guy *et al*., 1999 ; Baéza *et al*., 2005). En effet, pour des canards élevés jusqu'à 94 jours, Guy *et al*. (1999) ont obtenu des indices de consommation de 4,41 pour le canard Pekin, 4,08 pour le mulard et 3,22 pour le canard de Barbarie.

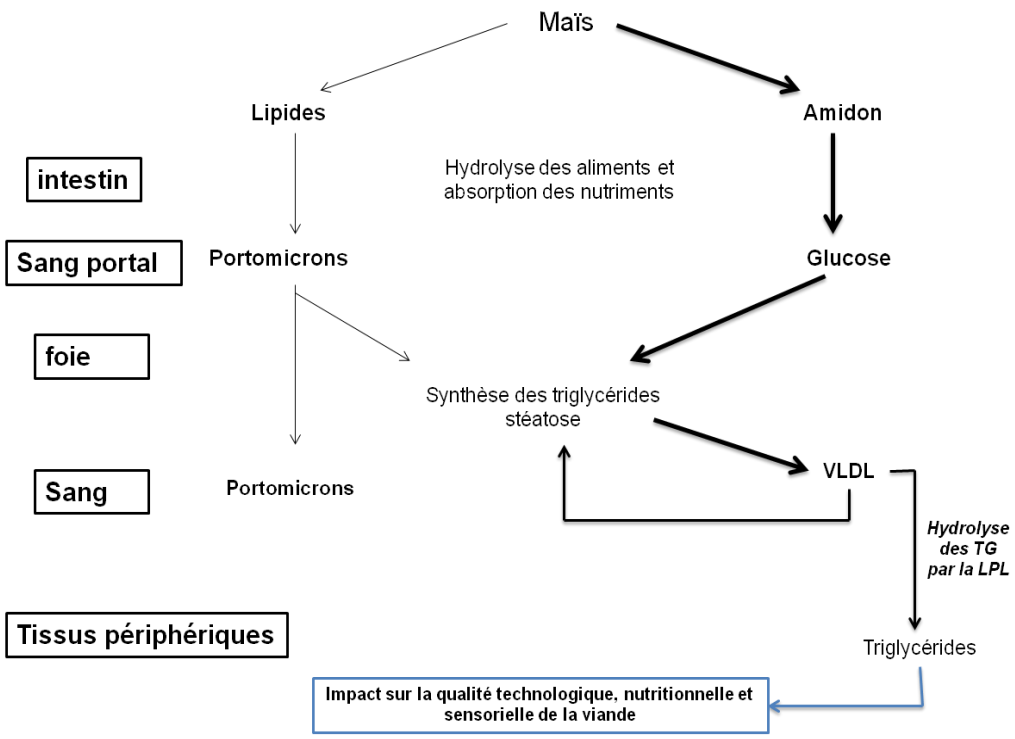
En gavage, lorsque les aliments ingérés sont proportionnels au poids vif de l'animal, Davail *et al*. (2003) montrent que le canard de Barbarie présente le poids de foie le plus élevé par rapport aux autres génotypes (Pékin et mulard). Mais lorsque les animaux sont gavés selon leur capacité maximale d'ingestion, le mulard présente le poids de foie le plus élevé (Guy *et al*., 1995) par rapport au Barbarie. Ces résultats laissent penser qu'en gavage, le canard le plus efficace est le canard de Barbarie ; la capacité du mulard à produire du foie gras est plus liée à sa capacité d'ingestion (Davail *et al*., 2003) qu'à son efficacité de transformation de l'aliment ingéré.

1.4.3 La composition corporelle et aptitude au gavage

La formation du foie gras chez le canard résulte de la stéatose hépatique, synthèse et stockage de lipides dans le foie provoqués par le gavage. La stéatose hépatique est complètement réversible chez les palmipèdes (Salichon, 1991). C'est la réponse de l'animal à un apport important de maïs (amidon). En effet, chez les oiseaux et contrairement aux mammifères, en vue d'un stockage d'énergie avant la migration, le foie est le site principal de la lipogenèse de novo (Léveille *et al*. 1975).

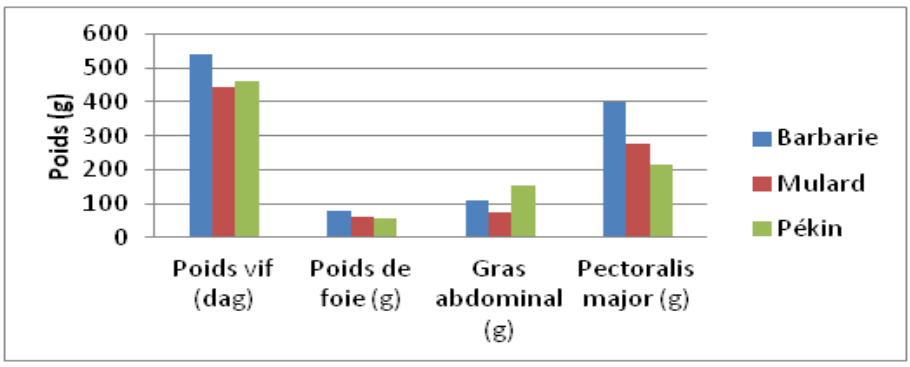
Les mécanismes physiologiques d'obtention de foie gras ont fait l'objet d'une revue bibliographique (Baéza *et al.,* 2005). L'alimentation des animaux pendant le gavage induit un apport massif de glucides (figure 3.4). Lorsque la capacité de lipogenèse induite par le gavage dépasse la capacité de synthèse et de sécrétion des VLDL (Very Low Density Lipoprotein ou lipoprotéine de très basse densité, groupe de lipoprotéines dont le rôle est de transporter le cholestérol dans le sang pour les apporter aux cellules), les triglycérides néo-synthétisés s'accumulent dans les hépatocytes. De plus, lors de ces afflux massifs de triglycérides sanguins sous forme de VLDL, la lipoprotéine lipase (LPL) n'est plus en capacité de tout hydrolyser au niveau des tissus périphériques : les triglycérides sanguins vont alors retourner au foie et contribuer à la stéatose hépatique (Hermier *et al*., 1999, Davail *et al*., 2003).

Figure 3.4 : Mécanismes de la stéatose hépatique chez les palmipèdes (Hermier *et al*., 1999)



Pour des animaux non gavés, le poids de foie du Barbarie est le plus important, le mulard présentant un poids intermédiaire par rapport à ses deux souches parentales (figure 3.5 ; Chartrin *et al.,* 2004).

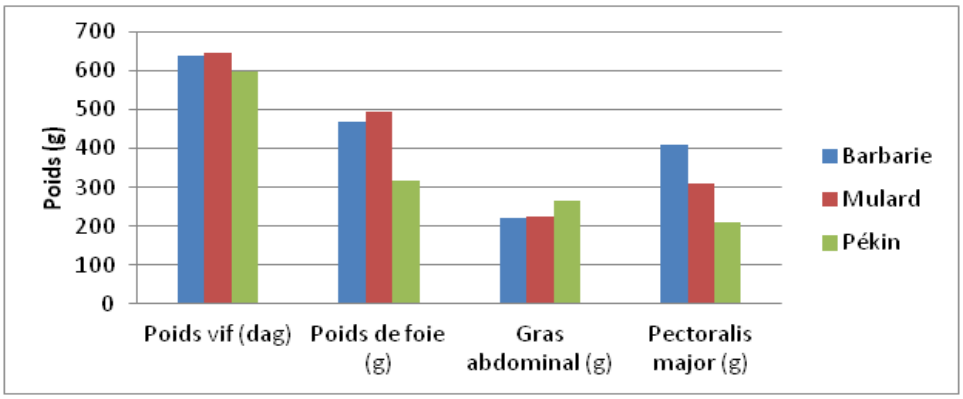
Figure 3.5 : Différence des poids vif, poids de foie, poids du gras abdominal et poids du Pectoralis major pour les différents génotypes – animaux non gavés (Chartrin *et al*., 2004)



Réglementairement, le poids du foie du canard doit être supérieur à 300g pour avoir l'appellation « foie gras » (décret 95-625). Grâce à l'utilisation de souches sélectionnées et la maîtrise des techniques de gavage, le poids de foie gras produit peut atteindre aujourd'hui les 700g (Theron, 2011), l'objectif moyen étant actuellement de 550g.

Lorsque les canards sont gavés selon leur capacité maximale d'ingestion, les canards mulards présentent le poids de foie le plus élevé (figure 3.6 ; Chartrin *et al.,* 2004). Le canard Pékin gavé présente toujours le poids de foie le plus faible en comparaison avec les autres génotypes, ce qui explique que celui-ci ne soit pas élevé pour la production de foie gras. Si on ajuste l'ingéré alimentaire au poids vif des animaux, c'est le canard de Barbarie qui présente le poids le plus élevé de foie gras, suivi du mulard puis du Pékin (Davail *et al.,* 2003). Cela suggère que la capacité d'ingestion du canard mulard a un effet immédiat sur sa capacité à produire du foie gras.

Figure 3.6 : Différence des poids vif, poids de foie, poids du gras abdominal et poids du *Pectoralis major* pour les différents génotypes – animaux gavés (Chartrin *et al.,* 2004)



A noter que le poids de foie comme le poids vif de l'animal ou le poids du gras abdominal sont très différents selon que l'animal est gavé ou pas (figures 3.5 et 3.6). Le gavage du canard a donc naturellement un impact sur sa composition corporelle. Enfin on note que le canard de Barbarie, gavé ou maigre, présente la proportion ou le poids de muscle le plus élevé (Chartrin *et al.,* 2004).

1.4.4 La qualité des produits

Les deux produits valorisés du canard mulard sont le foie gras, essentiellement vendu transformé (71% de foies gras prêt à consommer contre 29% de foie gras cru en 2009 - CIFOG, 2009) et la viande (magret, cuisse, aiguillette). La qualité du foie gras de mulard est souvent comparée à celle du foie des canards de Barbarie, le canard commun n'ayant pas d'intérêt zootechnique à être gavé.
Au-delà du poids, le principal critère de qualité du foie gras est le taux de fonte, critère technologique mesuré par la quantité de lipides exsudés après pasteurisation ou stérilisation. D'autres critères tels que la couleur et la texture sont utilisés pour apprécier la qualité du foie. Le foie gras du canard mulard est reconnu pour être le meilleur sur ce critère de rendement technologique. Différentes études menées pour comparer le taux de fonte des foies gras (taux de fonte mesuré sur un échantillon de 60g de foie placé pendant 50 minutes en autoclave à 105 °C) des deux génotypes de

canards gavés (le canard de Barbarie et le mulard) montrent, en moyenne, un taux de fonte supérieur de 10 points pour le canard de Barbarie : 45% pour le Barbarie et 35% pour le mulard selon Babilé *et al.,* 1988, et respectivement 55,6% contre 43,9% selon Salichon *et al.,* 1993. Dufour *et al.* (1993) ont observé une corrélation de 0,55 entre le poids de foie et le taux de fonte. En étudiant des foies de même poids et engraissement mais avec des taux de fonte différents, Theron (2011) met en avant une possible relation entre le niveau de fonte lipidique à la cuisson du foie gras et le stade physiologique de développement de la stéatose. Selon cet auteur, deux canards soumis au même rythme de gavage peuvent présenter des états différents de stéatose : l'un ayant atteint sa capacité maximale avant la fin de la période de gavage, stabilisant son poids de foie, mettrait en place des processus de défense cellulaire, alors que l'autre poursuit sa stéatose et présenterait préférentiellement des processus anaboliques. Ainsi, à l'abattage, pour des poids de foie similaires, le premier présenterait des qualités technologiques amoindries.

Actuellement, la texture du foie gras, appréciée par palpation, est prise en compte par les transformateurs dans l'évaluation de la qualité de ce dernier, et permet d'orienter les foies dans différentes voies technologiques (foies gras entier, bloc de foie gras…). Ils considèrent cette appréciation comme une mesure indirecte du taux de fonte (qui est une mesure destructive) : chez le canard mulard, plus le foie est ferme, plus il présentera un taux de fonte élevé (Robin *et al.,* 2004).

La couleur du foie gras à froid constitue aussi un critère d'évaluation : elle est appréciée à l'œil à l'abattoir. Elle peut être objectivée via le chromamètre qui traduit la couleur en valeurs numériques (système L*, a*, b*, norme CIE 1976, qui correspondent respectivement, à la luminance, l'indice de coloration rouge et l'indice de coloration jaune). Le type de maïs utilisé pour le gavage semble avoir une incidence sur l'indice de coloration jaune (Castaing *et al.*,1995) du foie, alors que l'indice de rouge signe principalement des défauts de présentation de types pétéchies, hématomes, saignées incomplètes…

La qualité de la viande de canard peut être appréciée selon différents critères : sa coloration, sa tendreté, son pH dans les muscles ainsi que les pertes sous vide ou à la cuisson. Pour la coloration, Baéza *et al. (*2005) ont montré que le canard de Barbarie, gavé ou pas, avait des filets plus clairs et plus jaunes et le canard Pékin, plus foncés et moins jaunes. Les canards mulards présentent eux des muscles avec les intensités de rouge les plus élevées et le Barbarie les moins élevées (tableau 3.1).

Tableau 3.1 : Effet du génotype et du gavage sur la couleur et les pertes à la cuisson des filets de canards âgés de 14 semaines (extrait de Baéza *et al.,* 2005)

Paramètre		Couleur du filet			Pertes à la cuisson (%)
		L*	a*	b*	
Pékin	gavés	41,51	13,79	13,63	21,08
	Non gavés	30,85	13,66	8,20	16,90
Mulard	gavés	41,18	14,21	14,33	15,86
	Non gavés	31,90	13,78	8,93	14,51
Barbarie	gavés	43,63	12,80	14,73	16,98
	Non gavés	34,97	13,63	11,43	14,63
Effet génotype		***	*	***	***
Effet gavage		***	ns	***	***

*, **, *** : Effet significatif à 5%, 1% ou 0,1% - ns : effet non significatif

La tendreté du muscle n'est pas appréciée en routine. Elle peut néanmoins être mesurée par la force de cisaillement ou par analyse sensorielle. Le test de cisaillement est réalisé sur la viande crue ou après congélation/décongélation/cuisson de la viande. Pour se faire, les échantillons de muscle sont découpés parallèlement au sens des fibres. Le cisaillement est effectué par une lame à ouverture triangulaire équipée de capteur de force. Selon Smith *et al.* (1992), pour un même âge à l'abattage (7 semaines), la viande de canard est plus ferme que celle du poulet. Entre différents génotypes de canard, le canard Pékin présente les filets les plus tendres. Les filets du canard de Barbarie sont eux, les plus durs (Larzul *et al.,* 2002). Les mêmes auteurs mettent en évidence que le pH *post-mortem* dans les magrets est sensiblement le même chez les différents génotypes de canard (Pékin, Barbarie et mulard). Les canards mulards présentent les taux de perte à la cuisson les plus bas comparés à ses deux souches parentales (tableau 3.1).

1.4.5 Le comportement

Pour le bien être de l'animal, différentes études sur le comportement (réactions de stress et d'émotivité) du mulard face à l'acte de gavage ont été menées (Guémené *et al.,* 2004). Il a été montré que les manipulations, comme la capture et le gavage du mulard, sont sources de stress, qui se traduisent par des hausses des taux de corticostérone sanguins chez le canard (Noirault *et al.,* 1999). Cependant, il

semble que le gavage seul n'induise pas une augmentation significative de taux de corticostérone notamment lorsque les canards sont élevés en cages individuelles (Guémené *et al.* 2007). Comparé à ses deux espèces parentales, le canard mulard présente une réactivité intermédiaire si on apprécie cette réactivité par la mesure du niveau de corticostérone. Mais si le caractère mesuré est l'évitement de l'Homme, le canard mulard présente des réactions extrêmes par rapport au Pékin et au Barbarie (Faure *et al.,* 2001). L'hypothèse d'un lien de causalité entre coloration blanche du plumage des canards mulards et expression de comportement évoquant une réponse de ces derniers à la peur a été étudiée (Guémené *et al.,* 2002 ; Guémené *et al.,* 2004). Ces auteurs notent une légère différence de réactions des canards (niveau de corticostérone pour le test de l'immobilité tonique par exemple) selon leurs colorations de plumage. Cette différence est toutefois insuffisante pour valider cette hypothèse de lien de causalité.

1.5 Les déterminismes génétiques des caractères d'intérêt chez le canard mulard

Sont passés en revue dans ce paragraphe, les paramètres génétiques (héritabilités et corrélations génétiques) calculés pour les différents types de caractères d'intérêt étudiés chez le canard mulard. Il s'agit notamment des caractères relatifs à la croissance du mulard, à son efficacité alimentaire, à son aptitude au gavage et à la qualité de ses produits mais aussi des caractères de reproduction de la cane commune. Le canard mulard étant un hybride stérile, il est impossible de créer une population de mulard pour calculer les paramètres génétiques directement sur cette dernière. Ces paramètres génétiques sont donc estimés dans les deux souches parentales pour des performances mesurées sur le mulard. Sachant qu'un parent donné ne transmet que la moitié de sa valeur génétique additive, l'héritabilité mesurée pour les canards mulards est en espérance 4 fois plus petite que celle mesurée pour l'une ou l'autre de ses populations parentales pures. De nouveaux logiciels permettent d'utiliser des modèles génétiques avec 2 matrices de parentés indépendantes (Commun versus Barbarie) pour calculer simultanément les paramètres génétiques dans les deux souches parentales. Ces modèles prennent en compte l'absence totale de lien de parenté inter-populations parentales (Chapuis et Larzul, 2006 ; Marie-Etancelin *et al*., 2011).

La connaissance des paramètres génétiques est nécessaire pour optimiser l'amélioration génétique des performances du mulard dans chacune des 2 populations parentales. Lorsque les caractères d'intérêt mesurés sur les mulards sont fortement corrélés génétiquement avec ces mêmes caractères mesurés sur les populations parentales, l'amélioration génétique des mulards pour lesdits caractères peut se faire en sélectionnant directement sur les caractères mesurés dans les populations parentales. Les paramètres génétiques des caractères du canard mulard ont fait l'objet d'une synthèse bibliographique par Marie-Etancelin *et al.* (2006).

1.5.1 Caractères liés à l'élevage des canards mulards

Différentes études ont été menées pour estimer les héritabilités des caractères de croissance du mulard dans les populations parentales : des corrélations génétiques ont été ainsi publiées entre les performances en pure (principalement chez le canard Pékin) et en croisement (Larzul *et al*., 2002 ; Brun et Larzul, 2003).

Sur les populations de canards de Barbarie, le poids vifs présente des héritabilités variant entre 0,24 à 0,58 (Mignon-Grasteau *et al.,* 1998 ; Hu *et al.,* 1999) selon l'âge où la pesée a été faite. Ces héritabilités varient aussi en fonction du sexe de l'animal. Ces mêmes auteurs mettent en évidences des héritabilités supérieures chez les femelles d'environ 0,1 points par rapport aux mâles, les caractères de poids étant considérées comme deux caractères différents dans les deux sexes à cause du dimorphisme sexuel important chez le canard de Barbarie.

Le poids vif du canard Pékin est tout aussi héritable : Larzul *et al*. (2002) et Brun *et al*. (2003) estiment des héritabilités respectives de 0,37 et 0,55 pour le poids à six semaines d'âge. Lorsque l'héritabilité de ce même caractère mesuré sur des canards mulards est estimée dans la population de canard commun, la valeur d'héritabilité obtenue est de l'ordre de 0,25 à 0,29 selon ces auteurs. La corrélation génétique mesurée dans la population Pékin, entre le poids de la cane Pékin et le poids vif du canard mulard est de 0,74 (Larzul *et al*., 2002). Cette corrélation est de 0,65 entre le poids de la cane Pékin et le poids de paletot (correspondant à la peau, le gras sous-cutané, les muscle et les os du dos du palmipède gras) du canard mulard (Larzul *et al*., 2002). La sélection du canard Pékin sur sa croissance améliore donc le poids vif et la quantité de muscle du canard mulard.

La variabilité génétique des caractères relatifs à l'efficacité alimentaire n'a pas été étudiée chez le canard mulard. Néanmoins, Larzul *et al.* (2004) ont mis en évidence qu'une amélioration de l'indice de consommation dans la population maternelle est en partie transmise aux descendants mulards.

Les réactions de stress du canard mulard, notamment au contact de l'Homme, ont fait l'objet de quelques études. Le niveau de corticostérone après le test d'open-field, et de contention dans le filet ainsi que la fréquence de déplacement lors du test d'open-field sont très héritables. Ainsi, en considérant ces caractères mesurés sur les mulards comme étant des mesures répétées attribuées à la cane commune, ils présentent des héritabilités respectives de 0,38 ; 0,37 et 0,33 pour les tests d'open-field, de contention et de déplacement (Guémené *et al*., 2004).

Pour ce qui est des effets d'hétérosis, il a été montré que la croissance des mulards est supérieure à celle des espèces parentales, avec, lorsque les mulards sont encore peu âgés, des effets d'hétérosis de l'ordre de 4% (Larzul *et al.,* 2006). Au delà de 16 semaines d'âge (après la fin de la phase de gavage), la croissance des mulards est bien inférieure à celle de ses parentaux, avec des effets d'hétérosis estimés entre -6% et -12% selon l'âge considéré (Tai et Rouvier, 1998).

1.5.2 Caractères liés au gavage des canards mulards

Le calcul des paramètres génétiques dans les populations parentales du canard mulard prend toute son importance ici. En effet les caractères liés au gavage (aptitude au gavage et surtout qualité des produits) ne sont pas mesurables dans la population de canes communes. La sélection des populations parentales sur ces caractères nécessite donc un testage sur descendance, c'est à dire des mesures de performances mesurées sur les mulards.

Le déterminisme génétique de l'aptitude au gavage du canard mulard dans les deux populations parentales a fait l'objet de plusieurs études. Le poids de foie est le caractère le plus étudié avec les poids de la carcasse et de ses composants, que cela soit en population pure de Barbarie gavé (Babilé, 1989 ; Mignon-Grasteau *et al.,* 1998) ou avec des performances mesurées sur des canards mulards (Poujardieu *et al.,* 1994, Larzul *et al.,* 2002 ; Chapuis et Larzul, 2006 ; Marie-Etancelin *et al.,* 2011). Chez le canard de Barbarie gavé, le poids de foie gras présente une héritabilité élevée de 0,39 (Mignon-Grasteau *et al.,* 1998). Une synthèse d'études publiées par Poujardieu *et al.* (1994) fait état d'héritabilités du poids de foie de Barbarie gavé variant entre 0,20 et 0,40. Les héritabilités estimées pour les poids de foie gras de mulard dans la population commune varie entre 0,06 et 0,18 (Poujardieu *et al.,* 1994 ; Larzul *et al.,* 2002 ; Chapuis et Larzul, 2006 ; Marie-Etancelin *et al*., 2011). Ces 2 derniers auteurs ont estimé simultanément dans les deux populations parentales l'héritabilité du poids de foie du mulard : ils ont obtenu des estimations d'héritabilité dans la voie Barbarie de 0,08 et 0,09 respectivement. Néanmoins, il n'existe pas d'étude estimant la corrélation génétique entre le poids de foie gras du Barbarie et celui du mulard, permettant de conclure ou pas à la possibilité de sélectionner le poids de foie du mulard directement via celui du Barbarie.

Le taux de fonte à la cuisson du foie gras, principal indicateur de qualité de ce dernier, présente une variabilité génétique moyenne selon différentes études menées (Babilé, 1989 ; Poujardieu *et al*., 1994). Babilé (1989) estime l'héritabilité de la perte d'exsudat lipidique à 0,23 avec des performances mesurées sur le Barbarie et Poujardieu *et al*. (1994) l'estiment à 0,18 pour le foie gras du canard mulard dans la population commune. Marie-Etancelin *et al*. (2011) confirme ce résultat avec une héritabilité estimé à 0,22 pour le taux de fonte du foie gras du mulard dans la population commune. Le poids de foie est très corrélé génétiquement avec le taux de fonte. Cette corrélation est estimée à 0,89 par Poujardieu *et al*. (1994) sur la voie commune et environ 0,80 par Marie-Etancelin *et al*. (2011) dans chacune des 2 voies parentales.

Ces derniers auteurs s'intéressent aussi à des caractères de coloration de foie gras du mulard ainsi que les taux de lipides, de collagène et de protéines dans le foie. Les héritabilités obtenus sont moyennes (0,10 pour la luminosité ; 0,11 pour la coloration rouge et 0,13 pour la coloration jaune) dans la population commune et généralement plus faibles dans la population de Barbarie (0,05 pour la luminosité ; 0,06 pour la coloration rouge), sauf pour la coloration jaune du foie qui présente une héritabilité similaire dans les 2 populations parentales (0,12 chez le Barbarie) (Marie-Etancelin *et al*., 2011). La coloration rouge et la luminosité du foie sont négativement corrélées dans les deux populations (-0,87 en commun et -0,77 en Barbarie) alors que la coloration jaune ne semble pas être liée aux deux premiers caractères (luminosité

et coloration rouge du foie) dans la population de Barbarie. Cette dernière affirmation est moins vraie dans la population commune (Marie-Etancelin *et al*., 2011). Selon ces mêmes auteurs, les taux de lipides et de protéines dans le foie présentent des héritabilités moyennes. Ces deux caractères sont très corrélés génétiquement avec le taux de fonte, le premier positivement (0,87 dans la population commun et 0,93 dans la population Barbarie) et le second négativement (supérieur à -0,85 dans les 2 populations). Le taux de collagène a lui une héritabilité très basse dans la population paternelle (0,02), et qui s'élève à 0,09 quand elle est estimée dans la population commune. Ce caractère est corrélé négativement avec le poids de foie et le taux de fonte (Marie-Etancelin *et al*., 2011).

Marie-Etancelin *et al*. (2011) ont étudié un grand nombre de caractères relatifs à la qualité de la viande du mulard gavé. Ils estiment des héritabilités moyennes (autour de 0,10 dans les deux populations) pour les caractères de coloration de la viande, sauf pour la coloration jaune dans la population de Barbarie qui est plus faible (0,06). Ces trois caractères de couleur présentent des corrélations très fortes avec le taux de lipides dans le magret, caractère héritable dans la population de commune (h^2=0,24) mais intermédiaire côté Barbarie (h^2=0,11) (Marie-Etancelin *et al*., 2011).

Peu d'articles relatifs à l'estimation des paramètres génétiques (héritabilités et corrélations génétiques races pures/croisement) ont été publiés, notamment pour les caractères de qualité du foie gras, le produit phare de la filière. A l'exception des caractères de croissance et de l'efficacité alimentaire, aucune étude ne démontre que l'amélioration des performances propres des populations parentales se traduit par une amélioration de performances du mulard.

Le canard mulard bénéficie pour un certain nombre de caractères d'aptitude au gavage, d'un effet d'hétérosis favorable. Ainsi, il est prouvé que le poids de foie gras et le gain de poids en gavage bénéficient d'un fort effet d'hétérosis mais aussi des effets génétiques additifs directs favorables de son père Barbarie et des effets maternels favorables de la cane Pékin (Larzul *et al.,* 2006). Ces résultats contribuent à expliquer la supériorité du mulard quant à son aptitude au gavage, comparé à ses deux souches parentales.

Les perspectives à envisager pour l'amélioration des performances, en termes d'aptitude au gavage et de qualités des produits notamment, pourraient être liées à la génétique moléculaire. En effet, les approches moléculaires pourraient permettre de sélectionner précocement les génotypes parentaux qui serviront à produire les canards mulards destinés au gavage. Cela permettra de contourner les problématiques posées par la sélection sur le canard mulard qui sont : 1-son statut d'hybride et sa stérilité associée, 2-le fait que les caractères d'intérêt chez le mulard ne soient pas tous directement mesurables dans les souches parentales 3-le coût élevé inhérent à la mesure de ces caractères sur le mulard.

2 Les QTLs identifiés pour les caractères d'intérêt chez le canard

2.1 Introduction

A ce jour, le seul dispositif de détection de QTL publié, relatif à des caractères d'intérêt zootechnique chez les canards, l'a été par Huang *et al.* (2007a et b), et reposait sur un croisement de canards Pekin destinés à la production de viande. Compte tenu du faible nombre de réarrangements inter-chromosomiques entre la poule et le canard (Fillon *et al.,* 2001), les QTLs détectés chez le poulet, pour des caractères d'intérêt chez le canard (tels que les caractères de croissance, de production de viande…) vont être étudiés ici comme base de comparaison pour nos propres résultats de détection de QTL. On présentera, donc, en plus des QTLs obtenus sur canards maigres (Huang *et al.*, 2007), les différents QTL mis en évidence chez la poule pour les différents caractères d'intérêt chez le canard.

A contrario chez la poule, il existe un grand nombre de protocoles dédiés à la détection de QTL. Toutefois, seule une partie des caractères d'intérêt chez la poule sont comparables aux caractères mesurés sur les canards à gaver. En effet, les conditions d'élevage et d'alimentation, entre autres, sont fortement différentes.

2.2 Les QTLs détectés pour les caractères d'intérêt

2.2.1 Introduction

Les premiers QTLs détectés, pour des caractères d'intérêt chez la poule datent de 1998 mais c'est à partir de 2002 que la recherche de ces derniers s'est intensifiée.

La plateforme National Animal Génome Research Program (www.animalgenome.org), qui recense les différents QTLs obtenus pour les différentes espèces domestiques, indiquait en Décembre 2011, 125 publications pour le poulet. Celles-ci ont permis, toujours selon les auteurs de la plateforme, de détecter 2 451 QTLs (dont certains sont redondants) sur un large éventail de 248 caractères. Il s'agit, entre autres, des caractères de croissance (Ambo *et al.*, 2009 ; Wahlberg *et al.*, 2009 ; Siwek *et al.*, 2004), des taux de glucose, cholestérol et triglycérides sanguins (Park *et al.,* 2006), des caractères de qualité de viande (Abasht *et al.*, 2006) ou des caractères liés au stress (Buitenhuis *et al.*, 2006 ; Jensen *et al.*, 2005).

Le dispositif de canards Pékin mis en place par Huang *et al.* (2007) est un protocole de détection de QTL de type « demi-frères » : 284 individus issus de 4 familles de pères provenant d'une lignée de Pékin sélectionnée pour son poids à 42 jours d'âge, et croisés avec 7 femelles sélectionnées pour leur production élevée d'œufs à 360 jours d'âge, ont été génotypés. Les caractères étudiés dans ce protocole sont essentiellement des caractères de croissance (poids corporel de l'éclosion, et de 1 à 7 semaines d'âge), de conformation (longueur du corps, du bréchet et du cou à 7 semaines d'âge), de pièces de carcasse (poids de la carcasse, de la tête, du cou, des pattes, des ailes, des cuisses, de la poitrine, du cœur, du foie, de la poitrine …) et de qualité de viande (perte par exsudation et à la cuisson du filet,

…). Dans cette étude, 14 QTLs hautement significatifs (seuil de 5% à l'échelle du génome) ont été détectés.

