Fine mapping of female fertility qtl in french dairy cattle

Slim Ben Jemaa

▶ **To cite this version:**

Slim Ben Jemaa. Fine mapping of female fertility qtl in french dairy cattle. Life Sciences [q-bio]. AgroParisTech, 2009. English. NNT : 2009AGPT0012 . pastel-00005162

HAL Id: pastel-00005162
https://pastel.hal.science/pastel-00005162

Submitted on 27 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N°/__/__/__/__/__/__/__/__/__/__/

THÈSE

pour obtenir le grade de

Docteur

de

L'Institut des Sciences et Industries du Vivant et de l'Environnement (Agro Paris Tech)

Spécialité : Génétique

présentée et soutenue publiquement
par

Slim Ben Jemaa

Le 05 Mars 2009

CARTOGRAPHIE FINE DE QTL DE FERTILITE FEMELLE CHEZ LES BOVINS LAITIERS FRANÇAIS

FINE MAPPING OF FEMALE FERTILITY QTL IN FRENCH DAIRY CATTLE

Directeur de thèse : ***André Eggen***

Codirecteur de thèse : ***Mathieu Gautier***

INRA, UR339 Laboratoire de Génétique biochimique et de Cytogénétique, F-78352 Jouy en Josas CEDEX.

Devant le jury :

Mme. Sandrine Lagarrigue, Professeur de génétique à Agrocampus de Rennes	**Rapporteur**
M. Gaudenz Dolf, Professeur de génétique à l'université de Bern	**Rapporteur**
Mme. Joelle Dupont, Directeur de recherche, à l'INRA de Tours	**Examinateur**
M. Laurent Journaux, Vice-président de l'UNCEIA	**Examinateur**
M. André Eggen, Directeur de recherche à l'INRA de Jouy en Josas	**Examinateur**
M. Etienne Verrier, Professeur à l'AgroParisTech	**Président**

A ma chère mère, Chérifa

A ma chère sœur Oumayma

Remerciements

Ce travail s'est déroulé au Laboratoire de Génétique biochimique et de Cytogénétique (LGbC) de l'INRA de Jouy-en-Josas. Je tiens donc à remercier Edmond-Paul Cribiu, le directeur, de m'avoir accueilli. Je remercie aussi le département de génétique animale (en particulier le chef de Département, M. Didier Boichard), à l'origine de ce projet de thèse, l'UNCEIA (en particulier son directeur Maurice Barbezant, son directeur adjoint, Alain Malafosse et le responsable du service génétique Laurent Journaux) ainsi que le CASDAR pour la participation au financement du projet ainsi que de ma bourse de thèse.

Je remercie notamment les membres du jury : M. Etienne Verrier de l'Ecole Doctorale ABIES, Mme Sandrine Lagarrigue et M. Gaudenz Dolf d'avoir accepté d'être les rapporteurs de ma thèse, Mme Joëlle Dupont et M. Laurent Journaux d'en être les examinateurs.

Je tiens ensuite à remercier mon directeur de thèse André Eggen pour m'avoir mis à disposition les moyens nécessaires pour que cette thèse se déroule dans les meilleures conditions possibles, pour m'avoir facilité les démarches administratives tout au long de ma thèse et pour son soutien moral. A propos d'administration, je souhaite aussi remercier vivement Patricia et Françoise pour leur aide et leur serviabilité et qui m'ont épargné beaucoup de tracasseries.

Je tiens également à remercier vivement Mathieu Gautier pour son encadrement scientifique (dès mon stage de master) rigoureux, les « petits » scripts qu'il me fournissait, ses conseils scientifiques et ses critiques, parfois dures, mais utiles et constructives. Mathieu, je pense que grâce à toi j'ai appris beaucoup de choses (les bonnes bien entendu) qui me seront utiles dans ma future carrière.

Mes remerciements vont aussi au groupe « Forza QTL » pour les réunions régulières organisées et pour les collaborations constructives : Sébastien Fritz et François Guillaume pour leur coopération et leur aide constante notamment lors du petit stage que j'ai passé à la SGQA, Tom Druet pour son encadrement scientifique et son aide dans le domaine de génétique quantitative.

Je remercie également mon comité de thèse composé de Joëlle Dupont et Thomas Heams, en plus de certaines personnes citées précédemment (Alain Malafosse, André Eggen, Mathieu Gautier, Sébastien Fritz et Tom Druet), qui ont porté un œil critique et qui ont orienté mon travail.

Je remercie particulièrement Thomas Heams, le correspondant de l'école doctorale ABIES pour son soutien moral, ses conseils et son aide et notamment son souci pour que cette thèse se déroule dans les meilleures conditions.

Je remercie également la direction scientifique de l'école doctorale ABIES, en particulier Mme Françoise Launay, de m'avoir facilité les démarches administratives qui étaient nécessaires lors de mon séjour en France.

J'exprime ma gratitude à toute « l'équipe des bovins », pour leur contribution scientifique et la bonne ambiance qui régnait dans le groupe. Merci à Mekki (Mekki ? Makki ? Mikey ?) pour son soutien scientifique et moral constant lors de ce projet de thèse. Merci à mes deux « voisins » qui occupent la même pièce que moi : Aurélien et Catherine pour la bonne ambiance qui s'est créée dans cette pièce. Je remercie Catherine notamment pour son aide dans les manips que je réalisais.

Merci Hélène pour la lecture critique et attentive de mes rédactions en Anglais. Merci au reste du groupe : Amandine, Bernard, Cécile, Maud, Sandrine et la nouvelle recrue Sabrina, pour leur soutien moral.

J'adresse aussi de vifs remerciements aux autres membres du LGbC pour la bonne humeur générale avec une attention particulière à Sead Chadi pour son aide et sa patience lors de l'envoi des séquences.

Mes remerciements vont aussi aux membres de l'équipe de football de l'INRA de Jouy-en-Josas, qui, grâce à eux j'ai passé de très bons moments au cours de mes quatre ans passés à l'INRA, en particulier à notre capitaine Jean-Pierre Bidanel, Sébastien Fritz, Nico1, Nico2, Nico3 (je leur laisse le soin de se mettre d'accord pour le classement des noms) et les autres membres de l'équipe.

Je tiens enfin à remercier toute ma famille proche et mes amis : mes chers parents, ma très, très chère sœur Oumayma et mes deux frères Amine et Sofiène, ma tante Nabiha, son mari Hédi et leurs fils Ghada, Salma et Saif, mes amis : Aimen, Dali, Houssine (les amis et seulement les amis peuvent aussi l'appeler Hassen), Khaled « le Japonais », Mouez, Mourad Saddam (il n'est en aucun cas un des proches de l'ex président Irakien), Naoufel, Nidhal1, Nidhal2, Zied,…ça fait tant d'années qu'on se voit plus comme avant mais notre amitié restera intacte et sans faille.

Merci à tous

Résumé

L'objectif de la thèse est la cartographie fine de QTL de fertilité femelle (FF) chez les races bovines Françaises : la Prim'Holstein, la Normande et la Montbéliarde. *La première étape* du projet consiste en une primo-localisation, sur un dispositif de 78 familles, de QTL de FF dans 12 régions génomiques. Six QTL de FF ont été détectés chez la Prim'Holstein et deux chez la Normande. *La deuxième étape* consiste en une cartographie des QTL détectés sur les chromosomes BTA01, BTA02, BTA03 et BTA21 sur un échantillon de 41 familles des trois races. La localisation de ces QTL a été confirmée et précisée pour les QTL détectés sur les chromosomes BTA01 et BTA03. Le QTL sur le chromosome BTA03 a été finement cartographié en utilisant 437 SNP ce qui a permis de réduire son intervalle de localisation à quelques centimorgans et de sélectionner six gènes candidats. Les régions flanquant les exons de ces six gènes ont été séquencées sur quatre pools d'ADN d'individus de phénotypes extrêmes afin de trouver des polymorphismes intéressants dans les régions séquencées qui pourront ensuite être validés sur l'ensemble du dispositif.

La mise à disposition récente d'une puce de 54000 SNP situés dans tout le génome bovin a changé la stratégie de cartographie de QTL de FF. Les résultats de détection obtenus avec cette puce ont permis de vérifier la sur-représentativité, par rapport à tout le génome bovin, de gènes situés dans les intervalles de localisation des QTL et impliqués dans des fonctions de reproduction. De plus, des gènes pouvant influencer la FF ont été sélectionnés et suggérés comme étant des gènes candidats positionnels et fonctionnels responsables de la dégradation de la FF chez les bovins laitiers.

Mots clés: QTL - Fertilité femelle - Analyse de liaison - gènes candidats

Abstract

The aim of this thesis was to map finely female fertility (FF) QTL in the three main French dairy cattle breeds Prim'Holstein, Normande and Montbeliarde. The first step consisted in detecting QTL on 78 families in 12 genomic regions. Six FF QTL were found in Prim'Holstein and two in Normande breeds. In the second step of the study, the QTL on bovine chromosomes BTA01, BTA02, BTA03 and BTA21 were mapped on a sub-sample of 41 families and the confidence interval for BTA01 and BTA03 QTL was reduced. Using a set of 437 SNP, the QTL location interval on BTA03 was reduced to a few centimorgans allowing us to select six candidate genes, which were partially sequenced for four DNA pools belonging to individuals with extreme phenotypes in order to find some intersting polymorphisms which can be validated on the whole pedigree. Several SNPs corresponding to the a priori status of the individuals for the QTL, were found in most of the candidate genes. However, since a bovine SNP chip with 54 000 SNP became available, no further validation step was carried out.

Based on fine mapping results, overrepresentation, with regard to the whole bovine genome, of genes located in QTL confidence intervals involved in biological functions in connection with reproduction, were checked. These results were also used to identify positional candidate genes that may influence dairy cattle FF.

Key words : QTL - Female fertility - Linkage analysis - candidate genes

Table des matières

Introduction Générale *14*

A. La fertilité chez les bovins laitiers: définition, mécanismes biologiques impliqués et liens avec la sélection pour les caractères de production laitière ---------- 18

I. La fertilité : définition, méthodes de mesure et importance dans la profitabilité d'un troupeau laitier ---------- 19

1. Notion de fertilité chez la vache laitière ---------- 19
2. Cycle de reproduction chez la vache laitière et méthodes de mesure de la fertilité femelle dans un troupeau laitier ---------- 19
   a) Période d'attente volontaire ou PAV ---------- 19
   b) Intervalle [première IA - Conception] ---------- 19
   c) Période de gestation ---------- 20
3. Importance économique de la fertilité femelle dans un troupeau de bovins laitiers ---------- 22

II. Mécanismes biologiques impliqués dans la baisse de fertilité femelle des bovins laitiers et relation avec le niveau de production laitière ---------- 22

1. Description du cycle ovarien chez la vache laitière et des mécanismes physiologiques sous-jacents --- 22
2. Importance de la détection d'œstrus dans la fertilité des vaches laitières ---------- 24
3. Principales composantes de la subfertilité au cours du cycle de reproduction de la vache laitière ------- 25
   a) Anomalies de la reprise de la cyclicité ---------- 25
   b) Pertes de produits de gestation chez les bovins laitiers ---------- 25
4. Liens potentiels entre l'augmentation de la production laitière et la baisse de fertilité chez la vache laitière ---------- 28

III. La baisse de fertilité femelle chez les bovins laitiers: un problème commun à plusieurs pays ---------- 31

IV. Liens entre la sélection pour les caractères de production laitière et la baisse de fertilité femelle chez les bovins laitiers ---------- 35

1. Evolution de la productivité des vaches laitières suite à une sélection intense pour les caractères de production et une optimisation des conditions d'élevage ---------- 35
2. La sélection pour les caractères de production s'est accompagnée d'une baisse de fertilité femelle des bovins laitiers ---------- 35

B. Amélioration de la fertilité par voie génétique ---------- 37

I. Historique de l'amélioration génétique de la FF en France ---------- 38

1. La fertilité : un caractère introduit récemment dans les objectifs de sélection ---------- 38
2. Choix de l'indice de fertilité femelle adéquat ---------- 38

II. Evaluation génétique de la fertilité : Méthodes d'indexation ---------- 39

1. Poids de la fertilité dans les objectifs de sélection de quelques pays ---------- 39
2. Modèles d'évaluation génétique de la FF en France ---------- 39
   a) Modèle d'évaluation mono-caractère ---------- 40
   b) Modèle d'évaluation multi-caractères ---------- 41
3. Indices de fertilité utilisés pour l'évaluation génétique de la FF dans d'autres pays ---------- 42
4. Indexation internationale de la fertilité ---------- 42

III. Limites des méthodes d'indexation actuelles de la fertilité femelle et nécessité d'une approche moléculaire pour l'amélioration génétique des bovins laitiers ---------- 43

1. La fertilité : caractère à faible héritabilité mais à variation génétique importante ---------- 43
   a) Héritabilité des indices classiques de fertilité ---------- 43
   b) Héritabilité d'indices endocriniens liés à la fertilité ---------- 45
   c) La fertilité caractère à variation génétique élevée ---------- 46
2. Difficulté d'amélioration génétique de la FF par les indexes de sélection classiques ---------- 47
   a) Nature des données utilisées pour l'élaboration des index de fertilité classiques ---------- 47
   b) Nature des caractères mesurant la fertilité femelle ---------- 48
3. Conclusion ---------- 48

C. Cartographie de QTL impliqués dans la variabilité génétique des caractères quantitatifs chez les animaux domestiques ---- 50

I. Notion de caractère quantitatif et de QTL ---- 51
1. Définition d'un caractère quantitatif ---- 51
2. Propriétés des caractères quantitatifs ---- 51
   a) Propriétés statistiques ---- 51
   b) Origine de la variation continue des caractères quantitatifs : le modèle infinitésimal ---- 51
   c) Application du modèle infinitésimal : effet d'un seul locus ---- 51
   d) Généralisation du modèle à un locus : le modèle polygénique - Notion de QTL et d'héritabilité 53
3. Effets d'interaction ---- 56
   a) Interaction génotype X génotype (épistasie) : ---- 56
   b) Interaction génotype x sexe ---- 57
   c) Interaction génotype x milieu ---- 57

II. Cartographie de QTL ---- 57
1. L'approche « gène candidat » ---- 57
2. L'approche de cartographie génétique en utilisant les marqueurs neutres ---- 58
   a) Cartographie de QTL par analyse de liaison ---- 58
   b) Cartographie de QTL par analyse d'association ---- 69
   c) Combiner l'étude de liaison et le DL pour la cartographie fine de QTL : l'approche LDLA ---- 70
3. Recherche et validation de mutations ayant un effet sur un caractère quantitatif ---- 71
   a) Notion de QTN ---- 71
   b) Etapes de validation d'un QTN ---- 71

III. Intérêts de la cartographie de QTL chez les animaux domestiques : la Sélection Assistée par Marqueurs (SAM) ---- 74
1. Principe de la SAM ---- 75
2. Types de sélection assistée par marqueurs : ---- 75
   a) SAM de première génération ou SAM1 ---- 75
   b) SAM de deuxième génération ou SAM2 ---- 76
   c) SAM de troisième génération ou SAM3 ---- 76
   d) La sélection génomique ---- 76

D. Matériel et méthodes ---- 78

I. Matériel Animal ---- 79
1. Pedigree d'étude ---- 79
   a) Pedigree de l'étude de primo-localisation ---- 79
   b) Pedigree de l'étude de cartographie ---- 79
2. Données phénotypiques ---- 80

II. Marqueurs moléculaires ---- 81
1. Choix des microsatellites ---- 81
2. Multiplexage des marqueurs pour le génotypage ---- 85
   a) Principe ---- 85
   b) Amplification et visualisation des ADN ---- 85
   c) Mise au point des multiplex et génotypage des individus ---- 86
3. Résultats des génotypages ---- 87
   b) Nombre de génotypages disponibles ---- 88
   c) Nombre de génotypages considérés pour la détection de QTL ---- 88
4. Vérification des erreurs de typage et construction de la carte génétique ---- 91

III. Cartographie des QTL par analyse de liaison ---- 91
1. Modèle d'analyse ---- 91
   a) Détection de QTL ---- 91
   b) Evaluation du statut hétérozygote des pères au QTL ---- 92

IV. Choix des gènes candidats positionnels et fonctionnels sur le BTA03 et sélection d'individus de phénotypes extrêmes ---- 93
1. Choix des régions à étudier sur le BTA03 ---- 93
2. Ancrage de la région à étudier de BTA03 sur les génomes murin et humain ---- 93
3. Sélection des gènes candidats à partir du génome murin ---- 93
4. Sélection et séquençage d'individus de phénotypes extrêmes pour les gènes candidats ---- 95
   a) Sélection des individus à partir de familles ségréguant le QTL ---- 95

b) Amplification de l'ADN génomique des animaux ---- 95
c) Séquençage des individus pour les gènes candidats sélectionnés à partir des résultats de l'analyse LDLA ---- 95

V. Inventaire des gènes candidats positionnels et fonctionnels en se basant sur les résultats de la puce Illumina 50K ---- 97

E. Résultats ---- 99

I. Résultats de l'étude de primo-localisation ---- 100

II. Résultats de l'étude de cartographie ---- 100

1. Chromosome BTA01 ---- 101
   a) Données de génotypage analysées ---- 101
   b) Carte génétique et densité en marqueurs ---- 101
   c) Résultats de l'analyse de liaison ---- 102
2. Chromosome BTA02 ---- 106
   a) Données de génotypage analysées ---- 106
   b) Carte génétique et densité en marqueurs ---- 106
   c) Résultats de l'analyse de liaison ---- 107
3. Chromosome BTA21 ---- 110
   b) Résultats de l'analyse de liaison ---- 111
4. Chromosome BTA03 ---- 114
   a) Données de génotypage analysées ---- 114
   b) Carte génétique et densité en marqueurs ---- 114
   c) Résultats de l'analyse de liaison ---- 115
5. Conclusion ---- 118

III. Cartographie fine de QTL de fertilité femelle ---- 118

1. Cartographie fine de QTL de fertilité femelle sur le BTA03 ---- 118
2. Cartographie fine de QTL de fertilité femelle à travers tout le génome en utilisant la puce Illumina « BovineSNP50 » ---- 119

IV. Etude des gènes candidats positionnels et fonctionnels ---- 122

1. Etude de gènes candidats sur le BTA03 ---- 122
   a) Ancrage de l'intervalle de localisation du QTL sur les génomes, humain et murin ---- 122
   b) Séquençage des gènes candidats sélectionnés ---- 123
   c) Etudes complémentaires ---- 125
2. Etude de gènes candidats sur l'ensemble du génome ---- 126
   a) Réseaux de gènes identifiés ---- 127
   b) Fonctions biologiques surreprésentées par la liste de gènes identifiés avec l'index de fertilité « vache » ---- 127
   c) Fonctions biologiques surreprésentées par la liste de gènes identifiés avec l'index de fertilité « génisse » ---- 128
3. Etude de gènes candidats sélectionnés à partir des résultats de cartographie fine, obtenus par la puce Illumina ---- 128
   a) Inventaire des gènes présents dans les intervalles de localisation des QTL identifiés ---- 128
   b) Sélection de gènes candidats positionnels et fonctionnels dans les intervalles de localisation identifiés ---- 130

F. Discussion et perspectives ---- 132

Liste des abréviations

ABCG2	ATP-Binding Cassette, sub-family G, member 2
ADN	Acide Désoxyribo-Nucléique
Blast	Basic Local Alignment Search Tool
BLUP	Best Linear Unbiased Predictor
BTA	Bos Taurus chromosome
CD	Coefficient de Détermination
CIM	Composite Interval Mapping
cM	centiMorgan
CMH	Complexe Majeur d'Histocompatibilité
CNG	Centre National de Génotypage
CSH	Chromosome Segment Homozygosity
CTIG	Centre de Traitement de l'Information Génétique
DD	Daughter Design
DGAT1	DiacylGlycérol AcylTransferase 1
DL	Déséquilibre de Liaison
DYD	Daughter Yield Deviation
EM:	Expectation Maximisation
FASTA	FAST Alignement
FDR	False Discovery Rate
FF	Fertilité Femelle
GDD	Grand Daughter Design
HSA	Homo Sapiens chromosome
IA	Insémination Artificielle
IAM	Introgression Assistée par Marqueurs
IBD	Identity By Descent
IC	Intervalle de Confiance
IF	Insémination Fécondante
IGF2	Insulin-like Growth Factor 2
INRA	Institut National de la Recherche Agronomique
kb	Kilo paires de Bases
LDLA	Linkage Disequilibrium & Linkage Analysis
LOD	Logarithm Of the Odds
LOD2pt	LOD score 2 point
LRT	Likelihood Ratio Test
MARC	Meat Animal Research center, Clay center, Nebraska, USA
MIM	Multiple Interval Mapping
MMU	Mus Musculus chromosome
PAV	Période d'Attente Volontaire
PCR	Polymerase Chain Reaction
QTL	Quantitative Trait Locus
QTN	Quantitative Trait Nucleotide
RH	Radiation Hybrid
RIL	Recombinant Inbred Line
SAM	Sélection Assistée par Marqueurs
SGQA	Station de Génétique Quantitative et Appliquée
SNP	Single Nucleotide Polymorphism
TNR	Taux de Non Retour
UNCEIA	Union Nationale des Coopératives agricoles d'élevage et d'insémination Animale
USDA	United States Department of Agriculture
UTR	Untranslated Region

Liste des figures

Figure A-1 : Schéma simplifié du cycle de reproduction d'une vache laitière. Le cycle est délimité par deux vêlages successifs et peut être divisé en une période ouverte (période pendant laquelle la vache n'est pas encore gravide) et une période de gestation. ----- 20

Figure A-2 : Représentation schématique de trois vagues de croissance folliculaire et des variations des concentrations d'oestradiol et de FSH dans un cycle ovarien typique de 21 jours. Le premier jour du cycle est celui où l'ovulation a lieu. Chaque vague de croissance folliculaire est caractérisée par la croissance simultanée, sous l'action de la FSH, d'un groupe de follicules primaires. Le nombre de follicules en croissance diminue progressivement jusqu'à ce qu'il ne reste qu'un seul follicule dominant qui, en l'absence du pic préovulatoire de LH, subit une atrésie à son tour. Seul le follicule dominant de la dernière vague folliculaire ovule (figure prise du site de Intervet : http://www.partners-in-reproduction.com/reproduction-cattle/follicular-development.asp). ---------- 23

Figure A-3 : Relation entre le niveau de production laitière quotidien et la durée d'oestrus d'un échantillon de 350 vaches laitières. Le niveau de production laitière quotidien est une moyenne calculée à partir des 10 premiers jours précédant l'œstrus (d'après Lopez et al., 2004). ---------- 24

Figure A-4 : Evolution de deux index génétiques de fertilité femelle (fertilité vache et fertilité génisse) chez les trois principales races bovines laitières Françaises entre les années 1990 et 2001 (Barbat et al., 2007). ---------- 33

Figure C-1 : Valeurs génotypiques et fréquences alléliques des trois classes génotypiques dérivant d'un locus biallélique Q. N_2 est le nombre d'allèles du second allèle Q_2 (a). Décomposition de Fisher : l'axe des abscisses représente N2. L'axe des ordonnées représente la valeur génotypique G. Celle-ci est régressée sur N_2. La droite de régression a pour pente α_2-α_1 et passe par les trois points représentant les valeurs génétiques. Les différences entre les valeurs génotypiques et celles génétiques sont les valeurs de dominance δ. (b) (d'après Falconer, 1996).---------- 53

Figure C-2 : Distribution des fréquences des effets apparents de QTL détectés chez le porc et les bovins laitiers dans une méta-analyse incluant les résultats issus d'expériences de détection de QTL chez ces deux espèces. La distribution des effets de QTL détectés chez le porc montre un plus grand nombre de QTL à effet moyen que de QTL à grand effet. Chez les bovins laitiers, un plus grand nombre de QTL à petit effet- moyen effet est détecté (d'après Hayes et Goddard, 2001).---------- 55

Figure C-3 : Principe de détection de QTL par analyse des effets de ségrégation d'un QTL à deux allèles chez un père hétérozygote sur la distribution des performances de sa descendance. Le suivi de la ségrégation du QTL se fait à travers celle des allèles d'un marqueur moléculaire situé au voisinage du QTL. Les distributions en pointillés présentent l'effet réel du QTL alors que celles en ligne continue présentent l'effet apparent du marqueur lié. La différence entre les deux types de distributions est due aux recombinaisons entre le marqueur et le QTL (Gautier, 2003). ---------- 62

Figure C-4 : Principes de détection de QTL en utilisant la méthode de cartographie par intervalles. Les notations sont conformes à celles utilisées dans le texte (Gautier, 2003). ---------- 63

Figure C-5 : Résultats de détection d'un QTL de naissance gémellaire sur le BTA05 chez la race bovine laitière Norvégienne, obtenus en utilisant trois méthodes différentes : l'analyse de liaison (LA), l'analyse d'association (LD) et la méthode LDLA. Cette dernière méthode a donné la localisation la plus précise en montrant un pic de localisation unique situé à 65 cM environ et un intervalle de localisation inférieur à 1 cM (Meuwissen et al., 2002). ---------- 71

Figure E-1 : Cartes génétiques des marqueurs microsatellites de BTA01 obtenues avec les données expérimentales (a), à partir de la base de données publique USDA (les distances sont exprimées en cM) (b) et carte physique obtenue par blast des microsatellites sur l'assemblage bovin (Btau_4.0) (les distances sont exprimées en Mégabases) (c). Certains marqueurs n'ont pas pu être ancrés sur cet assemblage. L'origine des cartes génétiques ((a) et (b)) est celle de la carte génétique de USDA. - 102
Figure E-2 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur le BTA01 chez la race Prim'Holstein. Les courbes de F-value correspondent aux TNR56 (courbe dessinée avec des carrés), TNR90 (courbe dessinée avec des triangles) et TNR281 (courbe dessinée avec des croix). L'informativité des marqueurs est décrite par la courbe en pointillés (l'axe des ordonnés de l'informativité est celui situé à droite de la figure). Les seuils de significativité correspondants ont été obtenus par 10000 permutations. L'intervalle de confiance est obtenu par bootstrap (10000 échantillonnages). Les marqueurs SAM utilisés lors de la primo-localisation sont indiqués par des flèches.---------- 104
Figure E-3 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur le BTA01 en race Montbéliarde. La légende est identique à celle de la figure E-2. Les seuils de significativité correspondants ont été obtenus par 1000 permutations.------- 104
Figure E-4 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur le BTA01 en race Normande. La légende est identique à celle de la figure E-2. Les seuils de significativité correspondants ont été obtenus par 1000 permutations. ---------- 105
Figure E-5 : Profil des courbes de F-value pour chaque couple de positions sous le modèle bi-QTL sur BTA01 en race Prim'Holstein. Le TNR90 a été utilisé comme phénotype de fertilité femelle. L'origine des abscisses correspond à la position de BM4307 (premier marqueur de la carte génétique utilisée dans l'étude de cartographie). ---------- 106
Figure E-6 : Cartes génétiques des marqueurs microsatellites de BTA02 obtenues avec les données expérimentales (a), à partir de la base de données publique USDA (les distances sont exprimées en cM) (b) et carte physique obtenue par blast des microsatellites sur l'assemblage bovin (Btau_4.0) (les distances sont exprimées en Mégabases) (c). Certains marqueurs n'ont pas pu être ancrés sur cet assemblage. L'origine des cartes génétiques ((a) et (b)) est celle de la carte génétique de USDA. - 107
Figure E-7 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA02 en race Prim'Holstein. Les courbes de F-value correspondent aux TNR à 56 (courbe dessinée avec des carrés), 90 (courbe dessinée avec des triangles) et 281 jours après IA (courbe dessinée avec des croix). L'informativité des marqueurs est décrite par la courbe en pointillés (l'axe des ordonnés de l'informativité est celui situé à droite de la figure). Les seuils de significativité correspondants ont été obtenus par 10 000 permutations. Les marqueurs SAM utilisés lors de la primo-localisation sont indiqués par des flèches.---------- 109
Figure E-8 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA02 en race Normande. La légende est identique à celle de la figure E-7. Les seuils de significativité correspondants ont été obtenus par 1000 permutations. ---------- 109
Figure E-9 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA02 en race Monbéliarde. La légende est identique à celle de la figure E-7. Les seuils de significativité correspondants ont été obtenus par 1000 permutations. ---------- 110
Figure E-10 : Cartes génétiques des marqueurs microsatellites du BTA21 obtenues avec les données expérimentales (a), à partir de la base de données publique USDA (les distances sont exprimées en cM) (b) et carte physique obtenue par blast des microsatellites sur l'assemblage bovin (Btau_4.0) (les distances sont exprimées en Mégabases) (c). CSSM018 n'a pas pu être ancré sur cet assemblage. ---------- 111
Figure E-11 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA21en race Prim'Holstein. Les courbes de F-value correspondent aux TNR à 56

(courbe dessinée avec des carrés), 90 (courbe dessinée avec des triangles) et 281 jours après IA (courbe dessinée avec des croix). L'informativité des marqueurs est décrite par la courbe en pointillés (l'axe des ordonnés de l'informativité est celui situé à droite de la figure). Les seuils de significativité correspondants ont été obtenus par 5000 permutations. Les marqueurs SAM utilisés lors de la primo-localisation sont indiqués par des flèches.---------- 112
Figure E-12 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA21 en race Normande. La légende est identique à celle de la figure E-11. Les seuils de significativité correspondants ont été obtenus par 5000 permutations.------- 113
Figure E-13 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA21 en race Montbéliarde. La légende est identique à celle de la figure E-11. Les seuils de significativité correspondants ont été obtenus par 5000 permutations.------- 113
Figure E-14 : Cartes génétiques des marqueurs microsatellites du BTA02 obtenues avec les données expérimentales (a), à partir de la base de données publique USDA (les distances sont exprimées en cM) (b) et carte physique obtenue par blast des microsatellites sur l'assemblage bovin (Btau_4.0) (les distances sont exprimées en Mégabases) (c). Certains marqueurs n'ont pas pu être ancrés sur cet assemblage. L'ordre des marqueurs de la carte génétique est en accord avec celui du dernier assemblage sauf au niveau de INRA006 et de INRA3015. ---------- 115
Figure E-15 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA03 en race Prim'Holstein. Les courbes de F-value correspondent aux TNR à 56 (courbe dessinée avec des carrés), 90 (courbe dessinée avec des triangles) et 281 jours après IA (courbe dessinée avec des croix). L'informativité des marqueurs est décrite par la courbe en pointillés (l'axe des ordonnés de l'informativité est celui situé à droite de la figure). Les seuils de significativité correspondants ont été obtenus par 10000 permutations. Les marqueurs SAM utilisés lors de la primo-localisation sont indiqués par des flèches.---------- 116
Figure E-16 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA03 en race Normande. La légende est identique à celle de la figure E-15. Les seuils de significativité correspondants ont été obtenus par 1000 permutations.------- 117
Figure E-17 : Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA03 chez la Montbéliarde. La légende est identique à celle de la figure E-15. Les seuils de significativité correspondants ont été obtenus par 1000 permutations.------- 117
Figure E-18 : Résultats de l'analyse LDLA obtenus sur un dispositif de 17 familles Prim'Holstein, avec le panel de 437 SNP et en utilisant le TNR90 comme phénotypes de fertilité femelle. La courbe LDLA montre huit pics situés dans l'intervalle compris entre 11,68 cM et 34,86 cM (Druet et al., 2008). ---------- 119
Figure E-19 : Résultats de l'ancrage de la région [10,93 Mb-15,70 Mb] du BTA03 sur les génomes humains et murin à partir de la base de données UCSC. Cette figure montre que la région étudiée est conservée principalement avec des régions du HSA01 (en couleur marron) et avec des régions du MMU01 et du MMU03 (en couleur marron pour le MMU01 et vert olive pour le MMU03). Les bornes exactes des régions, humaine et murine, orthologues à cette région du BTA03 ont été obtenues à partir de cette même base de données.---------- 123
Figure E-20 : Position des six gènes candidats sélectionnés par rapport à la courbe obtenue par analyse LDLA avec les 437 SNP sur le BTA03 (Druet et al., 2008). Les six gènes candidats sont situés entre le deuxième et le quatrième pic. ---------- 125
Figure E-21 : Représentation graphique des fonctions biologiques les plus surreprésentées dans la liste de gènes identifiée obtenue avec l'index de fertilité « vache ». Le seuil de significativité est indiqué par la ligne orange. Les valeurs de –log (p-value) montrent qu'il n'y a pas de fonctions biologiques significativement surreprésentées dans la liste introduite, par rapport au reste du génome.---------- 127

Figure E-22 : Représentation graphique des 10 fonctions biologiques les plus surreprésentées dans la liste de gènes identifiée obtenue avec l'index de fertilité « génisse ». Le seuil de significativité est indiqué par la ligne orange. Toutes les fonctions biologiques listées dans cette figure sont significativement surreprésentées dans la liste introduite, par rapport au reste du génome. ---------- 128
Figure F-1 : Figure représentant le résultat de détection de QTL de fertilité femelle sur le BTA03 avec la puce 50K en utilisant l'approche LDLA (courbe en bleu). La valeur maximale du LRT, ainsi que sa position et le pourcentage de variation génétique, V(G), expliqué par le QTL sont indiqués sur cette figure. Les deux droites indiquent respectivement, de haut en bas, les seuils à 1% et à 5%. Un $IC_{95\%}$ du QTL détecté avec le panel de microsatellites ($QTL_{microsat}$) et celui détecté avec le panel de 437 SNP (QTL_{437SNP}) sont indiqués avec les flèches doubles. L'intervalle des pics dépassant le seuil de 1% représente le QTL détecté avec la puce 50K (QTL_{50K}). L'axe des abscisses représente la position, en mégabases, sur le BTA03. ---------- 140
Figure F-2 : Représentation schématique simplifiée de l'interaction entre l'axe de reproduction (aire grisée A), et l'axe somatotrope (aire grisée B). Ces interactions se font par l'intermédiaire d'hormones, de métabolites et d'éléments nutritifs (d'après Chagas et al., 2007). ---------- 143

Introduction générale

Les bovins constituent une part importante de l'agriculture « animale » dans la plupart des pays du globe. Depuis leur domestication, il y a 9000 ans (Gautier, 2003), ils approvisionnent les populations humaines en plusieurs produits tels que le lait, la viande, les habits, les fertilisants…

Dans certains pays, en particulier les pays tropicaux, les élevages bovins sont destinés à la production de plusieurs produits à la fois. Dans d'autres pays, la production bovine est divisée en deux secteurs principaux : la production de lait et la production de viande. En Europe, les élevages répondent généralement à un objectif double: les mêmes animaux sont utilisés à la fois pour la production de lait et de viande. Certains états font exception à cette règle (*e.g.* la France, l'Australie, la Nouvelle-Zélande, les USA, le Canada, Israël …), avec des races plus spécifiques pour la production de lait ou de viande et une amélioration génétique des animaux ainsi orientée vers l'un ou l'autre des deux objectifs.

Dans les élevages laitiers, l'objectif principal est de produire du lait de bonne qualité au moindre coût. Pour répondre à cet objectif, les éleveurs jouent principalement sur deux paramètres :

- Le premier paramètre est le maintien, voire l'augmentation du niveau de production global et la réduction du nombre d'animaux dans une étable ; ce qui passe obligatoirement par une augmentation de la productivité individuelle des animaux. Jouer sur l'accroissement de la productivité individuelle est devenu indispensable après l'instauration des quotas laitiers en Europe suite à une situation productive largement excédentaire dont les conséquences se sont traduites par une chute de l'effectif du cheptel laitier accompagnée par une augmentation de la production laitière par vache, dans beaucoup de pays industrialisés en Europe. En France, par exemple, le nombre de vaches laitières est passé de 7 millions environ en 1984 (Gautier, 2003) à 3,6 millions environ en 2006 (Madelin, 2006). La productivité individuelle moyenne a évolué quant à elle de 5823 Kg/an en 1987 à 8823 Kg/an en 2007 (Le Mézec, 2009). Au Royaume Uni, le troupeau laitier, d'environ 2,6 millions de vaches laitières en 1995, s'est réduit à environ 1,9 millions de vaches en 2001 alors que la production individuelle est passée de 5380 l/an en 1995 à 6320 l/an en 2001 (Ball et Peters, 2004).

Le deuxième paramètre est le maintien d'un bon niveau de performance reproductive, crucial pour une production efficace. En effet, un système de production bovin efficace, quel que soit son objectif (lait ou viande), doit maximiser la productivité individuelle de l'animal par unité

d'aliment et par unité de temps. Une reproduction efficace résulte en des intervalles de vêlage optimaux qui, à leur tour, résultent en une production optimale de lait et de veaux par unité de temps. De plus, les animaux présentant un échec à la reproduction sont réformées, limitant ainsi le choix d'animaux disponibles capables de se reproduire, et limitant par conséquent le progrès génétique des caractères d'intérêt pour les éleveurs. La maîtrise de la fertilité femelle (FF) des bovins est donc un élément clé pour garantir la profitabilité et la durabilité de l'élevage. Malgré sa faible héritabilité, des études ont montré que la génétique est responsable de plus d'un tiers de la dégradation de la FF chez la race Prim'Holstein (Le Mézec *et al.,* 2007). L'amélioration génétique de la FF se fait par sélection des reproducteurs sur la base d'indices de valeur génétique ou index qui constituent des prédicteurs de leur niveau de fertilité. Ces indices aident les sélectionneurs à prendre des décisions dans le cadre de plans d'accouplements raisonnés et de schémas de sélection préétablis. Toutefois, les approches utilisées pour l'élaboration des index ignorent les mécanismes moléculaires sous-jacents qui interviennent dans la variabilité de la FF. La connaissance de tels mécanismes est devenue pourtant indispensable pour améliorer l'efficacité de l'amélioration génétique de ce caractère.

L'étude des bases moléculaires des caractères complexes est devenue possible grâce à la découverte de marqueurs moléculaires fréquents et très polymorphes dans le génome. La séquence complète du génome bovin disponible depuis peu, et qui nécessite encore quelques améliorations, est un nouvel outil qui a facilité et facilitera énormément la cartographie génétique de caractères d'intérêt zootechnique chez les bovins. Plusieurs régions génomiques impliquées dans la variation de ces caractères ont déjà été identifiées dans tout le génome (http://genomes.sapac.edu.au/bovineqtl/index.html).

Ces régions génomiques appelées aussi « Quantitative Trait Loci » ou QTL peuvent être utilisées en première approche pour une sélection assistée par marqueurs (SAM) qui combine l'information moléculaire et les indices de valeurs génétiques classiquement utilisés. Cependant, les intervalles de localisation de ces régions QTL sont larges (plusieurs centimorgans voire quelques dizaines de centimorgans) et nécessitent une cartographie plus fine afin d'identifier la (les) mutation(s) causale(s) ce qui permettrait une sélection beaucoup plus efficace, complétant les outils de sélection basés sur les indices de sélection classiques.

L'objectif de ce projet de thèse est de cartographier des QTL de FF, mesurée par le taux de non retour en chaleur (TNR) des vaches après leur insémination, chez les trois principales races bovines laitières Françaises : la Prim'Holstein, la Normande et la Montbéliarde. Ce projet de thèse fait partie d'un programme de cartographie fine de grande envergure dont l'objectif est la détection de QTL relatifs à plusieurs caractères d'intérêt pour les professionnels de l'élevage en vue de les incorporer dans les indices de valeur génétique des

animaux candidats à la sélection. Le programme mis en place exploite la structure familiale préexistante des bovins laitiers, organisée en plusieurs familles de demi- frères de même père. Ce programme associe l'INRA, le GIE LABOGENA et les professionnels de l'élevage représentés par l'UNCEIA (Union Nationale des coopératives Agricoles d'Elevage et d'Insémination Animale). Il a été mis en place suite aux résultats obtenus par un programme de détection de QTL initial (ce programme sera appelé par la suite programme QTL initial) couvrant tout le génome (Boichard *et al.*, 2003) et incluant 14 familles : neuf de la race Prim'Holstein, quatre de la race Normande et une Montbéliarde. A partir des résultats de ce programme QTL initial, 14 régions génomiques ayant un effet sur des caractères fonctionnels et de production ont été sélectionnées et incluses dans le programme SAM.

Dans ce projet de thèse, la cartographie des QTL de fertilité femelle s'est déroulée en trois étapes :

- Une primo-localisation des QTL de FF dans 12 des 14 régions SAM en utilisant des TNR calculés à la fin de la gestation : le TNR281 jours (cette mesure représente le pourcentage de vaches gestantes 281 jours après IA). Cette étude sera appelée ultérieurement « étude de primo-localisation » et a été faite sur un dispositif animal composé de 78 familles appartenant aux trois races laitières citées précédemment. Il est à noter que deux des 14 régions SAM initialement incluses dans le programme n'ont pas été analysées car elles n'ont été génotypées que pour un seul marqueur.
- Une cartographie à moyenne résolution de quatre QTL détectés sur les chromosomes BTA01, BTA02, BTA03 et BTA21 lors de l'étude de primo-localisation. Cette étude sera appelée ultérieurement « étude de cartographie » faite à l'aide de quelques marqueurs microsatellites génotypés sur un échantillon de 41 familles sélectionnées à partir des 78 familles de l'étude de primo-localisation.

Les deux premières étapes (étude de primo-localisation et de cartographie) ont fait appel à des méthodes d'analyse de liaison qui sont des méthodes robustes mais qui aboutissent à des intervalles de localisation de QTL très larges.

- Une cartographie fine utilisant un grand nombre de SNP et faisant appel à une approche de cartographie plus précise (l'approche LDLA pour Linkage Disequilibrium Linkage Analysis), faite d'abord sur le chromosome BTA03 puis sur tout le génome, à l'aide d'une puce contenant 54000 SNP. Des gènes candidats sur le BTA03 ont été sélectionnés et séquencés sur huit individus issus de pères supposés être hétérozygotes au QTL et appartenant à deux catégories opposées de phénotypes extrêmes (quatre individus de chaque catégorie). Le but est de maximiser les chances de trouver des polymorphismes intéressants qui pourraient être validés par la suite sur

l'ensemble du dispositif. Toutefois, la mise à disposition d'une puce Illumina de 54000 SNP a rendu peu rentable l'étape de validation des polymorphismes trouvés dans les gènes candidats du BTA03 puisque la disponibilité de la nouvelle puce Illumina permet une étude des SNP présents dans les régions les plus pertinentes à travers tout le génome bovin.

Une introduction sera consacrée à décrire la baisse de fertilité des vaches laitières : aspects, mécanismes biologiques impliqués et liens avec la sélection pour les caractères de production laitière. Le terme « subfertilité » sera utilisé pour désigner la baisse de fertilité chez une vache laitière puisque le terme « infertilité » veut dire une incapacité permanente et irréversible à se reproduire ce qui ne semble pas être le cas chez la plupart des vaches laitières. La deuxième partie sera consacrée à l'amélioration de la fertilité femelle par voie génétique: les limites des méthodes d'indexation actuelle et la nécessité d'approches plus fines seront discutées. Enfin, une troisième partie sera consacrée à la description des méthodologies utilisées pour la cartographie de QTL.

A. La fertilité chez les bovins laitiers: définition, mécanismes biologiques impliqués et liens avec la sélection pour les caractères de production laitière

I. La fertilité : définition, méthodes de mesure et importance dans la profitabilité d'un troupeau laitier

1. Notion de fertilité chez la vache laitière

La fertilité femelle des bovins est un caractère complexe qui se réfère à tout facteur directement lié à la génisse ou la vache, et pouvant impacter sa probabilité de devenir gestante. Elle mesure la capacité de la génisse ou de la vache à être fécondée. Théoriquement, chaque stade du cycle de reproduction de la femelle est susceptible d'impacter la fertilité de l'animal. Néanmoins, certains stades semblent être plus importants et plus critiques que d'autres pour une bonne fertilité (voir II-3-b de cette partie A).

2. Cycle de reproduction chez la vache laitière et méthodes de mesure de la fertilité femelle dans un troupeau laitier

D'un point de vue zootechnique, le cycle de reproduction d'une vache s'étale sur la période qui sépare deux vêlages successifs ; il peut être décomposé chez la vache laitière en quatre périodes (Figure A-1). L'idéal pour un éleveur est de tendre, pour son troupeau, vers un intervalle entre deux vêlages consécutifs ne dépassant pas les 13 mois car les coûts d'entretien sont très élevés dans les élevages laitiers en particulier dans les systèmes intensifs ; toute augmentation de la période entre deux vêlages consécutifs s'accompagne d'une augmentation de la période d'improductivité des vaches et par conséquent d'une réduction des revenus associés.

a) Période d'attente volontaire ou PAV

En prenant le vêlage comme début du cycle reproducteur d'une vache, la PAV constitue la première phase de ce cycle. C'est la durée de temps minimale que l'éleveur décide d'établir entre le vêlage et la première mise à la reproduction de la vache. Elle est généralement comprise entre 45 et 60 jours. La PAV représente un risque significatif pour le futur état de santé et de productivité de la vache laitière. En effet, c'est au cours de cette période que les risques de pathologies physiologiques telles que la rétention placentaire et la métrite sont les plus élevés chez la vache laitière.

b) Intervalle [première IA - Conception]

Cet intervalle, aussi appelé période de mise à la reproduction (PMR), reflète la rapidité d'une vache à devenir gestante après la première insémination.

La somme des deux intervalles [fin de PAV – première IA] et [première IA - Conception] constitue la période ouverte ou le nombre de jours ouverts.

c) Période de gestation

La durée moyenne de gestation pour la race Holstein est de 282 jours (Guerrier *et al.,* 2007) Les vaches portant des jumeaux ont, en moyenne, une période de gestation réduite de six à 10 jours (Nielen *et al*., 1989; Ryan et Boland, 1991). Environ, deux mois avant le vêlage, la vache entre en période de tarissement pendant laquelle la glande mammaire est mise au repos pour préparer la prochaine lactation. La période de gestation se termine par le vêlage.

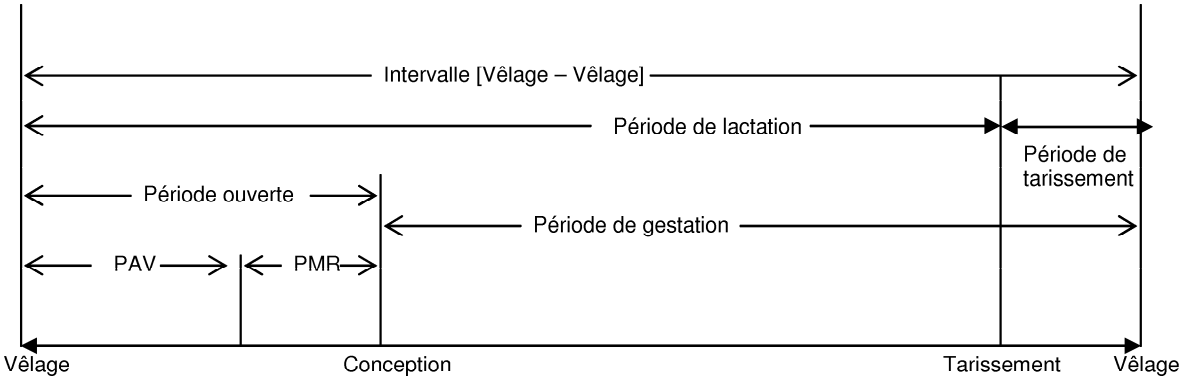


Figure A-1 : *Schéma simplifié du cycle de reproduction d'une vache laitière. Le cycle est délimité par deux vêlages successifs et peut être divisé en une période ouverte (période pendant laquelle la vache n'est pas encore gravide) et une période de gestation.*

Le suivi des performances de reproduction de son troupeau laitier est assuré par des indices de fertilité qui lui permettent de se fixer des objectifs reproductifs réalistes, d'identifier les problèmes reproductifs à un stade précoce et de tracer l'historique des problèmes de reproduction dans son troupeau. La plupart des indices de fertilité sont des performances moyennes calculées sur l'ensemble des vaches du troupeau. Ces indices sont classés principalement en deux catégories: les indices d'intervalle (étroitement liés aux quatre périodes définies précédemment) et les indices de taux de réussite à l'IA. Les indices d'intervalle les plus importants sont :

- L'intervalle [vêlage – vêlage] : c'est le nombre de jours entre deux vêlages successifs. Un IVV de 12 – 13 mois est idéal pour la plupart des troupeaux laitiers. Ce paramètre est cependant très global et reflète seulement le succès reproductif puisqu'il ne tient pas compte des échecs de reproduction.
- L'intervalle [vêlage – 1ère IA] : c'est le nombre de jours entre le dernier vêlage et la première IA. Il regroupe la PAV et l'intervalle [fin de PAV – première IA].
- L'intervalle [vêlage – conception] appelé aussi nombre de jours ouverts ou période ouverte: c'est le nombre de jours entre le dernier vêlage et la conception. Pour une vache, ils représentent l'intervalle [vêlage - conception] et regroupent donc la PAV, l'intervalle [fin de PAV – première IA] et l'intervalle [première IA - Conception]. Le

cas idéal (pour avoir un IVV ne dépassant pas les 13 mois), est d'avoir un nombre de jours ouverts compris entre 85 et 115 jours.

La deuxième catégorie d'indices comprend les indices de taux de réussite à l'IA dont les plus utilisés sont :

- Le taux de conception à un rang d'IA donné : c'est une mesure de l'efficacité de la conception. Il représente le rapport entre le nombre de vaches qui ont été diagnostiquées comme gestantes et le nombre total d'un rang spécifique d'IA. Le taux de conception à la première IA, par exemple, représente le rapport entre le nombre de vaches qui ont conçu et le nombre total de premières IA.
- Le nombre de services par conception : c'est une mesure de la fertilité des vaches inséminées avec succès et donc devenues gestantes. Ce paramètre est calculé en divisant le nombre total d'inséminations réalisées dans le troupeau par le nombre de vaches gestantes. Un troupeau laitier ayant un bon niveau de fertilité doit avoir un nombre de services par conception inférieur à 1,75 ; cependant des valeurs inférieures à 2 constituent un niveau de fertilité acceptable.
- Le taux de non retour en chaleur ou TNR : cet indice de fertilité a été défini par le comité international d'enregistrement des performances animales comme étant le pourcentage de femelles inséminées pour la première fois pendant une certaine période (un mois par exemple) et qui n'ont pas eu une deuxième insémination au cours d'un certain nombre de jours (par exemple 21, 28, 56, 90 jours) (International commitee for animal recording-ICAR, 2006). L'estimation du TNR à 21 jours ou à 28 jours après l'IA, n'est pas sans inconvénient. En effet, étant donné que seulement 50% des vaches qui ont une première ovulation après vêlage montrent des signes d'œstrus (Ball et Peters, 2004), le TNR à 21 ou à 28 jours tend à surestimer les taux de gestation. Des TNR calculés à 56 ou à 90 jours après IA sont plus fiables puisqu'ils permettent de mieux estimer le pourcentage de vaches gestantes.
- Le taux de gestation des filles ou « Daughter Pregnancy Rate » (DPR). Ce taux mesure le pourcentage de vaches gestantes tous les 21 jours. Il est fonction des jours ouverts et de la PAV et il est calculé selon la formule: DPR= 21/ (jours ouverts - PAV + 11).

Ces indices de fertilité mesurent des composantes différentes de la fertilité ; ainsi certains mesurent l'aptitude de la vache à revenir en chaleur tels que l'intervalle [vêlage – 1ère IA] tandis que d'autres traduisent l'aptitude de la vache à être gestante tels que le taux de non retour à 56 jours après 1ère IA. D'autres critères sont plus globaux et combinent les deux composantes à la fois tels que l'IVV ou les jours ouverts.

3. Importance économique de la fertilité femelle dans un troupeau de bovins laitiers

L'impact économique de performances de reproduction non optimales entraîne des pertes liées à :

- Des coûts supplémentaires directement observables par les dépenses d'IA supplémentaires, les interventions curatives et les préventions mises en place.
- Un allongement de l'IVV entraînant :
  - une diminution de la productivité en veau par vache et par an.
  - Une diminution de la productivité laitière par vache et par an qui varie en fonction de l'intensité de l'allongement de l'IVV.
  - Un déplacement de la période de vêlage chez les vaches laitières qui conduit à produire plus en période défavorable du calendrier fourrager.
- Des réformes liées aux problèmes de reproduction. Ces réformes génèrent une part importante des pertes attribuées aux problèmes de reproduction et représentent une part importante des motifs de réforme mentionnés par les éleveurs.

II. Mécanismes biologiques impliqués dans la baisse de fertilité femelle des bovins laitiers et relation avec le niveau de production laitière

1. Description du cycle ovarien chez la vache laitière et des mécanismes physiologiques sous-jacents

Dans ce paragraphe, le cycle ovarien et la période de gestation chez la vache ainsi que quelques mécanismes hormonaux clés sous-jacents seront abordés ; toutefois, détailler ces mécanismes n'est pas l'objet de ce manuscrit.

Le cycle ovarien chez la vache laitière dure en moyenne 21 jours. L'ovulation est souvent considérée comme le début du cycle. Un cycle ovarien est formé de deux à trois vagues de croissance et d'atrésie folliculaire (Fortune *et al*., 2001). Une vague de croissance folliculaire est caractérisée par le développement simultané, sous l'action d'une hormone hypophysaire, la FSH (follicle-stimulating hormone ou hormone folliculostimulante), d'un groupe de follicules primaires parmi lesquels un seul follicule, le follicule dominant, va continuer sa croissance pendant que les autres régressent (Figure A-2). Les vagues folliculaires émergent aux jours 2 et 11 du cycle pour les animaux présentant deux vagues et aux jours 2, 9 et 16 pour ceux présentant trois vagues (Evans, 2003). Les vaches laitières n'ont généralement que deux vagues folliculaires par cycle (Taylor et Rajamahendran, 1991; Townson *et al*., 2002) ;

alors que les génisses et les vaches allaitantes présentent deux ou trois vagues par cycle (Savio *et al.*, 1988; Ginther *et al.*, 1989). Townson *et al.* (2002) ont trouvé un taux de gestation plus élevé (P<0,058) chez les vaches ayant trois vagues par rapport à celles n'ayant que deux vagues de croissance folliculaire (Townson *et al.*, 2002). Inskeep et Dailey (2005) expliquent cette observation par le fait que les follicules qui se développent dans un cycle à deux vagues sont plus « vieux » que ceux se développant dans un cycle à trois vagues. Le processus de vieillissement dans ce contexte est un terme physiologique qui inclut la reprise de la méiose et probablement l'apoptose en attendant l'ovulation (Inskeep et Dailey, 2005).

Le follicule dominant continue à se développer et à produire, sous l'influence de la FSH, des quantités croissantes d'œstrogènes qui vont induire un pic de LH (luteinizing hormone ou hormone lutéinisante) nécessaire à l'ovulation (Ball et Peters, 2004). Les oestrogènes sont également impliqués dans le déroulement de l'œstrus de la vache (Galina et Orihuela, 2007).

L'ovulation, caractérisée par l'expulsion de l'ovocyte à partir du follicule dominant, annonce la fin de la première phase du développement folliculaire : la phase folliculaire et le début d'une deuxième phase : la phase lutéale. Au cours de cette phase, le follicule dominant (dépourvu de l'ovocyte) est transformé en corps jaune (ou corps lutéal) qui assure principalement la production de la progestérone : une hormone clé dans la reproduction qui intervient principalement dans le développement de l'endomètre et la survie de l'embryon en cas de fécondation (Webb *et al.*, 2002).

