

# Analyse de la diversité de quelques races et lignées porcines françaises à l'aide de marqueurs génétiques dans le cadre d'un vaste projet européen

Louis OLLIVIER (1)<sup>1</sup>, Magali SANCRISTOBAL (2), Jean-Louis FOULLEY (1), Christian LEGAULT (1), Guillaume LAVAL (9), Denis MILAN (2), Jean-Claude CARITEZ (3), Germaine BURGAUD (3), Joseph GRUAND (4), Marie-Yvonne BOSCHER (5), Yves AMIGUES (5), Julien BARRET (6), Claire HASSENFRAZT (6), Michel LUGUET (7), Florence LABROUE (8), Claude CHEVALET (2)

(1) INRA-Station de Génétique Quantitative et Appliquée, 78352 Jouy-en-Josas cedex

(2) INRA-Laboratoire de Génétique Cellulaire, BP 27, 31326 Castanet Tolosan cedex

(3) INRA-Domaine Pluridisciplinaire du Magneraud, 17700 Surgères

(4) INRA-Station Expérimentale de Sélection Porcine, 86480 Rouillé

(5) LABOGENA, Centre de Recherches de Jouy, 78352 Jouy-en-Josas cedex

(6) Agence de la Sélection Porcine, 149 rue de Bercy, 75595 Paris cedex 12

(7) Institut Technique du Porc, la Motte-au-Vicomte, 35650 Le Rheu cedex

(8) Chambre d'Agriculture, rue de la Géraudière, 44939 Nantes cedex

(9) Computational and Molecular Population Genetics Laboratory, Baltzerstrasse 6, 3012 Bern, Suisse

## Analyse de la diversité de quelques races et lignées porcines françaises à l'aide de marqueurs génétiques dans le cadre d'un vaste projet européen

Un échantillon de seize races (et lignées) porcines françaises a été étudié dans le cadre d'un projet européen de diversité génétique financé par l'Union Européenne et portant sur soixante-dix races, avec examen individuel à raison de cinquante individus par race environ. Les races françaises se répartissaient en cinq races locales, quatre variétés des grandes races *Large White*, *Landrace* et *Piétrain*, et enfin sept lignées commerciales de race pure ou synthétiques. Les marqueurs génétiques considérés étaient d'une part des microsatellites (50 locus), et d'autre part des AFLP (148 locus), qui sont un polymorphisme de longueur de fragments d'ADN amplifiés. Une grande variabilité intra-race est observée pour les microsatellites. L'hétérozygotie moyenne de 0,56 et le nombre moyen d'allèles de 4,6 par locus chez les races françaises sont très proches des moyennes de l'étude européenne. L'arbre phylogénétique des 70 races révèle des regroupements de la plupart des variétés et lignées autour de leur race de référence (*Large White*, *Landrace*, *Piétrain*, *Duroc*, *Hampshire*). Les races et lignées françaises s'insèrent dans ces regroupements d'une manière conforme à l'attente. L'analyse de la diversité des races européennes montre que les races locales contribuent pour 56 % à la diversité totale des microsatellites, et pour un peu moins à celle des AFLP. L'originalité des races locales françaises, notamment des porcs *Basque* et *Limousin*, vient confirmer l'étude pilote de LAVAL et al (2000). Une contribution notable est également apportée par la lignée synthétique *Tia Meslan*, connue pour inclure des gènes en provenance de deux races chinoises. L'article discute des enseignements qui peuvent être tirés de cette étude pour la conservation des races et la gestion de la variabilité génétique de l'espèce porcine, autour de la question centrale qui est de savoir dans quelle mesure la diversité des marqueurs génétiques utilisés reflète une variabilité génétique utile.