2.2.2 Les caractères de croissance

Un nombre élevé d'études de détection de QTL s'intéresse à des caractères de croissance. La facilité de la mesure de ces caractères explique, en partie, l'abondance des QTLs trouvés pour ces derniers. Les QTLs obtenus ne se concentrent pas sur un chromosome en particulier, mais sont répartis sur l'ensemble du génome. Parmi les auteurs ayant étudié ces caractères chez la poule, Ambo *et al.* (2009), Wahlberg *et al.* (2009) et Siwek *et al.* (2004) ont mis en place des protocoles de QTL avec des familles F2 de populations parentales sélectionnées sur des critères différents : Ambo *et al*. (2009) ont croisé une lignée de poulet de chair avec une lignée de poules pondeuses et ont mis en évidence l'implication des chromosomes GGA2, 3 et 4 dans la variabilité des poids corporels à 35 et 45 jours d'âge. Wahlberg *et al.* (2009) eux, ont repris avec une carte densifiée, une ancienne étude (Jacobsson *et al.,* 2005) croisant deux lignées de poulets issues d'une sélection divergente sur le poids corporel dans le jeune âge. Ils mettent en évidence 7 régions chromosomiques d'intérêt (répartis sur GGA1, 3, 4, 7 et 20) pour les caractères de croissance. Auparavant, Siwek *et al*. (2004) ont croisé deux lignées issues d'une sélection divergente sur la réponse immunitaire et à la résistance aux maladies. L'objectif de cette étude était de tester si l'amélioration de la réponse immunitaire diminuait les performances de croissance des animaux (deux caractères négativement corrélés selon ces auteurs) : ils concluent que les portions chromosomiques responsables de la variation de la réponse immunitaire sont différentes de celles contrôlant la variation de la croissance. Ils constatent également qu'il y a deux blocs de régions chromosomiques d'intérêt, l'une responsable des variations des poids corporels aux bas âges et l'autre contrôlant les variations des poids aux âges avancés.

Pour ce qui concerne la croissance du canard, Huang *et al*. (2007) ont mis en évidence 4 QTLs hautement significatifs (5% au niveau du génome) dont 2 localisés sur le groupe de liaison CAU1 (CAU pour **C**hina **A**gricultural **U**niversity, cela correspond à nos groupes de liaison avec le numéro de chromosome du poulet GGA1), groupe de liaison ancré sur le chromosome *Gallus Gallus* 1. Ils sont impliqués dans la variation du poids corporel à l'éclosion et du poids à six semaines d'âge. Ces résultats sont cohérents avec d'autres résultats obtenus chez le poulet sur GGA1, puisque plusieurs auteurs ont mis en évidence des zones chromosomiques impliquant le poids à six semaines d'âge (Atzmon *et al*., 2007 ; Carlborg *et al*., 2004 ; Liu *et al.,* 2008). Chez la poule, aucune étude ne rapporte de QTL pour le poids à l'éclosion sur ce chromosome mais Carlborg *et al*. (2003) ont détecté un QTL pour le poids corporel à un jour d'âge. Les deux autres QTLs sont localisés sur le groupe de liaison CAU2 (ancré sur GGA2). Il s'agit de QTLs contrôlant le poids corporel à l'éclosion et le poids à une semaine d'âge. A ce jour, des QTLs comparables n'ont pas été identifiés chez la poule.

2.2.3 Les caractères liés au métabolisme

2.2.3.1 Les caractères liés au stress

Le stress est mesuré par le taux de corticostérone dans le sang de l'animal avant et après soumission à un stress. Les différents stress auxquels les animaux sont généralement soumis sont le test d'open-field (Buitenhuis *et al*., 2006), le test de restriction alimentaire ou de changement d'alimentation (Jensen *et al*., 2005). Le test d'open-field consiste à placer l'animal seul dans un environnement nouveau et noter pendant un laps de temps ses activités, ses latences, les fréquences de ses cris et de ses sauts.

Chez la poule, 7 régions chromosomiques sont connues pour avoir un rôle dans la variabilité du niveau de corticostérone circulant dans le sang. Deux QTLs pour le niveau basal de corticostérone circulant dans le sang apparaissent situés sur les chromosomes 1 et 7, et 5 QTL pour le taux de corticostérone après soumission à un stress de l'animal sont identifiés sur les chromosomes 1, 5, 15, 18 et Z (Buitenhuis *et al.*, 2006 ; Jensen *et al.*, 2005).

2.2.3.2 Taux de cholestérol, de glucose et de triglycérides dans le sang

A ce jour, 19 QTLs ont été détectés pour les taux de cholestérol, de glucose et de triglycérides dans le sang chez le poulet : 11 pour le taux de glucose, 4 pour le taux de triglycérides et 4 pour le taux de cholestérol (Park *et al*., 2006 ; Zhou *et al*., 2007 ; Nadaf *et al*., 2009).

Zhou *et al*. (2007), dans un protocole de recherche de régions chromosomiques responsable de la variabilité de caractères liés au métabolisme, dans un croisement de Leghorn et de Fayoumi, ont localisé 5 QTLs à l'origine de la variation de taux de glucose dans le sang (chromosomes Z, GGA2, GGA4, GGA7 et GGA8). Aucune région ne se superpose sur les deux QTLs préalablement mis en évidence par Park *et al*. (2006) (chromosomes GGA20 et GGA 27). Le QTL localisé par ces derniers auteurs sur le GGA27 est le plus significatif de tous avec une significativité de 5% au niveau du génome. Enfin, les 4 autres QTLs (sur GGA3, GGA5, GGA6 et GGA13) impliqués dans la variation de taux de glucose dans le sang, chez la poule, sont localisés par Nadaf *et al*. (2009).

Les 4 QTLs détectés pour le taux de cholestérol dans le sang sont localisés sur les chromosomes 1, 3, 9 et 20 (Park *et al.*, 2006 ; D'andré Hirwa *et al.*, 2010). Park *et al.,* (2006) ont réalisé deux croisements réciproques de deux lignées de poulet résultant d'une sélection divergentes sur la croissance. Il constate que le type de croisement réciproque a eu un effet très significatif sur les concentrations plasmatiques de glucose et de cholestérol.

Enfin, Park *et al.* (2006) sont les premiers à détecter des QTLs pour le taux de triglycérides dans le sang : ils en cartographient une zone d'intérêt sur les chromosomes GGA2. Campos *et al*. (2009), dans un croisement entre une lignée de poulet à chair et une lignée de poules pondeuses, ont mis en évidence trois nouvelles régions chromosomiques responsables de la concentration des triglycérides sanguins sur GGA5, GGA23 et GGA27.

2.2.4 Les caractères de qualité des produits

Les premières recherches de QTLs pour les caractères de qualité de viande de poulet comme la coloration de la viande, son pH, son gras intramusculaire et le taux de perte à la cuisson ont été publiées en 2005 (Rabi *et al.*, 2005). A ce jour 30 régions d'intérêt pour des caractères de qualité de viande ont été détectés. Deux régions candidates pour des gènes qui contrôlent la variation de la teneur en protéines brutes du muscle de la poule sont mises en évidence sur les chromosomes GGA 9 (à 110 cM) et GGA 12 (90 cM) en 2009 (Nie *et al.*, 2009). Cette étude est menée sur 17 populations (poulet et canard) dans le but d'isoler des régions pouvant être à l'origine de la variation du taux de la Ghréline, hormone qui stimule l'appétit et qui contribue donc à l'obésité chez l'Homme.

Wright *et al.* (2006), dans un croisement de Junglefowl rouge et de Leghorn blanches et Nadaf *et al*. (2007), dans un croisement de deux lignées de poulets issus d'une sélection divergente pour la vitesse de croissance, ont effectué des recherches de QTLs contrôlant le pH de la viande. Ces derniers (Wright *et al.*, 2006 + Nadaf *et al*., 2007) ont mis en évidence, au total, 10 QTLs (localisés sur les chromosomes 1, 2, 4, 7, 12 et 20). En outre, Nadaf *et al.* (2007) observent dans leur étude que la sélection divergente pour la vitesse de croissance a entrainé la ségrégation de plusieurs caractères de qualité de la viande. Ainsi, ils mettent en évidence 2 QTLs pour la coloration rouge et jaune de la viande sur le même chromosome (GGA11) et quasiment à la même position. Cette région du génome a fait par la suite, l'objet d'une cartographie fine. Les résultats de Nadaf *et al*. (2007) combiné à une analyse eQTL d'expression de gènes ont permis d'identifier un gène causal sous-jacent aux 2 QTLs hautement significatifs (significativités de 1% au niveau du génome) de coloration de la viande localisés sur le chromosome GGA11 (Le Bihan-Duval *et al*., 2011). En plus de cette région du génome, 5 autres QTLs (sur GGA1, GGA2, GGA5, GGA8 et GGA10) ont été identifiées pour des caractères de coloration de viande (Van Kaam *et al*., 1999 ; Rabie *et al*., 2005 ; Wright *et al*., 2006).

Concernant la qualité de la viande de canard maigre, un QTL pour la perte à la cuisson du filet sur le CAU6 a été détecté par Huang *et al*. (2007). Par ailleurs, 8 QTLs (significatifs à 5% au niveau du génome) liés à des caractères de carcasse et de qualité de la viande ont été mis en évidence. Sept de ces QTLs sont localisés sur le groupe de liaison CAU4 (ancré sur GGA4) : il s'agit de QTLs contrôlant le poids du gras abdominal, du foie, du cou et de la cuisse ainsi que la perte en eau par exsudation du filet. Plusieurs QTLs contrôlant le poids de gras abdominal ont aussi été détectés sur le chromosome *Gallus Gallus* 4 chez la poule (Jennen *et al*., 2004 ; Zhou *et al.,* 2006 ; Atzmon *et al.,* 2008).

Dans leur étude, Huang *et al*. (2007) ont détecté 2 QTLs significatifs à 5% au niveau du génome, pour les caractères de conformation, l'un contrôlant la circonférence de la cuisse (sur CAU12) et l'autre la longueur du bréchet (sur CAU6). A cela s'ajoute 2 QTLs responsables de la variabilité du poids de foie, l'un sur le groupe de liaison CAU6, et l'autre, avec une significativité moindre (significatif à 5% au niveau du chromosome) sur le groupe de liaison CAU7 (ancré sur GGA7). Chez la poule, 2 QTLs impliqués dans la variabilité du poids du foie sont identifiés sur le chromosome GGA2 par Gao *et al*. (2009) et sur GGA1 par Wright *et al*. (2006). Ces derniers auteurs ont par ailleurs identifié un QTL hautement significatif (1% au niveau du génome) impliqué dans la variation de la pigmentation du foie.

Chapitre IV. Détection de QTL dans le protocole de canard gras

1 Introduction

Les résultats du premier dispositif expérimental mis en place pour la recherche de QTL impliqués dans la variabilité des caractères d'intérêt économique ou biologique chez le canard gras, sont présentés dans ce chapitre. Ce dispositif consiste en 7 familles informatives de type Backcross, obtenues en croisant 2 souches expérimentales, la souche INRA444 (canard Kaiya, résultat d'un croisement original Tsaiya x Pekin) et la souche synthétique de Pékin lourd (INRA37), elle-même issue d'un croisement original de 3 souches lourdes européennes.

Le choix de ce croisement est motivé par les résultats de Rouvier *et al*. (1994) qui ont montré que le poids corporel et le poids de foie des mulards issus des mères Pékin et Tsaiya et ceux issus de deux croisements réciproques sont significativement différents. Ces auteurs concluent que l'utilisation des mères résultant d'un croisement entre Tsaiya et Pékin pour la production des mulards permettent d'obtenir un foie gras de poids raisonnable et de meilleure qualité.

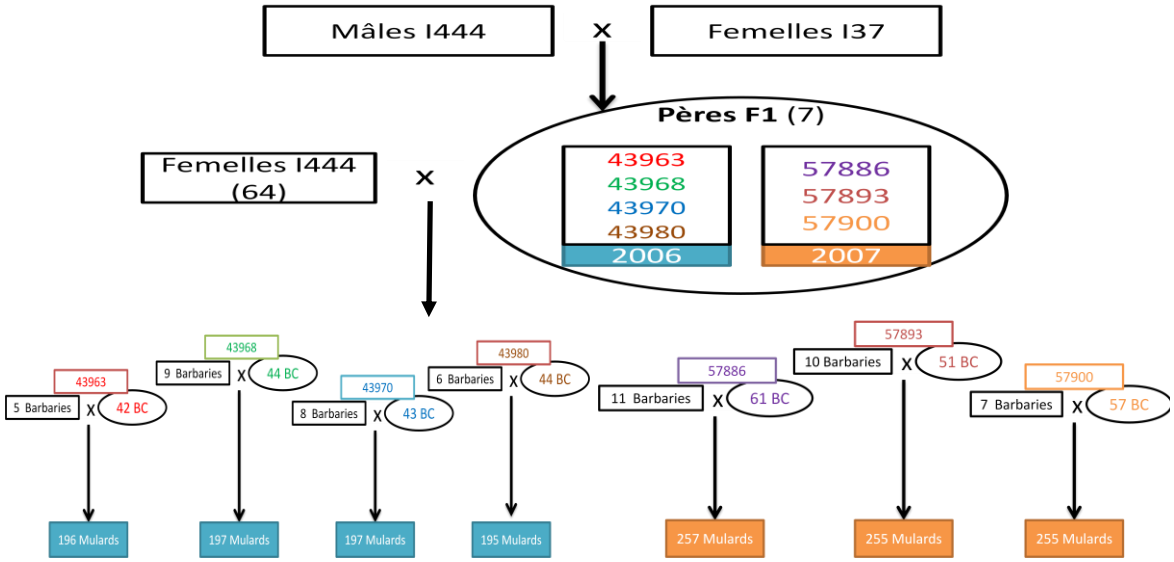
L'utilisation d'un protocole backcross est motivée, en plus d'un certain nombre de contraintes techniques qui ont exclu la mise en place d'un protocole F2, par un effectif nécessaire moins important lorsque les allèles au QTL sont l'un dominant, l'autre récessif. En effet, dans ce cas, le protocole backcross est plus puissant que le protocole F2 pour la détection de QTL s'il est réalisé sur le parent homozygote récessif (Darvasi, 1998), puisque le nombre de descendants nécessaire pour estimer la dominance est alors deux fois moins important que dans un protocole F2. Par contre, lorsque les allèles au QTL sont co-dominants, l'estimation des effets additifs dans un protocole F2 nécessite 30% moins de descendants par rapport à un protocole backcross.

Une large gamme de caractères originaux a été mesurée sur les 1600 canards mulards produits en croisant les femelles Backcross avec des mâles Barbaries. En effet, la plupart de ces caractères (notamment ceux relatifs au gavage) ne sont pas mesurables sur les canes communes Backcross : les mesures sont effectuées sur les mulards plutôt que sur les canes Backcross, ce qui n'est pas classiquement réalisé dans ce type de protocole.

2 Le dispositif détaillé

Le dispositif a été procréé sur 2 années consécutives, chaque année produisant spécifiquement des familles. Ainsi 4 familles de 43 canes BC ont été produites la 1ère année, et 3 familles de 57 canes la seconde année. Annuellement, les mulards issus de ces canes BC étaient produits en 2 lots d'éclosion successifs, chacun d'eux redivisé en 2 lots de gavage.

Figure 4.1 : Dispositif de détection de QTL



Dans notre schéma expérimental (figure 4.1) de production des mulards, différents effets fixés sont identifiés:

- les mulards sont issus de 2 lots d'éclosion avec comme objectif de répartir, pour chaque famille de père F1, la moitié des mulards est issue d'une première éclosion et l'autre moitié d'une deuxième éclosion.
- les mulards sont ensuite élevés (de 0 à 12 semaines) chaque année en 2 fois 8 lots d'élevage d'une cinquantaine de canards.
- Trois gaveurs ont réalisés le gavage des 1600 mulards, 2 gaveurs (G1A et G1B) la première année et 2 gaveurs (G2A et G2C) la deuxième année ; un gaveur (GA) étant commun aux 2 années. Les descendants d'une famille de père sont gavés pour moitié par un gaveur et pour moitié par un autre.

3 Les mesures des caractères

Les caractères enregistrés sur chacun des 1600 mulards peuvent être classés en 4 grands groupes. Il s'agit de caractères de croissance, de caractères liés au métabolisme pendant le gavage, de caractères d'aptitude au gavage et de caractères relatifs aux qualités des produits (foie gras et magret). De plus, quelques aspects comportementaux ont été enregistrés : en élevage, le taux de corticostérone basal et après un stress (pendaison par les pattes, tête en bas) a été dosé et en gavage, le comportement des mulards (facilité de gavage et d'attrapage) a été apprécié.

Pour l'ensemble des caractères, l'analyse de variance qui prend en compte les effets fixes précédemment identifiés (lot d'éclosion, lot d'élevage, gaveur) et la famille de père F1, met en évidence un effet très significatif de la famille du père F1. Les valeurs moyennes des caractères étudiés ainsi que les différences par famille sont présentées ci-après.

Caractères de croissance

Les caractères de croissance (tableau 4.1 ; figure 4.2) ont été mesurés par pesée des animaux à jeun, à différents âges (à 12, 28, 42 et 70 jours d'âge). La balance utilisée est une balance Balea (AGPA - Automate graphique de pesée animale), qui permet l'acquisition d'un poids fiabilisé : en effet, le poids retenu est une moyenne de N poids d'un même animal obtenu en quelques secondes, et dont l'écart-type est inférieur à un seuil S (N et S étant paramétrables). Si un animal bouge ou est dérangé par un autre individu, la prise de mesure ne sera pas faite : la pesée sera poursuivie jusqu'à ce que les conditions de précision soient remplies.

Pour ces différents âges, les poids moyens des animaux sont conformes aux poids de mulards de Kaiya habituellement rencontrés (tableau 4.1) : néanmoins la variabilité de ces poids (coefficient de variation de l'ordre de 20% à 12 jours) est plutôt élevée, surtout à 12 jours d'âge.

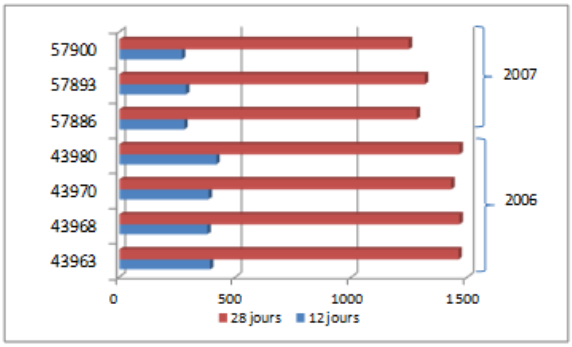
Tableau 4.1 : Descriptions des caractères de croissances.

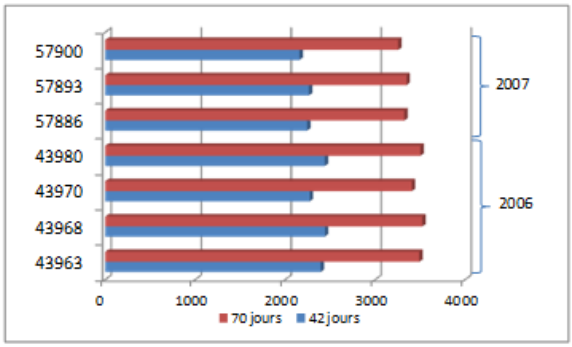
	N	Moyenne	Ecart-type
BW12[1] (g)	1551	337	69
BW28[1] (g)	1549	1373	150
BW42[1] (g)	1545	2306	206
BW70[1] (g)	1508	3396	288

[1] BW12, BW28, BW42, BW70 : poids corporel à 12, 28, 42 et 70 jours.

Au-delà de la variation entre père, un effet « année » (figure 4.2) apparait pour les poids corporels, principalement dans le jeune âge (à 12 et 28 jours d'âge). En effet les mulards nés en 2006 (A2006f1 à A2006f4) ont en moyenne un poids plus élevé que ceux nés en 2007 (A2007f5 à A2007f7). Par contre, les gains de poids moyens sont sensiblement identiques quelles que soient les familles de pères F1.

Figure 4.2 : Poids moyens par famille de père F1
(à 12 et 28 jours à gauche) (à 42 et 70 jours à droite)





Les caractères relatifs au métabolisme

Au début (deuxième repas), au milieu (dixième repas) et à la fin (vingtième repas) de la période du gavage, des échantillons de cinq millilitres de sang ont été collectés pour mesurer le taux de métabolites plasmatiques afin apprécier l'évolution de la stéatose hépatique durant la phase de gavage. Ces collectes, réalisées trois heures après les repas ont permis de déterminer la concentration en glucose, triglycérides et cholestérol dans le plasma sanguin (tableau 4.2 ; figure 4.3), en utilisant des kits spécifiques (tubes EDTA ou acide éthylène diamine tétraacétique).

Deux tubes de sang par canard sont prélevés, l'un des 2 tubes étant centrifugé à 3000 tours/min puis le plasma est aliquoté en 3 petits tubes. L'autre tube est conservé en l'état comme sécurité.

Quel que soit le caractère étudié (triglycérides, cholestérol, glucose), on observe une augmentation des valeurs de dosages sanguins avec l'avancée du gavage (tableau 4.2). De plus, la variabilité des mesures s'accroit aussi de façon importante avec le stade de gavage.

Tableau 4.2 : Descriptions des caractères liés au métabolisme.

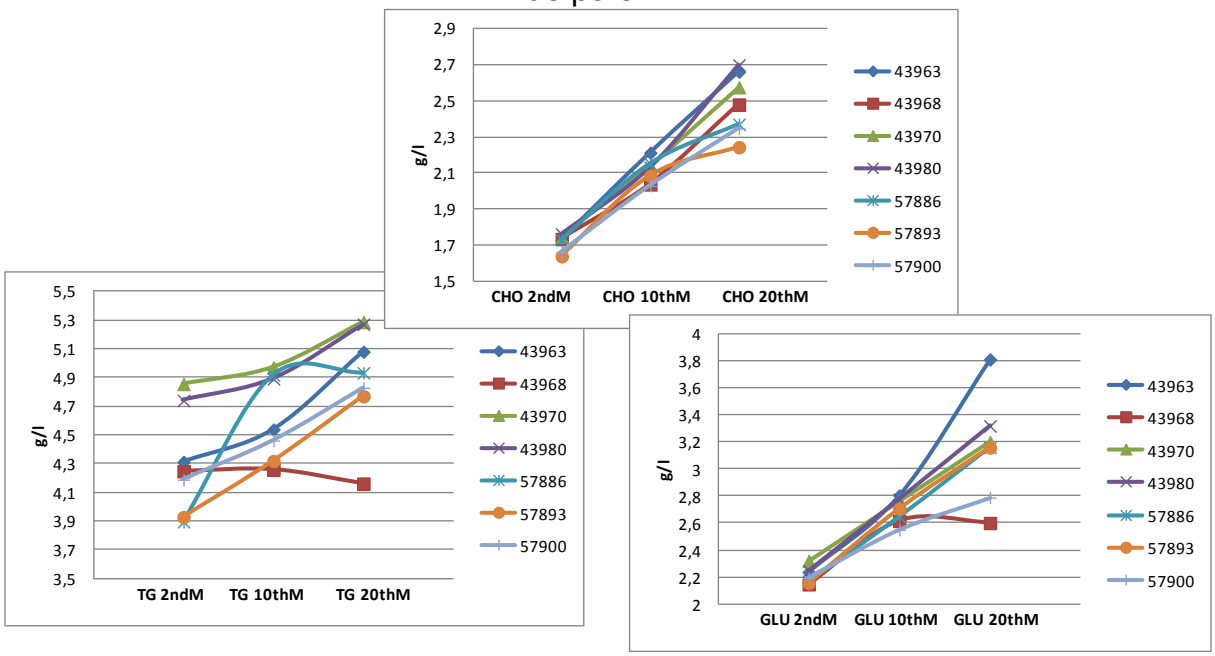
	N	Moyenne	Ecart-type
TG $2^{nd}M^1$ (g/l)	1499	4,27	1,00
TG $10^{th}M^1$ (g/l)	1498	4,61	1,08
TG $20^{th}M^1$ (g/l)	1443	4,89	1,52
CHO $2^{nd}M^1$ (g/l)	1501	1,71	0,25
CHO $10^{th}M^1$ (g/l)	1499	2,11	0,32
CHO $20^{th}M^1$ (g/l)	1433	2,46	0,47
GLU $2^{nd}M^1$ (g/l)	1500	2,20	0,27
GLU $10^{th}M^1$ (g/l)	1498	2,69	0,56
GLU $20^{th}M^1$ (g/l)	1451	3,13	1,08

[1] TG, CHO et GLU : taux de Triglycéride, de Cholestérol et de Glucose dans le sang au 2^{nd}, $10^{ème}$ et $20^{ème}$ repas.

Les taux de cholestérol sont assez comparables pour les différentes familles de père (figure 4.3), ainsi que leur évolution entre les 3 stades de mesures. Pour les taux de triglycérides, on notera toutefois 2 familles atypiques : la famille du père 57886 dont le taux de triglycérides était très élevé en milieu de gavage mais est resté stable ensuite, et la famille du père 43968 qui est la seule dont le taux du triglycéride a légèrement décru avec l'avancée du gavage. Pour les taux de glucose, la variabilité

entre les familles de pères apparait particulièrement en fin de gavage, avec une disparité dans les évolutions de concentration selon les familles : en effet, le taux de glucose entre le 10ème et le 20ème repas augmente très fortement pour le père 43963 alors qu'il reste constant pour le père 43968.

Figure 4.3 : Taux de Triglycéride, de Cholestérol et de Glucose sanguins par famille de père F1



Les caractères de taux de corticostérone avant et après stress

Pour mesurer la réponse au stress des animaux, un test de pendaison est organisé. L'animal est suspendu pendant 10 minutes par les pattes, la tête vers le bas. Des prélèvements sanguins sont effectués avant et après le test pour estimer les niveaux de corticostérone et évaluer la réponse de l'axe hypothalamo-hypophyso-surrénalien (tableau 4.3 ; figure 4.4).

Le niveau de corticostérone s'élève très fortement avec le stress, puisqu'il passe en moyenne de 10 ng/ml à 55 ng/ml (tableau 4.3). Néanmoins, la variabilité des mesures était plus importante avant le stress (coefficient de variation de 107%) qu'après le stress (coefficient de variation de 77%).

Tableau 4.3 : Descriptions des caractères liés au stress

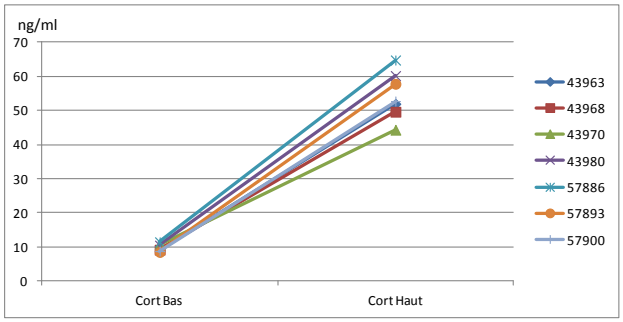
	N	Moyenne	Ecart-type
Cort Bas[1] (ng/ml)	1502	9,66	10,32
Cort Haut[1] (ng/ml)	1502	54,97	42,47
Delta Cort[1] (ng/ml)	1502	45,31	39,37

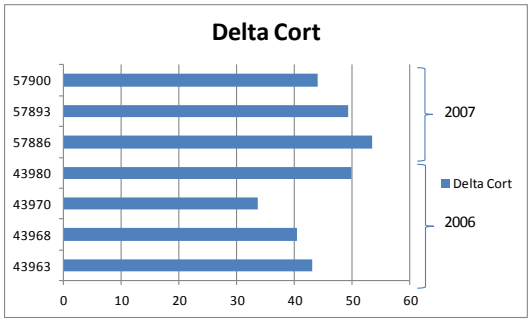
[1] CortL, CortH, DeltaC : Taux de Corticostérone basale, après stress et delta de corticostérone.

L'évolution de corticostérone avec le stress ne semble pas être affectée par un effet année (figure 4.4). Néanmoins, on peut remarquer que la famille du père 43970 présente une moindre augmentation de son niveau de corticostérone comparé

aux autres familles. Hormis ce père, le classement des familles selon le niveau de corticostérone avant ou après stress reste le même.

Figure 4.4 : Taux de corticostérones par famille de père F1





Les caractères d'aptitude au gavage

L'aptitude au gavage des canards a été mesurée par pesée des animaux en début et fin de la période de gavage, et des différentes pièces de la carcasse (le foie gras, les cuisses, le gras abdominal ou le magret qui correspond aux muscles pectoraux, sa peau et la graisse sous-cutanée qui le couvre). Aussi le poids de la carcasse ressuée (poids plumé, saigné, sans pattes et bout d'ailes) a été enregistré. La consommation journalière de maïs a également été enregistrée pour chaque animal durant le gavage : la dose-repas étant identique pour tous les animaux, sont enregistrés les repas sautés ou les demi-doses données (tableau 4.4 ; figure 4.5).

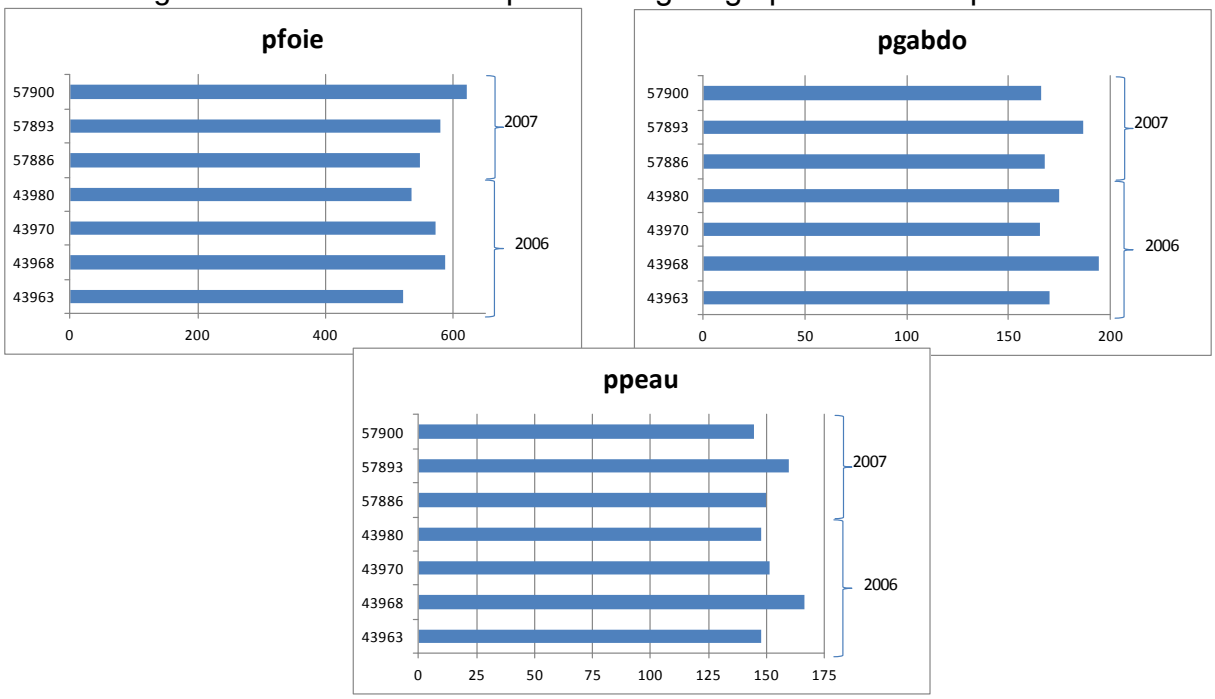
Tableau 4.4 : Description des caractères d'aptitude au gavage.