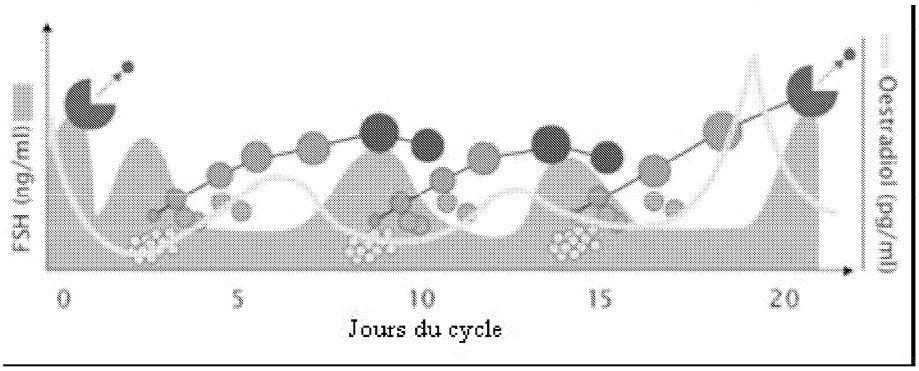


Figure A-2 : *Représentation schématique de trois vagues de croissance folliculaire et des variations des concentrations d'oestradiol et de FSH dans un cycle ovarien typique de 21 jours. Le premier jour du cycle est celui où l'ovulation a lieu. Chaque vague de croissance folliculaire est caractérisée par la croissance simultanée, sous l'action de la FSH, d'un groupe de follicules primaires. Le nombre de follicules en croissance diminue progressivement jusqu'à ce qu'il ne reste qu'un seul follicule dominant qui, en l'absence du pic préovulatoire de LH, subit une atrésie à son tour. Seul le follicule dominant de la dernière vague folliculaire ovule (figure prise du site de Intervet : http://www.partners-in-reproduction.com/reproduction-cattle/follicular-development.asp).*

2. Importance de la détection d'œstrus dans la fertilité des vaches laitières

L'œstrus constitue la seule manifestation comportementale nette du cycle ovarien. Dans la plupart des troupeaux laitiers, les éleveurs l'utilisent comme repère pour choisir le moment adéquat de l'insémination de leurs vaches. L'œstrus chez la vache laitière dure en moyenne 17 heures ± 5 heures et précède l'ovulation de 12 à 15 heures environ. Wiltbank *et al.* (2005) montrent que la durée de l'œstrus d'une vache laitière est influencée par son niveau de production laitière. Ils constatent que, pour les vaches laitières hautes productrices, la durée d'œstrus est plus courte ($P<0,001$) que pour les vaches ayant une production laitière plus faible. La figure A-3 présente la relation entre la durée de l'œstrus et le niveau de production laitière.

Un délai de quelques heures est nécessaire entre le moment de l'ovulation et celui de l'IA. Ce délai permet la maturation des spermatozoïdes dans le tractus génital femelle. Asdell (1964) rapporte que l'IA a plus de chances de réussir si elle est réalisée entre la deuxième moitié de la période d'œstrus et six heures après la fin de l'œstrus.

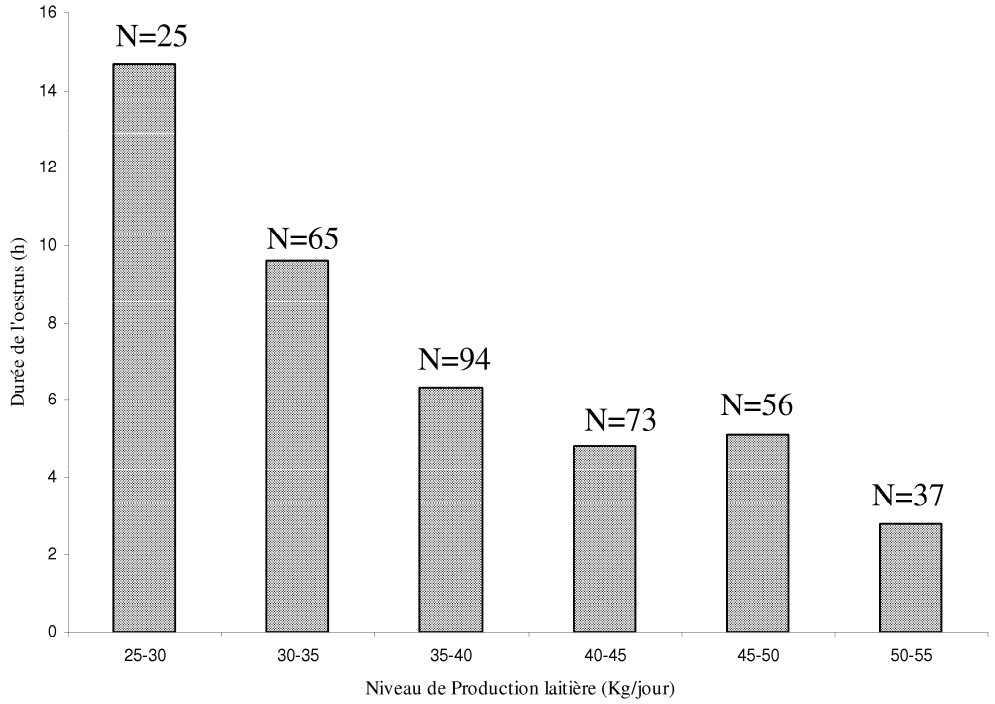


Figure A-3 : *Relation entre le niveau de production laitière quotidien et la durée d'oestrus d'un échantillon de 350 vaches laitières. Le niveau de production laitière quotidien est une moyenne calculée à partir des 10 premiers jours précédant l'œstrus (d'après Lopez et al., 2004).*

3. Principales composantes de la subfertilité au cours du cycle de reproduction de la vache laitière

Outre les problèmes de gestion contribuant à l'augmentation de l'IVV et par conséquent à la baisse de fertilité au niveau du troupeau laitier, plusieurs problèmes majeurs liés à la physiologie de la reproduction ainsi que l'impact du niveau de production laitière sur la reproduction de la vache laitière *post partum* peuvent être responsables d'une subfertilité (ou fertilité réduite). Un premier type de problèmes concerne la reprise de la cyclicité alors qu'un deuxième type concerne la période de gestation.

a) Anomalies de la reprise de la cyclicité

La période qui suit la parturition est une période critique pour la suite de la vie reproductive de la vache. En effet, au cours des quatre premières semaines *post partum*, le système immunitaire de la vache est très sollicité. La plupart des vaches développent une légère endométrite non pathologique (Lewis, 1997) et la capacité de l'utérus à prévenir et/ou à gérer les infections serait compromise par cette endométrite (Thatcher *et al*., 2006). Pendant cette période, les vaches laitières sont également sujettes à une absence de comportement d'œstrus (anœstrus) et/ou absence d'ovulation (anovulation). Les vaches qui ne sont pas observées en œstrus sont souvent considérées comme ayant une anomalie de leur activité ovarienne même si ce n'est pas toujours le cas puisque, dans un troupeau laitier, il peut toujours y avoir des ovulations silencieuses. Les cycles ovariens anormaux sont très fréquents dans les troupeaux laitiers et surtout chez les vaches laitières hautes productrices. Leur prévalence peut atteindre 50% dans certains troupeaux (Lucy, 2007). Plusieurs études montrent que les vaches ayant des cycles ovariens anormaux ont également des performances reproductives inférieures à celles ayant des cycles normaux (Mann *et al*., 2005; Petersson *et al*., 2006).

D'autre part, la rapidité de la reprise de la cyclicité est corrélée à la fertilité des vaches laitières. En effet, une association positive entre la précocité de rétablissement de l'ovulation et le taux de conception a été démontrée par plusieurs études (Butler *et al*. 2001; Darwash *et al*., 2001). Il semblerait que la minimisation de l'intervalle [vêlage -1ère ovulation] donne plus de temps à la vache pour accomplir plusieurs cycles ovariens avant l'IA ; ce qui, à son tour, améliore le taux de conception (Butler *et al*., 1989).

b) Pertes de produits de gestation chez les bovins laitiers

Le comité pour la standardisation de la nomenclature de la reproduction chez les bovins (1972) a proposé de segmenter la période de gestation en deux stades différents : un stade embryonnaire qui s'étend depuis la conception jusqu'au 42ème jour de gestation et un stade fœtal allant du 42ème jour jusqu'à la mise bas. Les risques d'échec de gestation ne sont pas forcément les mêmes entre ces deux stades : ils peuvent également varier à l'intérieur d'un

même stade de gestation. Selon l'ordre chronologique des évènements et l'importance des pertes de produits de gestation, on distingue :

(1) Les pertes de produits de gestation au cours de la fécondation

A ce stade de la gestation, la plupart des études réalisées chez les bovins laitiers rapportent des taux de fécondation élevés : estimé à 95% par Ayalon en 1978 (Ayalon, 1978) et variant entre 90 et 100% d'après Sreenan et Diskin (1986) (Sreenan et Diskin, 1986).

Santos *et al*. (2004) signalent une moyenne de 75% chez les bovins allaitants, avec un rang variant entre 60 et 100%. Chez les vaches laitières, des taux de fécondation similaires de l'ordre de 76,2% (variant entre 55,3 et 87,8%) ont également été rapportés. Les mêmes auteurs indiquent que les taux de fécondation chez les vaches sont similaires à ceux obtenus chez les génisses.

Sartori *et al*. (2002) aboutissent aux mêmes conclusions quant à une indépendance du taux de fécondation du niveau de production laitière (Sartori *et al*., 2002) en absence de stress thermique.

(2) Les pertes de produits de gestation au cours de la période embryonnaire précoce (0-24 jours après IF) :

C'est au cours de cette période que les pertes de produits de gestation sont les plus importantes. La mortalité embryonnaire est la plus élevée au cours des quatorze premiers jours de gestation; elle compte 40% de l'ensemble des pertes de produits de gestation (Dunne *et al*., 2000, Silke *et al*., 2002).

En France, Humblot (2001) a estimé, à partir de 44 troupeaux, un taux de mortalité embryonnaire précoce de l'ordre de 32%. Cette valeur ne tient pas compte du niveau de production laitière de la vache et de son stade physiologique. Ces deux facteurs sont vraisemblablement déterminants dans la viabilité du produit de fécondation ; ainsi, des moyennes de viabilité de 50, 57,9 et de 71,9% à 5-6 jours post-IA, ont été rapportées respectivement, chez les vaches laitières en lactation, celles qui ne sont pas en lactation et les génisses (Santos *et al*., 2004). De même, Sreenan et Diskin (1986) soulignent, à partir d'une synthèse de quatre études, qu'entre 17 et 28% des embryons ont une qualité médiocre dès le septième jour après insémination chez les vaches ayant une production laitière élevée ou moyenne alors que, 6% seulement des embryons ont une qualité médiocre chez les génisses.

(3) Pertes de produits de gestation au cours de la période embryonnaire tardive (25-50) jours après IF

La mortalité embryonnaire au cours de cette période est élevée mais moins marquée que celle de la période embryonnaire précoce. Le taux de mortalité embryonnaire à ce stade a été estimé à 12,8% avec des variations de 3,2% à 42,7%. Toutefois, le taux de mortalité de 42,7% a été obtenu chez des vaches laitières hautes productrices soumises à un stress thermique (Santos *et al*., 2004). Humblot (2001) rapporte un taux de mortalité embryonnaire tardif de 14,7%.

Sur le plan pratique, il est très difficile de situer le moment exact de la perte embryonnaire ; par conséquent, les causes de mortalité embryonnaire qui seront évoquées par la suite se rapportent à l'ensemble du stade embryonnaire (précoce et tardif). Ces causes sont d'origines multiples et dépendent à la fois de la capacité de l'embryon à se développer et de l'adéquation du milieu utérin maternel.

La capacité de l'embryon à se développer dépend principalement du bon déroulement du processus de la formation de follicules et de la fonction lutéale du corps jaune. Les embryons issus de la fécondation de vieux oocytes (ceux issus de deux vagues de croissance folliculaire par exemple) ou ceux issus d'oocytes ayant ovulé tôt après le vêlage (caractérisés par une faible production de progestérone) sont associés à un taux de mortalité embryonnaire précoce élevé. De même, les ovules de petite taille sont associés à un taux de mortalité embryonnaire plus élevé (Perry *et al*., 2005; Perry *et al*., 2007; Mussard *et al*., 2007).

Le milieu utérin maternel constitue la deuxième origine de mortalité embryonnaire et ceci à travers deux voies :

- ✓ *Une voie hormonale* : les dérégulations de la sécrétion des hormones stéroïdes et/ou celles métaboliques sont fréquentes chez la vache laitière *post partum*. Par exemple, l'effet d'un excès d'œstradiol sur la mortalité embryonnaire durant la reconnaissance maternelle de la gestation a été démontré par plusieurs études (MacMillan *et al*., 1986; Thatcher *et al*., 1989; Pritchard *et al*., 1994). De même, la progestérone est essentielle au développement normal de l'embryon. Green *et al*. (2005) ont démontré que la capacité de développement de l'embryon est étroitement liée à la concentration plasmatique de progestérone au cours des cinq premiers jours suivant l'insémination.
- ✓ *Une voie immunitaire* du fait de l'activation du système immunitaire suite aux pathologies très fréquentes après le vêlage. Gilbert *et al*. (2005) ont trouvé que, dans un effectif de 141 vaches Holstein issues de cinq troupeaux laitiers différents, 53% des animaux présentent des symptômes d'endométrite (inflammation utérine)

entre 40 et 60 jours *post partum*. Les vaches ayant une endométrite ont un taux de réussite à la première IA plus faible et nécessitent plus d'IA pour concevoir. Un lien a été aussi établi entre les mammites et la mortalité embryonnaire précoce (Chebel *et al.*, 2004). Les mécanismes par lesquels l'activation du système immunitaire et les réponses inflammatoires engendrent la mort embryonnaire ne sont pas encore clairs. Hansen *et al.* (2004) postulent que les cytokines ou les produits d'activation des cytokines peuvent entraîner la perte embryonnaire à travers l'hyperthermie (la température élevée bloquant le bon fonctionnement de l'oocyte et le développement embryonnaire), des effets toxiques sur le corps jaune et une interférence avec la maturation des oocytes et le développement embryonnaire. Les effets de l'activation du système immunitaire peuvent s'exercer au niveau de l'axe hypothalamo-hypophysaire, de l'ovaire, de l'utérus ou de l'embryon.

(4) Les pertes fœtales

Les pertes fœtales sont généralement beaucoup plus faibles que celles ayant lieu pendant la période embryonnaire précoce et tardive. Des études faites sur les vaches laitières et allaitantes montrent que les pertes fœtales sont variables, mais généralement faibles lorsque le diagnostic de la gestation se fait après 35 jours de gestation. Ces pertes sont inférieures à 10,7% chez les vaches laitières alors qu'elles sont de l'ordre de 4,2% chez les génisses. Environ la moitié des pertes fœtales est d'origine inconnue, l'autre moitié étant due à des lésions du fœtus et à des agents infectieux (Santos *et al.*, 2004).

4. Liens potentiels entre l'augmentation de la production laitière et la baisse de fertilité chez la vache laitière

La corrélation génétique négative entre caractères de production et différents indices de mesure de fertilité (voir II-2 page 22) ne reflète pas obligatoirement une relation de cause à effet. En effet, l'existence d'une corrélation génétique entre production et fertilité ne signifie pas nécessairement qu'un processus physiologique de production laitière interagit directement avec la reproduction. Une hypothèse largement répandue est celle d'une influence de la production laitière sur le bilan énergétique de la vache laitière lorsque l'ingestion alimentaire ne peut pas satisfaire ses besoins d'entretien et de lactation. En effet, l'évolution spectaculaire de la capacité de production individuelle des vaches laitières au cours des dernières décennies nécessite des quantités d'énergie importantes pour assurer la production laitière au détriment des besoins d'entretien et de reproduction.

Les besoins énergétiques de la vache laitière dépassent nettement l'énergie fournie par l'alimentation pendant la fin de la gestation et le début de lactation. La mobilisation du tissu adipeux de la vache est insuffisante pour couvrir les besoins énergétiques de l'animal, ce qui résulte en une période de bilan énergétique négatif (BEN). La sévérité et la durée du BEN sont influencées par le potentiel génétique de production de l'animal (Veerkamp *et al*., 1994; Buckley *et al*., 2000 ; Pryce *et al*., 2001 ;Veerkamp *et al*., 2003 ; Wathes *et al*., 2007).
Le BEN chez la vache laitière est estimé par son état corporel. Une échelle allant de 1 à 5 est utilisée pour l'appréciation de l'état corporel de la vache laitière. Le BEN agit sur la fertilité femelle des bovins laitiers en influençant:

(1) La cyclicité de la vache laitière :

Plusieurs études soulignent, en effet, que le BEN est associé à une plus grande incidence des cycles anormaux pouvant augmenter l'intervalle au premier service en retardant l'ovulation chez la vache laitière durant les trois à quatre premières semaines *post partum* (Butler et Smith, 1989; Beam *et al*., 1999) et en réduisant le taux de conception (Wathes *et al*., 2007).

(2) Effets sur la qualité de l'oocyte et sur le développement embryonnaire

Snijder *et al*. (2000) ont trouvé que les oocytes fécondés in vitro et issus de vaches ayant un faible score d'état corporel ont un taux de division et de développement inférieur à ceux provenant d'individus ayant un meilleur score d'état corporel.
L'influence du BEN sur la qualité de l'oocyte se fait principalement à travers deux voies principales :

- *Une voie endocrinienne :*

La capacité de développement de l'oocyte est intrinsèquement liée à la phase de croissance et à l'état du follicule. Le BEN entrave le processus de croissance folliculaire à travers l'axe hypothalamus–hypophyse–ovaire (Beam *et al*., 1997; Gong, 2002a; Lucy, 2003; Webb *et al*., 2004). L'une des voies les plus affectées est la dérégulation de la sécrétion de LH. Le pic préovulatoire de LH est un signal critique pour le passage de l'oocyte de la métaphase I à la métaphase II (Hyttel *et al*., 1997; Rizos *et al*., 2002; Humblot *et al*., 2005). D'autre part, le pic préovulatoire de LH et la fréquence des pulses de LH sont importants pour la maturation du cytoplasme de l'oocyte (Lindsey *et al*., 2002). Le BEN agit également sur la concentration de progestérone dans le sang ; celle-ci augmente progressivement durant les deux ou trois premiers cycles après vêlage. L'augmentation de la concentration de progestérone est diminuée/modérée par le BEN au début de la lactation. La progestérone est impliquée dans le processus de polyadénylation de l'ARNm maternel ; cette hormone est donc importante dans

la régulation de l'expression de gènes responsables du développement de l'oocyte (Leroy *et al.*, 2008).

- *Une voie métabolique :*

Le BEN influence la concentration sanguine de plusieurs métabolites qui assurent le lien entre la nutrition et la reproduction. Le glucose fait partie de ces métabolites. Le BEN est à l'origine de la diminution de la concentration sanguine en glucose. L'hypoglycémie chez la vache laitière agit :

- *Directement sur la qualité de l'oocyte* : Dans les cellules du cumulus, le glucose est métabolisé en pyruvate et en lactate, deux substrats pour la production d'ATP (Cetica *et al.*, 2002 ; Leroy *et al.*, 2008).

  D'autre part, le glucose est métabolisé dans l'oocyte pour la synthèse de l'ADN et de l'ARN (Sutton *et al.*, 2003; Leroy *et al.*, 2008).
- *Indirectement sur la qualité de l'oocyte* : En influençant la sécrétion de LH et la réponse ovarienne aux gonadotrophines (FSH et LH).

Le BEN influence la reproduction également à travers des hormones dites « métaboliques ». Celles-ci subissent des changements drastiques de leur concentration sanguine suite au BEN. Deux hormones métaboliques très importantes sont : l'insuline et l'IGF-I. Leur concentration diminue durant le BEN. L'hormone de croissance (GH) est une troisième hormone métabolique importante dont la concentration augmente suite au BEN. L'insuline et l'IGF-I influencent la sécrétion de LH et de GnRH et la sensibilité de la réponse ovarienne à LH et FSH (Lucy, 2003). De même, une faible concentration de IGF-I est à l'origine de l'augmentation de la concentration sanguine en GH. L'augmentation de la concentration de GH entraîne un accroissement de la gluconéogenèse dans le foie et de la lipolyse des tissus adipeux (entraînant une accumulation des acides gras non estérifiés ou AGNE). L'augmentation de la concentration des AGNE et de GH, entraîne une sorte de résistance à l'action de l'insuline qui empêche l'utilisation du glucose par les tissus non mammaires et oriente le glucose exclusivement vers la synthèse de lait (Lucy, 2007). Des concentrations élevées d'AGNE prédisent une dégradation de la fertilité chez les vaches laitières *post partum* (McDougall *et al.*, 2005; Lucy, 2007).

(3) Effets sur le statut sanitaire de la vache laitière : développement de pathologies métaboliques

La mobilisation intense des réserves lipidiques de la vache laitière, pour remédier aux besoins d'entretien, se traduit par une libération des AGNE à partir du tissu adipeux. Les AGNE sont capturés par le foie et seront totalement oxydés en dioxyde de carbone, qui servira comme

source d'énergie ou seront partiellement oxydés en corps cétoniques ou en acétate transporté pour être utilisé ailleurs dans le corps. La forme la plus fréquente des corps cétoniques est le béta-hydroxybutirate (BHB) et sa concentration dans le sang est un indicateur du degré d'oxydation des acides gras. L'augmentation de la concentration des corps cétoniques entraîne la cétose, une maladie métabolique très fréquente chez la vache laitière. En effet, 50% des vaches laitières sont sujettes à une cétose subclinique au cours du premier mois de lactation (Wathes *et al*., 2007). Les AGNE peuvent aussi être estérifiés en triacylglycérols (TAG) qui vont s'accumuler dans le foie pendant un certain temps avant d'être exportés. Par conséquent, il peut y avoir une accumulation des TAG dans le foie qui peut aboutir au déclenchement d'une pathologie métabolique : l'infiltration lipidique. Ainsi de fortes concentrations de TAG sont associées à des intervalles [vêlage-1ère ovulation] plus longs et une fertilité réduite (Butler et Smith, 1989; Rukkwamsuk *et al*., 1999; Jorristma *et al*., 2000).

III. La baisse de fertilité femelle chez les bovins laitiers: un problème commun à plusieurs pays

En France, les données d'IA sont centralisées dans le système d'information génétique Français (SIG) depuis 1995. Entre 1997 et 2006, plus de 37 millions de données d'IA ont ainsi été récoltées dans toute la France chez les trois principales races bovines laitières : la Prim'Holstein, la Normande et la Montbéliarde (Tableau A-1).

La figure A-4 montre l'évolution des index de fertilité femelle des taureaux d'IA Français agréés, sur une période de 12 années (1990-2001) pour les races Prim'Holstein, Normande et Montbéliarde et pour deux index de fertilité différents : index de fertilité vache et index de fertilité génisse. Les deux index de fertilité sont exprimés en % de réussite à une IA (quelque soit le rang de l'IA). L'évolution en dents de scie d'une année à l'autre est une caractéristique commune aux deux types d'index de fertilité (vache et génisse) chez les deux races Normande et Montbéliarde. Les index de fertilité vache chez la race Prim'Holstein ont, par contre, tendance à baisser progressivement. La variation de la courbe relative à l'index de fertilité génisse est, quant à elle, moins marquée et elle est en grande partie négative à partir de l'année 1993.

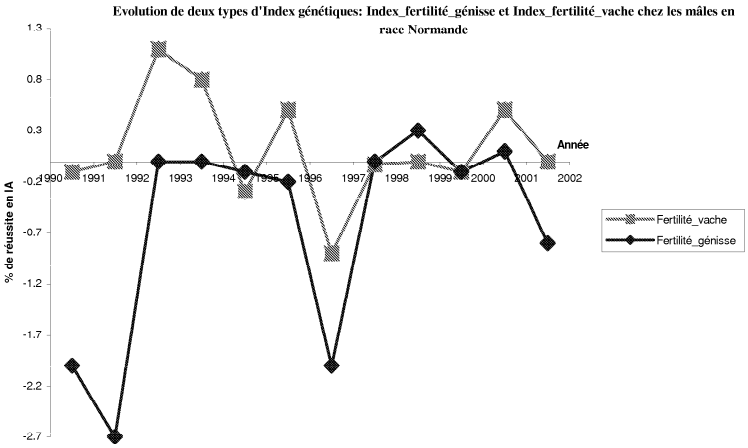
Tableau A-1 : *Données d'IA établies des campagnes 1997-2006 : le rang des IA varie de 1 à 5 et les performances de reproduction n'ont pas été retenues au-delà du quatrième vêlage (Le Mézec et al. 2007).*

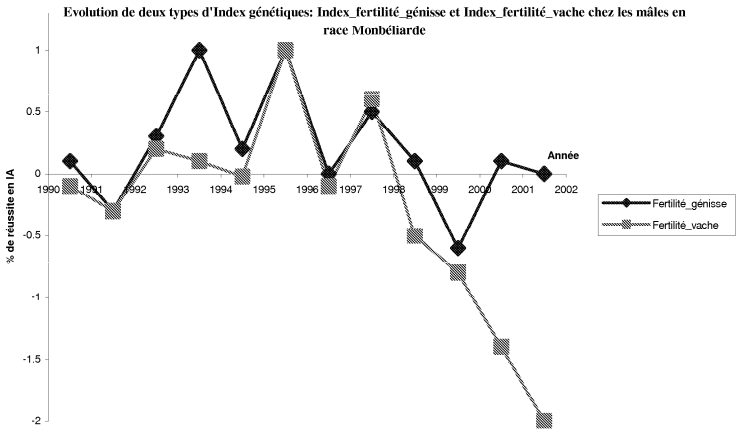
Race	**Toutes IA**	**Première IA**
Holstein	28 673 463	15 680 994
Normande	3 975 649	2 425 537
Montbéliarde	4 388 382	2 664 030

Le tableau A-2 montre le pourcentage de femelles ayant été inséminées plus de deux fois au cours des années 1998 et 2005 pour les races Holstein, Normande et Montbéliarde. Si le pourcentage de génisses inséminées plus de deux fois n'a pas beaucoup varié entre 1998 et 2005 pour les trois races, le pourcentage de vaches inséminées plus de deux fois présente une nette augmentation pour la race Holstein entre 1998 et 2005 alors qu'il reste presque constant en Normande et Montbéliarde.

Tableau A-2 : *Evolution, par race, des pourcentages de femelles nécessitant plus de deux IA pour être fécondées, entre les années 1998 et 2005 (Le Mézec et al. 2007).*

% de femelles à 3 IA ou plus	**Année**	**Montbéliarde**	**Normande**	**Holstein**
Génisses	1998	13%	13%	14%
	2005	13%	12%	15%
Vaches	1998	18%	18%	23%
	2005	18%	19%	28%





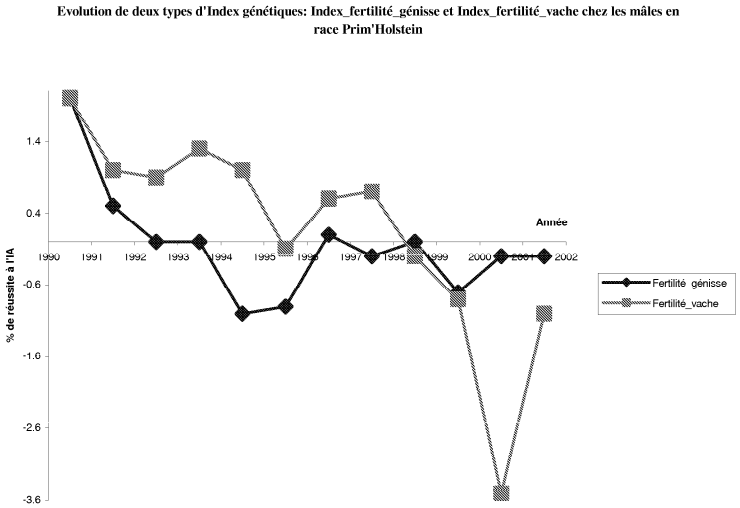


Figure A-4 *: Evolution de deux index génétiques de fertilité femelle (fertilité vache et fertilité génisse) chez les trois principales races bovines laitières Françaises entre les années 1990 et 2001 (Barbat et al., 2007).*

Le Mézec *et al*. (2007) ont trouvé que, chez la Holstein, l'intervalle [vêlage -vêlage] (IVV) est passé à 408 jours en 2005, alors que 10 ans auparavant il était de 395 jours. De plus, cet intervalle est plus long chez la Holstein que chez la Normande et la Montbéliarde, et ce, malgré une durée de gestation plus courte pour la Prim'Holstein (282 jours chez la Prim'Holstein contre 288 jours chez les races Normande et Monbéliarde).

Ces mêmes auteurs ont souligné que la part de dégradation d'origine génétique est importante chez la Holstein et est de l'ordre de 35 à 40% chez les primipares. Ainsi, l'index génétique de fertilité femelle des primipares Holstein nées entre 1993 et 2002 a chuté de 0,53 point d'index alors que l'index génétique fertilité « vache » a chuté de 0,2 point d'index. L'évolution de l'index génétique fertilité « génisse» a subi la même dégradation chez la Holstein avec une amplitude moins marquée (Le Mézec *et al.* 2007) alors qu'il est resté inchangé ou s'est amélioré chez les races Normande et Montbéliarde.

Plusieurs autres pays sont aussi concernés par la dégradation de la fertilité femelle des bovins laitiers. Au Royaume-Uni par exemple, le taux de vêlage après le premier service diminue de 1% par an et les troupeaux laitiers ont un taux de vêlage au premier service de seulement 40% (Royal *et al*., 2002).

De même, le taux de conception suite à la première IA est passé de 65% en 1951 à 40% en 1996 dans l'état de New York (Butler *et al*., 1998). Lucy (2001) rapporte que dans les années 50, les taux de conception étaient de l'ordre de 55% alors qu'il n'est que de 45% pour les inséminations qui ont lieu suite à un oestrus spontané dans les articles récents et est approximativement de 35% pour les IA à temps fixe. Le même auteur indique aussi une augmentation du nombre d'IA par conception de 1,62 en 1972 à 2,91 en 1996 et une augmentation de l'intervalle [vêlage- conception] et de l'intervalle [vêlage - 1ère IA] dans 143 troupeaux laitiers entre les années 1970 et 1999.

Aux Pays-Bas, l'IVV a augmenté de 0,2 jour par an entre les années 1981 et 1990 (Royal *et al*., 2002). D'autres pays tels que l'Irlande (Roche *et al*., 2000) et l'Australie (Macmillan *et al*., 1996) ont observé une baisse du taux de conception après la première IA.

IV. Liens entre la sélection pour les caractères de production laitière et la baisse de fertilité femelle chez les bovins laitiers

1. Evolution de la productivité des vaches laitières suite à une sélection intense pour les caractères de production et une optimisation des conditions d'élevage

La baisse de fertilité chez les bovins laitiers a été attribuée à une multitude de facteurs dont la plupart sont liées à l'intensification des systèmes d'élevage se traduisant principalement par l'augmentation de la productivité individuelle des animaux. Ainsi, pour réaliser leurs objectifs de production, les éleveurs ont essayé de maximiser la productivité de leurs animaux :

- A court terme, en agissant sur les facteurs environnementaux (principalement sur l'alimentation).
- A long terme, à travers une sélection génétique essentiellement orientée vers les caractères de production.

Ainsi, en France, la productivité individuelle moyenne a évolué quant à elle de 5823 Kg/an en 1987 à 8823 Kg/an en 2007 (Le Mézec, 2009). Aux Etats-Unis, la production moyenne quotidienne d'une vache laitière de race Holstein est passée de 37 Kg en 1960, à 116 Kg entre 1990 et 1996 (Hansen, 2000). Au Royaume-Uni, la production laitière des vaches laitières a augmenté de 200 Kg/an entre 1999 et 2003 (Pryce *et al*., 2004).

2. La sélection pour les caractères de production s'est accompagnée d'une baisse de fertilité femelle des bovins laitiers

Plusieurs études menées dans différents pays ont montré l'existence d'une corrélation génétique négative entre les caractères de production et différents indices de mesure de fertilité. Certaines études ont comparé des animaux sélectionnés pendant plusieurs années sur un haut potentiel génétique à d'autres sélectionnés pour de bas potentiels génétiques pour les caractères de production (Harrison *et al*., 1990; Hageman *et al*., 1991; Pryce *et al*., 1999; Roche *et al*., 2000; Gong *et al*., 2002b; Gutierrez *et al*., 2006). Toutes ces études ont montré que la fertilité est inversement liée au potentiel génétique des caractères de production des vaches laitières.

D'autres études ont montré l'existence de corrélations génétiques défavorables entre les deux types de caractères. Ces corrélations varient de 0,22 à 0,59 pour l'IVV, de 0,16 à 0,64 pour le nombre de jours ouverts et de 0,22 à 0,44 pour l'intervalle [vêlage-1ère IA] (Pryce *et al*., 2004). Ce qui veut dire que l'augmentation de la production laitière s'accompagne d'un allongement de ces intervalles. De même, des corrélations génétiques variant de -0,62 à 0,05, ont été trouvées entre production laitière et taux de conception au premier service (Pryce *et al*., 2004).

Il convient toutefois de souligner que la baisse de fertilité n'est pas spécifique à un système de production en particulier. En effet, elle concerne aussi bien les systèmes d'élevage exclusivement pastoraux (Harris *et al*., 2001; Diskin *et al*., 2006) que les systèmes d'élevage intensifs (Nebel et McGilliard, 1993; Beam et Butler, 1999).

B. Amélioration de la fertilité par voie génétique

La première partie de l'introduction a évoqué principalement les différents aspects mettant en évidence la baisse de fertilité femelle des bovins laitiers en France et dans certains pays développés ainsi que les liens existant entre la sélection pour les caractères de production et la baisse de fertilité femelle. Les aspects zootechniques et physiologiques sous-jacents à la baisse de fertilité femelle ont également été soulignés. La deuxième partie de l'introduction portera sur l'aspect génétique de ce caractère : historique de l'amélioration génétique de la fertilité femelle, les méthodes actuelles et leurs limites.

I. Historique de l'amélioration génétique de la FF en France

1. La fertilité : un caractère introduit récemment dans les objectifs de sélection

L'amélioration génétique de la fertilité femelle ne faisait pas partie des priorités des éleveurs il y a 20 ans. Depuis 1993, les données d'IA réalisées dans les troupeaux soumis au contrôle laitier sont collectées dans la base de données du centre de traitement de l'information génétique (CTIG) de l'INRA. Ces données brutes se limitent à l'identification de la vache inséminée, du taureau utilisé, de l'inséminateur réalisant l'acte, de la date d'IA et d'autres données optionnelles comme l'utilisation de la semence fraîche ou congelée, d'un traitement hormonal, d'une semence fractionnée…. Cette information brute est enrichie par la suite avec les données d'identification et de contrôle laitier, en particulier des données de mise bas pour construire différents indicateurs de fertilité. L'évaluation génétique des taureaux laitiers sur la fertilité *post partum* de leurs filles a débuté en juin 1998 ; cependant, les index de fertilité publiés ont souvent été des estimations moins précises que celles réalisées pour les autres caractères.

2. Choix de l'indice de fertilité femelle adéquat

Comme il a été précédemment évoqué (paragraphe B-II-2), la fertilité des vaches laitières est un caractère complexe qui peut être divisé en plusieurs sous-caractères mesurés par différents indices de fertilité ne traduisant pas les mêmes phénomènes physiologiques sous-jacents. De plus, la plupart de ces indices ont des propriétés peu satisfaisantes pour le généticien du fait de leur dépendance de facteurs non génétiques entraînant des biais difficiles à corriger lors de l'évaluation génétique. L'intervalle [vêlage-1ère IA], la durée de la période d'insémination et le nombre de jours ouverts, par exemple, dépendent de la politique de l'éleveur. Le choix du caractère reflétant au mieux la composante génétique de la fertilité de la vache laitière est donc primordial pour avoir un niveau acceptable de progrès génétique par génération. D'autre part, la fertilité des vaches laitières est déterminée par deux composantes principales : la reprise de l'activité sexuelle après le vêlage et l'aptitude à conduire une gestation traduite par le succès ou l'échec de chaque insémination. Cette deuxième composante paraît plus importante mais plus limitante chez les vaches laitières et moins influencée par la politique de l'éleveur. La réussite de l'IA a donc été retenue comme étant le caractère de fertilité utilisé pour l'évaluation génétique des taureaux d'IA en France. Cette donnée indique si une IA a abouti ou non à une fécondation. Deux alternatives sont envisageables pour calculer cette donnée : la première consiste à établir le statut de l'IA à une certaine période après l'insémination par un diagnostic de gestation ; dans ce cas, ce sont des valeurs de non retour en chaleur (TNR) qui sont calculées. Les TNR les plus utilisés sont calculés à 56 ou 90 jours

après IA. Ainsi, la variable de non retour prend la valeur 1 s'il n'y a pas de nouvelles IA, 56 (respectivement 90) jours après une IA et 0 dans le cas contraire pour le calcul du TNR56 (respectivement TNR90). La deuxième alternative est celle d'une IA dont le succès est statué dès lors qu'un vêlage est recensé suite à celle-ci. L'avantage de la première alternative par rapport à la deuxième, est que l'on garde toutes les IA (réussies ou non) et que l'on obtient une information plus précoce.

Les informations provenant de génisses sont considérées comme divergentes de celles provenant des vaches laitières, c'est pour cette raison que ces caractères sont souvent considérés comme différents. La fertilité des génisses laitières ne semble pas s'être aussi dégradée au cours des dernières années que celle des vaches laitières (Boichard *et al.*, 1998 ; Andersen-Ranberg et al., 2005).

II. Evaluation génétique de la fertilité : Méthodes d'indexation

1. Poids de la fertilité dans les objectifs de sélection de quelques pays

Les index de mérite global (IMG), analogue à l'ISU (Index de synthèse UPRA) en France, donnent les principaux objectifs de sélection d'une race dans un pays. La comparaison de neuf pays importants, en terme d'effectifs de taureaux et d'échange d'animaux, permet d'apprécier des situations diverses quant au poids relatif accordé à la fertilité. Les Pays-Bas arrivent en tête avec une pondération de la fertilité de 16% dans leur index global. La France arrive en deuxième position avec un poids de 12,5% dans l'ISU en races Holstein et Montbéliarde et de 10% en races Normande et brune. En Nouvelle Zélande, au Danemark et aux Etats-Unis, la fertilité compte pour 9% de l'index global. Le Royaume-Uni a récemment intégré la fertilité dans son index (Août 2007). Le Canada lui accorde un poids de 2% alors que l'Allemagne se limite à attribuer 1% du RZG à la fertilité. Par contre, l'Italie n'a pas encore intégré directement de caractères de fertilité dans son index global (Minery et al., 2008).

2. Modèles d'évaluation génétique de la FF en France

Depuis juin 2007, le modèle utilisé dans l'index officiel est un modèle animal pour lequel les valeurs génétiques des mâles et des femelles sont estimées simultanément en prenant en compte l'ensemble des liens de parenté. Ce modèle remplace le modèle « père **x** grand-père » de 1998 qui ne permettait d'estimer que la valeur génétique des mâles et considérait les

parentés uniquement entre mâles, en supposant une population de femelles de niveau génétique unique (Barbat *et al.,* 2007).

La nouvelle indexation permet également d'estimer conjointement la « fertilité génisse » et la « fertilité vache », ce qui devrait conduire à des index plus précis et plus stables, notamment lors du passage d'une forte proportion des filles du stade génisse au stade vache.

D'autres caractères fonctionnels considérés comme prédicteurs de fertilité tels que la longévité, l'inclinaison du bassin, … sont aussi pris en compte dans l'index de fertilité officiel pour améliorer sa précision.

En pratique, l'évaluation génétique de la fertilité est réalisée en plusieurs étapes. Dans un premier temps, l'analyse porte séparément sur les IA réalisées sur génisses et vaches et permet de calculer des performances corrigées pour l'ensemble des effets.

Ces performances corrigées sont ensuite analysées, d'une part pour fournir des index pour l'évaluation internationale réalisée par INTERBULL, d'autre part pour le calcul de l'ISU et de l'index de fertilité officiel (Barbat *et al.,* 2007).

a) Modèle d'évaluation mono-caractère

Ce modèle permet d'analyser le résultat de chaque IA séparément pour les génisses et les vaches. Le modèle inclut les facteurs suivants :

- effet génétique pour chaque femelle avec données, ainsi que tous les ascendants mâles ou femelles.
- Effets d'environnement permanent : c'est l'effet non génétique qui affectera toutes les IA d'une même vache tout au long de sa carrière.
- Effet d'interactions : troupeau **x** année, inséminateur **x** année, mois **x** région **x** année, jour de la semaine **x** région **x** année
- Age à l'IA pour les génisses.
- Race du taureau ayant réalisé l'IA.
- Rang du vêlage précédant l'IA pour les vaches.
- % de consanguinité de la vache.
- % de consanguinité du veau à naître.

Les deux derniers effets (% de consanguinité de la vache et du veau) ont été rajoutés au modèle de 2007 par rapport à celui de 1998. Les effets génétiques, d'environnement permanent, de l'interaction taureau d'IA **x** année sont traités en effets aléatoires, ce qui signifie que l'on suppose que ces effets suivent une distribution a priori de moyenne et de variance connues. Dans le cas de l'effet génétique, cette distribution prend en compte les relations de parenté entre les animaux. Les autres effets aléatoires sont considérés aléatoires entre eux. Les autres effets du modèle sont traités en effets fixés.

L'héritabilité du caractère est estimée à 2% (voir page 54 la définition de l'héritabilité). Ce modèle permet d'établir des index élémentaires : « fertilité vache » et « fertilité génisse ». Néanmoins, l'échantillon des inséminations réalisées sur vaches représente par définition un échantillon sélectionné puisque les génisses non gestantes au final sont éliminées. Cela entraîne un biais sur les index qui ne sont donc pas utilisés directement.
Afin de bien prendre en compte la sélection réalisée, les performances corrigées, calculées dans les évaluations mono-caractère sont ré-analysées conjointement, à l'aide de modèles multi-caractères faisant intervenir en outre, l'année de naissance des femelles (Barbat *et al.,* 2007).

b) Modèle d'évaluation multi-caractères

Deux types de modèles multi-caractères sont utilisés : un modèle bi-caractères pour l'évaluation internationale et un modèle multi-caractères pour l'indexation Française, celui utilisé pour les calculs de l'index de synthèse UPRa (ISU).

(1) Le modèle bi-caractères :

Ce modèle permet de combiner l'information sur vaches et sur génisses. Le modèle prend en compte la liaison génétique existant entre les deux caractères à travers leur corrélation génétique, estimée à 0,6. Les index obtenus sont utilisés par INTERBULL pour l'évaluation génétique internationale sur la fertilité. Au niveau national, bien que le mode de combinaison de l'information soit différent, l'index de fertilité vache calculé par ce modèle est proche de celui diffusé jusqu'en octobre 2006 (Barbat *et al.,* 2007).

(2) Le modèle multi-caractères :

Un modèle animal combinant des index élémentaires de fertilité (vache et génisse) et relatifs à d'autres caractères prédicteurs de la FF (comptage cellulaire, longévité fonctionnelle et certains caractères de morphologie) synthétise toute l'information connue pour le taureau candidat à la sélection et ses apparentés. La production laitière est incluse dans le modèle parce qu'elle est corrélée négativement à la fertilité et est fortement sélectionnée. L'index issu de cette combinaison est appelé « index FERc » (acronyme pour fertilité combinée).
Cet index est utilisé actuellment en tant qu'index officiel des taureaux pour la fertilité de leurs filles post-partum.
Les index sont exprimés en unités d'écart-type génétique par rapport à une base mobile de valeur 0 et sont compris entre -3 et +3. La base mobile regroupe les taureaux mis à l'épreuve

par les Unités de Sélection nés au cours des années (n-11) à (n-8) incluses (soit de 1998 à 2001 pour la base 2009, sans restriction sur le CD) (Institut de l'élevage, 2009).

3. Indices de fertilité utilisés pour l'évaluation génétique de la FF dans d'autres pays

Le nombre et la nature des indices de fertilité indexés sont différents selon les pays. Ils dépendent principalement des objectifs de sélection de chaque pays et de la disponibilité de certains types de données. Certains pays choisissent de séparer l'indexation des vaches de celle des génisses, d'autres ne le font pas. D'autres pays choisissent d'inclure dans leur indexation des caractères fortement corrélés avec la fertilité tels que l'état corporel, certains caractères morphologiques, la production laitière, etc...

Par exemple, le Danemark indexe cinq caractères : deux caractères pour les génisses et trois pour les vaches. Dans son index de synthèse fertilité, le Danemark donne plus d'importance à la reprise de l'activité sexuelle de la vache avec un poids de 54% donné à l'intervalle [Vêlage-1ère IA], 10% à l'intervalle [1ère – dernière IA] génisse et 36% à l'intervalle[1ère – dernière IA] vache. L'Allemagne calcule un index synthétique semblable au Danemark prenant en compte l'intervalle [1ère – dernière IA] (11%) et le TNR56 en génisse (11%). Le Royaume-Uni indexe quatre caractères en vache dans un modèle multi-caractères dans lequel l'état corporel et la production laitière sont utilisés comme prédicteurs ; l'index de synthèse de fertilité accorde un poids de 83% au TNR56 et de 17% à l'IVV. La Nouvelle Zélande a un système de production saisonnier dans lequel les éleveurs essaient de regrouper au maximum les vêlages. Les critères de mesure de la fertilité prennent en compte cette particularité. Ce sont le PM21 (qui mesure la réussite ou l'échec d'une vache à être présentée à l'accouplement dans les 21 premiers jours de la période de reproduction du troupeau) et le CR42 (qui reflète la réussite ou l'échec de la vache à produire un veau dans les 42 premiers jours de la période de vêlage du troupeau). Un modèle multi-caractères analyse à la fois le PM21 et le CR42 à partir des trois premières lactations ainsi que la production laitière et la note de l'état corporel.

Les Pays-Bas indexent trois caractères de fertilité : l'intervalle [Vêlage-1ère IA], l'intervalle [Vêlage-Vêlage] et le TNR56. Ils utilisent un modèle multi-caractères intégrant la production laitière, la quantité de matière grasse (MG), la matière protéique (MP) et l'état corporel en tant que prédicteurs de fertilité. Dans l'index de synthèse de fertilité, le poids de l'IVV est de 85%, celui du TNR de 15%.

4. Indexation internationale de la fertilité

Les premières études en vue d'une indexation internationale de la FF ont débuté en juillet 2004 en race Holstein. La première indexation internationale de la FF, faite par le service d'évaluation internationale des taureaux d'IA ou INTERBULL, a eu lieu en février 2007 en

race Holstein. Dix pays y ont participé : la Belgique, l'Espagne, le Royaume-Uni, l'Irlande, Israël, l'Italie, la Nouvelle-Zélande, les Pays-Bas, la Suisse et les Etats-Unis. La France a participé pour la première fois en août 2007. Actuellement, cinq catégories de caractères sont évaluées :

- *Catégorie 1 :* l'aptitude de la génisse à concevoir. Par exemple, par mesure du taux de conception.
- *Catégorie 2 :* l'aptitude de la vache à reprendre sa cyclicité après le vêlage. L'intervalle [vêlage-1ère IA] en est un exemple.
- *Catégorie 3 :* l'aptitude de la vache à concevoir exprimée en pourcentage de réussite telle que le TNR.
- *Catégorie 4 :* l'aptitude de la vache à concevoir exprimée en intervalles telle que l'intervalle [1ère IA- conception] ou l'intervalle [1ère – dernière IA].
- *Catégorie 5 :* des caractères qui mesurent à la fois l'aptitude de la vache à reprendre sa cyclicité et à concevoir, exprimés sous forme d'intervalles tels que : l'IVV ou le nombre de jours ouverts.

Les évaluations génétiques sont réalisées par INTERBULL qui calcule également des corrélations génétiques entre pays. Une corrélation génétique élevée entre deux pays signifie que les pays en question ont une définition assez proche du caractère et par conséquent, la valeur génétique dans l'un de ces deux pays sera utile pour l'autre. Certaines catégories sont mieux harmonisées que d'autres ; ainsi, les caractères en génisse et les intervalles mesurant la reprise de la cyclicité et les intervalles mesurant les deux aptitudes à la fois ont des corrélations entre pays élevées, dépassant 0,75. En revanche, les caractères mesurant l'aptitude de la vache à être gestante sont plus hétérogènes entre pays. Les pays utilisateurs du TNR56, en tant que prédicteur de FF, sont bien harmonisés (Italie, Pays-Bas, Suisse, Israël, Angleterre, Danemark, Finlande et Suède) (INTERBULL, 2008).

III. Limites des méthodes d'indexation actuelles de la fertilité femelle et nécessité d'une approche moléculaire pour l'amélioration génétique des bovins laitiers

1. La fertilité : caractère à faible héritabilité mais à variation génétique importante

a) Héritabilité des indices classiques de fertilité

La fertilité des bovins laitiers mesurée par les indices classiques ou traditionnels est un caractère à faible héritabilité. En effet, les études d'estimation des paramètres génétiques de différents indices de mesure de fertilité ont rapporté une héritabilité inférieure à 6% pour la

majorité de ces indices (tableau B-1). Pour un même indice de fertilité, des différences importantes entre les valeurs d'héritabilité ont été trouvées entre les différentes études. Ceci est principalement dû à la variabilité du nombre d'animaux inclus dans les études, à une différence des modèles génétiques utilisés et surtout au fait que la valeur d'héritabilité ne dépend pas que du caractère, mais aussi de la fréquence allélique des gènes contrôlant le caractère analysé dans la population.

La faible héritabilité des caractères de fertilité mesurés par les indices est due au fait qu'une grande partie de la variation résiduelle utilisée dans les modèles statistiques de prédiction des caractères de fertilité reste inexpliquée. En effet, les indices de fertilité, bien que corrigés pour les effets de l'environnement, sont des indices qui reflètent trop peu la capacité de la vache à établir une activité ovarienne normale après le vêlage, à montrer des signes d'œstrus ou à concevoir et maintenir une gestation lorsqu'elle est inséminée au bon moment.

Tableau B-1 : Valeurs d'héritabilité rapportées pour différents indices de fertilité chez la vache laitière

Indice de fertilité	**Héritabilité (h^2)**	**Nombre d'animaux**	**Référence**
IVV	3,3%	43029	Wall *et al.* (2003)
	2,4%	62443	Kadarmideen *et al.* (2003)
	5,5%	32359	Dal Zotto *et al.* (2007)
	4%	63160	González-Recio *et al.* (2005)
	4%	40736	De Haas *et al.* (2007)
I [V-1ère IA]	5%	63160	González-Recio *et al.* (2005)
	5,8%	30037	Weigel *et al.* (2000)
	3,7%	43029	Wall *et al.* (2003)
	2,5%	62443	Kadarmideen *et al.* (2003)
	8%	40736	De Haas *et al.* (2007)
	3,01%	1815581	Andersen-Ranberg *et al.* (2005)
Nombre de jours ouverts	7%	41170	González-Recio *et al.* (2006)
	4%	1454916	Chang *et al.* (2006)
	3,7%	> 16000000	Van Raden *et al.* (2004)
Taux de conception	2%	250215	Boichard et Manfredi (1994)
	1%	40736	De Haas *et al.* (2007)
Nombre d'IA/conception	1%	40736	De Haas *et al.* (2007)
	3,8%	63160	González-Recio *et al.* (2005)
	5%	63160	González-Recio *et al.* (2005)
TNR56	0,99%	1815581	Andersen-Ranberg *et al.* (2005)
	1,8%	43029	Wall *et al.* (2003)

b) Héritabilité d'indices endocriniens liés à la fertilité

(1) Héritabilité de la concentration de progestérone

Peu d'études se sont basées sur des paramètres physiologiques liés à la fertilité pour essayer d'estimer d'une manière plus précise et plus fiable la part de la composante génétique de la fertilité. En effet, de tels paramètres physiologiques reflètent d'une manière plus transparente le déroulement de l'activité ovarienne après le vêlage et pendant la gestation.

L'un des paramètres physiologiques qui a été suggéré par plusieurs études pour mesurer la fertilité est la concentration de progestérone dans le lait. En effet, la progestérone joue un rôle très important dans la reproduction des vaches laitières (A-I-2-b). Des concentrations anormales de progestérone sont associées à des anomalies de l'activité ovarienne (Lamming et Darwash, 1998; Royal *et al*., 2002) qui, à leur tour, sont associées à une baisse de fertilité chez la vache *post partum*. Dans ces études, des dosages de progestérone du lait à différentes périodes du cycle ovarien ont été utilisés pour mesurer la durée de l'intervalle entre le vêlage et la reprise de l'activité lutéale (indiquant un corps jaune actif). Cet intervalle est négativement associé à la fertilité chez la vache laitière. La plupart des études d'héritabilité (h^2) faites sur la concentration de progestérone ont trouvé des valeurs moyennes de h^2. Darwash *et al*. (1997) présentent une valeur de h^2 de 0,21 au Royaume-Uni (étude de 1137 animaux et 1737 lactations élevés entre 1975 et 1982). Veerkamp *et al*. (1998) calculent une valeur variant entre 0,14 et 0,20 aux Pays-Bas. Royal *et al*. (2002) signalent des h^2 de 0,16 (1080 vaches post partum et 1212 lactations). Ces valeurs d'héritabilité sont nettement supérieures à celles trouvées avec les paramètres de fertilité traditionnels. Toutefois, ces valeurs de h^2 ont été estimées à partir d'effectifs faibles et, par conséquent, sont peu précises. L'obtention de dosages de progestérone sur un grand effectif est coûteuse et laborieuse du fait notamment de la nécessité de prélèvements fréquents pour mesurer les concentrations de progestérone. En effet, des prélèvements individuels quotidiens ou même bi-hebdomadaires de progestérone à partir de filles de taureaux d'IA dans le cadre de schémas de testage sur descendance ne sont pas réalisables à cause de la dispersion de la descendance dans plusieurs troupeaux. Darwash *et al*., (1999) et Veerkamp *et al*., (1998) ont proposé des méthodes alternatives plus pratiques : ils ont postulé que la moyenne des profils de progestérone de taureaux d'IA issues de mesures mensuelles de progestérone sur leurs filles peut être utilisée comme mesure de FF lors de la sélection de taureaux d'IA dans un schéma de sélection sur descendance. Les concentrations de progestérone peuvent alors être mesurées mensuellement lors des contrôles laitiers. Van der Lende *et al*., (2004) ont proposé d'utiliser la moyenne des profils de progestérone pour calculer le stade de lactation le plus précoce où au moins 50%

des filles d'un taureau d'IA ont déjà commencé leur activité lutéale. Ces auteurs ont observé une large variation comprise entre 25 et 78 jours (moyenne de 41,9 jours et écart type de 13,6 jours) entre les taureaux de testage pour ce caractère. L'étude était réalisée sur les filles de 20 taurcaux d'IA (4470 échantillons dc lait).

(2) Héritabilité de quelques pathologies liées à la fertilité chez la vache laitière

La baisse de la fertilité chez les vaches laitières se manifeste principalement par des anomalies de la reprise de l'activité ovarienne qui se traduit généralement par des cycles anormalement courts et le développement de corps jaune persistant. Peu d'études ont été faites sur les paramètres génétiques de telles anomalies ovariennes. La plupart de ces études ont trouvé de faibles héritabilités. Citons l'exemple de trois études :

- Abdel-Azim *et al*. (2005) qui ont trouvé une h^2 de 0,03 pour l'anomalie de l'ovaire cystiquc (10 332 obscrvations) ct dc 0,11 pour l'anomalic du corps jaunc pcrsistant (8 489 observations). L'anomalie de l'ovaire cystique inclue trois catégories dont celle du corps jaune persistant.
- Zwald *et al*. (2004) ont trouvé une h^2 de 0,04 pour l'anomalie du corps jaune cystique (272 576 lactations étudiées).
- Harder *et al*. (2006) ont trouvé une valeur maximale de h^2 de 0,04 pour un ensemble de pathologies de fertilité incluant les anomalies ovariennes, l'endométrite, la rétention placentaire et d'autres types de pathologies. Cette étude comprenait 19 301 lactations de différents ordres.

En conclusion, les valeurs d'héritabilité obtenues par la plupart des études pour les pathologies liées à la fertilité sont globalement faibles. Ceci est probablement dû d'une part à une hétérogénéité du diagnostic de ces pathologies (différences entre les seuils de déclaration de la présence de pathologie chez l'animal) et d'autre part, au fait que certaines de ces études incluent simultanément plusieurs pathologies et supposent donc qu'elles ont les mêmes origincs génétiqucs, cc qui n'cst pas forcémcnt lc cas.

c) La fertilité caractère à variation génétique élevée

Toutes les études qui ont estimé les paramètres de FF mesurée par les paramètres traditionnels ont trouvé une faible héritabilité de ce caractère. Les faibles valeurs d'héritabilité sont expliquées principalement par la grande influence des effets environnementaux (y compris les décisions de l'éleveur) sur les paramètres de fertilité traditionnels. Toutefois, une héritabilité faible ne signifie pas absence de variabilité génétique, mais simplement que la part de la variation génétique additive est beaucoup plus faible que celle due à l'environnement.

