



**HAL**  
open science

## Diversité génétique des populations porcines françaises dans les régions chromosomiques soumises à la sélection : projet DIVQTL

Nathalie N. Iannuccelli, Juliette Riquet, Marie-José. Mercat, H. Legros, Magali San Cristobal, Anne Lacoste, Jean Pierre Bidanel, Denis Milan

### ► To cite this version:

Nathalie N. Iannuccelli, Juliette Riquet, Marie-José. Mercat, H. Legros, Magali San Cristobal, et al.. Diversité génétique des populations porcines françaises dans les régions chromosomiques soumises à la sélection : projet DIVQTL. Actes du BRG, 2006, 6, pp.111-128. hal-02653875

**HAL Id: hal-02653875**

**<https://hal.inrae.fr/hal-02653875v1>**

Submitted on 30 Jul 2022

**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.

## Diversité génétique des populations porcines françaises dans les régions chromosomiques soumises à la sélection : projet DIVQTL

Nathalie IANNUCELLI<sup>(1)</sup>, Juliette RIQUET<sup>(1)</sup>, Marie-José MERCAT<sup>(2)</sup>,  
Hélène LEGROS<sup>(3)</sup>, Magali SANCRISTOBAL<sup>(1)</sup>, Anne LACOSTE<sup>(2)</sup>,  
Jean-Pierre BIDANEL<sup>(4)</sup>, Denis MILAN<sup>(1)\*</sup>

<sup>(1)</sup> INRA, Laboratoire de Génétique Cellulaire, 31326, Castanet-Tolosan, France

<sup>(2)</sup> ITP, Pôle Génétique et Qualité des produits, 35651 Le Rheu, France

<sup>(3)</sup> Labogena, INRA, 78352 Jouy-en-Josas, France

<sup>(4)</sup> INRA, Station de Génétique Quantitative et Appliquée, 78352 Jouy-en-Josas, France

**Abstract: Genetic diversity of French pig populations in QTL regions under selection: the DIVQTL project.** The purpose of this study was to analyze the genetic diversity of the French porcine populations in several regions of the genome containing QTL and major genes. A first phase was carried out on 1000 animals with 10 markers microsatellites, followed by a second phase where a selection of 300 individuals representative of the various breeds was analyzed on 27 complementary microsatellites as well as a first set of 29 SNP.

This study enables us to show a panorama of the whole of French porcine breeds and lines, by showing at the same time their within-line diversity as well as the similarities between these lines. The tree of the relations between races presents a vision picture which seems very accurate, for the lines derived from the same genetic type (Large White, Landrace...) as well as for the relationship of the synthetic lines with their parental races. Among the local breeds for example, the similarities (although moderate) between the "Blanc de l'ouest" and the "Bayeux" seem logical within sight of the history of these races. For the various closed lines held by the various breeding companies, one clearly highlights those primarily made up with Large White or Landrace blood. The lines containing Piétrain blood are also gathered within a Piétrain branch, including the Tiameslan race and its Redone derivative also containing Chinese blood. Among the "European" breeds, the Créole one, which is indicated on the site of the BRG ([http://www.brg.prd.fr/brg/pages/rga/porcins/190\\_pag](http://www.brg.prd.fr/brg/pages/rga/porcins/190_pag)) as "issued from Iberian breeds then English (Large Black), American (Duroc) and French (Normand, Craonnais) breeds", is the most original line compared to the other ones.

With a little astonishment, it appears that the methods based on allelic frequencies make it possible to produce results which appear plausible for lines analyzed on very reduced numbers of individuals (6 animals only for some lines). The second part of this study (2nd set of microsatellites and markers SNP) relating to only 300 animals for 33 races, one could expect a less clear vision.

---

\* Correspondance et tirés à part : [denis.milan@toulouse.inra.fr](mailto:denis.milan@toulouse.inra.fr)

Beyond the comparisons between breeds, this program also aimed at carrying a first glance on diversity in several genomic regions containing genes influencing the traits of economic interest, whose variability was likely to have evolved under the influence of the animal breeding. It is clear that within this framework, the number of analyzed animals is too weak to be able to produce another thing that a qualitative impression. This program however made it possible to highlight areas for which the selection of the animals has privileged certain alleles. For some SNP where the two alleles are found in all the types of races, one indeed observes very different frequencies between lines, suggesting a different selection in these lines. It is for this reason remarkable to find a homogeneity rather strong among the male synthetic races containing Piétrain blood.

This program thus deserves to be continued on a larger scale by analyzing all 1000 animals on the whole of the markers of this program having a sufficient informativeness. A prolongation of this program will be carried out within the framework of the Agenae project Biomark that is cofunded by the ministry for research and Bioporc association representing the porcine breeding.

#### **diversity/ pig STL/ selection**

**Résumé :** Le projet DIVQTL avait pour but d'analyser la diversité génétique de la quasi-totalité des populations porcines françaises présente au sein de 8 régions du génome contenant des QTL ou des gènes majeurs. Une première phase a été réalisée sur 1 000 animaux avec 10 marqueurs microsatellites, suivie d'une deuxième phase où une sélection de 300 individus représentatifs des différentes races a été analysée sur 27 microsatellites complémentaires ainsi qu'un premier jeu de 29 SNP.

Cette étude nous permet de montrer un panorama de l'ensemble des races et lignées porcines françaises en montrant à la fois leur diversité intra race ainsi que les ressemblances entre ces races. Les races synthétiques apparaissent ainsi proches de leurs races fondatrices les plus importantes. Au sein des différentes lignées dérivées d'un même type, on peut également identifier des lignées plus différentes des autres, et donc méritant une attention particulière. Avec un peu d'étonnement, il apparaît que les méthodes basées sur des fréquences alléliques de marqueurs situés dans ces régions soumises à la sélection permettent de produire des résultats qui paraissent sensés pour des lignées analysées sur des nombres d'individus très réduits (6 animaux seulement pour certaines lignées).

#### **diversité/ porc/ QTL/ sélection**

### **1. INTRODUCTION**

Depuis les années 1990, l'essor de la génétique moléculaire a fourni de nouveaux outils permettant une meilleure connaissance des animaux par l'analyse de leur génome. Le développement de marqueurs génétiques a permis la cartographie de régions du génome contenant des gènes majeurs ou des QTL influençant de façon importante les performances des animaux.