	N	Moyenne	Std
DFI[1] (g/d)	1498	1325	74
BWbeg[1] (g)	1501	3830	300
BWend[1] (g)	1498	5805	378
OWG[1] (g/d)	1498	1974	247
CW[1] (g)	1474	4902	329
pfoie[1] (g)	1492	567,9	115,3
pgabdo[1] (g)	1476	175,0	28,8
pmMW[1] (g)	1476	256,5	23,6
ppeau[1] (g)	1476	152,4	20,3
TSW[1] (g)	1476	481,9	45,1

[1] DFI, BWbeg, BWend, OWG, CW, FLW, AFW, pmMW, pmSFW, TSW : consommation moyenne journalière, poids du début et fin gavage, gain de poids durant le gavage, poids de la carcasse, du foie, du gras abdominal, du magret, de la peau et de cuisse.

Si aucun effet « année de mesure » n'apparait pour les caractères d'aptitude au gavage, l'effet famille de pères F1 est particulièrement significatif pour 3 caractères : le poids de foie gras, le gras intra-abdominal et le poids de la peau du magret (figure 4.5). Parmi ces caractères en lien direct avec l'engraissement des animaux, les familles de père semblent se comporter de la même façon pour le gras abdominal et pour le poids de peau du magret, mais pas pour le poids de foie gras : ainsi les familles 57893 et 43968 ont un engraissement sous-cutané ou abdominal très développé alors que la famille avec le poids de foie le plus élevé est la 57900.

Figure 4.5 : Caractères d'aptitude au gavage par famille de père F1



En plus des caractères décrits ci-haut, des observations comportementales en gavage ont été effectuées par les gaveurs. La facilité d'attrapage et de gavage sont notées pour chaque canard mulard, la note est attribuée par le gaveur sur une échelle de 1 à 4.

Les caractères de qualité de produits

La qualité du foie gras (tableau 4.5 ; figure 4.6) et du magret (tableau 4.6 ; figure 4.7) a été appréhendée par des mesures de compositions et des mesures de propriétés technologiques.

Ainsi le taux de fonte du foie (pourcentage de perte de graisses après stérilisation de 60g d'un foie pendant 50 min à 105° C) et les pertes à la cuisson du muscle (15 min dans une eau bouillante de 85° C) sont mesurés. Pour le magret le pH à 20 minutes et 24 heures *post-mortem* sont mesurés. Le taux de lipides et de protéines ont été estimés par spectrophotomètre avec la technique *near infra red* (FOSS NIRSystem) sur des échantillons du foie gras. Le taux de lipides du magret est aussi estimé par la même technique. Pour le foie, le taux de collagène a été déterminé selon Woessner (1961). Les indices L*, a* et b* de coloration pour le foie gras et le magret sont mesurés par un chromamètre (CR 300 Minolta). La perte par exsudat sous vide des muscles enveloppés sous un film plastique a été mesurée après 6 jours de stockage à 6° C. Enfin, la tendreté de la viande crue a été mesurée avec le test de Warner-Bratzler : ce test mesure, à l'aide d'un appareil spécifique, la force et l'énergie nécessaire pour couper un morceau de viande.

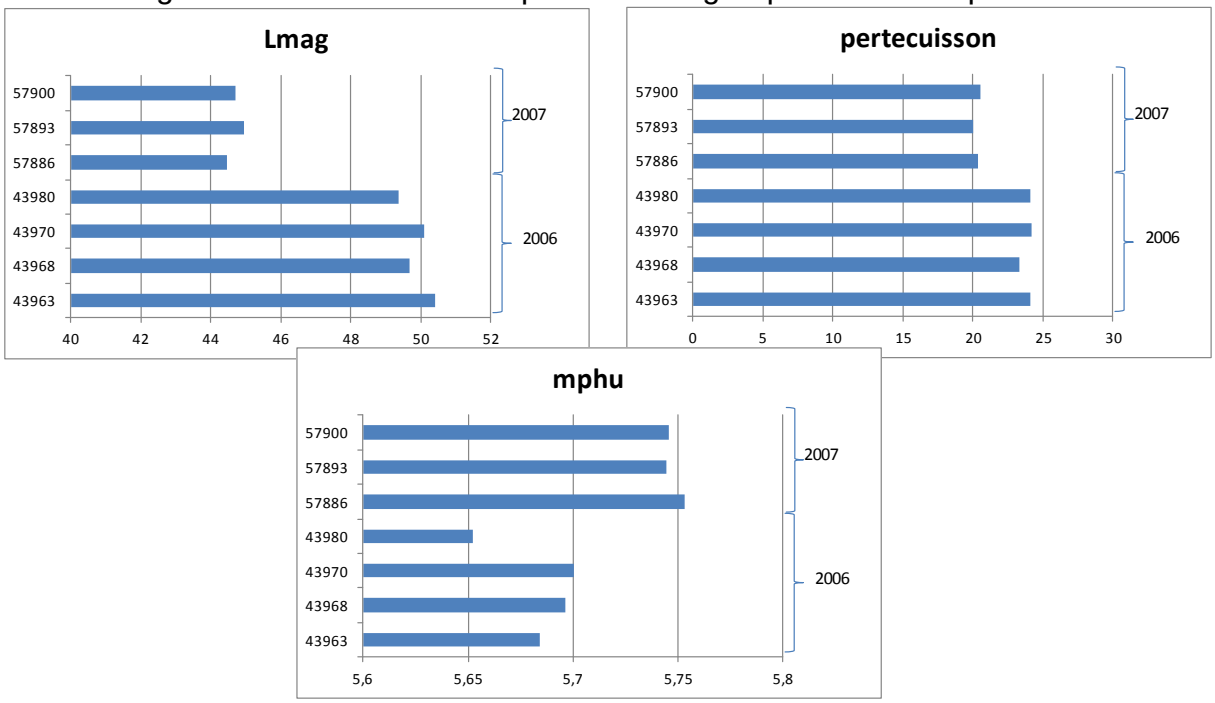
Tableau 4.5 : Description des caractères de qualité du magret.

	N	Moyenne	Std
Lmag[1]	1476	47,3	3,4
amag[1]	1476	20,4	2,5
bmag[1]	1475	7,62	1,46
MpH20[1]	1476	6,01	0,18
MpHu[1]	1476	5,72	0,14
MvacL (%)[1]	1462	1,58	0,84
pertecuisson (%)[1]	1437	22,10	3,84
Fmax[1]	1443	42,49	7,65
Energy (mJ) [1]	1442	149,61	40,16
MlipC (%)[1]	1476	4,93	0,70

[1]ML*, Ma*, Mb*, MpH20, MpHu, MvacL, MCookL, Fmax, Energy, MlipC : luminosité, indice rouge et jaune du magret, le pH du magret 20 post mortem et le pH ultime, Perte d'eau sous vide et à la cuisson du magret, force et énergie nécessaire pour couper un morceau de magret et taux de lipide du magret.

La plupart des caractères de qualité des produits présentent un effet année très marqué : c'est particulièrement le cas pour les 3 indices de couleur des magrets, les caractères de rhéologie et le pH ultime du magret (figure 4.6). Hormis pour le pH 20 minutes post mortem, on observe globalement une grande homogénéité de 3 familles de pères de 2007 et à contrario une plus grande hétérogénéité pour les familles de pères de 2006.

Figure 4.6 : Caractères de qualité du magret par famille de père F1



Un certain nombre de caractères ont aussi permis d'apprécier la qualité du foie gras tel que la texture au doigt sur foie refroidi. Aussi, des caractères comme l'extension des taches dues aux viscères sur le foie, la pigmentation du lobe du foie ou les zones de nécros sont estimés.

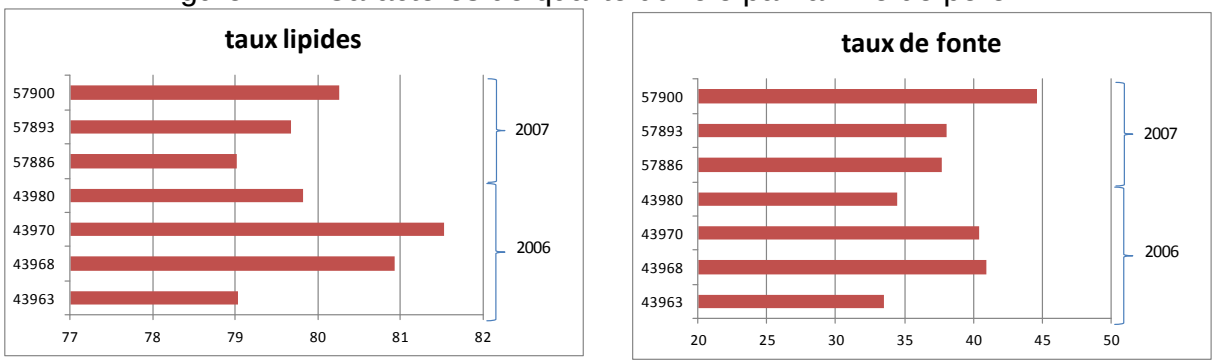
Tableau 4.6 : Description des caractères de qualité du foie.

	N	Moyenne	Std
Taux de fonte[1] (%)	1472	38,7	12,4
Lfoie[1]	1476	72,4	2,36
afoie[1]	1476	9,17	1,77
bfoie[1]	1476	31,2	2,9
Taux de collagène[1] (mg/g)	1436	1,27	0,27
Taux de protéine[1] (%)	1476	7,7	1,1
Taux de lipides[1] (%)	1476	52,4	4,3

[1]MR, LL*, La*, Lb*, LColC, LProtC, LLipC : taux de fonte du foie, luminosité, indice rouge et jaune du foie, taux de collagène, de protéine, de lipide du foie.

Le taux de fonte est le caractère de qualité du foie gras le plus variable avec une coefficient de variation de l'ordre de 32% (tableau 4.6). A l'inverse, la luminance du foie et le taux de lipides font partie des caractères les moins variables (CV inférieurs à 10%).

Figure 4.7 : Caractères de qualité du foie par famille de père F1



On observe des différences entre familles de pères F1 particulièrement marquées pour le taux de fonte et le taux de lipides des foies (figure 4.7). Ainsi, le taux de fonte varie de 33,5% pour la famille du père 43963 à 44,6% pour la famille du père 57900. Néanmoins les différences de taux de fonte d'une famille à l'autre ne semblent pas être directement expliquées par les différences de taux de lipides entre familles.

4 Résultats et discussion

4.1 Article 2

Detection of single and pleiotropic QTL controlling metabolism, meat and liver quality traits of the overfed inter-specific hybrid mule duck

M. Kileh-wais*, J.M. Elsen*, A. Vignal†, K. Fevès†, F. Vignoles†, X. Fernandez‡, H. Manse‡, S. Davail§, J.M. André§, D. Bastianelli#, L. Bonnal#, O. Filangi╟, E. Baéza¶, D. Guéméné¶,%, C. Genêt¡, M.D. Bernadet¤, F. Dubos¤, C. Marie-Etancelin*

* Institut National de la Recherche Agronomique, SAGA Station d'Amélioration Génétique des Animaux, UR631, 31 326 Castanet Tolosan, France
† Institut National de la Recherche Agronomique, LGC Laboratoire de Génétique Cellulaire, UR444, 31 326 Castanet Tolosan, France
‡ Université de Toulouse, Institut National Polytechnique de Toulouse, Ecole Nationale Vétérinaire de Toulouse, Ecole Nationale Supérieure Agronomique de Toulouse, Institut National de la Recherche Agronomique, TANDEM Tissu Animaux, Nutrition, Digestion, Ecosystème et Métabolisme, UMR1289, 31 326 Castanet Tolosan, France
§ Université de Pau et des Pays de l'Adour, IPREM-EMM Institut Pluridisciplinaire de Recherche sur l'Environnement et les Matériaux, Equipe Environnement et Microbiologie, UMR5254, 40 004 Mont de Marsan Cedex, France
# Centre de Coopération Internationale en Recherche Agronomique pour le Développement, SELMET Systèmes d'Elevage Méditerranéens et Tropicaux, UR112, 34 398 Montpellier Cedex 5, France
╟ Institut National de la Recherche Agronomique, PEGASE Physiologie, Environnement et Génétique pour l'Animal et les Systèmes d'Elevage, UMR1348, Agrocampus Ouest, 35000 Rennes, France
¶ Institut National de la Recherche Agronomique, URA Unité de Recherche Avicole, UR83, 37 380 Nouzilly, France
% Syndicat des Sélectionneurs Avicoles et Aquacoles Français SYSAAF, 37 380 Nouzilly, France
¡ Institut National de la Recherche Agronomique, GABI Génétique Animale et Biologie Intégrative, UR1313, 78352 Jouy-en-Josas, France
¤ Institut National de la Recherche Agronomique, UEPFG Unité Expérimentale des Palmipèdes à Foie Gras, UE89, 40 280 Benquet, France

Soumis le 24 Avril 2012 à *Journal of Animal Science*

QTL in overfed hybrid duck

Detection of single and pleiotropic QTL controlling metabolism, meat and liver quality traits of the overfed inter-specific hybrid mule duck[1]

M. Kileh-wais*, J.M. Elsen*, A. Vignal†, K. Feves†, F. Vignoles†, X. Fernandez‡, H. Manse‡, S. Davail§, J.M. André§, D. Bastianelli#, L. Bonnal#, O. Filangi‖, E. Baéza¶, D. Guéméné¶,%, C. Genêt¡, M.D. Bernadet¤, F. Dubos¤, C. Marie-Etancelin*[2]

* Institut National de la Recherche Agronomique, SAGA Station d'Amélioration Génétique des Animaux, UR631, 31 326 Castanet Tolosan, France
† Institut National de la Recherche Agronomique, LGC Laboratoire de Génétique Cellulaire, UR444, 31 326 Castanet Tolosan, France
‡ Université de Toulouse, Institut National Polytechnique de Toulouse, Ecole Nationale Vétérinaire de Toulouse, Ecole Nationale Supérieure Agronomique de Toulouse, Institut National de la Recherche Agronomique, TANDEM Tissu Animaux, Nutrition, Digestion, Ecosystème et Métabolisme, UMR1289, 31 326 Castanet Tolosan, France
§ Université de Pau et des Pays de l'Adour, IPREM-EMM Institut Pluridisciplinaire de Recherche sur l'Environnement et les Matériaux, Equipe Environnement et Microbiologie, UMR5254, 40 004 Mont de Marsan Cedex, France
# Centre de Coopération Internationale en Recherche Agronomique pour le Développement, SELMET Systèmes d'Elevage Méditerranéens et Tropicaux, UR112, 34 398 Montpellier Cedex 5, France
‖ Institut National de la Recherche Agronomique, PEGASE Physiologie, Environnement et Génétique pour l'Animal et les Systèmes d'Elevage, UMR1348, Agrocampus Ouest, 35000 Rennes, France
¶ Institut National de la Recherche Agronomique, URA Unité de Recherche Avicole, UR83, 37 380 Nouzilly, France
% Syndicat des Sélectionneurs Avicoles et Aquacoles Français SYSAAF, 37 380 Nouzilly, France

[1] This work was supported by the French National Research Agency GENANIMAL, by the Scientific Group of avian genome analysis AGENAVI and the Midi-Pyrénées and Aquitaine regional councils. M. Kileh Wais was supported by a grant from INRA Animal Genetics Division and the University of Djibouti. The technicians of UEPFG Experimental Unit deserve special thanks for the ducks breeding, the phenotypic traits recording and blood sampling. We thank G. Cardinet for her help for the development of the microsatellite markers.

[2] Corresponding author: Christel.Marie-Etancelin@toulouse.inra.fr

¡ Institut National de la Recherche Agronomique, GABI Génétique Animale et Biologie Intégrative, UR1313, 78352 Jouy-en-Josas, France
¤ Institut National de la Recherche Agronomique, UEPFG Unité Expérimentale des Palmipèdes à Foie Gras, UE89, 40 280 Benquet, France

ABSTRACT
The mule duck, an interspecific hybrid obtained by crossing Common duck (*Anas Platyrhynchos*) females with Muscovy (*Cairina moschata*) drakes, is widely used for fatty liver production. The purpose of the present study is to detect and map single and pleitropic QTL segregating in the Common duck species, having an influence on the expression of traits in their overfed mule ducks offspring. To this end, a Common duck back-cross (BC) design was generated by crossing Kaiya duck and heavy Pekin duck experimental lines, which differ notably in the bodyweight and overfeeding ability of their mule progeny. The BC females were mated to Muscovy drakes and on average, 4 male mule ducks per BC female - 1600 in total - were hatched and measured for growth, metabolism during growth and overfeeding period, overfeeding ability and breast meat and fatty liver qualities. The phenotypic value of the BC females was estimated for each trait by assigning the mean value of its offspring's phenotype, taking the variance, depending on the number of sons measured per BC and the heritability of the trait considered, into account. The genetic map used for QTL detection has 91 microsatellite markers aggregated into 16 linkage groups (LG), covering a total of 778 cM.
Twenty-two QTL were found significant at the 1% chromosome-wise threshold level, using the single trait detection option of the QTLMap software. Most QTL were detected for breast meat and fatty liver qualities: QTL for meat pH 20 minutes *post mortem* were mapped on LG4 (at 1% genome-wide level) and QTL for meat lipid content and cooking losses were found both on LG2a. For the fatty liver weight and composition in protein and lipid, QTL were mainly detected on LG2c and LG9 and multiple traits analyses highlighted pleiotropic effects of QTL in these chromosome regions. Apart for the strong QTL on chromosome Z for plasma triglyceride content at the end of overfeeding period detected in single trait analysis, all metabolic traits QTL were revealed with the multi-traits approach: QTL on LG14 and LG21 affected the plasma cholesterol and triglyceride contents whereas QTL on LG2a seemed to impact glycaemia and the basal plasma corticosterone content. A higher density genetic map will be needed for fine mapping of the QTL.

Key Words: hybrid mule duck, liver quality, meat quality, metabolic trait, pleiotropy, QTL.

INTRODUCTION

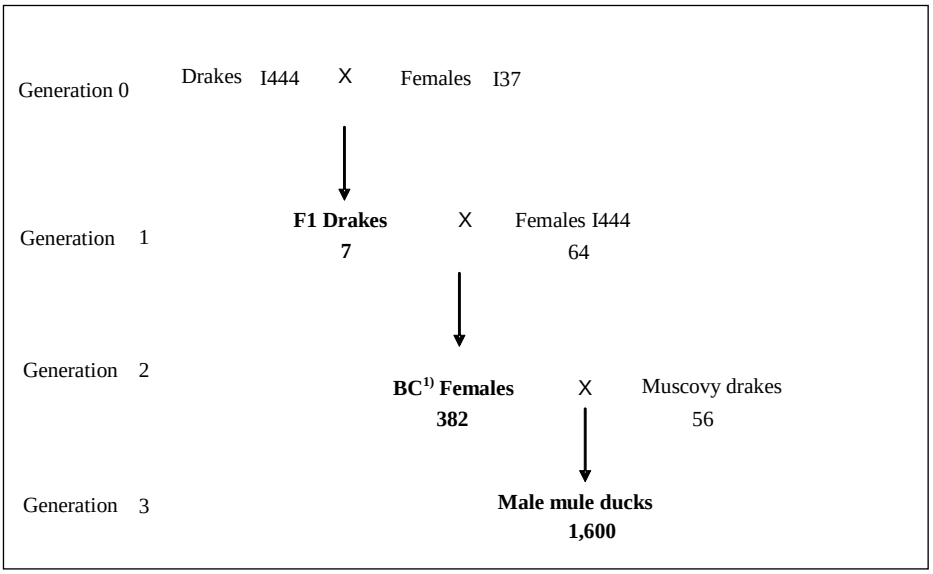
Ninety-five percent of the fatty liver production in France comes from mule ducks, the infertile inter-specific hybrid progeny of Common duck (*Anas Platyrhynchos*) females and Muscovy (*Cairina moschata*) drakes. Genetic improvement of mule duck performances is thus created by selecting both parental species for production traits which are recorded in their progenies. The parental lines are usually specialized in specific traits: Muscovy ducks are mainly selected for body weight and feed efficiency whereas genetic improvement in Common ducks is focused on reproductive traits (Marie-Etancelin *et al*., 2008). Therefore, the evaluation of the breeding values of

purebred reproducer duck candidates has to be done by progeny testing, lengthening the generation interval. Moreover, traits related to the quality of products need expensive measurements, which are destructive for the end products. In this context, it is particularly interesting to identify major genes or QTL in the parental populations having an impact on traits in the mule duck. To investigate QTL segregating in Common duck, an experimental back-cross (BC) using 2 experimental lines was designed. The phenotypic value of these BC females was estimated by averaging the performances recorded on their mule duck offspring. QTL thus measured were detected by genotyping the BC females with microsatellites markers.
The primary objective of the present study was to identify chromosomal regions influencing production traits, mainly fatty liver and breast muscle quality traits of overfed mule ducks. However, to dissect the underlying physiological mechanisms, metabolic traits related to lipid metabolism during overfeeding were also investigated. First QTL detections were done on a trait by trait basis, after which multi-trait analyses were performed, for all co-localized QTL and all correlated traits, in order to identify pleiotropic QTL.

MATERIALS AND METHODS

Animals and husbandry
Experimental procedures were performed in accordance with French National Guidelines for the care and use of animals for research purposes (Certificate of Authorisation to Experiment on Living Animals n° 7740, Ministry of Agriculture and Fish Products).
The experimental design, a grandson design with male mule ducks measured to estimate the value of their mothers, is similar in principle to the granddaughter design proposed for dairy cows (Weller *et al*., 1990), with four generations (Figure 1). Great-grandparents (generation 0) were recruited in two experimental strains of Common ducks - I444, a light Kaiya strain (the crossbreeding product of a Tsaiya duck and an Asian Pekin duck) and I37, a heavy Pekin strain (a synthetic strain created from 3 heavy European Pekin lines). The choice of this cross was motivated by previous results from Rouvier *et al*. (1994) showing that growth and fatty liver performances in mule ducks sons of Tsaiya ducks, of Pekin ducks or of the 2 reciprocal (Tsaiya x Pekin) or (Pekin x Tsaiya) crossbreds were significantly different. Seven “grandson” families were produced, in which the grandsires (generation 1) are thus 7 (I444 x I37) F1 individuals, crossed to a total of 64 Common female I444 ducks, to produce a total of 382 BC females (generation 2). Finally, the BC females were mated to Muscovy drakes, to produce the 1,600 mule ducks (generation 3), which were used to estimate their phenotypic values. However, the number of mule duck progenies per BC female was highly variable, ranging from 1 to 8. In this experiment, all phenotypic measurements were recorded on mule ducks.

Figure 1: Experimental design



1)BC= Back-cross

As described in Marie-Etancelin *et al*. (2011), mule ducklings were bred in batches of 50 animals. They were fed *ad libitum* from 0 to 6 weeks of age with a starting diet (2,820 kcal ME/kg and 17.5% crude proteins) and then were fed restricted from 6 to 10 weeks of age (230 g/day with a growing diet, 2,850 kcal ME/kg and 15.5% crude proteins). With the same growing diet, the “pre overfeeding” period started at 10 weeks of age with 5 days of restriction (200 g/day) and 6 days of gradual increase of feed amount (from 220 to 320 g/j). At 12 weeks of age, ducks were overfed during 12 days in two successive series of 200 animals -with a gap of 2 days between series- starting at 80 and 82 days of age, with 2 different crammers. During this period, animals were bred in collective cages of 4 or 5 individuals and were overfed twice a day with a mix of 35% corn-flour, 25% corn grain and 40% water: the average feed amount ingested by animal and meal varied from 410 g to 825 g. At the end of the overfeeding period, the animals were slaughtered at 92 and 94 days of age, respectively. They were bled after a “head-only” electrical stunning and plucked. The carcasses were refrigerated 24 hours at 4°C. Then, they were eviscerated: fatty liver, breast muscles, legs and abdominal fat were removed. To avoid confusion between a group of full-sibs mule ducks (progeny of a BC dam) and fixed effects, mule ducks were selected to be split up into the 2 annual hatching series, then into different breeding batches and cramming series. For each animal, the hatching, breeding and overfeeding batches were recorded.

Measurements recorded on mule ducks are detailed in Table 1, and classified in 6 types of traits: growth traits before the overfeeding period, corticosterone traits, body weights and metabolic traits during the overfeeding period, overfeeding ability traits measured after slaughter, liver quality traits and muscle quality traits. Body weights were recorded at 12, 28, 42 and 70 days of age. A hanging test, measuring the animal stress response, was organised: at six weeks of age, ducks were hanged by the legs head downwards on a string during 10 minutes, and blood samples were taken before and after the test in order to estimate the corticosterone levels and to assess the response of the HPA (hypothalamic-pituitary-adrenal) axis to this hanging stress. Measurements of plasma metabolites contents (glucose, triglyceride and cholesterol) at the beginning (after the second meal), the middle (after the tenth

meal) and the end (after the twentieth meal) of the 12 days overfeeding period were performed. Body weights at the beginning and the end of the overfeeding period and feed consumption during the whole overfeeding period were recorded. Behavioural information such as the ease of catching and overfeeding was also registered at the beginning, middle and end of the overfeeding period. To appreciate the duck overfeeding ability, the carcass and the component pieces (fatty liver, thigh, breast skin, breast muscle and abdominal fat) were dissected and weighted. Measurements related to liver quality such as melting rate, lipid, protein and collagen contents, colour (L*, a*, b* coordinates in the CIELAB system) were recorded. The external liver traits (texture, pigmentation of lobes, necrosis area, extent of red spots …) were also appraised. Lastly, muscle quality was estimated by measuring the pH 20 minutes *post mortem* and the ultimate pH, the cooking losses and the drip losses under vacuum, the colour descriptive L*, a*, b* values and the lipid content recording. The raw meat tenderness was measured with the maximal shear force and the energy at the maximum of the Warner-Bratzler test. All meat quality measurements were carried out on the breast meat (*Pectoralis major* muscle). Genetic parameters of all these traits are detailed in Marie-Etancelin *et al*. (2011).

Table 1: Traits descriptions: abbreviations, means and standard deviation (n ranging from 340 to 342)

Trait		Means	s.d.
Growth measurements			
BW12, kg	Body weights (at 12, 28, 42, 70 days of age)	0.34	0.06
BW28, kg		1.37	0.13
BW42, kg		2.31	0.16
BW70, kg		3.40	0.21
BWG12-28, g/d	Body weight gain	60.92	4.58
BWG12-42, g/d		63.49	4.17
BWG12-70, g/d		51.84	3.07
BWG28-42, g/d		62.11	4.82
BWG28-70, g/d		47.04	3.48
BWG42-70, g/d		37.64	4.36
BW_A, kg	Asymptotic body weight	4.07	0.25
T_i, d	Age at inflexion point	23.51	1.60
Corticosterone traits			
CortL, ng/ml	Corticosterone level before stress	9.77	5.80
CortH, ng/ml	Corticosterone level after stress	55.95	28.16
DeltaC, ng/ml	Corticosterone difference between before and after stress	46.18	26.27
Body weights and metabolic traits during overfeeding period			
DFI, kg/d	Daily feed intake	1.32	0.04
BWbeg, kg	Body weight at beginning of overfeeding period	3.83	0.22
BWend, kg	Body weight at end of overfeeding period	5.80	0.25
OWG, kg	Weight gain during the overfeeding period	1.97	0.15
TG 2^{nd}M, g/l	Plasma triglyceride content (at 2^{nd}, 10^{th} and 20^{th} meal)	4.30	0.64
TG 10^{th}M, g/l		4.60	0.65
TG 20^{th}M, g/l		4.90	0.88
CHO 2^{nd}M, g/l	Plasma cholesterol content (at 2^{nd}, 10^{th} and 20^{th} meal)	1.71	0.14
CHO 10^{th}M, g/l		2.11	0.19
CHO 20^{th}M, g/l		2.46	0.30
GLU 2^{nd}M, g/l	Plasma glucose content	2.21	0.15

GLU 10thM, g/l	(at 2nd, 10th and 20th meal)	2.69	0.29
GLU 20th M, g/l		3.12	0.67
FC 2ndM[1)]	Catching ease (at 2nd, 10th and 20th meal)	2.55	0.60
FC 10thM[1)]		2.57	0.64
FC 20thM[1)]		2.53	0.65
FO 2ndM[1)]	Overfeeding ease (at 2nd, 10th and 20th meal)	1.97	0.69
FO 10thM[1)]		1.31	0.36
FO 20thM[1)]		1.24	0.36
Overfeeding ability traits			
CW, kg	Bled-plucked carcass weight	4.90	0.22
FLW, kg	Fatty liver weight	0.57	0.07
pmMW, kg	*Pectoralis major* muscle weight	0.26	0.02
pmSFW, kg	Breast skin+subcutaneous fat weight	0.15	0.01
TSW, kg	Thigh+shank weight	0.48	0.03
AFW, kg	Abdominal fat weight	0.17	0.02
Liver quality traits			
MR, %	Liver melting rate	38.67	8.56
LL*	Liver lightness	72.40	1.48
La*	Liver redness	9.19	1.16
Lb*	Liver yellowness	31.20	2.19
LLipC, %	Liver lipid content	52.37	2.72
LProtC, %	Liver protein content	7.69	0.48
LColC, mg/g	Liver collagen content	1.27	0.16
TFL[1)]	Texture of fatty liver	1.29	1.09
ERS[1)]	Extent of red spots	1.11	0.43
EG[1)]	Type of grain	0.95	0.54
ESV[1)]	Extent of spots due to the viscera	1.45	0.52
PL[1)]	Pigmentation's lobes	0.41	0.29
Muscle quality traits			
Energy, mJ	Energy needed to cut the muscle	149.58	26.37
Fmax	Maximal shear force	42.37	5.13
ML*	Muscle lightness	47.34	3.91
Ma*	Muscle redness	20.38	2.39
Mb*	Muscle yellowness	7.64	1.00
MpH20	Muscle pH20 minutes *post mortem*	6.01	0.11
MpHu	Muscle ultimate pH	5.72	0.07
MCookL, %	Muscle cooking losses	22.20	2.50
MvacL, %	Muscle drip losses	1.57	0.42
MWC, %	Muscle water content	71.69	0.54
MLipC, %	Muscle lipid content	4.92	1.27

[1)] *Non-continuous traits.*

Genotyping and map construction

Microsatellite markers were selected from 2 data sets: (i) from 158 duck markers developed by the INRA-LGC laboratory (Marie-Etancelin *et al*., 2006), which were submitted to the European Nucleotide Archive in December 2010 (see Appendix 1)

and (ii) from the international nucleotide sequences data-bases (GenBank/EMBL/DDBJ). In order to optimize the markers choice along the duck genome, we took advantage of the great similarity between the duck and chicken karyotypes: all duck nucleotide sequences markers were located by sequence similarity in the chicken genome to predict their positions in the duck genome. The blast e-value cutoff for first-pass mapping of duck microsatellites onto a chicken chromosome was $1e^{-4}$; for locating more precisely markers included in a duck genetic LG already assigned to a chicken chromosome the value was equal to $1e^{-15}$.
One hundred and sixteen microsatellites markers were selected amongst the most informative markers (at least 3 of the 7 F1 sires heterozygous at the marker position) and well distributed throughout the chicken genome. They were used to genotype the BC female ducks, their parents (sires F1 and dams I444) and their paternal grand-parents. Fluorescent microsatellite analyses were performed on ABI3100 and ABI3730 DNA sequencers (Applied Biosystems). The analyses were performed using GeneMapper V4.0 software, and genotypes were integrated in a local INRA database (GEMMA, LIMS) in which correct Mendelian inheritance could be checked. Linkage groups (LG) were built using the Crimap software 2.4 (Green *et al*., 1990).