Boichard *et al*. (1998) ont trouvé un écart type génétique d'environ cinq points de réussite à l'IA, ce qui offre une marge non négligeable pour l'amélioration génétique de la fertilité. D'autres études concordent quant à l'importance de la variabilité génétique de la fertilité. Philipsson *et al*. (1981) ont conclu que l'ampleur de la variance génétique additive est considérable pour plusieurs caractères de fertilité, en dépit des valeurs d'héritabilité généralement basses pour les résultats individuels d'insémination. Kadarmideen *et al*. (2001) ont abouti aux mêmes conclusions quant à une variation génétique élevée des caractères de FF. Ils ont analysé l'IVV, le nombre de services par conception et le nombre de jours au premier service, comme caractères de FF et trouvent, sur un nombre total de données de 66 383 enregistrements, des coefficients de variation aussi importants que ceux des caractères de production et variant de 0,13 pour l'IVV à 0,36 pour le nombre de jours au premier service.
Royal *et al*. (2002) ont trouvé un écart type standard du début de l'activité lutéale mesurée par la concentration de progestérone dans le lait de 0,208, alors que l'écart type phénotypique n'est que de 0,52.
Toutes ces études soulignent clairement que la part de la variance génétique additive de certains « sous-caractères » de fertilité est non négligeable par rapport à la variance phénotypique totale de ces « sous-caractères ».

2. Difficulté d'amélioration génétique de la FF par les indexes de sélection classiques

L'amélioration génétique de la fertilité par les index de sélection classiques paraît difficile. En effet, plusieurs facteurs sont liés à la nature des données récoltées, d'autres sont liés à la nature des caractères utilisés pour la mesure de la fertilité. Plusieurs de ces facteurs ont été déjà abordés précédemment, mais seront listés de nouveau pour avoir une idée claire des difficultés d'amélioration génétique par indexation classique.

a) Nature des données utilisées pour l'élaboration des index de fertilité classiques

Les données collectées pour l'élaboration d'indices de fertilité sont souvent inadéquates pour l'analyse statistique. Des caractères tels que le taux de conception ou le TNR sont enregistrés comme étant des observations binaires (réussite ou échec) et l'application de modèles Gaussiens à de telles données peut aboutir à un biais, sauf si la probabilité de succès est proche de 0,5 et les composantes de variance homogènes entre les différents groupes (Boichard et Manfredi, 1994; Weigel et Rekaya, 2000). Les caractères d'intervalle (tels que l'IVV, les DO ou le nombre de jours à la première IA) et les données de comptage tels que le nombre d'IA/conception sont le plus souvent des données qui ont une distribution asymétriques, en plus d'être censurées (données manquantes). S'il est possible de corriger

l'asymétrie des données en faisant des transformations, les données censurées sont difficiles à corriger. La censure des données peut être due au fait que certaines vaches ne deviennent jamais gestantes, d'autres meurent ou sont réformées avant de devenir gestantes. La censure de ces données entraîne des biais importants dans l'estimation des paramètres génétiques et des valeurs génétiques des taureaux d'IA.

b) Nature des caractères mesurant la fertilité femelle

(1) La FF est un caractère tardif et limité à un seul sexe

La FF est un caractère limité à un seul sexe (les femelles) et s'exprimant tardivement dans la vie des bovins. Ainsi, pour avoir une bonne estimation des valeurs génétiques des taureaux d'IA, il faut attendre plusieurs années (le temps de les tester sur un grand nombre de leurs filles) avant que ces taureaux ne soient indexés. Le maintien de ces taureaux pendant toute la période de testage n'est pas sans coût.

(2) La faible héritabilité des caractères mesurant la fertilité femelle

La faible héritabilité est une caractéristique commune à tous les caractères mesurant la FF des bovins. En d'autres termes, cela signifie que les mesures faites pour ces caractères sont de mauvais prédicteurs de la valeur génétique réelle de l'animal. L'obtention de valeurs génétiques ayant un minimum de fiabilité, nécessite donc un grand nombre de descendants, ce qui entraîne une augmentation des coûts des schémas de sélection mis en place. L'utilisation de caractères plus directement dépendant de la physiologie de la reproduction des vaches est onéreuse pour ces schémas.

3. Conclusion

La sélection sur index classiques des caractères de fertilité femelle chez les bovins laitiers est peu efficace pour toutes les raisons citées précédemment, d'autant plus qu'en incluant d'autres caractères dans ces index, il est toujours difficile d'obtenir le progrès génétique permettant de rentabiliser les investissements de mise en place des schémas de sélection en cours. L'identification de régions génomiques ayant une influence sur la fertilité femelle à l'aide de marqueurs moléculaires judicieusement choisis dans le génome bovin et l'intégration de l'information moléculaire dans les index de sélection classiques dans le cadre d'une sélection assistée par marqueurs (SAM) pourrait être plus qu'une alternative pour une estimation plus précise de la valeur génétique des candidats à la sélection et la réalisation de progrès génétiques suffisamment rentables pour ce caractère. De telles régions génomiques sont appelées QTL (pour Quantitative Trait Loci). La grande variabilité d'origine génétique

de la fertilité femelle des bovins laitiers est un atout majeur qui justifie amplement l'utilisation de la voie génétique pour l'amélioration de ce caractère.

C. Cartographie de QTL impliqués dans la variabilité génétique des caractères quantitatifs chez les animaux domestiques

I. Notion de caractère quantitatif et de QTL

1. Définition d'un caractère quantitatif

Un caractère quantitatif est un caractère qui présente une variation continue ; autrement dit, pour un caractère quantitatif, les individus d'une population sont classés de façon continue dans une série graduée. Par contre, un caractère qualitatif présente une variation discrète : les individus d'une population sont classés en plusieurs catégories distinctes avec une faible ou une absence de connexion entre les catégories, Par nature, les caractères quantitatifs sont beaucoup plus nombreux du fait des mesures utilisées.

2. Propriétés des caractères quantitatifs

a) Propriétés statistiques

Chaque caractère mesurable et présentant une variation continue peut, en principe, être étudié comme un caractère quantitatif. La distribution des caractères quantitatifs est généralement normale du fait du nombre important de facteurs qui les contrôle et peut être décrite par une moyenne et une variance calculées à partir des données. Dans l'étude des caractères quantitatifs il est donc possible d'utiliser ces propriétés statistiques.

b) Origine de la variation continue des caractères quantitatifs : le modèle infinitésimal

Avant 1918, on pensait que les caractères à variation continue étaient hérités d'une manière différente des caractères qualitatifs selon la théorie de l'hérédité par « mélanges » : les performances des descendants étaient un mélange des performances des parents. En 1918, Fisher a mis en place la base théorique de la génétique quantitative en développant un modèle appelé infinitésimal dans lequel la variation continue d'un caractère quantitatif est expliquée par la ségrégation d'un grand nombre de loci (ce nombre tend vers l'infini) non liés et non épistatiques (Fisher, 1918).

c) Application du modèle infinitésimal : effet d'un seul locus

(1) Décomposition de la valeur génotypique au niveau d'un locus

Fisher a suggéré que la valeur génotypique, d'un parent peut être décomposée en une fraction transmissible à son descendant et une fraction non transmissible. La valeur transmissible à la descendance est appelée valeur génétique et est composée de la somme des effets moyens α_1 et α_2 des deux allèles. La valeur non transmissible est appelée valeur de dominance.

Dans le cas d'un locus biallélique Q (Q1 et Q2 sont les deux allèles de ce locus), la valeur génotypique G d'un individu est respectivement 2a, 1+k et 0 pour Q_2Q_2, Q_1Q_2 et Q_1Q_1.

G peut aussi être modélisée de la manière suivante :

$$G = \mu_G + \alpha_1 N_1 + \alpha_2 N_2 + \delta \qquad (E1)$$

μ_G: *valeur génotypique moyenne* $(\mu_G = \sum G_{ij} \times fréquence\ (Q_i Q_j))$

α_1: Effet moyen de l'allèle 1.

α_2: Effet moyen de l'allèle 2.

N_1 est le nombre de copies de l'allèle 1 du locus ($N_1 = 2 - N_2$).

N_2 est le nombre de copies de l'allèle 2 du locus.

δ est la valeur de dominance : déviation de la valeur génotypique pour le génotype Q_iQ_j de celle prédite par la contribution individuelle moyenne de chaque allèle.

G peut être écrite d'une autre manière et exprimée en fonction de N_2 par exemple.

$$G = \mu_G + 2\alpha_1 + (\alpha_2 - \alpha_1)N_2 + \delta \qquad (E2)$$

L'équation (E2) est une régression linéaire où $\mu_G + 2\alpha_1$ est l'ordonnée à l'origine, N_2 est la variable indépendante, $(\alpha_1-\alpha_2)$ est le coefficient de régression et δ est la résiduelle (Figure C-1).

$G = \lambda + \alpha N_2 + \delta$ (où $\alpha = \alpha_2 - \alpha_1$ et $\lambda = \mu_G + 2\alpha_1$).

En revenant à l'équation (E1), l'espérance de G est $\mu_G = \mu_G + \alpha_1 E(N_1) + \alpha_2 E(N_2) + 0$

D'où $\alpha_1 E(N_1) + \alpha_2 E(N_2) = 0$ (E3)

N_1 est une variable aléatoire qui peut prendre trois valeurs:

- 2 si le génotype de l'individu est Q_1Q_1.
- 1 si le génotype de l'individu est Q_1Q_2.
- 0 si le génotype de l'individu est Q_2Q_2.

$E(N_1) = 0*E(N_1=0) + 1*E(N_1=1) + 2*E(N_1=2)$

$E(N_1=1)$ est la fréquence des individus hétérozygotes dans la population, alors que $E(N_1=2)$ est la fréquence des individus homozygotes Q_1Q_1dans la population en supposant que la population est en équilibre de Hardy-Weinberg.

$E(N_1) = 2p(1-p) + 2p^2 = 2p$

De même, $E(N_2)=2q$

L'équation E3 devient alors $2p\alpha_1 + 2q\alpha_2 = 0$ et sachant que $\alpha = \alpha_2 - \alpha_1$, on aura : $\alpha_1 = -q\alpha$ et $\alpha_2 = p\alpha$.

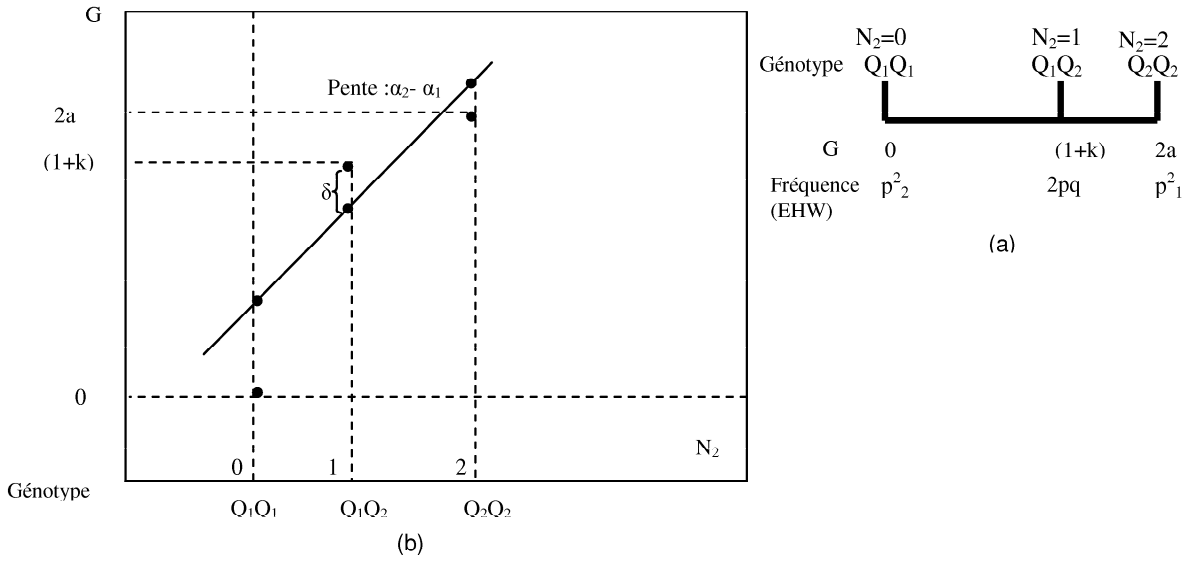


Figure C-1 : *Valeurs génotypiques et fréquences alléliques des trois classes génotypiques dérivant d'un locus biallélique Q. N_2 est le nombre d'allèles du second allèle Q_2 (a). Décomposition de Fisher : l'axe des abscisses représente N2. L'axe des ordonnées représente la valeur génotypique G. Celle-ci est régressée sur N_2. La droite de régression a pour pente α_2-α_1 et passe par les trois points représentant les valeurs génétiques. Les différences entre les valeurs génotypiques et celles génétiques sont les valeurs de dominance δ. (b) (d'après Falconer, 1996).*

(2) Variance génétique

Repartons de l'équation (E2) : $G = \mu_G + 2\alpha_1 + (\alpha_2 - \alpha_1) N_2 + \delta$

La variance génétique $\sigma^2_G = V(\mu_G + 2\alpha_1 + (\alpha_2 - \alpha_1) N_2 + \delta) = (\alpha_2 - \alpha_1)^2 \sigma^2 (N_2) + \sigma^2 (\delta))$

$\sigma^2_G = \alpha^2 2pq + \sigma^2 (\delta)$

Enfin, en considérant que la variance de la résiduelle δ est nulle (pas d'effets de dominance), on aboutit à $\sigma^2_G = 2pq\ \alpha^2$.

d) Généralisation du modèle à un locus : le modèle polygénique - Notion de QTL et d'héritabilité

(1) Généralisation du modèle à un locus : le modèle polygénique

En partant de l'équation $G = \mu_G + \alpha_1 N_1 + \alpha_2 N_2 + \delta$ établie dans le cas d'un seul locus biallélique, la variance génétique σ^2_G s'écrit sous la forme :

$\sigma^2_G = V(\mu_G + \alpha_1 N_1 + \alpha_2 N_2 + \delta)$ (E3)

L'équation (E3) peut être généralisée à k loci indépendants de contribution égale à la valeur génotypique :

$$\sigma^2_G = \sum_{j=1}^{k} \sigma^2 (\mu_G + \alpha_1 N_1 + \alpha_2 N_2) + \sum_{j=1}^{k} \sigma^2 (\delta)$$

Cette équation est souvent écrite sous la forme : $\sigma^2_G = \sigma^2_A + \sigma^2_D$. Les effets d'interaction entre les loci ne sont pas considérés dans cette équation.

La valeur phénotypique observée P d'un caractère quantitatif n'est pas seulement d'origine génotypique, mais contient une part environnementale E, la valeur phénotypique d'un caractère peut alors s'écrire P= G + E et sa variance est :

$\sigma^2_P = \sigma^2_G + \sigma^2_E$ (E4)

(2) Notion d'héritabilité

L'équation (E4) nous permet de définir l'héritabilité qui représente la part de la variance phénotypique qui est d'origine génétique. L'héritabilité est l'un des paramètres les plus importants pour l'étude génétique des caractères quantitatifs. Elle est définie de deux manières :

- L'héritabilité au sens large : H^2 qui est la variance génétique totale sur la variance phénotypique. $H^2 = \sigma^2_G / \sigma^2_P$
- L'héritabilité au sens strict h^2 qui est la variance génétique additive sur la variance phénotypique. $h^2 = \sigma^2_A / \sigma^2_P$.

La deuxième définition de l'héritabilité (celle au sens strict) est la plus couramment utilisée, étant donné que la valeur génétique additive est la seule composante génétique à être transmise à travers les générations et par conséquent elle présente plus d'intérêt pour les sélectionneurs.

A partir de la définition même de l'héritabilité, on peut noter que plus σ^2_A est importante, plus le caractère est améliorable par voie génétique puisque σ^2_A mesure la source de variabilité d'origine génétique transmissible à la descendance.

(3) Notion de QTL

Selon le modèle infinitésimal de Fisher, la variance génétique σ^2_G est due à la contribution égale d'un très grand nombre de loci (ce nombre tend vers l'infini). Toutefois, ce modèle est en désaccord avec la mise en évidence d'effets de gènes dits « majeurs » sur des caractères quantitatifs. Citons l'exemple de la mutation K232A dans le gène DGAT1 (Grisart *et al.*, 2002, Grisart *et al.*, 2004) et la mutation Y581S dans le gène ABCG2 (Cohen-Zinder *et al.*, 2005) : deux mutations influençant la composition du lait chez les bovins et la mutation 3072(G-A) dans l'intron 3 du gène IGF2 chez le porc (Van Laere AS *et al.*, 2003). Ces gènes majeurs expliquent une grande partie, mais pas toute la variation du caractère au sein de la population. A côté de ces gènes majeurs, il existe des QTL à effet intermédiaire sur le caractère dont la ségrégation des différents allèles ne se traduit pas nettement à cause d'une relation non univoque entre génotype au locus et phénotype (Hayes et Goddard, 2001).

Le modèle de Fisher a surtout servi de modèle mathématique pour l'étude de la variation continue des caractères quantitatifs, malgré tout assez peu réaliste d'un point de vue purement biologique.

(4) Nombre et effets de QTL

Shrimpton et Robertson (1988) ont étudié le caractère « nombre de poils sternopleuraux » chez la drosophile. En fonction de leurs effets sur le caractère, ils ont proposé une distribution en « L » du nombre de QTL; c'est-à-dire qu'il existe un nombre limité de QTL avec un grand effet et un plus grand nombre de QTL avec un faible effet sur le caractère étudié. Hayes et Goddard (2001) ont dérivé les distributions des effets de QTL à partir d'une méta-analyse des estimations des effets de QTL déjà publiés dans les espèces bovine et porcine. Ils n'ont considéré que les effets additifs des QTL. La distribution des effets de QTL était semblable à celle obtenue par Shrimpton et Robertson (1988) avec peu de QTL ayant un grand effet sur le caractèrc ct un grand nombrc dc QTL ayant un cffct faiblc ou moycn (Figurc C-2).

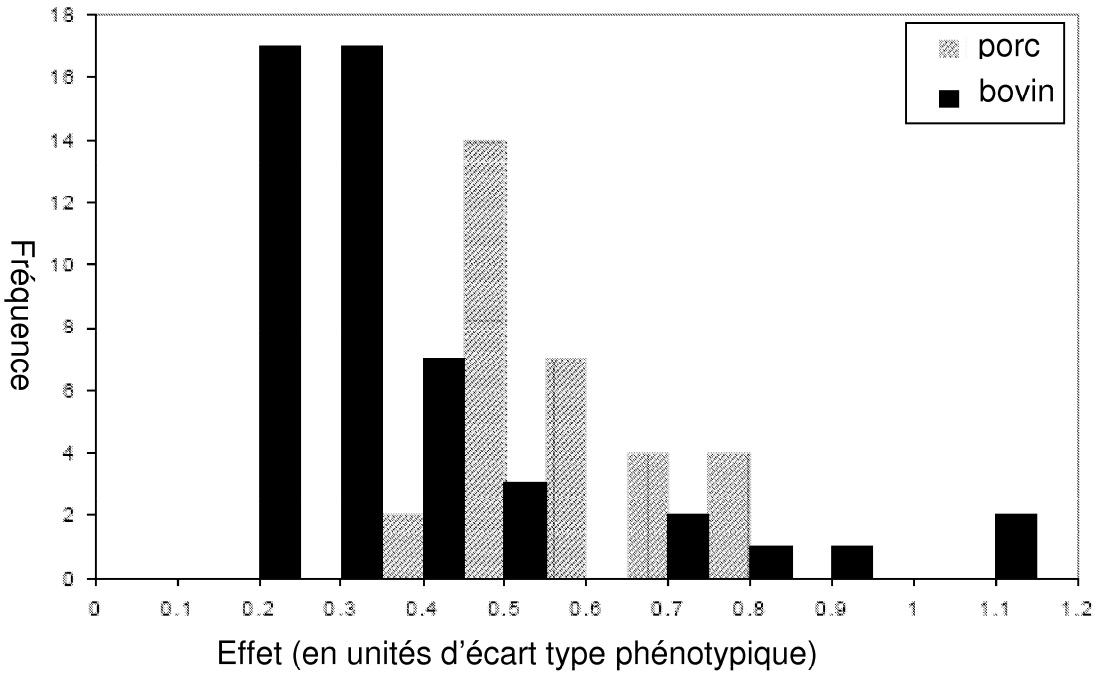


Figure C-2 : Distribution des fréquences des effets apparents de QTL détectés chez le porc et les bovins laitiers dans une méta-analyse incluant les résultats issus d'expériences de détection de QTL chez ces deux espèces. La distribution des effets de QTL détectés chez le porc montre un plus grand nombre de QTL à effet moyen que de QTL à grand effet. Chez les bovins laitiers, un plus grand nombre de QTL à petit effet- moyen effet est détecté (d'après Hayes et Goddard, 2001).

3. Effets d'interaction

Les effets de QTL varient en fonction du fond génétique, du sexe et de l'environnement extérieur ; cela suppose donc l'existence de plusieurs types d'interactions : génotype **x** génotype ou épistasie, génotype **x** environnement et génotype **x** sexe (Mackay, 2001).

a) Interaction génotype X génotype (épistasie) :

Dans son sens le plus large, l'épistasie implique que l'effet d'un génotype donné sur le phénotype dépend du fond génétique de l'individu. Dans sa forme la plus simple, ceci se réfère à une interaction entre les paires de loci dans lesquelles l'effet phénotypique d'un locus dépend du génotype du second locus et plus généralement l'effet phénotypique d'un locus dépend du génotype de plusieurs autres loci. La plupart des travaux de cartographie de QTL ont négligé les interactions épistatiques entre les différents QTL malgré leur importance dans la variation des caractères quantitatifs principalement à cause de difficultés liées à la taille des populations trop faible pour une puissance de détection suffisante. En effet, la détection d'interactions entre loci nécessite une grande taille de populations, comme le nombre de classes de génotypes considérés et des tests statistiques augmente rapidement avec le nombre de loci. Cependant, le degré d'implication des interactions épistatiques dans la variation des caractères quantitatifs n'est pas encore connu. Par conséquent, on ne peut pas supposer que l'épistasie peut être détectée dans tous les caractères et dans toutes les populations (Carlborg et Haley, 2004).

L'épistasie se rapporte aux effets d'interaction entre plusieurs loci pour déterminer un caractère donné. Le cas contraire, action d'un seul locus sur plusieurs caractères est aussi fréquent dans la nature ; ce type d'action est appelé *pléiotropie*. Un exemple de pléiotropie très connu chez les bovins laitiers est celui du gène DGAT1 qui affecte à la fois la quantité de lait, le taux de matière protéique et de matière grasse. Un autre exemple : les gènes impliqués dans le développement des soies sensorielles chez la drosophile sont aussi impliqués dans le développement du système nerveux central, des yeux et des ailes et dans le déterminisme du sexe (Mackay, 2001). Il existe souvent une confusion entre la distinction des effets de pléiotropie et ceux de liaison génétique. En effet, la corrélation génétique entre deux caractères donnés peut être soit le résultat de l'effet d'un seul QTL sur les deux caractères, soit le résultat de l'effet de deux QTL étroitement liés mais influençant indépendamment les deux caractères en question. La distinction entre les effets de pléiotropie et de liaison n'est possible que lorsque la (les) mutation(s) responsable(s) de la variation des deux caractères est (sont) identifiée(s).

b) Interaction génotype x sexe

Au niveau du caractère, une interaction significative entre le génotype et le sexe signifie que la corrélation génétique entre mâles et femelles pour un caractère donné est inférieure à 1 et qu'il y a un dimorphisme sexuel d'origine génétique pour le caractère. Une telle interaction peut être attribuée à des loci liés au sexe ayant les mêmes effets dans les deux sexes ou à des loci situés sur des autosomes avec des effets différents chez le mâle et la femelle.

c) Interaction génotype x milieu

Pour détecter des interactions génotype **x** milieu, il est nécessaire que les mêmes génotypes soient évalués dans des environnements différents. Cela est possible dans le cas de populations où le fond génétique est contrôlé telles que les lignées RIL (pour recombinant inbred lines) ou les populations F2 où les lignées parentales sont des lignées pures. Cependant, cela semble difficile dans le cas de populations où les individus sont génétiquement hétérogènes telles les populations de bovins laitiers en particulier.

Pour tester l'interaction génotype **x** milieu, les premières études ont essayé de tester si le même QTL est significatif dans différents environnements. Paterson *et al*. (1991), ont détecté 29 QTL chez la tomate dans trois environnements de culture différents : quatre QTL détectés dans les trois environnements, 10 dans deux environnements différents et 15 dans un seul environnement (Paterson *et al*., 1991).

Chez la drosophile, une interaction génotype **x** milieu a été observée pour des QTL affectant le nombre de soies sensorielles (Gurganus *et al*., 1998) et la durée de vie (Leips et Mackay, 2000; Vieira *et al*. 2000). Chez les bovins, plusieurs études basées sur des approches de génétique quantitative ont rapporté l'existence d'une interaction génotype **x** milieu pour plusieurs caractères ; citons l'exemple de la quantité de matière protéique (Calus *et al*., 2002).

II. Cartographie de QTL

Les approches de cartographie de QTL appartiennent globalement à deux catégories principales :

- L'approche « gène candidat » : cette approche ne sera pas détaillée car elle n'est pas l'approche qui a été utilisée dans cette thèse.
- L'approche de cartographie génétique en utilisant les marqueurs neutres.

1. L'approche « gène candidat »

Le principe de l'approche « gène candidat » repose sur l'identification de gènes connus pour leur implication dans des fonctions bien déterminées (soit chez l'espèce étudiée, soit chez d'autres espèces). En plus, cette approche tire profit des analyses statistiques d'association réalisée pour la restriction des domaines de recherche à certaines parties du génome. Cette

approche repose sur le fait d'avoir prédit l'identité du ou des « bon(s) » gène(s) dans les régions génomiques prédéfinies et suppose, par conséquent, une bonne connaissance du (des) phénomène(s) physiologique(s) sous-jacent(s). Dans le cas contraire et c'est en général le cas, ce type d'approches est inadéquat car la vraie fonction biologique n'est bien connue que pour quelques gènes.

2. L'approche de cartographie génétique en utilisant les marqueurs neutres

Le deuxième type d'approches de cartographie de QTL a recours à un balayage de tout le génome (en Anglais : « genome scan ») à l'aide de marqueurs neutres. Ce deuxième type d'approches de cartographie de QTL peut être divisé en trois sous-approches :

- *Cartographie de QTL par analyse de liaison :* cette approche nécessite une densité de marqueurs moléculaires faible à moyenne et la connaissance de la généalogie des individus.
- *Cartographie de QTL par étude d'association :* cette approche nécessite une forte densité en marqueurs. La connaissance de la généalogie des individus n'est pas forcément nécessaire.
- *Cartographie de QTL par LDLA* (pour Linkage Disequilibrium and Linkage Analysis) : cette approche est une combinaison des deux premières approches et nécessite donc une forte densité en marqueurs.

La cartographie de QTL par analyse de liaison et celle par LDLA sont deux approches complémentaires dans le sens où la première est robuste mais peu résolutive alors que la deuxième est hautement résolutive mais résulte en un nombre élevé de faux positifs dû principalement aux tests multiples réalisés. Les deux types d'approches font partie des étapes du clonage positionnel dont l'objectif est l'identification de gènes influençant un caractère donné à partir de la seule connaissance de sa localisation chromosomique (et permet donc de s'affranchir d'une connaissance préalable de la biologie du caractère). La première étape du clonage positionnel consiste à réaliser un balayage du génome par analyse de liaison permettant de mettre en évidence de larges régions chromosomiques ayant une influence sur le caractère en question. Ensuite, l'approche LDLA est utilisée pour réduire l'intervalle de localisation du QTL. L'étape suivante consiste à étudier les gènes candidats fonctionnels et positionnels dans les régions génomiques ayant une influence sur le caractère donné identifiées par l'approche LDLA.

a) Cartographie de QTL par analyse de liaison

La cartographie de QTL par analyse de liaison nécessite la présence de pedigrees (familles) et de marqueurs moléculaires polymorphes. Elle consiste à tester à chaque position *x* du génome

définie par rapport aux marqueurs moléculaires adjacents, l'hypothèse de ségrégation de QTL dans le pedigree ; ce qui revient à comparer les performances de groupes de descendants établis selon l'allèle reçu de leurs parents à la position *x* considérée. La probabilité de transmission des allèles à la position *x* est calculée conditionnellement aux génotypes aux marqueurs flanquants**.**

(1) Les marqueurs moléculaires

(a) Les qualités nécessaires des marqueurs moléculaires pour la cartographie de QTL

Pour les généticiens, les marqueurs moléculaires utilisés pour la cartographie de QTL doivent:

- Etre polymorphes, c'est à dire présenter plusieurs variants alléliques permettant ainsi de suivre les évènements de recombinaisons au fil des générations.
- Avoir une répartition homogène sur l'ensemble du génome.
- Etre co-dominants pour que tous les génotypes puissent être identifiés sans ambiguïté (homozygotes et hétérozygotes).

(b) Notion d'informativité des marqueurs dans la détection de QTL

La détection de QTL nécessite que les marqueurs utilisés soient informatifs c'est-à-dire qu'ils permettent de tracer leur transmission des parents aux descendants sans ambiguïté. En d'autres termes, s'ils permettent d'identifier, au niveau du descendant, lequel des allèles est hérité du père et lequel est hérité de la mère. Par exemple, dans le cas où les deux parents sont hétérozygotes et leur descendant est homozygote ou dans le cas où les parents sont homozygotes et leur descendant est hétérozygote, le marqueur est informatif. Par contre, lorsque les parents et le fils ont le même génotype, l'origine des allèles reçus par le descendant n'est plus identifiable et le marqueur est non informatif.

(c) Types de marqueurs utilisés dans la cartographie de QTL

Les marqueurs moléculaires les plus utilisés actuellement dans la cartographie de QTL sont les microsatellites et les SNP.

(i) Les microsatellites

Les microsatellites sont des répétitions d'un nombre variable de di, tri, ou tétranucléotides ($(CA)_n$ ou $(CAT)_n$ ou $(ACGG)_n$ par exemple). Ils sont très fréquents dans les génomes des mammifères et très polymorphes du fait d'un taux de mutation élevé et qui est estimé chez l'homme à 10^{-4} et 10^{-2} par locus et par génération (Lai et Sun, 2003). Au cours des 15 dernières années, les microsatellites ont été les marqueurs moléculaires de choix puisqu'ils

sont très polymorphes, répartis de façon homogène dans le génome (un microsatellite tous les 100 Kb en moyenne), et co-dominants. Cependant, les microsatellites présentent l'inconvénient que leur génotypage n'est pas facilement automatisable.

(ii) Les SNP

Les SNP sont des variations nucléotidiques d'une seule paire de bases du génome. Généralement, ce sont des polymorphismes bialléliques. Ils sont de loin les polymorphismes les plus fréquents dans le génome. Les SNP présentent l'avantage que leur génotypage est facilement automatisable par rapport aux microsatellites et multiplexable sur des supports de type puce à ADN, avec plusieurs dizaines de milliers de génotypes réalisées en parallèle.

(2) Dispositifs animaux utilisés dans la cartographie des QTL par analyse de liaison chez les animaux domestiques

(a) Les croisements expérimentaux

Ce genre de croisements nécessite un court intervalle entre les générations et un nombre de descendants moyen ou élevé.

Le principe des croisements expérimentaux est le suivant: une génération F1 est créée en croisant deux races pures (ou lignées) très divergentes pour le caractère d'intérêt. La population F1 est ensuite soit autocroisée pour créer une génération F2 (trois classes de génotypes sont alors créées 11- 12 – 22 [en se plaçant dans le cas d'un QTL biallélique]) soit incluse dans un croisement de retour ou backcross avec l'une des races parentales (dans ce cas il y a deux classes de génotypes créés : 11 et 12). En sélection animale, le choix entre la F2 et le croisement de retour est le plus souvent dicté par des considérations pratiques. Un pedigree de croisement de retour nécessite la génération de peu de mâles F1 qui peuvent ensuite être accouplés à des femelles de l'une des races parentales. Une F2 nécessite, par contre, une génération d'un plus grand nombre de femelles F1.

Certains croisements expérimentaux ont aussi été générés chez les bovins tels que le dispositif F2 de l'INRA du Pin au Haras (Domaine de Borculo, Orne) créé à partir de deux races bovines laitières: la Normande et la Holstein (Larroque *et al.*, 2002). Toutefois, les croisements expérimentaux sont très longs à mettre en œuvre (du fait de l'intervalle de génération élevé). C'est pour cette raison que la plupart des programmes de recherche de QTL ont cherché à identifier des QTL à partir de structures familiales préexistantes.

(b) Les structures familiales préexistantes : exemple des bovins laitiers

Les bovins laitiers ont une structure familiale constituée de grandes familles de demi-frères grâce à la large diffusion de l'IA qui a permis l'instauration de différents modèles familiaux

adaptés à la cartographie de QTL. Deux types de structures familiales fréquemment utilisés ont été décrits pour la détection de QTL :

- ✓ Le protocole « filles » ou « daughter design » (DD) qui consiste en des familles de demi-sœurs. Les filles sont génotypées et évaluées sur la base de leurs propres performances.
- ✓ Le protocole « petites filles » ou « grand daughter design » (GDD) qui consiste en des familles de taureaux de demi-frères (fils) issus du même taureau d'IA (père). Les fils sont génotypés et leur phénotype correspond à leur valeur génétique évaluée sur la base des performances de leurs filles. Le phénotype des fils peut être exprimé sous forme de « daughter yield deviation » ou DYD : c'est la moyenne des performances des filles du taureau corrigée pour les effets environnementaux et la valeur génétique de leur mère. Le GDD est plus puissant, dans la détection des QTL que le DD puisque les performances utilisées dans le GDD sont corrigées pour un plus grand nombre d'effets indépendants, la variabilité résiduelle est donc réduite.

(3) Principes de l'analyse de liaison dans un dispositif expérimental du type GDD

(a) Principes de l'analyse de liaison à un marqueur

L'allèle reçu au marqueur (M_1 ou M_2) permet de discriminer deux groupes dans la population : P_1 et P_2. La mise en évidence d'un QTL lié au marqueur considéré repose sur la confrontation de la moyenne des performances des deux groupes de descendants identifiés selon leur génotype au marqueur (Figure C-3). Toutefois, plus le marqueur est loin du QTL et plus le nombre de recombinaisons augmente et par conséquent plus l'effet apparent du marqueur sera faible.

La moyenne des performances de la population P_1 est $\mu_{P1} = ½\alpha (1-\theta) - ½ \alpha\theta = ½\alpha (1-2\theta)$.

La moyenne des performances de la population P_2 est $\mu_{P2} = ½\alpha\theta - ½\alpha (1-\theta) = -½\alpha (1-2\theta)$.

L'effet apparent du marqueur est donc $\delta = \mu_{P1} - \mu_{P2} = \alpha (1-2\theta)$.

- Si le marqueur est complètement lié au QTL ($\theta=0$), l'effet apparent du marqueur correspond à celui du QTL.
- Si le marqueur est indépendant du QTL ($\theta=1/2$), l'effet apparent du marqueur est nul.

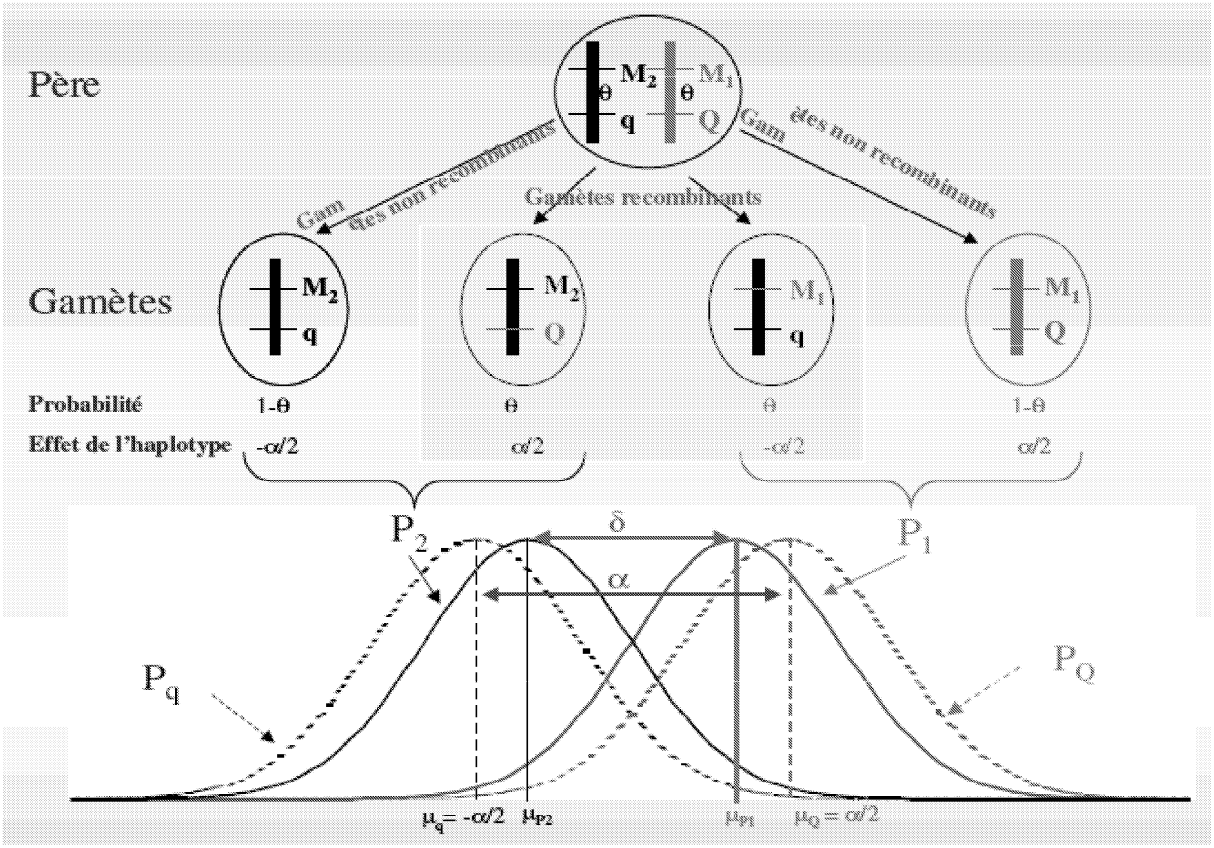


Figure C-3 : *Principe de détection de QTL par analyse des effets de ségrégation d'un QTL à deux allèles chez un père hétérozygote sur la distribution des performances de sa descendance. Le suivi de la ségrégation du QTL se fait à travers celle des allèles d'un marqueur moléculaire situé au voisinage du QTL. Les distributions en pointillés présentent l'effet réel du QTL alors que celles en ligne continue présentent l'effet apparent du marqueur lié. La différence entre les deux types de distributions est due aux recombinaisons entre le marqueur et le QTL (Gautier, 2003).*

(b) Principe de la cartographie par intervalles (« interval mapping »)

La cartographie par intervalles a été proposée par Lander et Botstein en 1989. Elle permet de faire une distinction entre l'effet du QTL et sa position relativement aux marqueurs.

La transmission d'un haplotype (c'est à dire de la combinaison d'allèles de plusieurs loci qui sont transmis ensemble sur le même chromosome) constitué de deux marqueurs successifs M et N est considérée au lieu de celle d'un seul marqueur. La probabilité de présence de l'un des deux allèles alternatifs du QTL est testée conditionnellement aux génotypes des deux marqueurs flanquants les plus proches (en supposant l'absence d'interférence entre ces deux marqueurs). Cette probabilité tient compte d'une éventuelle recombinaison entre le QTL et les deux marqueurs flanquants.

Selon l'haplotype reçu du père, on distingue quatre classes de descendants : deux classes de recombinants (ayant reçu Mn ou mN) et deux classes de parentaux (ayant reçu MN ou mn) (Figure C-4).

L'effet apparent du QTL est donné par la différence entre les moyennes des performances des groupes parentaux (δ_{par}).

La moyenne attendue du groupe ayant reçu MN est :

$$\frac{(1-\theta 1)(1-\theta 2)\alpha/2+\theta 1\theta 2\times 0}{(1-\theta)}=\frac{(1-\theta 1)(1-\theta 2)\alpha}{2\times(1-\theta)}$$

La moyenne attendue du groupe ayant reçu (mn) est :

$$\frac{(1-\theta 1)(1-\theta 2)\times 0+\theta 1\theta 2\times \alpha/2}{(1-\theta)}=\frac{\theta 1\theta 2\alpha}{2\times(1-\theta)}$$

La différence entre les moyennes des 2 groupes parentaux (MN) et (mn) est $\mu_{MN} - \mu_{mn}$:

$$\frac{(1-\theta 1)(1-\theta 2)\alpha}{2\times(1-\theta)}-\frac{\theta 1\theta 2\alpha}{2\times(1-\theta)}=\frac{\alpha(1-(\theta 1+\theta 2))}{2\times(1-\theta)}=\frac{\alpha(1-\theta)}{2\times(1-\theta)}=\alpha/2$$

De même, la différence entre les moyennes des 2 groupes de recombinants est $\mu_{Mn} - \mu_{mN}$:

$$\frac{\alpha}{2}\times\frac{[\theta 2(1-\theta 1)-\theta 1(1-\theta 2)]}{\theta}=\frac{\alpha}{2}\times\frac{(\theta 2-\theta 1)}{\theta}$$

En supposant la stricte additivité des distances génétiques entre les 2 marqueurs, $\theta_1 + \theta_2 = \theta$. Connaissant θ, on obtient alors un système à deux équations et à deux inconnues qui permettra de séparer l'effet du QTL (α/2) de sa position relativement aux deux marqueurs.

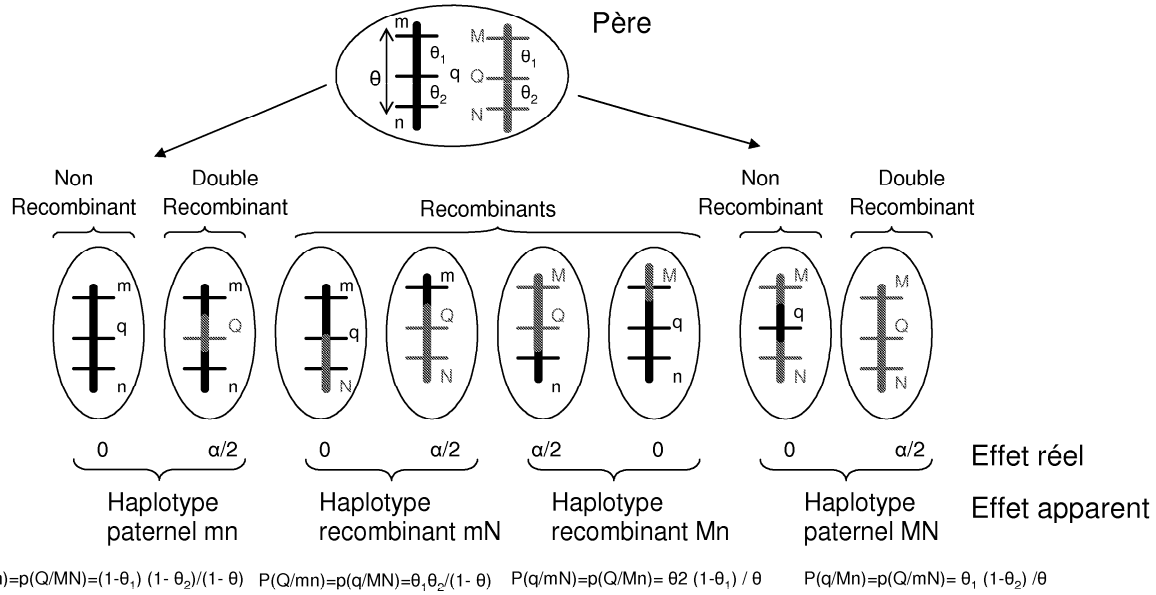


Figure C-4 : *Principes de détection de QTL en utilisant la méthode de cartographie par intervalles. Les notations sont conformes à celles utilisées dans le texte (Gautier, 2003).*

❖ Importance de l'informativité des marqueurs dans la cartographie par intervalles

L'informativité des marqueurs est un paramètre très important dans les méthodes de cartographie par intervalles. En effet, dans les principes de détection décrits précédemment, la détection du QTL se fait grâce aux marqueurs flanquant la position testée du QTL et plus précisément grâce à l'identification de l'origine parentale de ces marqueurs. De plus, les recombinants apportent la plus grande part de l'information pour la cartographie de QTL. Les marqueurs doivent donc permettre d'identifier d'une manière univoque l'haplotype reçu du père par les descendants (en ne prenant en considération que la voie paternelle). Dans le cas où deux marqueurs consécutifs sont informatifs, la probabilité de transmission, de père à fils, d'un fragment d'ADN délimité par deux marqueurs et portant l'un des deux allèles au QTL (en se mettant dans le cas d'un QTL biallélique) est très élevée et tend vers 1, alors que celle dc l'autrc cst trcs faiblc ct tcnd vcrs 0 ; par contrc, si lcs dcux marqucurs flanquant lc fragment d'ADN contenant l'un des deux allèles du QTL ne sont pas informatifs, les deux probabilités de transmission des deux allèles alternatifs du QTL sont égales. L'informativité au niveau de tout le pedigree est la généralisation de la mesure de cette probabilité de transmission au niveau de tous les intervalles formés par deux marqueurs consécutifs. Le choix des marqueurs doit donc conduire à maximiser l'informativité du protocole.

(4) Mise en place du modèle d'analyse de liaison

La cartographie de QTL par analyse de liaison consiste à mettre en place un modèle permettant de tester deux hypothèses :

- Une hypothèse nulle H_0 : « il n'y pas de QTL à la position *x* testée » qui se traduit par un effet nul du QTL à la position testée :H_0 : « $\alpha=0$ »
- Une hypothèse alternative H_1 : « il y a un QTL à la position *x* testée » qui se traduit par un effet non nul du QTL à la position testée :H_0 : « $\alpha\neq0$ »

Généralement, les méthodes de cartographie de QTL par analyse de liaison supposent aussi :

- Un QTL biallélique (les deux allèles seront notés Q et q)
- Une distribution normale des performances (phénotypes) des individus homozygotes aux deux allèles du QTL ($Y_{ij}/Q \sim N(\mu_i + \alpha/2, \sigma_i^2)$ et $Y_{ij}/Q \sim N(\mu_i - \alpha/2, \sigma_i^2)$.
- Les mères sont homozygotes au QTL (ne transmettent que l'allèle q par exemple).

La probabilité de la présence du QTL est testée conditionnellement aux génotypes des deux marqueurs informatifs flanquants les plus proches.

(5) Estimation des paramètres du modèle :

(a) Estimation des paramètres du modèle par régression linéaire

Le modèle par régression linéaire a été proposé par Haley et Knott (1992) et appliqué aux dispositifs du type DD ou GDD par Knott et al. (1996). Il s'écrit, pour une famille i, de la manière suivante :

$$Y_{ij} = \mu_i^x + p_{ij}^x \left(\frac{\alpha_i^x}{2}\right) + \left(1 - p_{ij}^x\right)\left(-\frac{\alpha_i^x}{2}\right) + \varepsilon_{ij}^x = \mu_i^x + x_{ij}\left(\frac{\alpha_i^x}{2}\right) + \varepsilon_{ij}^x \; avec \; x_{ij} = 2p_{ij}^x - 1$$

Où :

- μ_{ij} est la valeur phénotypique moyenne du fils j du père i.
- p^x_{ij} est la probabilité de transmission de l'un des deux allèles du QTL conditionnellement aux génotypes aux marqueurs flanquants ($1-p^x_{ij}$ est la probabilité de transmission de l'autre allèle du QTL conditionnellement aux génotypes aux marqueurs flanquants).
- $\alpha^x_i/2$ est la moitié de l'effet de substitution du QTL chez le père i.
- ε^x_{ij} est l'erreur résiduelle, relative au fils j du père i, du modèle.

μ_i^x et $\alpha^x_i/2$ sont estimés indépendamment pour chaque famille i.

Les valeurs de μ_i et $\alpha_i/2$ correspondent à celles qui minimisent la somme des carrés des résidus :

$$SSE_i = \sum_{J_i} (Y_{ij} - \widehat{Y_{ij}})^2$$

$$\hat{\alpha} = \frac{2Cov(Y_i, Z_i)}{Var(Z_i)} \; et \; \hat{\mu_i} = \overline{Y_i} - \frac{\alpha_i}{2}\overline{Z} \text{ où (pour une famille ayant J fils) :}$$

$$Y_i = \begin{bmatrix} Y_{i1} \\ \vdots \\ Y_{iJ} \end{bmatrix}$$

et

$$Z_i = \begin{bmatrix} 2p_{i1} - 1 \\ \vdots \\ 2p_{iJ} - 1 \end{bmatrix}$$

La significativité de l'effet du QTL à l'intérieur d'une famille i est testée par la statistique de test :

$$F_i = (J_i - 2) \times SSR_i / SSE_i$$

Où, SSR_i est la somme des carrés des écarts dus à l'effet du QTL dans la famille i.

$SSR_i = \sum_{ij} (\widehat{Y_{ij}} - \overline{Y_i})^2$ avec $\widehat{Y_{ij}} = \hat{\mu}_i + x_{ij}(\hat{\alpha}_{ij}/2)$.

Sous l'hypothèse d'absence de QTL (H_0 : α_i=0), la statistique de test F_i est distribuée selon une loi normale de 1 et de J-2 degrés de liberté.

La statistique de test F calculée pour l'ensemble des I familles (sous l'hypothèse d'une indépendance des familles) est :

$$F = \frac{\sum_I SSR_i / I}{\sum_I SSE_i / (J_i - 2I)}$$

Sous l'hypothèse d'absence de QTL (H_0: α_i=0), la statistique de test F est distribuée selon une loi normale de I et de J degrés de liberté.

Si la probabilité critique p_c correspondant à la valeur de la statistique du test F est faible (>0,01 par exemple), l'hypothèse H_0 de la non nullité de l'effet du QTL testée à la position x, est rejetée avec une probabilité (un risque) p_c de se tromper. La question de la significativité des résultats sera abordée plus loin.

(b) Estimation des paramètres du modèle par maximum de vraisemblance

Cette méthode a été proposée par Lander et Bostein (1989) pour tester la présence de QTL dans un intervalle délimité par deux marqueurs chez deux lignées consanguines divergentes pour un caractère quantitatif. Elle a été appliquée par Georges *et al*. (1995) dans un dispositif GDD. Cette méthode consiste à calculer sur l'ensemble du pedigree et en chaque position *x*, les vraisemblances L(*x*) de l'observation des performances Y_k des fils pour chaque famille de père j sous les deux hypothèses alternatives citées précédemment :

- H_0 ($L_0(x)$) : « α=0 (père homozygote au QTL)» : cette hypothèse a une vraisemblance $L_0(x)$.
- H_1 ($L_1(x)$) : « α≠0 (père hétérozygote au QTL) » : cette hypothèse a une vraisemblance $L_1(x)$.

La vraisemblance de la famille d'un père hétérozygote i est :

$$L_i^{Qq} = \prod_j \left[\left(p_{ij} \times \frac{1}{2\pi\sigma} \times e^{\frac{-(y_{ij} - \bar{y}_i + \alpha/2)^2}{2\sigma^2}} \right) + \left((1 - p_{ij}) \times \frac{1}{2\pi\sigma} \times e^{\frac{-(y_{ij} - \bar{y}_i - \alpha/2)^2}{2\sigma^2}} \right) \right]$$

De la même manière on calcule la vraisemblance d'un père homozygote L_i^{QQ} et L_i^{qq}.

La vraisemblance pour l'ensemble du dispositif est :

$$L(\alpha, \sigma, f_q) = \prod_I [f_Q^2 \times L_i^{QQ} + f_Q f_q \times L_i^{Qq} + f_q f_Q \times L_i^{qQ} + f_q^2 \times L_i^{qq}]$$

L'estimation des paramètres du modèle se fait par maximisation de la vraisemblance qui peut être réalisée par des algorithmes EM (Acronyme pour Expectation Maximisation).

La comparaison entre les deux hypothèses s'effectue en utilisant le rapport de vraisemblance (LOD score en Anglais) : $\text{lod}(x)=\log_{10}(L_1(x)/L_0(x))$.

Le plus souvent, on utilise le Likelihood Ratio Test (LRT) :

$\text{LRT}(x)= 2\ln(L_1(x)/L_0(x)) = 2\ln(10)\text{Lod}(x) = 4{,}6\ \text{Lod}(x)$

LRT suit un χ^2 d'un nombre de degrés de liberté égal à la différence entre le nombre de paramètres estimés sous H_1 et H_0. Le LRT est calculé en chaque point x du chromosome considéré.

(6) Estimation des seuils de significativité

La détection de QTL fait appel à des tests multiples dont le nombre est égal au nombre de positions testées sur le génome. Dans chacun de ces tests (correspondant à chacune des positions testées, chaque position est délimitée par deux marqueurs), l'hypothèse nulle « pas de QTL en ségrégation en cette position » est testée. Normalement, l'hypothèse nulle est rejetée lorsque la probabilité du test statistique sous l'hypothèse nulle est au-dessus d'un certain seuil (déterminé par l'expérimentateur). Toutefois, si plusieurs marqueurs sont testés, plusieurs hypothèses nulles peuvent satisfaire à ce critère de rejet, par simple chance, conduisant ainsi à la détection de faux QTL positifs, le nombre de faux positifs augmentant avec le nombre de tests réalisés. Les seuils de rejet doivent donc être ajustés pour la multiplicité des tests. Deux méthodes principales sont utilisées pour gérer le problème des tests multiples : celle des tests de permutation (Doerge et Churchill, 1996) et celle du contrôle des taux de fausses découvertes ou « false discovery rate » (FDR) (Benjamini et Hochberg, 1995).

(a) Les tests de permutations (Doerge et Churchill, 1996)

Cette méthode consiste à associer aléatoirement les phénotypes aux génotypes de manière à les dissocier en faisant des milliers de permutations entre les génotypes et les phénotypes. Des « genome scan » sont ensuite réalisés sur des milliers (10000 par exemple) de permutations. Les valeurs des tests statistiques obtenues avec les données réelles sont comparées avec la distribution des plus grandes valeurs obtenues pour chaque permutation pour déterminer son niveau de significativité. Un QTL est dit :

- Suggéré s'il est rencontré au moins une fois par « genome scan »

- Significatif s'il est détecté une fois tous les 20 « genome scan » (5 faux positifs tous les 100 « genome scan » ; ce qui correspond à une p-value de 5%).

(b) Le taux de fausses découvertes ou « false discovery rate » (FDR)

Le FDR (Benjamini et Hochberg, 1995) est la proportion de rejets erronés de l'hypothèse nulle H_0 par rapport au nombre total de rejets. Benjami et Hochberg (1995) ont montré que lorsque les tests statistiques sont indépendants, le FDR est contrôlé à un seuil q déterminé par l'équation suivante : $m_0/m <= q$ où m est le nombre total d'hypothèses testées et m_0 est le nombre d'hypothèses nulles vraies.

(7) Méthodes d'estimation de l'intervalle de confiance de localisation d'un QTL

L'intervalle de confiance (IC) d'un QTL donne une idée de la fiabilité de sa position prédite. Les IC des QTL peuvent être estimés par deux types de méthodes principales : celles par « bootstrapping » et celles par « Lod drop off »

(a) Méthodes par « bootstrapping » (Visscher *et al.*, 1996)

Cette approche consiste à créer un jeu de données *in silico*, par n tirages avec remise, des N observations expérimentales (chacune consiste en un phénotype lié à un génotype). Pour chaque tirage, on cherche et on stocke la position la plus probable (celle qui maximise la statistique de test). Les N positions sont ensuite classées par ordre croissant. L'$IC_{x\%}$ est délimité par la ((100-x/2)*N/100)ème et la ((x/2)*N/100)ème position entre lesquelles se trouvent x% des solutions.