Des programmes de cartographie fine de ces régions peuvent alors être entrepris de façon à cartographier finement les régions concernées, et si possible identifier le gène et la mutation responsable des effets observés. Dans un premier temps ces travaux de cartographie de QTL ont été essentiellement entrepris sur des familles expérimentales ou produites dans le cadre des programmes d'amélioration génétique des animaux. Ces techniques permettent de détecter les régions du génome porteuses des gènes d'intérêt, la précision de la cartographie ainsi obtenue est assez faible, le gène étant cartographié dans un intervalle de 20 voire 40 cM [1], [2]. D'autres techniques de cartographie sont alors nécessaires comme l'étude d'animaux porteurs de chromosomes recombinants dans la région étudiée [3], [4], la comparaison fine d'haplotypes caractérisés comme porteurs ou non de la mutation [5], ou l'étude du déséquilibre de liaison [6] permettant une cartographie très précise de l'emplacement de la mutation.

À côté des familles permettant la cartographie initiale des QTL, l'étude des populations sélectionnées fournit des résultats très instructifs à plusieurs titres. En effet, si un QTL ayant un effet favorable sur l'un des caractères pris en compte dans l'index de sélection est en ségrégation dans une population sélectionnée [7], on s'attend à ce que la fréquence de l'allèle favorable augmente dans la population. Au sein de ces populations, on s'attend également à ce que le déséquilibre de liaison, observé sur l'allèle sélectionné, s'étende sur une région plus grande que celui observé sur les autres allèles [8].

Dans un autre ordre d'idée, la génétique moléculaire permet d'étudier la diversité génétique des populations animales. L'analyse des fréquences de marqueurs alléliques et leur comparaison entre races permet de définir des distances génétiques entre races, et d'identifier les races les plus similaires. Des premières études de la diversité des races porcines ont pu ainsi être réalisées dans le cadre des programmes européens PiGMaP [9] puis PigBio-Div I [10], au moyen de marqueurs microsatellites.

Dans le cadre du programme DIVQTL, nous avons réalisé une première analyse de la diversité de 33 races et lignées porcines françaises, en analysant des marqueurs situés dans les régions contenant les QTL les plus importants sur lesquels nous réalisons actuellement des travaux de cartographie fine sur les chromosomes 1, 2, 4 et 7, ou à proximité de quelques gènes ayant un effet majeur sur les chromosomes 6 et 15.

Les objectifs de ce projet étaient de : comparer les lignées entre elles (et notamment les lignées synthétiques relativement à leurs lignées parentales), tenter de recenser la diversité génétique présente dans les différentes races dans les régions étudiées, combiner l'apport de microsatellites et de SNP, recenser pour les haplotypes connus leur fréquence dans les différentes lignées.

## 2. MATÉRIEL ET MÉTHODES

### 2.1. Animaux

Le tableau I présente la liste des 33 types génétiques échantillonnés, avec l'indication pour les lignées synthétiques des races utilisées pour les établir. Des informations complémentaires sur ces lignées peuvent être obtenues sur le site de l'agence de la sélection porcine (<http://www.asp.asso.fr>).

En règle générale l'objectif était de disposer de 20 animaux des races locales et de 40 animaux des races commerciales. Dans certains cas (Large White et Landrace), la population a été subdivisée en un certain nombre de sous-échantillons. Ce choix des animaux a été coordonné par l'ITP.

Les 300 animaux indiqués au tableau I dans la colonne « 2<sup>e</sup> phase » sont des animaux appartenant également à la 1<sup>e</sup> phase. Après analyse du premier jeu de marqueurs nous avons retenus de 8 à 12 animaux par race en cherchant à sélectionner des animaux représentant à l'état hétérozygote les allèles moins fréquents de l'une des régions analysées. Pour pouvoir déterminer si possible la phase des allèles analysés, nous avons dans chaque région retenu également un animal homozygote pour l'allèle le plus fréquent.

### 2.2. Marqueurs

Deux jeux de marqueurs microsatellites ont été analysés. Un premier jeu constitué des marqueurs suivants a été analysé sur tous les animaux échantillonnés : Sw1301 (chr1), Swc9 (chr2), S0214 et S0217 (chr4), CRC et Sw1129 (chr6), LRA1, Sw1856 et S0101 (chr7), S1002 (chr15). Un jeu complémentaire de 27 microsatellites a été analysé sur le sous-ensemble de 300 individus : SW216, SW501, SW1828, MCS410D1A, MCS647E6A et MCS455C8A (chr1), MCS1002, MCS141B2A, MCS218, SW2623 et MCS364 (chr2), S0001, SW969, SW1492, SW1364, SW714 et SW1996 (chr4), SW1067 (chr6), SW2019, BMP5, DAXX, WAF, NFY, S0102 et CBGR1 (chr7), S1006 et S1008 (chr15). Le génotypage des microsatellites a été réalisé sur séquenceur ABI3730 par Labogena (<http://www.labogena.fr>).