Statistical methods
Growth traits were modeled following the Weibull model according to Maruyama *et al*., (1999, 2001). Parameters -the asymptotic weight (BW_A) and the age at the inflexion point (t_i) of the growth curve- were estimated with the following model:

$$BW_t = BW_A - (BW_A - B)exp^{\left[-\left(\frac{C-1}{C}\right)\left(\frac{t}{t_i}\right)^C\right]}$$

where BW_t is the body weight at age t, BW_A is the asymptotic weight and t_i is the age at the inflexion point. The parameters B and C have no biological interpretation. These parameters were estimated by non-linear regression with the NLIN procedure of SAS (2002) taking into account all available weights from birth to slaughter, i.e. at 12 days, 28 days, 42 days and 70 days of age.
Before the linkage analysis aiming at locating QTL, all mule duck traits (recorded or computed traits) were previously corrected for environmental fixed effects with a GLM procedure of SAS (2002). For all traits, the “hatching batch” effect were taken into account, adding the “breeding batch” effect for traits related to growth, or the “overfeeding batch” effect for traits related to overfeeding, or the “batch of measurement set” effect for traits linked to corticosterone. The residual effects of the previous linear model were kept and the performance for each BC female was computed as the average of its mule duck sons’ residual. A few traits (catching ease, overfeeding ease, liver texture, liver pigmentation's lobes or extent of red spots …) were not distributed on a continuous scale (Table 1), but were categorical data. Nevertheless, when mule duck residuals were averaged to be assigned to a given BC female, these traits could be considered as continuous. Thus, threshold traits were treated as normal traits.
QTL detection was carried out with the QTLMap software (Elsen *et al*., 1999; Gilbert *et al*., 2008; Filangi *et al*., 2010) in order to implement a linkage analysis according to the interval mapping method (Lander and Botstein, 1989). For each chromosome, probabilities of each possible phase of the F1 male founders were first estimated from their progenies’ marker information. The most likely sire phases were assumed to be the correct ones: for a set of tested positions (practically every 1 cM), the probabilities of corresponding chromosomal segments transmission to the offspring

were estimated. Then, QTL detection was carried out by within-sire linear regression (Knott *et al*., 1996). The model was the following:

$$Y_{ij} = s_i + (2p_{ij} - 1)a_i + e_{ij}$$

where the dependent variable Y_{ij} is the average performances (previously corrected for fixed effects) of the n_{ij} mule ducks sons of BC *j* of sire *i*. For each location on the genome, s_i is the male founder *i* effect, a_i half the substitution effect of the putative QTL carried by the sire *i*, and p_{ij} the probability, for daughter (BC) *j,* of inheriting one arbitrarily defined QTL allele from her sire *i*, given the marker information. The residual variance e_{ij} was defined within sire families to improve robustness to unlinked QTL segregation between families (Goffinet *et al*., 1999).

In our "grandson design" with recorded phenotypes belonging only to the third generation, phenotypes assigned to the BC generation had a variance expressed as:

$$var(z_j) = \sigma^2 \frac{1 + \frac{1}{4}(n-1)h^2}{\frac{1}{4}n}$$

with $\sigma^2 = \sigma_g^2 + \sigma_e^2$, *n* the number of sons of BC and σ_g^2, σ_e^2, σ^2, the genetic, residual and total variances, and $h^2 = \sigma_g^2/\sigma^2$ the heritability of the trait (Kileh-Wais and Elsen, 2012). Those variances were computed with the Varcomp procedure of SAS software (2002). For the multi-trait approach, we took into account the covariance between traits S and T as

$$cov(z_j) = \frac{n_{ST} + \frac{1}{4}(n_S n_T - n_{ST})r_{ST}}{n_S n_T}$$

with n_S the number of mule ducks, sons of the BC j for trait S, n_T this number for trait T, n_{ST} the number of sons of BC j for both traits S ant T and r_{ST} the ratio between genetic and phenotypic covariances between traits S and T.

An adapted version of the QTLMap software (Elsen *et al*., 1999) was developed to take into account the phenotypic variance heterogeneity in single and in multi-trait analyses (Filangi *et al*., 2010).

For each trait and each linkage group, 1,000 within-family permutations were performed to estimate empirical chromosome-wide significance level of the test statistics (Churchill and Doerge, 1994). The conservative genome-wide thresholds were derived from chromosome-wide significance levels, using an approximate Bonferroni correction:

$$P_{genome\ -wide} = 1 - (1 - P_{chromosome\ -wide})^{1/r}$$

where r is the ratio between the length of a specific linkage group and the length of the genome considered for QTL detection (778 cM). The 95% confidence intervals of the QTL locations were estimated by lod drop-off. In practice, the bounds of the interval are the two locations where likelihood was equal to the maximum likelihood minus 3.84 (= $\chi^2_{(1,0.05)}$). The QTL effect was expressed in phenotypic deviation units

(SD), and estimated as: $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$, where *SD* is the phenotypic standard deviation and *n* is number of sires (Roldan *et al*., 2010).

QTL detection was first carried out on a single trait basis. As it has been shown that in the case of pleiotropic QTL (named plQTL), the fact of using simultaneously phenotypic information from different correlated traits increases the QTL location precision and possibly increases the power of evidencing effects too small to be detected in single trait analysis (Gilbert and Le Roy, 2003), multiple traits QTL detection were also performed. Traits to be included in the multiple trait analyses were grouped according to two main criteria: (1) the "co-location criteria" (Co-located QTL: CQ) for traits for which two or more significant QTL (at 5% chromosome-wide threshold) were located in the same chromosome area by single trait approaches and (2) the "correlation criteria" (Phenotypic Correlation: PC) for traits presenting phenotypic correlations higher than 0.55 (see Appendix 2). The CQ criteria aims at refining the QTL location if the single traits analyses indeed correspond to a single plQTL, whereas the PC criteria is used to detect new plQTL that could not be detected by the single traits approach. Finally, as an additional phenotypic approach, metabolism traits during overfeeding were grouped by meal (2^{nd}, 10^{th} or 20^{th} meal) and component (CHO, TG and GLU) in multiple traits QTL detection. The multi-traits detections were carried out by recursive backward approaches: among the n co-located or correlated traits for which a significant QTL was found, all sub-groups of traits (2 by 2, 3 by 3, etc ...) were tested to identify the underlying pleiotropic QTL. The most significant combinations traits which *P*-value were higher than the highest single trait *P*-value, were retained and presented in the multi-traits QTL Tables.

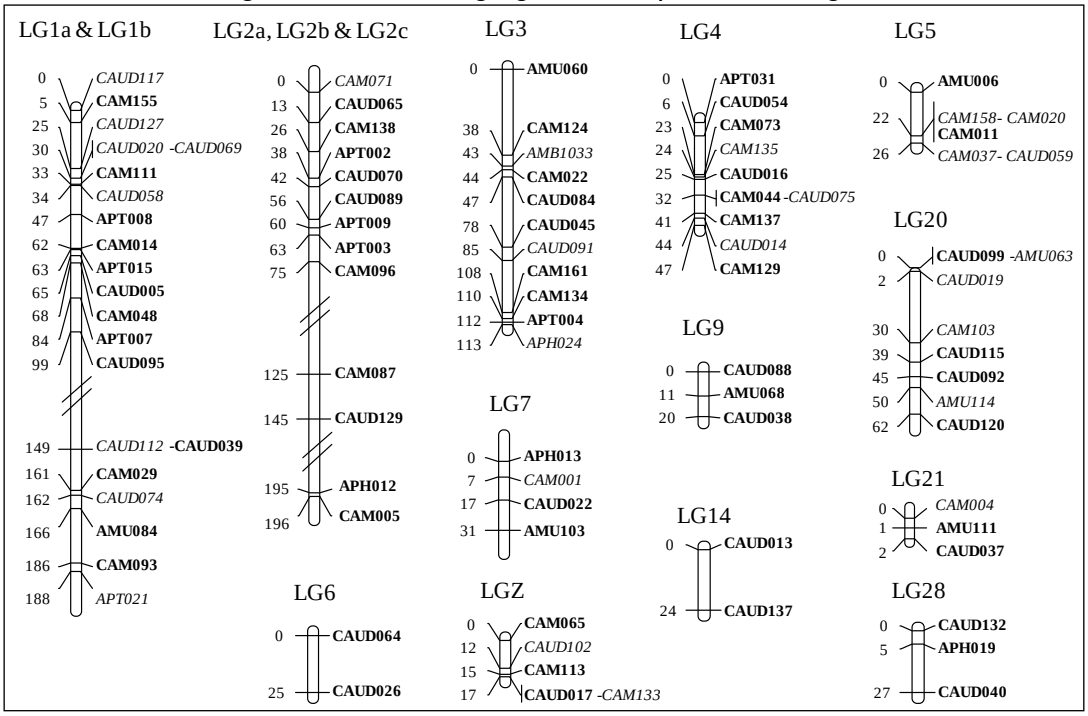
RESULTS

Genetic maps

A genetic map consisting of 91 markers aggregated into 16 linkage groups (LG) and covering a total of 778 cM was obtained (Figure 2). The sequence alignment by BLAST analysis of the duck microsatellite markers to the chicken genome sequence allows the assignment of the 16 duck LG to 13 different chicken chromosomes. Previous reports showed the strong conservation of synteny between duck and chicken (Fillon *et al*., 2007; Skinner *et al*., 2009), allowing to propose a nomenclature of duck chromosomes based on that of chicken (Skinner *et al*., 2009).

In these studies, the major rearrangement between the duck and chicken karyotypes found is that the ancestral chromosomes 4 and 10, fused in the chicken lineage to give GGA4, remain separate in duck. Apart for this only interchromosomal rearrangement, conservation of synteny between the two species was demonstrated, including all microchromosomes analyzed. We thus consider here that duck chromosomes APL1 to APL9 correspond to chicken chromosomes GGA1 to GGA9, then APL10 corresponds to GGA4p and finally, the rest of the karyotype is offset by one, with GGA10 corresponding to APL11 and so on. Therefore, the number of the LG in our genetic map was the number of corresponding duck chromosome (APL). As some of our independent LG were located by BLAST search on the same chicken chromosome, we used letters in the nomenclature to distinguish them (e.g. LG1a and LG1b for two independent LG assigned to GGA1).

Figure 2: Sex-average genetic map in centiMorgan

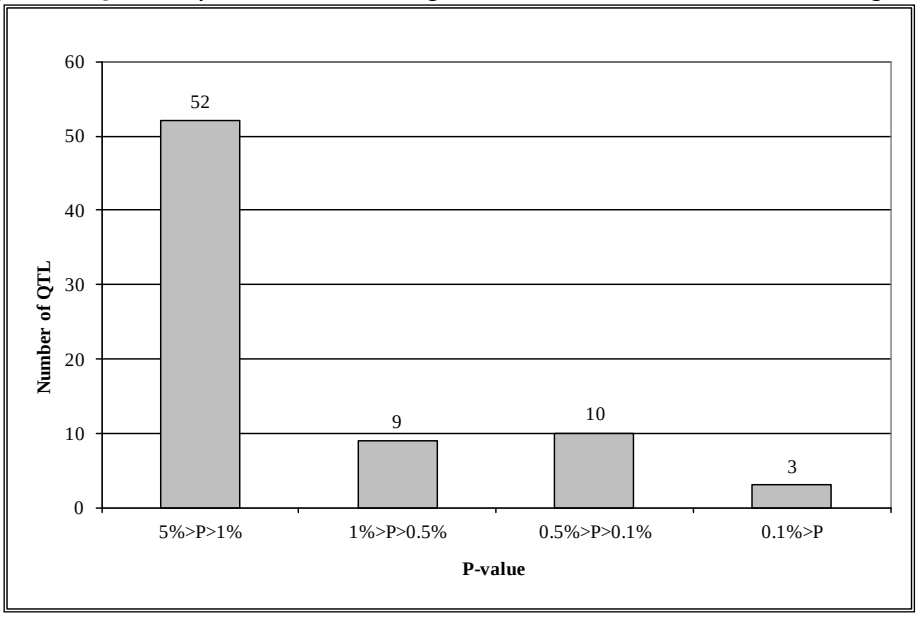


LG: Linkage Group
Markers in boldface belong to the framework map and markers in italics belong to the putative map.

QTL detection

Single trait QTL analysis on all the measured traits allowed the detection of 74 QTL (Figure 3). Out of these, 22 were significant at 1% threshold at the chromosome-wide level. The presentation of the results will focus on these most significant QTL, most of which mainly concern overfeeding ability and product quality traits. QTL will be first presented as three groups of traits involved in functions which are agronomically related and for which physiological links can be suspected: “growth”, “body weight and metabolism during overfeeding” and “overfeeding ability and quality of products”. These will be presented using first single trait and subsequently multi-traits approaches, the latter allowing the detection of plQTL. Finally, multi-trait detections will be presented for plQTL concerning more than one of our three groups of traits. QTL upper the 1% chromosome-wide threshold (in single trait and multi-traits detections) are detailed in Tables 2 to 8 and QTL under the 1% chromosome-wide threshold (in single trait detections) presented in Appendix 3.

Figure 3: QTL frequencies according to the chromosome-wide level significance



Traits related to growth

QTL for growth traits are presented in Table 2 for the single trait approach and Table 3 for the multi-traits plQTL approach.

Two QTL significant at 0.5% threshold at chromosome-wide (and 5% at genome-wide) - one for body weight at 28 days of age and one for body weight gain between 12 and 28 days - were detected by the single trait analyses on linkage group 3. Six others QTL significant at the 5% threshold chromosome-wide, for BW42, BW70, BWG12-28, BWG12-42, BWG12-70, BWG28-42, and BWG28-70, were co-localized on LG2b (suppl. 3). However, none of these reached the 1% chromosome-wide threshold. Thanks to the Weibull curve approach, 3 new QTL were detected: 2 QTL for Ti respectively on LG7 (at 1% threshold) and LG1b (at 5% threshold – suppl. 3) and 1 QTL for BW_A on LG20 (at 5% threshold – suppl. 3). The Ti QTL on LG7 presented a relatively strong QTL effect of 0.30 in phenotypic deviation units. The lack of highly significant QTL for BW_A is consistent with the absence of QTL for the bodyweight at 70 days, since both traits are highly correlated (0.78).

Table 2: Growth traits – in a single trait analysis

LG	Traits[1)]	Flanking markers[2)]	Location[3)] (cM)	LRT_x[4)]	*P*-value[5)]	Effect of QTL substitution[6)]	Confidence interval (cM)
3	BW28	CAM022-CAUD084	46	29.10	** / †	0.14	33 - 64
3	BWG12-28	CAUD084-CAUD045	51	27.75	** / †	0.19	35 - 65
7	Ti	APH013	0	21.93	*	0.30	0 – 6

[1)] See Table 1 for the definitions of the traits.
[2)] Flanking markers of the most probable QTL position.
[3)] Most probable QTL position.
[4)] Maximum likelihood ratio for x locus.
[5)] Level of significance of P-value - at chromosome-wide: 0.01>P>0.005; ** 0.005>P>0.001; *** 0.001>P*
- at genome-wide: †0.05>P>0.01; †† 0.01>P
[6)] QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

In multi-traits analyses, neither the co-localized QTL approach nor the correlation between traits approach allowed to confirm QTL or to identify a new one. The multi-traits detection for BW28 and BWG12-28, which are strongly correlated (r = 0.92) and presented 2 co-localized QTL on LG3, did not reveal any plQTL reaching 1% threshold on LG3. However, by combining the two Weibull parameters (Ti and BW_A) in a multi-traits analysis, a new plQTL significant at 1%, appeared in LG21 (Table 3). This is consistent with the observation that in single trait analyses the LRTs on LG21 for Ti and BW_A respectively are nearly significant.

Table 3: Growth traits – in a multiple traits analysis

LG	Traits group & type[1)]	Traits[2)]	Flanking markers[3)]	Location[4)] (cM)	LRT_x[5)]	*P*-value[6)]	Effect of QTL substitution[7)]	Confidence interval (cM)
21		BW_A	CAM004	0	25.96	*	0.12	0 - 2
	PC	Ti					0.11	

[1)] Criteria to select the traits for multiple trait analysis: CQ for QTL Co-located – PC for Phenotypic Correlation.
[2)] See Table 1 for the definitions of the traits.
[3)] Flanking markers of the most probable QTL position.
[4)] Most probable QTL position.
[5)] Maximum likelihood ratio for x locus.
[6)] Level of significance of P-value - at chromosome-wide: 0.01>P>0.005; ** 0.005>P>0.001; *** 0.001>P*
- at genome-wide: †0.05>P>0.01; †† 0.01>P
[7)] QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

Traits related to bodyweight and metabolism during the overfeeding period

QTL related to metabolism traits during overfeeding (P<0.01) are presented in Tables 4 and 5 for single and multiple traits analysis, respectively.

Table 4: Metabolic and body weight traits during the overfeeding phase – in a single trait analysis

LG	Traits[1)]	Flanking markers[2)]	Location[3)] (cM)	LRT_x[4)]	*P*-value[5)]	Effect of QTL substitution[6)]	Confidence interval (cM)
2a	FO 20th	CAM138-APT002	28	24.67	*	0.34	15 - 37
4	FC 2nd	CAM137	41	34.71	*** / ††	0.23	38 - 44
Z	TG 20th	CAM065-CAUD102	1	26.54	**	0.48	0 - 8

[1)] See Table 1 for the definitions of the traits.
[2)] Flanking markers of the most probable QTL position.
[3)] Most probable QTL position.
[4)] Maximum likelihood ratio for x locus.
[5)] Level of significance of P-value - at chromosome-wide: 0.01>P>0.005; ** 0.005>P>0.001; *** 0.001>P*
- at genome-wide: †0.05>P>0.01; †† 0.01>P
[6)] QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

Only one strong QTL was found for plasma triglyceride content after the 20th overfeeding meal on LGZ (corresponding to the Z gonosome) with a *P*-value lower than 0.5% at the chromosome-wide level and a QTL substitution effect of 0.48 phenotypic standard deviations. None of the other traits, related to blood plasma contents (glucose, triglyceride or cholesterol) or to the stage of the overfeeding process (2nd, 10th or 20th meal), reached the 1% threshold at chromosome-wide level.

For the 2 traits related to the ease of handling the animals during overfeeding (overfeeding per se and catching), 2 QTL reached the 1% threshold: 1 on LG2a for overfeeding ease at the 20th meal and one highly significant QTL (P<0.001) on LG4 for catching ease at the 2nd meal. Finally, for this group of metabolic and behavior traits during overfeeding, 15 QTL, significant at 5% threshold at the chromosome-wide level (suppl. 3), were scattered on 9 different LG: 4 linkage groups with co-localized QTL (LG2a, 2c, 14, and 21) and 5 with isolated QTL (LG1b, 3, 6, 7 and 28). To sum up, the plasma cholesterol contents were affected by QTL on LG2c, 3, 14 and 21, while the plasma triglyceride contents were influenced by QTL on 2 of these 4 LG (LG14 and LG21). The main QTL involved in the genetic determinism of glucose contents were located on 3 others LG (LG2a, 6 and 28). Finally, LG7 and LG28 carried QTL affecting the body weights at the beginning of the overfeeding period, whereas 2 QTL on LG9 and LG21 affected respectively the weight gain during the overfeeding period and the weight at the end of overfeeding period.
In multiple trait analyses, the co-location criteria on LG2a, LG14 and LG21, and the correlation criteria between cholesterol and triglyceride contents after the 10th meal revealed 4 potential plQTL areas, 2 of which are located on LG14. A strong plQTL (P<0.001) was revealed on LG2a by combining corticosterone level before stress and glucose content at the 2nd meal. Another strong plQTL reaching the same *P*-value threshold (P<0.001) was confirmed on LG21 for cholesterol and triglyceride contents at the 10th overfeeding meal. This plQTL was revealed both by the analysis of two co-located uni-trait QTL on LG21 and by the analysis of correlated traits. The correlated-traits approach for triglyceride and cholesterol contents after the 10th meal also revealed a weaker but new plQTL on LG14, this result being consistent with the plQTL (at 0.5% threshold) obtained on LG14 after the analysis based QTL co-localisation of cholesterol content after the 10th meal and triglyceride content after the 2nd meal. For this last plQTL, the correlated-traits approach has also reduced the confidence interval to 0-17 cM, whereas it was respectively 0-22 cM and 0-24 cM for TG2nd and CHO10th in single trait detections. Thus, two plQTL were found for triglyceride content after the 2nd or the 10th meal, combined with cholesterol contents after the 10th meal, on LG 14: indeed, both single trait QTL for TG2nd and CHO10th have *P*-values around 0.03 whereas for TG10th the *P*-value reached only 0.06. The close location of the 2 plQTL suggests that they may be confounded into one plQTL concerning the three traits.

Table 5: Metabolic and body weight traits during the overfeeding phase – in a multiple traits analysis

LG	Traits group & type[1)]	Traits[2)]	Flanking markers[3)]	Location[4)] (cM)	LRT_x[5)]	*P*-value[6)]	Effect of QTL substitution[7)]	Confidence interval (cM)
2a	CQ (2a)	GLU2ndM	CAM071-CAUD065	12	40.55	***/†	0.11	0 - 18
		CortL					0.15	
14	CQ(14)	CHO10thM	CAUD013-CAUD137	6	35.31	**	0.11	0 - 17
		TG 2ndM					0.15	
14		CHO10thM	CAUD013-CAUD137	20	27.33	*	0.11	6 - 24
	PC	TG 10thM					0.08	
21	CQ(21)	CHO10thM	CAUD037	2	37.48	***	0.10	0 - 2
	PC	TG 10thM					0.10	

[1)] Criteria to select the traits for multiple trait analysis: CQ for QTL Co-located – PC for Phenotypic Correlation.
[2)] See Table 1 for the definitions of the traits.
[3)] Flanking markers of the most probable QTL position.

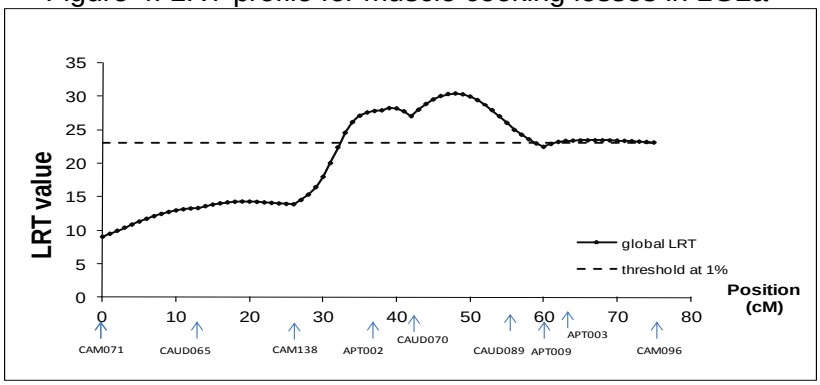
[4] Most probable QTL position.
[5] Maximum likelihood ratio for x locus.
[6] Level of significance of P-value - at chromosome-wide: 0.01>P>0.005; ** 0.005>P>0.001; *** 0.001>P*
- at genome-wide: †0.05>P>0.01; †† 0.01>P
[7] QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

Traits related to overfeeding ability and quality of products

QTL related to the overfeeding ability and product quality traits are presented in Tables 6 and 7 for the single and multiple trait analyses, respectively.

In single trait detection, 16 QTL were identified at the 1% chromosome-wide threshold in 7 different LG, with LG2a gathering QTL related to muscle quality traits, whereas LG2c and LG9 are particularly involved in QTL for fatty liver quality traits. The 2 strongest QTL (at 1% at genome-wide level of significance) were for muscle pH 20 minutes *post mortem* on LG4 and for muscle cooking losses on LG2a (Figure 4).

Figure 4: LRT profile for muscle cooking losses in LG2a



Still on LG2a, but not at the same position, 2 QTL for lipid contents in muscle and, with a smaller *P*-value, for muscle maximal shear force, were revealed. Finally for breast meat quality, a QTL on LG6, significant at 0.5% and explaining 0.23 phenotypic standard units of muscle cooking losses variation was found. The variability in fatty liver weight was partly controlled by 3 QTL on LG21, LG9 and LG2c. Moreover, the latter 2 linkage groups (LG9 and LG2c) have also QTL related to fatty liver quality. On LG2c, 2 QTL for lipid and protein contents were co-localized with fatty liver weight QTL and interestingly, these 3 traits are highly correlated. The QTL substitution effect for liver protein content on LG2c reached 0.31 phenotypic standard deviations. On LG9, 2 QTL for protein content and liver redness were detected at 1% threshold but only the QTL for LprotC was co-localized with the QTL chromosome-wide level, for carcass weight, fatty liver texture and fatty liver weight.

Table 6: The overfeeding ability and product quality traits – in a single trait analysis

LG	Traits[1)]	Flanking markers[2)]	Location[3)] (cM)	LRT_x[4)]	*P*-value[5)]	Effect of QTL substitution[6)]	Confidence interval (cM)
2a	MLipC	CAUD065-CAM138	18	27.40	**/†	0.10	3 - 26
2a	MFmax	CAM138-APT002	28	25.37	*	0.23	22 - 33
2a	MCookL	CAUD070-CAUD089	48	30.43	***/††	0.18	34 - 55
2c	FLW	APH012	0	19.10	*	0.21	0 - 1
2c	LLipC	APH012	0	20.10	*	0.22	0 - 1
2c	LProtC	APH012	0	21.26	**	0.31	0 - 1
3	ESV	AMU060-CAM124	9	25.13	*	0.20	0 - 26
4	MpH20	APT031	0	30.27	***/††	0.22	0 - 5
6	ESV	CAUD064-CAUD026	19	17.28	*	0.20	0 - 25
6	MCookL	CAUD026	25	21.39	**	0.23	7 - 25
9	La*	CAUD088	0	25.75	**	0.22	0 - 10
9	LProtC	CAUD038	20	23.28	*	0.25	0 - 20
9	FLW	CAUD038	20	21.64	*	0.20	5 - 20
21	CW	AMU111	1	22.47	**	0.25	0 - 2
21	TFL	CAUD037	2	23.49	**	0.06	0 - 2
21	FLW	CAUD037	2	23.25	**	0.21	0 - 2

1) See Table 1 for the definitions of the traits.
2) Flanking markers of the most probable QTL position.
3) Most probable QTL position.
4) Maximum likelihood ratio for x locus.
5) Level of significance of P-value - at chromosome-wide: 0.01>P>0.005; ** 0.005>P>0.001; *** 0.001>P*
- at genome-wide: †0.05>P>0.01; †† 0.01>P
6) QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

The multiple traits analysis revealed 9 pleiotropic QTL, 8 of which related to products quality and mainly to liver quality. Four plQTL on LG2c, LG3 and LG9 with very significant thresholds (reaching 0.5% or 0.1% threshold at the chromosome-wide) were suggested by the co-localization of the single-trait QTL, whereas 5 others plQTL were new ones, identified on LG1b, LG3, LG5 and LG7 by the joint analysis of 3 couples of traits. However, most of these 5 plQTL only reached the 1% threshold. These multi-traits analyses confirmed the importance of LG2c for the variability of the liver melting rate, when this trait is associated with the liver structural composition (such as protein contents or collagen contents) and, to a lesser extent, that of LG9 for liver composition variability. LG3 seemed to contain three plQTL: one for liver yellowness and spots on liver due to viscera at the beginning of the linkage group, for which the confidence interval (between 0 to 16 cM) was reduced as compared to single trait detections (confidence intervals between 0 to 19 cM for Lb* and between 0 to 26 for ESV) ; and 2 others in the middle (at 85 cM when combining melting rate and liver protein contents) and at the end (at 108 cM when combining liver protein and lipid contents). One strong plQTL for liver composition traits appeared (at 0.1% threshold) on LG7, which is a totally new information since none of the traits forming this group has a QTL (even at 5%) with the single trait approach. Lastly, the combination of 2 muscle rheological traits (energy needed to cut the breast muscle and maximal shear force) allowed the discovery of a plQTL on LG5.