(b) Méthodes par « Lod drop off »

Elle s'appuie sur l'utilisation de la propriété du LRT qui suit asymptotiquement un Chi-deux (χ^2) d'un nombre de degrés de liberté égal à la différence entre les nombres de paramètres estimés sous H1 et H0. En conséquence, si on fixe la position du QTL en x_0 (position du maximum de vraisemblance L(x0)), la variable aléatoire $X=\sup_x (L(x)-L(x_0))$ suit un χ^2 à un degré de liberté (la position). Pratiquement, pour un niveau de seuil recherché de θ% (par exemple 95%), il suffit de retrancher la valeur du percentile Z(100-θ)% d'un χ^2 à un degré de liberté (dans notre cas, $Z_{0.05}=3.84$), pour déterminer les bornes de l'intervalle qui contiendra toutes les positions pour lesquelles la valeur L(x) sera comprise entre $L(x_0)$ et $L(x_0) - Z_\theta$. L'usage courant du « Lod Score » qui équivaut au LRT à un facteur 4.6 (=2ln10) près, conduit à une modification d'échelle à l'origine de la valeur 1 (arrondi de 3.84/4.6=0.83) souvent utilisée pour bâtir des intervalles de confiance à 95 %.

b) Cartographie de QTL par analyse d'association

La cartographie de QTL par analyse de liaison est une méthode robuste mais peu résolutive car elle n'exploite qu'une seule génération de recombinaisons et considère que les familles sont indépendantes entre elles ; ce qui n'est pas le cas en réalité car il y a des liens de parenté inter-familles négligés dans les approches d'analyse de liaison. De plus, les phases marqueur/QTL peuvent varier selon les familles ; par conséquent, les allèles des marqueurs ne renseignent pas sur l'effet de l'allèle au QTL.

Une approche complémentaire à la cartographie de QTL par analyse de liaison est celle par étude d'association. Cette approche exploite l'association permanente existant entre le QTL et un certain nombre de marqueurs du fait de la présence de petits fragments chromosomiques non recombinants identiques dans toute la population et issus d'un ancêtre commun. Si ces régions chromosomiques contiennent un QTL quelconque, les allèles de ce QTL seront associés en permanence avec le même haplotype formé par les marqueurs. L'association permanente entre allèles du QTL et ceux des marqueurs est connue sous le nom de déséquilibre de liaison entre marqueurs et QTL. Les segments chromosomiques non recombinants issus d'un ancêtre commun sont alors appelés segments IBD (pour Identical By Descent).

La disponibilité de panels de SNP constitués de plusieurs dizaines de milliers de SNP pour plusieurs espèces domestiques et la baisse des coûts de génotypage associés a rendu possible le génotypage de centaines d'animaux pour ces panels de SNP. De plus, des cartes génétiques de bonne qualité sont désormais disponibles pour plusieurs espèces domestiques dont les bovins. Tous ces avancements ont rendu possible la cartographie directe de QTL par des études d'association, à travers tout le génome.

❖ Notion de déséquilibre de liaison :

Le déséquilibre de liaison (DL) ou déséquilibre gamétique est un terme utilisé pour désigner l'association non aléatoire entre les allèles de deux ou plusieurs loci. Ces associations non aléatoires peuvent être dues à une proximité physique ou à différentes forces évolutives qui ont affecté la population au cours du temps. Parmi ces forces, on peut citer : la migration, la mutation, la sélection, la dérive génique.

Plusieurs formules abordant différents aspects de l'association non aléatoire ont été proposées pour mesurer le DL entre les allèles au niveau de deux loci. La plupart de ces formules dépendent, pour une paire de loci bialléliques, de la valeur D calculée :

D = freq (A1_B1) * freq (A2_B2) - freq (A1_B2) * freq (A2_B1) (1) où freq (A1_B1) est la fréquence de l'haplotype A1_B1 (et de même pour les autres haplotypes).

c) Combiner l'étude de liaison et le DL pour la cartographie fine de QTL : l'approche LDLA

L'analyse de liaison et le DL ont, tous les deux, des inconvénients lorsqu'ils sont analysés séparément. L'analyse de liaison est une méthode puissante pour détecter les QTL ; cependant, elle résulte en de larges intervalles de confiance des QTL détectés. Les études d'associations donnent par contre, des estimations plus précises des intervalles de confiance, mais résultent en un grand nombre de faux positifs à cause des multiples tests réalisés.

Une approche intermédiaire entre l'analyse de liaison et celle par association consiste à combiner les deux types d'approches en une seule. En effet, l'exploitation simultanée de l'analyse de liaison et du déséquilibre de liaison pour la cartographie de QTL a été proposée par plusieurs auteurs (Farnir *et al*., 2000; Meuwissen *et al*. 2002; Lund *et al*., 2003). Ce type d'approche est appelé LDLA (pour Linkage Disequilibrium and Linkage Analysis).

Brièvement, la cartographie par LDLA consiste en trois étapes :

- ✓ Estimation des phases de liaison des animaux génotypés (distinguer entre les haplotypes hérités du père et ceux hérités de la mère).
- ✓ Calcul de la matrice IBD, G_i. Cette matrice est composée de deux parties : une première sous-matrice décrivant les probabilités IBD au niveau des positions putatives du QTL entre les haplotypes des animaux fondateurs calculées à travers le DL dont l'information provient des haplotypes maternels. La deuxième sous-matrice décrit les probabilités IBD entre les animaux du pedigree à travers l'analyse de liaison dont l'information est apportée par la transmission des haplotypes paternels. Les deux sous-matrices résultent en une matrice G_i contenant les probabilités IBD entre tous les haplotypes aux positions évaluées du QTL et suppose donc que tous les animaux sont issus d'une population fondatrice unique.
- ✓ La vraisemblance d'un modèle contenant un QTL et un effet polygénique, est comparée à celle d'un modèle ne contenant que l'effet polygénique. La différence entre les deux modèles est calculée à travers la différence du Log vraisemblance des deux modèles.

Grâce à cette approche, Meuwissen *et al*. (2002) ont pu réduire l'intervalle de localisation d'un QTL affectant le taux de naissances gémellaires chez les bovins laitiers Norvégiens, à moins de 1 cM. La figure C-5 montre les différences de détection de ce QTL obtenues avec l'analyse de liaison, l'analyse d'association et la méthode LDLA (Meuwissen *et al*., 2002).

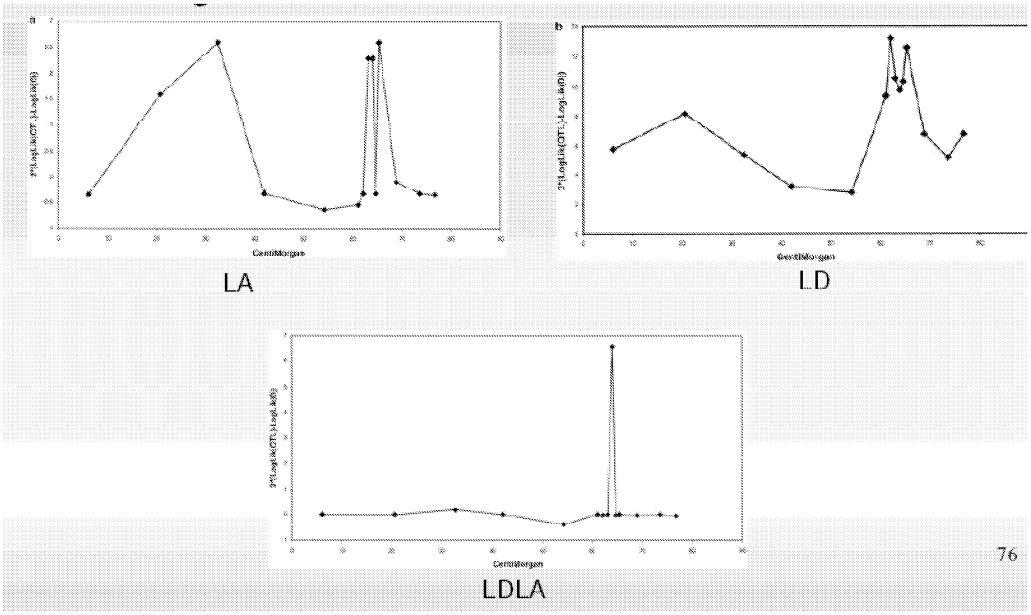


Figure C-5 : *Résultats de détection d'un QTL de naissance gémellaire sur le BTA05 chez la race bovine laitière Norvégienne, obtenus en utilisant trois méthodes différentes : l'analyse de liaison (LA), l'analyse d'association (LD) et la méthode LDLA. Cette dernière méthode a donné la localisation la plus précise en montrant un pic de localisation unique situé à 65 cM environ et un intervalle de localisation inférieur à 1 cM (Meuwissen et al., 2002).*

3. Recherche et validation de mutations ayant un effet sur un caractère quantitatif

a) Notion de QTN

La dernière étape de la cartographie de QTL consiste en la recherche et la validation de polymorphismes qui ont des effets sur le caractère étudié. Ces polymorphismes sont généralement des mutations ponctuelles appelées QTN (acronyme pour Quantitative Trait Nucleotide).

Avant d'être reconnu en tant que tel, un QTN doit d'abord être validé ; cela consiste à apporter toutes les preuves nécessaires pour démontrer sa causalité dans la variation phénotypique du caractère étudié.

b) Etapes de validation d'un QTN

Trois étapes sont nécessaires pour la validation des QTN : vérification de la concordance des QTN avec le génotype au QTL des individus, une validation statistique et une validation fonctionnelle de ces QTN.

(1) Vérification de la concordance des QTN

Cette étape suit celle de la détection de polymorphismes compatibles avec le statut a priori des individus dans des gènes candidats sélectionnés (QTN putatif), l'étape suivante est de valider les polymorphismes trouvés sur l'ensemble du pedigree. La validation se fait en

vérifiant la concordance des SNP avec les génotypes au QTL chez l'ensemble des individus du pedigree. La concordance est complète si :

- Tous les individus homozygotes au QTL le sont aussi pour le polymorphisme.
- Tous les individus hétérozygotes au QTL le sont aussi pour le polymorphisme.
- Le même allèle du QTL est associé avec le même allèle du QTN putatif pour tous les animaux hétérozygotes.

Il est probable que le QTN putatif ne soit pas en concordance complète avec le génotype au QTL pour plusieurs raisons :

- Une erreur de classement des génotypes au QTL de pères de familles à cause d'un effet faible du QTL ou d'un nombre de descendants faible.
- La concordance complète n'est attendue que si l'effet du QTL est dû à un seul site dimorphe.

Par ailleurs, la concordance ne peut être considérée comme une preuve de la détection de QTN que lorsque la probabilité de concordance par chance dans l'IC est suffisamment faible pour qu'elle puisse être rejetée statistiquement. Cohen-Zinder *et al*. (2005) ont proposé une formule généralisée par Ron et Weller (2007) pour calculer la probabilité (P_c) qu'un polymorphisme montre une concordance :

$P_c = \int_0^1 2[p\,(1-p)]^n\,[1-2p\,(1-p)]^{m)}\,dp$ où p est la probabilité d'un allèle et 1-p est celle de l'autre allèle du SNP ; n est le nombre de pères de familles hétérozygotes au QTL et m est le nombre d'homozygotes.

(2) Validation statistique des QTN

Les méthodes statistiques utilisées pour la validation de QTN ont été abordées par Cohen-Zinder *et al*. (2005). Globalement, ces méthodes essaient de prouver que :

- L'effet du QTN putatif explique toute la variabilité observée ; ce qui implique qu'il n'y ait pas d'autres polymorphismes en DL avec le QTL avec des effets significatifs sur le caractère. Ainsi, Grisart *et al*. (2004) ont trouvé des valeurs élevées de LOD scores au niveau de 10-12 marqueurs avec une valeur plus élevée pour la mutation K232A pour la quantité de matière grasse (MG). Cependant, lorsque le génotype K232A a été pris comme effet fixe dans le modèle, aucun des effets associés à d'autres marqueurs n'a été significatif. Ce qui suggère que le polymorphisme K232A explique à lui seul tout l'effet du QTL. Ces résultats ont néanmoins été contredits par Bennewitz *et al*. (2004), Kühn *et al*. (2004) et Gautier et al. (2007).
- Le même QTN est en ségrégation dans d'autres races : si l'effet du QTN putatif est maintenu dans les diverses races, il est plus vraisemblable qu'il soit causal plutôt

qu'en DL avec le QTN causal (Ron et Weller, 2007). L'effet de DGAT1 a été démontré dans les populations Holstein des Pays-Bas, d'Allemagne, d'Israël et de Nouvelle Zélande (Grisart *et al.*, 2002 ; Winter *et al.*, 2002 ; Weller *et al.*, 2003).

- Le changement des fréquences alléliques du QTN correspondent aux changements attendus dus à la sélection dans la population. Falconer (1996) a estimé le changement attendu de la fréquence allélique d'un QTL dû à la sélection, Δq, par : $\Delta q = -iaq(1-q)/\sigma_p$ où i est l'intensité de sélection, a est l'effet de substitution du QTL, q est la fréquence d'un des deux allèles du QTL (en supposant que le QTL est biallélique) et σ_p est l'écart type phénotypique standard (Falconer, 1996). Un suivi de l'évolution de la fréquence allélique de QTN a été rapporté : Weller *et al.* (2003) présentent une diminution de la fréquence de l'allèle K (allèle de DGAT1 améliorateur de la quantité de MG), dans la population Holstein d'Israël entre 1981 et 1990, de 15% à 5% et depuis elle a augmenté à 10% (Weller *et al.*, 2003) ; alors que Grisart *et al.* (2004) ont trouvé que la fréquence de l'haplotype portant l'allèle K de DGAT1 a subi une sélection positive (Grisart *et al.*, 2004). De même, Van Laere *et al.* (2003) rapportent une fréquence de l'allèle favorable Q de IGF2 plus élevée dans les populations commerciales porcines de l'Europe que dans les populations indigènes de la Chine (Van Lacre *et al.* 2003). Cette différence de fréquence allélique a été attribuée à une sélection intensive pour la croissance et les caractères de carcasse dans les populations commerciales (Yang *et al.*, 2006). De même, la fréquence de l'allèle 581Y de ABCG2 a diminué de 0,75 à 0,62, entre 1982 et 1990 chez les populations de vaches israéliennes. Cette fréquence a ensuite augmenté à 0,77 en 2002 (Cohen-Zinder *et al.*, 2005).

(3) Validation fonctionnelle des QTN

La validation fonctionnelle d'un QTN constitue un élément de preuve important pour démontrer la causalité du QTN.

- Une manière de valider fonctionnellement un QTN putatif, est de démontrer des différences entre le produit des allèles alternatifs du QTN :

- Au niveau des ARNm, c'est le cas de la mutation IGF2 chez le porc où il a été démontré que les individus portant la mutation ont une quantité d'ARNm, dans le muscle, trois fois plus importante.
- Au niveau des protéines, c'est le cas de DGAT1 : les allèles alternatifs de DGAT1 associés aux allèles + et – du QTL n'ont pas montré de différences dans la quantité d'ARNm produite (Grisart *et al.*, 2004), mais des différences significatives ont été

trouvées dans la quantité de triglycérides synthétisés, entre des cellules Sf9 exprimant l'allèle K et celles exprimant l'allèle A. Un autre exemple est celui de la mutation g+ 6723 (G-A) située dans l'UTR de la myostatine où un ratio de 3 a été trouvé, entre la quantité de myostatine produite par des animaux de phénotype sauvage et des animaux hétérozygotes (G/A) (Clop *et al*., 2006).

- Une autre manière de valider un QTN putatif est de démontrer des différences dans la fonction des deux protéines codées par les deux allèles. Chez DGAT1, par exemple, l'expression de protéines recombinantes pour DGAT1 qui diffèrent seulement au niveau de la mutation K232A, montre que cette mutation affecte la vitesse maximale (V_{max}) de l'enzyme concordante avec le phénotype observé.

Sur le plan pratique, pour la plupart des animaux domestiques ces deux types de validation peuvent être démontrés en utilisant des cultures cellulaires extraites à partir de tissus appropriés.

- L'étude de la conservation de la protéine (produit du gène candidat) entre plusieurs espèces éloignées phylogénétiquement, est aussi considérée comme un élément de validation fonctionnelle. L'étude fonctionnelle d'IGF2 a montré que la mutation a lieu dans un îlot CpG évolutivement conservé, qui de plus est hypométhylé dans le muscle squelettique.

III. Intérêts de la cartographie de QTL chez les animaux domestiques : la Sélection Assistée par Marqueurs (SAM)

L'une des applications principales de la cartographie de QTL pour les professionnels de l'élevage est la sélection assistée par marqueurs. Pour les espèces présentant un long intervalle entre les générations, la rentabilité de la SAM par rapport aux méthodes de sélection classiques, est plus évidente dans le cas de caractères à faible héritabilité, limités à un seul sexe ou mesurés d'une manière tardive au cours de la vie de l'animal. La fertilité femelle des bovins laitiers satisfait aux trois caractéristiques citées précédemment. La rentabilité de la SAM dépend principalement du progrès génétique réalisé (à court et à long terme) des caractères sélectionnés par SAM par rapport à celui assuré par les méthodes de sélection traditionnelles.

Le progrès génétique réalisé ΔG, d'un caractère donné s'écrit :

$\Delta G=i*a*\sigma/\Delta t$ où i est l'intensité de sélection, a est la précision des index de sélection, σ est l'écart type du caractère et Δt est l'intervalle entre deux générations successives. La SAM

permet ainsi d'augmenter le progrès génétique (ΔG) en augmentant la précision des index de sélection (a) et en diminuant l'intervalle de temps (Δt).

1. Principe de la SAM

La théorie de base de l'incorporation des marqueurs moléculaires ayant un effet direct sur un caractère quantitatif dans les index de sélection a été développée pour la première fois par Neimann-Sorensen et Robertson (1961) et Smith (1967).

D'une manière générale, la SAM consiste à utiliser les marqueurs moléculaires comme cofacteurs pour améliorer la prédiction des valeurs génétiques individuelles d'un caractère quantitatif d'intérêt. D'un point de vue pratique, cela consiste à trouver le(s) QTL le(s) plus significatif(s) présent(s) dans le génome et à sélectionner pour ce(s) QTL défini(s) par les deux marqueurs flanquants.

L'information issue des marqueurs est incluse dans un modèle BLUP (acronyme de Best Linear Unbiased Predictor), souvent utilisé pour l'estimation des valeurs génétiques des animaux candidats à la sélection (Fernando et Grossman, 1989) :

$Y=Zu + Qq + e$

où Y est le vecteur des phénotypes, u est le vecteur des effets polygéniques. Z et Q sont les matrices d'incidence, q est le vecteur des effets du QTL. Le modèle de Fernando et Grossman suppose que un effet aléatoire du QTL c'est-à-dire que les effets du QTL varient d'un animal à l'autre. La corrélation entre les effets du QTL de deux individus quelconque dans le pedigree est égale à la probabilité IBD entre ces deux individus. La variance des effets des QTL, V(q) est alors égale à $G\sigma_q^2$ où G est la matrice IBD entre tous les allèles du QTL et σ_q^2 est la variance due au QTL.

2. Types de sélection assistée par marqueurs :

Il n'est pas nécessaire d'identifier les polymorphismes ayant un effet sur des caractères d'intérêt pour mettre en place la SAM. D'ailleurs très peu d'études sont arrivées jusqu'à l'identification de QTN. Les intervalles de localisation de QTL délimités par des marqueurs moléculaires génotypés sur des dispositifs animaux de détection de QTL sont souvent utilisés pour faire de la SAM. Selon la proximité des intervalles de localisation du QTN recherché, on peut distinguer trois générations de SAM.

a) SAM de première génération ou SAM1

Lorsque les marqueurs utilisés dans la SAM sont loin du QTN, la SAM mise en place est alors une SAM de première génération. Elle peut être mise en place en se basant sur les résultats d'un simple balayage du génome à l'aide de trois à quatre marqueurs polymorphes par chromosome (distants de 15 à 50 cM). Cependant, la SAM1 n'a pas une efficacité

optimale étant donné que les marqueurs utilisés sont loin du QTN et par conséquent présentent de fortes chances de recombinaison entre marqueurs et QTN.

b) SAM de deuxième génération ou SAM2

La SAM2 utilise des marqueurs moléculaires qui sont en déséquilibre de liaison avec le QTN recherché dans la population. Contrairement au premier type de marqueurs, les marqueurs utilisés dans ce type de SAM sont suffisamment proches du QTN pour qu'il y ait un DL entre le marqueur et le QTL à l'échelle de toute la population (de 1 à 5 cM selon l'étendue du DL qui dépend de la structure et de l'histoire de la population). La SAM2 a une bonne efficacité étant donné le nombre limité de recombinaisons qu'il peut exister entre le marqueur et le QTN. Cependant, la SAM2 nécessite un grand nombre de marqueurs pour qu'elle soit appliquée à l'échelle de tout le génome ; ce qui est devenu possible pratiquement grâce à la disponibilité des SNP et à la baisse des coûts de génotypage.

c) SAM de troisième génération ou SAM3

La SAM3 exploite directement les QTN identifiés. Son utilisation est justifiée par le fait que la SAM2, malgré son efficacité élevée, présente deux limites majeures :

- La liaison qui existe entre les allèles des marqueurs et ceux d'un QTL ne fournissent des informations, ni sur la structure des séquences, ni sur l'organisation des QTN car les QTL et les marqueurs qui les délimitent sont situés dans des régions génomiques différentes.
- L'inférence de la position des QTN à partir des marqueurs qui les entourent est affectée par le degré de correspondance entre le génotype des marqueurs et le phénotype (autrement dit de l'héritabilité du caractère), par la densité en marqueurs et par la population étudiée.

L'identification de QTN est plus adéquate pour l'incorporation des données moléculaires dans les indices de sélection car elle permet de cibler directement les polymorphismes responsables de la variation du caractère et de s'affranchir ainsi de la nature de la population étudiée.

d) La sélection génomique

La raison pour laquelle une faible fraction de la variance génétique est expliquée par le(s) QTL détecté(s) est due à la stringence des tests statistiques réalisés afin d'éviter les faux positifs. L'idée de la sélection génomique est de contourner l'obstacle du seuil de significativité en estimant simultanément les effets de toutes les portions chromosomiques (chaque portion est délimitée par deux marqueurs successifs) sur le caractère de manière à pouvoir sélectionner directement sur les effets estimés de chaque portion chromosomique (Meuwissen *et al.*, 2001).

Grâce à la disponibilité de marqueurs génétiques très fréquents dans le génome et aux avancements réalisés dans les techniques de biologie moléculaire, des cartes génétiques à haute densité sont désormais disponibles pour l'espèce bovine. L'application de la sélection génomique est devenue possible.

Meuwissen *et al*. (2001) ont montré, en utilisant des méthodes Bayésiennes (qui supposent une distribution a priori, de la variance associée à chaque portion chromosomique), que la précision de prédiction de la sélection génomique serait de 0,84, alors qu'elle n'est que de 0,32 lorsqu'il n'y a que les effets les plus grands inclus dans le modèle.

D. Matériel et méthodes

I. Matériel Animal

1. Pedigree d'étude

a) Pedigree de l'étude de primo-localisation

Un dispositif GDD composé de 78 familles: 47 Prim'Holstein, 18 Normandes et 13 Montbéliardes, a été utilisé dans cette étape de primo-localisation. Les 78 familles sont décrites dans le tableau M-1 et d'une façon plus détaillée dans l'annexe 2. Ces familles ont été choisies parmi les autres familles disponibles selon deux critères principaux :

- Maximisation du nombre de descendants par famille.
- Minimisation du lien de parenté entre familles.

Les animaux intégrés à l'étude appartiennent à deux grands programmes de détection de QTL :

- Un programme QTL initial (Boichard *et al.*, 2003) dont le dispositif comprenait 14 familles : neuf Prim'Holstein, trois Normandes et deux Montbéliardes (Annexe 2). Tous les individus de ces familles sont nés entre 1988 et 1998. Plusieurs QTL, notamment des QTL de fertilité femelle sur les chromosomes BTA01, BTA07 et BTA21, ont été détectés dans ces familles.
- Un programme SAM (Boichard *et al.*, 2002) : ce programme représente une double valorisation des résultats du programme QTL initial. Il offre pour la profession une première information moléculaire pour appliquer une SAM de première génération afin d'assister les méthodes classiques d'évaluation de la valeur génétique des taureaux d'IA. Ce programme offre aussi un avantage pour les programmes de recherche en cours en leur permettant d'inclure de nouvelles familles dans le dispositif de détection.

Un total de 4993 taureaux d'IA a été utilisé lors de cette étude. Au début de la thèse, tous les ADN étaient déjà disponibles puisque les individus sont issus de programmes déjà préexistants. Les extractions d'ADN avaient été réalisées à partir de doses de semence et plus rarement à partir d'échantillons de sang.

b) Pedigree de l'étude de cartographie

Un ensemble d'animaux composé de 41 familles appartenant aux trois races a été sélectionné à partir des 78 familles de l'étude de primo-localisation pour affiner la localisation des quatre QTL détectés sur le BTA01, BTA02, BTA03 et BTA21. Une description plus détaillée des 41 familles de l'étude de cartographie est présente dans l'annexe 3. Deux critères ont servi pour le choix des familles :

- Le nombre de familles informatives pour les QTL détectés dans le programme SAM a été maximisé (fertilité femelle mais aussi caractères de production et de résistance aux maladies).
- La surreprésentation de familles étroitement liées a été évitée.

Les 41 familles sont composées de 26 familles Prim'Holstein, de neuf familles Normandes et de six familles Montbéliardes pour un nombre total de 2925 individus. Les 14 familles utilisées dans le programme QTL initial ont été incluses dans cet échantillon. Un résumé des caractéristiques des familles pour les trois races est présent dans le tableau M-1.

Tableau M-1 : *Caractéristiques des familles du dispositif de l'étude de primo-localisation et de celui de l'étude de cartographie pour chacune des trois races.*

Dispositif	**Race**	**Nombre de familles**	**Nombre moyen de fils/famille**	**Taille de la plus petite famille**	**Taille de la plus grande famille**
Primo-localisation	Prim'Holstein	47	70	21	239
	Normande	18	55	34	74
	Montbéliarde	13	59	25	98
Cartographie	Prim'Holstein	26	68	84	36
	Normande	9	59	63	40
	Montbéliarde	6	64	70	41

2. Données phénotypiques

Les performances des fils utilisées dans le modèle de détection de QTL correspondent à des DYD, dérivées des valeurs génétiques calculées par l'équipe « bovins laitiers » de la SGQA (Station de Génétique Quantitative et Appliquée, INRA, Jouy-en-Josas) sur la base des résultats (réussite/échec) à l'IA de leurs filles après correction pour les effets environnementaux. La précision du prédicteur des indices est évaluée, pour chaque fils, par le CD (ou coefficient de détermination) de valeur génétique et dépend du nombre de filles utilisées pour l'évaluation du fils en question (plus ce nombre est élevé et meilleure est la précision : le CD tend vers 1).

La primo-localisation des QTL s'est fait en utilisant les DYD du TNR calculé à 281 jours après IA (TNR281). Ensuite, pour l'étude de cartographie, une caractérisation plus fine de la fertilité femelle a été réalisée : d'autres indices de fertilité dérégressés calculés à partir de taux de non retour à 56 et 90 jours après IA, ont été élaborés lors d'un stage à la SGQA (Guillaume *et al.*, 2007). L'étude de cartographie a donc utilisé trois TNR différents : TNR56, TNR90 et TNR281.

II. Marqueurs moléculaires

1. Choix des microsatellites

Au cours de l'étude de primo-localisation, les données de génotypage pour les 12 régions SAM étaient déjà disponibles à partir du programme SAM (Boichard *et al.*, 2006). Au total 39 marqueurs microsatellites, avec deux à cinq microsatellites par région génomique, ont été utilisés pour la primo-localisation (Annexe 3b).

Sur la base des résultats de primo-localisation, de nouveaux marqueurs microsatellites ont été choisis dans les régions d'intérêt des trois chromosomes BTA01, BTA02 et BTA21 à partir de la carte génétique USDA (http://www.marc.usda.gov/genome). Deux critères ont été utilisés pour le choix des marqueurs :

- Les marqueurs doivent assurer une couverture des intervalles de primo-localisation la plus uniforme possible.
- Les marqueurs doivent être le plus polymorphes possible.

En ce qui concerne le BTA03, ce chromosome a fait l'objet d'une étude préalable (Guillaume *et al.*, 2007) pour laquelle 19 microsatellites au total ont été utilisés pour cartographier le QTL sur l'échantillon des 41 familles de l'étude de cartographie. Ensuite, pour affiner l'intervalle de localisation du QTL, cinq autres marqueurs microsatellites ont été ajoutés. Ces microsatellites (INRA3001, INRA3002, INRA3003, BZ847539 et CC582410) ont été sélectionnés et développés à partir de l'assemblage de la séquence bovine Btau_2.0 (ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/fasta/Btau20050310-freeze/README.2.0.txt), seul assemblage disponible à l'époque. Le choix de développer des microsatellites à partir de cet assemblage permettait une définition quasi-optimale de leur position (les cartes génétiques publiées étaient par ailleurs limitées en microsatellites pour cette zone du BTA03). Leur position attendue sur le BTA03 a été confirmée par cartographie d'hybrides irradiés (voir page 82 pour une description du principe et du travail de cartographie RH de ces microsatellites).

Trois des cinq marqueurs initialement définis ont été éliminés à cause d'un profil illisible sur gel lors de la cartographie RH (INRA3002, INRA3003 et BZ847519). Les deux autres microsatellites : INRA3001 et CC582410, en plus de URB006, ont été génotypés sur les 41 familles. La vérification des génotypes des trois races a montré qu'un microsatellite (CC582410) est monomorphe pour les animaux de race Prim'Holstein et a donc été éliminé de l'analyse pour cette race.

Un seul marqueur sur le BTA03 a finalement été ajouté pour la race Prim'Holstein, ce qui s'est avéré insuffisant pour affiner la localisation du QTL sur ce chromosome. D'autre part, la détection du QTL du BTA03, chez la seule race Prim'Holstein, et le peu de marqueurs

localisés dans la région de localisation du QTL (région centromérique de BTA03), a rendu obligatoire l'ajout de nouveaux microsatellites. Ainsi, 29 microsatellites supplémentaires ont été sélectionnés à partir d'un assemblage de la séquence bovine plus récent : Btau_3.1 (ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/fasta/Btau20060815-freeze/ReadMeBovine.3.1.txt). L'objectif était de définir un seul multiplex qui serait génotypé sur les 26 familles de la race Prim'Holstein. La position et l'ordre des 29 microsatellites ont été établis par cartographie d'hybrides irradiés (RH). Six des 29 microsatellites ont été éliminés à cause de la persistance de produits non spécifiques dans les réactions d'amplification des séquences correspondantes. La position et l'ordre des 23 microsatellites restants ont été confirmés par cartographie RH en utilisant le logiciel CarthaGène (Schiex et Gaspin, 1997; http://www.inra.fr/mia/T/CarthaGene/) (voir page 83 pour la description du travail réalisé avec CarthaGène). Quatre des 23 microsatellites présentaient des données brutes RH identiques deux à deux, ce qui signifie que ces quatre microsatellites sont étroitement liés deux à deux. L'un des deux microsatellites identiques a donc été éliminé. Vingt et un microsatellites développés à partir de la séquence bovine Btau3.1 sont finalement disponibles. De plus, trois microsatellites supplémentaires (BMS2904, DIK069 et DIK4193) ont été sélectionnés à partir de la base de données publique USDA (http://www.marc.usda.gov/genome). Comme il a été précédemment précisé, un seul multiplex devait être défini pour le génotypage des individus des 26 familles Prim'Holstein. D'autres microsatellites ont ainsi été éliminés. Au total, neuf microsatellites ont été retenus pour la mise au point d'un nouveau multiplex nommé BTA03B:

- Les trois microsatellites de la carte USDA ont été gardés parce qu'ils présentaient un bon niveau de polymorphisme.
- Six des 21 microsatellites développés à partir de la séquence bovine ont été choisis de telle manière à ce qu'il y ait une répartition la plus homogène possible de marqueurs dans l'intervalle de localisation du QTL.

La composition de ce multiplex est décrite dans l'annexe 1.

Au total, 26 microsatellites ont été inclus dans les analyses de détection de QTL de fertilité femelle sur le BTA03 pour la race Prim'Holstein et 20 microsatellites pour les races Normande et Montbéliarde.

❖ Cartographie par analyse d'hybrides d'irradiation ou cartographie RH

✓ Principe de la cartographie RH

La cartographie RH, tout comme la cartographie génétique, a pour objectif de déterminer l'ordre et la distance séparant deux marqueurs génétiques. Toutefois, au lieu d'utiliser les recombinaisons méiotiques, la cartographie RH utilise les cassures induites par irradiation

afin de déterminer la distance séparant deux marqueurs. Le principe de base de la cartographie par analyse d'hybrides d'irradiation est le suivant : pour une dose d'irradiation initiale *x* donnée (exprimée en Rad de rayon gamma), plus deux marqueurs sont proches, plus rarement ils seront séparés par une cassure chromosomique et vice versa. La cartographie RH possède deux avantages principaux par rapport à la cartographie génétique :

- Elle ne nécessite pas de marqueurs polymorphes comme c'est le cas en cartographie génétique.
- Elle présente l'avantage d'être cumulative c'est-à-dire qu'une fois un locus est génotypé pour un panel RH, il n'est plus nécessaire de le génotyper de nouveau lorsque que d'autres loci sont ajoutés et analysés. Les données issues de tous les marqueurs peuvent ainsi être analysées simultanément.

✓ <u>Confirmation de la position et de l'ordre des microsatellites sélectionnés à partir de l'assemblage bovin sur le BTA03</u>

Trente quatre microsatellites (cinq microsatellites (INRA3001, INRA3002, INRA3003, BZ847539 et CC582410) en plus des 29 microsatellites sélectionnés ultérieurement) ont été génotypés pour un panel RH construit dans le cadre du projet européen ComRad Map (Williams *et al*., 2002).

Le panel RH est contenu dans une plaque de 96 puits (94 puits contenant chacun 5 µl de solutions d'ADN d'une lignée d'hybrides irradiés, et un puit contenant un témoin positif, un autre contenant un témoin négatif).

Dans chacun des puits, 10 µl de mix PCR contenant 3,3 µl de solution tampon 5X flexi (promega) sont mélangés à 1,32 µl de $MgCl_2$, 0,13 µl de dNTP et 6 µl d'eau. Un volume de 0,1 µl d'enzyme « Go Taq Flexi » (promega) est ensuite distribué dans chacun des 96 puits avant de lancer la PCR. Les conditions PCR sont les suivantes :

- Température d'hybridation variant entre 58°C et 64°C (selon le marqueur testé)
- Temps d'hybridation : 20 secondes
- Temps d'extension : 20 secondes
- Temps d'élongation : 20 secondes
- Nombre de cycles : 30

Les produits PCR sont déposés sur gels d'agarose à 1,5% en présence d'une solution bleu de charge afin de vérifier les profils obtenus. Les données brutes obtenues à l'issue des réactions PCR sont introduites dans le logiciel CarthaGène. L'ordre des marqueurs est ensuite établi en utilisant les différentes options de CarthaGène. CarthaGène a été conçu pour la construction de cartes génétiques et de cartes d'hybrides irradiés. En règle générale, la première étape de construction de la carte RH d'un chromosome consiste à effectuer une analyse multiple par

paires de marqueurs pour laquelle la vraisemblance (L) d'une carte RH serait le produit de la vraisemblance des données pour chaque paire de marqueurs. Des « Lod Score 2 points » (Lod2pt) sont calculés pour chaque paire de marqueurs. La vraisemblance de la carte est calculée à partir du nombre de clones RH qui sont ++, --, +- et -+. En supposant que la probabilité de rétention P est constante tout au long du chromosome, l'expression de L en fonction de la probabilité de cassure θ et de rétention P s'écrit :

$L(\theta,P) = [P^2\theta + P(1-\theta)]^a [PQ\theta]^{b+c} [Q^2\theta + Q(1-\theta)]^d$ (Lawrence et Morton, 1992).

Où a, b, c et d représentent respectivement, le nombre de clones RH ++, +-, -+ et -- dans un échantillon de n= a+b+c+d clones étudiés pour les deux loci.

Sur la base des Lod2pt obtenus, les marqueurs sont regroupés dans des groupes de liaison. La commande « group » de CarthaGène permet de fixer une valeur (LOD score) seuil de regroupement.

Une première stratégie généralement adoptée consiste à commencer par un seuil faible (LOD score de 3 par exemple) pour identifier les marqueurs isolés. Les seuils sont ensuite augmentés progressivement de manière à réduire les groupes à une taille compatible avec une utilisation efficace des algorithmes de détermination de l'ordre des marqueurs (généralement une vingtaine de marqueurs par groupe). Une fois que l'on a plusieurs groupes de liaison, une analyse de l'ordre des marqueurs à l'intérieur de chaque groupe est réalisée. La commande « nicemapl » permet de faire un premier niveau d'amélioration en utilisant des Lod2pt. D'autres algorithmes exécutés par d'autres commandes sont généralement utilisés pour améliorer l'ordre des marqueurs : les commandes « annealing », « greedy » et « polish ».

Une fois l'ordre des marqueurs le plus vraisemblable établi, l'option « flips » permet de vérifier et valider cet ordre. Cette option effectue des permutations, par fenêtres de n marqueurs, de l'ordre des marqueurs afin de vérifier s'il existe un ordre meilleur.

Une deuxième stratégie consiste à établir en premier lieu une carte « framework » avec quelques marqueurs c'est-à-dire une carte où l'ordre des marqueurs introduits est sûr. L'option « buildfw » permet de créer un tel type de cartes. CarthaGène construit des cartes « framework » en utilisant une méthode d'insertion progressive en introduisant les marqueurs non inclus dans tous les intervalles possibles de la carte « framework ». Pour un marqueur donné, la différence de la « log vraisemblance » entre le meilleur ordre et le deuxième meilleur ordre, est utilisée pour qualifier le marqueur ; si cette différence est supérieure à un seuil déjà fixé par l'expérimentateur, le marqueur est inséré dans la carte à cette position ; s'il y a d'autres ordres dont la différence de la « log vraisemblance » avec le meilleur ordre est inférieure à un deuxième seuil (fixé aussi par l'expérimentateur), ils sont alors retenus comme éventuels points de départ pour le prochain marqueur à insérer.

Finalement, lorsque tous les marqueurs insérables sont introduits dans la carte de départ, la nouvelle carte est retenue et les autres marqueurs sont provisoirement insérés puis enlevés, un par un, dans toutes les positions possibles de la carte. Pour chaque marqueur, la meilleure position est marquée avec un signe « + ».

Remarque :

Avant toute analyse avec CarthaGène, il est important de visualiser les marqueurs ayant les mêmes profils de rétention et de ne garder que ceux qui diffèrent afin d'alléger au maximum les calculs ultérieurs. La commande « mrkdouble » permet d'identifier ces marqueurs.

2. Multiplexage des marqueurs pour le génotypage

a) Principe

Le multiplexage des marqueurs c'est-à-dire la présence simultanée de plusieurs marqueurs dans un seul milieu réactif (ou dans un même multiplex), est une pratique courante et qui permet d'économiser des quantités importantes d'ADN et de réactifs nécessaires. Cette technique consiste à équilibrer le milieu réactif, en tenant compte de la taille des différents allèles des marqueurs, de telle sorte que tous les allèles puissent être distingués sans ambiguïté. Une contrainte majeure à la création de multiplex est la taille des allèles des différents marqueurs d'un multiplex. En effet, pour qu'ils soient distinguables les uns des autres, les intervalles de variation des tailles de ces allèles ne doivent pas être chevauchants. L'utilisation de différents fluorochromes (voir paragraphe suivant) permet, néanmoins de résoudre en partie ce problème. En effet, il existe trois fluorochromes différents qui peuvent être utilisés pour distinguer les allèles de microsatellites ayant des intervalles de variation chevauchants (outre celui utilisé pour le marqueur de taille interne).

b) Amplification et visualisation des ADN

L'amplification des marqueurs est réalisée avec l'une des deux amorces (forward ou reverse) marquée en 5' par un fluorochrome. Trois fluorochromes différents ont été utilisés pour le marquage des fragments amplifiés et un fluorochrome pour le marquage du marqueur de taille ; ce dernier correspond à un ensemble de fragments d'ADN de taille connue nécessaires à l'identification des différents allèles lors des analyses ultérieures.

Les quatre fluorochromes sont :

- ✓ Tamra : fluorochrome de couleur rouge utilisé pour le marqueur de taille.
- ✓ 6-Fam : fluorochrome de couleur bleue.
- ✓ Hex : fluorochrome de couleur jaune.
- ✓ Tet : fluorochrome de couleur verte.

Les amorces fluorescentes ont été synthétisées par la société MWG Biotech AG.
Les produits d'amplification sont ensuite déposés sur gel d'acrylamide dans un séquenceur automatique (ABI prism 377, Applied BioSystems) lors de l'étape de mise au point et injectés dans un séquenceur à 96 capillaires (MegaBace 1000, Molecular Dynamics – GE HealthCare Life Sciences) lors de l'étape de génotypage sur l'ensemble du dispositif. Dans les deux types de séquenceurs, un rayon laser de longueur d'onde prédéterminée est utilisé au cours de la migration, pour exciter les fluorochromes contenus dans les produits d'amplification et produire ainsi un signal lumineux de longueur d'onde spécifique. Les données brutes sont ensuite analysées à l'aide d'un logiciel de génotypage permettant une meilleure interprétation des profils. Ces logiciels sont « Genotyper » pour ABI prism 377 et « Genetic Profiler v1.5» pour le MegaBace 1000.

c) Mise au point des multiplex et génotypage des individus

Les marqueurs des chromosomes BTA01, BTA02, BTA03 et BTA21 ont été regroupés en huit multiplex différents en fonction de la gamme de variation des allèles et de la chronologie des mises au point pour le BTA03 (deux multiplex ont été mis au point à deux moments différents pour le BTA03).Une marge de 20 pb a été le plus souvent retenue entre la taille maximale de l'allèle du microsatellite le plus petit et la taille minimale de l'allèle du microsatellite immédiatement plus grand. La température d'amplification varie d'un multiplex à une autre. Les caractéristiques de chacun des huit multiplex ainsi que les marqueurs éliminés sont décrits dans l'annexe 1**.** Chacun des groupes de marqueurs définissant un multiplex a été mis au point à partir d'échantillons d'ADN de concentration calibrée (20 ng/µl) et en utilisant la quantité d'amorces de chacun des marqueurs comme variable d'optimisation. Les produits PCR ont été ensuite déposés sur un séquenceur ABI377 (Applied BioSystems) pour être visualisés et analysés à l'aide du logiciel Genotyper (Applied BioSystems).

(1) Amplification par PCR

Le protocole de PCR pour réaliser deux mélanges réactionnels indépendants a été suivi afin de limiter les amplifications non spécifiques :

- Le premier mélange réactionnel contient le mélange d'amorces (amorces reverse et forward) des différents marqueurs du multiplex avec les concentrations définies lors de la mise au point.
- Le deuxième mélange réactionnel est un mélange « général » commun à tous les multiplex. Il contient du $MgCl_2$ (2 mM), les dNTP (0,25 mM), la solution tampon (1x) et la « Go Taq Flexi » polymérase (Promega) (0,1 unité) ajoutée avant le lancement de la réaction PCR.

Les réactions de PCR sont réalisées dans un volume final de 12 µl (2 µl d'ADN de concentration variant de 10 à 20 ng/µl, 5 µl de mélange amorces et 5 µl de mélange général).

(2) Dépôt sur gel, visualisation des produits PCR et analyse des profils

Les mises au point des multiplex sont réalisées sur un séquenceur de type ABI Prism 377. Dans le protocole, le produit PCR est dilué au 1/5ème (2 µl d'ADN dans 8 µl d'eau) et est ajouté à un mélange de dépôt contenant le marqueur de taille (kit Genscan-500 Rox), de la formamide (agent dénaturant) et 3 µl de solution bleu de charge. Un volume final de 8 µl est déposé dans le gel d'acrylamide vertical après dénaturation des produits de PCR pendant 3 min à 94°C.

Le gel d'acrylamide est préparé selon les méthodes standards à partir de 15 ml de solution d'acrylamide, dans laquelle sont introduits 100 µl de persulfate d'ammonium (10%) et 10 µl dc Tcmcd afin dc catalyscr la rćaction dc polymćrisation. L'analysc dcs gćnotypcs dcs échantillons est réalisée avec le logiciel Genotyper.

L'ensemble des individus a été ensuite génotypé à l'aide d'un séquenceur à 96 capillaires (MegaBace 1000, Molecular Dynamics – GE HealthCare Life Sciences). Après PCR, les produits sont dilués au 1/10, puis purifiés sur colonne G50 (résine Sephadex, Pharmacia). Enfin, 3 µl de marqueur de taille (kit ET-Rox 400, Molecular Dynamics – GE HealthCare Life Sciences) sont ajoutés à 5 µl de produit PCR purifié avant dénaturation (2 min à 95°C) et injection dans le séquenceur. L'analyse des génotypes est réalisée à l'aide du logiciel Genetic Profiler v1.5.

3. Résultats des génotypages

(a) Répartition, par race, des marqueurs retenus pour l'analyse et des marqueurs éliminés

Au total 47 marqueurs ont été génotypés chez la race Prim'Holstein et 42 marqueurs chez les races Normandes et Montbéliardes. Un marqueur : CC582410, a été génotypé chez ces deux dernières races et ne l'a pas été chez la Prim'Holstein car non polymorphe pour cette race. Les marqueurs du multiplex «MultiBTA03_B » (annexe 1) n'ont été génotypés que pour la race Prim'Holstein.

Deux, 10 et 2 marqueurs ont été respectivement éliminés de la détection de QTL sur les chromosomes BTA21, BTA03 et BTA02. Il s'agit des marqueurs DIK2492 et INRA103 pour BTA21, des marqueurs MNS-74, BMS2790, CA065, DIK4403, BZ847519, INRA3002, INRA3003, INRA3020, INRA3022 et INRA3029 pour BTA03 et des marqueurs DIK4618 et BMS2024 pour BTA02.

Ces marqueurs ont été éliminés soit suite à la vérification de leur position par cartographie RH (pour les marqueurs du BTA03), soit au cours de la mise au point des multiplex, soit lors de la vérification des génotypes des individus suite à la détection de génotypes ambigus (profils illisibles avec le logiciel Genetic Profiler) ou absence de polymorphismes (annexe 1).

b) Nombre de génotypages disponibles

Au total, 158 835 typages ont été réalisés pour les trois races (en comptant les marqueurs utilisés lors de la primo-localisation). La répartition, par race et par type de marqueurs (SAM ou non SAM), des typages est détaillée dans le tableau M-2.

Tableau M-2 : *Répartition, par marqueur, des données de génotypage disponibles chez les trois races : Prim'Holstein (PH), Normande (NO) et Montbéliarde (MO).*

Chromosome	**BTA01**			**BTA02**			**BTA03**			**BTA21**		
	PH	NO	MO	PH	NO	MO	PH	NO	MO	PH	NO	MO
Nombre de génotypages des marqueurs SAM	6804	1844	1319	3171	986	569	6169	1577	1074	4712	1384	914
Nombre de génotypages des marqueurs « non SAM »	17381	4611	2900	25235	6580	4396	45780	9290	6650	3774	1037	678
Total trois races/ Chromosome	34859			40937			70540			12499		

c) Nombre de génotypages considérés pour la détection de QTL

La répartition, par race, des génotypes pour chacun des marqueurs ainsi que certaines caractéristiques des microsatellites utilisés pour chacun des chromosomes BTA01, BTA02, BTA03 et BTA21 sont indiqués dans le tableau M-3.

Parmi les génotypes disponibles, certains marqueurs n'ont pas été considérés dans les analyses de détection de QTL. En effet, deux marqueurs sur le BTA02 (DIK1081 et DIK4334) n'ont pas été inclus : la construction de la carte génétique sur le chromosome BTA02 a montré qu'ils étaient trop éloignés des autres marqueurs sur ce chromosome (Figure E-6) et donc n'apportaient pas d'information sur la région considérée. De même, deux marqueurs n'ont été génotypés que sur un nombre limité d'individus d'une même race : il s'agit de INRA135 sur le chromosome BTA02 et HUJI177 sur BTA03. En éliminant ces marqueurs des analyses de détection de QTL, 150 975 génotypes ont été considérés pour la détection de QTL.

Tableau D-3 : *Répartition, par race, des génotypes et caractéristiques des microsatellites utilisés pour chacun des chromosomes BTA01, BTA02, BTA03 et BTA21.*

BTA	**Marqueur**	**Position**	Individus Montbéliard génotypés (%)	Individus Normand génotypés (%)	Individus Prim'Holstein génotypés (%)	Multiplex	Nombre d'allèles
BTA01	BM4307	1	361(91%)	463(83%)	1926(90%)	SAM	8
	BMS4000	2	235(59%)	497(89%)	1949(92%)	multiA	13
	BM1312	3	369(93%)	507(91%)	2066(97%)	SAM	9
	BMS4030	4	285(72%)	479(86%)	1888(89%)	multiC	9
	DIK2886	5	243(61%)	489(87%)	1992(94%)	multiA	7
	INRA073	6	369(93%)	525(94%)	2007(94%)	SAM	4
	BM8246	7	367(93%)	531(95%)	2026(95%)	multiA	12
	DIK4587	8	365(92%)	515(92%)	2045(96%)	multiA	8
	BM864	9	355(89%)	515(92%)	1955(92%)	multiA	15
	BMS4028	10	371(93%)	526(94%)	2021(95%)	multiA	10
	DVEPC32	11	326(82%)	464(83%)	1971(93%)	multiD	8
	BMS4041	12	359(90%)	523(94%)	1845(87%)	multiA	4
	BM1824	13	222(56%)	288(52%)	974(46%)	SAM	5
BTA02	DIK1081	1	383(96%)	521(93%)	2027 (95%)	Multiplex BTA02	6
	DIK4334	2	376(94%)	521(93%)	2019 (95%)	Multiplex BTA02	4
	DIK2496	3	392(98%)	529(95%)	2048 (96%)	Multiplex BTA02	2
	DIK4673	4	254(64%)	516(92%)	2024 (95%)	Multiplex BTA02	6
	BM4440	5	372(93%)	526(94%)	1987 (93%)	Multiplex BTA02	8
	DIK2719	6	302(76%)	495(89%)	2027 (95%)	Multiplex BTA02	7
	DIK4972	7	266(67%)	474(85%)	1902 (89%)	multiB	3
	BMS778	8	295(74%)	533(95%)	1996 (94%)	SAM	9
	RM041	9	357(90%)	528(94%)	1920 (90%)	multiB	6
	DIK4208	10	378(95%)	532(95%)	1999 (94%)	multiB	4
	BMS1866	11	353(89%)	502(90%)	1933 (91%)	multiC	9
	BM1223	12	359(90%)	474(85%)	1981 (93%)	multiB	5
	INRA135	13	35(9%)	83(15%)	202 (9%)	SAM	8
	BMS2519	14	351(88%)	504(90%)	1968 (92%)	multiB	10
	BM2113	15	240(60%)	357(64%)	1029 (48%)	SAM	7
	IDVGA-2	16	264(66%)	475(85%)	1715 (81%)	multiB	9

BTA	Marqueur	Position	Individus Montbéliard génotypés (%)	Individus Normand génotypés (%)	Individus Prim'Holstein génotypés (%)	Multiplex	Nombre d'allèles
BTA21	BMS1117	1	322 (81%)	496 (89%)	1884 (89%)	multiC	5
	DIK2492	2	344 (86%)	525 (94%)	1883 (89%)	multiC	4
	ILSTS103	3	341 (86%)	525 (94%)	1874 (88%)	SAM	3
	TGLA122	4	341 (86%)	525 (94%)	1937 (91%)	SAM	8
	CSSM018	5	217 (54%)	525 (94%)	924 (43%)	SAM	7
BTA03	BMS871	1	366 (92%)	535 (96%)	2091 (98%)	SAM	6
	URB006	2	323 (81%)	472 (84%)	2022 (95%)	multiD	5
	INRA3001	3	285 (71%)	435 (78%)	1888 (89%)	multiD	7
	INRA3015	4	0 (0%)	0 (0%)	1998 (94%)	BTA03B	5
	INRA3012	5	0 (0%)	0 (0%)	2006 (94%)	BTA03B	6
	INRA3013	6	0 (0%)	0 (0%)	2003 (94%)	BTA03B	4
	INRA006	7	377 (95%)	541(97%)	2093 (98%)	multi3B	3
	DIK4193	8	0 (0%)	0 (0%)	2034 (96%)	BTA03B	3
	ILSTS096	9	355 (89%)	408 (73%)	1999 (94%)	SAM	8
	BMS2904	10	0 (0%)	0 (0%)	2023 (95%)	BTA03B	4
	DIK2101	11	379 (95%)	542 (97%)	2079 (98%)	multi3B	6
	RM019	12	346 (87%)	499 (89%)	1998 (94%)	multi3A	4
	BMS2522b	13	270 (68%)	539 (96%)	2067 (97%)	multi3A	9
	MNB-65	14	365 (92%)	526 (94%)	2032 (95%)	multi3B	6
	DIK4196	15	365 (92%)	533 (95%)	2000 (94%)	multi3A	8
	BMS963	16	376 (94%)	542 (97%)	2044 (96%)	multi3A	6
	BMS482	17	378 (95%)	543 (97%)	2065 (97%)	multi3B	8
	DIK2434	18	370 (93%)	543 (97%)	2043 (96%)	multi3A	11
	MNB-86	19	351 (88%)	478 (85%)	1908 (90%)	multi3B	12
	BL41	20	366 (92%)	510 (91%)	1994 (94%)	multi3B	6
	DIK069	21	0 (0%)	0 (0%)	2016 (95%)	BTA03B	2
	DIK2609	22	377 (95%)	542 (97%)	2112 (99%)	multi3A	2
	BM4129	23	312 (78%)	543 (97%)	1941 (91%)	multi3A	3
	INRA003	24	333 (84%)	482 (86%)	1831 (86%)	multi3B	11
	ILSTS029	25	369 (93%)	530 (95%)	2030 (95%)	multi3A	7
	HUJI177	26	147 (37%)	307 (55%)	1237 (58%)	SAM	7

4. Vérification des erreurs de typage et construction de la carte génétique

Une vérification des génotypes obtenus est réalisée à partir des informations généalogiques connues : elle permet de mettre en évidence des incompatibilités de filiation. Deux programmes développés en interne sont utilisés pour vérifier les erreurs des génotypes à cet effet. Un premier programme a été développé pour la vérification des erreurs en ne tenant compte que de la voie paternelle : ce programme reconstruit les génotypes des pères à partir de ceux de leurs descendants et ceci pour les marqueurs informatifs. Une fois les erreurs de typage par voie paternelle corrigées, un deuxième programme vérifie les erreurs de typage à travers la voie maternelle. Ce programme utilise les génotypes déjà reconstitués des pères et ceux des fils pour détecter les éventuelles erreurs de filiation à travers la voie maternelle.

La construction des cartes génétiques a été réalisée avec le logiciel CRIMAP 2.4 (Green *et al.*, 1990). L'ordre initial des marqueurs était par défaut celui de la carte USDA (http://www.marc.usda.gov/genome) pour les marqueurs issus de cette base de données et celui issu de l'analyse avec CarthaGène pour les marqueurs définis à partir de l'assemblage de la séquence bovine. En ce qui concerne la carte des SNP du BTA03 (voir IV-1), l'ordre initial des marqueurs a été celui du dernier assemblage bovin (version Btau_4.0).

La fiabilité de ces ordres a été ensuite testée en exécutant l'option « flips » de CRIMAP 2.4 en réalisant des permutations de l'ordre des marqueurs en utilisant des fenêtres de quatre marqueurs. Ensuite, l'option « chrompic » a été utilisée pour vérifier les doubles recombinants. La présence de doubles recombinaisons fréquents entre deux marqueurs assez proches dans une même famille traduit en effet soit des erreurs de génotypage des descendants, soit une erreur de reconstitution de l'haplotype du père de la famille. Les distances génétiques ont été calculées en utilisant la fonction cartographique de Haldane.

III. Cartographie des QTL par analyse de liaison

1. Modèle d'analyse

a) Détection de QTL

Un modèle d'analyse de liaison par régression linéaire (partie C-II-2-a-(5)-(a) de l'introduction) a été utilisé pour la cartographie de QTL de fertilité femelle lors des deux études de primo-localisation et de cartographie.

Ce modèle est de la forme :

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

Où:

- Y_{ij} : DYD du fils j issu du père i.