Diversité génétique des populations porcines françaises sélectionnées

**Tableau I** : Animaux échantillonnés dans le cadre du programme DIVQTL.

Type racial	Race ou lignée	Acronyme	Nb animaux	
			1 <sup>e</sup> phase	2 <sup>e</sup> phase*
Duroc	Duroc	Du	43	8/2/-
Duroc	DRB (Du, LW, LR) F	DRB	40	8/3/-
Duroc	DRC (Du, LW, LR) M	DRC	40	8/1/-
Locale	Basque	Ba	20	8/-/-
Locale	Blanc de l'Ouest	BO	20	8/-/-
Locale	Bayeux	By	20	8/-/-
Locale	Créole	Cr	20	8/-/-
Locale	Gascon	Ga	20	8/1/-
Locale	Limousin	Li	20	8/1/-
Landrace	Carélie (Landrace Finlandais)	Car	37	8/1/-
Landrace	FH012 (LR)	FH12	40	8/1/-
Landrace	Landrace	LR	20	8/1/-
Landrace	Landrace 1977	LR77	23	8/-/-
Landrace	Landrace français	LRF	6	5/-/-
Landrace	P321 (LR, chinois)	P321	40	8/-/-
Large White	FH025 (LW)	FH25	39	8/1/-
Large White	Gallia (LW, LR)	Gal	40	8/1/-
Large White	Large White	LW	22	8/2/-
Large White	Large White 1977	LW77	21	8/-/-
Large White	Large White Femelle	LWF	9	7/-/-
Large White	Large White Male	LWM	26	8/1/-
Large White	P121 (LW)	P121	36	8/-/-
Synthétique Femelle	FH004 (synthé F)	FH4	40	8/1/-
Piétrain	FH016 (Piétrain)	FH16	40	8/2/-
Piétrain	Piétrain	Pi	18	8/2/-
Chinoise	Meishan	MS	20	8/4/3
Synthétique chinoise	Tiameslan (MS, Laco)	TML	27	8/2/-
Synthétique chinoise	Redone (TML, Carélie)	Re	29	8/3/-
Synthétique chinoise	Taizumu (LWH, MS)	TZM	37	8/2/-
Synthétique Mâle	FH019 (LW, Ha, Du)	FH19	41	8/1/-
Synthétique Mâle	Laconie (LW, Ha, Pi)	Laco	42	8/3/-
Synthétique Mâle	Musclor (Pi, LW, DRC)	Muscl	40	8/1/-
Synthétique Mâle	Penshire (Ha, LW, Du)	PSR	41	8/-/-
Total			977	260/37/3

\* X/Y/Z : X animaux choisis pour l'hétérozygotie / Y animaux choisis pour leur homozygotie pour un allèle majoritaire / Z animaux parentaux d'un programme QTL (Bidanel *et al.* 2001) choisis comme témoin.

Un ensemble de 29 SNP a également été analysé par la technique Amplifluor sur le sous-ensemble de 300 animaux. Ces marqueurs ont été choisis pour leur localisation à proximité des QTL d'intérêt. Parmi ces marqueurs, on peut citer les 4 SNP mis en évidence dans le gène de la CBG, modifiant la protéine CBG [11]. Le génotypage a été réalisé sur la plateforme CRGS du Génopole Toulouse Midi-Pyrénées (<http://crgs.genopole-toulouse.prd.fr>).

### 2.3. Distance génétique et mesure de dispersion des marqueurs

Afin d'estimer l'information apportée par chaque marqueur, et de tenter d'identifier les marqueurs pour lesquels une différenciation forte des races existe, une mesure de la dispersion des races a été calculée pour chaque marqueur  $X$ , comme la moyenne des distances de Reynolds de chaque population à une population virtuelle incluant tous les individus :

$$Disp(X) = \frac{\sum_{pop} D_{pop}(X)}{nbpop} \quad \text{où} \quad D_{pop}(X) = \frac{\sum_{allèlei=1}^n (f_{pop}(x_i) - f_{globale}(x_i))^2}{1 - \sum_{allèlei=1}^n (f_{globale}(x_i))^2}$$

$f_{globale}(x_i)$  étant la fréquence de l'allèle  $x_i$  du marqueur  $X$  calculée sur l'ensemble des animaux génotypés,  $f_{pop}(x_i)$  étant la fréquence de l'allèle  $x_i$  du marqueur  $X$  calculée sur la population des animaux génotypés.

Pour analyser la diversité génétique entre races, des calculs de distance entre races basées sur la distance de Reynolds ainsi qu'une classification automatique des races ont été réalisés comme décrit précédemment [10].

Le diagramme « heatmap » réalisé sous R représente l'apparement moyen intra et inter races  $\sum_i f_{pop1}(x_i) \times f_{pop2}(x_i)$  après réarrangement des populations (lignes et colonnes) selon les résultats de la classification automatique. Sur la diagonale, cette valeur correspond à l'hétérozygotie théorique de la race. Une couleur foncée indique un fort niveau d'apparement.

### 2.4. Déséquilibre de liaison

Le déséquilibre de liaison entre marqueurs a été estimé suivant les deux paramètres classiques  $D'$  et  $r^2$  [12], [13] :

$$D' = \sum_i \sum_j p_i p_j |D'_{ij}| \quad \text{et} \quad r^2 = \sum_i \sum_j \frac{D_{ij}}{p_i q_j} \quad \text{où}$$

$$D'_{ij} = \frac{D_{ij}}{D_{\max}} \quad \text{avec} \quad D_{ij} = h_{ij} - p_i q_j$$

où  $h_{ij}$  fréquence de l'haplotype porteur des allèles  $i$  et  $j$  aux locus  $X$  et  $Y$

et  $D_{\max} = \min(p_i q_j, (1 - p_i)(1 - q_j))$  si  $D_{ij} < 0$ ,

et  $D_{\max} = \min(p_i(1 - q_j), (1 - p_i)q_j)$  si  $D_{ij} > 0$ .

avec  $p_i$  fréquence de l'allèle  $i$  du premier marqueur, et  $q_j$  fréquence de l'allèle  $j$  du deuxième marqueur.

Pour cette analyse, ne disposant pas de la phase des marqueurs, seules les paires de génotypes où l'un au moins des deux marqueurs était homozygote ont été pris en compte.

### 3. RÉSULTATS ET DISCUSSION

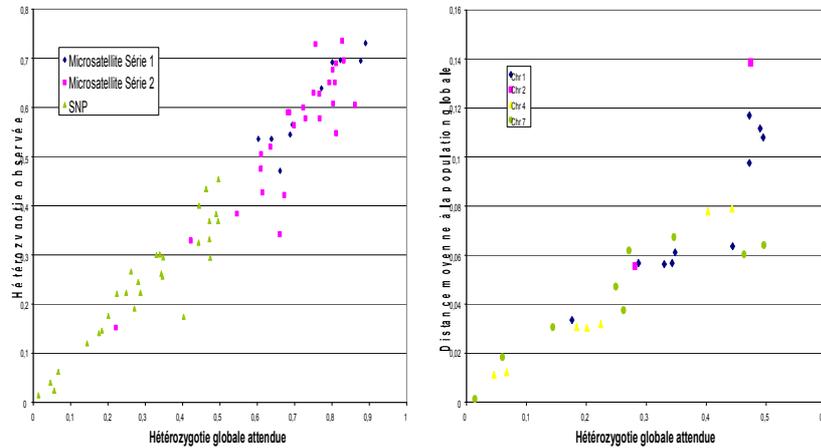
#### 3.1. Polymorphisme et informativité des marqueurs

Sur les 33 races ou lignées analysées, nous avons dans un premier temps étudié par région un ou deux marqueurs microsatellites très informatifs afin de disposer d'un premier aperçu de la diversité existante dans les différentes races. Le budget imparti au projet ne nous permettant pas d'analyser la totalité des animaux sur un grand nombre de marqueurs, nous avons dû réduire le jeu d'animaux à 300 individus représentatifs des populations. Pour chaque région, nous avons sélectionné des animaux porteurs à l'état hétérozygote des allèles plus rares et parfois originaux, ainsi que des animaux homozygotes porteurs des allèles les plus fréquents.