Table 7: The overfeeding ability and product quality traits – in a multiple traits analysis

LG	Traits group & type[1)]	Traits[2)]	Flanking markers[3)]	Location[4)] (cM)	LRT_x[5)]	*P*-value[6)]	Effect of QTL substitution[7)]	Confidence interval (cM)
1b	PC	MR LProtC	APT021	39	32.21	*	0.10 0.07	32 - 39
2c	CQ(2c) PC	MR LProtC	APH012	0	35.34	***	0.11 0.14	0 - 1
2c	CQ(2c)	MR LcolC	APH012	0	27.33	**	0.12 0.11	0 - 1
3	PC	LProtC LLipC	CAM134-APT004	108	35.26	*	0.12 0.10	98 - 110
3	CQ(3)	Lb* ESV	AMU060-CAM124	6	47.41	***/††	0.11 0.14	0 - 16
3	PC	MR LProtC	CAUD091	85	34.34	*	0.10 0.13	78 - 94
5	PC	Fmax Energy	CAM037	26	27.51	*	0.11 0.07	10 - 26
7	PC	LProtC LLipC	APH013-CAM001	6	33.67	***/†	0.07 0.11	0 - 12
9	CQ(9) PC	LProtC LLipC	CAUD038	20	34.94	**	0.12 0.11	0 - 20

1) Criteria to select the traits for multiple trait analysis: CQ for QTL Co-located – PC for Phenotypic Correlation.
2) See Table 1 for the definitions of the traits.
3) Flanking markers of the most probable QTL position.
4) Most probable QTL position.
5) Maximum likelihood ratio for x locus.
6) Level of significance of P-value - at chromosome-wide: 0.01>P>0.005; ** 0.005>P>0.001; *** 0.001>P*
- at genome-wide: †0.05>P>0.01; †† 0.01>P
7) QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

plQTL influencing phenotypes across groups of traits

The most significant multi-traits plQTL detections, acting simultaneously on two or more traits belonging to different groups of traits "growth", "bodyweight and metabolism during overfeeding" and "overfeeding ability and quality of products" quality), are presented in Table 8. As these traits can concern very different physiological functions, correlations between traits were usually low (all under 0.55) so the multi-traits QTL detection across group was only performed for traits having co-localised QTL.

Four plQTL were obtained on LG7, LG9 and LG21, among which 3 reached the 0.1% threshold. Association of weight gain during the overfeeding period and fatty liver weight traits on LG9 allowed to detect a strong plQTL (at 0.1% threshold) with a smaller QTL confidence interval, reduced from positions 5 - 20 cM and 8 - 20 CM for FLW and OWG respectively, down to positions 11 - 20 cM for the plQTL. On LG21, a plQTL was detected by combining carcass weight and body weight at the end of overfeeding period. Although the Bonferroni correction for multiple tests is very conservative, the maximum likelihood ratio for this plQTL reached the 5% genome-wide threshold. This plQTL is probably mostly determined by CW for which a QTL was also detected on LG21 at 1% threshold in single trait analysis. Still on LG21, another plQTL, significant at 0.1% threshold, was observed by associating overfeeding ease measured at the 10th meal and liver collagen contents. This plQTL

is characterized by a sharp increase in the detection power since the single trait detection threshold was only 5% for both traits. However, the biological explanation for this is unclear. Finally, thanks to the combination of body weight at the beginning of overfeeding period and body weight gain between 42 and 70 days, a plQTL on LG7 appeared while this linkage group was not previously identified as impacting late growth.

Table 8: mixed group – in a multiple traits analysis

LG	Traits group & type[1)]	Traits[2)]	Flanking markers[3)]	Location[4)] (cM)	LRT_x[5)]	*P*-value[6)]	Effect of QTL substitution[7)]	Confidence interval (cM)
7		BWbeg	APH013	0	25.00	*	0.18	0 – 14
	CQ(7)	BWG42-70					0.17	
9		OWG	CAUD038	20	39.02	***/†	0.10	11 – 20
	CQ (9)	FLW					0.11	
21		BWend	AMU111	1	53.57	***	0.18	0 - 2
	CQ(21)	CW					0.17	
21		FO10th	CAM004	0	30.19	***	0.12	0 - 2
	CQ(21)	LColC					0.10	

[1)] Criteria to select the traits for multiple trait analysis: CQ for QTL Co-located – PC for Phenotypic Correlation.
[2)] See Table 1 for the definitions of the traits.
[3)] Flanking markers of the most probable QTL position.
[4)] Most probable QTL position.
[5)] Maximum likelihood ratio for x locus.
[6)] Level of significance of P-value - at chromosome-wide: 0.01>P>0.005; ** 0.005>P>0.001; *** 0.001>P*
- at genome-wide: †0.05>P>0.01; †† 0.01>P
[7)] QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

DISCUSSION

This paper reports the first QTL detection work carried out on overfed inter-specific mule ducks. Single trait analyses, as a first step of our exploration, allowed to detect 22 QTL significant at 1% threshold at the chromosome-wide level. The ratios between the number of QTL detected and the number of traits recorded for a physiological function are very different: very few QTL were detected on traits related to metabolism (blood lipid metabolism traits during overfeeding and corticosterone level during growth) or growth traits, while, in proportion, many QTL were detected for liver and muscle qualities traits. Subsequently, with the multi-traits detection, 18 pleiotropic QTL (plQTL) were identified, with at least 7 new chromosomal regions not previously identified by the univariate analyses. Indeed, the multi-traits QTL analyses were done either i) on groups consisting of correlated traits, or ii) on traits QTL located close to one another after the single QTL detection analyses. With this second type of grouping, the estimated positions of the pleiotropic QTL were similar to one of the positions found in the single trait QTL analyses, and moreover, the confidence interval was generally reduced. This is consistent with Gilbert and Le Roy (2003) who demonstrated that using information from different correlated traits sharing a pleiotropic QTL increases the precision of the estimated location of this QTL. This approach highlighted the most interesting chromosomal regions especially for fatty liver and breast muscle quality traits. The other grouping method (correlation

criteria) allowed to identify new QTL. These QTL could not be detected in single trait analyses because of their small effects: the information gain obtained thanks to the combination of these correlated traits was particularly interesting for duck metabolism traits.
The statistical analyses were performed with some approximations. The data were pre-corrected for significant nuisance effects evidenced by SAS-GLM. This two steps procedure greatly simplified the calculations, in particular for multi-trait analyses, and is generally considered as safe in terms of QTL location estimation (Mc Rae *et al*., 2005; Lillehammer *et al*., 2009). As demonstrated by Crooks *et al*. (2009) for polygenic effect, when data are pre-corrected, possible confusion between certain levels of a nuisance effect and transmitted QTL alleles could result in underestimation of QTL effect and loss of power. It must be emphasized that this approach is conservative and should not result in false QTL detection. The traits analyzed being means of BC offspring (mule ducks) performances: the variability of their variances had to be considered in the model. Both heritabilities and genetic to phenotypic covariance ratios were given as parameters and not estimated with QTL effects. This is clearly an approximation which supposed that the QTL effects are not too strong. This two steps procedure is in the line of the GRAMMAR model for association analysis developed by Aulchenko *et al*. (2007). The backward approach used for multi-trait detection generated many statistical tests, and the significance of the results must be often considered with caution. However, the improvement of QTL localization generated by this approach is probably free of this drawback.

Twelve among the 22 QTL concerned original traits such as liver quality (7 QTL) and muscle quality (5 QTL) for overfed ducks. Some of these traits, especially the latter correspond to traits studied in chicken. The karyotypes of the two species are very similar, with only one inter-chromosomal rearrangement detected to date (Fillon *et al*., 2007; Skinner *et al*., 2009), suggesting that QTL located on orthologous chromosomes could be equivalent. However, further investigations to refine duck QTL location and chicken-duck intra-chromosomal rearrangements will have to be done. The strongest QTL related to breast muscle quality traits (concerning cooking losses, lipid content and maximal shear force to cut the muscle) were mainly gathered on linkage group 2a: QTL for MCookL and for MLipC reached respectively 1% and 5% genome-wide thresholds. In Beijing-You chickens, Ye *et al*. (2010) assessed the association of single nucleotide polymorphisms (SNP) in the fatty acid binding proteins (*A-FABP*) gene, located on GGA2, with the content of intramuscular fat (IMF). Another QTL for muscle cooking losses was detected on LG6: Huang *et al*. (2007a) have already found a QTL for MCookL trait on CAU6, although the breast muscle was not fattening by overfeeding. Regarding pH 20 minutes *post mortem*, 1 strong QTL reaching the 1% genome-wide threshold, was revealed on LG4. On chicken, Wright *et al*. (2006) and Nadaf *et al*. (2007) already mapped QTL for breast meat ultimate pH with QTL Express software on chromosome 4 at 233 cM and 201 cM, respectively. Although this is not exactly the same trait involved in the 2 poultry species (ultimate pH in chicken versus pH 20 minute *post mortem* in duck), this same QTL location on LG4 between species is surprising especially as the chicken is a white meat type species and the duck is a red meat type. Finally, the multi-traits analysis brought up a pleiotropic QTL for meat texture (energy needed to cut the muscle and maximal shear force traits) on LG5, but to our knowledge, there is no publication on this type of trait on chicken.

More specifically, QTL related to fatty liver quality cannot be compared with any previously published results, since the present work is the first design for QTL detection in overfed waterfowl. On chromosomal regions 2c, 7 and 9, remarkable clusters of QTL impacting simultaneously at least 2 of these 4 traits (liver melting rate, lipid, protein or collagen contents) are noticeable. These trait associations are very consistent as collagen and proteins are considered important for structuring the fatty liver tissue: a variation of this liver "extracellular matrix" would logically impact on lipids storage capacity and on the liver melting rate. Indeed, recent work has highlighted the role of the protein fraction in the determinism of the technological quality of fatty livers (Théron *et al*., 2011). On LG2c, the strongest improvement of the QTL detection power was obtained by associating melting rate and liver protein content, whereas the pleiotropic QTL on LG7 and LG9 were more related to liver composition (lipid and protein contents).
Regarding the overfeeding ability traits, two of the 3 QTL influencing the fatty liver weight were on LG2c and on LG9, co-localized with liver quality trait QTL. However, higher *P*-values of pleiotropic QTL were obtained by associating two co-localized liver quality traits rather than by associating liver weight and a liver quality trait. The co-localization of QTL for traits measured independently consolidates the reliability of our results, but the chromosomal regions that we identified didn't match with previously published results in other birds. Indeed, on duck or on chicken, most of the liver weight QTL detected did not intersect our QTL localizations: 3 significant areas on LG4, LG6 and CAU7 (corresponding to our LG20) in lean Pekin ducks (Huang *et al*., 2007a) and 3 others on GGA1, GGA8 and GGA12 (corresponding to our LG13) for chicken (Wright *et al*., 2006). Only Gao *et al*. (2009) revealed a liver weight QTL on the GGA2 which a confidence interval covering half of the chromosome in a F2 cross between Silkies fowl and White Plymouth Rock chicken.

For the growth traits, only 3 QTL were detected: 2 for growth between 12 and 28 days of age on LG3 (at 5% genome-wide threshold) and 1 on LG7 (at 1% chromosome-wide threshold). These results are consistent with those found in chicken, as GGA3 is an interesting chromosome for growth traits, with more than 50 QTL identified (www.animalgenome.org, Hu *et al.,* 2010) and numerous publications highlighting QTL related to young ages (Siwek at al., 2004; Ambo *et al*., 2009; Wahlberg *et al*., 2009; Atzmon *et al*., 2008). Maruyama *et al*. (2001) and Vitezica *et al*. (2010) demonstrated that the Weibull function is the best model for fitting duck body weight data, compared to others models such as Gompertz, logistic, von Bertalanffy, Morgan-Mercer-Flodin or Spline. Performing QTL detections from Weibull parameters has highlighted a new area of interest for growth traits in LG7, which has a pleiotropic QTL for growth between 42 and 70 days of age and bodyweight at the beginning of the overfeeding period. Moreover, GGA7 is also known to contain 28 growth traits QTL in chicken (Zhou *et al*., 2007; Atzmon *et al*., 2008; Ambo *et al*., 2009). Nevertheless, the Pekin duck experimental design analyzed by Huang *et al*. (2007b) revealed 4 chromosomal areas (at 5% genome-wide threshold - 2 QTL on LG1 and 2 others QTL on LG2) impacting growth traits which were different from ours QTL localizations.
The multi-traits approach within growth traits has only brought one new QTL when gathering both Weibull parameters (Ti and BW_A) on LG21. By comparing this result with the QTL detected for the bodyweight at the end of overfeeding and the carcass weight on LG21, we could hypothesize that this QTL area modified both the growth curve and the final product as carcass weight. In a specific intercross between two

chicken lines divergently selected for juvenile body-weight, Wahlberg *et al*. (2009) identified 7 regions harboring QTL influencing growth (GGA1, GGA3, GGA4, GGA7 and GGA20) including a network of interacting loci on GGA7. Knowing that GGA20 corresponds to APL21 and the short arm of GGA4 to APL10 (Skinner *et al*., 2009) for which we have no marker yet, we could highlight the consistency between our linkage groups containing QTL for growth traits in young age, and the results by Wahlberg *et al*. (2009) in chicken.

Regarding metabolism traits, we detected only one QTL for triglyceride contents in blood at the end of the overfeeding period. This QTL, located on LGZ (at 1% on the chromosome-wide level) was not identified in chicken QTL literature: only a QTL for glucose content was identified by Zhou *et al*. (2007) on this sexual chromosome. Thanks to the multi-traits analyses, 3 new areas of interest for metabolic traits during the overfeeding phase appeared on LG2a, LG21 and, with a lower likelihood on LG14. On LG14 and LG21, chromosome areas appeared to be involved in the control of lipid metabolism downstream of the liver. On LG2a, the plQTL impacting plasma corticosterone level before stress and glucose contents, is consistent with the fact that corticosterone is involved in the glucose metabolism.
The QTL results for metabolic traits published in chicken are very diverse. To our knowledge, no publications referenced a QTL for corticosterone response to a constraint on GGA2, although Buitenhuis *et al*. (2003) detected their strongest significant QTL for severe feather peaking on GGA2 in a chicken F2 population established from a cross between two commercial lines of laying hens differing in their propensity to feather pecking. Usually, the corticosterone response to a given stress is considered to be associated with feather pecking behavior (Buitenhuis *et al*., 2003). For blood glucose content, the strongest QTL published are spread out on GGA2 and GGA7 (Zhou *et al*., 2007- in a broiler-Fayoumi cross) and on GGA20 and GGA27 (Park *et al*., 2006- in a cross between low and high weights White Plymouth Rock). These last authors have already detected a QTL for cholesterol levels on GGA20 (corresponding to APL21) and a QTL for plasma triglyceride contents on GGA2.
Lastly, Nadaf *et al*. (2009) found a QTL for plasma glucose with a genome-wide level of significance on GGA13 (orthologous chromosome of APL14). However, no triglyceride or cholesterol assay had been made in this study, making any comparison difficult. All these results suggested that metabolic traits are regulated by a large number of QTL with small effects and it was therefore difficult to highlight areas of interests shared by ducks and chickens.

Finally, although multi-trait analyses done with correlated traits revealed new QTL regions, multi-trait analyses done with traits whose QTL were co-localized allowed us to better dissect the results previously obtained in single trait analyses. Whatever the traits, this approach evidenced the 2 main traits contributing to the pleiotropic effect. It is interesting to note that most significant likelihoods were obtained with a combination of two traits but never more than 2.
In single or in multiple trait analyses, the scarcity of QTL obtained for growth traits in relation to the number of recorded traits, yet very heritable (Marie-Etancelin *et al*., 2011), may be explained by a highly polygenic genetic determinism of growth and the absence of QTL with strong effects for this trait. In contrast, we hypothesized that the more complex traits such as glucose, cholesterol and triglyceride contents – particularly weakly heritable at the beginning of the overfeeding period (Marie-

Etancelin *et al*., 2011) - were depending on numerous small QTL. Even if we have obtained here a large number of QTL for liver and meat quality traits, we need to keep in mind that the high number of traits tested for QTL detection (a total of 1008 analyses computed in single trait approach and 366 computations in multi-traits analysis) generates an increased risk of false positive QTL.

CONCLUSION

This duck experimental design, the first one for overfed waterfowl, and the first with trait measurements in a sterile inter-specific hybrid, was based on a crossing of two Common ducks lines with progeny testing. It allowed the detection of 74 QTL significant at 5% threshold at chromosome-wide level, 22 of them reaching the 1% threshold. Most of them affected original traits such as the product quality traits of overfed ducks or their overfeeding ability. The multi-traits QTL approach was effective since it allowed to reveal new chromosomal areas not previously identified by the univariate approach. Moreover it suggested QTL with pleiotropic effects and sometimes reduced the QTL confidence interval.
Objectives of future work will be to increase the genetic map density with SNP markers in order to cover at least duck chromosomes 8, 10, 11 and 12 missing in the present analysis and to complete macro-chromosomes 5, 6 and 9. We will also plan to refine the map for QTL affecting fatty liver quality traits which should be studied as a priority, because of their economic importance and the complexity of their measures.

LITERATURE CITED

Atzmon, G., S., Blum, M., Feldman, A., Cahaner, U., Lavi, J., Hillel. 2008. QTL detected in a multigenerational resource chicken population. J. Hered. 99:528–38.

Aulchenko, Y.S., D.J., de Koning, C., Haley. 2007. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. Genetics 177:577-585.

Ambo, M., A.S., Moura, M.C., Ledur, L.F., Pinto, E.E., Baron, D.C., Ruy, K., Nones, R.L., Campos, C., Boschiero, D.W., Burt, L.L., Coutinho. 2009. Quantitative trait loci for performance traits in a broiler x layer cross. Anim. Genet. 40:200–8.

Buitenhuis, A.J., T.B., Rodenburg, Y.M., van Hierden, M., Siwek, S.J.B., Cornelissen, M.G.B., Nieuwland, R.P.M.A.,Crooijmans, M.A.M., Groenen, P., Koene, S.M., Korte, H., Bovenhuis, J.J., van der Poel. 2003. Mapping quantitative trait loci affecting feather pecking behavior and stress response in laying hens. Poult. Sci. 82(8):1215-1222.

Churchill, G., and R., Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963–971.

Crooks, L., G., Sahana, D.J., de Koning, M., Lund, O., Carlborg. 2009. Comparison of analyses of the QTLMAS XII common dataset. II: genome-wide association and fine mapping. BMC Proc. 3(1):S2.

Elsen, J.M., B. Mangin, B., Goffinet, D., Boichard, P., Le Roy. 1999. Alternative models for QTL detection in livestock. I General introduction. Genet. Sel. Evol. 31:213–224.

Elsen, J.M., 2001. Méthodes de détection : approche multiQTL. Séminaire, "QTL. De la détection à l'utilisation", Batz-sur-Mer, 24-26 septembre 2001, Séance 1 : Méthodes de détection", 11-21.

Filangi, O., C., Moreno, H., Gilbert, A., Legarra, P., Le Roy, J.M., Elsen. 2010. QTLMap, a Software for QTL Detection in Outbred Populations. Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany. Com. 0787.

Fillon, V., M. Vignoles, R. P. Crooijmans, M. A. Groenen, R. Zoorob, and A. Vignal. 2007. Fish mapping of 57 bac clones reveals strong conservation of synteny between galliformes and anseriformes. Anim Genet 38: 303-307.

Gao, Y., Z.Q., Du, W.H., Wei, X.J., Yu, X.M., Deng, C.G., Feng, J., Fei, J.D., Feng, N., Li, X.X., Hu. 2009. Mapping quantitative trait loci regulating chicken body composition traits. Anim. Genet. 40:952-954.

Gilbert, H., and P., Le Roy. 2003. Comparison of three multi-traits methods for QTL detection. Genet. Sel. Evol. 35:281–304.

Gilbert, H., P., Le Roy, C., Moreno, D., Robelin, J.M., Elsen. 2008. QTLMAP, a software for QTL detection in outbred population. Ann. Hum. Genet. 72(5):694.

Goffinet, B., P., Le Roy, D., Boichard, J.M., Elsen, B., Mangin. 1999. Alternative models for QTL detection in livestock. III. Heteroskedastic model and models corresponding to several distributions of the QTL effect. Genet. Sel. Evol. 31:341-350.

Green, P., K., Falls, S., Crooks. 1990. Documentation for CRIMAP*. In Washington University,* St Louis, WA*:* Washington University.

Hu, Z.L., C.A., Park, E.R., Fritz, J.M., Reecy. 2010. QTLdb: A Comprehensive Database Tool Building Bridges between Genotypes and Phenotypes. Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany. Invited paper 17.

Huang, Y., C.S., Haley, F., Wu, S., Hu, J., Hao, C., Wu, N., Li. 2007a. Genetic mapping of quantitative trait loci affecting carcass and meat quality traits in Beijing ducks (*Anas platyrhynchos*). Anim. Genet. 38:114–119.

Huang, Y., C.S., Haley, S., Hu, J., Hao, C., Wu, N., Li. 2007b. Detection of quantitative trait loci for body weights and conformation traits in Beijing ducks. Anim. Genet. 38:525-526.

Kileh-Wais, M., and J.M., Elsen. 2012. QTL detection from regression analysis of “generalized de-regressed proof” information. J. Anim. Breed. Genet. DOI: 10.1111/j.1439-0388.2011.00981.x

Knott, S.A., J.M., Elsen, C.S., Haley. 1996. Methods for multiple marker mapping of quantitative trait loci in half-sib populations. Theor. Appl. Genet. 93:71-80.

Lander, E.S., and D., Botstein. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199.

Lillehammer, M., B.J., Hayes, T.H.E., Meuwissen, M.E., Goddard. 2009. Gene by environment interactions for production traits in Australian dairy cattle. J. Dairy Sci. 92:4008-4017.

McRae, A.F., S.C., Bishop, G.A., Walling, A.D., Wilson, P.M., Visscher. 2005. Mapping of Multiple Quantitative Trait Loci for Growth and Carcass Traits in a Complex Commercial Sheep Pedigree. Anim. Sci. 80:135-141.

Marie-Etancelin, C., K., Feve, C., Genet, G., Cardinet, F., Vignoles, F., Pitel, A., Vignal. 2006. Microsatellite DNA markers for duck *(Anas platyrhynchos and Cairina moschata).* Symp. Scientific Cooperation in Agriculture between COA (Tainan) and INRA (France) Taiwan, 6-10 november 2006, 165-168.

Marie-Etancelin, C., H., Chapuis, J.M., Brun, C., Larzul, M.M., Richard, R., Rouvier. 2008. Genetics and selection of mule ducks in France: a review. World Poult. Sci. J. 64:187-207.

Marie-Etancelin, C., B., Basso, S., Davail, K., Gontier, X., Fernandez, Z.G., Vitezica, D., Bastianelli, E., Baéza, M.D., Bernadet, G., Guy, J.M., Brun, A., Legarra. 2011. Genetic parameters of product quality and hepatic metabolism in fattened mule ducks. J. Anim. Sci. 89(3):669-679.

Maruyama, K., M.K., Akbar, C.M., Turk. 1999. Growth pattern and carcase development in male ducks selected for growth rate. Br. Poult. Sci. 40:233–239.

Maruyama, K., B., Vinyard, M.K., Akbar, D.J., Shafer, C.M., Turk. 2001. Growth curve analyses in selected duck lines. Br. Poult. Sci. 42:574–582.

Nadaf, J., H., Gilbert, F., Pitel, C., Berri, K., Feve, C., Beaumont, M., Duclos, A., Vignal, T.E., Porter, J., Simon, S.E., Aggrey, L.A., Cogburn, E., Le Bihan-Duval. 2007. Identification of QTL controlling meat quality traits in an F2 cross between two chicken lines selected for either low or high growth rate. BMC Genomics 8:155-162.

Nadaf, J., F., Pitel, H., M.J., Duclos, F., Vignoles, C., Beaumont, A., Vignal, T.E., Porter, L.A., Cogburn, S.E., Aggrey, J., Simon, E., Le Bihan-Duval. 2009. QTL for several metabolic traits map to loci controlling growth and body composition in an F2 intercross between high- and low-growth chicken lines. Physiol. Genomics 38:241–249.

Park, H.B., L., Jacobsson, P., Wahlberg, P.B., Siegel, L., Andersson. 2006. QTL analysis of body composition and metabolic traits in an intercross between lines divergently selected for growth. Physiol. Genomics 25:216-223.

Roldan, D.L., A.M., Dodero, F., Bidinost, H.R., Taddeo, D., Allain, M.A., Poli, J.M., Elsen. 2010. Merino sheep: a further look at quantitative trait loci for wool production. Animal 4:1330-1340.

Rouvier, R., G., Guy, D., Rousselot-Paillet, B., Poujardieu. 1994. Genetic parameters from factorial cross breeding in two duck strains (Anas platyrhyncos) Brown Tsaiya and Pekin, for growth and fatty liver traits. Br. Poult. Sci. 35:509-517.

SAS. 2002. SAS version 9.1.3 SAS Inst. Inc. Cary, NC.

Siwek, M., S.J., Cornelissen, A.J., Buitenhuis, M.G., Nieuwland, H., Bovenhuis, R.P., Crooijmans, M.A., Groenen, H.K., Parmentier, J.J., van der Poel. 2004. Quantitative trait loci for body weight in layers differ from quantitative trait loci specific for antibody responses to sheep red blood cells. Poult. Sci. 83:853–861.

Skinner, B.M., L.B.W., Robertson, H.G., Tempest, E.J., Langley, D., Ioannou, K.E., Fowler, R.P.M.A., Crooijmans, A.D., Hall, D.K., Griffin, M., Völker. 2009. Comparative genomics in chicken and Pekin duck using FISH mapping and microarray analysis. BMC Genomics 10:357-367.

Théron, L., X., Fernandez, N., Marty-Gasset, C., Pichereaux, M., Rossignol, C., Chambon, D., Viala, T., Astruc, C., Molette. 2011. Identification by proteomic analysis of early *post mortem* markers involved in the variability in fat loss during cooking of mule duck 'foie gras'. J. Agric. Food Chem. 59:12617-12628.

Vitezica, Z.G., C., Marie-Etancelin, M.D., Bernadet, X., Fernandez, C., Robert-Granie. 2010. Comparison of nonlinear and spline regression models for describing mule duck growth curves. Poult. Sci. 89:1778-1784.

Wahlberg, P., O., Carlborg, M., Foglio, X., Tortoir, A.C., Syvänen, M., Lathrop, I.G., Gut, P.B., Siegel, L., Andersson. 2009. Genetic analysis of an F(2) intercross between two chicken lines divergently selected for body-weight. BMC Genomics 10:248-260.

Weller, J.L., Y., Kashi, M., Soller. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. J. Dairy Sci. 73:2525-2537.

Wright, D., S., Kerje, K., Lundström, J., Babol, K., Schütz, P., Jensen, L., Andersson. 2006. Quantitative trait loci analysis of egg and meat production traits in a red junglefowl x white leghorn cross. Anim. Genet. 37:529-534.

Ye, M.H., J.L., Chen, G.P., Zhao, M.Q., Zheng, J., Wen. 2010. Associations of A-FABP and H-FABP markers with the content of intramuscular fat in Beijing-you Chicken. Anim. Biotech. 21:14–24.

Zhou, H., C.M., Evock-Clover, J.P., McMurtry, C.M., Ashwell, S.J., Lamont. 2007. Genome-wide linkage analysis to identify chromosomal regions affecting phenotypic traits in the chicken. IV. Metabolic traits. Poult. Sci. 86:267-276.