- S_i : effet du père i.

- p_{ij} : probabilité de transmission d'un allèle au QTL

- a_i : moitié de l'effet de substitution du QTL.

- c_{ij} : résiduelle du modèle.

Ce modèle est implémenté dans QTLmap (Boichard *et al.*, 2003), logiciel utilisé lors de l'étude de primo-localisation et dans QTLexpress (Seaton *et al.* 2002, http://qtl.cap.ed.ac.uk/), logiciel utilisé lors de l'étude de cartographie.

Dans les deux études, l'analyse se décompose en trois étapes :

- Estimation, pour chaque père, de la probabilité de chacun des haplotypes possibles conditionnellement aux génotypes des descendants aux marqueurs.

- Estimation, pour chaque individu, de la probabilité de transmission des allèles paternels (probabilité d'avoir reçu l'un ou l'autre fragment chromosomique du père à la position considérée) par pas de 1 cM sur l'ensemble de la région considérée à partir de l'information génétique apportée par les marqueurs. Ce calcul permet également d'évaluer l'informativité du dispositif en calculant à chaque position la valeur absolue de la différence entre les deux probabilités pour chaque individu.

- Pour chacune des familles considérées comme indépendantes et à chaque position *x* analysée (*x* variant par pas de 1 cM), les performances des fils sont régressées sur les probabilités de transmission en supposant la présence d'un QTL biallélique.

Le modèle 2-QTL (implémenté dans QTLexpress) a été testé à chaque fois où deux pics distincts et significatifs ont été détectés pour un TNR donné.

b) Evaluation du statut hétérozygote des pères au QTL

Pour chaque famille i, l'équation (E) permet de calculer la statistique de test $Ti = \sqrt{n_i}\, a_i / \sigma_{\hat{a}}$ où a_i est l'effet moyen de substitution du QTL dans la famille i, $\hat{\sigma}_a^2$ est la variance génétique additive estimée et n_i le nombre de descendants dans la famille i). Sous l'hypothèse H0 (absence de QTL donc $a_i = 0$), cette statistique suit une distribution de Student à n_{i-1} degrés de liberté. Comme n_i est grand, on peut raisonnablement approcher cette distribution par une loi normale. Ainsi, en la position testée, si $|Ti| > 1{,}96$ (seuil nominal de significativité à 5%), on peut rejeter l'hypothèse nulle ($a_i = 0$) et déclarer le père i hétérozygote.

IV. Choix des gènes candidats positionnels et fonctionnels sur le BTA03 et sélection d'individus de phénotypes extrêmes

1. Choix des régions à étudier sur le BTA03

Le choix des régions génomiques à analyser sur le BTA03 a été dicté par une analyse utilisant l'approche LDLA, faite à la SGQA sur un dispositif de 17 familles de race Prim'Holstein faisant partie des 26 familles Prim'Holstein du dispositif de cartographie. Quatre cent trente-sept SNP répartis sur tout le chromosome BTA03 ont été utilisés dans cette analyse. Le TNR90 a été utilisé comme phénotype de fertilité femelle. Cette analyse LDLA a révélé un intervalle de localisation du QTL de 3,8 cM ; celui délimité par les deux SNP : rstoul_bta3_snp_299 (16,1 cM) et rsiccare_bta3_835 (19,9 cM) (Tableau M-4) (S. Fritz, communication personnelle). Cet intervalle correspond à une distance physique de 4,77 Mb ; celui compris entre 10,93 Mb et 15,70 Mb.

Tableau M-4 : *Séquence des deux SNP délimitant l'intervalle de localisation d'un QTL FF sur BTA03 obtenu par analyse LDLA.*

Nom du SNP	Séquence
rstoul_bta3_snp_299	AGTCTGCTTGCCATGATGGTCATTGCTGCTGC ACTTCTGGGGTATTG[C/T]AGACCCTGGAGGAAAAACGGT TAGTAACTCTTTTTGGCCTTCTTGGTTA
rsiccarc_bta3_835	CCCCAGAGCCGATGCTGATTGATGAACTGTATGAGCTGGTGGT GGATGCCATCTTTGGCTTCAGCTTCACGGGTGAGGTTCGGGAGCC ATTCCG[A/C]AGCATCCTGAGTGTCCTGAATGGGCTC ACTGTGCCCATTGCGAGCAT

2. Ancrage de la région à étudier de BTA03 sur les génomes murin et humain

La région définie par l'intervalle précédemment décrit a été ancrée sur les génomes humain et murin afin de profiter de l'avancée des travaux de séquençage et d'annotation de gènes chez ces deux espèces. L'ancrage de la région d'intérêt a été fait en consultant la base de données de l'UCSC (http://genome.ucsc.edu/). Un inventaire des références de séquences humaines (RefSeq HSA) de la région a permis de dresser une liste des gènes présents à partir de la même base de données.

3. Sélection des gènes candidats à partir du génome murin

Afin de sélectionner des gènes candidats pouvant affecter d'une manière ou d'une autre la fertilité femelle, la base de données MGI (http://www.informatics.jax.org/) a été consultée

pour l'ensemble des gènes identifiés à partir des séquences de référence murines identifiées dans la région d'intérêt. Cette base de données contient une description des phénotypes identifiés de souris Knock Out (KO) ou mutantes pour quelques milliers de gènes. Les gènes candidats ont été choisis sur la base de deux critères principaux:

- L'absence d'expression du gène, ou une mutation, est à l'origine de mortalité embryonnaire précoce chez la souris.
- L'absence d'expression du gène, ou une mutation, influence un processus lié directement à la reproduction (mécanismes hormonaux, régulation de produits géniques liés à la reproduction…).

En se basant sur des phénotypes de souris KO ou mutantes décrites dans la base de données MGI (http://www.informatics.jax.org/), six gènes candidats ont été choisis parmi les 149 références : progestin and adiporeceptor family member 6, metaxin 1, spectrin alpha 1, Src domain homology2 containing transforming protein, adenosine deaminase, RNA specific et tropomyosin 3, gamma. La position de ces gènes par rapport aux pics de la courbe obtenue avec l'analyse LDLA est décrite dans la figure E-20. Leurs caractéristiques sont décrites dans le tableau M-5.

Tableau D-5 : *Positions physiques sur les génomes bovin et murin et phénotypes de souris KO et/ou processus physiologiques des six gènes candidats sélectionnés sur le BTA03.*

Gène	Symbole	Position Physique sur le BTA03 (Btau_4.0) (Mb)	Position Physique MMU (Mb) (Chr)	Processus physiologique/ phénotype souris KO
Progestin and adiporeceptor family member 6	**PAQR6**	15,88	88,450 (MMU03)	Intermédiaires des actions non génomiques des hormones stéroïdes.
Metaxin 1	**MTX1**	16,80	89,288 (MMU03)	Létalité embryonnaire (les homozygotes meurent entre 4 et 9 jours)
Spectrin alpha 1	**SPNA1**	11,97	176 (MMU01)	Létalité embryonnaire
Src Domain homology2 containing transforming protein	**SHC1**	16,99	89,50 (MMU03)	Voie de signalisation des récepteurs à œstrogène, à glucocorticoïdes et à insuline / Mortalité embryonnaire durant l'organogenèse (E11,5 - E12,5 j)
Adenosine deaminase, RNA specific	**ADAR1**	17,37	89,800 (MMU03)	Mortalité embryonnaire durant l'organogenèse (E12 j)
Tropomyosin 3, gamma	**TPM3**	17,70	90,150 (MMU03)	Mortalité embryonnaire avant l'implantation

4. Sélection et séquençage d'individus de phénotypes extrêmes pour les gènes candidats

a) Sélection des individus à partir de familles ségréguant le QTL

Seize fils de phénotypes extrêmes ont été présélectionnés à partir de quatre pères hétérozygotes identifiés lors de l'analyse de liaison (Annexe 4). A partir de ces 16 fils, huit ont été finalement sélectionnés pour constituer quatre pools d'ADN d'individus de phénotype extrême : deux pools +/+ et deux individus -/- (Tableau M-6).

Tableau M-6 : *Identifiant des individus et de leur père et valeurs et CD des TNR90 des huit individus des quatre pools de phénotypes extrêmes sélectionnés*

Race	ID_2005	TNR90j	ID_père	CD_90j	Classe	numéro du pool
66	1963	0,0275	3536	0,6556	+/+	P1
66	4093	0,0295	4373	0,656	+/+	
66	4422	0,0007	4582	0,6815	+/+	P2
66	2249	0,066	3537	0,6426	+/+	
66	2882	-0,0289	3536	0,6292	-/-	P3
66	4108	-0,0683	4373	0,6046	-/-	
66	1518	-0,0557	3537	0,592	-/-	P4
66	4423	-0,0469	4582	0,9716	-/-	

b) Amplification de l'ADN génomique des animaux

Etant donné le peu d'ADN disponible pour certains animaux et les besoins en ADN pour des génotypages supplémentaires, l'ADN génomique de 16 animaux présélectionnés ont été amplifiés à partir des quantités restantes en utilisant le Kit d'amplification « GenomiPhi HY DNA » et en suivant les instructions du fournisseur (GE Healthcare). Les réactions d'amplification ont été réalisées dans 20 µl de volume final avec des concentrations moyennes de 3 µg/µl. Les ADN amplifiés ont été ensuite dilués au 1/100ème pour la réalisation des réactions de PCR.

c) Séquençage des individus pour les gènes candidats sélectionnés à partir des résultats de l'analyse LDLA

(1) Régions séquencées

Les régions génomiques choisies pour le séquençage correspondent aux régions comprenant les exons des gènes candidats sélectionnés. Si ces derniers contiennent peu d'exons, toute la région génomique encadrant les exons est alors séquencée. Dans le cas contraire, seuls les

exons pour lesquels des mutations ont été décrites chez la souris ont été séquencés. Ce choix est justifié par le fait que l'objectif principal est d'identifier des polymorphismes en déséquilibre de liaison avec une possible mutation causale et pas nécessairement de détecter la (les) mutations causale(s) en elle(s)-même.

(2) Préparation des échantillons à séquencer

Les amorces commandées pour les régions à séquencer ont d'abord été testées sur l'ADN d'individus contrôle avant de les utiliser pour amplifier, par PCR, l'ADN génomique des quatre pools (Tableau M-6). Au cours du test, les conditions PCR ont été optimisées pour éliminer les bandes supplémentaires (résultats d'amplification aspécifique) en jouant sur trois paramètres :

- La température d'hybridation
- Le temps d'élongation
- Le nombre de cycles

Les réactions de PCR sont faites dans un volume final de 10 µl de mix PCR contenant 3 µl de solution tampon 5X Taqflexi (Promega), 1,2 µl de $MgCl_2$ (25 mM), 0,12 µl de dNTP (100 mM), 0,045 de chaque amorce (forward et reverse) et 5,15 µl d'eau. L'enzyme « Go Taq Flexi » (5 u/µl) :0,075 µl (promega) est distribué juste avant le lancement de la PCR dans chacun des puits contenant 2 µl d'ADN-test.

Une fois les réactions de PCR optimisées, les quatre pools d'ADN sont amplifiés dans un volume final de 55 µl (12 µl de solution tampon 5X Taqflexi, 4,8 µl de $MgCl_2$, 0,48 µl de dNTP, 0,18 de chaque amorce (forward et reverse) et 37,06 µl d'eau) et 5 µl d'ADN génomique.

5 µl des produits de PCR ainsi amplifiés sont ensuite déposés sur gel d'agarose à 1% pour vérifier les profils des échantillons.

Les produits PCR sont ensuite purifiés afin d'éliminer au maximum les oligonucléotides utilisés lors de l'amplification et éviter ainsi d'obtenir des séquences multiples. 50 µl d'eau sont ajoutés aux 50 µl de produit PCR et les 100 µl ainsi obtenus sont transférés dans une plaque commerciale de 96 puits (Millipore). La plaque est soumise à un vide d'air (10 bars environ) pendant 30 min puis 40 µl d'eau sont ajoutés au produit séché dans chacun des puits. La plaque est ensuite placée dans un agitateur (30 min). 10 µl de chaque produit purifié sont déposés dans un puits d'une plaque de 96 puits et envoyés à séquencer, chez un prestataire de service (Qiagen), avec une plaque miroir contenant chacune 10 µl de l'amorce (forward ou reverse) correspondante à chacun des puits de la plaque des échantillons.

L'analyse des séquences est réalisée avec le logiciel NovoSnp2.0.3 (http://www.molgen.ua.ac.be/bioinfo/novosnp/, Weckx S. *et al.*, 2005).

V. Inventaire des gènes candidats positionnels et fonctionnels en se basant sur les résultats de la puce Illumina 50K

Au cours de ce projet de thèse (plus précisément vers la fin de la thèse), une puce bovine (bovineSNP50 Genotyping Beead Chip) contenant 54000 SNP environ, a été commercialisée par Illumina en collaboration avec l'USDA ARS, l'Université du Missouri et l'Université d'Alberta. La distance moyenne séparant deux SNP consécutifs est de 51,5 Kb à travers tout le génome bovin (http://www.illumina.com/pages.ilmn?ID=256). Cette puce a été utilisée pour génotyper un dispositif animal comprenant environ 1800 taureaux d'IA de race Prim'Holstein, 800 taureaux d'IA de race Normande et de 750 taureaux d'IA de race Montbéliarde. Les SNP de la puce couvrent l'ensemble des autosomes bovins ainsi que le chromosome X. Une étude LDLA a été faite sur tout le génome à la SGQA. Deux types d'index génétiques de fertilité femelle ont été utilisés : un index de fertilité calculé pour les vaches (index de fertilité « vache ») et index de fertilité calculé pour les génisses (index de fertilité « génisse »).

Le résultat des analyses LDLA faites avec les deux types d'index sont présentés sous formes de courbes LDLA où les QTL sont identifiés par la présence de pics dépassant la valeur seuil de 1%. Pour exploiter ces résultats au niveau génomique, un inventaire des identifiants de gènes et de protéines situés dans des régions de 1 Mb de part et d'autre de la position de chacun des pics (l'intervalle de localisation de chaque pic étant donc équivalent à 2 Mb) identifiés a été réalisé. Un script développé en interne a permis de récupérer tous les identifiants de gènes et de protéines existant dans différentes bases de données (ENSEMBL, UCSC et NCBI) ainsi que leur localisation et leur position sur le chromosome bovin. L'identifiant et les positions des SNP localisés à l'intérieur de ces gènes ont été aussi récupérés dans un fichier. Le logiciel « Ingenuity Pathways Analysis » (IPA) (http://www.ingenuity.com/) a été utilisé pour analyser la liste de ces gènes. IPA est un logiciel, à fonctions multiples, qui permet d'analyser, d'identifier des réseaux de gènes impliqués dans des fonctions biologiques et de consulter la bibliographie relative aux molécules composant les réseaux identifiés à partir de différents types de données (relatives à des séquences nucléotidiques, des données de transcriptomique ou autre). Les sources d'informations d'IPA sont issues surtout des avancées des recherches acquises chez l'homme, le rat et la souris. Les objectifs de cette étape sont :

- L'identification de réseaux de gènes qui ont des fonctions biologiques surreprésentées dans l'ensemble des intervalles de localisation des QTL identifiés et des molécules qui pourraient avoir un rôle dans la reproduction de la vache laitière.
- L'identification de gènes individuels qui peuvent avoir un rôle dans la reproduction.

Le choix de ces gènes candidats a été basé sur leur influence sur la fertilité femelle à travers l'une des voies qui ont été citées dans l'introduction c'est-à-dire sur la base de :

✓ Leur influence directe (système reproductif, système endocrinien, période de gestation) ou

✓ Leur influence indirecte sur la fertilité et principalement des processus métaboliques qui peuvent être à l'origine du bilan énergétique négatif chez les vaches laitières hautes productrices. De tels processus métaboliques peuvent être liés à la synthèse de triacylglycérol, d'acide gras ou d'hormones métaboliques.

E. Résultats

I. Résultats de l'étude de primo-localisation

Les résultats de primo-localisation de QTL de fertilité femelle dans les 12 régions génomiques SAM sont exposés, par race, dans le tableau E-1. Soixante dix-huit familles au total ont été incluses dans cette analyse : 47 familles Prim'Holstein, 18 de race Normande et 13 de race Montbéliarde. Le nombre de marqueurs utilisés pour chaque région est aussi donné dans ce tableau.

Les cartes génétiques utilisées pour chaque région chromosomique sont présentées dans l'annexe 3b. Les distances génétiques entre marqueurs sont celles publiées dans la base de données USDA. Six QTL influençant le TNR281 ont été détectés chez la race Prim'Holstein : trois QTL significatifs (P<0,01) sur les chromosomes BTA01, BTA02 et BTA03 et trois QTL suggérés (P<0,05) sur les chromosomes BTA06, BTA20 et BTA23. En race Normande, un QTL significatif a été détecté sur le BTA01 et un QTL suggéré a été détecté sur le BTA21. Aucun QTL n'a été détecté en race Montbéliarde.

Tableau E-1 : *Résultats de détection de QTL, par race et en utilisant le TNR281, réalisés dans les 12 régions chromosomiques SAM.*

Chromosome	Prim'Holstein	Normande	Montbéliarde	Nombre de marqueurs markers
BTA01	**	**	N	4
BTA02	**	N	N	2
BTA03	**	N	N	4
BTA06	*	N	N	3
BTA07	N	N	N	3
BTA14	N	N	N	2
BTA15	N	N	N	3
BTA19	N	N	N	4
BTA20	*	N	N	5
BTA21	N	*	N	3
BTA23	*	N	N	2
BTA26	N	N	N	4

** : P<0.01
* : P<0.05
N : Non significatif

II. Résultats de l'étude de cartographie

L'étude de primo-localisation a permis de mettre en évidence la présence de certains QTL dans les 12 régions SAM. Une deuxième étude, utilisant un plus grand nombre de marqueurs, a été nécessaire pour affiner l'intervalle de localisation des QTL détectés. Pour cela, les régions génomiques significatives ont été choisies parmi les autres régions pour réaliser une

étude de cartographie à moyenne résolution en utilisant des marqueurs microsatellites. Ces régions concernent les chromosomes BTA01, BTA02, BTA03 et BTA21.

Il est à noter que, pour les quatre chromosomes analysés, seuls les individus génotypés pour au moins 60% des marqueurs ont été inclus dans l'analyse. Le nombre total d'individus génotypés et ceux inclus dans l'analyse sont décrits, par chromosome et par race, dans le tableau E-2. Les valeurs du test-t (Test de Student) décrit précédemment (matériel et méthodes partie III-1-b) pour tous les pères d'une race donnée où le QTL est significatif au seuil de 5%, sont données dans l'annexe 5.

Tableau E-2 : *Description, par race et par chromosome, du nombre d'individus génotypés et du nombre d'individus inclus dans les détections de QTL.*

	Prim'Holstein		**Normande**		**Montbéliarde**		**Trois races**
	Individus totaux	Individus génotypés sur plus de 60% de marqueurs (%)	Individus totaux	Individus génotypés sur plus de 60% de marqueurs (%)	Individus totaux	Individus génotypés sur plus de 60% de marqueurs (%)	Total des Individus génotypés sur plus de 60% de marqueurs
BTA01	2095	2033 (97%)	559	528 (94%)	383	367 (96%)	2928
BTA02	2099	2051 (98%)	558	538 (96%)	401	373 (93%)	2962
BTA03	2152	1854 (86%)	552	486 (88%)	395	344 (87%)	2684
BTA21	1974	1868 (95%)	536	522 (97%)	353	343 (97%)	2733

1. Chromosome BTA01

a) Données de génotypage analysées

Trois mille trente-sept individus ont été génotypés pour les trois races sur le chromosome BTA01. Les analyses de détection de QTL ont inclus 2928 individus (génotypés pour au moins 60% des marqueurs).

b) Carte génétique et densité en marqueurs

Neuf marqueurs ont été sélectionnés à l'intérieur de la zone de primo-localisation et génotypés sur un échantillon de 41 familles appartenant aux trois races pour affiner l'intervalle de localisation du QTL chez les races Prim'Holstein et Normande. Les analyses de liaison se sont finalement appuyées sur 13 marqueurs microsatellites (en comptant les marqueurs SAM). La densité en marqueurs est passée d'un marqueur tous les 20 cM à un marqueur tous les 6 cM, en moyenne. Les distances génétiques entre les différents marqueurs, obtenues avec les données du dispositif expérimental (race Prim'Holstein), celles issues de la

carte de l'USDA et celles issues de la carte physique de l'assemblage Btau 4.0, sont indiquées dans la figure E-1. Les distances génétiques deux-points entre les marqueurs sont relativement similaires entre la carte génétique obtenue avec les données expérimentales et celle de la carte USDA.

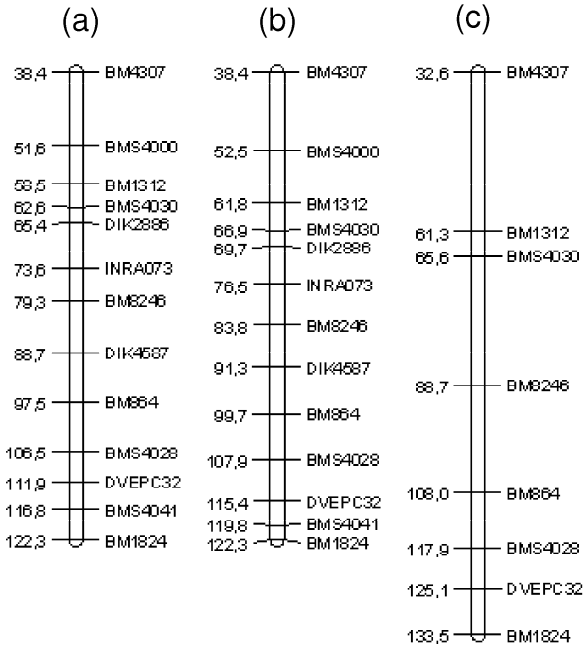


Figure E-1 : *Cartes génétiques des marqueurs microsatellites de BTA01 obtenues avec les données expérimentales (a), à partir de la base de données publique USDA (les distances sont exprimées en cM) (b) et carte physique obtenue par blast des microsatellites sur l'assemblage bovin (Btau_4.0) (les distances sont exprimées en Mégabases) (c). Certains marqueurs n'ont pas pu être ancrés sur cet assemblage. L'origine des cartes génétiques ((a) et (b)) est celle de la carte génétique de USDA.*

c) Résultats de l'analyse de liaison

(1) Informativité du dispositif

Les courbes d'informativité globale du dispositif calculée pour chacune des races et par pas de 1 cM, sont représentées dans les figures E-2, E-3 et E-4. L'informativité moyenne des marqueurs est de 0,77 ± 0,051, 0,81 ± 0,06, 0,73 ± 0,10 respectivement pour les races Prim'Holstein, Normande et Montbéliarde.

(2) Résultats de détection de QTL

(a) Modèle uni-QTL

(i) Courbes de F-value

Les courbes de F-value correspondant aux TNR56, TNR90 et TNR281 pour chacune des trois races sont représentées dans les figures E-2, E-3 et E-4. Un QTL de fertilité femelle sur le BTA01 a été détecté chez la race Prim'Holstein en utilisant les TNR90 et TNR281. La courbe de F-Value de ces deux indices de fertilité présente deux pics dont le plus marqué est situé à 97 cM (à côté du marqueur BM864). Les courbes de F-value obtenues avec les deux indices de fertilité ont la même allure (notamment en ce qui concerne la position des deux pics). Toutefois, le QTL détecté avec le TNR90 est significatif (P<0,009) alors qu'il devient suggéré (P< 0,019) lorsque les données du TNR281 sont utilisées pour la détection. L'estimation, par la méthode de bootstrap (10 000 échantillonnages), de l'intervalle de confiance du QTL détecté a donné une localisation moyenne du QTL à 95,5 cM et un intervalle de confiance à 95% ($IC_{95\%}$) compris entre 71 et 122 cM c'est-à-dire un intervalle de 51 cM. Aucun QTL n'a été détecté pour les races Montbéliarde et Normande.

(ii) Statut des pères au QTL

Les pères déclarés hétérozygotes (p<0,05), en se basant sur le test de Student, sont listés dans le tableau E-3 pour les TNR90 et TNR281 où le QTL est significatif. En comparant les deux types de données, cinq pères sont déclarés hétérozygotes avec le TNR90 et quatre sont déclarés hétérozygotes avec le TNR281. Un père de famille, le 6865, est déclaré hétérozygote avec le TNR90 alors qu'il ne l'est pas avec le TNR90.

Tableau E-3 : *Valeurs des effets de substitution allélique et du test-t des familles hétérozygotes (p<0,05) pour les TNR90 et TNR281 sur le BTA01 en race Prim'Holstein.*

Indice de fertilité	Famille	Effet de substitution	Valeur du test-t
TNR90	3442	0,0311	2,5875
	3481	0,017	2,5596
	3498	0,0141	2,9849
	6865	-0,0217	2,1744
	7210	0,0376	3,5225
TNR281	3442	0,0402	3,1328
	3481	0,0177	2,5127
	3498	0,0152	3,0098
	7210	0,0325	2,8939

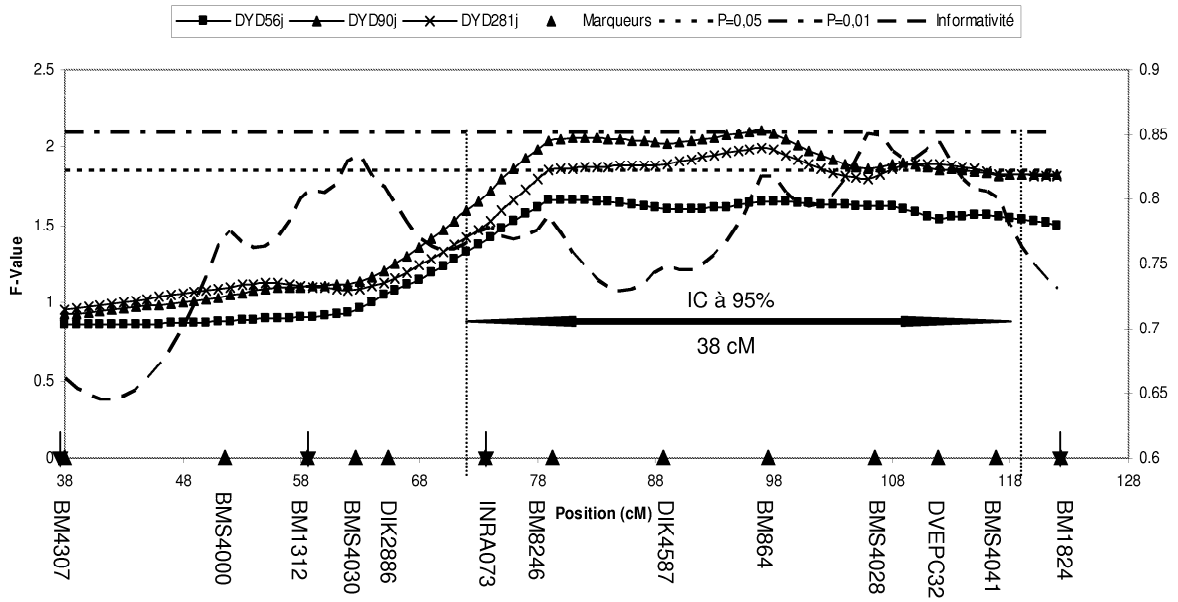


Figure E-2 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur le BTA01 chez la race Prim'Holstein. Les courbes de F-value correspondent aux TNR56 (courbe dessinée avec des carrés), TNR90 (courbe dessinée avec des triangles) et TNR281 (courbe dessinée avec des croix). L'informativité des marqueurs est décrite par la courbe en pointillés (l'axe des ordonnés de l'informativité est celui situé à droite de la figure). Les seuils de significativité correspondants ont été obtenus par 10000 permutations. L'intervalle de confiance est obtenu par bootstrap (10000 échantillonnages). Les marqueurs SAM utilisés lors de la primo-localisation sont indiqués par des flèches.*

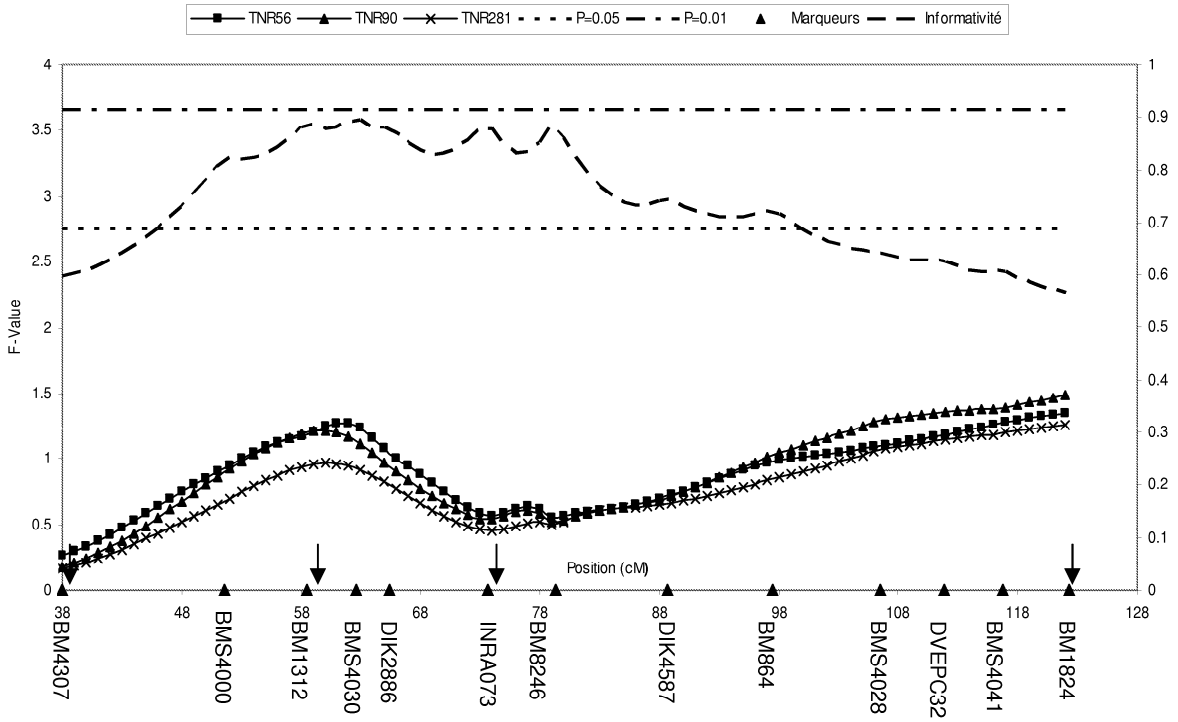


Figure E-3 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur le BTA01 en race Montbéliarde. La légende est identique à celle de la figure E-2. Les seuils de significativité correspondants ont été obtenus par 1000 permutations.*

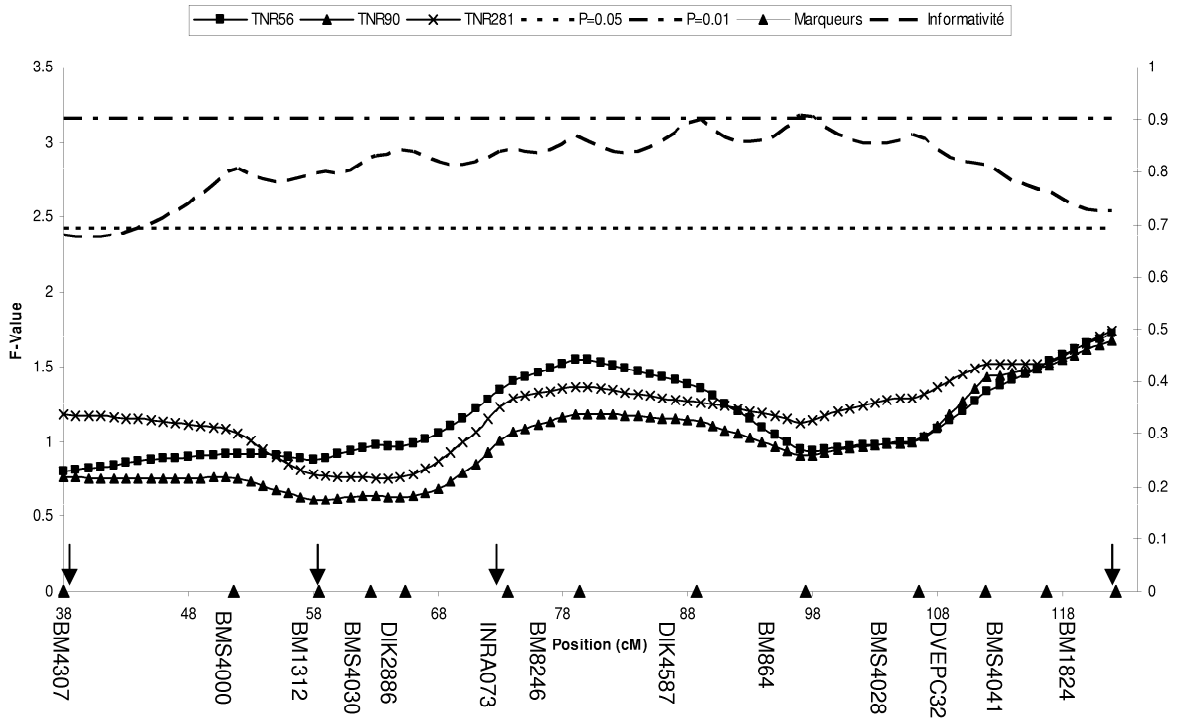


Figure E-4 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur le BTA01 en race Normande. La légende est identique à celle de la figure E-2. Les seuils de significativité correspondants ont été obtenus par 1000 permutations.*

(b) Modèle 2-QTL

L'allure de la courbe de F-value obtenue pour la race Prim'Holstein avec les TNR90 et TNR281 montre l'existence de deux pics : le premier est situé à la position 82 cM (proche de BM8246), le deuxième à 97 cM (proche de BM864). L'hypothèse de l'existence de deux QTL qui ségréguent a donc été testée. Les résultats de cette analyse sont présentés dans la figure E-5. Les positions les plus vraisemblables des deux QTL sont respectivement 82 et 98 cM. La valeur de la statistique F qui teste le modèle bi-QTL contre le modèle uni-QTL (tableau E-4) n'a pas atteint le seuil de significativité minimal de 5% et donc l'hypothèse de l'existence de deux QTL a été rejetée.

Tableau E-4 : *Valeurs des statistiques F des deux modèles « 2 QTL contre 0 QTL » et « 2 QTL contre 1 QTL » concernant les 26 familles Prim'Holstein du dispositif de l'étude de cartographie*

Modèle	Valeur de la statistique F	Significativité
2 QTL *vs* 0 QTL	1,87	Non significatif
2 QTL *vs* 1 QTL	1,63	Non significatif

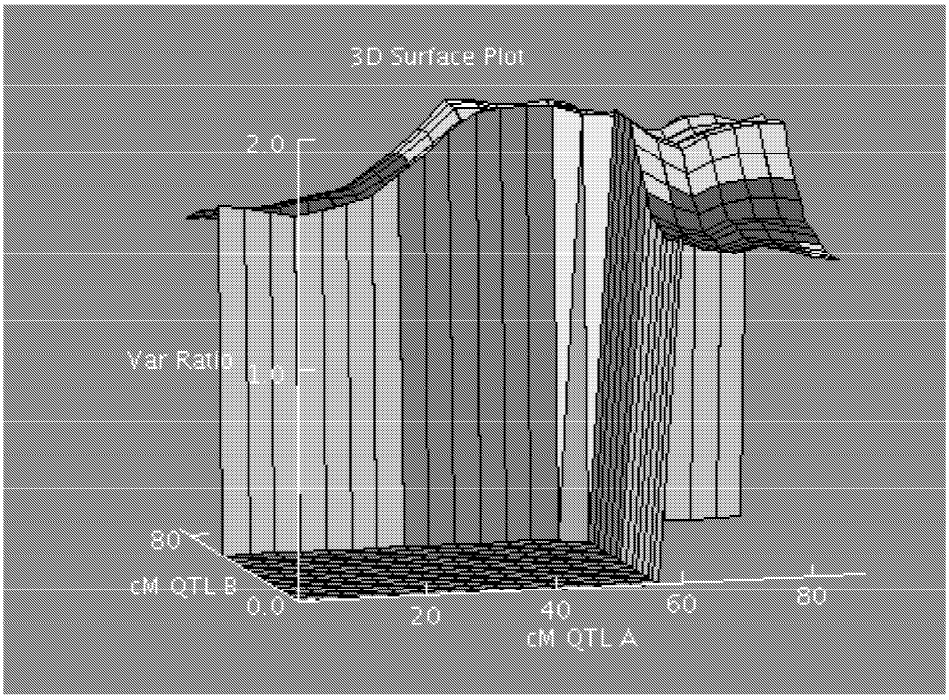


Figure E-5 : *Profil des courbes de F-value pour chaque couple de positions sous le modèle bi-QTL sur BTA01 en race Prim'Holstein. Le TNR90 a été utilisé comme phénotype de fertilité femelle. L'origine des abscisses correspond à la position de BM4307 (premier marqueur de la carte génétique utilisée dans l'étude de cartographie).*

2. Chromosome BTA02

a) Données de génotypage analysées

Trois mille cinquante-huit individus ont été génotypés pour les trois races et pour les 13 microsatellites sur le chromosome BTA02. Deux mille neuf cent soixante-deux individus sont inclus dans les analyses (individus génotypés pour au moins 60% des marqueurs ; Tableau E-2).

b) Carte génétique et densité en marqueurs

Treize marqueurs ont été sélectionnés dans une région incluant celle de la primo-localisation et ont été génotypés sur les 41 familles du dispositif de cartographie. En incluant les marqueurs SAM, l'analyse de liaison sur ce chromosome s'est finalement appuyée sur 16 marqueurs microsatellites augmentant ainsi la région couverte par les marqueurs de 30 cM à 114 cM. La densité moyenne en marqueurs est d'environ 7 cM. Toutefois, la construction de la carte génétique sur ce chromosome a montré l'existence d'un intervalle de 20 cM environ entre les deux premiers marqueurs : DIK1081 et DIK4334 d'une part et le troisième marqueur DIK2496 d'autre part (Figure E-6). Les deux premiers marqueurs ont donc été éliminés de l'analyse de liaison lors de l'étude de cartographie. La région couverte par les marqueurs est

désormais de 89 cM. Les distances génétiques entre les différents marqueurs, obtenues avec les données du dispositif expérimental (race Prim'Holstein), celles issues de la carte de l'USDA et celles issues de la carte physique de l'assemblage Btau_4.0, sont indiquées dans la figure E-6. Les distances deux-points entre les marqueurs sont relativement similaires entre la carte génétique expérimentale et celle de la carte USDA pour toutes les paires de marqueurs adjacents sauf pour la paire de marqueurs BM2113- IDVGA-2 où la différence entre les deux types de données est de 9 cM environ (Figure E-6).

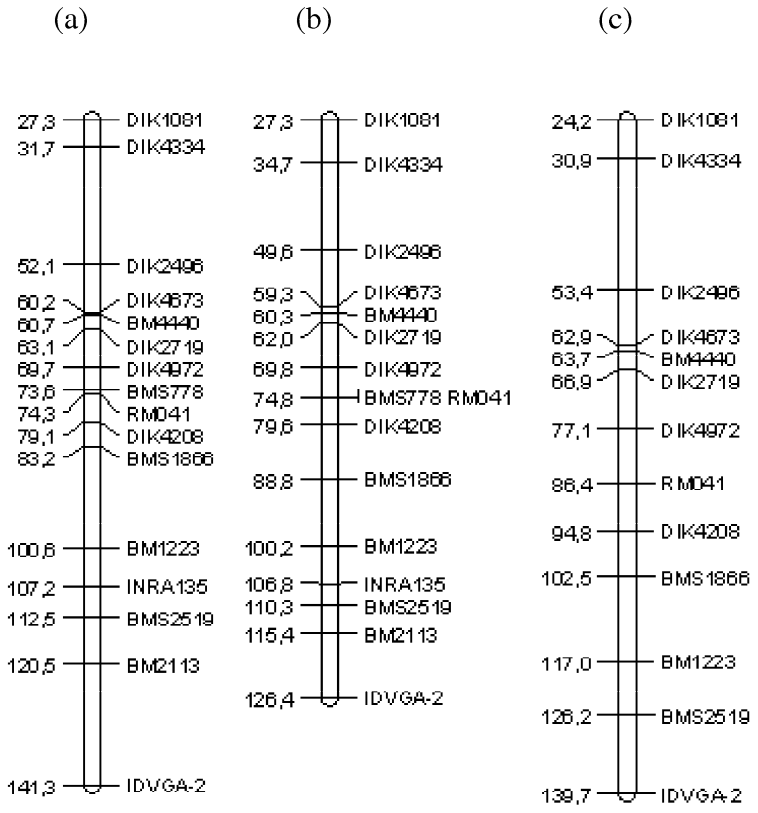


Figure E-6 : *Cartes génétiques des marqueurs microsatellites de BTA02 obtenues avec les données expérimentales (a), à partir de la base de données publique USDA (les distances sont exprimées en cM) (b) et carte physique obtenue par blast des microsatellites sur l'assemblage bovin (Btau_4.0) (les distances sont exprimées en Mégabases) (c). Certains marqueurs n'ont pas pu être ancrés sur cet assemblage. L'origine des cartes génétiques ((a) et (b)) est celle de la carte génétique de USDA.*

c) Résultats de l'analyse de liaison

(1) Informativité du dispositif

La courbe d'informativité globale du dispositif, calculée pour chacune des races et par pas de 1 cM le long de la région couverte par les deux marqueurs extrêmes, est représentée dans les figures E-7, E-8 et E-9. L'informativité moyenne des marqueurs est de 0,73 ± 0,11, 0,70 ± 0,16, 0,73 ± 0,08, respectivement pour les races Prim'Holstein, Normande et Montbéliarde.

(2) Résultats de détection de QTL

(a) Courbes de F-value

Les courbes de F-value correspondant aux TNR56, TNR90 et TNR281 pour chacune des trois races sont représentées dans les figures E-7, E-8 et E-9. Deux QTL différents ont été détectés chez la race Prim'Holstein : le premier, un QTL suggéré (P<0,047), a été détecté avec le TNR56 à la position de 70 cM, à côté des marqueurs BMS778 et RM041.

Les courbes de F-value obtenues avec les TNR90 et TNR281 ont la même allure à la même position mais n'atteignent pas le seuil de significativité de 5% (Figure E-7). Un deuxième QTL suggéré est détecté avec le TNR90 (P<0,048) et le TNR281 (P<0,034) à une position plus télomérique. La position de ce pic se situe à 115 cM (à côté de BM2113). L'estimation, par la méthode de bootstrapping (10 000 échantillonnages), de l'intervalle de confiance du QTL détecté avec le TNR281, identifie une localisation moyenne du QTL à 87 cM et un $IC_{95\%}$ compris entre 49 et 135 cM c'est-à-dire sur tout l'intervalle couvert par les marqueurs.

(b) Statut des pères au QTL

Six pères sont déclarés hétérozygotes (p<0,05) pour le QTL détecté avec le TNR56 : 3468, 3481, 3502, 5245, 6109 et 6865. Les pères 3480, 3514 et 6865 sont déclarés hétérozygotes (P<0,05), à la fois pour les TNR90 et TNR281 alors que le père 3487 n'a été déclaré hétérozygote que pour le TNR90 (tableau E-5)

Tableau E-5 : *Valeurs des effets de substitution allélique et du test-t des familles hétérozygotes (p<0,05) pour les TNR90 et TNR281 sur le BTA02 en race Prim'Holstein.*

Indice de fertilité	Famille	Effet de substitution allélique	Valeur du test-t
TNR56	3468	0,0166	2,7265
	3481	-0,0108	2,2134
	3502	-0,0105	1,966
	5245	-0,0162	2,0852
	6109	-0,0305	2,6167
	6865	0,0209	2,5433
TNR90	3487	0,0196	2,3549
	3480	-0,0267	2,5939
	3514	-0,0277	2,7497
	6865	0,0228	2,5949
TNR281	3468	0,0152	1,9769
	3480	-0,03	2,8793
	3514	-0,0267	2,5084
	6865	0,0257	2,7049

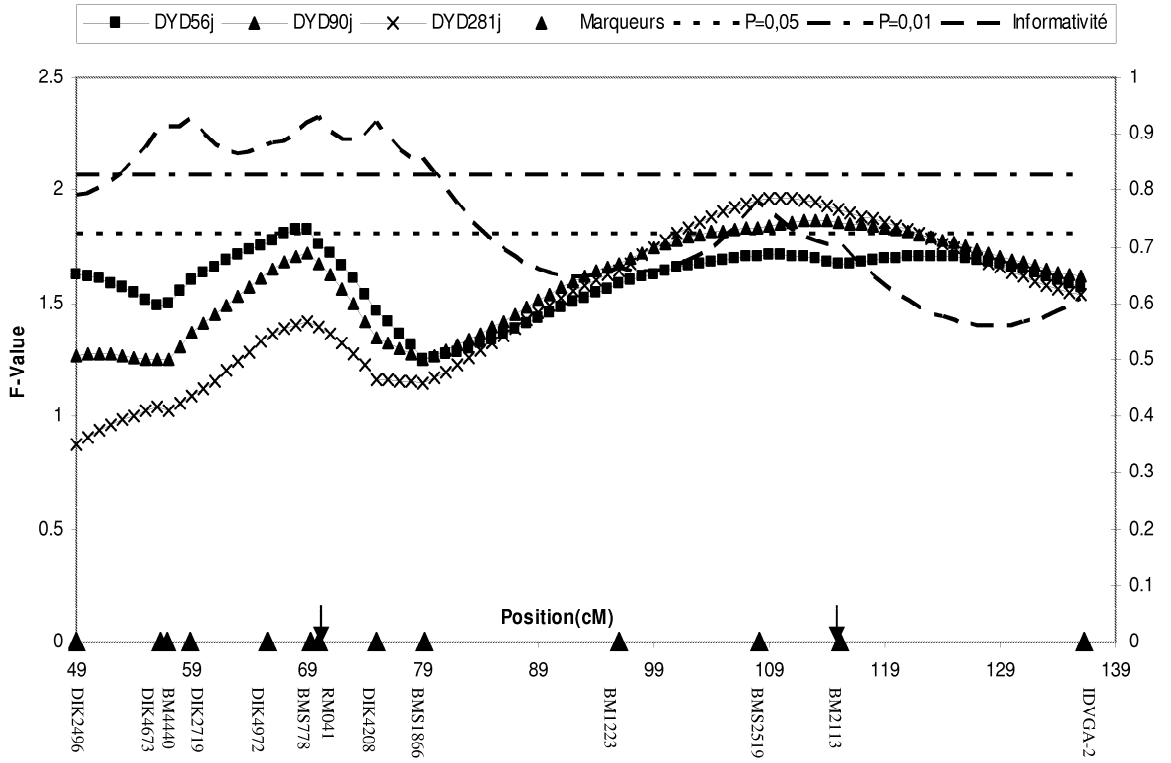


Figure E-7 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA02 en race Prim'Holstein. Les courbes de F-value correspondent aux TNR à 56 (courbe dessinée avec des carrés), 90 (courbe dessinée avec des triangles) et 281 jours après IA (courbe dessinée avec des croix). L'informativité des marqueurs est décrite par la courbe en pointillés (l'axe des ordonnés de l'informativité est celui situé à droite de la figure). Les seuils de significativité correspondants ont été obtenus par 10 000 permutations. Les marqueurs SAM utilisés lors de la primo-localisation sont indiqués par des flèches.*

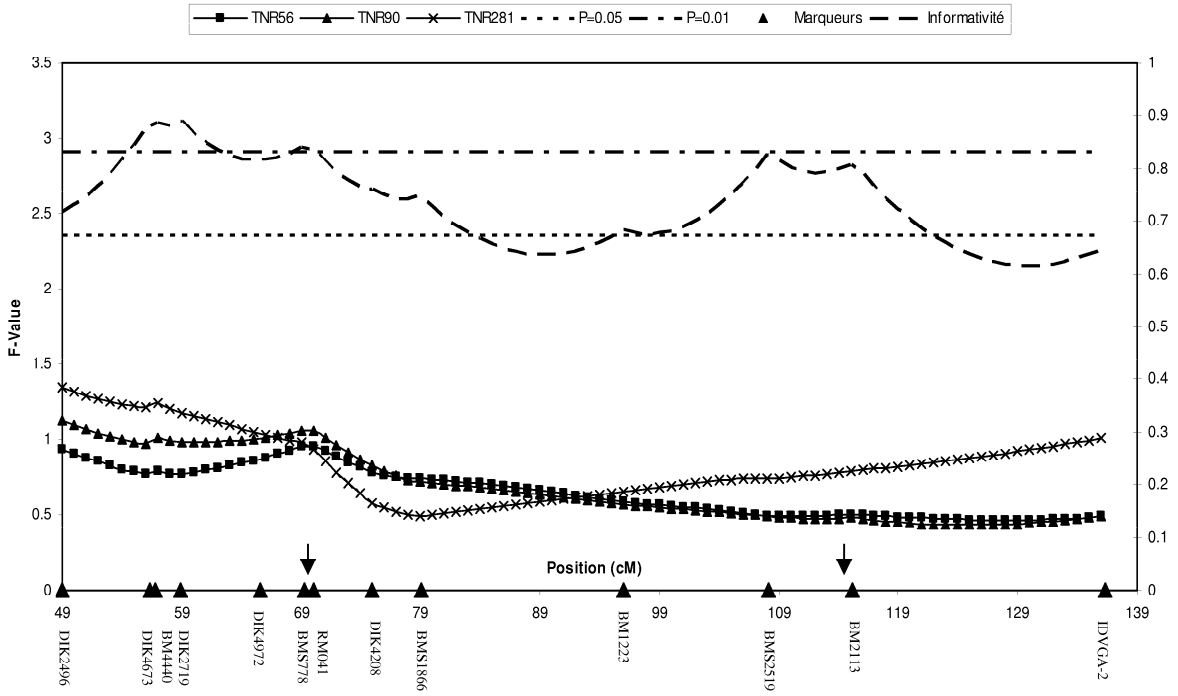


Figure E-8 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA02 en race Normande. La légende est identique à celle de la figure E-7. Les seuils de significativité correspondants ont été obtenus par 1000 permutations.*

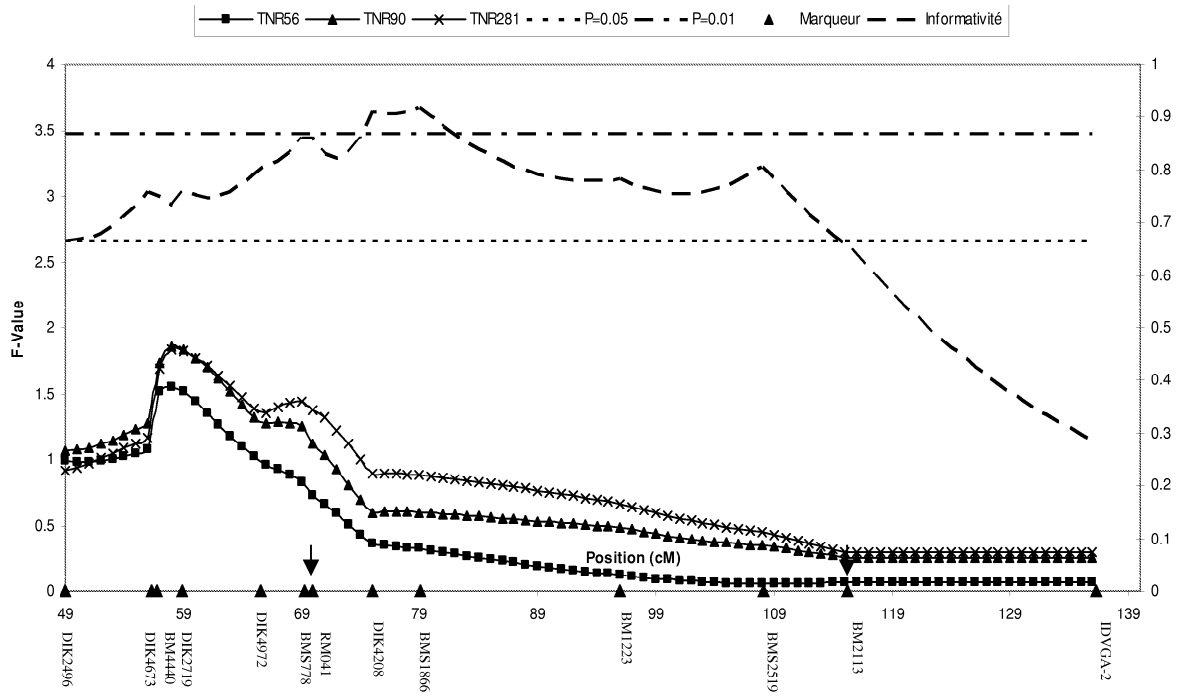


Figure E-9 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA02 en race Monbéliarde. La légende est identique à celle de la figure E-7. Les seuils de significativité correspondants ont été obtenus par 1000 permutations.*

3. Chromosome BTA21

(a) Données de génotypage analysées

Deux mille huit cent soixante-trois individus ont été génotypés pour les trois races sur le chromosome BTA21. Deux mille sept cent trente-trois individus sont inclus dans les analyses (individus génotypés pour au moins 60% des marqueurs ; Tableau E-2).

(b) Carte génétique et densité en marqueurs

Deux marqueurs supplémentaires ont été génotypés pour 41 familles du dispositif de l'étude de cartographie (Figure E-10). Au total, cinq marqueurs sont disponibles pour l'étude de cartographie (en comptant les microsatellites SAM). La taille de la région couverte par les marqueurs est passée de 34,5 cM à 65,5 cM.

La densité en microsatellites quant à elle, n'a pas augmenté étant donné que seulement deux marqueurs ont été ajoutés pour une distance ajoutée de 31 cM. La densité en marqueurs est passée de un marqueur tous les 11,5 cM à un marqueur tous les 13,5 cM. Les distances génétiques entre les différents marqueurs, obtenues avec les données du dispositif expérimental (race Prim'Holstein), celles issues de la carte de l'USDA et celles issues de la carte physique de l'assemblage Btau 4.0, sont présentées dans la figure E-10. Les distances

deux-points entre les marqueurs sont relativement similaires entre la carte génétique expérimentale et celle de la carte USDA pour toutes les paires de marqueurs adjacents.

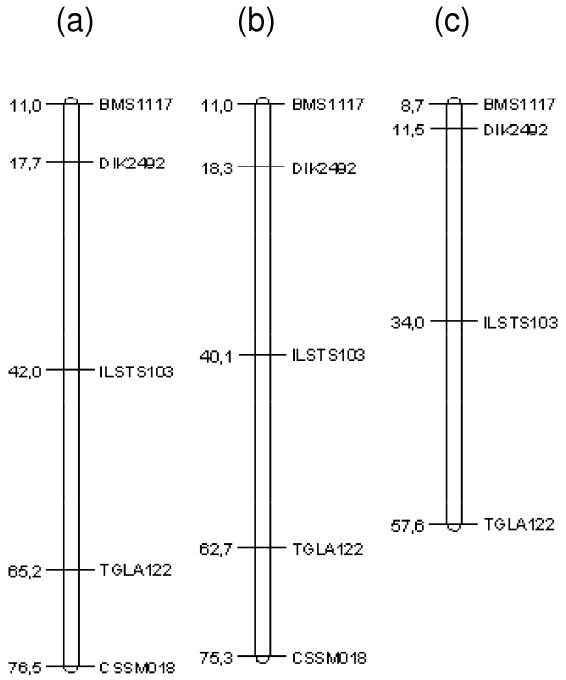


Figure E-10 : *Cartes génétiques des marqueurs microsatellites du BTA21 obtenues avec les données expérimentales (a), à partir de la base de données publique USDA (les distances sont exprimées en cM) (b) et carte physique obtenue par blast des microsatellites sur l'assemblage bovin (Btau_4.0) (les distances sont exprimées en Mégabases) (c). CSSM018 n'a pas pu être ancré sur cet assemblage.*

b) Résultats de l'analyse de liaison

(1) Hétérozygotie des pères pour les marqueurs et informativité du dispositif

La courbe d'informativité globale du dispositif calculée pour chacune des races et par pas de 1 cM le long de la région couverte par les deux marqueurs extrêmes, est représentée dans les figures E-11, E-12 et E-13. L'informativité moyenne des marqueurs est de 0,71 ± 0,17, 0,69 ± 0,14 et 0,77 ± 0,12 respectivement pour les races Prim'Holstein, Normande et Montbéliarde.