## Analysis of the diversity of some French breeds and lines of pigs using genetic markers in the framework of a wide European survey

A sample of sixteen French breeds (and lines) of pigs was investigated in the framework of an EU-funded diversity project including seventy breeds, about fifty individuals per breed having been individually examined. The French sample was made of five local breeds, four French varieties of major international breeds such as *Large White*, *Landrace*, and *Piétrain*, and seven commercial lines either purebred or synthetic. The genetic markers were microsatellites (50 loci) and AFLP (148 loci), the latter being a length polymorphism of amplified DNA fragments. A large within-breed variability was observed in microsatellites. The average within-breed expected heterozygosity (0.56) and number of alleles per locus (4.6) in the French breeds were close to the European averages. The phylogenetic tree drawn from the Reynolds distances among the seventy breeds showed that the national varieties of major breeds and the commercial lines were mostly clustered around their breeds of reference (*Large White*, *Landrace*, *Piétrain*, *Duroc* and *Hampshire*). The French breeds and lines clustered in accordance with expectation. The analysis of diversity showed that the local breeds accounted for 56 % of the total European microsatellite diversity, and slightly less for AFLP. The genetic originality of the French local breeds, particularly *Basque* and *Limousin*, confirmed the pilot study of LAVAL et al (2000). The important diversity contribution of the *Tia Meslan* synthetic line was noted, this line being known to be hosting genes from two Chinese breeds. The paper discusses what can be learned from this study and applied to the conservation of breeds and management of genetic variability in the pig, around the central question which is to know to what extent the marker diversity reflects some useful genetic variability.

<sup>1</sup> Détaché auprès de l'INRA par la Fédération Européenne de Zootechnie (Via Nomentana 134, 00100 Rome, Italie)

## INTRODUCTION

Le concept de biodiversité est aujourd'hui au cœur de préoccupations qui s'expriment dans tous les pays du monde. Pour ce qui est des animaux, et en particulier du porc, on sait que les pressions économiques ont conduit à mettre l'accent sur un nombre relativement limité de lignées spécialisées. Beaucoup de races moins bien adaptées aux besoins actuels ont ainsi vu leur effectif décroître au point de mettre en danger leur survie même. La conservation de la diversité que représentent ces races est cependant nécessaire pour pouvoir satisfaire les besoins alimentaires du futur et aussi pour sauvegarder le riche patrimoine agricole des diverses régions du monde. Un survol historique des ressources génétiques animales et des inquiétudes que soulève leur avenir prévisible, aussi bien dans le monde développé que dans les pays en voie de développement, a été présenté par BARKER (2002). Une évaluation aussi précise que possible des ressources disponibles pour faire évoluer chaque espèce dans les directions souhaitées est donc d'abord nécessaire. Plusieurs actions dans ce sens ont été menées récemment en Europe dans l'espèce porcine (OLLIVIER et al, 2003). L'objet de cet article est d'en résumer les principaux enseignements. Il faut aussi souligner l'importante participation française à ces actions financées par l'Union Européenne. L'accent sera donc mis sur les races et lignées françaises évaluées, dont les résultats seront placés à la fois dans une perspective nationale et une plus large perspective européenne. Des résultats plus détaillés font actuellement l'objet de publications en préparation (SANCRISTOBAL et al, 2004 ; PLASTOW et al, 2004 ; OLLIVIER et al, 2004).

## 1. MATÉRIELS ET MÉTHODES

### 1.1. Les races

L'étude porte sur un total de 70 races ou lignées porcines, au rang desquelles figurent 16 races françaises. L'ensemble représente 68 races/lignées domestiques strictement européennes, auxquelles s'ajoutent la race chinoise *Meishan* (partiellement échantillonnée à l'INRA) et un échantillon de sanglier européen. Comme indiqué au tableau 1, ce vaste ensemble recouvre des entités de nature assez diverse, qui, pour le porc domestique strictement européen, ont été classées en trois catégories : 29 races locales, 18 variétés nationales de « grandes races », et 21 lignées commerciales, de race pure ou composites, et généralement dérivées de la catégorie précédente.

### 1.2. Les marqueurs

Deux techniques de marquage ont été mises en œuvre, d'une part les microsatellites (répétition d'une séquence ADN courte) et d'autre part les AFLP (polymorphisme de longueur de fragments d'ADN obtenus par amplification). Ces techniques sont décrites en détail par PITEL et RIQUET (2000). Les caractéristiques générales de ces marqueurs sont très différentes. Les microsatellites sont des marqueurs multi-alléliques, co-dominants et hautement polymorphes, contrairement aux AFLP qui sont bi-alléliques, dominants et faiblement polymorphes.