Pour tester si la réduction du nombre d'animaux induisait un biais important des fréquences alléliques, nous avons comparé les fréquences calculées pour les 33 races sur les deux jeux rassemblant au total 977 et 300 animaux. Sur les 1 249 fréquences alléliques intra races non nulles sur le jeu des 977 animaux, la différence moyenne par rapport aux mêmes fréquences sur le jeu de 300 animaux n'est que de 5,0 % ; 85,7 % des fréquences diffèrent de moins de 10 %, 96,7 % de moins de 15 %, 98,8 % de moins de 20 %. Ces différences sont de plus inmanquablement exacerbées pour les microsatellites qui présentent en moyenne 10,2 allèles par rapport aux SNP bialléliques. On peut donc penser que la manière dont nous avons établi le jeu réduit de 300 animaux pourra permettre de limiter les erreurs d'estimation des fréquences alléliques intra-race pour les marqueurs SNP bialléliques.

Vingt-sept marqueurs microsatellites complémentaires ont été analysés sur les 300 animaux sélectionnés. Un jeu de 29 SNP cartographiés à proximité des marqueurs microsatellites a également été analysé sur les 300 animaux.

Sur les deux jeux de microsatellites constitués respectivement de 10 et 27 microsatellites, des hétérozygoties moyennes de 61 % et 55 % ont été observées. Sur le jeu de SNP, une hétérozygotie moyenne de 24 % a été observée (fig. 1A).



**Figure 1 :** A) Hétérozygotie des marqueurs analysés : hétérozygotie observée (pourcentage d'animaux hétérozygotes) et attendue ( $1 - \Sigma p_i^2$ ) ; B) Mesure de la dispersion pour les SNP, en fonction de l'hétérozygotie attendue.

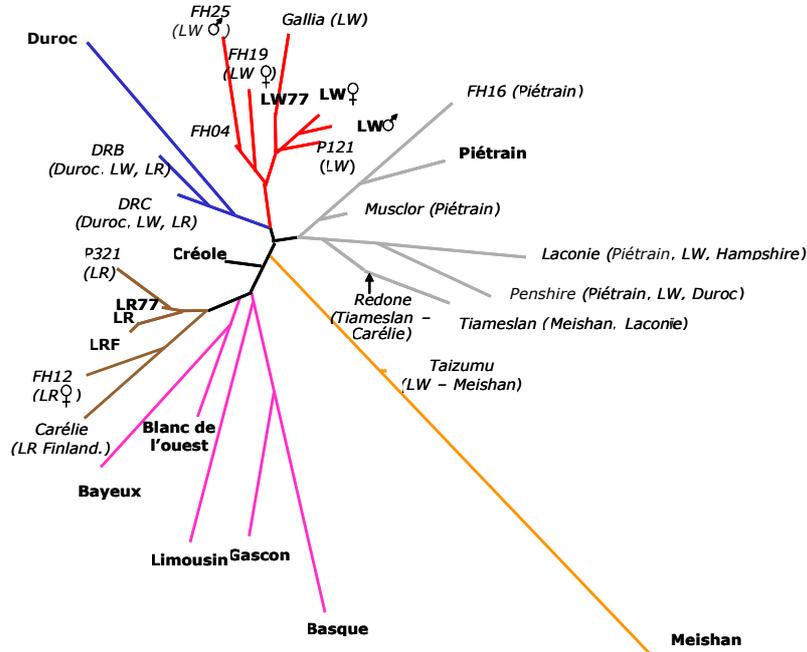
Afin d'identifier les marqueurs SNP pour lesquels les races semblaient présenter des différences de fréquences alléliques importantes, signe potentiel de sélection différentielle de certains allèles dans certaines lignées, nous avons calculé pour chaque marqueur une mesure de la dispersion *Disp* des races exprimée comme la moyenne des distances de Reynolds pour chacune des 33 populations à une population virtuelle composée de l'ensemble des animaux analysés. Une valeur de cette mesure sera importante pour des marqueurs où d'une race à l'autre des fréquences alléliques très différentes sont observées. Pour un marqueur, pour lesquels les fréquences alléliques seraient les mêmes dans toutes les races, cette mesure de dispersion prendrait une valeur nulle. Cependant, pour que des fréquences très différentes puissent être observées dans différentes races, il est nécessaire que les fréquences des deux allèles soient globalement assez équilibrées.

La figure 1B montre la valeur de cette mesure de dispersion des races relativement à l'informativité potentielle du marqueur, exprimée ici par l'hétérozygotie globale attendue calculée sur la population virtuelle globale. Cette mesure est ainsi fortement corrélée à l'hétérozygotie du marqueur pour les marqueurs faiblement ou modérément hétérozygotes (hétérozygotie < 40 %). Pour ces marqueurs, une différenciation importante pourra être toutefois éventuellement observée sur quelques races, pour des marqueurs comme les marqueurs du gène *CBG* où certains allèles sont spécifiques des races chinoises.

Pour les marqueurs présentant une hétérozygotie théorique supérieure à 40 %, on peut potentiellement s'attendre à ce que les deux allèles soient présents à des fréquences équilibrées dans beaucoup de races. Pour ces marqueurs, on peut voir de fortes différences allant jusqu'à un facteur 2 pour cette mesure de dispersion, mettant en évidence des différences de répartition des allèles entre races. Même si le nombre de marqueurs de ce type est pour l'instant faible (10 marqueurs), une certaine corrélation semble apparaître entre des valeurs de même hétérozygotie située dans une même région chromosomique. Pour prendre deux exemples : 1) Le marqueur du chromosome 2, pour lequel la dispersion des races est la plus forte est un marqueur situé à proximité immédiate du gène *IGF2* pour lequel une mutation influençant le taux de muscle a été identifiée. La mutation de ce gène d'expression paternelle, est particulièrement sélectionnée dans les lignées mâles. 2) De même le marqueur MC4R (chr1, région 1A fig. 4), correspondant à une mutation affectant potentiellement l'engraissement et la croissance des animaux [14] présente des fréquences alléliques assez différentes en fonction des races.