Appendix 1: microsatellite markers sequences and accession number

Table 1a: *Anas Plathyrynchos*

Accession	Locus	Sequence Primer 1	Sequence Primer 2	Microsatellite
FR750405	CAM001	AAGGTCTTTCCAGCCTAAATG	CATCAGCATCAAAGCAGAAC	$(CCTT)^{15}$
FR750406	CAM003	ACAGCCTAAGGTGTTTTTCC	TTTTCCTTTGGTCAGTGTGG	$(AC)^{11}$
FR750407	CAM004	AGCAGCTGAGGACTGTATGC	CAAAGGCACTTTTTGTTTGC	$(AC)^{7}$
FR750408	CAM005	TGAATAGACATTGTACTGCTGAGG	TGAAAGTGAGGACAGAAATCC	$(GT)^{4}$
FR750409	CAM006	GGTAATTCTGGCTTTTGTCC	TGCGAAGAATTAGACTTGAGC	$(GT)^{6}$
FR750410	CAM007	TGTCAGGTGTTGTGAGAAAGG	TGGGAGTTTCAGAAGTTTATGG	$(TG)^{3}TA(TG)^{4}$
FR750411	CAM010	TTGAATTTATCAACTGCCTTAGAA	TCATTGGACTGAGGAAGGTC	$(AC)^{13}$
FR750412	CAM011	AGTATCATTTGCTACTGCACGC	CCAGGGCTTCAAATGGAGAAT	$(AC)^{11}$
FR750413	CAM012	TTGTTTTGGCCTGTCTGTATGT	TGCAAGAGATATGGTGCACAAA	$(CA)^{9}$
FR750414	CAM013	GGAATGATGTTTACAGTTCATAGG	TGCTTTGAAGCACTACATGG	$(GTTT)^{4}$
FR750415	CAM014	GCAGAGAAAATGAAATCCTACC	AGCTTAATTCTGCAAGCCTA	$(TG)^{11}$
FR750416	CAM021	GGGAAGACACAAACAGCCTT	TTTGTCCGCTGCATCCTTG	$(CAGA)^{3}$
FR750417	CAM022	TGGTCACTGCCTTACCATTT	GTTGACTGAGAACATGGCATT	$(AAAC)^{3}$
FR750418	CAM023	GCAAAGACTGGTACTTGGAACA	TCCCTTGCCACTTCTCCTATT	$(TAAA)^{5}$
FR750419	CAM026	GAGGTAGGTCTGCAAAGAAACA	GGGGCGATTGAATGAGTGTAT	$(GGA)^{8}$
FR750420	CAM027	GGGGGAAAGAGAATACGCATAT	CACCAGACCACAAGGACATCT	$(TGAA)^{3}$
FR750421	CAM028	ACAGATGTCAGTACTGCCGG	TTAGGGGCAAGTTAGGAGCC	$(GT)^{3}GA(GT)^{6}$
FR750422	CAM029	ATAGGGCCATGGATTGTTTTGC	CCGGTGGATTTCATGCTATTGT	$(GTT)^{4}$
FR750423	CAM030	GTTGACAGATGCAAGCTGTAAC	GCACGTGAGATGATTCTTTGG	$(GACA)^{4}$
FR750424	CAM031	TTGCTGTTGTAGCTGTGAGATT	GCTTTGCTGCTGTTTATTCAGT	$(AAAT)^{3}$
FR750425	CAM034	AGCGTGTTTCTTTAAGCTCTGT	CAGTGTAATTCCAGCCTGCTG	$(TG)^{6}(GT)^{3}(T)^{4}(GT)^{5}$
FR750426	CAM037	TGGGCTCAGGATTTGAAGAAT	ACCAGTTTGCATGATAGGTGT	$(CA)^{10}$
FR750427	CAM040	CTTCTTGCTGCAGAATGAAT	TCACCAAATTTGAGCTTTGG	$(GT)^{5}AT(GT)^{10}$
FR750428	CAM044	GCATGGGAAGCCAATGTTTT	CGCATGATTTCTGCTTCTCTAT	$(GT)^{13}$
FR750429	CAM047	AGTGACTTGAAGGTGAAGGT	CTCCCCAAACCCCATATAAG	$(TC)^{5}$
FR750430	CAM052	GACTGGCAATCCAAAGCA	CACAGCTCCAGAAACAGTCC	$(GT)^{10}$
FR750431	CAM053	GCCATGGCTCCACAGCAG	ATCTCCCACATTGGGTCACG	$(GT)^{15}$
FR750432	CAM062	TTACTTTGTGTGGTCGTCTT	CTTCCACCTATATGTGAGCTG	$(AC)^{5}AT(AC)^{4}(AT)^{5}$
FR750433	CAM065	CCCATACATGTATATGCTGTGC	CTATGCAATGGGCTTCGTAGT	$(AAGG)^{22}AAAG(AAGG)^{3}$
FR750434	CAM071	GCAATATCCCTGGCAGTACC	ACCAGAAATGAGAAGCAGACC	$(GT)^{16}$
FR750435	CAM073	AGGATGCAGTCTACATTTGCA	CCAGTTTCACCAGTTGAGAAG	$(CA)^{14}$
FR750436	CAM076	CTTCCACAAGTGGCAGTGAC	GGAAGAGGGGATTAGACAAAA	$(GT)^{4}GA(GT)^{12}$
FR750437	CAM080	TCCAAACTTAAGCAATTTTTC	CGGTGACCGACCTACCTA	$(GGAA)^{6}(GA)^{11}$
FR750438	CAM081	CACTTAACACTCCAGACCGATA	TTGGGAATTTGTCTCCTTGT	$(AC)^{10}$
FR750439	CAM082	TTTTCATGCTGTGTTTGAGAA	GGTTCTGGGAAGAAAATCAA	$(TC)^{4}$
FR750440	CAM085	TCCTGAGAAGATGGAAATCTG	TCCTTGTCAGCAATTGAGAA	$(CA)^{8}$
FR750441	CAM087	TCCACTGTGCATTTGAGTAAA	CACTTAAAGCTCAGGCAAGG	$(CA)^{21}$
FR750442	CAM093	TGTATCGACTCATTTTGCTCA	GCCTGACTCTTTTTATTTTCCA	$(CTTC)^{6}(CTTT)^{8}$
FR750443	CAM101	GTCAAGTCAGTGCCTCACAA	GTATCTCTTCCAGGCCAACC	$(TTTTG)^{3}(TTTG)^{2}$
FR750444	CAM103	CGGGAAGCTGTAGATAAGCA	TGGCTGGAGAGAGCACCT	$(T)^{6}A(T)^{5}$
FR750445	CAM106	AACACCACCACCAAAGAAAT	TAAACCGGTGGAATCTTTCT	$(TAT)^{4}$
FR750446	CAM111	GTGGCTAGTGCTGATGAATG	TGAACCACTTTCTCACAGGA	$(GT)^{13}$
FR750447	CAM113	TGCTCGGAGGACTGTTAGAT	GCTGCATGCATCTTCTCTTA	$(GT)^{12}$
FR750448	CAM119	CAATGACCAGAAAAGCAACA	AAGAACTTCCCCCTCTTCTG	$(GT)^{10}$
FR750449	CAM121	ACATCTGGAAAAAGCCAGAG	CATTTCAGGTCCTGTTTCCT	$(AG)^{9}$
FR750450	CAM127	TGTAGACTGAAGTGCAACCAA	AGGAATCATGCTGGAAAAAT	$(AT)^{7}$
FR750451	CAM128	CAGCCCTAGAGCTCAAAAAG	GCACTTAAGAGCATGCACAC	$(CA)^{15}$
FR750452	CAM130	TTGCTCCCTTTTGTAGTTCC	GTGGTGGATTGCACTGTAAG	$(ATTT)^{5}$
FR750453	CAM131	CAACGTTCCTATGTCTCACG	AGCTACAGGGTAACTGCACA	$(GT)^{10}$
FR750454	CAM133	CTGTACTTACCCCTGCCATC	CTTTCCTAACCTGCGTTTTC	$(CA)^{10}TA(CA)^{5}$
FR750455	CAM134	ACCCAAGATTTCAGTCTCCA	CAGTATCCAGAGGTGGAAGC	$(CT)^{12}$
FR750456	CAM135	CCATTCAAAGACTCAATGGTT	TGTACACTCACATAGAAGAA	$(TTCC)^{3}CTCC(TTCC)^{5}$ & $TTCT(TTCC)^{7}$
FR750457	CAM138	AGGTTCAAAAGGCAGTTGTC	GCTGAAGCTTGGAGAATACC	$(AAAC)^{2}GAAC(AAAC)^{3}$
FR750458	CAM150	TGCACGTGGTATTTTACACA	GATGCTCTGCTCATACCTGA	$(AG)^{5}$
FR750459	CAM151	GTATGCTTTGTGGTCCATGA	TGCTTGAGATGAAACACTCC	$(GA)^{3}A(GA)^{3}$
FR750460	CAM153	GAGATTTCCAGCTTGCATTT	GGAAGCAGAAGTCTGAAAAGA	$(TG)^{6}T(TG)^{3}$
FR750461	CAM155	GGCAGCAAATTAAGGTTGAC	GACAAGGCTATCCTGGCTTA	$(TG)^{14}$
FR750462	CAM157	ATAACGGGAATCAGCTCTTG	GGGGATGGTTCACATTTTTA	$(CT)^{6}TTTCC(CT)^{3}$
FR750463	CAM158	TGCACTCAGGTTTTCTTTGA	ATCCTTTGCCTGCTTGATAA	$(TC)^{12}$
FR750464	CAM161	TAAAGCAGAGCCAGGCAGT	CACTGAGACACTTTTGTGCAG	$(GGAA)^{16}$

Table 1a: *Caïrina Moschata*

Accession	Locus	Sequence 1	Sequence 2	Microsatellite
FR750465	CAM002	TTTCATTCCCTCCCCATC	CGGGTTTATAATATCATTTTGAGG	$(AC)^{8}$
FR750466	CAM009	ACCTGGCTACAAAGCATCAC	AATAGATGGACAAAAGATGCAA	$(GGAA)^{13}$
FR750467	CAM015	CCACATGTCACTTTGTGAACAC	TGGAACGTGTCAGGTGGAA	$(CA)^{14}$
FR750468	CAM018	TCTCCTACTTTCTGCTTACATGG	AATTTTGCCAGGCACCTATG	$(GT)^{12}$
FR750469	CAM019	TGGTCACCCTTTTCCATCTT	CATTCTGTCTCGTTTACTTTCTCC	$(CA)^{10}$
FR750470	CAM020	AGCAACTGAGCTACCAGAAAAA	GTCTCCACCAGCCTATTTCTG	$(CA)^{9}$
FR750471	CAM025	TGAGGGTTGACTTTTAAGGCTT	TGTGGCACTAGACTTTGCAAA	$(AATT)^{3}$
FR750472	CAM032	GTGTTGACTAGCTCTCCGTTTA	GACGCTCACGAACACTTAGG	$(ATTT)^{2}$
FR750473	CAM033	TCCAACCCAAAGGTCTCTTCA	GCTTCATAGAGGCCCAGGTT	$(AAAAAAG)^{3}$
FR750474	CAM035	GCCCCACTAGGGTTGGTAAG	CTTCTGGTCGTCTCCTGGTA	$(GT)^{10}$
FR750475	CAM036	GGTCAGAGTCAAATACGGAGTC	GCACTCCACCCACTTGTTT	$(CA)^{5}AA(CA)^{7}$
FR750476	CAM039	CTGAGAGGCCCCTGCTGT	CACAGCTCCTCCCCAGGT	$(CA)^{3}TG(CA)^{5}CGCA$ &$(CCCA)^{3}CA(CCCA)^{3}(CA)^{8}$
FR750477	CAM041	AACCAAGAAGAATTTCAGCAAA	ATTGCAGCAGGAACTTCATT	$(AAGG)^{7}$
FR750478	CAM042	CTGTGTCTGTGTGAAGCCTA	CCCCATGGAAAGAACTGACAAA	$(TG)^{10}$
FR750479	CAM043	TTACCAGCTCTTGGAAAAT	AAGCATTTGCTTCTTGTGAC	$(ATT)^{3}$
FR750480	CAM045	CTTACCCCTGGAAACTCCTT	TCACGAGAAAAGAAATGCAG	$(GGAA)^{14}$
FR750481	CAM046	GCTGCTTTCTGTTCATTTAGCT	TCTGACCACATGCTTTCTCAT	$(AACATTTGAAG)^{2}$
FR750482	CAM048	GAGGCAGTTGGTTAGTGGTT	CAAATGCCCACAATATAGCA	$(GCCT)^{4}(TCCT)^{6}$ & $(TACT)^{2}(TCCT)^{4}$
FR750483	CAM049	TCCTTGCTCTCCTGCTTCTG	ACCTCACAGTAGCAGAAAAC	$(GATTTGAG)^{2}$
FR750484	CAM054	CTCAGTAAACCTGGTACCTT	CTGTGCTAAGTAACAGCTCAC	$(TG)^{10}$
FR750485	CAM055	TCATTAAAACAAAAATAAAATGCAG	CAAGCAAATGTTCGTGTGTT	$(TG)^{6}TA(TG)^{11}$
FR750486	CAM057	TTCTCAGTTTGCTTGCTCTG	CGTGGACGTATAAACTGCTG	$(TCTCT)^{2}$
FR750487	CAM058	TTCTGTATGGAAAGCAAGATTAAC	ATCGAGCATGTGGAGATTTT	$(TTCC)^{10}$
FR750488	CAM061	CAACTCACAGGTTGAGATAAGG	ACGTGGTCAATCATCTCTCATT	$(AAT)^{5}$
FR750489	CAM063	CGGTTTGCTGTATTTTCACT	GGGAGAGTTGTAAAATGAAGC	$(CT)^{11}$
FR750490	CAM064	CAGGTTGCAAGTCAACAGAC	GTCCTGGAAGTGAAGAGAAGT	$(CA)^{10}$
FR750491	CAM067	GCAAACCGATAACAATGGGA	TAGAGCCTCATGGTACAGGC	$(AAGG)^{3}$
FR750492	CAM068	AGAGCAGAGATTGGGGCAAT	TTTGTCAGCCCTTGGGACA	$(TG)^{9}$
FR750493	CAM070	TGGGTAGTCAATTTGCCTTAGT	CTACAGCTGGTGGGACTATAG	$(TG)^{9}$
FR750494	CAM072	TTTACACAGCTAGAAGGCTTGA	TGGCTCCACATTAACAGATG	$(ATTT)^{5}$
FR750495	CAM088	TGCGTCAACAGCTACATACA	ATAGCAGCCATTCCCATAGA	$(TTG)^{5}$
FR750496	CAM092	GTGCAGACTTGCTGAAAGAA	TGCAGTCTGTAGCTTGTGGT	$(GTT)^{5}$
FR750497	CAM096	CAGCATGAGATTGTTTAGAGGA	AAGCTTTGTGCAATTGATGA	$(AG)^{15}(AAAG)^{20}$
FR750498	CAM098	TAACAGCAGCAAGTCGAAAG	CTTCGGTGCTCCTAGAGGT	$(CA)^{7}CC(CA)^{3}CC(CA)^{6}$
FR750499	CAM099	TCATTTGGGAAGGTCTCAAC	CTCGGAACTCAATGTTAGCC	$(GA)^{14}$
FR750500	CAM100	AAATGTTATTCTCCCGTTGG	GAACAGACCTCTGCCAAAAG	$(TTTTA)^{4}$
FR750501	CAM104	GGAACCTACCGCAATTTATG	TCTCCCTGTTGACAGGTAATC	$(AAC)^{3}$
FR750502	CAM107	GCCCAAATTTAAAGGTTCTG	GAAAGAGGGAGAGTGTGAGG	$(CTTT)^{3}(CT)^{13}$
FR750503	CAM108	CCAAAAGGTAACCAAGGAAC	TGTTGACCAGGAAAATATGG	$(CCTT)^{13}$
FR750504	CAM110	TGTGTGTGACAAACCTCTGA	TGTTTTCCTTTTCAAGCACA	$(AC)^{13}$
FR750505	CAM117	ATGATGCATGAAGCAGAGTG	CCTAAGGTTTGCACAGGTTT	$(GT)^{10}$
FR750506	CAM118	GTGCAGACTTGCTGAAAGAA	TGCAGTCTGTAGCTTGTGGT	$(TG)^{6}$
FR750507	CAM120	TCTCCCAAAGAAAAATCTGG	AGAGGGAGAAAGAGTGAGAGC	$(CCTT)^{19}$
FR750508	CAM124	TCAACTTTGCAAACATAGGG	TTAGTCCCTTTCCCAACAGA	$(AG)^{20}(AAAG)^{5}$
FR750509	CAM126	AATGCTCCTTCTCTGAGCTG	GGGTATTCCATGCTGCTATT	$(AG)^{5}$
FR750510	CAM129	GAAGCTGTTTGTGTTTGCAC	TGTGTAGCTGCATGTTTCCT	$(CCTT)^{22}$
FR750511	CAM132	CAATCTGCTGCTTCTTCTGA	AGCCAGCTGCAAAGAGAATA	$(TTCC)^{21}$
FR750512	CAM136	GAAACAAAGGGCGACATAAT	ATCTGATGGGGTTTTGACAC	$(GA)^{12}$
FR750513	CAM137	CCTTGCTGAGGTCTAAATCC	TGTAAAAGCCATTGTGGTCA	$(TG)^{11}$
FR750514	CAM139	ATGGCAATGGTTGTCCTACT	TTCTGAACTTGCATTGTTTGA	$(TTCC)^{20}$
FR750515	CAM141	TAGCCTAAGAAGAGCCCAGA	TTAAATGCTTGGCTGAACTG	$(AG)^{8}$
FR750516	CAM142	TTCATCTTGAGGTACCTTGG	CATGTTCCTGAGAAGCACAG	$(GGAGA)^{2}$
FR750517	CAM143	CTTTAACACCAGTGGGTTCC	CCCTTGACAATTAGCCTTTG	$(ATT)^{3}G(ATT)^{2}$
FR750518	CAM144	AATGCCTTGTGGAAACTGAT	TTTTATTGGGCAACCTAACC	$(GTTTT)^{6}$
FR750519	CAM146	TGGAAGCAAGATATGTGAGC	TGCAGGATTAATGTCAGCAT	$(TTCC)^{27}$
FR750520	CAM147	TTGTAGCTCCCTGCACAGTA	GGTTCTCTGAGTTTCCTTGC	$(CTG)^{4}$
FR750521	CAM149	AAGCCATAGATTCCTCAGATG	ATATGCTGTGTGCTGGAAAA	$(AAGG)^{29}$
FR750522	CAM152	GAATTAAAATCCCTTTGCTTCT	AAGACATGCAGGCTATTTCA	$(CTCCT)^{8}$
FR750523	CAM154	CGATGAATGCAGTCTCTGAC	AGGTACATCTCACAAGCAACC	$(AT)^{6}$

Appendix 2: Tables representing the correlation coefficients between the different traits studied classified by their physiological functions.

Table 1a: growth traits

	BWG 12-28	BWG 12-42	BWG 28-42	BWG 12-70	BWG 28-70	BWG 42-70	BW12	BW28	BW42	BW70
BWG12-28	1	**0.87**	0.50	0.58	0.19	-0.07	**0.62**	**0.92**	**0.92**	**0.70**
BWG12-42		1	**0.86**	**0.72**	0.42	0.01	0.38	**0.72**	**0.93**	**0.75**
BWG28-42			1	**0.66**	0.54	0.08	0.01	0.31	**0.68**	0.58
BWG12-70				1	**0.90**	**0.69**	0.26	0.49	**0.67**	**0.96**
BWG28-70					1	**0.87**	-0.01	0.11	0.33	**0.79**
BWG42-70						1	-0.03	-0.05	0.00	**0.60**
BW12							1	**0.87**	**0.68**	0.53
BW28								1	**0.91**	**0.70**
BW42									1	**0.80**
BW70										1

Table 1b: traits related to stress

	cortL	cortH	deltaC
cortL	1	0.42	0.23
cortH		1	**0.98**
deltaC			1

Table 1c: metabolic traits

	CHO 2^{nd}M	CHO 10^{th}M	CHO 20^{th}M	TG 2^{nd}M	TG 10^{th}M	TG 20^{th}M	GLU 2^{nd}M	GLU 10^{th}M	GLU 20^{th}M
CHO 2^{nd}M	1	0.36	0.37	0.22	0.19	0.00	0.11	0.02	0.09
CHO 10^{th}M		1	0.39	0.09	0.27	0.16	-0.09	0.15	0.25
CHO 20^{th}M			1	0.28	0.22	0.31	0.02	0.13	0.34
TG 2^{nd}M				1	0.37	0.25	0.30	0.19	0.18
TG 10^{th}M					1	0.38	0.10	0.19	0.32
TG 20^{th}M						1	0.16	0.12	0.37
GLU 2^{nd}M							1	0.12	0.06
GLU 10^{th}M								1	0.31
GLU 20^{th}M									1

Table 1d: carcass traits

	FLW	AFW	TSW	pmMW	pmSFW	BWbeg	BWend	CW
FLW	1	0.13	-0.12	-0.12	0.07	-0.05	0.34	0.35
AFW		1	0.44	0.20	0.58	0.49	0.52	0.56
TSW			1	0.48	0.52	**0.70**	**0.68**	**0.67**
pmMW				1	0.28	**0.69**	0.57	0.54
pmSFW					1	0.51	0.57	0.59
BWbeg						1	0.81	**0.80**
BWend							1	**0.97**
CW								1

Table 1e: product quality traits

	Fmax	Energy	MWC	MLipC	Mb*	ML*	Ma*	MpH20	MpHu	MvacL	MCookL	MR	Lb*	LL*	La*	LProtC	LLipC	LColC
Fmax	1	**0.82**	-0.01	-0.20	0.25	-0.47	0.49	-0.35	0.20	0.21	-0.18	0.22	-0.21	0.04	0.13	0.38	0.07	-0.15
Energy		1	-0.07	-0.10	0.29	-0.38	0.47	-0.22	0.12	0.13	-0.16	0.19	-0.21	0.04	0.11	-0.35	0.06	-0.11
MWC			1	**-0.79**	-0.47	0.00	-0.30	0.23	0.02	0.39	0.12	-0.12	-0.11	-0.09	0.08	0.13	-0.03	0.15
MLipC				1	0.36	0.41	-0.04	-0.25	-0.14	-0.38	0.18	0.09	0.25	0.20	-0.22	0.02	0.11	-0.14
Mb*					1	-0.29	**0.68**	-0.12	0.26	-0.10	-0.38	0.08	-0.24	-0.02	0.15	-0.37	-0.08	-0.06
ML*						1	**-0.84**	-0.07	-0.45	-0.14	0.65	-0.11	0.40	0.23	-0.38	0.47	0.15	0.00
Ma*							1	-0.08	0.43	0.01	**-0.66**	0.16	-0.43	-0.12	0.33	-0.55	-0.12	-0.05
MpH20								1	-0.08	-0.02	-0.05	-0.10	-0.17	-0.17	0.10	0.06	-0.08	0.10
MpHu									1	0.01	-0.49	0.03	-0.30	-0.10	0.15	-0.32	-0.04	0.03
MvacL										1	-0.03	0.07	-0.18	0.04	0.06	-0.17	0.08	-0.06
MCook											1	-0.07	0.39	0.17	-0.27	0.40	0.12	0.01
MR												1	-0.15	0.38	-0.35	**-0.75**	**0.79**	-0.38
Lb*													1	-0.05	-0.13	0.43	0.05	-0.04
LL*														1	**-0.79**	-0.37	0.46	-0.41
La*															1	0.18	-0.46	0.32
LProtC																1	**-0.63**	0.42
LLipC																	1	-0.42
LColC																		1

Appendix 3: QTL detection in single trait analysis with *P*-value ranging from 1% to 5% chromosome-wide level

Growth traits

LG	Traits[1)]	Flanking markers[2)]	Location[3)] (cM)	LRT_x[4)]	*P*-value[5)]	Effect of QTL substitution[6)]	Confidence interval (cM)
1a	Ti	APT015	36	20.55	0.050	0.49	20 - 43
2a	BW70	APT003-CAM096	64	20.38	0.031	0.58	35 - 75
2a	BWG12-70	APT003	63	19.86	0.036	0.57	33 - 75
2a	BWG12-28	APT009-APT003	61	17.74	0.050	0.49	37 - 75
2a	BWG12-42	APT003-CAM096	66	17.18	0.050	0.48	35 - 75
2a	BWG28-42	APT003-CAM096	66	16.26	0.050	0.5	31 - 75
2a	BWG28-70	APT003	63	16.95	0.050	0.56	32 - 75
2a	BW42	CAUD089-APT009	59	17.69	0.050	0.46	37 - 75
3	BWG12-42	CAUD084-CAUD045	62	23.25	0.018	0.31	41 - 74
7	BWG42-70	APH013	0	18.32	0.024	0.75	0 - 13
14	BW12	CAUD137	24	17.56	0.033	0.32	0 - 24
20	BW_A	CAUD099	0	15.57	0.016	0.55	0 - 16
28	BW70	APH019	5	13.72	0.050	0.61	0 - 27

1) See Table 1 for the definitions of the traits.
2) Flanking markers of the most probable QTL position.
3) Most probable QTL position.
4) Maximum likelihood ratio for x locus.
5) Level of significance of P-value at chromosome-wide.
6) QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

Metabolic and body weight traits during the overfeeding phase

LG	Traits[1)]	Flanking markers[2)]	Location[3)] (cM)	LRT_x[4)]	*P*-value[5)]	Effect of QTL substitution[6)]	Confidence interval (cM)
1b	DFI	CAUD039-CAM029	8	19.70	0.032	0.21	0 - 16
1b	FC 10^{th}M	CAM093	37	23.60	0.013	0.25	26 - 39
2a	CortL	CAM071-CAUD065	10	19.91	0.035	0.24	0 - 31
2a	GLU 2^{nd}M	CAM071	0	20.84	0.028	0.23	0 - 18
2c	CHO 2^{nd}M	CAM005	1	15.41	0.030	0.13	0 - 1
2c	FC 2^{nd}M	CAM005	1	13.57	0.050	0.17	0 - 1
3	CHO 10^{th}M	CAUD084-CAUD045	71	22.35	0.024	0.48	47 - 82
6	GLU 2^{nd}M	CAUD064-CAUD026	1	13.45	0.044	0.24	0 - 20
7	BWbeg	APH013	0	19.29	0.018	0.29	0 - 25
7	FO 2^{nd}M	APH013	0	18.11	0.025	0.21	0 - 26
9	OWG	CAUD038	20	17.12	0.044	0.16	9 - 20
14	CHO 10^{th}M	CAUD013-CAUD137	7	17.77	0.031	0.64	0 - 24
14	CortL	CAUD137	24	15.92	0.050	0.11	0 - 24
14	TG 2^{nd}M	CAUD013-CAUD137	7	17.08	0.038	0.14	0 - 22
21	CHO 10^{th}M	CAUD037	2	18.59	0.014	0.64	0 - 2
21	FO 10^{th}M	CAM004	0	18.60	0.014	0.25	0 - 2
21	BWend	CAM004	0	19.60	0.010	0.14	0 - 2
21	TG 10^{th}M	CAUD037	2	15.19	0.045	0.30	0 - 2
28	GLU 2OthM	APH019-CAUD040	26	15.73	0.030	0.29	0 - 27
28	BWbeg	APH019-CAUD040	14	13.60	0.050	0.26	0 - 27

1) See Table 1 for the definitions of the traits.
2) Flanking markers of the most probable QTL position.
3) Most probable QTL position.
4) Maximum likelihood ratio for x locus.
5) Level of significance of P-value at chromosome-wide.
6) QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

The overfeeding ability and product quality traits

LG	Traits[1)]	Flanking markers[2)]	Location[3)] (cM)	LRT_x[4)]	*P*-value[5)]	Effect of QTL substitution[6)]	Confidence interval (cM)
1a	Lb*	APT008-CAUD058	55	21.77	0.037	0.15	50 - 89
1a	MLipC	CAUD095	0	20.35	0.050	0.08	0 - 5
2a	La*	CAM071-CAUD065	2	19.64	0.038	0.28	0 - 27
2a	Mb*	CAM071-CAUD065	6	19.56	0.039	0.24	0 - 27
2a	pmMW	APT002-CAUD070	40	21.27	0.025	0.16	37 - 49
2c	Lb*	CAM005	1	14.53	0.029	0.15	0 - 1
2c	MR	APH012	0	16.60	0.018	0.15	0 - 1
2c	LColC	APH012	0	14.28	0.044	0.15	0 - 1
2c	MpHu	CAM005	1	14.92	0.037	0.10	0 - 1
3	Lb*	AMU060-CAM124	2	21.26	0.033	0.11	0 - 17
6	MpH20	CAUD064-CAUD026	15	13.31	0.046	0.24	0 - 25
7	Ma*	AMU103	31	15.34	0.050	0.02	20 - 31
7	MpH20	APH013	0	18.56	0.022	0.22	0 - 13
9	LLipC	CAUD088	0	21.43	0.011	0.21	0 - 20
9	MR	CAUD088-AMU068	5	17.30	0.042	0.18	0 - 20
14	LL*	CAUD013	0	18.18	0.026	0.25	0 - 10
21	LColC	AMU111	1	16.48	0.030	0.16	0 - 2
21	Fmax	CAUD037	2	15.25	0.045	0.01	0 - 2
Z	MpHu	CAM113	15	20.44	0.019	0.01	9 - 15

[1)] See Table 1 for the definitions of the traits.
[2)] Flanking markers of the most probable QTL position.
[3)] Most probable QTL position.
[4)] Maximum likelihood ratio for x locus.
[5)] Level of significance of P-value at chromosome-wide.
[6)] QTL effect in phenotypic deviation units (SD), and estimated as $\alpha = \frac{1}{SD} \times \frac{1}{n} \sum_{i=1}^{n} |\alpha_i|$.

4.2 QTLs significatifs à 5% au niveau du chromosome

Dans l'article 2, seuls les QTLs significatifs au seuil de 1% au niveau du chromosome sont discutés. Néanmoins, 52 autres QTLs significatifs entre 5% et 1% au niveau du chromosome, obtenus via l'approche uni-caractère et listés dans l'appendix 3 de l'article 2 seront discutés ici.

Pour garder une cohérence dans les notations, les noms des caractères sont toujours en anglais. Aussi, la nomenclature des groupes de liaison et des marqueurs est homogénéisée avec celle de l'article 2.

4.2.1 Les caractères de croissance

Treize QTLs de croissance, significatifs à 5% au niveau du chromosome ont été détectés (appendix 3 – tableau 1). Ces QTLs sont répartis sur 6 groupes de liaisons, mais le GL2a regroupe à lui seul 7 de ces QTLs. Chez la poule, le chromosome homologue Gallus Gallus 2 est connu pour être impliqué dans la variabilité des caractères de croissance. Par exemple Zhou *et al*. (2006) ont mis en évidence plusieurs QTLs pour des caractères de croissance (poids à 42 jours d'âge, ou gain de poids entre 14 et 28 jours d'âge ou entre 28 et 42 jours d'âge) sur le chromosome GGA2. Sur ce même chromosome Ou *et al*. (2009) ont identifié un QTL impliqué dans la variabilité du poids à 70 jours d'âge.

Contrairement aux analyses uni-caractères, aucun QTL n'est mis en évidence sur le groupe de liaison 2a avec les analyses multi-caractères. Sachant que seuls les QTLs significatifs au seuil de 1% ont été retenus dans le cadre de ces dernières analyses, on peut conclure que le regroupement des caractères (quelque soit le critère de regroupement) n'a aucun apport en termes de puissance par rapport aux analyses uni-caractères.

4.2.2 Les caractères liés au métabolisme durant le gavage et les caractères liés au stress (taux de corticostérone avant et après stress)

Onze QTLs significatifs à 5% au niveau du chromosome ont été détectés pour des caractères liés au métabolisme durant le gavage (appendix 3 – tableau 2). Il s'agit des taux de cholestérol, de glucose et de triglycéride dans le sang et le taux basal de corticostérone. Des QTLs ont été identifiés pour tous ces caractères chez la poule. En effet, Zhou *et al*. (2007) et Nadaf *et al*. (2009) ont respectivement mis en évidence des zones chromosomiques impliquées dans la variation du taux de glucose dans le sang sur les chromosomes GGA2 et GGA6. Les QTLs obtenus pour le taux de triglycérides ne semblent pas être connus chez la poule et seul un QTL pour le taux de cholestérol sur le groupe de liaison 3 est déjà identifié chez la poule par Park *et al*. (2006). Aucun QTL pour le taux basal de la corticostérone n'est identifié pour la poule sur les chromosomes GGA2 et GGA13, chromosomes qui correspondent aux groupes de liaison de nos deux QTLs pour ce caractère.

4.2.3 Les caractères d'aptitude au gavage

Dix QTLs significatifs à 5% au niveau du chromosome sont détectés pour des caractères d'aptitude au gavage des animaux (Appendix 3 – tableau 2).

Hormis le poids de foie mesuré sur des canards maigres ou sur du poulet, les caractères d'aptitude au gavage ne peuvent pas être comparés avec les résultats des poules, celles-ci n'étant pas gavées. Dans notre analyse aucun QTL de poids de foie gras significatif entre 1 et 5% au niveau du chromosome n'a été détecté.

4.2.4 Les caractères de qualité de produits

Dix-huit QTLs significatifs à 5% au niveau du chromosome, contrôlant des caractères de qualité du foie gras et du magret sont présentés dans le tableau 3 de l'appendix 3. Dix de ces QTLs sont relatifs à la qualité du foie gras et 8 à celle du magret.

Les groupes de liaison 2c et 9 semblent être des GL d'intérêt pour la qualité du foie gras car en plus des QTLs significatifs détaillés dans l'article 2, on y retrouve 5 QTLs significatifs à 5% au niveau de chromosome pour ce type de caractère : 3 QTLs relatifs au taux de fonte, au taux de collagène et à la coloration jaune apparaissent sur LG2c et 2 QTLs impactant le taux de fonte de le taux de lipide du foie sur LG9.

Au-delà de LG2c, 2 autres QTLs significatifs à 5% au niveau du chromosome sont cartographiés pour la coloration jaune du foie (GL1a et GL3). Ce caractère associé à l'extension des taches due aux viscères permet d'améliorer en puissance et en précision le QTL sur le GL3 dans l'analyse multi-caractère (Lb* et ESV). À ce jour, le seul QTL identifié chez la poule pour le caractère de coloration du foie l'est sur le chromosome GGA4 (Wright *et al*., 2006).