Le nombre de locus typés est de 50 pour les microsatellites. Ces locus ont été choisis pour leur polymorphisme et de manière à couvrir l'ensemble du génome. Le choix a égale-

**Tableau 1** - Races échantillonnées

PAYS d'origine	Catégorie de population			Autre	Total	
	Race locale	Race internationale	Lignée commerciale		Population	Echantillon d'ADN
<b>EUROPE</b>						
Allemagne	4	3	4	-	11	544
Belgique	-	1	-	-	1	46
Danemark	1	2	-	-	3	132
Espagne	4	-	-	-	4	170
Finlande	-	1	-	-	1	56
France	5	4	7	-	16	782
Islande	-	1	-	-	1	35
Italie	4	3	-	-	7	277
Norvège	-	1	-	-	1	50
Pays-Bas	-	1	-	-	1	30
Pologne	1	-	-	1 <sup>1</sup>	2	60
Portugal	1	-	-	-	1	60
République Tchèque	1	-	-	-	1	50
Royaume-Uni	7	-	-	-	7	312
Suède	1	1	-	-	2	60
<b>CHINE</b>	-	-	-	1 <sup>2</sup>	1	61
<b>INTERNATIONAL (PIC)</b>	-	-	10	-	10	498
<b>TOTAL</b>	<b>29</b>	<b>18</b>	<b>21</b>	<b>2</b>	<b>70</b>	<b>3223</b>

<sup>1</sup> Sanglier

<sup>2</sup> Meishan

ment tenu compte de l'adaptation des marqueurs à des typages en série automatisés. Les 50 marqueurs retenus assurent la couverture des 19 chromosomes du porc, avec deux ou trois marqueurs par chromosome (voir GROENEN et al, 2003 et tableau 3 plus loin). Le travail a été partagé entre LABOGENA (pour 34 locus, sur séquenceur ABI 3700 multi-capillaire) et l'Université de Wageningen (pour 16 locus, sur séquenceur ABI 373).

Contrairement aux microsatellites, la position des locus AFLP sur le génome du porc n'est pas connue. Cette imprécision est cependant compensée par la possibilité de typer un grand nombre de locus. Dans cette étude, 148 locus AFLP ont été étudiés. Le travail a été réalisé par la firme PIC (Pig Improvement Company) avec l'aide de l'établissement néerlandais Keygene (PLASTOW et al, 2003).

### 1.3. L'analyse génétique

L'analyse a d'abord porté sur la variabilité génétique interne à chaque race par les méthodes de RAYMOND et ROUSSET

(1995) et le logiciel Genetix de BELKHIR et al (1998). La différenciation entre races a été évaluée par les indices de fixation de Wright, calculés selon WEIR et COCKERHAM (1984). Les distances génétiques entre races ont été estimées en utilisant les distances classiques de Reynolds (REYNOLDS et al, 1983) et de Nei standard (NEI, 1972). Deux utilisations principales des distances obtenues seront présentées : d'une part la construction d'arbres phylogénétiques (HARTL et CLARK, 1997) et d'autre part l'analyse de la diversité selon la méthode proposée par WEITZMAN (1993). La mise en œuvre de cette méthode devient très lourde dès que le nombre de populations dépasse 30. Un logiciel a été spécifiquement écrit pour y remédier (DERBAN et al, 2002).

## 2. RÉSULTATS

### 2.1. Diversité intra-race

Les résultats des analyses intra-race sont présentés au tableau 2 pour les microsatellites. Les niveaux d'hétérozygotie varient de 0,4 pour le porc *Basque* à 0,7 pour le porc

**Tableau 2** - Polymorphisme intra-race des microsatellites

(N : nombre d'échantillons d'ADN, n : nombre de locus, H : hétérozygotie attendue, Ao : nombre d'allèles observés, Ap : nombre d'allèles « privés »\*,  $F_{IS}$  : indice de fixation intra-race de WEIR et COCKERHAM, 1984)