Résultats de détection de QTL

(a) Courbes de F-value

Les courbes de F-value correspondant aux TNR56, TNR90 et TNR281 pour chacune des trois races sont représentées dans les figures E-11, E-12 et E-13. Aucun QTL de fertilité femelle n'a été détecté pour les différents TNR en races Prim'Holstein et Montbéliarde. En race Normande, un QTL de fertilité femelle est détecté pour le TNR56 ($P<0,048$). Le pic de la

courbe de F-value obtenu avec le TNR56 est situé au début du BTA21, à 15 cM à côté du marqueur DIK2492 (Figure E-12).

L'estimation, par la méthode de bootstrap (10 000 échantillonnages), de l'intervalle de confiance du QTL détecté a donné une localisation moyenne du QTL à 32 cM environ et un intervalle de confiance à 95% ($IC_{95\%}$) compris entre 0 et 64 cM.

(b) Statut des pères au QTL

Un seul père de famille : le 1537 est déclaré hétérozygote (P<0,05) pour le QTL détecté sur le chromosome BTA21 avec le TNR56 (Tableau E-6).

Tableau E-6 : *Valeurs des effets de substitution allélique et du test-t de la seule famille hétérozygote (p<0,05) pour le TNR56 sur le BTA21 en race Normande.*

Indice de fertilité	Famille	Effet de substitution allélique	Valeur du test-t
TNR56	1537	0,0263	3,0306

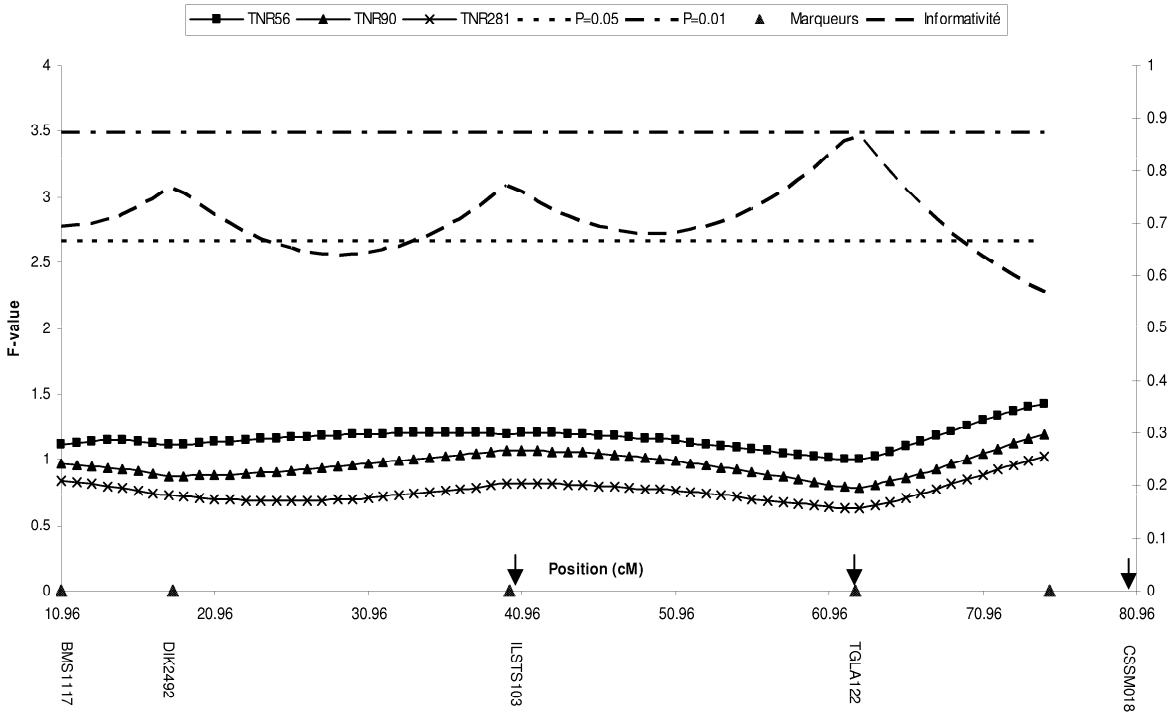


Figure E-11 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA21en race Prim'Holstein. Les courbes de F-value correspondent aux TNR à 56 (courbe dessinée avec des carrés), 90 (courbe dessinée avec des triangles) et 281 jours après IA (courbe dessinée avec des croix). L'informativité des marqueurs est décrite par la courbe en pointillés (l'axe des ordonnés de l'informativité est celui situé à droite de la figure). Les seuils de significativité correspondants ont été obtenus par 5000 permutations. Les marqueurs SAM utilisés lors de la primo-localisation sont indiqués par des flèches.*

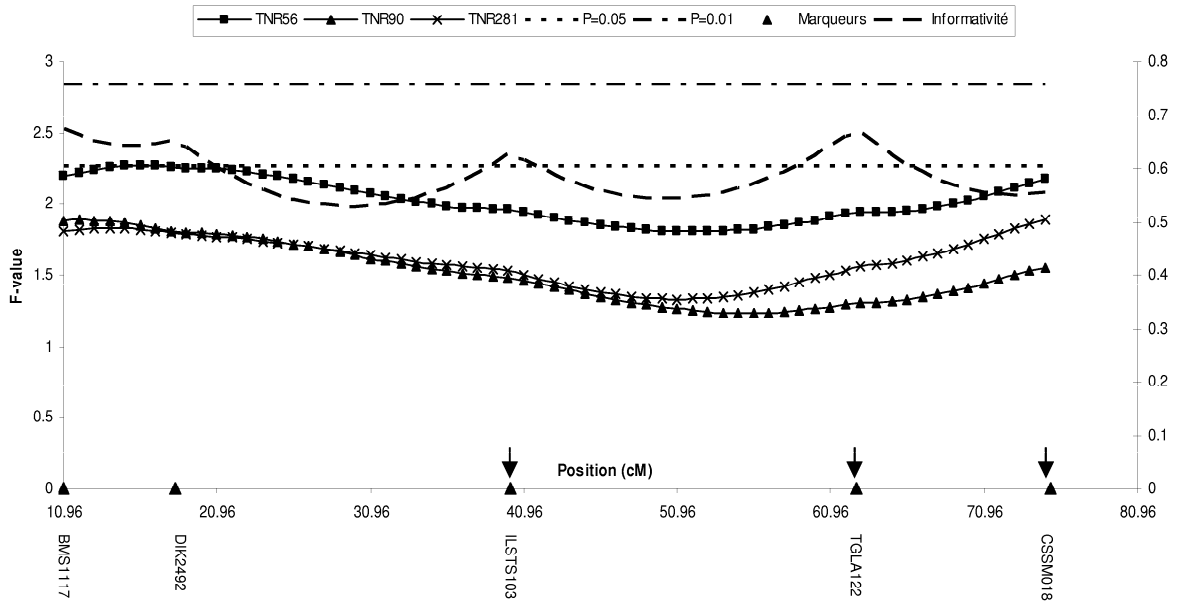


Figure E-12 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA21 en race Normande. La légende est identique à celle de la figure E-11. Les seuils de significativité correspondants ont été obtenus par 5000 permutations.*

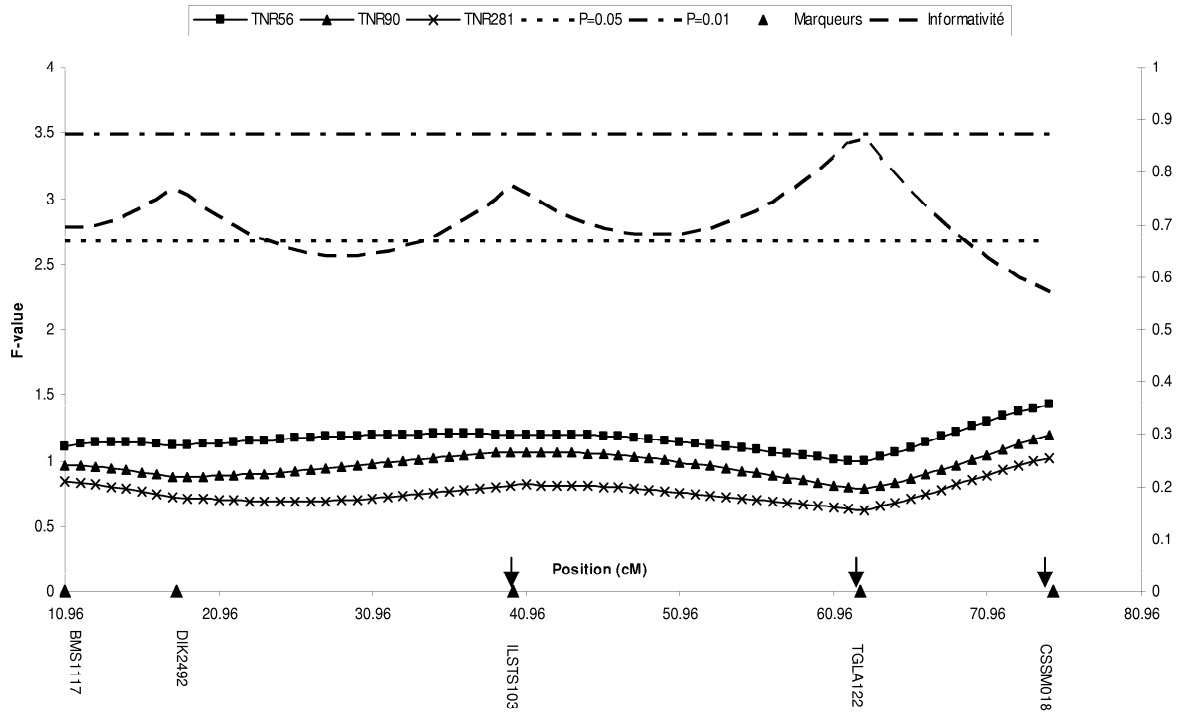


Figure E-13 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA21 en race Montbéliarde. La légende est identique à celle de la figure E-11. Les seuils de significativité correspondants ont été obtenus par 5000 permutations.*

4. Chromosome BTA03

a) Données de génotypage analysées

Trois mille quatre-vingt dix-neuf individus ont été génotypés pour les trois races sur chromosome BTA03. Toutefois, certains marqueurs de ce chromosome n'ont été génotypés que pour la race Prim'Holstein (INRA3015, INRA3012, INRA3013, DIK4193, BMS2904 et DIK069), alors qu'un seul marqueur a été génotypé chez les races Normandes et Montbéliardes mais pas en Prim'Holstein ; il s'agit de CC582410. Deux mille six cent quatre-vingt quatre individus sont inclus dans les analyses (individus génotypés pour au moins 60% des marqueurs ; Tableau E-2).

b) Carte génétique et densité en marqueurs

Vingt-deux, 18 et 18 marqueurs ont été sélectionnés, à l'intérieur de la région de primo-localisation et génotypés respectivement, en race Prim'Holstein, Normande et Monbéliarde. En incluant les marqueurs SAM, l'analyse de liaison sur ce chromosome s'est finalement appuyée sur 26, 22 et 22 marqueurs microsatellites augmentant ainsi la densité en marqueurs de un marqueur tous les 22,5 cM à un marqueur tous les 3,5 cM chez la Prim'Holstein et un marqueur tous les 4 cM en races Normande et Montbéliarde.

La position physique de tous les microsatellites du BTA03, sur la dernière version de la séquence bovine (Bta 4.0) a été déterminée par blast sauf pour RM019 et INRA023.

Les distances génétiques entre marqueurs, obtenues avec les données du dispositif expérimental (race Prim'Holstein), celles issues de la carte de l'USDA et celles issues de la carte physique de l'assemblage Btau_4.0, sont indiquées dans la figure E-14. L'ordre des marqueurs sur la carte génétique est en accord avec celui de la carte physique pour tous les marqueurs sauf pour INRA006 et INRA3015.

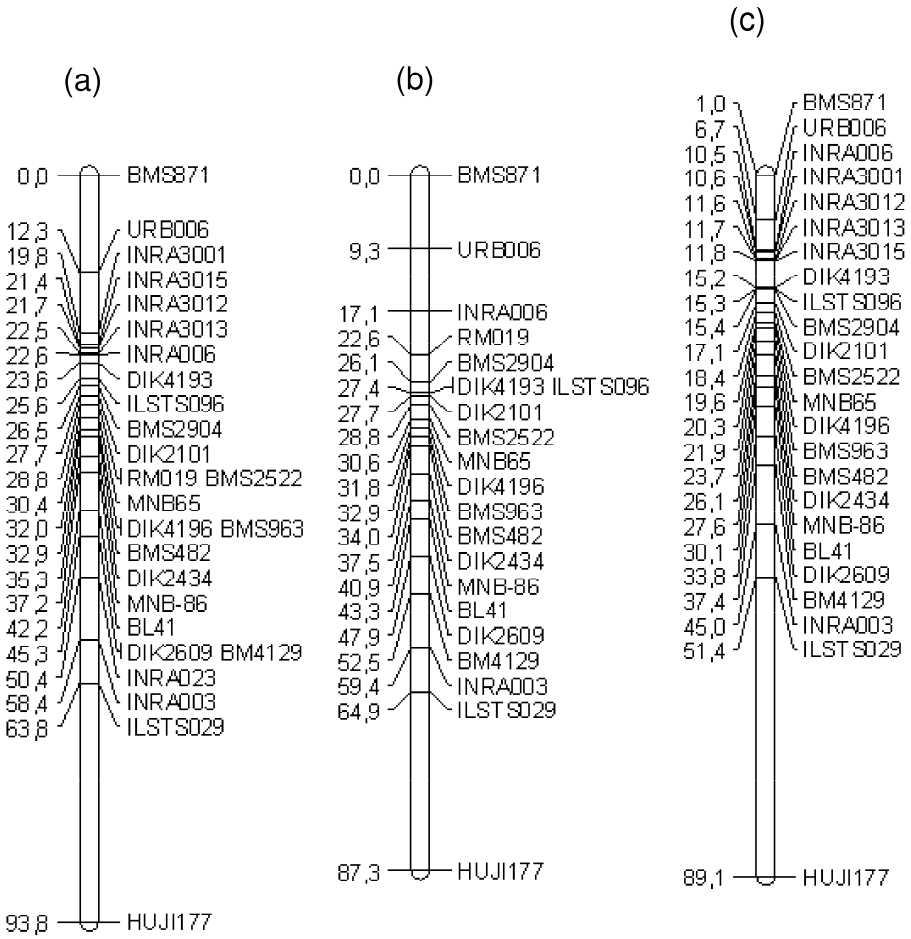


Figure E-14 : *Cartes génétiques des marqueurs microsatellites du BTA02 obtenues avec les données expérimentales (a), à partir de la base de données publique USDA (les distances sont exprimées en cM) (b) et carte physique obtenue par blast des microsatellites sur l'assemblage bovin (Btau_4.0) (les distances sont exprimées en Mégabases) (c). Certains marqueurs n'ont pas pu être ancrés sur cet assemblage. L'ordre des marqueurs de la carte génétique est en accord avec celui du dernier assemblage sauf au niveau de INRA006 et de INRA3015.*

c) Résultats dc l'analysc dc liaison

(1) Informativité du dispositif

La courbe d'informativité globale du dispositif, calculée pour chacune des races et par pas de 1 cM le long de la région couverte par les deux marqueurs extrêmes, est représentée dans les figures E-15, E-16 et E-17. L'informativité moyenne des marqueurs est de 0,71 ± 0,17, 0,69 ± 0,14 et 0,77 ± 0,12 respectivement pour les races Prim'Holstein, Normande et Montbéliarde.

(a) Résultats de détection de QTL

(i) Courbes de F-value

Les courbes de F-value correspondant aux TNR56, TNR90 et TNR281 pour chacune des trois races sont représentées dans les figures E-15, E-16 et E-17. Un QTL de fertilité femelle a été détecté en race Prim'Holstein avec le TNR90 (P<0,015) et le TNR281 (P<0,05). Les pics des

courbes de F-value obtenus avec les deux indices ont la même position : 25 cM (à côté de ILSTS096) (Figure E-15). L'estimation, par la méthode de bootstrap (10 000 échantillonnages), de l'intervalle de confiance du QTL détecté avec le TNR90 a donné une localisation moyenne du QTL à 25 cM et un $IC_{95\%}$ compris entre 0 et 58 cM.

(ii) Statut des pères au QTL

Les pères de famille 4442, 4729, 6145 et 6151 sont déclarés hétérozygotes (P<0,05) pour le QTL détecté avec le TNR90 et TNR281. Le père 3514 est déclaré hétérozygote avec le TNR90, mais pas avec le TNR281 (Tableau E-7).

Tableau E-7 : *Valeurs des effets de substitution allélique et du test-t des familles hétérozygotes (p<0,05) pour les TNR90 sur BTA03 en race Prim'Holstein lors de l'étude de cartographie.*

Indice de fertilité	Famille	Effet de substitution allélique	Valeur du test-t
TNR90	3514	0,0187	1,96
	4442	0,0217	2,89
	4729	0,0178	2,05
	6145	-0,0213	2,45
	6151	-0,0279	2,84
TNR281	4442	0,0195	2,4544
	4729	0,0224	2,3997
	6145	-0,0197	2,1512
	6151	-0,0278	2,6872

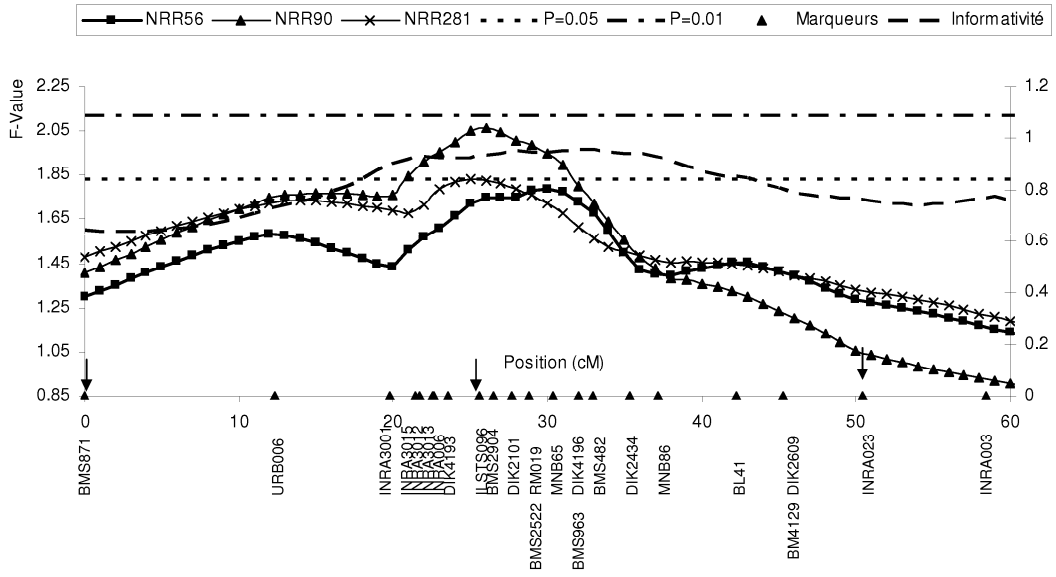


Figure E-15 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA03 en race Prim'Holstein. Les courbes de F-value correspondent aux TNR à 56 (courbe dessinée avec des carrés), 90 (courbe dessinée avec des triangles) et 281 jours après IA (courbe dessinée avec des croix). L'informativité des marqueurs est décrite par la courbe en pointillés (l'axe des ordonnés de l'informativité est celui situé à droite de la figure). Les seuils de significativité correspondants ont été obtenus par 10000 permutations. Les marqueurs SAM utilisés lors de la primo-localisation sont indiqués par des flèches.*

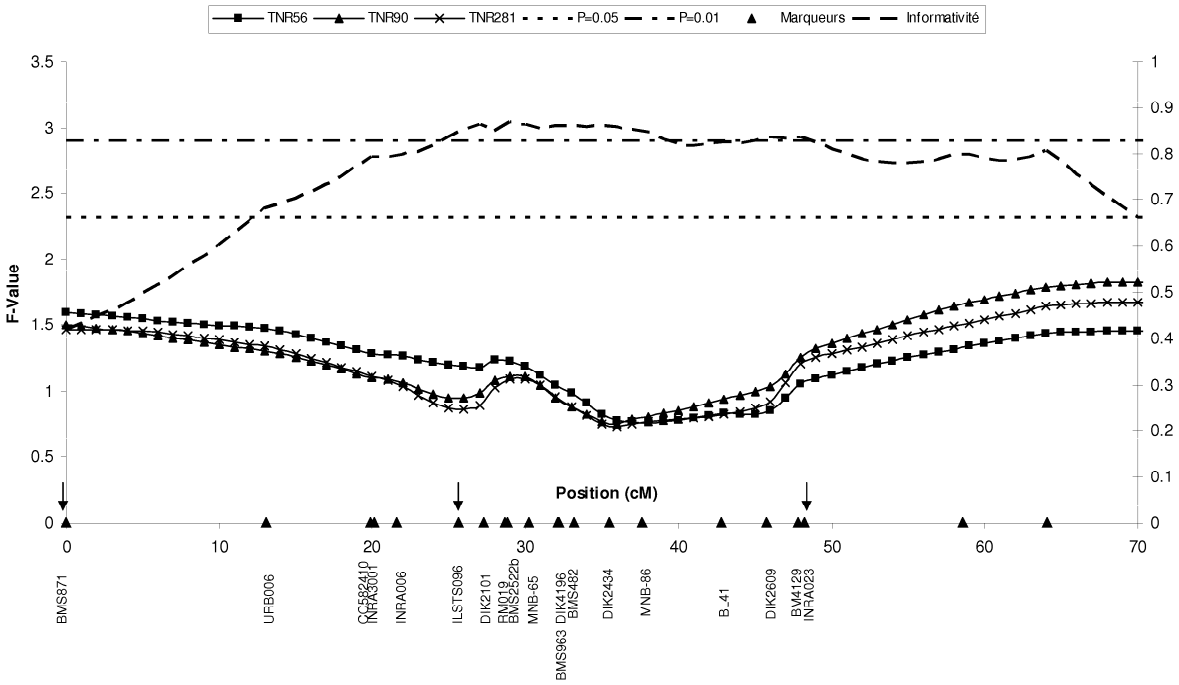


Figure E-16 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA03 en race Normande. La légende est identique à celle de la figure E-15. Les seuils de significativité correspondants ont été obtenus par 1000 permutations.*

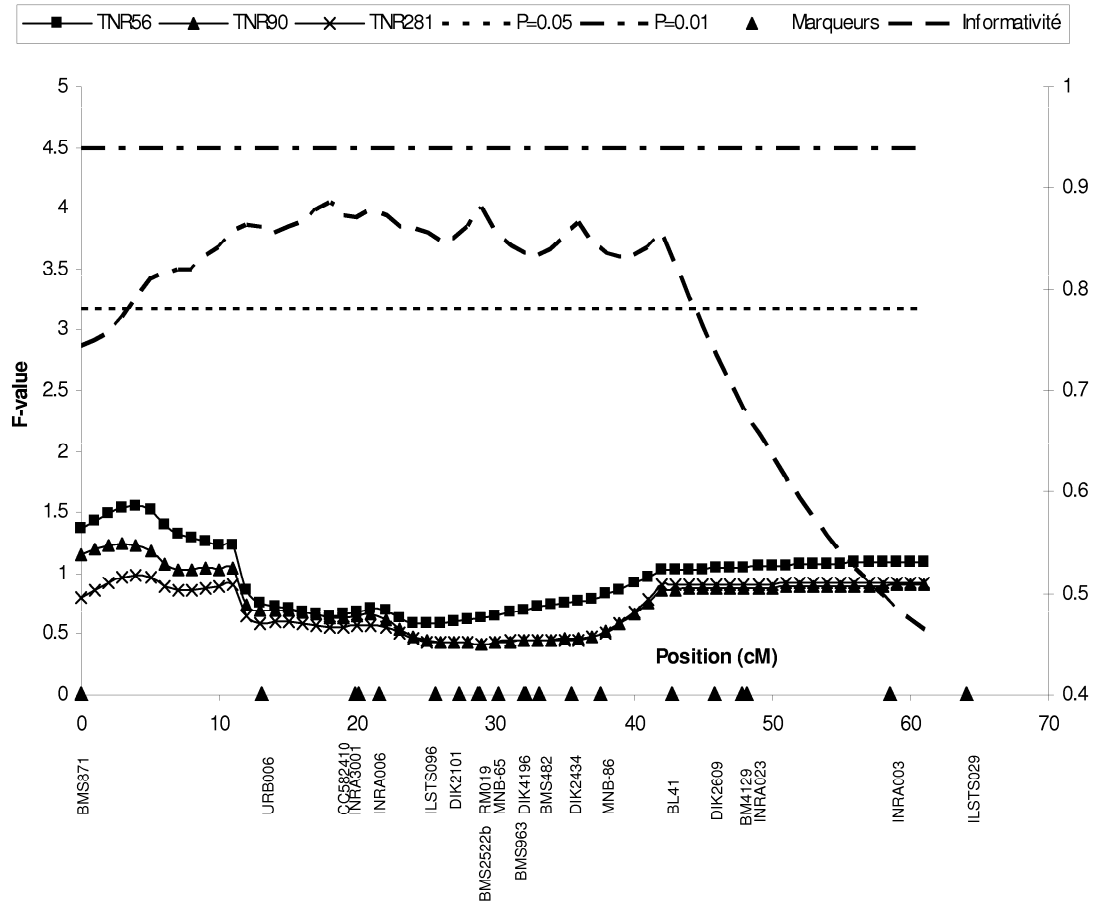


Figure E-17 : *Résultats de cartographie, par analyse de liaison, du QTL de fertilité femelle sur BTA03 chez la Montbéliarde. La légende est identique à celle de la figure E-15. Les seuils de significativité correspondants ont été obtenus par 1000 permutations.*

5. Conclusion

L'étape de cartographie à moyenne résolution a permis, dans un premier temps, de réduire les intervalles de localisation des QTL sur les chromosomes BTA01, BTA02, BTA03. Ce travail a fait l'objet d'une publication (Ben Jemaa *et al.*, 2008). Les intervalles de localisation des QTL demeurent toutefois larges et permettent de mettre en place une SAM de première génération, peu efficace pour assurer un progrès génétique rentable des caractères d'intérêts chez les bovins, en particulier ceux de faible héritabilité tels que la fertilité femelle. L'étape suivante a consisté à réaliser une cartographie fine des QTL détectés afin de pouvoir affiner la région de localisation des QTL permettant de réaliser des études de gènes candidats.

III. Cartographie fine de QTL de fertilité femelle

L'évolution considérable des technologies de génotypage à haut débit utilisant les SNP et la baisse des coûts, a permis de mettre à disposition de nouveaux outils pour la cartographie fine des QTL détectés par le génotypage à moyenne résolution. Le chromosome BTA03 a été choisi en premier lieu pour faire de la cartographie fine : les travaux en cours sur ce chromosome étaient en effet plus avancés que pour les autres chromosomes. Ensuite, grâce à la mise à disposition de la puce « Bovine SNP50 » d'Illumina, tous les autres chromosomes ont été couverts à l'aide de cette puce. Le travail de cartographie fine s'est donc effectué en utilisant une approche LDLA : toutes les analyses ont été prises en charge par la SGQA.

1. Cartographie fine de QTL de fertilité femelle sur le BTA03

La cartographie fine du QTL de fertilité femelle sur le BTA03 a été faite en utilisant 437 SNP qui ont été génotypés sur un dispositif de 17 familles de race Prim'Holstein appartenant aux 26 familles Prim'Holstein du dispositif de cartographie.

Ce travail a fait l'objet d'une publication (Druet *et al.*, 2008). Dans cet article, la courbe obtenue par analyse LDLA montre huit pics qui s'étendent sur la région comprise entre 11,5 et 35 cM du BTA03 (Tableau E-8, figure E-18). Le regroupement d'haplotypes des pères hétérozygotes a toutefois permis de considérer que ce sont les pics 1, 2, 4 et 5 de la courbe LDLA (situés respectivement à 11,68, 16,87, 23,12 et 25,55 cM) qui sont les plus intéressants (Druet *et al.*, 2008). L'analyse LDLA mentionnée dans la partie IV-1 de « matériel et méthodes » (S. Fritz, communication personnelle) situe le QTL de fertilité femelle dans la région comprise entre 16,1 cM et 19,9 cM c'est-à-dire dans la région comprise entre le deuxième et le troisième pic de la publication mentionnée ci-dessus.

Tableau E-8 : *Position des huit pics obtenus par analyse LDLA, en utilisant 437 SNP du BTA03 et le TNR90 comme phénotype de fertilité femelle (d'après Druet et al., 2008).*

Numéro du pic	Position (cM)
1	11,68
2	16,87
3	19,68
4	23,12
5	25,55
6	26,93
7	28,03
8	34,86

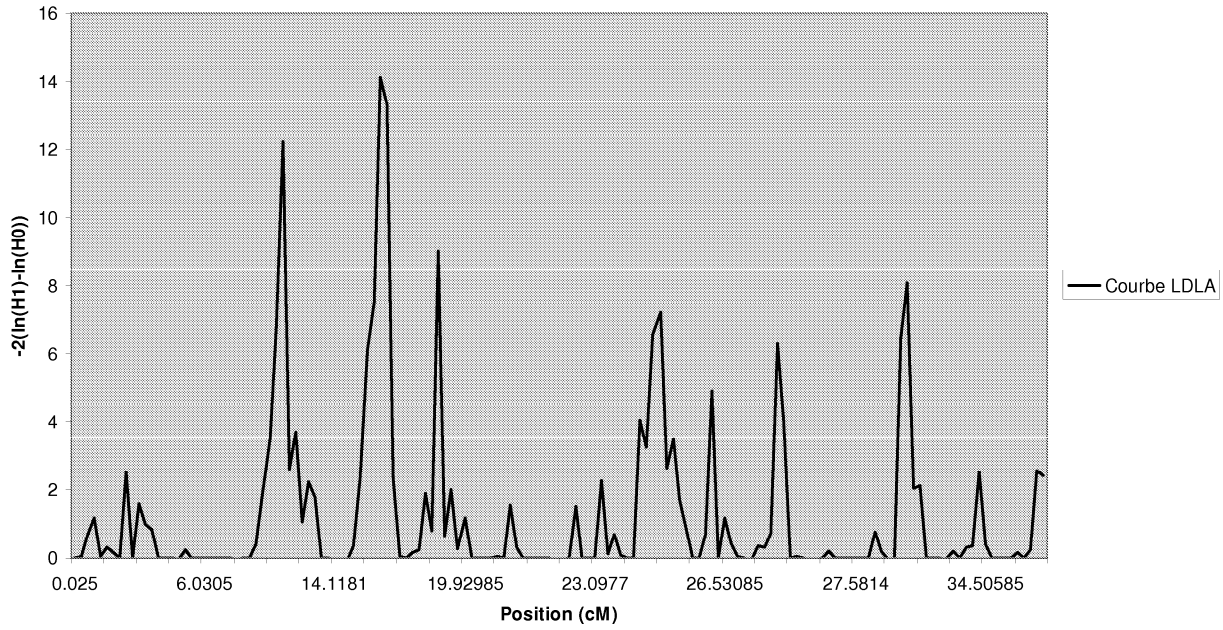


Figure E-18 : *Résultats de l'analyse LDLA obtenus sur un dispositif de 17 familles Prim'Holstein, avec le panel de 437 SNP et en utilisant le TNR90 comme phénotypes de fertilité femelle. La courbe LDLA montre huit pics situés dans l'intervalle compris entre 11,68 cM et 34,86 cM (Druet et al., 2008).*

2. Cartographie fine de QTL de fertilité femelle à travers tout le génome en utilisant la puce Illumina « BovineSNP50 »

Au cours de cette thèse et plus précisément pendant la phase de séquençage des six gènes candidats cités précédemment, une puce commerciale comprenant 54000 SNP a été rendue publique. Il s'est donc avéré plus judicieux d'utiliser cette puce pour faire des détections de QTL influençant plusieurs caractères dont la fertilité femelle des bovins. La puce BovineSNP50 d'Illumina (appelée par la suite puce Illumina) présente l'avantage de couvrir le génome dans son ensemble. Cette puce a été utilisée sur un dispositif composé de 1800 taureaux d'IA Prim'Holstein, de 800 taureaux d'IA Normands et de 750 taureaux d'IA de race Montbéliarde (Projet CartoFine financé par l'ANR et APIS-GENE). Seuls les résultats de

détection chez la race Prim'Holstein seront présentés. Plusieurs pics obtenus avec l'analyse LDLA et dépassant le seuil de 1% ont été trouvés sur différentes régions génomiques avec les deux types d'indices. Un résumé des résultats obtenus avec les deux types d'indices est présenté dans les tableaux E-9 et E-10.

Des pics significatifs ont été détectés sur les chromosomes 1, 2, 3, 4, 6, 8, 9, 11, 14, 15, 18, 19 et 28 avec l'index de fertilité « vache »et sur les chromosomes 2, 5, 7, 9, 10, 12, 14 et 22 avec l'index de fertilité « génisse ».

Tableau E-9 : *Résumé des résultats de l'étude LDLA obtenus avec la puce Illumina et les indices de fertilité femelle « vache » réalisée par la SGQA. Ce tableau décrit la position des SNP où des pics significatifs (P<0,01) ont été détectés. La première colonne décrit le numéro du chromosome, la deuxième décrit l'identifiant du SNP correspondant au pic détecté, les troisième et quatrième colonnes décrivent respectivement, la position (en Mb) et la valeur du LRT du pic.*

BTA	Identifiant_SNP	Position du pic de la courbe LDLA (Mb)	LRT	Variance expliquée par le QTL
1	156	7,2375	13,4727	0,0489
2	120	6,1163	10,9175	0,0276
2	149	8,3859	13,6488	0,0306
2	160	8,9484	11,0619	0,0304
2	224	13,4194	24,5736	0,0441
2	251	14,9454	22,5508	0,0426
2	268	15,8212	22,3349	0,0423
2	293	17,5517	15,8554	0,0221
2	312	18,665	16,0475	0,0422
3	1746	114,5163	11,1143	0,0327
3	1782	116,3449	9,3812	0,0325
3	1798	117,399	10,3619	0,0209
4	1064	67,1258	15,6844	0,0506
4	1099	69,1303	10,367	0,0217
6	1612	101,2925	9,1619	0,0359
6	1636	102,6684	11,0335	0,0395
6	1668	104,2663	10,3138	0,0359
6	1692	105,5721	9,5139	0,0246
8	88	5,2775	9,5597	0,0347
9	774	53,8881	9,6335	0,0286
9	850	58,2004	9,2131	0,0382
9	1004	67,2926	9,5778	0,0276
11	1534	92,7193	9,8122	0,0297
14	626	32,005	10,3859	0,0369
14	666	35,3497	16,2236	0,0486
14	686	36,5238	10,8659	0,0326
14	700	37,6375	10,9267	0,032
15	376	24,9308	9,8072	0,0255
15	552	33,6629	13,2345	0,0482
15	578	35,2243	17,1255	0,0416
15	636	38,7236	10,3714	0,0314
18	392	21,9917	12,0671	0,0246
18	430	23,8241	11,3307	0,035
18	712	41,6038	10,1668	0,0311
18	772	45,5755	12,4079	0,0225
18	838	49,2959	10,3146	0,0216
18	940	55,0699	14,4156	0,0475
18	956	56,1081	12,6902	0,0341
18	976	57,3928	15,7446	0,0287
19	722	45,0939	10,1885	0,027
19	764	47,2974	10,2703	0,0349
28	606	36,3384	10,4283	0,0354

Tableau E-10 : *Résumé des résultats de l'analyse LDLA obtenus avec la puce Illumina et l'index de fertilité « génisse » réalisée par la SGQA. Le descriptif des colonnes est le même que celui du tableau E-9.*

BTA	Identifiant_SNP	Position (Mb)	LRT	Variance expliquée par lc QTL
2	120	6,1163	11,8233	0,0374
5	488	33,5356	10,9677	0,1431
5	606	43,6431	11,8143	0,2727
7	197	10,8374	10,1581	0,2378
9	68	3,4463	11,4278	0,0725
10	654	41,0514	10,2323	0,0706
10	700	44,2194	11,862	0,1296
10	772	47,7023	17,3644	0,1451
10	786	48,2647	12,7732	0,1459
10	848	51,6276	11,0737	0,0978
10	864	52,3224	9,346	0,0759
12	374	25,0485	9,8965	0,0621
12	430	29,2955	9,2243	0,0372
12	514	35,6608	9,7643	0,0628
12	814	55,3957	12,2426	0,0385
12	836	56,5325	10,5442	0,0358
12	874	58,7053	12,1125	0,0553
14	176	6,601	11,1547	0,0899
22	946	58,1517	9,0274	0,0527

IV. Etude des gènes candidats positionnels et fonctionnels

1. Etude de gènes candidats sur le BTA03

a) Ancrage de l'intervalle de localisation du QTL sur les génomes, humain et murin

L'ancrage de la région [10,93 Mb-15,70 Mb], l'intervalle de localisation du QTL du BTA03 obtenu par analyse LDLA, sur les génomes humain et murin a montré que cette région du BTA03 est conservée avec ces deux espèces. Ce bloc de synténie conservé correspond globalement à une région de 3,15 Mb environ, sur le chromosome 1 humain (HSA01) : celui compris entre 157,97 et 154,82 Mb du HSA01. Le bloc de synténie conservé avec la région d'intérêt du BTA03 chez l'homme, contient 78 séquences de référence (RefSeq).

Les données de cartographie comparée entre l'intervalle [10,93 Mb-15,70 Mb] du BTA03 et le génome de la souris ont montré que ce bloc du BTA03 est principalement conservé avec deux portions de chromosomes murins : MMU01 et MMU03. L'intervalle [10,92 Mb-12,61 Mb] du BTA03 est conservé avec l'intervalle [167,06 Mb-175,20Mb] du MMU01 tandis que l'intervalle [12,66 Mb-15,70 Mb] du BTA03 est conservé avec l'intervalle [86,80 Mb-87,86 Mb] du MMU03. Toutefois, il existe des micros remaniements qui peuvent aussi être dus à des problèmes d'assemblage de séquences, entre quelques portions de cette région du BTA03

et d'autres chromosomes murins. Le bloc de synténie conservé avec la région d'intérêt du BTA03 chez la souris, contient 151 séquences de référence murine (RefSeq MMU).

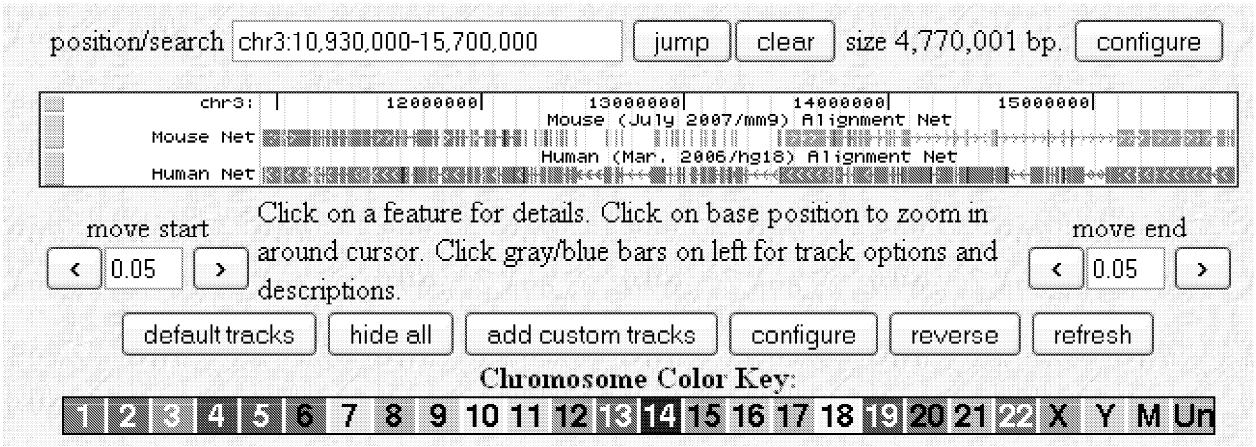


Figure E-19 : *Résultats de l'ancrage de la région [10,93 Mb-15,70 Mb] du BTA03 sur les génomes humains et murin à partir de la base de données UCSC. Cette figure montre que la région étudiée est conservée principalement avec des régions du HSA01 (en couleur marron) et avec des régions du MMU01 et du MMU03 (en couleur marron pour le MMU01 et vert olive pour le MMU03). Les bornes exactes des régions, humaine et murine, orthologues à cette région du BTA03 ont été obtenues à partir de cette même base de données.*

b) Séquençage des gènes candidats sélectionnés

Les résultats détaillés de séquençage des six gènes candidats sont représentés dans le tableau E-11. Un récapitulatif de ces résultats est présenté dans le tableau E-12. Deux, 13, 20, 1, 8 et 0 SNP ont été détectés respectivement pour les gènes PAQR6, MTX1, SPNA1, SHC1, ADAR1 et TPM3. Tous les polymorphismes détectés dans les régions séquencées sont situés dans des régions introniques, à l'exception de ceux détectés dans les exons 1 et 4 de ADAR1 et celui détecté dans la région régulatrice du gène PAQR6 (Tableau E-11). Parmi les mutations exoniques du gène ADAR1, une seule mutation est non synonyme et elle est à l'origine d'un changement de l'acide glutamique (E) (acide aminé acide et chargé négativement) en lysine (K)(acide aminé basique et neutre).

Au total, 39 polymorphismes ont été trouvés dans des régions introniques. Parmi les SNP trouvés, 10 SNP sont concordants avec le statut, à priori, des individus au QTL : 1 SNP dans la région régulatrice du PAQR6, 1 SNP dans la MTX1, 7 SNP dans SPNA1 et 1 SNP dans SHC1 (Tableau E-13).

Tableau E-11 : *Résultats détaillés de séquençage des six gènes candidats*

Gène-Exon	SNP	Type	Pools hétérozygotes	Position	Région	SNP à profil compatible Avec le statut au QTL	Nature de la mutation (pour les SNP exoniques)
MTX1_EX1	2	G>C G>T	P1, P2, P3, P4 P1, P2, P3, P4	790 802	Intronique Intronique	0	
MTX1_EX2-3	3	G>A G>A C>T	P1 P2 P1	1188 1326 1662	Intronique Intronique Intronique	0	
MTX1_EX4-5-6	7	C>T G>A G>C G>A C>T G>A C>T	P1, P2, P3, P4 P2, P3 P1, P2 P2 P2 P2 P1, P3, P4	2948 2984 3142 3498 3505 3524 3546	Intronique Intronique Intronique Intronique Intronique Intronique Intronique	1	
MTX1_EX7-8	1	C>T	P1, P2, P3, P4	3955	Intronique	0	
PAQR6_Région régulatrice	1	G>A	P1, P2	268		1	
PAQR6_EX1-2	0	-			-	-	
PAQR6_EX3-4	1	C>T			Intronique	0	
PAQR6_EX5-6	0	-			-	-	
TPM3_EX4-5	0	-			-	-	
TPM3_EX6	0	-			-	-	
TPM3_EX7-8	0						
TPM3_EX9	0	-			-	-	
SPNA1_EX16	5	G>A C>T A>C G>T C>T	P3, P4 P3 P3, P4 P3 P3	20544 20600 20644 20682 20752	Intronique Intronique Intronique Intronique Intronique	2	
SPNA1_EX28-29	3	C>T A>C G>A	P1, P2 P1, P2, P3 P1, P2	69980 70019 70272	Intronique Intronique Intronique	2	
SPNA1_EX28-29	2	C>T C>T	P1, P2, P3 P1, P2, P3	75172 75392	Intronique Intronique	0	
SPNA1_EX33-34	2	G>C A>T	P1, P2, P3 P2, P4	75454 75530	Intronique Intronique	0	
SPNA1_EX36	2	G>A C>T	P1, P2 P1, P2, P3, P4	81068 81071	Intronique Intronique	1	
SPNA1_EX_37-38	6	C>T G>A C>T C>T A>C G>A	P1, P2 P1, P2, P3, P4 P1, P2 P1, P2, P3, P4 P1, P2, P3, P4 P1, P2, P3, P4	83402 83465 83533 83579 83606 83667	Intronique Intronique Intronique Intronique Intronique Intronique	2	
SHC1_EX4-5-6	1	G>A	P1, P2	2167	Intronique	1	
SHC1_EX7	0						
ADAR1_EX1a	1	C>T	P1, P2, P3, P4	634	Exonique	0	Synonyme
ADAR_EX1b	2	G>A G>A	P1, P4 P2	1468 1552	Exonique Intronique	0	Non synonyme (E en K)
ADAR1_EX3	0	-	-	-	-	-	
ADAR1_EX4-5	4	G>C G>A C>T A>C	P1, P2, P3, P4 P1, P2, P3, P4 P1, P2, P3, P4 P1, P2, P3, P4	4734 4905 5209 16977	Intronique Exonique Exonique Intronique	0	Synonyme Synonyme
ADAR1_EX9	1	G>A	P2, P4	17034	Intronique	0	

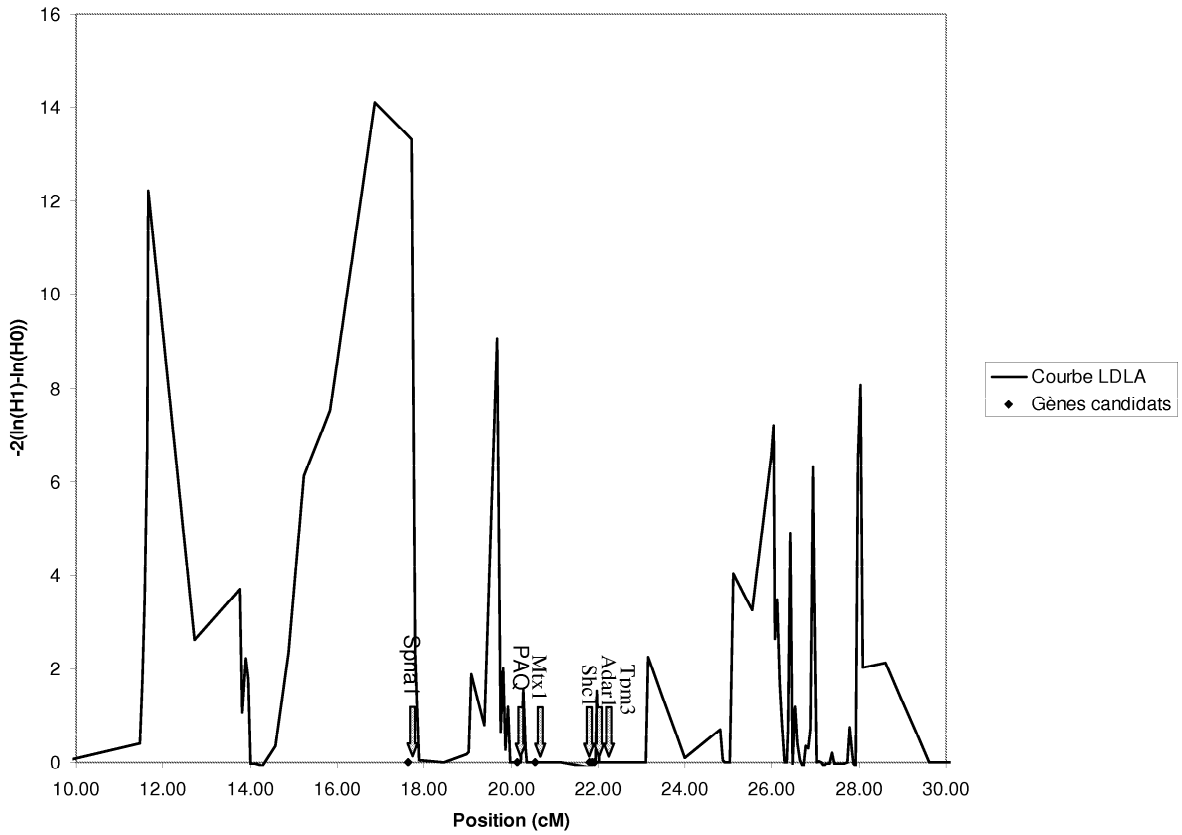


Figure E-20 : *Position des six gènes candidats sélectionnés par rapport à la courbe obtenue par analyse LDLA avec les 437 SNP sur le BTA03 (Druet et al., 2008). Les six gènes candidats sont situés entre le deuxième et le quatrième pic.*

Tableau E-12 : *Récapitulatif des résultats de séquençage des six gènes candidats.*

Gène	Nombre d'exons du gène (taille en Kb)	Taille séquencée (Kb)	Nombre de SNP trouvés	Nombre de SNP à profil compatible avec le statut au QTL des individus	SNP exoniques
PAQR6	7 (18,8)	2,8	2	1	0
MTX1	8 (4,8)	2,8	13	1	0
SPNA1	50 (75,6)	5,6	20	7	0
SHC1	12 (8,3)	2,8	1	1	0
ADAR1	15 (22,6)	5,6	8	0	4
TPM3	9 (20,9)	2,8	0	0	0

c) Etudes complémentaires

Afin de vérifier si les neuf SNP situés dans des régions introniques et concordants avec le statut au QTL des individus, se situent dans des sites putatifs de fixation de facteurs de transcription, une étude de comparaison a été faite entre les séquences contenant les 10 SNP et des sites de fixation de facteurs de transcription chez des organismes modèles tels que la

souris ou la drosophile à l'aide du logiciel TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html).

Cette étude a montré que les neuf SNP sont situés dans des régions homologues à des sites de fixation de trois types de facteurs de transcription dont les plus importants sont (tableau E-13) :

- *HSF (Heat Shock Factor)* : facteur de transcription qui régule l'expression d'un groupe de protéines : les « heat shock proteins » surexprimées lorsque les cellules sont exposées à des températures élevées ou à d'autres types de stress.
- *ADR1 (Alcohol Dehydrogenase gene Regulator 1)* : facteur de transcription contrôlant un grand nombre de fonctions régulatrices dont le gène ADH2 (glucose repressible alcohol dehydrogenase) impliqué dans le métabolisme du glycérol.
- *GATA-3 (GATA-binding factor 3)* : En consultant la base de données de IPA, ce facteur de transcription est impliqué dans plusieurs fonctions biologiques dont : la biosynthèse de norépinéphrine (neurotransmetteur impliqué dans la réponse au stress au niveau de l'hypothalamus), la réponse aux oestrogènes et les réponses immunitaires.

Tableau E-13 : *Résultats de l'étude d'homologie entre les séquences contenant les neuf SNP et des sites de fixation de facteurs de transcription chez des organismes modèle.*

Gène-Exon	Position du SNP (profil compatible Avec le statut au QTL)	Nom du facteur de transcription (espèce)	% d'homologie
MTX1_EX4-5-6	3142	-	-
SPNA1_EX16	20544 20644	HSF (*Drosophila melanogaster*) -	100 -
SPNA1_EX28-29	69980 70272	- GATA-3 (*Homo Sapiens*)	- 87,3
SPNA1_EX36	81068		-
SPNA1_EX_37-38	83402 83533	Adr1 (*Drosophila melanogaster*) -	100 -
SHC1_EX4-5-6	2167	HSF (*Mus musculus*)	87,3

2. Etude de gènes candidats sur l'ensemble du génome

Les résultats des analyses LDLA faites avec les deux index de fertilité « vache » et « génisse » en race Prim'Holstein a permis de faire une étude à travers le génome des gènes candidats qui pourraient être impliqués dans la baisse de la fertilité femelle chez cette race. Les gènes ont été sélectionnés dans une région de 2 Mb encadrant les pics des courbes LDLA identifiés selon les critères décrits.

a) Réseaux de gènes identifiés

A partir de la liste des références de gènes introduites, IPA est capable de sortir un classement des réseaux de gènes impliqués dans des fonctions biologiques surreprésentées dans la liste introduite par rapport au reste du génome, le classement est fait sur la base de « p-values » qui sont calculées sur la base de deux critères principaux :

- ✓ Le nombre de gènes de la liste qui participent dans la fonction biologique en question. Plus ce nombre est grand, moins l'association entre ces gènes est due au simple hasard et plus la « p-value » est significative.
- ✓ Le nombre total de gènes qui sont connus pour être associés, selon la base de données de IPA, à la fonction biologique en question.

Les figures E-21 et E-22 décrivent les fonctions biologiques les plus surreprésentées dans la liste de gènes introduite à IPA avec les deux index de fertilité « vache » (Figure E-21) et « génisse » (Figure E-22).

b) Fonctions biologiques surreprésentées par la liste de gènes identifiés avec l'index de fertilité « vache »

Plusieurs réseaux de gènes impliqués dans différentes fonctions biologiques, identifiés à partir de la liste de gènes introduite à IPA avec l'index fertilité « vache », apparaissent ; toutefois, les « p-value » de ces fonctions biologiques ne dépassent pas la valeur seuil fixée par IPA ; autrement dit, les fonctions biologiques identifiées avec la liste de gènes introduite à IPA en utilisant l'index de fertilité « vache », n'est pas surreprésentée par rapport au reste du génome.

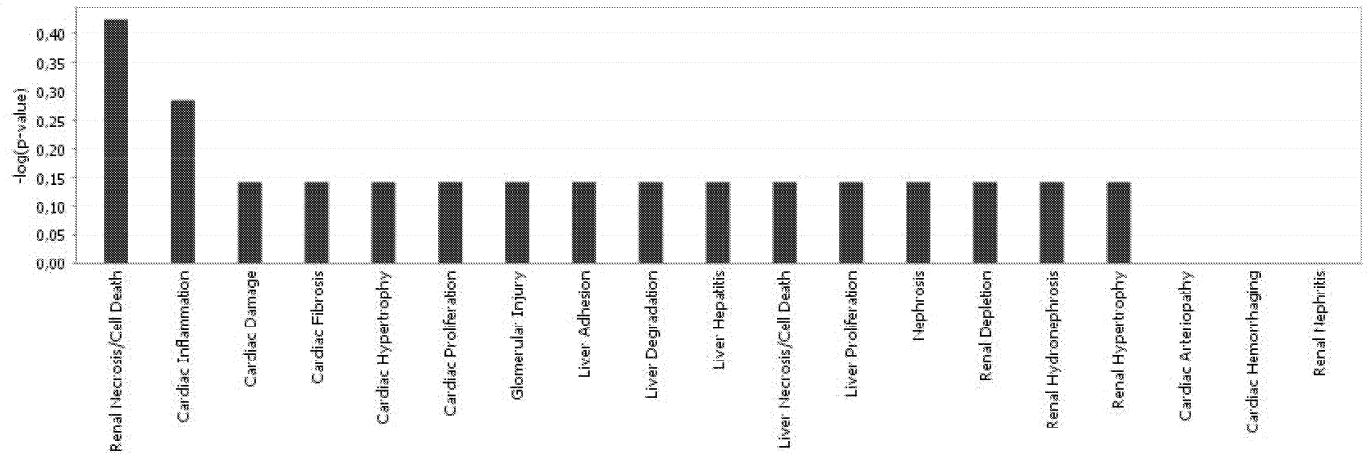


Figure E-21 : *Représentation graphique des fonctions biologiques les plus surreprésentées dans la liste de gènes identifiée obtenue avec l'index de fertilité « vache ». Les valeurs de – log (p-value) montrent qu'il n'y a pas de fonctions biologiques significativement surreprésentées dans la liste introduite, par rapport au reste du génome.*

c) Fonctions biologiques surreprésentées par la liste de gènes identifiés avec l'index de fertilité « génisse »

Toutes les « p-value » des 10 fonctions biologiques surreprésentées dans la liste de gènes identifiés, par rapport au reste du génome, obtenue avec l'index de fertilité « génisse » dépassent le seuil de significativité fixé. Les gènes impliqués dans les maladies du système reproductif et ceux impliqués dans les maladies infectieuses sont en quatrième position du point de vue sur-représentativité et sont représentés par 30 molécules dont la plupart sont impliquées dans le cancer des ovaires, de l'endomètre et du dysfonctionnement du placenta.