### 3.2. Comparaison des races étudiées

Même si un certain nombre de marqueurs étaient liés entre eux, l'ensemble du jeu de données a été utilisé pour analyser la diversité entre races. Un arbre a ainsi été réalisé sur les races en utilisant la méthode du plus proche voisin sur les distances de Reynolds calculées entre races [10]. L'arbre obtenu est particulièrement cohérent (fig. 2), les différentes races apparentées étant regroupées. Les différentes lignées Large White ou Landrace se regroupent entre elles, même lorsqu'un faible nombre d'individus avait pu être échantillonné (6 animaux dans certaines lignées seulement). Les lignées fermées le plus souvent synthétiques développées par les sélectionneurs privés se regroupent avec les races ayant eu la contribution majeure à l'établissement de ces lignées (lignées FH04, FH16, FH19, FH25, P121, P321, DRB, DRC, Musclor). Les différentes lignées mâles établies à partir d'animaux Piétrain se regroupent entre elles (FH16, Musclor, Laconie, Panshire...). Les lignées sino-européennes dérivées de la lignée Laconie (Tiameslan, Redone) se regroupent avec la Laconie, elle-même regroupée avec les animaux Piétrain, suggérant que les animaux dérivés de lignées apparentées au Piétrain et d'animaux Meishan, présentent dans les régions QTL un génome plus proche du génome Piétrain. Au contraire, la lignée Taizumu établie à partir de truies Meishan et de verrats de la lignée Large White femelle, est située sur la branche Meishan de l'arbre. L'ensemble des races locales, se regroupent entre elles, mais sont assez distinctes les unes des autres.



**Figure 2 :** Arbre selon la méthode du plus proche voisin, basé sur les distances de Reynolds entre populations (les races sont indiquées en gras, et les lignées en italique avec l'information des races à partir desquelles elles ont été établies).

À partir des mêmes données, une classification automatique des races, et un diagramme de type « Heatmap » a également été réalisé (fig. 3). Cette classification donne une image assez semblable à la précédente, en montrant une forte similarité des lignées de même type. Les souches Large White apparaissent particulièrement semblables. Parmi les races locales (qui sont regroupées entre elles), le Bayeux semble également proche des animaux Landrace. Le Meishan est quant à lui très différent des autres, pratiquement pas apparenté aux autres populations à l'exception des lignées synthétiques dérivées de cette race (Taizumu, Tiameslan, Redone). La classification automatique regroupe entre elles ces lignées dérivées du Meishan, au sein des races européennes, proches des lignées Duroc.

## Diversité génétique des populations porcines françaises sélectionnées

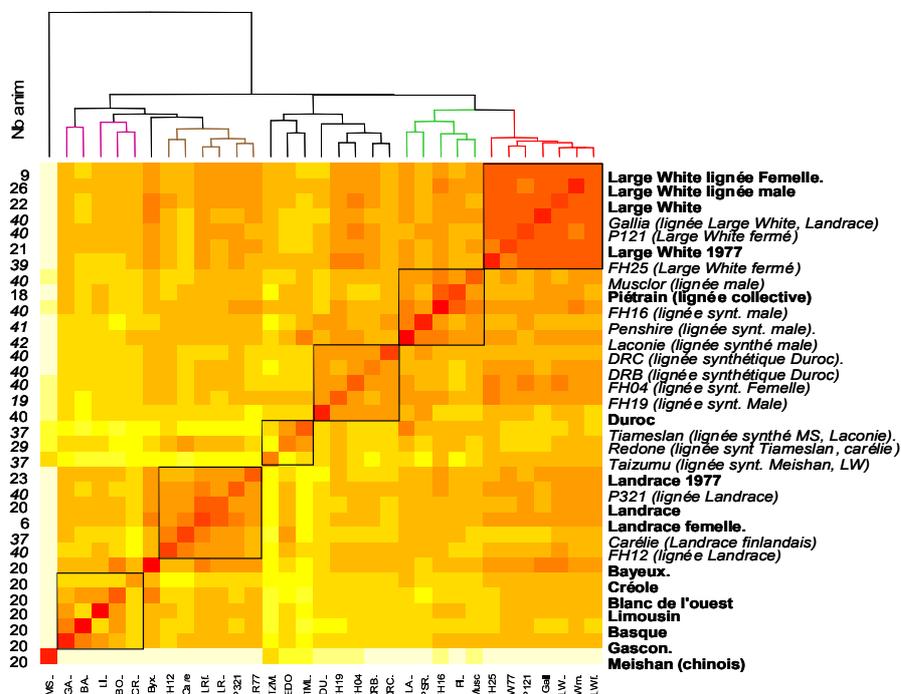
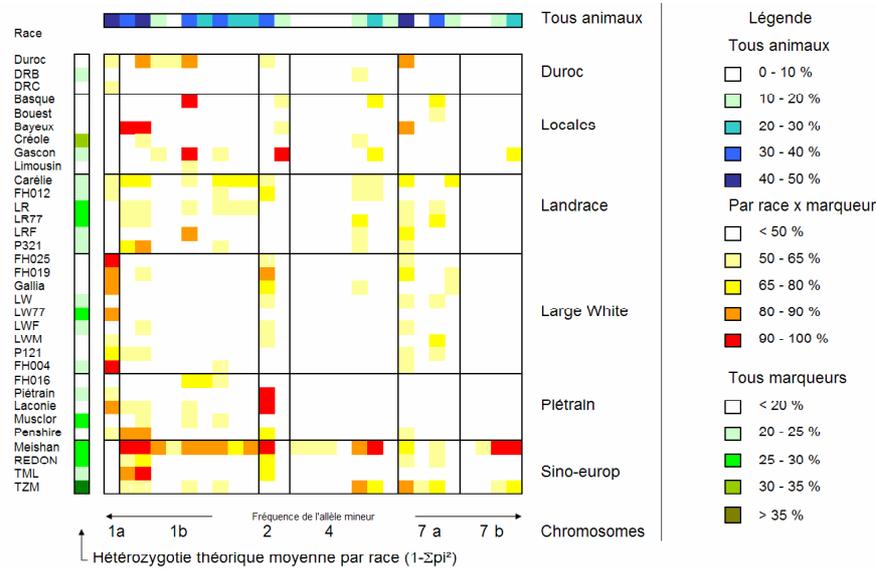


Figure 3 : Classification des races porcines basée sur l'analyse conjointe des génotypes microsatellites et SNP.