RACE <sup>1</sup>	N	n	H	Ao	Ap	$F_{IS}$ et test de Hardy-Weiberg <sup>2</sup>
FRBA01	47	26	0,391	3,35	0	-0,03 ns
FRCR01	44	50	0,684	7,16	16	0,10**
FRGA01	56	26	0,520	4,27	0	0,10**
FRLI01	56	26	0,470	3,58	1	0,02 ns
FRNO01	52	26	0,526	4,58	1	-0,03 ns
FRLR01	51	50	0,610	5,62	0	0,00 ns
FRLW01	50	50	0,607	5,12	1	0,02 ns
FRLW12	34	50	0,619	4,68	0	0,00 ns
FRPI02	50	50	0,580	4,98	2	0,00 ns
FRDR01	50	50	0,606	4,78	1	0,01 ns
FRLA01	49	50	0,548	4,04	1	0,02 ns
FRLR13	51	50	0,579	4,32	0	0,00 ns
FRLW08	50	50	0,529	4,14	2	-0,03 ns
FRLW09	46	50	0,561	4,32	3	-0,03 ns
FRPI05	48	50	0,516	3,94	0	0,01 ns
FRTM01	49	49	0,647	5,00	3	0,04 ns
Moyenne française	48,9	43,9	0,562	4,62	1,9	0,01
Moyenne générale	46,0	45,5	0,556	4,49	1,7	0,02

\* Allèles privés : comptés sur les 70 races de l'étude européenne

<sup>1</sup> Code race :

Races locales : BA Basque, CR Créole, GA Gascon, LI Limousin, NO Normand (ou Blanc de l'Ouest)

Races internationales : LR01 Landrace français, LW01 Large White français « femelle », LW12 Large White français « mâle », PI02 Piétrain français

Lignées commerciales : DR01 DRB (lignée synthétique SCAPAAG), LA01 Laconie (lignée synthétique Pen Ar Lan), LR13 FH012 (lignée Landrace France Hybrides), LW08 FH025 (lignée Large White France Hybrides), LW09 Gallia (lignée Large White Pen Ar Lan), PI05 FH016 (lignée Piétrain France Hybrides), TM01 Tia Meslan (lignée synthétique Pen Ar Lan)

<sup>2</sup> ns : non significatif

\*\* : significatif au seuil de 1%

**Tableau 3** - Indices de fixation F par locus selon WEIR et COCKERHAM (1984) (N : nombre de populations, A nombre d'allèles,  $F_{IS}$  et  $F_{ST}$  définis dans le texte)

Chromosome	Locus	N	A	$F_{IS}$	$F_{ST}$
1	CGA	59	42	0,00	0,16
	S0155	70	9	-0,00	0,23
	SW1828	59	13	-0,01	0,23
2	SW240	70	16	0,04	0,17
	S0226	70	19	-0,01	0,24
3	SW72	68	10	0,04	0,25
	SW902	59	19	0,03	0,28
	S0002	70	18	0,01	0,22
4	S0227	70	10	0,04	0,18
	S0301	54	8	0,02	0,16
	S0217	59	9	-0,01	0,20
	S0097	57	22	-0,01	0,20
5	S0005	70	30	0,01	0,18
	IGF1	70	13	-0,01	0,22
6	SW2406	59	15	0,04	0,19
	SW1067	59	18	0,02	0,25
	SW122	70	10	-0,01	0,21
	S0228	68	15	-0,01	0,24
7	S0025	54	13	0,01	0,17
	SW632	70	15	0,03	0,25
	S0101	68	13	0,01	0,20
8	SW2410	58	15	0,02	0,21
	S0225	70	15	0,04	0,23
	S0178	70	13	0,00	0,24
9	SW911	70	12	0,02	0,24
	SW174	54	7	-0,01	0,34
10	SW830	59	17	-0,00	0,21
	S0070	59	25	-0,01	0,24
	SW951	70	8	0,08	0,23
11	SW2008	59	10	0,04	0,22
	S0386	68	10	0,19	0,26
12	S0143	59	12	-0,02	0,16
	S0090	70	10	-0,00	0,22
13	SWR1941	59	11	0,00	0,18
	S0068	68	21	0,01	0,23
	SW769	59	17	0,01	0,21
	S0215	70	15	0,04	0,42
14	SW857	70	10	0,01	0,20
	SW295	58	16	0,00	0,23
15	S0355	68	12	-0,00	0,23
	SW1111	56	15	0,03	0,23
	SW936	68	12	0,01	0,18
16	SW742	58	21	-0,03	0,25
	S0026	69	9	-0,00	0,26
17	SWR1004	54	13	0,01	0,20
	SW24	70	16	0,00	0,25
18	SW1023	55	11	0,02	0,18
	SW787	57	10	-0,02	0,20
X	SW2476	56	11	0,06	0,25
	S0218	70	14	0,02	0,36
Moyenne		63,7	14,50	0,02	0,23

*Créole*. La moyenne des races françaises est assez proche de la moyenne de l'ensemble des races étudiées. On note aussi une tendance à une hétérozygotie et un nombre d'allèles

moindres dans les races locales (si on excepte le porc Créole) par rapport aux autres races. Les marqueurs AFLP sont nettement moins polymorphes avec une hétérozygotie moyenne de 0,11 et un nombre moyen d'allèles de 1,4 (non montré).