Sur le groupe de liaison 2a, en plus des 3 QTLs significatifs à 1% au niveau du chromosome, 2 QTLs pour des caractères de qualité de magret (pour la coloration jaune du magret et le poids du magret) significatifs à 5% sont détectés. On détecte aussi sur le GL7, deux QTLs significatifs à 5% au niveau du chromosome pour la coloration rouge du magret et le pH de celui-ci 20min post mortem. A noter qu'aucun QTL significatif à 1% n'est cartographié sur le GL7 pour les caractères de qualité du magret.

Les chromosomes GGA2 et GGA7 semblent être des chromosomes d'intérêt pour les caractères de qualité de la viande de volaille. En effet, trois QTLs impliqués dans la variabilité du pH (pH mesuré 20 minutes *post mortem* et le pH ultime) dans le muscle et dans la coloration du magret sont identifiés chez la poule (Wright *et al*., 2006). Nadaf *et al*. (2007) ont confirmé le QTL obtenu sur GGA2 pour le pH de la viande. Pour la coloration de la viande, le QTL identifié sur le groupe de liaison 2a semble être connu chez la poule : Van Kaam *et al*, 1999 ont mis en évidence une zone chromosomique impliquée dans la variabilité de la coloration de la viande de la poule sur GGA2.

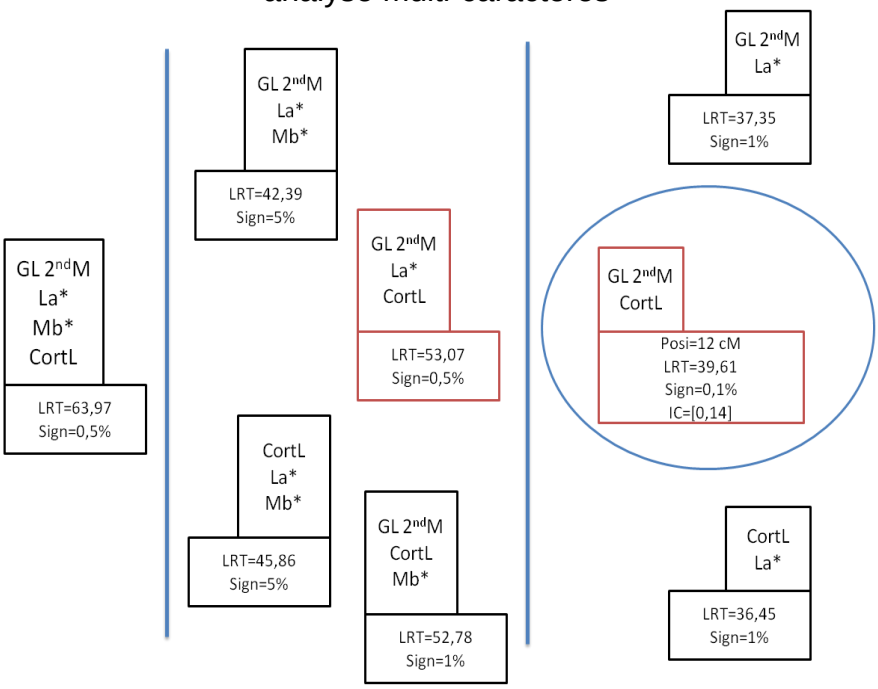
4.3 Les analyses multi-caractères

Comme décrit dans l'article, une analyse multi-caractères a été menée en complément de l'analyse uni-caractère. Les caractères ont été regroupés suivant deux critères : i) caractères très corrélés phénotypiquement, ii) caractères présentant des QTL à 1% co-localisés. L'objectif de cette approche multi-caractère était d'identifier les QTLs pléiotropes. Les résultats de ces analyses multi-caractères ont été présentés dans l'article 2, mais la démarche suivie pour aboutir à ces résultats mérite d'être détaillée.

Initialement, des groupes de 2, 3, 4 ou 5 caractères sont constitués selon les critères précédemment décrits. Lorsqu'il y a N caractères impliqués (avec N supérieur à 2), il est nécessaire de savoir combien de caractères parmi ces N caractères sont impliqués dans la pléiotropie. Des détections de QTL sont effectuées par étape afin d'identifier la combinaison des caractères pour lequel le QTL est le plus significatif. On retiendra au final ce sous-groupe de caractères.

Ci-dessous l'exemple du GL2a est détaillé (figure 4.8). Initialement, 4 QTLs à 5% étaient co-localisés en début de groupe de liaison grâce aux analyses uni-caractères (tableaux 2 et 3 de l'appendix 3) et concernaient les caractères de taux de glucose sanguin au 2ème repas de gavage, le taux de corticosterone sanguin basal à 6 semaines d'age, l'indice de rouge des foies gras et l'indice de jaune des magrets.

Figure 4.8 : Exemple de la démarche suivie pour arriver au résultat final d'une analyse multi-caractères



Description : La première colonne correspond à la première étape de l'analyse où tous les caractères sont analysés ensemble, la deuxième colonne correspond à l'étape de l'analyse où l'on considère les caractères 3 par 3 et la dernière colonne correspond à l'étape de l'analyse où l'on décortique, en effectuant des analyses bi-caractères, le groupe de caractère qui possède le plus grand LRT dans l'étape

précédente. Chaque analyse est représentée ici par deux rectangles. Dans le rectangle du haut sont recensés les caractères analysés ensemble et dans celui du bas on trouve les résultats de l'analyse (maximum LRT et la significativité atteinte).

Dans la première colonne l'analyse simultanée de tous les caractères ayant des QTLs co-localisés en uni-caractère (au seuil de 5 % au niveau du chromosome), soit 4 caractères différents, est effectuée. Puis, les 4 combinaisons de ces caractères pris 3 par 3 font l'objet de nouvelles détections multi-caractères (résultats dans la deuxième colonne) : sera retenu le sous-groupe qui présente la plus grande valeur LRT à condition que la significativité de ce QTL soit supérieure à celle du QTL obtenu pour le groupe complet (à significativité égale, on retient le groupe qui possède le moins de caractères). Ce sous-groupe de 3 caractères fera à son tour l'objet du même traitement : les 3 combinaisons de ces caractères pris 2 à 2 seront testées. Au final, l'association de GL 2^{nd}M et CortL permet d'obtenir un QTL significatif à 0,1% au niveau du chromosome (et 5% au niveau du génome).
On conclut ici, qu'il existe un QTL pléiotrope qui impacte préférentiellement le taux de glucose dans le sang au deuxième repas de gavage (GL 2^{nd}M) et le taux de corticostérone basale (CortL) en élevage avant stress (cf l'article 2).

Remarque : s'agissant d'un groupe constitué avec le critère de co-localisation, les autres combinaisons à deux caractères ont aussi fait l'objet d'analyses QTL pour voir s'il n'y avait pas un deuxième QTL pléiotrope. Ces analyses n'ont donné aucun QTL.

Cette démarche permet au final de retenir le groupe de caractères qui permet de maximiser la significativité du QTL et/ou de réduire l'intervalle de confiance. Elle nous a aussi permis d'épurer des groupes, les caractères qui n'apportent rien en termes de significativité et/ou de précision à l'éventuel QTL pléiotrope. Pour les groupes de caractères formés sur le critère « niveau de corrélation, cette méthode a permis de mettre en évidence des nouveaux QTLs (par rapport à l'analyse uni-caractère). En effet, la combinaison des caractères a dans ce cas permis de mettre en évidence un QTL pléiotrope alors qu'aucun des caractères composant ce groupe, pris isolément ne franchissait le seuil de 5% au niveau du chromosome. Le principal inconvénient est la multiplicité des détections de QTL engendrée par cette approche (au moins 8 analyses pour un groupe de 4 caractères), qui risque de générer la création de faux-positifs.

Toujours sur le même groupe de liaison GL2a mais dans la partie distale, les analyses uni-caractères ont permis de cartographier 7 QTLs relatifs à des caractères de croissance (tableau 1 Appendix 3) et significatifs à 5% au niveau du chromosome à des positions très proches (tous dans l'intervalle [59cM-66cM]). Cependant aucune analyse multi-caractère n'a permis d'améliorer en terme de significativité ces résultats, ce qui nous laisse penser qu'il ne s'agit pas d'un QTL pléiotrope mais de plusieurs QTLs positionnés dans cette courte portion du génome. Les corrélations très fortes (pouvant atteindre 0,96) entre les caractères de croissance peuvent aussi expliquer le peu d'apport des analyses multi-caractères : l'apport d'information d'un deuxième caractère très fortement corrélés au premier est faible.

L'approche multi-caractère est aussi un outil permettant de renforcer nos résultats uni-caractères. Il en a été ainsi des groupes de liaison 2c et 9 qui semblent être très impliqués dans la variabilité de la qualité du foie. En effet on détecte dans le

cadre des analyses uni-caractères 6 QTLs relatifs à la qualité du foie sur le GL2c (qui sont le poids de foie, les taux de lipides, proteines et collagène dans le foie, le taux de fonte et l'indice de jaune) et 5 QTLs sur le GL9 (qui sont le poids de foie, l'indide ce rouge, les taux de protéines et lipides ainsi que le taux de fonte). Ces QTLs sont confirmés par des analyses multi-caractères, les regroupements de ces caractères ont même permis de renforcer les significativités des QTLs (tableau 6 de l'article 2). Néanmoins, ces deux groupes de liaison sont très faiblement couverts par des marqueurs dans notre étude : deux marqueurs distants de 1cM pour le GL2c et trois marqueurs répartis sur 20 cM pour le GL9. Ces deux groupes de liaison devront donc être densifiés en marqueurs pour confirmer (ou infirmer) nos résultats.

5 Conclusions et perspectives

Au total les analyses uni et multi-caractères ont permis l'identification de 81 régions chromosomiques impliquées dans la variabilité des caractères d'intérêt chez le canard gras. Pour les analyses uni-caractères, 74 QTLs sont cartographiés dont 22 sont significatifs à 1% au niveau du chromosome et les 52 autres à 5%. Les analyses multi-caractères ont permis d'identifier 7 nouvelles régions chromosomiques non identifiées par les analyses uni-caractères.

Pour poursuivre ce travail, il s'agit de tenter d'identifier les gènes sous-jacents à ces QTLs. Des choix s'imposent car il est inenvisageable d'étudier l'ensemble de ces QTL. Nous avons choisi de considérer à la fois les QTLs supposés robustes et qui impactent des caractères particulièrement intéressants pour la filière « foie gras » : les aspects qualités du foie gras avec les QTLs concernant les groupes de liaisons 2c et GL9 ont été retenus. Ainsi, dans le cadre d'une thèse en cours (Yoannah François, 2011-2014), les mulards issus des familles où ségrègent ces QTL relatifs à la qualité du foie gras feront l'objet d'analyses transcriptomiques et protéomiques, ce qui permettra de réaliser des détections de QTL d'expression et de protéome (eQTL et pQTL).

Par ailleurs, le taux de fonte à la cuisson du foie, caractère de rendement technologique qui est le critère de qualité primordial, est aussi au cœur d'un second projet de recherche : il s'agit de vérifier si le polymorphisme observé d'un gène, qui chez le poulet influence la couleur de la viande (Le Bihan-Duval *et al*. 2011) impacte ou non la qualité technologique du foie gras. Une partie du projet consistera à réaliser des détections de QTLs, qui sont rendues possibles grâce aux résultats méthodologiques et aux développements informatiques (QTLMap) réalisés dans le cadre de cette thèse.

Certains groupes de liaison pourront faire l'objet d'une densification en marqueurs car très impliqués dans la variabilité des caractères d'intérêt chez le canard gras. On peut parler du groupe de liaison 21 sur lequel sont cartographié deux QTLs hautement significatifs grâce aux analyses multi-caractères : le premier impliqué dans la variabilité des taux de cholestérol et de triglycéride dans le sang mesurés au dixième repas et le second dans la variabilité de poids de carcasse et de poids de fin de gavage. Toujours dans le cadre de la densification en marqueurs, les groupes de liaison faiblement couverts comme les macro-chromosomes 5 (6 marqueurs pour 26cM) et 6 (2 marqueurs pour 25cM) et sur lesquels des QTLs ont été identifiés pour des caractères d'intérêt chez le canard, devront être considérés. En effet, Huang *et al*, 2007 dans le cadre d'un protocole de canards non gavés, ont cartographié des QTLs pour des caractères de qualité du foie et du magret sur les deux chromosomes CAU5 (perte sous vide du magret) et CAU6 (perte à la cuisson du magret, poids du foie).

Conclusions générales

L'objectif principal de la thèse consistait à mettre en évidence des QTLs impliqués dans la variabilité des caractères d'intérêt chez le canard gras dans un dispositif de croisement et pour ce faire, répondre aux différentes questions méthodologiques posées. Ce travail a été mené en deux étapes, i) développements méthodologiques et ii) analyse de QTL des données du dispositif.

Dans un premier temps les problématiques méthodologiques ont été abordées. Il s'agissait de proposer une stratégie de prise en compte du mérite génétique d'une part et trouver la meilleure façon (utilisation du modèle normal ou du modèle à seuil) d'analyser les caractères discrets d'autre part. La performance affectée à une cane backcross génotypée était la moyenne des performances de ses descendants mulards, mais le nombre de descendants mulards était variable d'une cane à l'autre : les performances des backcross ont donc du être pondérées par un coefficient de détermination (CD) fonction du nombre de descendants et de l'héritabilité du caractère étudié pour les analyses uni-caractères. Pour la détection de QTL en multi-caractères, le CD est toujours fonction des nombres de descendants phénotypés pour les différents caractères mais intègre les covariances génétiques entre ceux-ci. Ces approches uni et multi-caractères ont été programmées dans le logiciel QTLMap afin d'être utilisables en pratique.

Cette démarche peut être généralisée à la prise en compte d'un mérite génétique dans la détection de QTL, étendant la méthode de Weller *et al*. (1990) usuellement utilisée pour les dispositifs à trois générations (protocoles petites filles) et celle de Tribout *et al*. (2008), quand des phénotypes sont disponibles chez les descendants et petits descendants. Nous avons proposé de considérer, en généralisant ce type de dispositif, toutes les données phénotypiques disponibles dans les générations postérieures aux individus génotypés. Ainsi, pour un individu génotypé donné, le mérite génétique est défini comme la synthèse de ses performances propres si elles existent, et de celles de tous les individus issus de cet individu dans des générations postérieures et possédant des performances. Cet apport en données phénotypiques dans le modèle de détection de QTL permet d'augmenter sa puissance, surtout pour les protocoles où les tailles de familles sont réduites. Ces travaux ont été publiés dans Journal of Animal Breeding and Genetics.

Le deuxième point méthodologique abordé dans cette thèse est la prise en compte des caractères discrets dans la détection de QTL. En effet, plusieurs caractères d'intérêt mesurés dans le cadre du dispositif présentent des distributions non normales. Ces caractères, dont la plupart possède des modalités distribuées de façon équilibrée, peuvent le plus souvent être traités comme des caractères normaux en utilisant donc une fonction de pénétrance gaussienne dans le modèle de détection de QTL. Par simulation, nous avons en effet pu prouver que l'utilisation d'un modèle à seuil, qui tient compte de la non normalité des observations, n'avait pas d'apport significatif pour les caractères discrets possédant des modalités distribuées de façon homogène. Cette conclusion est cohérente avec la littérature (paragraphe 2, chapitre II). En effet, il est largement admis dans la littérature que l'utilisation d'un modèle à seuil ou d'un modèle normal pour l'analyse QTL des caractères discrets présentent des caractéristiques équivalentes. A contrario, nous avons montré que dans des cas particuliers (plus de deux modalités en fréquences très déséquilibrées), le modèle à seuil est plus efficace.

Le temps de calcul était aussi un obstacle pour l'utilisation du modèle à seuil. En effet, ce dernier est deux fois plus gourmand en temps de calcul par rapport au

modèle normal de par sa construction non linéaire (et sans doute du manque d'optimisation des routines élaborées face aux cas gaussiens en place depuis longtemps dans le logiciel utilisé).

Dans un second temps, le travail de la thèse consistait à analyser les données du dispositif QTL mis en place pour les canards gras. Vingt deux QTLs significatifs à 1% au niveau du chromosome ont été détectés dans le cadre de l'analyse caractère par caractère. Il s'agit essentiellement de QTLs impliqués dans la variabilité des caractères d'aptitude au gavage et de qualité du foie gras et du magret. En effet 16 des 22 QTLs significatifs à 1% au niveau du chromosome concernent les caractères d'aptitude au gavage et de qualité (soit 72%). Or l'ensemble de ces caractères (caractères d'aptitude au gavage et de qualité de produits) ne constitue que 40% (29 caractères sur 73) des mesures effectuées dans le cadre du dispositif. Les caractères de qualité de foie sont particulièrement surreprésentés en terme de QTL détectés, ce qui est intéressant quand on connait l'intérêt économique de la qualité du foie gras pour les professionnels. De plus, des zones chromosomiques spécifiques ont été révélées pour ces caractères : grâce aux analyses uni-caractères, 6 QTLs apparaissent co-localisés sur le GL2c et 5 QTLs co-localisés sur le GL9 pour la qualité du foie. Ces QTLs sont confirmés par des analyses multi-caractères. Ces résultats ont permis de jeter les bases d'une nouvelle thèse (Yoannah François, 2011-2014), dont l'objet est l'analyse protéomique et transcriptomique des individus portant ces QTL pleïotropiques.

L'approche uni-caractère a été complétée par des analyses multi-caractères dans l'espoir d'identifier de nouveaux QTLs grâce au regroupement de caractères corrélés mais aussi de confirmer et préciser des QTLs pléïotropes par l'analyse conjointe de caractères présentant des co-localisations de QTL. Ces analyses multi-caractères ont permis de cartographier 7 nouveaux QTLs significatifs à 1% au niveau du chromosome (via l'approche basé sur le critère de corrélation). L'approche multi-caractère a permis d'une part d'augmenter la puissance de détection, puisque 8 QTLs à 0,1% au niveau du chromosome ont été détectés contre 3 seulement en uni-caractère. D'autre part, ces analyses multi-caractères ont permis de réduire les intervalles de confiances des QTLs obtenus en uni-caractère.

Au total, 81 QTLs ont été mis en évidence dans le cadre de cette thèse. Soixante quatorze QTLs (dont 22 significatifs à 1% au niveau de chromosome et 52 significatifs à 5%) sont cartographiés grâce aux analyses uni-caractères. A cela s'ajoute les 7 nouveaux QTLs obtenus grâce aux analyses multi-caractères. L'ensemble de ce travail de détection de QTL a été soumis à la revue Journal of Animal Science fin avril 2012.

La suite à donner aux travaux méthodologiques réalisés est :

i) pour la première partie relative au mérite génétique, la prise en compte d'éventuels liens de parenté entre les pères fondateurs. En effet cette piste est à approfondir pour apporter des améliorations au modèle. La faisabilité de la prise en compte des performances des apparentés aux individus marqués pour réaliser la détection des QTLs peut aussi être étudiée. Cela permettrait d'augmenter considérablement la puissance des protocoles basés sur des individus marqués ayant des petits nombres de descendants grâce à la prise en compte des performances des apparentés (mais aussi des descendants de ces derniers).

ii) pour le deuxième point méthodologique abordé dans cette thèse, la prise en compte des caractères discrets dans la détection de QTLs, les résultats peuvent être confirmés en réalisant un nombre plus élevé de simulations. Des paramètres nouveaux peuvent être considérés, comme l'héritabilité du caractère étudié qui était fixé dans nos simulations.

Pour le protocole « canard à gaver », la suite directe de ces travaux consistera à densifier la carte génétique avec des marqueurs SNP afin de couvrir i) les chromosomes non couverts dans la carte canard actuelle (chromosomes 8, 10, 11 et 12), ii) les macro-chromosomes 5 et 6 connus pour leur implication dans la variabilité des caractères d'intérêt chez le canard et très faiblement couvert dans notre étude, et iii) le groupe de liaison 2c et le chromosome 9 qui hébergent un nombre important de QTLs affectant des caractères de qualité de foie, caractères de grande importance dans la filière « foie gras ».

Références bibliographiques

Abasht, B., Dekkers, J. C. M., & Lamont, S. J. (2006). Review of quantitative trait loci identified in the chicken. *Poultry Science*, 85, 2079-2096.

Almasy, L., and Blangero, J., (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. Am. J. Hum. Genet. 62: 1198–1211.

Ambo, M., Moura, A. S. A. M. T., Ledur, M. C., Pinto, L. F. B., Baron, E. E., Ruy, D. C., et al. (2009). Quantitative trait loci for performance traits in a broiler × layer cross. *Animal Genetics*, 40, 200-208.

Atzmon, G., Blum, S., Feldman, M., Cahaner, A., Lavi, U., & Hillel, J. (2008). QTLs detected in a multigenerational resource chicken population. *Journal of Heredity*, 99, 528-538.

Atzmon, G., Blum, S., Feldman, M., Lavi, U., & Hillel, J. (2007). Detection of agriculturally important QTLs in chickens and analysis of the factors affecting genotyping strategy. *Cytogenetic and Genome Research*, 117, 327-337.

Babilé, R. 1989. La production de foie gras de canards de barbarie (cairina moschata) : Aspects génétiques, nutritionnels et technologiques. *Institut polytechnique de Toulouse* (p. 315).

Babilé, R., Matheron, G, Poujardieu, B. (1988). La production de foie gras de canards mâles de barbarie. Relation entre la croissance et la production de foie. *Ann Zootech*, 37, 219-252

Baéza, E., Rideau, N., Chartrin, P., Davail, S., Hoo-Paris, R., Mourot, J., et al. (2005). Canards de barbarie, pékin et leurs hybrides : Aptitude à l'engraissement. *INRA Production Animale*, 18, 131-141.

Baeza, E., Salichon, M.R., Marche, G., Juin, H., 1997. Effet du sexe sur la croissance et sur les caractéristiques technologiques et sensorielles du pectoralis major (filet) chez le canard de barbarie. *2èmes Journées de la Recherche Avicole* pp. 191-194). Tours, France.

Baéza, E., Williams, J., Guémené, D., & J. Duclos, M. (2001). Sexual dimorphism for growth in muscovy ducks and changes in insulin-like growth factor i (igf-i), growth hormone (gh) and triiodothyronine (t3) plasma levels. *Reprod. Nutr. Dev.*, 41, 173-179.

Brun, J. M., & Larzul, C. (2003). Inheritance of reproductive traits of female common ducks (*anas platyrhynchos*) in pure breeding and in inter-generic crossbreeding with muscovy ducks (*cairina moschata*). *British Poultry Science*, 44, 40-45.

Brun, J. M., Richard, M. M., Marie-Etancelin, C., Rouvier, R., & Larzul, C. (2005). Le canard mulard : Déterminisme génétique d'un hybride intergénérique. *INRA Prod. Anim.*, 18, 295-308.

Brun, J. M., Sellier, N., Hu, Y.H., Rouvier, R. (1999). Fertility and embryonic mortality in the intergeneric cross between the common duck and the muscovy duck: Candling vs. Opening the eggs. *1st World Conf. Waterfowl* pp. 193-198. Taichung, Taiwan.

Buitenhuis, A. J., Rodenburg, T. B., Van Hierden, Y. M., Siwek, M., Cornelissen, S. J., Nieuwland, M. G., et al. (2006). Erratum to "Mapping quantitative trait loci affecting feather pecking behavior and stress response in laying hens". *Poultry Science*, 85, 1115-1116.

Campos, R. L. R., Nones, K., Ledur, M. C., Moura, A. S. A. M. T., Pinto, L. F. B., Ambo, M., et al. (2009). Quantitative trait loci associated with fatness in a broiler–layer cross. *Animal Genetics*, 40, 729-736.

Carlborg, O., Andersson, L. and Kinghorn, B. P., (2000). The use of a genetic algorithm for simultaneous mapping of multiple interacting quantitative trait loci. Genetics 155: 2003-2010.

Carlborg, O., Kerje, S., Schütz, K., Jacobsson, L., Jensen, P., & Andersson, L. (2003). A global search reveals epistatic interaction between qtl for early growth in the chicken. *Genome Research*, 13, 413-421.

Castaing, J., Albar, J., Serin, J. F. (1995). Valorisation par le porc charcutier de co-produits d'industries agro-alimentaires du sud de la france : Farine de maïserie, remoulage demi-blanc, farine basse de riz*Journées Rech. Porcine* pp. 223-230). France.

Chapuis, H., Larzul, C. (2006). Comment estimer simultanément les paramètres génétiques des caractères de gavage dans les deux lignées parentales du mulard en vue d'une sélection plus efficace ? *JRPFG*. Arcachon, France: 18-19 octobre.

Chartrin, P., Schiavone, A., Bernadet, M. D., Guy, G., Mourot, J., Duclos, M. J., et al. 2004. Effet du génotype et du gavage sur les dépôts de lipides intramusculaires dans le filet de canard. *6e Journées de la Recherche sur les Palmipèdes à Foie Gras* pp. 49-52). Arcachon, France.

Churchill, G., Doerge, R., (1994). Empirical threshold values for quantitative trait mapping. Genetics, 138:963–971.

D'André Hirwa, C., Yan, W., Wallace, P., Nie, Q., Luo, C., Li, H., et al. (2010). Effects of the thyroid hormone responsive spot 14α gene on chicken growth and fat traits. *Poultry Science*, 89, 1981-1991.

Darvasi, A., Weinreb, A., Minke, V., Weller, J. I. and Soller, M., (1993). Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics. 134: 943–951.

Darvasi, A. (1998). Experimental strategies for the genetic dissection of complex traits in animal models. Nat Genet 18, 19-24.

Davail, S., Rideau, N., Guy, G., André, J.-M., Hermier, D., & Hoo-Paris, R. (2003). Hormonal and metabolic responses to overfeeding in three genotypes of ducks. *Comparative Biochemistry and Physiology - Part A: Molecular and Integrative Physiology*, 134, 707-715.

Dufour, S., Durand, T., Gallis, J.L., Vuillaume, A., Canioni, P. 1993. Étude par rmn à basse résolution du foie gras de canard mulard : Application à la mesure du taux de fonte*1ère journée de la Recherche sur les palmipèdes à foies gras*.

Duhart, F. (2004). Pour une ethnozoologie historique des palmipèdes en Europe : la naissance du foie gras moderne (XVIe- XIXe s.). *6e Journ. Rech. Palmipèdes à Foie Gras* pp. 9-12. Arcachon, France.

Elsen, J.M., Mangin, B., Goffinet, B., Boichard, D., Le Roy, P., (1999). Alternative models for QTL detection in livestock. I General introduction. Genet. Sel. Evol. 31:213-224.

Faure, J. M., Guémené, D., Guy, G. (2001). Is there avoidance of the force feeding of ducks and geese? *Anim. Res.*, 50, 157-164.

Filangi, O., Moreno, C., Gilbert, H., Legarra, A., Le Roy, P., Elsen, J.M., (2010). QTLMap, a Software for QTL Detection in Outbred Populations. In Proceedings of the 9th WCGALP: 1-6 August; Leipzig Germany.

Fillon, V., Vignoles, M., Garrigues, A., Croojmans, R. P. M. A., Groenen, M. A. M., Séguela, A., et al. (2001). Comparative avian cytogenetics : Establishment of chromosome correspondence by fluorescent in situ hybridisation. *British Poultry Science*, 42, S23-S24.

Foulley, J. L., Gianola, D., & Thompson, R. (1983). Prediction of genetic merit from data on binary and quantitative variates with an application to calving difficulty, birth weight and pelvic opening. *Genetics Selection Evolution*, 15, 401 - 424.

Gao, Y., Du, Z. Q., Wei, W. H., Yu, X. J., Deng, X. M., Feng, C. G., et al. (2009). Mapping quantitative trait loci regulating chicken body composition traits. *Animal Genetics*, 40, 952-954.

Genizi, M. S. a. A. (1978). The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations. *Biometrics*, 34, 47-55.

George, A. W., Visscher, P. M. and Haley, C.,. S (2000). Mapping quantitative trait loci in complex pedigrees: a two-step variance component approach. Genetics 156: 2081–2092.

Gibson, G., Muse, S.V., (2004). Précis de génomique. de Boeck.

Gilbert, H. and Le Roy, P., (2001). Comparison of three multitrait methods for QTL detection. EAAP, Budapest, Hongria.

Gilbert, H. 2003. Multidimensionnalité pour la détection de gènes influençant des caractères quantitatifs. Application à l'espèce porcine. Institut National Agronomique Paris-Grignon (p. 188).

Gilbert, H. and Le Roy, P. (2003). Comparison of three multitrait methods for QTL detection. Genet. Sel. Evol., 35:281–304.

Gilbert, H., Le Roy, P. (2007). Methods for the detection of multiple linked QTL applied to a mixture of full and half sib families. Genet. Sel. Evol., 39:139–158.

Gille, U., Salomon, F.V. (1994). Heart and body growth in ducks. *Growth Dev. Aging*, 58, 75-81.

Goffinet, B., Le Roy, P., Boichard, D., Elsen, J. M. & Mangin, B., (1999). Alternative models for QTL detection in livestock. III.Heteroskedastic model and models corresponding to several distributions of the QTL effects. Genetic Selection Evolution 31,341-350.

Green, P., Falls K. and Crooks, S. (1990). CRI-MAP Documentation, Version 2.4., Washington University School of Medicine, St. Louis.

Guémené, D., & Guy, G. (2004). The past, present and future of forcefeeding and “foie gras” production. *World’s Poultry Science*, 60, 211-222.

Guémené, D., Guy, G., Mirabito, L., Sevère, J., & Faure, J.-M. (2007). Bien-être et élevage des palmipèdes. *INRA Prod. Anim.*, 20 53-58.

Guémené, D., Chapuis, H., Boulay, M., Garreau-Mills, M. (2002). Plasma corticosterone levels in male chickens from two divergently selected lines for featherpecking*11th European Poultry Conference*. Bremen (DEU).

Guy, G., Rousselot-Pailley, D., Gourichon, D. (1995). Comparaison des performances de l’oie, du canard mulard et du canard de barbarie soumis au gavage. *Ann. Zootech.*, 44, 297-305.

Guy, G., Hermier, D., Davail, S., Bely, M., André, J.M., Hoo-Paris, R. (1999). Meat production and force-feeding ability of different types of ducks. 1st World Waterfowl Conference, Taichung, Taïwan, 1-4 déc. 1999, 462-468.

Gianola, D., & Foulley, J. L. (1983). Sire evaluation for ordered categorical data with a threshold model. Genetics Selection Evolution, 15, 201 - 224.

Hackett CA, W. J. (1995). Genetic mapping of quantitative trait loci for traits with ordinal distributions. Biometrics, 51, 1252-1263.

Haldane, J.B.S., (1919). The combination of linkage values and the calculation of distance between the loci of linked factors. J. Genetics. 8:299-309.