### 3.3. Analyse du polymorphisme des régions

Les SNP que nous avons étudiés ont été choisis en raison de leur position connue à proximité de QTL d'intérêt (QTL chr 1, 2, 4, 7) ou pour le polymorphisme décrit au sein de gènes dont l'effet sur le caractère est connu (CBG, RN...). La figure 4 présente les Fréquences de l'Allèle Minoritaire globalement (FAM) race par race. Ces marqueurs présentent des taux de polymorphisme très différents (voir ligne « tous animaux » de la figure 4). Un allèle globalement rare, peut toutefois être retrouvé à une forte fréquence dans une race donnée, soit en raison de l'apparition de cette mutation dans une race particulière, soit en raison d'une gestion différente de cette race (sélection sur des critères ou intensités différents ...). Au contraire des microsatellites qui présentent des taux de mutations plus élevés, on considère que si un même allèle SNP est identifié dans deux animaux différents, ces deux copies dérivent d'un même haplotype muté ancestral.



**Figure 4 :** Hétérozygotie théorique moyenne par race, et fréquence de l'allèle minoritaire des SNP (globale et intra race).

La figure 4 met notamment en valeur une fois de plus la spécificité de la race chinoise Meishan, les allèles majoritairement présents dans cette race étant le plus souvent des allèles globalement minoritaires.

Sur la région « 1a » du chromosome 1 (région du gène *MC4R*), les lignées Large White présentent le plus souvent l'allèle Asp298 minoritaire à l'échelle de l'ensemble de la population. Certaines lignées Large White, présentent au contraire majoritairement l'allèle 298Asn, comme les lignées Landrace. Cet allèle Asp298 est associé à un effet plutôt dominant accélérant la croissance et l'engraissement des animaux [14], [16]. Des études complémentaires sur un nombre d'animaux plus importants méritent d'être entreprises pour approfondir ces premiers résultats et envisager si ces différences peuvent s'expliquer par des compositions d'index de sélection différents selon les lignées. Sur la région « 1b » du chromosome 1 contenant le QTL influençant l'engraissement des animaux, les allèles minoritaires globalement sont retrouvés en Meishan, mais aussi dans certaines races locales et dans une moindre mesure la race Duroc.

Pour la région IGF2 (chr2) contenant un QTL influençant le taux de muscle des animaux, l'allèle minoritaire du premier marqueur, situé à proximité immédiate du gène *IGF2* est rencontré en race Piétrain comme en Meishan, races pourtant totalement opposées en terme de taux de muscle. Cette constatation peut être rapprochée de l'identification en race Meishan

d'un haplotype totalement identique à l'haplotype muté (à l'exception de la mutation) observé en race Piétrain [5].

Pour le gène *CBG* (chr7), gène candidat très fort pour un QTL de cortisolémie [11], différentes mutations ont été mises en évidence (fig. 5). Certaines ont été rencontrées à l'état hétérozygote chez des verrats hétérozygotes au QTL, mais il n'a pas été possible d'identifier une seule mutation potentiellement responsable. Les mutations S15I et I265V pourraient toutefois avoir un rôle dans le QTL observé, induisant une cortisolémie plus importante, corrélée avec un engraissement plus important des animaux. Dans ce cas, les haplotypes favorables, qu'il pourrait être intéressant de sélectionner notamment en lignée mâle, seraient les haplotypes LW1, LW2 et LW3. Notre étude montre que parmi les lignées commerciales, l'allèle LW4 est essentiellement rencontré dans différentes lignées Piétrain ou dérivées du Piétrain, du Meishan ou du Duroc, ainsi que dans une moindre mesure dans les lignées Large White ; il semble par contre rare en Landrace. Par ailleurs, un nouvel haplotype recombinant a pu être mis en évidence (appelé ici Euro11) rencontré dans des races aussi différentes que la Créole, le Gascon, ou des synthétiques Piétrain.

Haplotypes selon Jusova et al.	Exon 2	Exon 3	Exon 4	Allèle				
	316I	I267M	266V	3307R	Majoritaire dans les races	Fréquent dans les races	Absent (pour les allèles fréquents)	Allèle rare, mais observé en
LW1, LW2, LW3	G	C	A	A	Taizumu		Meishan	
LW4, MS8, MS9	G	C	A	G	Besq, B. Ouest, Gasc. Lim.		Piétrain, Meishan, LRs	
MS5, MS6, MS7	G	C	G	G	Meishan, Synthe MS & Duroc, Gasc., Piétrains, 1 LW		Besq, B. Ouest, Bayeux, LRs	
MS10	T	C	G	G	Meishan			Meishan, Taizumu
Euro 11	G	C	G	A				5x Synthés Piétrain, 2xGasc., 1xGr.

Figure 5 : Haplotypes *CBG* rencontrés (selon nomenclature [11]).

Des études similaires sont actuellement menées dans les autres régions où la densité de marqueurs SNP le permet.

### 3.4. Variabilité intra-race

La variabilité intra race a été estimée à partir des hétérozygoties moyennes théoriques calculées sur les données microsatellites ou SNP. Les données race par race sont présentées graphiquement pour les SNP en première colonne de la figure 4. Les valeurs calculées par type de race (locale, majeure, synthétique européenne ou sino-européenne) sont présentées en tableau II pour les microsatellites et les SNP. Si les deux types de données vont dans le même sens, les données observées avec les marqueurs SNP présentent une situation plus tranchée.