L'équilibre de Hardy-Weinberg est vérifié sur les microsatellites à l'exception de deux races locales (*Créole* et *Gascon*). Les valeurs positives du paramètre dans ces deux races suggèrent la possibilité d'homogamie ou de consanguinité. Sept autres races locales européennes sont dans le même cas. Notons que les AFLP ne permettent pas d'effectuer le test de Hardy-Weinberg puisque les hétérozygotes ne sont pas identifiés.

## 2.2. Indices de fixation

Les deux indices de fixation ( $F_{IS}$  et  $F_{ST}$ ) du tableau 3 expriment respectivement la consanguinité intra-race et la différenciation des races, prenant en compte les 70 races de l'étude européenne. La valeur moyenne très faible de  $F_{IS}$  indique que sur l'ensemble des races et des marqueurs l'hypothèse de Hardy-Weinberg peut être acceptée, et ce malgré les exceptions relevées ci-dessus. Notons cependant une valeur exceptionnellement élevée de  $F_{IS}$  pour le marqueur S0386. Ceci correspond à un déficit d'hétérozygotes qui ne serait pas dû à la consanguinité mais à un artefact de la méthode de typage conduisant à la non détection de certains allèles, dits « nuls ». Les indices  $F_{ST}$  du tableau 3 montrent une forte individualisation des races étudiées. On relève également une variation assez forte de ce paramètre en fonction du marqueur, puisqu'il va de 0,156 (S0143) à 0,423 (S0215). Un autre indice d'originalité génétique est le nombre d'allèles présents dans une race à l'exclusion de toute autre race, ce qu'on appelle les allèles « privés ». Leur nombre est indiqué pour chaque race au tableau 2.

## 2.3. Distances génétiques

Le tableau 4 montre une grande similitude, en valeur relative, entre les distances obtenues avec les microsatellites par les deux méthodes, Reynolds et Nei standard. On voit que cette dernière est à peu près le double de la première. Sur l'ensemble des races européennes, les deux distances sont en effet liées par une relation approximativement linéaire (voir SANCRISTOBAL et al, 2002, pour 58 races). Dans l'ensemble, c'est la race *Basque* qui est la plus éloignée de toutes les autres races françaises (distances Reynolds allant de 0,26 à 0,40), suivie de la lignée *Tia Meslan* (distances Reynolds allant de 0,18 à 0,35). On voit aussi que les diverses variétés françaises de *Large White*, *Landrace* et *Piétrain* sont assez voisines de leurs homologues européennes. Certains voisinages de nos races locales ou lignées composites sont cependant inattendus. On notera en particulier la faible distance séparant deux races appartenant à des îles géographiquement très éloignées, le *Créole* de Guadeloupe et le *Noir de Sicile*. Les distances AFLP (non montrées) confirment les résultats microsatellites pour les races proches, mais peuvent être très différentes pour les races éloignées entre elles. Cependant, la race la plus proche du *Créole* pour les AFLP est la race locale tchèque *Presticke*, avec Reynolds aussi bien qu'avec Nei standard.