Haley, C. S., & Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity, 69, 315-324.

Haley, C.S., Knott, S.A. and Elsen, J.M., (1994). Mapping quantitative trait loci in crosses between outbred lines using least squares. Genetics 136:1195–1207.

Harville, D. A. a. M., R.W. (1984). A mixed model procedure for analyzing ordered categorical data. Biometrics, 40, 393–408.

Hermier, D., Salichan, M. R., Guy, G., Peresson, R., Mourot, J., & Lagarrigue, S. (1999). La stéatose hépatique des palmipèdes gavés : Bases métaboliques et sensibilité génétique. *INRA Production Animale*, 12, 265-271.

Hu, Y., Poivey, J., Rouvier, R., Liu, S., & Tai, C. (2004). Heritabilities and genetic correlations of laying performance in Muscovy ducks selected in Taiwan. *Br Poult Sci.*, 45, 180-185.

Hu, Y. H., Poivey, J.P., Rouvier, R., Wang, C.T., Tai, C. (1999). Heritabilities and genetic correlations of body weights and feather length in growing Muscovy selected in Taïwan. *Br. Poult. Sci.*, 40, 605-612.

Huang, Y., Haley, C. S., Hu, S., Hao, J., Wu, C., & Li, N. (2007a). Detection of quantitative trait loci for body weights and conformation traits in beijing ducks. *Animal Genetics*, 38, 525-526.

Huang, Y., Haley, C. S., Wu, F., Hu, S., Hao, J., Wu, C., et al. (2007b). Genetic mapping of quantitative trait loci affecting carcass and meat quality traits in beijing ducks (*anas platyrhynchos*). *Animal Genetics*, 38, 114-119.

Imbert, B. 2002. Comparaison génétique des caractéristiques biochimiques du magret de canard. *Diplôme de Recherche Universitaire de l'Institut National Polytechnique de Toulouse* (p. 63).

Jacobsson, L., Park, H.-B., Wahlberg, P., Fredriksson, R., Perez-Enciso, M., Siegel, P. B., et al. (2005). Many qtls with minor additive effects are associated with a large difference in growth between two selection lines in chickens. *Genetical Research*, 86.

Jamrozik, J., Schaeffer, L. R., Burnside, E. B., & Sullivan, B. P. (1991). Threshold models applied to holstein conformation traits. *Journal of Dairy Science*, 74, 3196-3201.

Jennen, D. G., Vereijken, A. L., Bovenhuis, H., Crooijmans, R. P., Veenendaal, A., van der Poel, J. J., et al. (2004). Detection and localization of quantitative trait loci affecting fatness in broilers. *Poultry Science*, 83, 295-301.

Jensen, P., Keeling, L., Schütz, K., Andersson, L., Mormède, P., Brändström, H., et al. (2005). Feather pecking in chickens is genetically related to behavioural and developmental traits. *Physiology & Behavior*, 86, 52-60.

Jansen, R.C., 1992. A general mixture model for mapping quantitative trait loci by using molecular markers. Theor Appl Genet, 85, p 252-260.

Jansen, R.C., 1993. Interval mapping of multiple quantitative trait loci. Genetics, 135, p 205-211.

Jansen, R.C., and Stam, P., (1994). High resolution of quantitative traits into multiple loci via interval mapping. Genetics, 136, p 1447-1455.

Jansen, R.C., (1994). Controlling the type I and type II errors in mapping quantitative trait loci. Genetics, 138, p 871-881.

Jansen, R.C., (1996). A general Monte Carlo method for mapping multiple quantitative trait loci. Genetics, 142, p 305-311.

Jansen, R.C., Johnson, D.L. et van Arendonk J.A.M., (1998). A mixture model approach to the mapping of quantitative trait loci in complex populations with an application to multiple cattle families. Genetics, 148, p 391-399.

Kao, C.-H., Zeng, Z.-B., Teasdale, R.D., (1999). Multiple interval mapping for quantitative trait loci. Genetics, 152:1203-1216.

Knapp, S.J., Bridges, W.C., Birkes J.and D., (1990). Mapping quantitative trait loci using molecular marker linkage maps. Theor Appl Genet 79: 583-592.

Knizetova, H., Hyanek, J., Hyankova, L., Belicek, P., (1995). Comparative study of growth curves in poultry. *Genet. Sel. Evol.*, 27, 365-375.

Knizetova, H., Hyanek, J., Knize, B., Prochatzova, H., (1991). Analysis of growth curve of fowl. II. Ducks. *Br. Poult. Sci.*, 22, 1027-1038.

Knott, S., Elsen, J.-M., Haley, C., (1996). Methods for multiple marker mapping of quantitative trait loci in half-sibs populations. Theor. Appl. Genet., 93:71–80.

Korol, A.B., Ronin, Y.I., Kirzhner, V.M. (1995). Interval mapping of quantitative trait loci employing correlated trait complexes. Genetics, 140:1137-1147.

Korol, A.B., Ronin, Y.I., Itskovich, A.M., Peng, J., Nevo, E., (1995). Enhanced efficiency of quantitative trait loci mapping analysis based on multivariate complexes of quantitative traits. Genetics, 175:1789-1803.

Kosambi, D.D., (1944). The estimation of map distance from recombination values, Annals of Eugenics. 12(3). 172-175

Lander, E. S., & Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using rflp linkage maps. *Genetics*, 121, 185-199.

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J,. Lincoln, S.E., Newburg, L., (1987). MAPMAKER : an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1, 174-181

Larzul, C. 2002. Paramètres génétiques du gavage dans une population de canards communs. *5è Journ. Rech. Palmipèdes à Foie Gras* pp. 33-36). Pau, France.

Larzul, C., Guy, G., Bernadet, M.D. (2004). Feed efficiency, growth and carcass traits in female mule ducks. *Arch. Gefügelk.*, 6, 265-268.

Larzul, C., Imbert, B., Bernadet, M.-D., Guy, G., & Rémignon, H. (2006). Meat quality in an intergeneric factorial crossbreeding between muscovy (cairina moschata) and pekin (anas platyrhynchos) ducks. *Anim. Res.*, 55, 219-229.

Le Bihan-Duval, E., Nadaf, J., Berri, C., Pitel, F., Graulet, B., Godet, E., et al. (2011). Detection of a cis eqtl controlling BCMO1 gene expression leads to the identification of a qtg for chicken breast meat color. *PLoS ONE*, 6, e14825.

Le Roy, P., Elsen, J.-M., Boichard, D., Mangin, M., Bidanel, J.P., Goffinet, B., (1998). An algorithm for QTL detection in mixture of full and half sib families. In 6th WCGALP. Armindale, 26: 257-260.

Léveille, G. A., Romsos, D. R., Yeh, Y. Y., & O'Hea, E. K. (1975). Lipid biosynthesis in the chick. A consideration of the site of synthesis, influence of diet and possible regulatory mechanisms. *Poultry Sci.*, 54, 1075-7093.

Liu, X., Zhang, H., Li, H., Li, N., Zhang, Y., Zhang, Q., et al. (2008). Fine-mapping quantitative trait loci for body weight and abdominal fat traits: Effects of marker density and sample size. *Poultry Science*, 87, 1314-1319.

Luo, Z.W. and M.J. Kearsey. (1989). Maximum likelihood estimation of linkage between a marker gene and a quantitative locus. Heredity 63: 401-408.

Mangin, B., Thoquet, P., Grimsley, N.H., (1998). Pleiotropic QTL analysis. Biometrics, 54:88-99.

Mangin, B., Goffinet, B., Le Roy, P., Boichard, D. & Elsen, J. M., (1999). Alternative models for QTL detection in livestock. II.Likelihood approximations and sire marker genotype estimations. Genetic Selection Evolution 31, 225-237.

Marie-Etancelin, C., Basso, B., Davail, S., Gontier, K., Fernandez, X., Vitezica, Z. G., et al. (2011). Genetic parameters of product quality and hepatic metabolism in fattened mule ducks. *Journal of Animal Science*, 89, 669-679.

Marie-Etancelin, C., Chapuis, H., Brun, J. M., Larzul, C., Mialon-Richard, M. M., & Rouvier, R. (2008). Genetics and selection of mule ducks in France: A review. *World's Poultry Science Journal*, 64, 187-208.

Marie-Etancelin, C., Larzul, C., Mialon-Richard, M.M., Brun J.M. (2006). Génétique du canard mulard : Synthèse bibliographique 2-paramètres génétiques dans les populations parentales. *7èmes Journées de la Recherche sur les Palmipèdes à Foie Gras*. Arcachon, France.

Martinez, O., Curnow, R.N., (1992). Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet., 85: 480-488.

Maruyama, K., Akbar, M.K., Turk, C.M., (1999). Growth patterns and carcass development in male ducks selected for growth rate. *Br. Poult. Sci.*, 40, 233-239.

Migliore, L., Romboli, I., Fedeli-Avanzi, C. (1988). Génétique du plumage du canard de barbarie (cairina moschata). Une mise au point. *Les colloques de l'INRA* pp. 29-38).

Mignon-Grasteau, S., Beaumont, C., Poivey, J.-P., & de Rochambeau, H. (1998). Estimation of the genetic parameters of sexual dimorphism of body weight in 'label' chickens and muscovy ducks. *Genetics Selection Evolution*, 30, 481 - 491.

Misztal, I. (1999). *Interbull Bull*, 20, 33-42.

Nadaf, J., Gilbert, H., Pitel, F., Berri, C., Feve, K., Beaumont, C., *et al*. (2007). Identification of qtl controlling meat quality traits in an f2 cross between two chicken lines selected for either low or high growth rate. *BMC Genomics*, 8, 155.

Nadaf, J., Pitel, F. d. r., Gilbert, H. l. n., Duclos, M. J., Vignoles, F., Beaumont, C., et al. (2009). Qtl for several metabolic traits map to loci controlling growth and body composition in an F2 intercross between high- and low-growth chicken lines. *Physiological Genomics*, 38, 241-249.

Nakamichi, R., Ukai, Y., Kishino, H., (2001). Detection of closely linked multiple quantitative trait loci using a genetic algorithm. Genetics, 158:463-475.

Nie, Q., Fang, M., Xie, L., Peng, X., Xu, H., Luo, C., *et al*. (2009). Molecular characterization of the ghrelin and ghrelin receptor genes and effects on fat deposition in chicken and duck. *Journal of Biomedicine and Biotechnology*.

Noirault, J., Guemene, D., Guy, G., Faure, J.M. (1999). Corticosterone plasma concentration in male mule ducks: Effects of sampling sites, repeated samplings and acth injections. *Br. Poult. Sci.*, 40 304-308.

Olver, M. D., Preez, J.J., Kuyper, M.A., Mould, D.J. (1977). The carcass composition and growth of the "mule" ducklings compared to purebred ducklings. *Agroanimalia*, 9, 7-12.

Ou, J. T., Tang, S. Q., Sun, D. X., & Zhang, Y. (2009). Polymorphisms of three neuroendocrine-correlated genes associated with growth and reproductive traits in the chicken. *Poultry Science*, 88, 722-727.

Park, H.-B., Jacobsson, L., Wahlberg, P., Siegel, P. B., & Andersson, L. (2006). Qtl analysis of body composition and metabolic traits in an intercross between chicken lines divergently selected for growth. *Physiological Genomics*, 25, 216-223.

Pong-Wong, R., George, A.W., Woolliams, J.A., Haley, C.S., (2001). A simple and rapid method for calculating identity-by-descent matrices using multiple markers. Genet. Sel. Evol., 33:453-471.

Poujardieu, B., Guichard, F., & Laventure, P. (1994). Parametres genetiques de croissance et de gavage de la cane commune. *Genetics Selection Evolution*, 26, 463 - 472.

Rabie, T. S. K. M., Crooijmans, R. P. M. A., Bovenhuis, H., Vereijken, A. L. J., Veenendaal, T., van der Poel, J. J., et al. (2005). Genetic mapping of quantitative trait loci affecting susceptibility in chicken to develop pulmonary hypertension syndrome. *Animal Genetics*, 36, 468-476.

Rebaï, A., Goffinet, B. and Mangin, B., (1995). Comparing power of different methods for QTL detection. Biometrics. 51: 87–99.

Robin, N., Babilé, R., Peyhorgue, A., Dubois, J.P., Leprettre, S. (2004). Facteurs de production et qualité des foies gras de canards et d'oies. *6èmes Journ. Rech. Palmipèdes à Foie Gras* pp. 157-165. Arcachon, France.

Ronin, Y.I., Kirzhner, V.M., Korol A.B., (1995). Linkage between loci of quantitative traits and marker loci : multitrait analysis with a single marker. Theor. Appl. Genet., 90:776-786.

Rouvier, R. (1992). L'amélioration génétique en france : Le contexte et les acteurs - les palmipèdes. *INRA Prod. Anim.*, hors série, 39-43.

Rouvier, R., Mialon, M.-M., Salzmann, F., Poujardieu, B., & Combebiac, R. (1988). Fertilité et éclosabilité des oeufs d'une souche de canes pékin (anas platyrhynchos) en croisement interspécifique avec le barbarie (cairina moschata) par insémination artificielle. *Ann. Zootech.*, 37, 73-86.

Rouvier, R., Tai Lui, J.J., Tai C. (1984). L'insémination artificielle des canes communes pour la production de mulards à taïwan. La situation actuelle. In : Insémination artificielles et amélioration génétique : Bilan et perspectives critiques. *Les colloques de l'INRA*, 29, 360-367.

Rouvier R, Guy G, Rousselot-Paillet D., Poujardieu B, 1994. Genetic parameters from factorial cross breeding in two duck strains (Anas platyrhyncos) Brown Tsaiya and Pekin, for growth and fatty liver traits. British Poultry Science 35:509-517.

Salichon, M. R., Guy, G., Rousselot, D., Blum, J.C. (1993). Composition de trois types de foies : Oie, canard mulard et canard de barbarie. *1ères Journ. Rech. Palmipèdes à Foie Gras* pp. 13-21). Bordeaux, France.

Salichon, Y. (1991). *Elevage et gavage des palmipèdes pour la production du foie gras.* Paris.

Seaton, G., Haley, C.S., Knott, S.A., Kearsey, M., Visscher, P.M., (2002). QTL Express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics, 18(2), 339-40.

Setiawan, I., Babilé, R, Auvergne, A, *et al.* (1994). Growth and carcass composition of 2 genetic types in mule ducklings. *Ann. Zootech.*, 43, 185-196.

Siwek, M., Cornelissen, S. J., Buitenhuis, A. J., Nieuwland, M. G., Bovenhuis, H., Crooijmans, R. P., et al. (2004). Quantitative trait loci for body weight in layers differ from quantitative trait loci specific for antibody responses to sheep red blood cells. *Poultry Science*, 83, 853-859.

Smith, D. P., Fletcher, D. L., & Papa, C. M. (1992). Post-mortem biochemistry of pekin duckling and broiler chicken pectoralis muscle. *Poultry Science*, 71, 1768-1772.

Soller, M., Brody T. and Genizi, A., (1976). Power of Experimental Designs for Detection of Linkage between Marker Loci and Quantitative Loci in Crosses between Inbred Lines. Theoretical and Applied Genetics 47: 35-39.

Soller, M., and Genizi, A., (1978). Efficiency of Experimental-Designs for Detection of Linkage between a Marker Locus and a Locus Affecting a Quantitative Trait in Segregating Populations. Biometrics 34: 47-55.

Tai, C., Rouvier, R. (1998). Crossbreeding effect on sexual dimorphism of body weight in intergeneric hybrids obtained between muscovy and pekin duck. *Genet. Sel. Evol.*, 30, 163-170.

Tai Liu, J. J., Tai, C. (1984). Studies on the artificial insemination of ducks. 3. A comparison of fertility of pooled semen and individual male semen in the crosses between muscovy (*cairina moschata*) and tsaiya duck (*anas platyrhynchos var domestica)*. *J. Taiwan Livest. Res.*, 17, 85-89.

Théron, L. 2011. Déterminisme biologique de la variabilité de la fonte lipidique à la cuisson du foie gras de canard. *Institut polytechnique de Toulouse* (p. 161).

Thomsen, H., Reinsch, N., Xu, N., Looft, C., Gruppe, S., Kühn, C., Brockmann, G. A., Schwerin, M., Leyhe-Horn, B., Hiendleder, S., Erhardt, G., Medjugorac, I., Russ, I., Förster, M., Brenig, B., Reinhardt, F., Reents, R., Blümel, J., Averdunk, G. and Kalm., E., (2001). Comparison of estimated breeding values, daughter yield deviations and de-regressed proofs within a whole genome scan for QTL. J. Anim. Breed. Genet. 118:357–370.

Tribout, T., Iannuccelli, N., Druet, T., Gilbert, H., Juliette, R., Ronan, G., et al. (2008). Detection of quantitative trait loci for reproduction and production traits in large white and french landrace pig populations (open access publication). *Genetics Selection Evolution*, 40, 61 - 78.

Van Arendonk, J. A. M., Tier, B. and Kinghorn, B. P., (1994). Use of multiple genetic markers in prediction of breeding values. Genetics 137: 319–329.

Van der Beek, S., and J. A. M. van Arendonk, (1995). Marker assisted selection in an outbred poultry breeding nucleus. Anim. Sci. 62:171–180.

Van Kaam, J. B., Groenen, M. A., Bovenhuis, H., Veenendaal, A., Vereijken, A. L., & Van Arendonk, J. A. (1999). Whole genome scan in chickens for quantitative trait loci affecting carcass traits. *Poultry Science*, 78, 1091-1099.

VanRaden, P.M., and Wiggans, G.R., (1991). Derivation, calculation, and use of national animal model information. J. Dairy Sci. 74:2737-2746.

Visscher, P.M., Thompson R. and Haley, C.S. (1996) Confidence intervals in QTL mapping by bootstrapping. Genetics 143:1013-1020.

Vitezica, Z. G., Marie-Etancelin, C., Bernadet, M. D., Fernandez, X., & Robert-Granie, C. (2010). Comparison of nonlinear and spline regression models for describing mule duck growth curves. *Poultry Science*, 89, 1778-1784.

Wahlberg, P., Carlborg, O., Foglio, M., Tordoir, X., Syvanen, A.-C., Lathrop, M., et al. (2009). Genetic analysis of an f2 intercross between two chicken lines divergently selected for body-weight. *BMC Genomics*, 10, 248.

Wang, T., Fernando, R. L., Van Der Beek, S., Grossman M., and Van Arendonk, J. A. M., (1995). Covariance between relatives for a marked quantitative trait locus. Genet. Sel. Evol. 27: 251–274.

Watanabe, M. (1961). Experimental studie on the artificial insemination of domestic ducks with special reference to the production of mule ducks. *J. Fac. Fish. Anim. Hush. Hiroshima univ.*, 3, 439-478.

Weller, J. I., & Gianola, D. (1989). Models for genetic analysis of dystocia and calf mortality. *Journal of Dairy Science*, 72, 2633-2643.

Weller, J. I., Kashi, Y., & Soller, M. (1990). Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *Journal of Dairy Science*, 73, 2525-2537.

Weller, J. I., Misztal, I., & Gianola, D. (1988). Genetic analysis of dystocia and calf mortality in israeli-holsteins by threshold and linear models. *Journal of Dairy Science*, 71, 2491-2501.

Weller, J. I., & Ron, M. (1992). Genetic analysis of fertility traits in israeli holsteins by linear and threshold models. *Journal of Dairy Science*, 75, 2541-2548.

Weller, J. I., Saran, A., & Zeliger, Y. (1992). Genetic and environmental relationships among somatic cell count, bacterial infection, and clinical mastitis. *Journal of Dairy Science*, 75, 2532-2540.

Weller, J.I., Wiggans, G.R., VanRaden, P.M., Ron, M., (1996). Application of a canonical transformation to detection of quantitative trait loci with the aid of genetic markers in a multi-trait experiment. Theor. Appl. Genet., 92:998–1002.

Woessner, J. F., Jr. (1961). The determination of hydroxyproline in tissue and protein samples containing small proportions of this imino acid. *Arch Biochem Biophys*, 93, 440-447.

Wright, D., Kerje, S., Lundström, K., Babol, J., Schütz, K., Jensen, P., et al. (2006). Quantitative trait loci analysis of egg and meat production traits in a red junglefowl × white leghorn cross. *Animal Genetics*, 37, 529-534.

Wright, S. (1934). Genetics of abnormal growth in the guinea pig. *Cold Spring Harbor Symposia on Quantitative Biology*, 2, 137-147.

Wright, D., Kerje, S., Lundström, K., Babol, J., Schütz, K., Jensen, P., et al. (2006). Quantitative trait loci analysis of egg and meat production traits in a red junglefowl × white leghorn cross. *Animal Genetics*, 37, 529-534.

Xu, C., Li, Z., & Xu, S. (2005). Joint mapping of quantitative trait loci for multiple binary characters. *Genetics*, 169, 1045-1059.

Xu S, A. W. (1996). Mapping quantitative trait loci for complex binary diseases using line crosses. *Genetics*, 143, 1417-1424.

Yi, N., & Xu, S. (1999). Mapping quantitative trait loci for complex binary traits in outbred populations. *Heredity*, 82, 668-676.

Zeng, Z., (1994): Precision mapping of quantitative trait loci. Genetics, 136:1457–1468.

Zeng, Z.-B., Kao, C.-H., Basten, C.J. (1999). Estimating the genetic architecture of quantitative traits. Genet. Res. 74:279-289. Kao et al. (1999)

Zhou, H., Deeb, N., Evock-Clover, C. M., Ashwell, C. M., & Lamont, S. J. (2006). Genome-wide linkage analysis to identify chromosomal regions affecting phenotypic traits in the chicken. I. Growth and average daily gain. *Poultry Science*, 85, 1700-1711.

Zhou, H., Evock-Clover, C. M., McMurtry, J. P., Ashwell, C. M., & Lamont, S. J. (2007). Genome-wide linkage analysis to identify chromosomal regions affecting phenotypic traits in the chicken. Iv. Metabolic traits. *Poultry Science*, 86, 267-276.

Table des figures

Figure 1.1 : Transmission des allèles au marqueur M et au QTL dans une population BC sans recombinaison 15
Figure 1.2 : Protocoles de détection de QTL chez les populations « inbred » sans prise en compte de recombinaison 17
Figure 1.3 : Courbes de distributions des phénotypes en fonction des allèles au marqueur reçu 18
Figure 1.4 : Courbes de distributions des phénotypes en fonction des allèles au QTL reçu 18
Figure 1.5 : Dispositifs expérimentaux impliquant deux générations 21
Figure 1.6 : Dispositifs expérimentaux impliquant trois générations 21
Figure 1.7 : Exemple de profil de vraisemblance (Lander et Botstein, 1989) 26
Figure 1.8 : Relation entre variable discrète et la variable sous-jacente 36

Figure 2.1 : Illustration d'un caractère à seuils avec 3 modalités 56

Figure 3.1 : Courbe de croissance théorique du canard mulard pendant la durée d'élevage 68
Figure 3.2 : Courbe de gavage théorique – modèle « palier tardif » 69
Figure 3.3 : Courbes des moyennes de poids mesurés à différents âges pour chacun des sexes des trois génotypes (Tai et Rouvier, 1998) 70
Figure 3.4 : Mécanismes de la stéatose hépatique chez les palmipèdes (Hermier *et al*., 1999) 72
Figure 3.5 : Différence des poids vif, poids de foie, poids du gras abdominal et poids du Pectoralis major pour les différents génotypes – animaux non gavés (Chartrin *et al*., 2004) 72
Figure 3.6 : Différence des poids vif, poids de foie, poids du gras abdominal et poids du *Pectoralis major* pour les différents génotypes – animaux gavés (Chartrin *et al.,* 2004) 73

Figure 4.1 : Dispositif de détection de QTL 86
Figure 4.2 : Poids moyens par famille de père F1 88
Figure 4.3 : Taux de Triglycéride, de Cholestérol et de Glucose sanguins par famille de père F1 89
Figure 4.4 : Taux de corticostérones par famille de père F1 90
Figure 4.5 : Caractères d'aptitude au gavage par famille de père F1 91
Figure 4.6 : Caractères de qualité du magret par famille de père F1 92
Figure 4.7 : Caractères de qualité du foie par famille de père F1 93
Figure 4.8 : Exemple de la démarche suivie pour arriver au résultat final d'une analyse multi-caractères 127

Table des tableaux

Tableau 1.1 : Nombre de descendants nécessaires pour la détection d'un QTL dans un croisement entre lignées homozygotes (« inbred ») 19
Tableau 1.2 : Nombre de descendants nécessaires pour la détection d'un QTL dans le cadre d'une population « outbred » (cas d'un ensemble de familles de pleins-frères de type « BC» ou « F2 ») 20
Tableau 1.3 : Génotype au QTL conditionnellement à l'information au marqueur (Mackinnon et Georges, 1992) 24
Tableau 1.4 : Espérance des coefficients des composantes de a et d pour toutes les combinaisons de génotypes possibles dans une population F2. D'après Haley et Knott, 1992. 28

Tableau 2.1 : Les caractéristiques des différentes populations simulées 58
Tableau 2.2 : paramètres estimés pour les deux modèles comparés grâce aux données simulées du cas 1 60
Tableau 2.3 : paramètres estimés pour les deux modèles comparés grâce aux données simulées du cas 5 60
Tableau 2.4 : paramètres estimés pour les deux modèles comparés grâce aux données simulées du cas 9 61
Tableau 2.5 : paramètres estimés pour les deux modèles comparés grâce aux données simulées du cas 13 62

Tableau 3.1 : Effet du génotype et du gavage sur la couleur et les pertes à la cuisson des filets de canards âgés de 14 semaines (extrait de Baéza *et al.,* 2005) 75

Tableau 4.1 : Descriptions des caractères de croissances. 87
Tableau 4.2 : Descriptions des caractères liés au métabolisme. 88
Tableau 4.3 : Descriptions des caractères liés au stress 89
Tableau 4.4 : Description des caractères d'aptitude au gavage. 90
Tableau 4.5 : Description des caractères de qualité du magret. 92
Tableau 4.6 : Description des caractères de qualité du foie. 93

Annexes

Annexe A : Les différentes simulations et leurs résultats (discret/normal) :

Cas1 : caractère à deux modalités (50%/50%)
QTL à effet fort situé au milieu du chromosome
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,49	0,49
Ecart-type	-	0,18	0,15
Erreur moyenne quadratique	-	0,03	0,02
puissance (%)	-	68,91	76

Cas2 : caractère à deux modalités (50%/50%)
QTL à effet fort situé à 20cM (extrémité du chromosome)
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,2	0,25	0,24
Ecart-type	-	0,21	0,19
Erreur moyenne quadratique	-	0,05	0,04
puissance (%)	-	70,06	72,00

Cas3 : caractère à deux modalités (50%/50%)
QTL à effet faible situé au milieu du chromosome
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,48	0,5
Ecart-type	-	0,32	0,31
Erreur moyenne quadratique	-	0,10	0,10
puissance (%)	-	8,3	9,20

Cas4 : caractère à deux modalités (50%/50%)
QTL à effet faible situé à 20cM (extrémité du chromosome)
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,2	0,44	0,45
Ecart-type	-	0,32	0,33
Erreur moyenne quadratique	-	0,16	0,17
puissance (%)	-	7,07	9,60

Cas5 : caractère à deux modalités (10%/90%)
QTL à effet fort situé au milieu du chromosome
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,45	0,48
Ecart-type	-	0,22	0,19
Erreur moyenne quadratique	-	0,05	0,04
puissance (%)	-	46,60	52,60

Cas6 : caractère à deux modalités (10%/90%)
QTL à effet fort situé à 20cM (extrémité du chromosome)
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,2	0,26	0,27
Ecart-type	-	0,22	0,23
Erreur moyenne quadratique	-	0,05	0,06
puissance (%)	-	48,83	53,20

Cas7 : caractère à deux modalités (10%/90%)
QTL à effet faible situé au milieu du chromosome
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,45	0,49
Ecart-type	-	0,32	0,31
Erreur moyenne quadratique	-	0,10	0,10
puissance (%)	-	5,91	6,86

Cas8 : caractère à deux modalités (10%/90%)
QTL à effet faible situé à 20cM (extrémité du chromosome)
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,2	0,43	0,47
Ecart-type	-	0,33	0,34
Erreur moyenne quadratique	-	0,16	0,19
puissance (%)	-	7,44	6,28

Cas9 : caractère à trois modalités (30%/30%/40%)
QTL à effet fort situé au milieu du chromosome
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,49	0,5
Ecart-type	-	0,19	0,16
Erreur moyenne quadratique	-	0,04	0,03
puissance (%)	-	59,2	64,8

Cas10 : caractère à trois modalités (30%/30%/40%)
QTL à effet fort situé à 20cM (extrémité du chromosome)
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,2	0,26	0,26
Ecart-type	-	0,20	0,22
Erreur moyenne quadratique	-	0,04	0,05
puissance (%)	-	66,20	65,20

Cas11 : caractère à trois modalités (30%/30%/40%)
QTL à effet faible situé au milieu du chromosome
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,51	0,49
Ecart-type	-	0,30	0,33
Erreur moyenne quadratique	-	0,09	0,11
puissance (%)	-	9,20	9,80

Cas12 : caractère à trois modalités (30%/30%/40%)
QTL à effet faible situé à 20cM (extrémité du chromosome)
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,2	0,45	0,45
Ecart-type	-	0,34	0,32
Erreur moyenne quadratique	-	0,18	0,17
puissance (%)	-	7,60	7,80

Cas13 : caractère à trois modalités (10%/10%/80%)
QTL à effet fort situé au milieu du chromosome
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,50	0,51
Ecart-type	-	0,16	0,23
Erreur moyenne quadratique	-	0,03	0,05
puissance (%)	-	67,40	36,69

Cas14 : caractère à trois modalités (10%/10%/80%)
QTL à effet fort situé à 20cM (extrémité du chromosome)
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,2	0,24	0,31
Ecart-type	-	0,17	0,26
Erreur moyenne quadratique	-	0,03	0,08
puissance (%)	-	67,13	34,52

Cas15 : caractère à trois modalités (10%/10%/80%)
QTL à effet faible situé au milieu du chromosome
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,5	0,49	0,50
Ecart-type	-	0,31	0,33
Erreur moyenne quadratique	-	0,09	0,11
puissance (%)	-	6,80	5,88

Cas16 : caractère à trois modalités (10%/10%/80%)
QTL à effet faible situé à 20cM (extrémité du chromosome)
Héritabilité : 0.35

Paramètres	Vraie valeur	Modèle discret	Modèle normal
Position du QTL (M)	0,2	0,43	0,46
Ecart-type	-	0,32	0,33
Erreur moyenne quadratique	-	0,16	0,18
puissance (%)	-	8,00	2,95