La population la plus diverse est la lignée synthétique Taizumu (36 % d'hétérozygotie moyenne théorique des SNP), établie récemment à partir de deux races très différentes la Meishan et le Large White lignée femelle. D'une façon générale, au sein de chaque type (Large White, Landrace, Piétrain), les lignées collectives présentent une diversité moyenne de 23 % su-

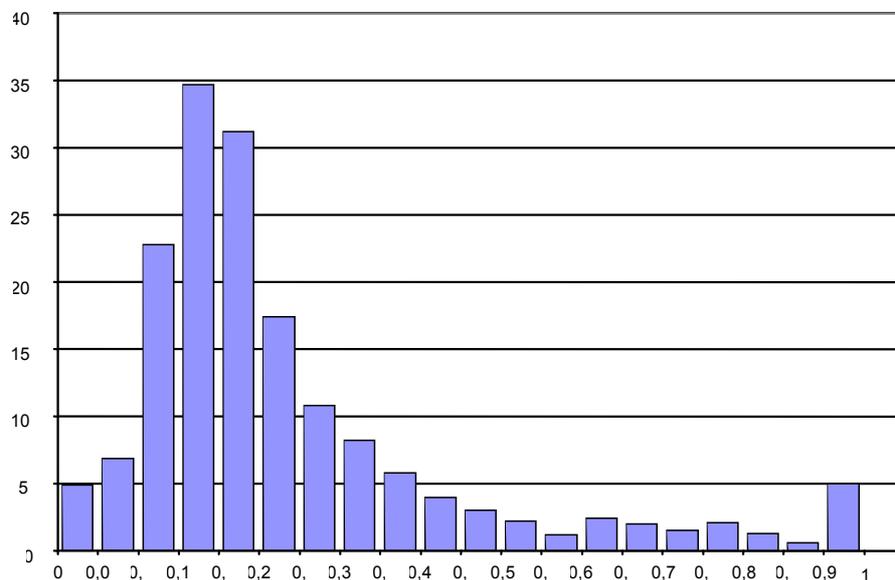
périeure à celle des lignées privées européennes (19,7 %) dont le nombre de reproducteurs est plus limité. Les races synthétiques sino-européennes présentent en moyenne une diversité supérieure due à l'utilisation de deux types de races très différentes. Les races locales présentent des diversités très hétérogènes, la Créole étant l'une des races les plus hétérozygotes (31 %) alors que les races Limousine, Basque et Bayeux sont parmi les races les plus homozygotes.

**Tableau II** : Hétérozygotie moyenne des marqueurs par type de race.

	Races			
	Locales	Majeures	Synthétiques	
			Européennes	Sino-européennes
Microsatellites				
Moyenne	50,7 %	53,9 %	53,5 %	58,4 %
Min	43,6 %	48,5 %	46,4 %	49,6 %
Max	65,3 %	58,9 %	60,9 %	63,6 %
SNP				
Moyenne	20,4 %	23,1 %	19,7 %	28,4 %
Min	15,7 %	19,2 %	14,1 %	20,3 %
Max	31,2 %	26,5 %	26,1 %	36,5 %

### 3.5. Déséquilibre de liaison

Si la densité assez faible des marqueurs analysés ne permet pas d'étudier des haplotypes pour toutes les régions, il est possible d'étudier le déséquilibre de liaison entre marqueurs situés sur des chromosomes différents (fig. 6). Le déséquilibre de liaison observé, bien qu'affecté par la répartition des animaux en différentes sous-populations isolées, présente une distribution globalement similaire à celle observée par Farnir *et al.*, sur une population bovine [17]. Les quelques cas de déséquilibre total pour des marqueurs situés sur des chromosomes différents s'explique ici par la présence de quelques allèles rencontrés spécifiquement dans une même race, sur deux chromosomes différents. Ces études devront être reprises sur des nombres plus conséquents d'individus.



**Figure 6 :** Distribution du déséquilibre de liaison D entre marqueurs situés sur des chromosomes différents.

L'étude de nombreux marqueurs proches permettra également d'étudier le déséquilibre de liaison rencontré sur un même chromosome dans les régions QTL.

#### 4. CONCLUSION

Cette étude nous permet de dresser un panorama de l'ensemble des races et lignées porcines françaises en montrant à la fois leur diversité intra race ainsi que les ressemblances entre ces races. L'arbre des relations entre races présente une vision qui semble très juste, tant pour les lignées dérivées d'un même type génétique (les Large White, les Landrace...) que pour le positionnement des lignées synthétiques par rapport aux races parentales. Au sein des races locales par exemple, l'apparentement (bien que modéré) entre le Blanc de l'Ouest et le Bayeux semble logique au vu de l'historique de ces races. Pour les différentes lignées fermées détenues par les différents sélectionneurs, on met clairement en évidence celles constituées essentiellement à partir de sang Large White ou Landrace. Les lignées contenant du sang Piétrain sont elles aussi regroupées entre elles, et regroupées au sein d'un rameau Piétrain, y compris la race Tiameslan et son dérivé Redone contenant également du sang chinois. Parmi les races « européennes », la race Créole, qui est indiquée sur le site du BRG ([http://www.brg.prd.fr/brg/pages/rga/porcins/190\\_pag](http://www.brg.prd.fr/brg/pages/rga/porcins/190_pag))

comme « issue de races ibériques puis anglaises (Large Black), américaines (Duroc) et françaises (Normand, Craonnais) », est la lignée la plus originale par rapport aux autres races.

Avec un peu d'étonnement, il apparaît que les méthodes basées sur des fréquences alléliques permettent de produire des résultats qui paraissent sensés pour des lignées analysées sur des nombres d'individus très réduits (6 animaux seulement pour certaines lignées). La deuxième partie de cette étude (2<sup>e</sup> jeu microsatellites et marqueurs SNP) ne concernant que 300 animaux pour 33 races, on pouvait s'attendre à une vision moins claire.

Au-delà des comparaisons entre races, ce programme visait également à porter un premier regard sur la diversité dans certaines régions génomiques contenant des gènes influençant les caractères d'intérêt économique, dont la variabilité était susceptible d'avoir évolué sous l'influence de la sélection animale. Il est clair que dans ce cadre-là le nombre d'animaux analysés est trop faible pour pouvoir produire autre chose qu'une impression qualitative. Ce programme a toutefois permis de mettre en évidence des régions pour lesquelles la sélection des animaux a privilégié certains allèles. Pour certains SNP où les deux allèles sont retrouvés dans tous les types de races, on observe en effet des fréquences très différentes entre races, suggérant une sélection différente dans ces races. Il est à ce titre étonnant de trouver une homogénéité assez forte parmi les races synthétiques mâles contenant du sang Piétrain.