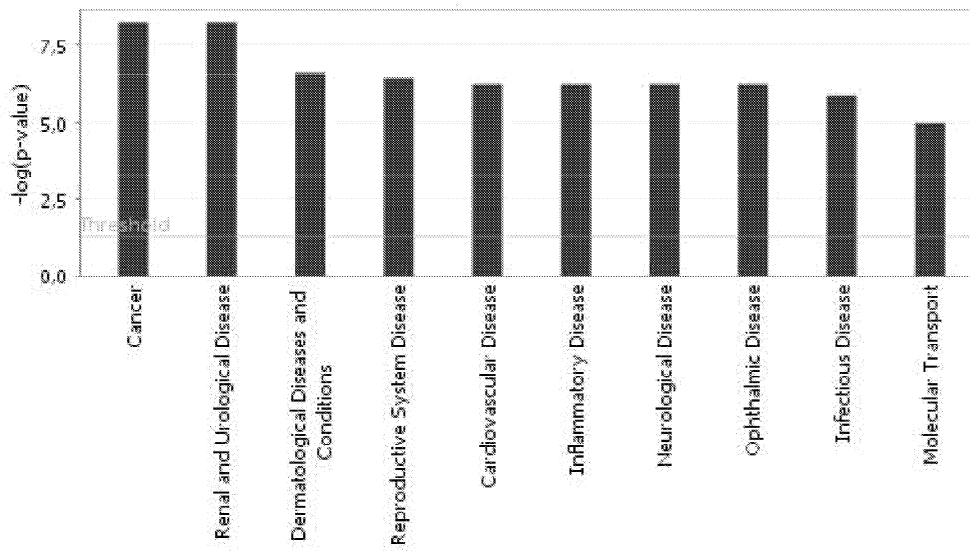


Figure E-22 : *Représentation graphique des 10 fonctions biologiques les plus surreprésentées dans la liste de gènes identifiée obtenue avec l'index de fertilité « génisse ». Le seuil de significativité est indiqué par la ligne orange. Toutes les fonctions biologiques listées dans cette figure sont significativement surreprésentées dans la liste introduite, par rapport au reste du génome.*

3. Etude de gènes candidats sélectionnés à partir des résultats de cartographie fine, obtenus par la puce Illumina

a) Inventaire des gènes présents dans les intervalles de localisation des QTL identifiés

(1) Inventaire des gènes présents dans les intervalles de localisation des QTL identifiés avec l'index de fertilité « vache »

Un total de 5652 identifiants de gènes a été récupéré à partir des différentes bases de données dans les régions encadrant 1 Mb de part et d'autre des pics obtenus par analyse LDLA avec l'index de fertilité « vache » sur tout le génome bovin. La répartition du nombre de pics et du nombre d'identifiants, par chromosome est indiquée dans le tableau E-14.

La liste des identifiants a été introduite dans IPA afin d'identifier des réseaux de gènes qui sont surreprésentés. A partir des 5652 identifiants introduits dans IPA, 4006 ont été assignés à des gènes (et donc cartographiés). A partir de ces 4006 identifiants, 1349 molécules de différents types (enzymes, cytokines, des canaux ioniques, protéines…) ont été identifiées. La différence entre le nombre d'identifiants assignés (4006) et le nombre de molécules identifiées (1349) sont principalement dues à la redondance de la liste des identifiants en input puisque plusieurs de ces identifiants désignent un même gène/molécule.

Tableau E-14 : *Nombre de pics (P<0,01) des courbes LDLA et d'identifiants de gènes récupérés, par chromosome, sur tout le génome bovin avec l'index de fertilité « vache » (S. Fritz, communication personnelle).*

Chromosome	Nombre de pics	Nombre d'identifiants de gènes
1	1	102
2	8	646
3	3	437
4	2	160
6	4	331
8	1	50
9	3	162
11	1	47
14	4	147
15	4	232
18	10	2797
19	2	496
28	1	45
Total	**44**	**5652**

(2) Inventaire des gènes présents dans les intervalles de localisation des QTL identifiés avec l'index de fertilité « génisse »

Un total de 1960 identifiants de gènes a été récupéré dans les régions encadrant 1 Mb de part et d'autre des pics obtenus, par analyse LDLA avec la puce Illumina 50K, sur tout le génome bovin. La répartition du nombre de pics des courbes LDLA (P<0,01) et du nombre d'identifiants, par chromosome est indiquée dans le tableau E-15. A partir des 1960 identifiants introduits à IPA, 1294 ont été reconnus par IPA. Parmi ces 1294 identifiants assignés, 660 molécules de différents types ont été identifiées.

Tableau E-15 : *Nombre de pics (P<0,01) des courbes LDLA et d'identifiants de gènes récupérés, par chromosome, sur tout le génome bovin avec l'index de fertilité « génisse » (S. Fritz, communication personnelle).*

Chromosome	Nombre de pics	Nombre d'identifiants de gènes
2	1	147
5	2	354
7	1	337
9	1	51
10	6	487
12	6	421
14	1	55
22	1	108
Total	**19**	**1960**

b) Sélection de gènes candidats positionnels et fonctionnels dans les intervalles de localisation identifiés

(1) Sélection de gènes candidats positionnels et fonctionnels identifiés avec l'index de fertilité « vache »

Un examen rapide des 1349 molécules identifiées prises individuellement a révélé l'existence de plusieurs produits géniques qui sont particulièrement intéressants car ils interviennent directement dans la reproduction et peuvent par conséquent affecter directement la fertilité femelle. La liste de ces gènes ainsi que leur rôle, sont indiqués dans le tableau E-16. Au total 14 gènes candidats positionnels et fonctionnels ont été identifiés lors de cette étude.

Tableau E-16 : *Nom, localisation et processus biologique des gènes candidats positionnels et fonctionnels identifiés avec l'index de fertilité « vache » en utilisant les résultats de cartographie fine obtenus avec la puce 50K.*

Nom du gène	Symbole	BTA	Position moyenne sur le BTA en Mb	Processus biologique
1-acylglycerol-3-phosphate O-acyltransferase 9	AGPAT9	6	102,214	Synthèse de phospholipides et de triacylglycérol
Stearoyl-CoA desaturase	SCD	6	101,324	métabolisme lipidique; biosynthèse d'acides gras.
Hydroxysteroid (17-beta) dehydrogenase 11	HSD17B11	6	106,322	biosynthèse de stéroides/ régulation des androgènes.
Hydroxysteroid (17-beta) dehydrogenase 13	HSD17B13	6	106,273	métabolisme des androgènes et des oestrogènes.
Hydroxysteroid (17-beta) dehydrogenase 14	HSD17B14	18	55,258	métabolisme lipidique/processus catabolique de stéroïdes.
Hepatocyte nuclear factor 4, gamma	HNF4G	14	35,620	récepteur d'hormones stéroïdes.

Nom du gène	Symbole	BTA	Position moyenne sur le BTA en Mb	Processus biologique
-	AKT	18	56,172	voie de signalisation des récepteurs d'insuline.
Nuclear receptor subfamily 1, group H, member 2	NR1H2	18	56,498	récepteur d'hormones stéroïdes/facteur de transcription.
Luteinizing hormone beta	LHB	18	55,580	régulation de progestérone/ ovulation
Lipolysis stimulated lipoprotein receptor	LSR	18	45,398	Dégradation des lipoprotéines dans le foie/ rôle dans les organes steroïdogènes/ développement du foie chez les embryons de souris.
Sulfotransferase family, cytosolic, 2B, member 1	SULT2B1	18	55,064	biosynthèse de stéroides/ régulation des androgènes
Synaptotagmin III	SYT3	18	56,669	neurotransmetteur/régulation de INS1
Upstream transcription factor 2, c-fos interacting	USF2	18	45,411	facteur de transcription/transcription de FSHR
Corticotropin releasing hormone receptor 1	CRHR1	19	47,316	Récepteur de la CRH : implantation de l'embryon dans l'utérus/ régulation des glucocorticoïdes (métabolisme protidique et glucidique).

(2) Sélection de gènes candidats positionnels et fonctionnels identifiés avec l'index de fertilité « génisse »

L'examen des 660 molécules identifiées prises individuellement a révélé l'existence de plusieurs produits géniques qui interviennent dans la reproduction et peuvent par conséquent affecter la fertilité des génisses. La liste de ces gènes et leurs rôles biologiques est indiquée dans le tableau E-17. Au total cinq gènes candidats positionnels et fonctionnels ont été identifiés lors de cette étude.

Tableau E-17 : *Nom, localisation et processus biologique des gènes candidats positionnels et fonctionnels identifiés avec l'index de fertilité « génisse » en utilisant les résultats de cartographie fine obtenus avec la puce 50K.*

Nom du gène	Symbole	BTA	Position moyenne sur le chromosome (Mb)	Processus biologique
Peroxisome proliferator-activated receptor gamma	PPARG	22	58,344	Vascularisation du placenta/activité de récepteur d'hormones stéroïdes
V-raf-1 murine leukemia viral oncogene homolog 1	RAF1	22	58,101	Vascularisation du placenta/activité de récepteur d'hormones stéroïdes
Spastic ataxia of Charlevoix-Saguenay (sacsin)	SACS	12	34,897	Sous influence de la HCG, SACS est surexprimé dans les cellules atteintes d'endométrite.

F. Discussion et perspectives

Au cours des derniers siècles, l'homme, à travers la sélection artificielle, a sélectionné des populations d'animaux domestiques uniques et caractérisés par des phénotypes très particuliers. Ainsi, sans aucune connaissance moléculaire des chromosomes, des gènes, et même des QTL, les éleveurs ont « manipulé » le génome des animaux domestiques par des approches de sélection classique. Aujourd'hui, les différentes populations animales représentent une ressource unique permettant de mieux appréhender et comprendre les bases moléculaires de la variation phénotypique. La découverte de marqueurs moléculaires abondants et/ou polymorphes a permis de développer les outils de la génétique moléculaire, ce qui a stimulé les sélectionneurs à investir dans l'identification de QTL influençant les caractères d'intérêt. L'identification de ces QTL a été principalement destinée à pratiquer et améliorer l'efficacité de la SAM. Toutefois, d'autres applications ont été développées à partir des données moléculaires produites. Elles consistent à identifier les bases moléculaires de la diversité phénotypique observée qui pourrait être exploitée pour une meilleure compréhension de la relation phénotype-génotype et d'étendre ces connaissances à d'autres espèces grâce aux outils de génomique comparée.

Un objectif de la détection de QTL est la SAM :

L'objectif principal de la cartographie de QTL influençant les caractères d'intérêt économique chez les animaux domestiques, est de faire de la SAM qui permet d'améliorer l'efficacité de la sélection en assistant les méthodes de sélection classiques basées sur les index génétiques, soit par des polymorphismes moléculaires directement responsables de la variation des caractères en question, soit par des marqueurs moléculaires liés à des polymorphismes responsables de la variation de ces caractères.

Pourquoi faire de la SAM ?

L'information issue des marqueurs moléculaires améliore le progrès génétique des caractères d'intérêt économique principalement en augmentant la précision des index de sélection classiques (faible pour les caractères à faible héritabilité) et en diminuant l'intervalle entre générations. La fertilité femelle des bovins laitiers est un caractère qui présente une faible héritabilité et par conséquent une faible précision des index de sélection classiques. De plus, ce caractère est exprimé tardivement au cours de la vie de l'animal. L'utilisation des informations moléculaires pourrait augmenter le progrès génétique réalisé pour la fertilité femelle, en prédisant la valeur génétique des candidats à la sélection, à un âge précoce. De plus, la fertilité femelle fait partie des caractères considérés comme secondaires dans les objectifs de sélection des bovins laitiers et ayant une corrélation génétique négative avec les caractères de production considérés comme primaires. L'utilisation des méthodes de sélection

classiques basées sur une approche globale qui s'affranchit totalement de l'information génomique ne tient que trop peu compte de l'effet négatif de l'amélioration des caractères de production sur les caractères secondaires.

L'utilisation de la SAM est également justifiée pour les caractères dont la variation génétique est masquée par une large variation phénotypique. La connaissance des bases moléculaires de ces caractères permettrait de mieux appréhender le caractère (ses bases physiologiques par exemple) et de mieux le définir afin de mieux le sélectionner (en choisissant des mesures du caractère moins biaisées par les effets environnementaux).

L'utilisation de la SAM pour l'amélioration génétique de la fertilité femelle est donc amplement justifiée puisque ce caractère présente les deux caractéristiques citées précédemment (antagonisme avec les caractères primaires dans les objectifs de sélection et large variabilité environnementale).

Facteurs influençant l'efficacité de la SAM

L'efficacité de la SAM et sa rentabilité économique sont d'autant plus grandes que la prédiction de la valeur génétique des reproducteurs par les marqueurs moléculaires est plus fiable. Celle-ci dépend principalement, pour un caractère donné, de la liaison entre les marqueurs utilisés en SAM et les polymorphismes causaux : plus la liaison est étroite, plus la SAM est efficace. Le degré de liaison entre les marqueurs moléculaires et les QTN dépend principalement de l'état d'avancement des travaux de cartographie de QTL, limité à son tour par la disponibilité d'une densité suffisante en marqueurs, de cartes génomiques bien élaborées et d'une population animale d'étude adéquate.

Evolution des moyens de génotypage et disponibilité d'une bonne qualité de l'assemblage

Au cours de ce projet de thèse, les différentes méthodes et moyens disponibles pour la cartographie de QTL de fertilité femelle ont fortement évolué avec notamment la disponibilité de la séquence complète du génome bovin, de dizaines de milliers de SNP et de technologies de génotypage à haut débit. En effet, le programme QTL initial (Boichard *et al.*, 2003) mis en place dès 1995 et dont les premiers résultats étaient disponibles dès 1999, a été réalisé à l'aide de 169 microsatellites répartis sur tout le génome ce qui revient à une densité en marqueurs de 1 microsatellite tous les 18,7 cM (en considérant une longueur totale du génome bovin d'environ 3160 cM (Ihara et al., 2004)). L'utilisation de SNP a permis de mettre à disposition des marqueurs génétiques très fréquents dans le génome, faciles à génotyper grâce à l'automatisation des procédures de génotypage ce qui a complètement changé l'échelle de résolution des méthodes de cartographie utilisées. L'utilisation de la puce 50K par exemple,

permet de passer à une densité de 1 SNP tous les 50 Kb environ, augmentant ainsi la puissance de résolution des dispositifs mis en place pour la cartographie de QTL.

La suite de la discussion s'intéressera aux résultats des différentes études réalisées tout en mettant l'accent sur l'impact de l'évolution des technologies mises en œuvre pour la production de données génomiques et des méthodes d'analyse de ces données sur les résultats de cartographie fine de QTL de fertilité femelle chez les bovins laitiers.

Importance du programme QTL initial dans l'élargissement du dispositif de détection de QTL

Le programme QTL initial (Boichard et al., 2003) a permis de mettre en évidence environ 120 QTL contrôlant 24 caractères enregistrés sur 14 familles des trois principales races bovines laitières Françaises : la Prim'Holstein, la Normande et la Montbéliarde. Trois QTL de fertilité femelle, mesurée par la réussite/échec de chaque IA à la fin de la période de gestation, ont été détectés sur ce dispositif : un QTL sur le BTA01, un QTL sur le BTA07 et un QTL sur le BTA21.

Les premiers résultats du programme QTL initial ont été valorisés dans un programme SAM de première génération où 14 régions génomiques ont été sélectionnées et génotypées sur de nouvelles familles (environ 8-10 000 individus par an). Le programme SAM a donc, permis aux programmes de recherche en cours, de disposer de nouvelles familles pour la cartographie de QTL. Ainsi, 78 familles appartenant aux trois races citées précédemment ont été disponibles pour la primo-localisation (détection) de QTL influençant différents caractères dont la fertilité femelle.

Discussion des résultats de l'étude de primo-localisation

L'élargissement du dispositif de détection de QTL à 78 familles des trois races a permis de vérifier si la détection des QTL lors du programme QTL initial est biaisée ou non par les deux facteur « échantillonnage » et « race ». En effet, les 78 familles choisies sont plus représentatives de la population globale pour chaque race, d'autant plus que les détections de QTL ont été réalisées indépendamment pour chaque race. Globalement, il y a eu des QTL qui n'ont pas été confirmés chez les trois races et de nouveaux QTL détectés en race Prim'Holstein, alors qu'en race Normande et Montbéliarde, aucun nouveau QTL n'a été détecté. En particulier, l'étude de primo-localisation n'a permis de confirmer qu'un seul des trois QTL de fertilité femelle détectés dans le programme QTL initial pour deux races majoritaires du dispositif (Prim'Holstein et Normande) de ce programme: celui du BTA01. Le QTL détecté sur le BTA21 n'a été confirmé que dans la race Normande. Il est plus

vraisemblable que ce QTL ne ségrégue que dans cette race. Le QTL sur le BTA07 n'a été confirmé chez aucune des races. Il est probable que ce QTL soit un faux positif.

D'autre part, l'ajout de nouvelles familles aux 14 familles du programme QTL initial des trois races a permis de détecter quatre nouveaux QTL chez les deux races Prim'Holstein et Normande : deux QTL significatifs (P<0,01) cartographiés sur les chromosomes BTA02 et BTA03 et deux QTL suggérés (P<0,05) sur les chromosomes BTA20 et BTA23 en race Prim'Holstein. A notre connaissance, une seule étude a rapporté des QTL affectant le TNR dans l'une de ces régions chromosomiques. En effet, Schrooten et al. (2000) ont rapporté un QTL affectant le TNR56 sur le BTA02 (à côté du marqueur BM2113) chez la race Prim'Holstein (Schrooten et al., 2000). Aucun QTL supplémentaire n'a été détecté chez la race Normande. Malgré l'ajout de cinq nouvelles familles Montbéliardes par rapport au programme QTL initial, aucun QTL de fertilité femelle n'a été détecté chez cette race. Il semblerait qu'aucun QTL de fertilité femelle ne soit en ségrégation dans les 14 régions génomiques SAM en race Montbéliarde.

Le QTL détecté sur le BTA01 en race Normande avec les 18 familles de l'étude de primo-localisation n'a pas été confirmé avec un échantillon de neuf familles choisies parmi les 18 familles initiales. Ceci est principalement dû au fait que deux des trois pères hétérozygotes, pour ce QTL, du dispositif de l'étude de primo-localisation n'ont pas été inclus dans le dispositif de l'étude de cartographie (tableau D-1). En race Prim'Holstein, par contre, les QTL initialement détectés sur les 47 familles de l'étude de primo-localisation sur les chromosomes BTA01, BTA02 et BTA03, ont été aussi détectés sur l'échantillon des 26 familles sélectionnées à partir des 47 familles initiales. Le maintien de la présence des QTL entre les deux études chez cette race est dû principalement au fait qu'une grande partie des pères hétérozygotes du dispositif de primo-localisation a été incluse dans l'échantillon des 26 familles du dispositif qui a servi pour faire l'étude de cartographie (tableau D-2).

En conclusion, il semblerait, d'après l'étude de primo-localisation, que les trois QTL détectés sur les chromosomes BTA01, BTA02 et BTA03 ne ségréguent que dans la race Prim'Holstein, alors que celui détecté sur le BTA21 ne ségrégue que dans la race Normande.

Tableau D-1 : *Description du nombre total de pères hétérozygotes dans le dispositif de primo-localisation et du nombre de pères hétérozygotes qui ont été inclus dans le dispositif de cartographie sur le chromosome BTA01 chez la race Normande.*

Chromosome	Nombre de Pères hétérozygotes de l'étude de primo-localisation	Nombre de pères hétérozygotes inclus dans l'étude de cartographie
BTA01	3	1

Tableau D-2 : *Description du nombre total de pères hétérozygotes(P<0,05) dans le dispositif de primo-localisation et du nombre de pères hétérozygotes qui ont été inclus dans lé dispositif de cartographie sur les trois QTL BTA01, BTA02 et BTA03 chez la race Prim'Holstein.*

Chromosome	Nombre de Pères hétérozygotes de l'étude de primo-localisation	Nombre de pères hétérozygotes inclus dans l'étude de cartographie
BTA01	9	6
BTA02	5	2
BTA03	10	8

Les régions étudiées dans l'étude de primo-localisation ne couvrent que 17% de l'ensemble du génome (en faisant le rapport entre les distances couvertes par l'ensemble des marqueurs des régions SAM et la longueur totale du génome bovin exprimée en cM). D'autres QTL affectant le TNR, existent vraisemblablement dans les autres régions du génome non couvertes pas les 39 microsatellites de l'étude de primo-localisation. Cette hypothèse est confirmée par les résultats de la puce 50K qui ont montré l'existence d'autres régions génomiques influençant la fertilité femelle en race Prim'Holstein.

Discussion des résultats de l'étude de cartographie

L'ajout de nouveaux microsatellites aux quatre régions génomiques les plus significatives identifiées dans l'étude de primo-localisation a permis d'augmenter l'informativité du dispositif et d'affiner l'intervalle de localisation des QTL détectés sur les chromosomes BTA01 et BTA03 dans un échantillon de 26 familles (appartenant aux 47 familles Prim'Holstein du dispositif de l'étude de primo-localisation). Les $IC_{95\%}$ des QTL sur les chromosomes BTA02 et BTA21 n'ont pas été réduits car les régions couvertes par les nouveaux microsatellites sont beaucoup plus larges que celles couvertes par les marqueurs SAM. Les $IC_{95\%}$ des QTL sur le BTA01 et BTA03 (respectivement 51 cM et 58 cM, pour les chromosomes BTA01 et BTA03) demeurent larges et ne permettent pas de faire une étude des gènes candidats dans ces régions vu le grand nombre de gènes présents. L'ampleur des $IC_{95\%}$

estimés est en partie due, à la méthode utilisée pour les calculer ; en effet, la méthode de « bootstrap » est une méthode conservatrice.

Les intervalles de localisation des QTL de fertilité femelle sur les chromosomes BTA01 et BTA03 permettent de faire une SAM de première génération dont l'efficacité est limitée puisqu'elle ne peut être faite qu'à l'intérieur d'une même famille car les allèles des marqueurs associés à un allèle donné du QTL diffèrent d'une famille à une autre. De plus, la probabilité de recombinaison entre les marqueurs utilisés dans la SAM et le QTL est très élevée.

L'utilisation de la puce 50K a permis de faire une cartographie fine des QTL de fertilité femelle dans tout le génome bovin et a permis d'avoir des intervalles de localisation des QTL, délimités par deux SNP successifs, de quelques centaines de kilobases. Les SNP délimitant des intervalles de localisation sont en déséquilibre de liaison avec le QTL dans toute la population, ce qui permet de faire une SAM de deuxième génération.

Apport de l'utilisation de phénotypes de fertilité femelle plus précis au cours de la gestation

Au cours de l'étude de cartographie à moyenne résolution, des TNR calculés à 56, 90 et 281 jours après IA ont été utilisés pour la détection de QTL. L'objectif étant de pouvoir distinguer entre les évènements affectant ce caractère au cours des différentes périodes de la gestation qui sont vraisemblablement influencées par des gènes différents. L'utilisation de ces TNR devrait donc améliorer la détection de QTL de fertilité femelle en affinant les phénotypes utilisés.

L'utilisation des TNR90 et TNR281 n'a pas contribué à la distinction entre différents effets de QTL: les courbes d'analyse de liaison obtenues avec les deux types de données ont globalement les mêmes allures avec des seuils de significativité parfois différents (tels que celles du BTA01 où le QTL était significatif ($P<0,01$) avec le TNR90 et seulement suggéré ($P<0,05$) avec le TNR281). Par contre, la distinction entre le TNR56 d'une part et les TNR90 et TNR281 d'autre part a permis de détecter deux QTL différents affectant le TNR avant et après 56 jours de l'IA sur le BTA02. La similitude entre les courbes de détection avec les TNR90 et TNR281 et la différence entre les courbes de détection avec le TNR56 d'une part et les TNR90 et TNR281 d'autre part est due aux corrélations entre ces TNR : elle est de 0,94 entre TNR90 et TNR281, alors qu'elle n'est que de 0,81 entre les TNR56 et TNR281; celle entre les TNR90 et TNR281 est de 0,91 (Guillaume, communication personnelle).

Le QTL reporté par Schrooten et al. (2000) affectant le TNR56 est localisé dans la même position que celui affectant les TNR90 et TNR281 obtenus avec nos données. Il pourrait s'agir du même QTL qui affecte les deux types de données dans les deux études.

Discussion des résultats de cartographie fine avec les 437 SNP sur le BTA03

L'analyse LDLA réalisée à la SGQA à l'aide d'un panel de 437 SNP a permis de réduire l'intervalle de localisation du QTL du BTA03 à une région de 4,77 Mb comprise entre 10,93 Mb et 15,70 Mb. Les pics des courbes LDLA sur le BTA03, obtenus en utilisant la puce 50K sont situés à l'extrémité télomérique de ce chromosome (figure F-1). Ces positions sont différentes de celles obtenues, par analyse LDLA, avec le panel de 437 SNP. Les différences entre les pics de localisation des QTL entre les deux études s'expliquent principalement par :

- ✓ Une augmentation de l'informativité du dispositif due à l'augmentation de la densité en marqueurs ; ainsi des familles non informatives pour des régions situées à la fin du BTA03 (à cause d'une densité en marqueurs plus faible) sont devenues informatives dans cette région permettant ainsi de détecter des régions ayant une influence sur les phénotypes de fertilité utilisés.
- ✓ La différence entre les dispositifs utilisés dans les deux études : la première (celle utilisant les 437 SNP) utilisait 17 familles alors que la deuxième (celle utilisant la puce 50K) 42 familles. Ainsi, les nouvelles familles ajoutées au dispositif pourraient être informatives pour les nouvelles régions détectées avec la puce 50K et moins informatives pour les régions qui étaient significatives lors de la première analyse.
- ✓ Les phénotypes de fertilité utilisés dans l'identification de QTL, avec la puce 50K, ne sont pas des TNR. Le changement des phénotypes de fertilité femelle entre les deux études pourrait aussi être en partie à l'origine de la différence entre les deux analyses.

Les trois explications présentées ci-dessus ont vraisemblablement contribué conjointement à une différence entre les résultats de détection obtenus avec les 437 SNP et ceux obtenus avec la puce 50K pour le BTA03.

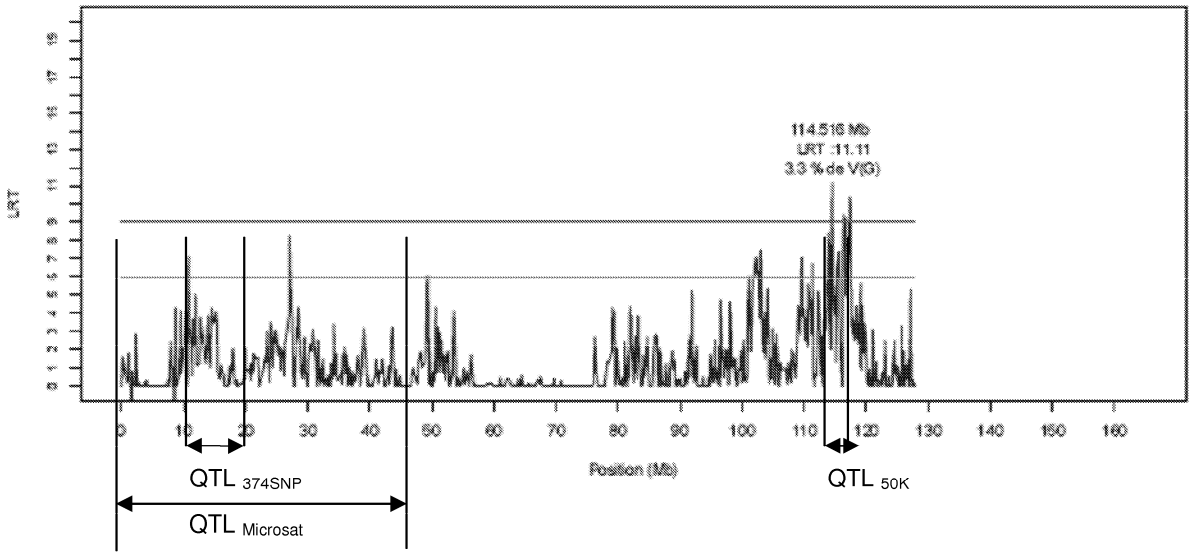


Figure F-1 : *Figure représentant le résultat de détection de QTL de fertilité femelle sur le BTA03 avec la puce 50K en utilisant l'approche LDLA (courbe en bleu). La valeur maximale du LRT, ainsi que sa position et le pourcentage de variation génétique, V(G), expliqué par le QTL sont indiqués sur cette figure. Les deux droites indiquent respectivement, de haut en bas, les seuils à 1% et à 5%. Un $IC_{95\%}$ du QTL détecté avec le panel de microsatellites ($QTL_{microsat}$) et celui détecté avec le panel de 437 SNP (QTL_{437SNP}) sont indiqués avec les flèches doubles. L'intervalle des pics dépassant le seuil de 1% représente le QTL détecté avec la puce 50K (QTL_{50K}). L'axe des abscisses représente la position, en mégabases, sur le BTA03.*

Discussion des résultats de séquençage des gènes candidats

Les résultats de cartographie fine du QTL de fertilité femelle sur le chromosome BTA03 obtenus avec le panel de 437 SNP, ont été utilisés pour le choix de six gènes candidats positionnels et fonctionnels sur ce chromosome. Ce choix s'est fait en exploitant les données de cartographie comparée avec les génomes humain et murin. Le séquençage de certaines régions incluant les exons de ces gènes chez des individus de phénotypes extrêmes issus de pères hétérozygotes visait à maximiser les chances d'identifier des polymorphismes intéressants qui pourraient être validés ultérieurement sur l'ensemble du dispositif.

Parmi les 10 SNP identifiés dont le statut est concordant avec celui des individus au QTL, neuf SNP sont situés dans des introns. Les neuf sites polymorphes identifiés font partie de sites de fixation de trois types de facteurs de transcription chez des espèces modèles. Des mutations dans chacun de ses facteurs de transcription identifiés pourraient affecter d'une manière ou d'une autre la fertilité des vaches laitières *post partum* :

- ✓ Le premier type de facteurs de transcription (HSF) est présent chez les mammifères sous la forme de trois isoformes : HSF-1, HSF-2 et HSF-4. Les souris femelles KO pour HSF-1 ont des problèmes de fertilité femelle et des anomalies du placenta (Xiao *et al.*, 1999). Les souris KO pour HSF-2 sont

subfertiles et présentent une production anormale d'ovules, une réduction du nombre de follicules et la présence de follicules cystiques hémorragiques dus à des défauts méiotiques. Ces souris présentent en plus des défauts de leur réponse hormonale et des taux anormaux des ARNm des récepteurs à LH (Kallio *et al.*, 2002).

- ✓ Le deuxième type de facteurs de transcription (Adr1) code pour un activateur transcriptionnel impliqué dans l'expression de gènes régulés par la répression du glucose. Adr1 intervient dans l'expression de gènes impliqués le métabolisme du glycérol et des acides gras (http://db.yeastgenome.org/cgi-bin/locus.pl?locus=ADR1). Des dérégulations de ces deux types de métabolismes sont à l'origine de plusieurs maladies métaboliques associées à une fertilité réduite chez les vaches laitières (page 30).
- ✓ Le troisième type de facteurs de transcription (GATA-3) est impliqué dans au moins deux processus physiologiques qui peuvent affecter la fertilité femelle des bovins laitiers : la biosynthèse de norépinéphrine et la réponse aux oestrogènes.

Le dixième SNP situé dans la région régulatrice du gène PAQR6 pourrait être intéressant à examiner et à valider sur le reste du dispositif. Ce gène appartient à une famille de gènes : les récepteurs de progestine et de adipoQ (PAQR pour progestin and adipoQ receptors) qui sont très conservés entre espèces. Récemment, il a été démontré que le gène PAQR6 est impliqué dans la réponse à la progestérone (Smith *et al.*, 2008) et qu'il a clairement subi une sélection positive chez plusieurs races bovines (Gautier *et al.*, sous presse) ce qui a été observée lors d'études qui s'intéressent aux signatures de sélection chez ces races. Le gène PAQR6 est un gène candidat sérieux qui pourrait expliquer une part de la baisse de fertilité chez les races bovines laitières Françaises.

Aucun polymorphisme identifié dans des exons du gène ADAR1 n'est concordant avec le statut des individus au QTL. Ce gène est largement exprimé chez les mammifères et son rôle biologique n'est pas totalement élucidé. Toutefois, on sait que ce gène code pour une enzyme responsable de l'édition de l'ARNm et qu'il est impliqué dans des maladies génétiques liées au métabolisme des purines. Des souris mutantes pour deux allèles de ADAR1 subissent une désagrégation du foie et meurent tôt au cours de l'embryogenèse (jours 11,5-12,5) (Hartner *et al.*, 2004). Il est important de souligner que le statut au QTL des individus n'est pas certain. Les autres polymorphismes trouvés non concordants avec le statut prédéfini des individus, en particulier ceux trouvés dans les exons de ADAR1, ne doivent pas systématiquement être exclus de l'étape de validation (qui doit suivre celle de la recherche de polymorphismes).

Discussion des critères de choix des gènes candidats

Le choix des six gènes candidats s'est fait sur la base de leur implication dans deux processus biologiques principaux :

- Leur rôle dans le développement embryonnaire et notamment si l'absence du gène en question ou sa mutation est à l'origine de mortalité embryonnaire chez la souris.
- Leur rôle dans la reproduction : le gène candidat influence un processus directement lié à la reproduction (mécanismes hormonaux, régulation de produits géniques impliqués dans la reproduction…) chez la souris.

Ce choix est logique étant donné que la fertilité femelle est directement affectée par ces deux processus biologiques. Toutefois, en prenant en compte les causes de la baisse de fertilité liées au détournement des ressources d'énergie vers les besoins de la production laitière au détriment de la reproduction, le choix de gènes candidats codant pour des produits impliqués dans les voies d'interaction du système reproductif avec d'autres systèmes, semble également être justifié. La figure F-2 représente un exemple de ce type d'interactions. Elle présente d'une manière simplifiée le « dialogue mutualisé » existant entre le système reproductif (représenté principalement par l'ovaire et l'utérus), le système nerveux central et le système somatotrope. Les messagers principaux de ce dialogue sont les métabolites (TAG, cholestérol …) et les hormones (progestérone, oestradiol, insuline, leptine …). Dans le cas d'une fécondation, une quatrième composante, celle de l'embryon, interagit avec les trois systèmes précédemment cités. La relation entre ces composantes devient encore plus complexe et l'équilibre entre elles est plus fragile. Le choix de gènes candidats affectant l'une de voies connectant ces quatre composantes est donc justifié.

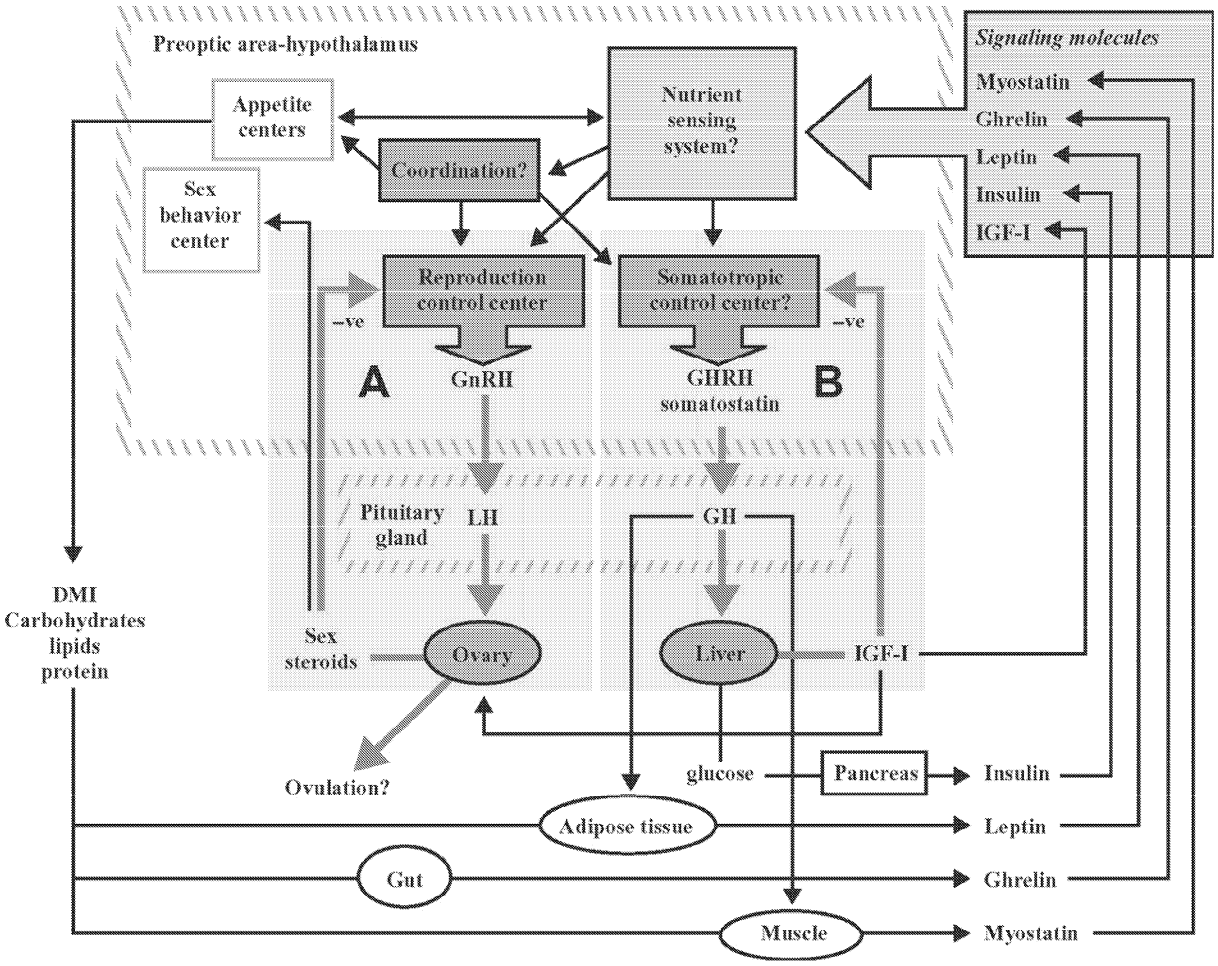


Figure F-2 : *Représentation schématique simplifiée de l'interaction entre l'axe de reproduction (aire grisée A), et l'axe somatotrope (aire grisée B). Ces interactions se font par l'intermédiaire d'hormones, de métabolites et d'éléments nutritifs (d'après Chagas et al., 2007).*

Toutefois, la prise en compte de ce type d'interactions dans le choix de gènes candidats contribue à augmenter considérablement le nombre de gènes à étudier et par conséquent leur séquençage devient une opération onéreuse. La multitude de critères qui peuvent servir à sélectionner les gènes candidats provient en partie de l'ambiguïté des phénotypes utilisés dans la détection qui est à son tour due à la globalité de ces phénotypes.

Limites des phénotypes utilisés et proposition de phénotypes alternatifs

Les TNR sont des phénotypes de fertilité femelle qui sont la résultante d'un ensemble de phénomènes physiologiques multiples, même pour le plus précoce d'entre eux (TNR56), d'autant plus que ces TNR ne traduisent pas un phénomène physiologique très précis. En plus, les données brutes à partir desquelles les différents TNR ont été élaborées sont celles des résultats d'IA enregistrés dans les élevages : ces mesures peuvent être biaisées par des effets liés à la gestion du troupeau tels que les erreurs commises lors de l'enregistrement des données d'IA ou les inséminations qui ont lieu au mauvais moment du cycle oestral.

Tous ces facteurs font que les TNR (et plus généralement ceux basés sur les enregistrements des données d'IA) sont peu fiables quant à la correspondance phénotype-génotype des individus.

Il est donc important que des phénotypes plus précis basés sur des mesures de phénomènes physiologiques liés à la reproduction soient utilisés pour le calcul d'index de fertilité femelle ; cela permettrait d'éliminer une bonne part de la variabilité environnementale (due aux aléas de la gestion du troupeau) et par conséquent d'avoir une meilleure correspondance phénotype-génotype permettant d'augmenter les valeurs de l'héritabilité du caractère.

Le système endocrinien joue un rôle crucial dans la reproduction des mammifères et la baisse de fertilité de la vache laitière se traduit souvent par des dérégulations hormonales. Des dosages hormonaux périodiques pourraient servir à une caractérisation plus pertinente des phénomènes physiologiques liés à la reproduction chez la vache laitière et à définir des phénotypes de fertilité précis. Le dosage de la concentration de progestérone dans le lait, par exemple, est utilisé pour caractériser des phénomènes physiologiques normaux tels que l'intervalle [vêlage -début de l'activité lutéale] ou la durée de la première activité lutéale après vêlage. Le dosage de progestérone peut être aussi utilisé pour caractériser des anomalies de reproduction telles que celle des corps jaunes persistants. Le choix d'exploiter des dosages de progestérone est justifié par le rôle dominant de cette hormone dans les cycles ovariens et dans le processus de gestation de la vache laitière.

Toutefois, les données de dosages hormonaux (de la progestérone en particulier) ne sont pas enregistrées en routine dans les contrôles laitiers en France ; par conséquent l'obtention de telles données va engendrer des coûts supplémentaires qui pourront être largement justifiés si l'estimation des paramètres génétiques de caractères dérivant de ce type de données aboutit à des valeurs d'héritabilité assez élevées. Pour que la caractérisation des stades physiologiques déterminées soit fiable, une fréquence de dosage de progestérone de deux à trois fois par semaine est nécessaire, ce qui peut paraître difficile à réaliser dans le cadre de contrôles laitiers de routine ; toutefois des dosages mensuels de cette hormone aboutissent à des niveaux de fiabilité acceptables. Petersson *et al.* (2008) ont montré que l'analyse de données issues de prélèvements de progestérone deux ou trois fois par semaine aboutit à une fiabilité de détection de retard de la reprise de l'activité ovarienne après vêlage (dans les 60 jours qui suivent le vêlage) de l'ordre de 0,94-0,99, alors que des prélèvements mensuels aboutissent à des valeurs de l'ordre de 0,85-0,88 (Petersson *et al.*, 2008).

Résultats préliminaires de la puce 50K

L'étape de validation de certains SNP trouvés dans les gènes candidats n'a pas pu avoir lieu du fait de la commercialisation de la puce 50K. Le séquençage de petites portions du génome

(celles contenant les polymorphismes détectés) n'est plus rentable par rapport aux perspectives qu'offre la nouvelle puce. En effet, le génotypage de la puce 50K sur le dispositif animal utilisé pour faire une détection de QTL de fertilité femelle offre l'opportunité de cartographier des QTL sur l'ensemble du génome bovin au lieu de s'intéresser à des régions limitées du génome. Sachant que les caractères de fertilité ont une faible héritabilité et une variation génétique élevée, les QTL contrôlant ces caractères sont vraisemblablement répartis sur différentes régions du génome pouvant être en interaction les unes avec les autres.

Les premières analyses de la puce 50K faites avec les deux index de fertilité « vache » et « génisse » chez la race Prim' Holstein ont permis de faire une cartographie fine des QTL de fertilité femelle dans tout le génome bovin et d'avoir des intervalles de localisation des QTL, délimités par deux SNP successifs, de quelques centaines de kilobases. Les SNP délimitant des intervalles de localisation sont en déséquilibre de liaison avec le QTL dans toute la population, ce qui permet de faire une SAM de deuxième génération. Le nombre de pics obtenus pour chacun des index est en accord avec la multifactorialité de ce caractère. La localisation des pics de détection est différente entre les deux index de fertilité « vache » et « génisse », soulignant des différences au niveau des mécanismes moléculaires sous-jacents entre multipares et primipares au moins pour la race Prim'Holstein.

Résultats préliminaires de l'analyse des réseaux de gènes

L'analyse IPA a permis d'illustrer la représentativité de réseaux de gènes impliqués dans des fonctions biologiques liées à la reproduction, dans les listes de gènes introduites, par rapport à l'ensemble du génome. Il faut toutefois noter que les fonctions biologiques identifiées proviennent de la base de données de IPA, dont la principale source d'informations provient de trois génomes : l'Homme, la souris et le rat. Vu les différences physiologiques évidentes, entre l'espèce bovine et ces trois espèces au niveau des causes de la baisse de fertilité (causes vraisemblablement liées à la production laitière), les informations issues de ces génomes ne sont pas entièrement satisfaisantes pour appréhender les causes de la baisse de fertilité femelle chez l'espèce bovine.

Le fait qu'aucune fonction biologique (représentée par un réseau de gènes donné) ne soit significative avec l'index de fertilité « vache », n'exclut pas qu'il puisse y avoir des molécules qui peuvent affecter la reproduction de la vache et qui sont incluses dans l'un des réseaux mentionnés. L'examen de l'ensemble des molécules de tous les réseaux de gènes identifiés est donc justifié.

Cet examen a abouti à un groupe de molécules candidates pouvant influencer la fertilité femelle des bovins laitiers, sélectionnées pour chaque type d'indices, à partir d'une liste de molécules identifiées par IPA : la sélection a concerné les molécules ayant un lien net avec la

fertilité femelle des bovins laitiers et en se basant sur ce qui a été rapporté dans la bibliographie pour expliquer les causes de la baisse de fertilité des bovins laitiers. D'autres molécules ayant un lien moins explicite avec la fertilité n'ont pas été sélectionnées.

Les molécules sélectionnées avec l'index de fertilité « vache » peuvent être divisées en trois catégories : des molécules impliquées dans un rôle métabolique, des molécules impliquées dans le système endocrinien contrôlant la reproduction et des molécules impliquées dans le développement embryonnaire ; alors que les molécules sélectionnées avec l'index de fertilité « génisse » se classent dans une seule catégorie : celles impliquées dans le système endocrinien contrôlant la reproduction.

Conclusion et perspectives :

Ce projet de thèse a été témoin de l'évolution spectaculaire des moyens et des techniques de génétique moléculaire mises à la disposition des chercheurs pour l'identification de gènes d'intérêt. En effet, les nouveaux moyens de génotypage (la puce 50K) conjugués à une amélioration de la qualité de l'assemblage de la séquence bovine, ont eu un impact évident sur la précision de résolution des approches de cartographie de QTL de fertilité femelle. En effet, les intervalles de localisation des QTL sont passés de l'ordre de quelques dizaines de centimorgans à seulement quelques centaines de kilobases. En plus, les nouveaux moyens de génotypage ont permis de cartographier des QTL de fertilité femelle dans tout le génome bovin avec une résolution élevée ; ce qui n'était pas le cas avec les microsatellites où on ne peut s'intéresser qu'à un nombre limité de régions avec une résolution moyenne.

La densité en SNP de cette puce permet de simplifier les approches de cartographie : des études d'association sont désormais possibles pour la cartographie de QTL. Toutefois, les modèles d'association utilisés doivent tenir compte de la forte structuration de la population de bovins laitiers afin d'éviter les associations trompeuses qui peuvent en résulter.

Pour les professionnels, la disponibilité de cartes génétiques de plus en plus denses en marqueurs moléculaires permettra d'envisager la mise en place d'une sélection génomique de leurs reproducteurs pour la fertilité femelle. Toutefois, l'amélioration des phénotypes de fertilité femelle utilisés semble être le facteur le plus limitant pour une meilleure résolution des méthodes de cartographie. En effet, ce caractère tel qu'il est mesuré aujourd'hui en France, est un caractère très complexe qui peut être dû à de multiples facteurs d'origine génétique différente. Il est donc important, pour améliorer l'efficacité des moyens de génétique moléculaire disponibles actuellement d'affiner la définition des phénotypes de fertilité femelle. La possibilité de l'introduction de la pratique des dosages hormonaux dans les contrôles laitiers est à évaluer sérieusement.

Références bibliographiques

Abdel-Azim G. A., A. E. Freeman, M. E. Kehrli, Jr., S. C. Kelm, J. L. Burton, A. L. Kuck et S. Schnell, 2005. Genetic basis and risk factors for infectious and noninfectious diseases in US Holsteins. I. Estimation of genetic parameters for single diseases and general health. J Dairy Sci 88: 1199-1207.

Andersen-Ranberg I. M., G. Klemetsdal, B. Heringstad et T. Steine, 2005. Heritabilities, genetic correlations, and genetic change for female fertility and protein yield in Norwegian Dairy Cattle. J Dairy Sci 88: 348-355.

Asdell S. A., 1964 Patterns of mammalian reproduction., London.

Ayalon N., 1978. A review of embryonic mortality in cattle. J Reprod Fertil 54: 483–493.

Ball P. J. H. et A. R. Peters, 2004 Reproduction in cattle.

Barbat A., A. Ginon et V. Ducrocq, 2007. L'évaluation génétique de la fertilité en France. BTiA: La revue Française de génétique et de reproduction. 126: 19-22.

Beam S. W. et W. R. Butler, 1997. Energy balance and ovarian follicle development prior to the first ovulation postpartum in dairy cows receiving three levels of dietary fat. Biol Reprod 56: 133-142.

Beam S. W. et W. R. Butler, 1999. Effects of energy balance on follicular development and first ovulation in postpartum dairy cows. J Reprod Fertil Suppl 54: 411-424.

Benjamini Y. et Y. Hochberg, 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Statistic Soc B 57: 289-300.

Ben Jemaa S., S. Fritz, F. Guillaume, T. Druet, C. Denis, A. Eggen et M. Gautier, 2008. Detection of quantitative trait loci affecting non-return rate in French dairy cattle. J Anim Breed Genet 125: 280-288.

Bennewitz J., N. Reinsch, S. Paul, C. Looft, B. Kaupe, C. Weimann, G. Erhardt, G. Thaller, C. Kuhn, M. Schwerin, H. Thomsen, F. Reinhardt, R. Reents et E. Kalm, 2004. The DGAT1 K232A mutation is not solely responsible for the milk production quantitative trait locus on the bovine chromosome 14. J Dairy Sci 87: 431-442.

Boichard D., A. Barbat et M. Briend, 1998 Evaluation génétique des caractères de fertilité femelle chez les bovins laitiers., pp. 103-106 in 5es rencontres autour des recherches sur les ruminants edited by P. INSTITUT DE L'ÉLEVAGE, FRANCE., Paris.

Boichard D., S. Fritz, M. N. Rossignol, M. Y. Boscher, A. Malafosse et J. J. Colleau, 2002 Implementation of marker-assisted selection in Frensh dairy cattle. In 7th World Congress on Genetics Applied to Livestock Production., Montpellier-France.

Boichard D., S. Fritz, M. N. Rossignol, F. Guillaume, J. J. Colleau et T. Druet, 2006 Implementation of marker-assisted selection: practical lessons from dairy cattle. In 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte.

Boichard D., C. Grohs, F. Bourgeois, F. Cerqueira, R. Faugeras, A. Neau, R. Rupp, Y. Amigues, M. Y. Boscher et H. Leveziel, 2003. Detection of genes influencing economic traits in three French dairy cattle breeds. Genet Sel Evol 35: 77-101.

Boichard D. et E. Manfredi, 1994. Genetic analysis of conception rate in French Holstein dairy cattle. Acta Agric Scand, Sect A 44: 138-145.

Buckley F., P. Dillon, M. Rath et R. F. Veerkamp, 2000. The relationship between genetic merit for yield and live weight, condition score, and energy balance of spring calving Holstein Friesian dairy cows on grass based systems of milk production. J Dairy Sci 83: 1878-1886.

Butler W. R., 1998. Review: effect of protein nutrition on ovarian and uterine physiology in dairy cattle. J Dairy Sci 81: 2533-2539.

Butler W. R., 2001. Nutritional effects on resumption of ovarian cyclicity and conception rate in postpartum dairy cows. Anim Sc. Occas Publ 26: 133-145.

Butler W. R., 2003. Energy balance relationships with follicular development, ovulation and fertility in postpartum dairy cows. Livest Prod Sci 83: 211-218.

Butler W. R. et R. D. Smith, 1989. Interrelationships between energy balance and postpartum reproductive function in dairy cattle. J Dairy Sci 72: 767-783.

Calus M. P., A. F. Groen et G. De Jong, 2002. Genotype x environment interaction for protein yield in Dutch dairy cattle as quantified by different models. J Dairy Sci 85: 3115-3123.

Carlborg O. et C. S. Haley, 2004. Epistasis: too often neglected in complex trait studies? Nat Rev Genet 5: 618-625.

Cetica P., L. Pintos, G. Dalvit et M. Beconi, 2002. Activity of key enzymes involved in glucose and triglyceride catabolism during bovine oocyte maturation in vitro. Reproduction 124: 675-681.

Chagas L. M., J. J. Bass, D. Blache, C. R. Burke, J. K. Kay, D. R. Lindsay, M. C. Lucy, G. B. Martin, S. Meier, F. M. Rhodes, J. R. Roche, W. W. Thatcher et R. Webb, 2007. Invited review: New perspectives on the roles of nutrition and metabolic priorities in the subfertility of high-producing dairy cows. J Dairy Sci 90: 4022-4032.

Chang Y. M., I. M. Andersen-Ranberg, B. Heringstad, D. Gianola et G. Klemetsdal, 2006. Bivariate analysis of number of services to conception and days open in Norwegian red using a censored threshold-linear model. J Dairy Sci 89: 772-778.

Chebel R. C., J. E. Santos, J. P. Reynolds, R. L. Cerri, S. O. Juchem et M. Overton, 2004. Factors affecting conception rate after artificial insemination and pregnancy loss in lactating dairy cows. Anim Reprod Sci 84: 239-255.

Clop A., F. Marcq, H. Takeda, D. Pirottin, X. Tordoir, B. Bibe, J. Bouix, F. Caiment, J. M. Elsen, F. Eychenne, C. Larzul, E. Laville, F. Meish, D. Milenkovic, J. Tobin, C. Charlier et M.

Georges, 2006. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. Nat Genet 38: 813-818.

Committee on bovine reproductive nomenclature, 1972. Recommendations for standardizing bovine reproductive terms. Cornell Vet 62: 216–237.

Cohen-Zinder M., E. Seroussi, D. M. Larkin, J. J. Loor, A. Everts-Van Der Wind, J. H. Lee, J. K. Drackley, M. R. Band, A. G. Hernandez, M. Shani, H. A. Lewin, J. I. Weller et M. Ron, 2005. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Res 15: 936-944.

Coppieters W., A. Kvasz, J. J. Arranz, B. Grisart, J. Riquet, F. Farnir et M. Georges, 1999. The great-grand-daughter design : a simple strategy to increase the power of a grand-daughter design for qtl mapping. Genet Res 74: 189-199.

Dal Zotto R., M. De Marchi, C. Dalvit, M. Cassandro, L. Gallo, P. Carnier et G. Bittante, 2007. Heritabilities and genetic correlations of body condition score and calving interval with yield, somatic cell score, and linear type traits in Brown Swiss cattle. J Dairy Sci 90: 5737-5743.

Darwash A. O., G. E. Lamming et M. D. Royal, 2001. A protocol for initiating oestrus and ovulation early postpartum in dairy cows. Anim Sci 72: 539–546.

Darwash A. O., G. E. Lamming et J. A. Woolliams, 1997. Estimation of genetic variation in the interval from calving to postpartum ovulation of dairy cows. J Dairy Sci 80: 1227-1234.

Darwash A. O., G. E. Lamming et J. A. Woolliams, 1999. The potential for identifying heritable endocrine parameters associated with fertility parameters associated with fertility in postpartum dairy cows. Anim Sci 68 333–347.

De Haas Y., L. L. Janss et H. N. Kadarmideen, 2007. Genetic correlations between body condition scores and fertility in dairy cattle using bivariate random regression models. J Anim Breed Genet 124: 277-285.

Diskin M. G. et D. G. Morris, 2008. Embryonic and early foetal losses in cattle and other ruminants. Reprod Domest Anim 43 Suppl 2: 260-267.

Diskin M. G., J. J. Murphy et J. M. Sreenan, 2006. Embryo survival in dairy cows managed under pastoral conditions. Anim Reprod Sci 96: 297-311.

Doerge R. W. et G. A. Churchill, 1996. Permutation tests for multiple loci affecting a quantitative character. Genetics 142: 285-294.

Druet T., S. Fritz, M. Boussaha, S. Ben-Jemaa, F. Guillaume, D. Derbala, D. Zelenika, D. Lechner, C. Charon, D. Boichard, I. G. Gut, A. Eggen et M. Gautier, 2008. Fine mapping of quantitative trait loci affecting female fertility in dairy cattle on BTA03 using a dense single-nucleotide polymorphism map. Genetics 178: 2227-2235.