Ce programme mérite donc d'être poursuivi à une échelle plus vaste en analysant l'ensemble des 1 000 animaux sur l'ensemble des marqueurs de ce programme présentant une informativité suffisante. Un prolongement de ce programme sera réalisé dans le cadre du projet Agenae Biomark, cofinancé par le ministère de la recherche et l'association Bioporc représentant la filière porcine.

## REMERCIEMENTS

Les auteurs remercient l'ensemble des sélectionneurs porcins partenaires du projet pour la fourniture de sang ou d'ADN des animaux de leurs lignées. Les auteurs remercient le département de génétique animale pour le financement de la première phase d'analyse des marqueurs microsatellites et la plateforme CRGS de la Génopole Toulouse Midi Pyrénées pour la mise à disposition du matériel d'analyse des SNP.

## RÉFÉRENCES

- [1] Bidanel J.P., Milan D., Iannuccelli N., Amigues Y., Boscher M.Y., Bourgeois F., Caritez J.C., Gruand J., Le Roy P., Lagant H., Quintanilla R., Renard C., Gellin J., Ollivier L., Chevalet C., Detection of quantitative trait loci for growth and fatness in pigs. *Genet Sel Evol.* 33 (2001) p. 289-309.
- [2] Milan D., Bidanel J.P., Iannuccelli N., Riquet J., Amigues Y., Gruand J., Le Roy P., Renard C., Chevalet C., Detection of quantitative trait loci for carcass composition traits in pigs. *Genet Sel Evol.* 34 (2002) p. 705-728.
- [3] Marklund L., Nystrom P.E., Stern S., Andersson-Eklund L., Andersson L., Confirmed quantitative trait loci for fatness and growth on pig chromosome 4. *Heredity.* 82 (Pt 2) (1999) p. 134-141.
- [4] Sanchez M.P., Riquet J., Iannuccelli N., Gogue J., Billon Y., Demeure O., Caritez J.C., Burgaud G., Feve K., Bonnet M., Pery C., Lagant H., Le Roy P., Bidanel J.P., Milan D., Effects of quantitative trait loci on chromosomes 1, 2, 4, and 7 on growth, carcass, and meat quality traits in backcross Meishan x Large White pigs. *J Anim Sci.* 84 (2006) p. 526-537.
- [5] Van Laere A.S., Nguyen M., Braunschweig M., Nezer C., Collette C., Moreau L., Archibald A.L., Haley C.S., Buys N., Tally M., Andersson G., Georges M., Andersson L., A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature.* 425 (2003) p. 832-836.
- [6] Milan D., Jeon J.T., Looft C., Amarger V., Robic A., Thelander M., Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H., Lundstrom K., Reinsch N., Gellin J., Kalm E., Roy P.L., Chardon P., Andersson L., A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science.* 288 (2000) p. 1248-1251.
- [7] Milan D., Demeure O., Laval G., Iannuccelli N., Genet C., Bonnet M., Burgaud G., Riquet J., Gasnier C., Bidanel J.P. Identification de régions du génome répondant à la sélection dans une lignée porcine sino-européenne : la Tai zumu. in 35èmes Journées de la Recherche porcine en France, Paris, 2003.
- [8] Sabeti P.C., Reich D.E., Higgins J.M., Levine H.Z., Richter D.J., Schaffner S.F., Gabriel S.B., Platko J.V., Patterson N.J., McDonald G.J., Ackerman H.C., Campbell S.J., Altshuler D., Cooper R., Kwiatkowski D., Ward R., Lander E.S., Detecting recent positive selection in the human genome from haplotype structure. *Nature.* 419 (2002) p. 832-837.
- [9] Laval G., Iannuccelli N., Legault C., Milan D., Groenen M.A., Giuffra E., Andersson L., Nissen P.H., Jorgensen C.B., Beeckmann P., Geldermann H., Foulley J.L., Chevalet C., Ollivier L., Genetic diversity of eleven European pig breeds. *Genet Sel Evol.* 32 (2000) p. 187-203.
- [10] SanCristobal M., Chevalet C., Haley C.S., Joosten R., Rattink A.P., Harlizius B., Groenen M.A.M., Amigues Y., Boscher M.-Y., Russell G., Law A., Davoli R., Russo V., Désautés C., Alderson L., Fimland E., Bagga M., Delgado J.V., Vega-Pla J.L., Martinez A.M., *et al.*, Genetic diversity within and between European pig breeds using microsatellite markers. *Animal Genetics.* (sous presse) (2006).
- [11] Ousova O., Guyonnet-Duperat V., Iannuccelli N., Bidanel J.P., Milan D., Genet C., Llamas B., Yerle M., Gellin J., Chardon P., Emptoz-Bonneton A., Pugeat M.,

- Mormede P., Moisan M.P., Corticosteroid binding globulin: a new target for cortisol-driven obesity. *Mol Endocrinol.* 18 (2004) p. 1687-1696.
- [12] Hedrick P.W., Gametic disequilibrium measures: proceed with caution. *Genetics.* 117 (1987) p. 331-341.
- [13] Abdallah J.M., Mangin B., Goffinet B., Cierco-Ayrolles C., Perez-Enciso M., A comparison between methods for linkage disequilibrium fine mapping of quantitative trait loci. *Genet Res.* 83 (2004) p. 41-47.
- [14] Kim K.S., Larsen N., Short T., Plastow G., Rothschild M.F., A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mamm Genome.* 11 (2000) p. 131-135.
- [15] Hernandez-Sanchez J., Visscher P., Plastow G., Haley C., Candidate gene analysis for quantitative traits using the transmission disequilibrium test: the example of the melanocortin 4-receptor in pigs. *Genetics.* 164 (2003) p. 637-644.
- [16] Tribout T., Mercat M.J., Le Roy P., Bouffaud M., Barret J., Estimation dans des populations porcines françaises de l'effet du génotype MC4R sur les performances de quelques caractères de croissance, de composition de la carcasse et de qualité de la viande. *Techni Porc.* 28 (2005) p. 11-16.
- [17] Farnir F., Coppieters W., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Mni M., Nezer C., Simon P., Vanmanshoven P., Wagenaar D., Georges M., Extensive genome-wide linkage disequilibrium in cattle. *Genome Res.* 10 (2000) p. 220-227.