Dunne L. D., M. G. Diskin et J. M. Sreenan, 2000. Embryo and foetal loss in beef heifers between day 14 of gestation and full term. Anim Reprod Sci 58: 39-44.

Evans A. C., 2003. Characteristics of ovarian follicle development in domestic animals. Reprod Domest Anim 38: 240-246.

Falconer D. S. et F. C. Mackay, 1996 Introduction to quantitative genetics, Edinburgh.

Farnir F., W. Coppieters, J. J. Arranz, P. Berzi, N. Cambisano, B. Grisart, L. Karim, F. Marcq, L. Moreau, M. Mni, C. Nezer, P. Simon, P. Vanmanshoven, D. Wagenaar et M. Georges, 2000. Extensive genome-wide linkage disequilibrium in cattle. Genome Res 10: 220-227.

Fernando R. L. et M. Grossman, 1989. Marker assisted selection using best linear unbiased prediction. Genet Sel Evol 21: 467-477.

Fisher R. A., 1918. The correlation between relatives on the supposition of Mendelian inheritance. Trans R Soc Edinb 52: 399-433.

Fortune J. E., G. M. Rivera, A. C. Evans et A. M. Turzillo, 2001. Differentiation of dominant versus subordinate follicles in cattle. Biol Reprod 65: 648-654.

Galina C. S. et A. Orihuela, 2007. The detection of estrus in cattle raised under tropical conditions:What we know and what we need to know. Horm Behav 52: 32–38.

Gautier M., 2003. Cartographie fine de QTL chez les bovins laitiers, pp. 284 in Laboratoire de Génétique Biochimique et de Cytogénétique - Département de Génétique Animale. AgroParisTech, Jouy-en-Josas.

Gautier M., A. Capitan, S. Fritz, A. Eggen, D. Boichard et T. Druet, 2007. Characterization of the DGAT1 K232A and variable number of tandem repeat polymorphisms in French dairy cattle. J Dairy Sci 90: 2980-2988.

Georges M., D. Nielsen, M. Mackinnon, A. Mishra, R. Okimoto, A. T. Pasquino, L. S. Sargeant, A. Sorensen, M. R. Steele, X. Zhao et al., 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics 139: 907-920.

Gilbert R. O., S. T. Shin, C. L. Guard, H. N. Erb et M. Frajblat, 2005. Prevalence of endometritis and its effects on reproductive performance of dairy cows. Theriogenology 64: 1879-1888.

Ginther O. J., L. Knopf et J. P. Kastelic, 1989. Temporal associations among ovarian events in cattle during oestrus cycles with two and three follicular waves. J Reprod Fertil 87: 223-230.

Gong J. G., 2002a. Influence of metabolic hormones and nutrition on ovarian follicle development in cattle: practical implications. Domest Anim Endocrinol 23: 229-241.

Gong J. G., W. J. Lee, P. C. Garnsworthy et R. Webb, 2002b. Effect of dietary-induced increases in circulating insulin concentrations during the early postpartum period on reproductive function in dairy cows. Reproduction 123: 419-427.

Gonzalez-Recio O. et R. Alenda, 2005. Genetic parameters for female fertility traits and a fertility index in Spanish dairy cattle. J Dairy Sci 88: 3282-3289.

Gonzalez-Recio O., R. Alenda, Y. M. Chang, K. A. Weigel et D. Gianola, 2006. Selection for female fertility using censored fertility traits and investigation of the relationship with milk production. J Dairy Sci 89: 4438-4444.

Green M. P., M. G. Hunter et G. E. Mann, 2005. Relationships between maternal hormone secretion and embryo development on day 5 of pregnancy in dairy cows. Anim Reprod Sci 88: 179-189.

Green P., 1992. Construction and comparison of chromosome 21 radiation hybrid and linkage maps using CRI-MAP. Cytogenet Cell Genet 59: 122-124.

Grisart B., W. Coppieters, F. Farnir, L. Karim, C. Ford, P. Berzi, N. Cambisano, M. Mni, S. Reid, P. Simon, R. Spelman, M. Georges et R. Snell, 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Res 12: 222-231.

Grisart B., F. Farnir, L. Karim, N. Cambisano, J. J. Kim, A. Kvasz, M. Mni, P. Simon, J. M. Frere, W. Coppieters et M. Georges, 2004. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. Proc Natl Acad Sci U S A 101: 2398-2403.

Guerrier J., L. Journaux, Y. M. Chatelin et H. Ledos, 2007. Durée de gestation des races bovines françaises. From http://217.167.235.86/html28/IMG/pdf/2007_03_Genetique_13_Guerrier.pdf.

Guillaume F., M. Gautier, S. Ben Jemaa, S. Fritz, A. Eggen, D. Boichard et T. Druet, 2007. Refinement of two female fertility QTL using alternative phenotypes in French Holstein dairy cattle. Anim Genet 38: 72-74.

Gurganus M. C., J. D. Fry, S. V. Nuzhdin, E. G. Pasyukova, R. F. Lyman et T. F. Mackay, 1998. Genotype-environment interaction at quantitative trait loci affecting sensory bristle number in Drosophila melanogaster. Genetics 149: 1883-1898.

Gutierrez C. G., J. G. Gong, T. A. Bramley et R. Webb, 2006. Selection on predicted breeding value for milk production delays ovulation independently of changes in follicular development, milk production and body weight. Anim Reprod Sci 95: 193-205.

Hageman W. H., 2nd, G. E. Shook et W. J. Tyler, 1991. Reproductive performance in genetic lines selected for high or average milk yield. J Dairy Sci 74: 4366-4376.

Haley C. S. et S. A. Knott, 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69: 315-324.

Hansen L. B., 2000. Consequences of selection for milk yield from a geneticist's viewpoint. J Dairy Sci 83: 1145-1150.

Hansen P. J., P. Soto et R. P. Natzke, 2004. Mastitis and fertility in cattle - possible involvement of inflammation or immune activation in embryonic mortality. Am J Reprod Immunol 51: 294-301.

Harder B., J. Bennewitz, D. Hinrichs et E. Kalm, 2006. Genetic parameters for health traits and their relationship to different persistency traits in German Holstein dairy cattle. J Dairy Sci 89: 3202-3212.

Harris B. L. et E. S. Kolver, 2001. Review of holsteinization on intensive pastoral dairy farming in New Zealand. J Dairy Sci 84 (E. Suppl.) E56-E61.

Harrison R. O., S. P. Ford, J. W. Young, A. J. Conley et A. E. Freeman, 1990. Increased milk production versus reproductive and energy status of high producing dairy cows. J Dairy Sci 73: 2749-2758.

Hayes B. et M. E. Goddard, 2001. The distribution of the effects of genes affecting quantitative traits in livestock. Genet Sel Evol 33: 209-229.

Humblot P., 2001. Use of pregnancy specific proteins and progesterone assays to monitor pregnancy and determine the timing, frequencies and sources of embryonic mortality in ruminants. Theriogenology 56: 1417-1433.

Humblot P., P. Holm, P. Lonergan, C. Wrenzycki, A. S. Lequarre, C. G. Joly, D. Herrmann, A. Lopes, D. Rizos, H. Niemann et H. Callesen, 2005. Effect of stage of follicular growth during superovulation on developmental competence of bovine oocytes. Theriogenology 63: 1149–1166.

Hyttel P., T. Fair, H. Callesen et T. Greve, 1997. Oocyte growth, capacitation and final maturation in cattle. Theriogenology 47: 23-32.

International commitee for animal recording (ICAR), 2006. International agreement of recording practices. Section 6: fertility recording: 211-213. From http://www.icar.org/Documents/Rules%20and%20regulations/Guidelines/Guidelines_2007.pdf

Ihara N., A. Takasuga, K. Mizoshita, H. Takeda, M. Sugimoto, Y. Mizoguchi, T. Hirano, T. Itoh, T. Watanabe, K. M. Reed, W. M. Snelling, S. M. Kappes, C. W. Beattie, G. L. Bennett et Y. Sugimoto, 2004. A comprehensive genetic map of the cattle genome based on 3802 microsatellites. Genome Res 14: 1987-1998.

Inskeep E. K. et R. A. Dailey, 2005. Embryonic death in cattle. Vet Clin North Am Food Anim Pract 21: 437-461.

Institut de l'élevage, 2009. Résultats des contrôles de descendance des taureaux Prim'holstein production laitière – morphologie caractères fonctionnels. From http://www.inst-elevage.asso.fr/html1/IMG/pdf_INTRO__PH__091_Fra_.pdf

INTERBULL, 2008. Interbull routine genetic evaluation for female fertility traits. From http://www.interbull.org/Female_fert/framesida-fert.htm.

Jorritsma R., H. Jorritsma, Y. H. Schukken et G. H. Wentink, 2000. Relationships between fatty liver and fertility and some periparturient diseases in commercial Dutch dairy herds. Theriogenology 54: 1065-1074.

Kadarmideen H. N., R. Thompson, P. C. Michael et M. A. Kossaibati, 2003. Genetic parameters and evaluations from single- and multiple-trait analysis of dairy cow fertility and milk production. Livest Prod Sci 81: 183–195.

Kadarmideen H. N., R. Thompson et G. Simm, 2001. Linear and threshold model genetic parameters for disease, fertility and milk production in dairy cattle. Anim Sci 71: 411–419.

Kallio M., Y. Chang, M. Manuel, T. P. Alastalo, M. Rallu, Y. Gitton, L. Pirkkala, M. T. Loones, L. Paslaru, S. Larney, S. Hiard, M. Morange, L. Sistonen et V. Mezger, 2002. Brain abnormalities, defective meiotic chromosome synapsis and female subfertility in HSF2 null mice. Embo J 21: 2591-2601.

Khatkar M. S., F. W. Nicholas, A. R. Collins, K. R. Zenger, J. A. Cavanagh, W. Barris, R. D. Schnabel, J. F. Taylor et H. W. Raadsma, 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics 9: 187.

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

Kuhn C., G. Thaller, A. Winter, O. R. Bininda-Emonds, B. Kaupe, G. Erhardt, J. Bennewitz, M. Schwerin et R. Fries, 2004. Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with major effect on milk fat content in cattle. Genetics 167: 1873-1881.

Lai Y. et F. Sun, 2003. The relationship between microsatellite slippage mutation rate and the number of repeat units. Mol Biol Evol 20: 2123-2131.

Lamming G. E. et A. O. Darwash, 1998. The use of milk progesterone profiles to characterise components of subfertility in milked dairy cows. Anim Reprod Sci 52: 175-190.

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

Larroque H., Y. Gallard, H. Thaunat, D. Boichard et J. J. Colleau, 2002 A crossbreeding experiment to quantitative trait loci in dairy cattle.,. In 7th world congress on genetics applied to livestock production, Montpellier, France.

Lawrence, S. and N. Morton (1992). "Physical mapping by multiple pairwise analysis." Cytogenet Cell Genet 59(2-3): 107-9.

Le Mézec P., 2007 Un répit dans la baisse de la fertilité., pp. 15-18 in BTiA: La revue Française de génétique et de reproduction.

Le Mézec P., 2009. Le cheptel laitier Français: évolution génétique et phénotypique 1988-2008 -prévision d'évolution génétique 2008-2014. From http://www.inst-elevage.asso.fr/html1/IMG/pdf_CR_010972007.pdf.

Leips J. et T. F. Mackay, 2000. Quantitative trait loci for life span in Drosophila melanogaster: interactions with genetic background and larval density. Genetics 155: 1773-1788.

Leroy J. L., G. Opsomer, A. Van Soom, I. G. Goovaerts et P. E. Bols, 2008. Reduced fertility in high-yielding dairy cows: are the oocyte and embryo in danger? Part I. The importance of negative energy balance and altered corpus luteum function to the reduction of oocyte and embryo quality in high-yielding dairy cows. Reprod Domest Anim 43: 612-622.

Lewis G. S., 1997. Uterine health and disorders. J Dairy Sci 80: 984-994.

Lewontin R. C., 1964. The Interaction of selection and linkage. I. General considerations; heterotic models. Genetics 49: 49-67.

Lindsey B. R., L. J. Maclellan, T. R. Whyte, J. E. Kinder et M. J. D'occhio, 2002. Differential requirement for pulsatile LH during the follicular phase and exposure to the preovulatory LH surge for oocyte fertilization and embryo development in cattle. Theriogenology 58: 1651-1662.

Lopez H., L. D. Satter et M. C. Wiltbank, 2004. Relationship between level of milk production and estrous behavior of lactating dairy cows. Anim Reprod Sci 81: 209-223.

Lucy M. C., 2001. Reproductive loss in high-producing dairy cattle: where will it end? J Dairy Sci 84: 1277-1293.

Lucy M. C., 2003. Mechanisms linking nutrition and reproduction in postpartum cows. Reprod Suppl 61: 415-427.

Lucy M. C., 2007. Fertility in high-producing dairy cows: reasons for decline and corrective strategies for sustainable improvement. Soc Reprod Fertil Suppl 64: 237-254.

Lund M. S., P. Sorensen, B. Guldbrandtsen et D. A. Sorensen, 2003. Multitrait fine mapping of quantitative trait loci using combined linkage disequilibria and linkage analysis. Genetics 163: 405-410.

Mackay T. F., 2001. The genetic architecture of quantitative traits. Annu Rev Genet 35: 303-339.

Macmillan K. L., I. J. Lean et C. T. Westwood, 1996. The effects of lactation on the fertility of dairy cows. Aust Vet J 73: 141-147.

Macmillan K. L., V. K. Taufa et A. M. Day, 1986. Effects of an agonist of gonadotrophin releasing hormone (Buserelin) in cattle. III. Pregnancy rates after a post-insemination injection during metoestrus or dioestrus. Anim Reprod Sci 11: 1-10.

Madelin V., 2006. Stabilité de la production bovine attendue pour 2006 et 2007. Agreste Conjoncture - Productions animales – Bovins. From http://agreste.agriculture.gouv.fr/IMG/pdf/bovin0607note.pdf.

Mann G. E., S. J. Mann, D. Blache et R. Webb, 2005. Metabolic variables and plasma leptin concentrations in dairy cows exhibiting reproductive cycle abnormalities identified through milk progesterone monitoring during the post partum period. Anim Reprod Sci 88: 191-202.

Mcdougall S., D. Blache et F. M. Rhodes, 2005. Factors affecting conception and expression of oestrus in anoestrous cows treated with progesterone and oestradiol benzoate. Anim Reprod Sci 88: 203-214.

Mcevoy T. G., J. J. Robinson, R. P. Aitken, P. A. Findlay, R. M. Palmer et I. S. Robertson, 1995. Dietary-induced suppression of pre-ovulatory progesterone concentrations in

superovulated ewes impairs the subsequent in vivo and in vitro development of their ova. Anim Reprod Sci 39: 89-107.

Mcrac A. F., J. C. Mccwan, K. G. Dodds, T. Wilson, A. M. Crawford et J. Slatc, 2002. Linkage disequilibrium in domestic sheep. Genetics 160: 1113-1122.

Meuwissen T. H., B. J. Hayes et M. E. Goddard, 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819-1829.

Meuwissen T. H., A. Karlsen, S. Lien, I. Olsaker et M. E. Goddard, 2002. Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. Genetics 161: 373-379.

Minery S., M. Brochard, et S. Mattalia, 2008. Use of national and international EBVs of fertility in total merit index. From http://www-interbull.slu.se/bulletins/bulletin38/31_Minery.pdf.

Mussard M. L., C. R. Burkc, E. J. Bchlkc, C. L. Gasscr et M. L. Day, 2007. Influcncc of premature induction of a luteinizing hormone surge with gonadotropin-releasing hormone on ovulation, luteal function, and fertility in cattle. J Anim Sci 85: 937-943.

Nebel R. L. et M. L. Mcgilliard, 1993. Interactions of high milk yield and reproductive performance in dairy cows. J Dairy Sci 76: 3257-3268.

Neimann-Sorensen A. et A. Robertson, 1961. The association between blood groups and several production characters in three Danish cattle breeds. Acta Agric Scand 11: 163-196.

Nielen M., Y. H. Schukken, D. T. Scholl, H. J. Wilbrink et A. Brand, 1989. Twinning in dairy cattle: A study of risk factors and effects. Theriogenology 32: 845-862.

Paterson A. H., S. Damon, J. D. Hewitt, D. Zamir, H. D. Rabinowitch, S. E. Lincoln, E. S. Lander et S. D. Tanksley, 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. Genetics 127: 181-197.

Perry G. A., M. F. Smith, M. C. Lucy, J. A. Green, T. E. Parks, M. D. Macneil, A. J. Roberts et T. W. Geary, 2005. Relationship between follicle size at insemination and pregnancy success. Proc Natl Acad Sci U S A 102: 5268-5273.

Perry G. A., M. F. Smith, A. J. Roberts, M. D. Macneil et T. W. Geary, 2007. Relationship between size of the ovulatory follicle and pregnancy success in beef heifers. J Anim Sci 85: 684-689.

Petersson K. J., H. Gustafsson, E. Strandberg et B. Berglund, 2006. Atypical progesterone profiles and fertility in Swedish dairy cows. J Dairy Sci 89: 2529-2538.

Petersson K. J., E. Strandberg, H. Gustafsson, M. D. Royal et B. Berglund, 2008. Detection of delayed cyclicity in dairy cows based on progesterone content in monthly milk samples. Prev Vet Med 86: 153-163.

Philipsson J., 1981. Aspects genétiques de la fertilité des femelles chez bovins laitiers. Livest Prod Sci: 307-319.

Pritchard J. Y., F. N. Schrick et E. K. Inskeep, 1994. Relationship of pregnancy rate to peripheral concentrations of progesterone and estradiol in beef cows. Theriogenology 42: 247-259.

Pryce J. E., M. P. Coffey et G. Simm, 2001. The relationship between body condition score and reproductive performance. J Dairy Sci 84: 1508-1515.

Pryce J. E., B. L. Nielsen, R. F. Veerkamp et G. Simm, 1999. Genotype and feeding system effects and interactions for health and fertility traits in dairy cattle. Livest Prod Sci 57: 193-201.

Pryce J. E., M. D. Royal, P. C. Garnsworthy et I. L. Mao, 2004. Fertility in the high-producing dairy cow. Livest Prod Sci 86: 125-135.

Rizos D., F. Ward, P. Duffy, M. P. Boland et P. Lonergan, 2002. Consequences of bovine oocyte maturation, fertilization or early embryo development in vitro versus in vivo: implications for blastocyst yield and blastocyst quality. Mol Reprod Dev 61: 234-248.

Roche J. F., D. Mackey et M. D. Diskin, 2000. Reproductive management of postpartum cows. Anim Reprod Sci 60-61: 703-712.

Ron M. et J. I. Weller, 2007. From QTL to QTN identification in livestock--winning by points rather than knock-out: a review. Anim Genet 38: 429-439.

Royal M. D., J. E. Pryce, J. A. Woolliams et A. P. Flint, 2002. The genetic relationship between commencement of luteal activity and calving interval, body condition score, production, and linear type traits in Holstein-Friesian dairy cattle. J Dairy Sci 85: 3071-3080.

Rukkwamsuk T., T. Wensing et T. A. Kruip, 1999. Relationship between triacylglycerol concentration in the liver and first ovulation in postpartum dairy cows. Theriogenology 51: 1133-1142.

Ryan D. P. et M. P. Boland, 1991. Frequency of twin births among Holstein-Friesian cows in a warm dry climate. Theriogenology 36: 1-10.

Santos J. E., W. W. Thatcher, R. C. Chebel, R. L. Cerri et K. N. Galvao, 2004. The effect of embryonic death rates in cattle on the efficacy of estrus synchronization programs. Anim Reprod Sci 82-83: 513-535.

Sartori R., R. Sartor-Bergfelt, S. A. Mertens, J. N. Guenther, J. J. Parrish et M. C. Wiltbank, 2002. Fertilization and early embryonic development in heifers and lactating cows in summer and lactating and dry cows in winter. J Dairy Sci 85: 2803-2812.

Savio J. D., L. Keenan, M. P. Boland et J. F. Roche, 1988. Pattern of growth of dominant follicles during the oestrous cycle of heifers. J Reprod Fertil 83: 663-671.

Schiex T. et C. Gaspin, 1997. CARTHAGENE: constructing and joining maximum likelihood genetic maps. Proc Int Conf Intell Syst Mol Biol 5: 258-267.

Schrooten C., H. Bovenhuis, W. Coppieters et J. A. Van Arendonk, 2000. Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. J Dairy Sci 83: 795-806.

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

Seegers H. et B. Grimard, 2003 La performance de reproduction d'un troupeau laitier., pp. 5-9 in BTiA: La Revue Française de Génétique et de Reproduction.

Shrimpton A. E. et A. Robertson, 1988. The Isolation of polygenic factors controlling bristle score in Drosophila melanogaster. I. Allocation of third chromosome sternopleural bristle effects to chromosome sections. Genetics 118: 437-443.

Silke V., M. G. Diskin, D. A. Kenny, M. P. Boland, P. Dillon, J. F. Mee et J. M. Sreenan, 2002. Extent, pattern and factors associated with late embryonic loss in dairy cows. Anim Reprod Sci 71: 1-12.

Smith C., 1967. Improvement of metric traits through specific genetic loci. Anim Prod 9: 349-358.

Smith J. L., B. R. Kupchak, I. Garitaonandia, L. K. Hoang, A. S. Maina, L. M. Regalla et T. J. Lyons, 2008. Heterologous expression of human mPRalpha, mPRbeta and mPRgamma in yeast confirms their ability to function as membrane progesterone receptors. Steroids 73: 1160-1173.

Sreenan J. M. et M. G. Diskin, 1986. The extent and timing of embryonic mortality in cattle. In: Sreenan JM, Diskin MG (eds), Embryonic mortality in farm animals. Martinus Nijhoff, CEC, Bruxelles, Belgique, pp. 142–158.

Sutton M. L., R. B. Gilchrist et J. G. Thompson, 2003. Effects of in-vivo and in-vitro environments on the metabolism of the cumulus-oocyte complex and its influence on oocyte developmental capacity. Hum Reprod Update 9: 35-48.

Taylor C. C. et R. Rajamahendran, 1991. Follicular dynamics, corpus luteum growth and regression in lactating dairy cattle. Can J Anim Sci 71: 61-68.

Thatcher W. W., T. R. Bilby, J. A. Bartolome, F. Silvestre, C. R. Staples et J. E. Santos, 2006.

Strategies for improving fertility in the modern dairy cow. Theriogenology 65: 30-44.

Thatcher W. W., K. L. Macmillan, P. J. Hansen et M. Drost, 1989. Concepts for regulation of corpus luteum function by the conceptus and ovarian follicles to improve fertility. Theriogenology 31: 149-164.

Townson D. H., P. C. Tsang, W. R. Butler, M. Frajblat, L. C. Griel, Jr., C. J. Johnson, R. A. Milvae, G. M. Niksic et J. L. Pate, 2002. Relationship of fertility to ovarian follicular waves before breeding in dairy cows. J Anim Sci 80: 1053-1058.

Van Der Lende T., L. M. Kaal, R. M. Roelofs, R. F. Veerkamp, C. Schrooten et H. Bovenhuis, 2004. Infrequent milk progesterone measurements in daughters enable bull selection for cow fertility. J Dairy Sci 87: 3953-3957.

Van Laere A. S., M. Nguyen, M. Braunschweig, C. Nezer, C. Collette, L. Moreau, A. L. Archibald, C. S. Haley, N. Buys, M. Tally, G. Andersson, M. Georges et L. Andersson, 2003. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. Nature 425: 832-836.

Vanraden P. M., A. H. Sanders, M. E. Tooker, R. H. Miller, H. D. Norman, M. T. Kuhn et G. R. Wiggans, 2004. Development of a national genetic evaluation for cow fertility. J Dairy Sci 87: 2285-2292.

Veerkamp R. F., B. Beerda et T. Van Der Lende, 2003. Effects of genetic selection for milk yield on energy balance, levels of hormones, and metabolites in lactating cattle, and possible links to reduced fertility. Livest Prod Sci 83: 257-275.

Veerkamp R. F., J. K. Oldenbroek et T. Van Der Lende, 1998 The use of milk progesterone measurements for genetic improvement of fertility traits in dairy cattle., pp. 62–67. in Proc. Int. Worksh. on Genetic Improvement of Functional Traits in Cattle; Fertility and Reproduction. Interbull, Grub, Germany.

Veerkamp R. F., G. Simm et J. D. Oldham, 1994. Effects of interaction between genotype and feeding system on milk production, feed intake, efficiency and body tissue mobilization in dairy cows. Livest Prod Sci 39: 229-241.

Vieira C., E. G. Pasyukova, Z. B. Zeng, J. B. Hackett, R. F. Lyman et T. F. Mackay, 2000. Genotype-environment interaction for quantitative trait loci affecting life span in Drosophila melanogaster. Genetics 154: 213-227.

Visscher P. M., R. Thompson et C. S. Haley, 1996. Confidence intervals in QTL mapping by bootstrapping. Genetics 143: 1013-1020.

Wall E., S. Brotherstone, J. A. Woolliams, G. Banos et M. P. Coffey, 2003. Genetic evaluation of fertility using direct and correlated traits. J Dairy Sci 86: 4093-4102.

Wathes D. C., M. Fenwick, Z. Cheng, N. Bourne, S. Llewellyn, D. G. Morris, D. Kenny, J. Murphy et R. Fitzpatrick, 2007. Influence of negative energy balance on cyclicity and fertility in the high producing dairy cow. Theriogenology 68 Suppl 1: S232-241.

Webb R., P. C. Garnsworthy, J. G. Gong et D. G. Armstrong, 2004. Control of follicular growth: local interactions and nutritional influences. J Anim Sci 82 E-Suppl: E63-74.

Webb R., K. J. Woad et D. G. Armstrong, 2002. Corpus luteum (CL) function: local control mechanisms. Domest Anim Endocrinol 23: 277-285.

Weckx S., J. Del-Favero, R. Rademakers, L. Claes, M. Cruts, P. De Jonghe, C. Van Broeckhoven et P. De Rijk, 2005. novoSNP, a novel computational tool for sequence variation discovery. Genome Res 15: 436-442.

Weigel K. A. et R. Rekaya, 2000. Genetic parameters for reproductive traits of Holstein cattle in California and Minnesota. J Dairy Sci 83: 1072-1080.

Weller J. I., M. Golik, E. Seroussi, E. Ezra et M. Ron, 2003. Population-wide analysis of a QTL affecting milk-fat production in the Israeli Holstein population. J Dairy Sci 86: 2219-2227.

Williams J. L., A. Eggen, L. Ferretti, C. J. Farr, M. Gautier, G. Amati, G. Ball, T. Caramorr, R. Critcher, S. Costa, P. Hextall, D. Hills, A. Jeulin, S. L. Kiguwa, O. Ross, A. L. Smith, K. Saunier, B. Urquhart et D. Waddington, 2002. A bovine whole-genome radiation hybrid panel and outline map. Mamm Genome 13: 469-474.

Wiltbank M., L. Lopez, R. Sartori et A. Gument, 2005 : The underlying physiology changing reproduction in lactating dairy cows. In Mid-South Ruminant Nutrition Conference, Arlington, Texas.

Winter A., W. Kramer, F. A. Werner, S. Kollers, S. Kata, G. Durstewitz, J. Buitkamp, J. E. Womack, G. Thaller et R. Fries, 2002. Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. Proc Natl Acad Sci U S A 99: 9300-9305.

Xiao X., X. Zuo, A. A. Davis, D. R. Mcmillan, B. B. Curry, J. A. Richardson et I. J. Benjamin, 1999. HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice. Embo J 18: 5943-5952.

Yang G. C., J. Ren, Y. M. Guo, N. S. Ding, C. Y. Chen et L. S. Huang, 2006. Genetic evidence for the origin of an IGF2 quantitative trait nucleotide in Chinese pigs. Anim Genet 37: 179-180.

Zwald N. R., K. A. Weigel, Y. M. Chang, R. D. Welper et J. S. Clay, 2004. Genetic selection for health traits using producer-recorded data. II. Genetic correlations, disease probabilities, and relationships with existing traits. J Dairy Sci 87: 4295-4302.

Annexes

Annexe 1

Multi3A

Composition en marqueurs :

Nom marqueur	Taille Inférieure	Taille Supérieure	Fluorescence	Quantité d'amorces (20 pM pour 1 individu)	Chromosome
BM4129	72	102	FAM	0,04	BTA03
BMS2522	129	167	TET	0,04	BTA03
BMS963	134	154	FAM	0,02	BTA03
DIK2434	244	278	FAM	0,06	BTA03
DIK2609	188	192	FAM	0,02	BTA03
DIK4196	187	212	TET	0,04	BTA03
ILSTS029	156	164	HEX	0,02	BTA03
RM019	85	91	HEX	0,02	BTA03

Quantité d'eau ajoutée pour 10 individus : 37,4 µl.

Marqueurs éliminés :

- BMS2904 : éliminé de ce multiplex puis réintégré dans le multiplex MultiBTA03_B
- MNS-74 : éliminé lors de la mise au point.
- BMS2790 : éliminé lors de la mise au point
- Nombre de cycles : 30

Conditions PCR :

- Température d'hybridation : 60°C
- Temps d'hybridation : 30 secondes
- Temps d'extension : 30 secondes
- Temps d'élongation : 30 secondes

Dépôt sur gel :

- ABI prism 377 : 8 µl de mélange (5 µl de produit PCR + 3 µl de bleu de charge)
- MegaBace 1000

Multi3B

Composition en marqueurs :

Nom marqueur	Taille Inférieure	Taille Supérieure	Fluorescence	Quantité d'amorces (20 pM pour 1 individu)	Chromosome
INRA006 (MB101)	106	120	FAM	0,04	BTA03
DIK2101	195	213	TET	0,03	BTA03
MNB-65	162	184	FAM	0,04	BTA03
BMS482	137	157	HEX	0,04	BTA03
MNB-86	178	217	HEX	0,045	BTA03
BL41	236	258	TET	0,02	BTA03
INRA003 (MB099)	96	120	HEX	0,045	BTA03

Quantité d'eau ajoutée pour 10 individus : 38,6 µl.

Marqueurs éliminés :

- CA065 : éliminé car monomorphe.
- DIK4403 : éliminé lors de la mise au point

Conditions PCR :

- Température d'hybridation : 60°C
- Temps d'hybridation : 30 secondes
- Temps d'extension : 30 secondes
- Temps d'élongation : 30 secondes
- Nombre de cycles : 30

Dépôt sur gel :

- ABI prism 377 : 8 µl de mélange (5 µl de produit PCR + 3 µl de bleu de charge)
- MegaBace 1000

MultiBTA03_B

Composition en marqueurs :

Nom marqueur	Taille Inférieure	Taille Supérieure	Fluorescence	Chromosome
BMS2904	83	91	FAM	BTA03
DIK069	194	197	HEX	BTA03
DIK4193	199	205	FAM	BTA03
INRA3012	252	279	FAM	BTA03
INRA3013	255	270	TET	BTA03
INRA3015	250	265	HEX	BTA03

Conditions PCR :

- Température d'hybridation : 60°C
- Temps d'hybridation : 30 secondes
- Temps d'extension : 30 secondes
- Temps d'élongation : 30 secondes
- Nombre de cycles : 30

Marqueurs éliminés :

- INRA3020 : éliminé lors de la mise au point
- INRA3022 : éliminé lors de la vérification des génotypes car monomorphe.
- INRA3029 : éliminé lors de la mise au point

Dépôt sur gel :

- ABI prism 377 : 8 µl de mélange (5 µl de produit PCR + 3 µl de bleu de charge)
- MegaBace 1000

Multiplex A

Composition en marqueurs :

Nom marqueur	TailleInf	TailleSup	Fluorescence	Quantité amorces (20 pM pour 1 individu)	Chromosome
BMS4028	92	120	FAM	0,025	BTA01
DIK4587	206	224	FAM	0,025	BTA01
BMS4000	117	147	HEX	0,020	BTA01
DIK2886	179	197	HEX	0,020	BTA01
BM864	214	274	HEX	0,030	BTA01
BM8246	167	197	TET	0,020	BTA01
BMS4041	224	230	TET	0,025	BTA01

Quantité d'eau ajoutée pour 10 individus : 49,67 µl.

Marqueurs éliminés : aucun marqueur éliminé

Conditions PCR :

- Température d'hybridation : 60°C
- Temps d'hybridation : 30 secondes
- Temps d'extension : 30 secondes
- Temps d'élongation : 30 secondes
- Nombre de cycles : 30

Dépôt sur gel :

- ABI prism 377 : 8 µl de mélange (5 µl de produit PCR + 3 µl de bleu de charge)
- MegaBace 1000

Multiplex B

Composition en marqueurs :

Nom marqueur	Taille Inférieure	Taille Supérieure	Fluorescence	Quantité d'amorces (20 pM pour 1 individu)	Chromosome
DIK4208	156	168	FAM	0,015	BTA02
BMS2519	99	127	HEX	0,03	BTA02
BM1223	163	177	HEX	0,02	BTA02
DIK4972	222	252	HEX	0,015	BTA02
RM041	74	94	TET	0,05	BTA02
IDVGA-2	119	147	TET	0,015	BTA02

Quantité d'eau ajoutée pour 10 individus : 49,67 µl.

Marqueurs éliminés :

- DIK4618 : éliminé car les génotypes avaient des profils illisibles.

Conditions PCR :

- Température d'hybridation : 57°C
- Temps d'hybridation : 30 secondes
- Temps d'extension : 30 secondes
- Temps d'élongation : 30 secondes

Dépôt sur gel :

- ABI prism 377 : 8 µl de mélange (5 µl de produit PCR + 3 µl de bleu de charge)
- MegaBace 1000

Multiplex C

Composition en marqueurs :

Nom marqueur	Taille Inférieure	Taille Supérieure	Fluorescence	Quantité d'amorces (20 pM pour 1 individu)	Chromosome
BMS1117	83	103	FAM	0,03	BTA21
BMS1866	134	158	FAM	0,01	BTA02
DIK2492	211	235	HEX	0,015	BTA21
BMS4030	162	180	TET	0,015	BTA01

Quantité d'eau ajoutée pour 10 individus : 49,82 µl.

Marqueurs éliminés :

- BMS2024 : éliminé lors de la mise au point pour manque d'amplification.
- BZ847519 : éliminé lors du génotypage car il est monomorphe.
- INRA3003 éliminé lors du génotypage car il est monomorphe.
- DIK2492 : éliminé lors du génotypage car il a un profil illisible.

Conditions PCR :

- Température d'hybridation : 57°C
- Temps d'hybridation : 30 secondes
- Temps d'extension : 30 secondes
- Temps d'élongation : 30 secondes

Marqueur éliminé : INRA103 (BTA21)

Dépôt sur gel :

- ABI prism 377 : 8 µl de mélange (5 µl de produit PCR + 3 µl de bleu de charge)
- MegaBace 1000

Multiplex D

Composition en marqueurs :

Nom marqueur	Taille Inférieure	Taille Supérieure	Fluorescence	Quantité d'amorces (20 pM pour 1 individu)	Chromosome
DVEPC32	162	176	FAM	0,02	BTA01
INRA3001	121	121	HEX	0,02	BTA03
CC582410	233	233	TET	0,015	BTA03
URB006	161	173	TET	0,02	BTA03

Quantité d'eau ajoutée pour 10 individus : 49,85 µl.

Marqueurs éliminés :
- INRA3002 : éliminé lors de la mise au point
- INRA103 : éliminé lors de la mise au point.

Conditions PCR :
- Température d'hybridation : 60°C
- Temps d'hybridation : 30 secondes
- Temps d'extension : 30 secondes
- Temps d'élongation : 30 secondes

Dépôt sur gel :
- ABI prism 377 : 8 µl de mélange (5 µl de produit PCR + 3 µl de bleu de charge)
- MegaBace 1000

Multiplex BTA02

Composition en marqueurs :

Nom marqueur	Taille Inférieure	Taille Supérieure	Fluorescence	Quantité d'amorces (20 pM pour 1 individu)	Chromosome
BM4440	100	130	FAM	0,02	BTA02
DIK2719	138	156	FAM	0,02	BTA02
DIK1081	84	108	TET	0,02	BTA02
DIK4673	207	244	TET	0,02	BTA02
DIK2496	177	184	HEX	0,04	BTA02
DIK4334	233	244	HEX	0,03	BTA02

Quantité d'eau ajoutée pour 10 individus : 47 µl.

Marqueurs éliminés : aucun marqueur éliminé.

Conditions PCR :

- Température d'hybridation : 58°C
- Temps d'hybridation : 30 secondes
- Temps d'extension : 30 secondes
- Temps d'élongation : 30 secondes

Dépôt sur gel :

- ABI prism 377 : 8 µl de mélange (5 µl de produit PCR + 3 µl de bleu de charge)
- MegaBace 1000 :

Annexe 2

Tableau 1 : *Répartition par père du nombre de femelles accouplées et de fils dans les 47 familles Prim'Holstein du dispositif de l'étude de primo-localisation. Les 9 pères faisant partie du programme QTL initial sont en gras italique.*

Numéro de famille	Identifiant père	Nombre de femelles	Nombre de fils
1	1	58	61
2	179	32	34
3	278	65	75
4	484	98	113
5	794	72	83
6	***1022***	***56***	***65***
7	1200	48	54
8	1351	43	51
9	1489	20	21
10	1551	89	99
11	1829	68	70
12	2036	37	52
13	2163	22	22
14	2230	51	66
15	2399	33	44
16	2510	30	36
17	2607	58	64
18	2788	36	39
19	2900	44	53
20	3042	31	33
21	3138	41	46
22	***3267***	***107***	***136***
23	3618	51	60
24	***3781***	***51***	***67***
25	***3951***	***144***	***194***
26	***4434***	***73***	***96***
27	4677	28	42
28	4776	63	74
29	***4977***	***170***	***239***
30	***5557***	***74***	***99***
31	***5805***	***65***	***83***
32	***6019***	***140***	***173***
33	6473	26	30
34	6556	38	46
35	6679	104	121
36	7009	45	59
37	7159	59	67
38	7345	41	52
39	7480	51	67
40	7650	31	39
41	7752	21	28
42	7823	20	23
43	7887	24	28
44	7964	90	103
45	8248	43	55
46	8390	31	38
47	8491	70	75

Tableau 2 : *Répartition par père du nombre de femelles accouplées et de fils dans les 18 familles Normandes du dispositif de l'étude de primo-localisation.* Les 3 pères faisant partie du programme QTL initial sont en gras italique.

Numéro de famille	Père	Nombre de femelles	Nombre de fils
1	1	33	35
2	103	32	36
3	204	58	65
4	***386***	***55***	***59***
5	556	71	75
6	774	25	26
7	851	65	69
8	1051	67	72
9	1258	56	58
10	1429	41	44
11	1556	32	36
12	1657	58	66
13	1840	60	66
14	***2027***	***69***	***77***
15	2243	36	37
16	***2353***	***70***	***73***
17	2567	54	58
18	2734	50	54

Tableau 3 : *Répartition par père du nombre de femelles accouplées et de fils dans les 13 familles Monbéliardes du dispositif de l'étude de primo-localisation. Les 2 pères faisant partie du programme QTL initial sont en gras italique.*

Numéro de famille	Père	Nombre de femelles	Nombre de fils
1	1	62	66
2	192	39	40
3	311	68	72
4	520	23	25
5	592	23	25
6	***664***	***88***	***96***
7	937	38	41
8	1055	45	46
9	1192	87	98
10	1465	59	67
11	***1651***	***69***	***78***
12	1868	38	44
13	1989	61	64

Annexe 3

Tableau 1 : *Répartition, par père, du nombre de femelles accouplées et de fils dans les 41 familles du dispositif de l'étude de confirmation. Les 14 pères faisant partie du programme QTL initial sont en gras italique.*

Race	Numéro	Pere	Nombre de femelles	Nombre de fils
Prim'Holstein	1	2108	48	58
Prim'Holstein	2	3442	35	37
Prim'Holstein	3	3465	41	47
Prim'Holstein	***4***	***3468***	***108***	***138***
Prim'Holstein	***5***	***3480***	***52***	***68***
Prim'Holstein	***6***	***3481***	***142***	***187***
Prim'Holstein	***7***	***3487***	***72***	***95***
Prim'Holstein	8	3492	28	41
Prim'Holstein	***9***	***3498***	***169***	***237***
Prim'Holstein	***10***	***3500***	***76***	***94***
Prim'Holstein	***11***	***3501***	***63***	***83***
Prim'Holstein	***12***	***3502***	***135***	***168***
Prim'Holstein	13	3514	38	49
Prim'Holstein	14	4442	72	84
Prim'Holstein	***15***	***4729***	***56***	***66***
Prim'Holstein	16	5245	69	75
Prim'Holstein	17	5970	88	98
Prim'Holstein	18	5972	68	70
Prim'Holstein	19	5974	37	51
Prim'Holstein	20	6109	30	36
Prim'Holstein	21	6143	103	104
Prim'Holstein	22	6145	45	59
Prim'Holstein	23	6146	58	66
Prim'Holstein	24	6151	51	68
Prim'Holstein	25	6865	48	54
Prim'Holstein	26	7210	43	51
Normande	1	1406	58	65
Normande	***2***	***1528***	***55***	***59***
Normande	***3***	***1537***	***72***	***76***
Normande	4	1611	62	66
Normande	***5***	***1788***	***70***	***79***
Normande	6	1855	50	51
Normande	7	1961	38	40
Normande	8	2409	64	68
Normande	9	2618	58	66
Montbéliarde	***1***	***1103***	***93***	***101***
Montbéliarde	2	1298	60	68
Montbéliarde	***3***	***1312***	***72***	***82***
Montbéliarde	4	1368	62	65
Montbéliarde	5	1531	61	65
Montbéliarde	6	1736	38	41

Tableau 2 : *Répartition, par père, du nombre de femelles accouplées et de fils dans les 17 familles Prim'Holstein du dispositif de l'analyse LDLA avec les 437 SNP (Druet et al., 2008).*

Numéro	Pere	Nombre de femelles	Nombre de fils
1	2108	48	58
2	3442	35	37
3	3465	41	47
4	3492	28	41
5	3514	38	49
6	4442	72	84
7	5245	69	75
8	5970	88	98
9	5972	68	70
10	5974	37	51
11	6109	30	36
12	6143	103	104
13	6145	45	59
14	6146	58	66
15	6151	51	68
16	6865	48	54
17	7210	43	51

Annexe 3b

Cartes génétiques des 12 régions génomiques SAM utilisées lors de l'étude de primo-localisation.

Chromosome	Marqueur	Distance génétique (carte USDA) (cM)
BTA01	BM4307	30
BTA01	BM1312	50
BTA01	INRA073	68
BTA01	BM1824	100
BTA02	BMS778	0
BTA02	BM2113	30
BTA03	BMS871	0
BTA03	ILSTS096	30
BTA03	INRA023	60
BTA03	HUJI177	84
BTA06	BM143	59
BTA06	DIK82	64
BTA06	INRAK	95
BTA07	INRABERN192	84
BTA07	ILSTS006	114
BTA07	INRA053	128
BTA14	ILSTS039	0
BTA14	CSSM066	5
BTA15	SPS115	8
BTA15	INRA224	14
BTA15	BMS2684	41
BTA19	CSSM065	85
BTA19	BMS1069	93
BTA19	ETH3	105
BTA19	BMC1013	125
BTA20	BM1225	6
BTA20	TGLA126	33
BTA20	DIK2695	44
BTA20	BM4107	50
BTA20	UWCA026	78
BTA21	ILSTS103	52
BTA21	TGLA122	80
BTA21	CSSM018	100
BTA23	BOLA	41
BTA23	BM7233	57
BTA26	BMS907	12
BTA26	HEL11	21
BTA26	INRA081	27
BTA26	RM026	40

Annexe 4

Valeurs des TNR90 et des CD des 16 fils présélectionnés à partir de quatre pères hétérozygotes.

ID_2005	ID_Pere_2005	Classe	TNR90	CD_90j
2245	3536	-/-	-0,0701	0,6857
2882	3536	-/-	-0,0289	0,6292
1518	3537	-/-	-0,0557	0,592
4088	4373	-/-	-0,0414	0,5681
4108	4373	-/-	-0,0683	0,6046
4143	4373	-/-	-0,0806	0,6166
4418	4582	-/-	-0,0274	0,9813
4423	4582	-/-	-0,0469	0,9716
1963	3536	+/+	0,0275	0,6556
2238	3536	+/+	0,0281	0,6586
2249	3537	+/+	0,066	0,6426
4093	4373	+/+	0,0295	0,656
4136	4373	+/+	0,0308	0,6494
4146	4373	+/+	0,0465	0,6666
4422	4582	+/+	0,0007	0,6815
4441	4582	+/+	0,0419	0,9793

Annexe 5

BTA01 :

Tableau 1 : *Effets de substitution allélique, erreur standard et valeur du test-T pour chacun des 26 pères Prim'Holstein de l'étude de cartographie à moyenne résolution pour le TNR90 sur le BTA01. Les pères hétérozygotes (P<0,05) sont en gras et en couleur rouge.*

Identifiant_père	Effet de substitution allélique	Erreur standard	Valeur du test-T
2108	0,0052	0,010194	0,5074
3442	**0,0311**	**0,012037**	**2,5875**
3487	-0,0027	0,007355	0,3727
3492	0,0121	0,010806	1,1224
3468	0,0064	0,006604	0,9641
3481	**0,017**	**0,006639**	**2,5596**
3480	0,0010	0,008436	0,1203
3498	**0,0141**	**0,004728**	**2,9849**
3500	0,0013	0,007432	0,1725
3502	0,0034	0,005631	0,6046
3465	-0,0014	0,009722	0,1418
3501	-0,0029	0,009521	0,3089
3514	-0,0103	0,010541	0,9816
4442	0,0026	0,008288	0,3153
4729	0,0118	0,01062	1,1114
5245	0,0117	0,008168	1,4304
5970	-0,0051	0,007092	0,7157
5972	0,0016	0,009031	0,1757
6143	-0,0136	0,007285	1,8666
6145	-0,0081	0,009834	0,8206
6146	0,0128	0,009158	1,3926
5974	-0,0054	0,009273	0,5838
6151	-0,0099	0,009198	1,0797
6109	0,0029	0,022442	0,1306
6865	**-0,0217**	**0,00998**	**2,1744**
7210	**0,0376**	**0,010663**	**3,5225**

Tableau 2 : *Effets de substitution allélique, erreur standard et valeur du test-T pour chacun des 26 pères Prim'Holstein de l'étude de cartographie à moyenne résolution pour le TNR281 sur le BTA01. Les pères hétérozygotes (P<0,05) sont en gras et en couleur rouge.*

Identifiant_père	Effet de substitution allélique	Erreur standard	Valeur du test-T
2108	0,0062	0,010878	0,5692
3442	**0,0402**	**0,012828**	**3,1328**
3487	-0,0019	0,007855	0,2376
3492	0,0074	0,01147	0,644
3468	0,0053	0,006994	0,7536
3481	**0,0177**	**0,007041**	**2,5127**
3480	-4,0E-4	0,008922	0,0411
3498	**0,0152**	**0,005036**	**3,0098**
3500	0,0011	0,007867	0,1398
3502	0,0018	0,005982	0,3087
3465	3,0E-4	0,010346	0,0302
3501	-0,0059	0,010123	0,5817
3514	-0,0072	0,011183	0,6417
4442	0,0021	0,008763	0,2367
4729	0,0179	0,011265	1,5846
5245	0,0128	0,008636	1,4843
5970	-0,0055	0,007539	0,7267
5972	-0,0014	0,009644	0,1462
6143	-0,0109	0,007792	1,4029
6145	-0,0093	0,010418	0,8943
6146	0,0165	0,009767	1,6938
5974	-0,0016	0,009843	0,1662
6151	-0,0131	0,009692	1,3568
6109	0,0095	0,02363	0,4016
6865	-0,0189	0,01063	1,7734
7210	**0,0325**	**0,011216**	**2,8939**

BTA02 :

Tableau 3 : *Effets de substitution allélique, erreur standard et valeur du test-T pour chacun des 26 pères Prim'Holstein de l'étude de cartographie à moyenne résolution pour le TNR90 sur le BTA02. Les pères hétérozygotes (P<0,05) sont en gras et en couleur rouge.*

Identifiant_père	Effet de substitution allélique	Erreur standard	Valeur du test-T
2108	0,0145	0,012016	1,204
3442	-0,0036	0,012493	0,289
3487	**0,0196**	**0,00833**	**2,3549**
3492	0,0108	0,010066	1,0707
3468	0,014	0,007725	1,8151
3481	8,00E-04	0,008142	0,095
3480	**-0,0267**	**0,010301**	**2,5939**
3498	0,0017	0,004895	0,3435
3500	0,0075	0,00869	0,8634
3502	-0,0119	0,00642	1,8572
3465	-0,0122	0,01041	1,1725
3501	-0,0091	0,009722	0,9317
3514	**-0,0277**	**0,010083**	**2,7497**
4442	0,0089	0,009796	0,9121
4729	8,00E-04	0,01043	0,0763
5245	-0,0069	0,009055	0,7573
5970	-0,0046	0,008606	0,5397
5972	0,0013	0,01286	0,1008
6143	-0,0025	0,006259	0,3954
6145	0,002	0,009263	0,2209
6146	-0,0069	0,008292	0,8321
5974	0,0026	0,009148	0,2815
6151	0,0105	0,009156	1,1511
6109	-0,0177	0,013275	1,3354
6865	**0,0228**	**0,008798**	**2,5949**
7210	-0,0199	0,010434	1,9107

Tableau 4 : *Effets de substitution allélique, erreur standard et valeur du test-T pour chacun des 26 pères Prim'Holstein de l'étude de cartographie à moyenne résolution pour le TNR281 sur le BTA02. Les pères hétérozygotes (P<0,05) sont en gras et en couleur rouge.*

Identifiant_père	Effet de substitution allélique	Erreur standard	Valeur du test-T
2108	0,0135	0,01202	1,1196
3442	-4,0E-4	0,013319	0,0294
3487	0,0159	0,008266	1,9207
3492	0,01	0,010803	0,9254
3468	**0,0152**	**0,007694**	**1,9769**
3481	-9,0E-4	0,008792	0,1008
3480	**-0,03**	**0,010406**	**2,8793**
3498	0,0019	0,004993	0,3762
3500	0,0079	0,008651	0,9124
3502	-0,0123	0,0064	1,9273
3465	-0,0107	0,011015	0,972
3501	-0,0025	0,0097	0,2622
3514	**-0,0267**	**0,010652**	**2,5084**
4442	0,0097	0,009981	0,9725
4729	9,0E-4	0,010421	0,0864
5245	-0,0166	0,010182	1,6313
5970	-0,0029	0,008611	0,3348
5972	0,0074	0,012922	0,5703
6143	-0,0026	0,006629	0,391
6145	-0,0025	0,009881	0,2559
6146	-0,0156	0,008765	1,7743
5974	0,0029	0,009792	0,2941
6151	0,0067	0,009711	0,6933
6109	-0,0222	0,014159	1,5691
6865	**0,0257**	**0,009488**	**2,7049**
7210	-0,022	0,01218	1,8103

BTA21 :

Tableau 5 : *Effets de substitution allélique, erreur standard et valeur du test-T pour chacun des 9 pères Normands de l'étude de cartographie à moyenne résolution pour le TNR56 sur le BTA21. Le père hétérozygote (P<0,05) est en gras et en couleur rouge.*

Identifiant_père	Effet de substitution allélique	Erreur standard	Valeur du test-T
1406	0,0057	0,008987	0,6314
1528	-0,0168	0,009055	1,8522
1537	0,0263	0,008684	3,0306
1788	-0,0068	0,012739	0,5303
1611	0,0048	0,0101	0,4763
1961	0,0214	0,013494	1,5842
1855	0,0304	0,019336	1,5699
2409	-0,009	0,013471	0,6648
2618	0,0271	0,01774	1,5296

BTA03

- ✓ Etude de cartographie à moyenne résolution

Tableau 6 : *Effets de substitution allélique, erreur standard et valeur du test-T pour chacun des 26 pères Prim'Holstein de l'étude de cartographie à moyenne résolution pour le TNR90 sur le BTA03. Les pères hétérozygotes (P<0,05) sont en gras et en couleur rouge.*

Identifiant_père	Effet de substitution allélique	Erreur standard	Valeur du test-T
2108	0,0075	0,008918	0,8439
3442	0,0015	0,011835	0,1229
3487	-0,0085	0,00712	1,2006
3492	-0,0133	0,010152	1,3109
3468	0,0013	0,005952	0,218
3481	-0,0023	0,004953	0,4661
3480	-0,0056	0,008784	0,6428
3498	-0,0055	0,004264	1,2796
3500	-0,0086	0,007026	1,2276
3502	9,00E-04	0,005588	0,1597
3465	0,0048	0,009654	0,4955
3501	0,0072	0,008496	0,8427
3514	**0,0187**	**0,009536**	**1,9623**
4442	**0,0217**	**0,007491**	**2,896**
4729	**0,0178**	**0,008636**	**2,059**
5245	-0,0059	0,008141	0,7285
5970	-0,0108	0,006513	1,6524
5972	0,0101	0,00859	1,1711
6143	-0,0025	0,005697	0,4358
6145	**-0,0213**	**0,008645**	**2,4586**
6146	0,0143	0,007702	1,8583
5974	0,0087	0,00908	0,962
6151	**-0,0279**	**0,009796**	**2,8457**
6109	0,0091	0,012488	0,7249
6865	0,0081	0,010061	0,8079
7210	0,0122	0,009257	1,318

Tableau 7 : *Effets de substitution allélique, erreur standard et valeur du test-T pour chacun des 26 pères Prim'Holstein de l'étude de cartographie à moyenne résolution pour le TNR281 sur le BTA03. Les pères hétérozygotes (P<0,05) sont en gras et en couleur rouge.*

Identifiant_père	Effet de substitution allélique	Erreur standard	Valeur du test-T
2108	0,0107	0,009524	1,1185
3442	-0,0010	0,012733	0,0813
3487	-0,0091	0,007609	1,2011
3492	-0,0126	0,010824	1,1619
3468	5,0E-4	0,006305	0,0731
3481	-0,0023	0,005301	0,4296
3480	-0,0066	0,009478	0,7013
3498	-0,0069	0,00455	1,5231
3500	-0,0098	0,007492	1,3079
3502	-6,0E-4	0,005906	0,0984
3465	0,0069	0,010421	0,665
3501	0,0102	0,008991	1,1365
3514	0,0149	0,010154	1,4662
4442	**0,0195**	**0,007948**	**2,4544**
4729	**0,0224**	**0,009354**	**2,3997**
5245	-0,0033	0,00876	0,3749
5970	-0,0080	0,0070	1,1379
5972	0,013	0,009356	1,3943
6143	-0,0016	0,006124	0,2603
6145	**-0,0197**	**0,009165**	**2,1512**
6146	0,0155	0,008229	1,8891
5974	0,0087	0,009666	0,9018
6151	**-0,0278**	**0,010335**	**2,6872**
6109	0,0112	0,013176	0,8505
6865	0,0059	0,010926	0,5445
7210	0,0156	0,009843	1,584

✓ Etude de cartographie fine

Tableau 8 : *Effets de substitution allélique, erreur standard et valeur du test-T pour chacun des 17 pères Prim'Holstein de l'étude de cartographie fine sur le BTA03 pour le TNR90. Les pères hétérozygotes (P<0,05) sont en gras et en couleur rouge.*

Identifiant_père	Effet de substitution allélique	Erreur standard	Valeur du test-T
2108	0,0046	0,009899	0,4606
3442	0,0055	0,013403	0,4077
3492	0,01	0,01054	0,9454
3465	-0,0037	0,011139	0,335
3514	**0,0216**	**0,010524**	**2,0521**
4442	**0,0253**	**0,008075**	**3,1349**
5245	0,0022	0,008531	0,2637
5970	0,0086	0,006531	1,3173
5972	-0,0039	0,008476	0,4549
6143	-0,0033	0,00574	0,5802
6145	**0,0264**	**0,008961**	**2,9448**
6146	**0,0195**	**0,007909**	**2,4717**
5974	0,0077	0,009065	0,8442
6151	**0,0249**	**0,010051**	**2,4803**
6109	-0,0211	0,013336	1,5827
6865	0,0056	0,010078	0,5587
7210	-0,0142	0,009732	1,4635