**Tableau 4** - Distances génétiques entre les 16 races et lignées françaises calculées selon FELSENSTEIN (2000). En gras : distances françaises les plus grandes ; en italique : distances les plus faibles. Au-dessus de la diagonale : distance de Reynolds - Au-dessous de la diagonale : distance standard de Nei

RACE <sup>1</sup>	FRLW01	FRLW08	FRLW09	FRLW12	FRLR01	FRLR13	FRPI02	FRPI05	FRLA01	FRTM01	FRDR01	FRCR01	FRBA01	FRGA01	FRLI01	FRNO01	<sup>2</sup>
FRLW01	-	0,09	0,07	0,03	0,16	0,2	0,18	0,18	0,18	0,22	0,14	0,14	0,33	0,18	0,25	0,26	DELW02 (0,06)
FRLW08	0,14	-	0,09	0,09	0,19	0,22	0,21	0,23	0,22	0,25	0,17	0,18	0,36	0,24	0,29	0,29	DELW02 (0,10)
FRLW09	0,11	0,13	-	0,06	0,18	0,21	0,19	0,2	0,2	0,24	0,15	0,15	0,31	0,19	0,26	0,25	DELW02 (0,06)
FRLW12	0,05	0,13	0,09	-	0,15	0,18	0,18	0,18	0,18	0,2	0,14	0,14	0,3	0,19	0,23	0,24	DELW02 (0,04)
FRLR01	0,33	0,34	0,35	0,32	-	0,1	0,16	0,21	0,21	0,23	0,16	0,13	0,32	0,21	0,22	0,15	ITLR03 (0,04)
FRLR13	0,42	0,42	0,42	0,39	0,17	-	0,19	0,23	0,23	0,24	0,19	0,15	0,35	0,24	0,24	0,2	ITLR03 (0,09)
FRPI02	0,39	0,38	0,36	0,37	0,31	0,38	-	0,1	0,19	0,23	0,17	0,13	0,32	0,19	0,25	0,21	DEPI03 (0,10)
FRPI05	0,32	0,38	0,32	0,32	0,39	0,43	0,13	-	0,21	0,27	0,21	0,17	0,37	0,24	0,3	0,27	DEPI03 (0,10)
FRLA01	0,34	0,39	0,36	0,34	0,42	0,47	0,34	0,34	-	0,22	0,2	0,16	0,34	0,19	0,29	0,26	BEPI01 (0,15)
FRTM01	0,61	0,6	0,59	0,59	0,64	0,66	0,59	0,67	0,52	-	0,21	0,18	0,35	0,25	0,29	0,28	CZPR01 (0,18)
FRDR01	0,27	0,3	0,28	0,28	0,33	0,39	0,35	0,39	0,4	0,56	-	0,11	0,3	0,19	0,24	0,19	DELW02 (0,12)
FRCR01	0,34	0,38	0,33	0,33	0,29	0,33	0,27	0,33	0,34	0,54	0,25	-	0,26	0,15	0,18	0,15	ITNS01 (0,08)
FRBA01	0,58	0,59	0,48	0,49	0,57	0,63	0,52	0,57	0,54	<b>0,73</b>	0,53	0,46	-	0,29	<b>0,40</b>	0,33	CZPR01 (0,27)
FRGA01	0,32	0,41	0,3	0,34	0,4	0,46	0,3	0,37	0,27	0,55	0,36	0,28	0,4	-	0,29	0,25	CZPR01 (0,16)
FRLI01	0,44	0,49	0,44	0,39	0,37	0,39	0,43	0,48	0,45	0,64	0,45	0,31	0,67	0,48	-	0,26	CZPR01 (0,19)
FRNO01	0,54	0,58	0,46	0,5	0,25	0,34	0,37	0,46	0,46	0,7	0,37	0,29	0,5	0,45	0,41	-	ITNS01 (0,14)
<sup>2</sup>	DELW02 (0,09)	DELW02 (0,14)	DELW02 (0,09)	DELW02 (0,07)	ITLR03 (0,07)	SELR07 (0,11)	DEPI03 (0,08)	DEPI03 (0,13)	BEPI01 (0,22)	CZPR01 (0,47)	ITDU01 (0,21)	ITNS01 (0,18)	GBLW06 (0,31)	GBLW06 (0,23)	SEWPO1 (0,27)	SELR07 (0,17)	

<sup>1</sup> Code race : Races françaises : voir tableau 2 - Races européennes : Duroc italien ITDU01, Landrace italien ITLR03, Landrace suédois SELR07, Large White allemand DELW02, Large White de la firme Pig Improvement Company GBLW06, Noir de Sicile ITNS, Piétrain allemand DEPI03, Piétrain belge BEPI01, Presticke de République tchèque CZPR01, Sanglier européen SEWPO1

<sup>2</sup> Race européenne la plus proche, (distance entre parenthèses)

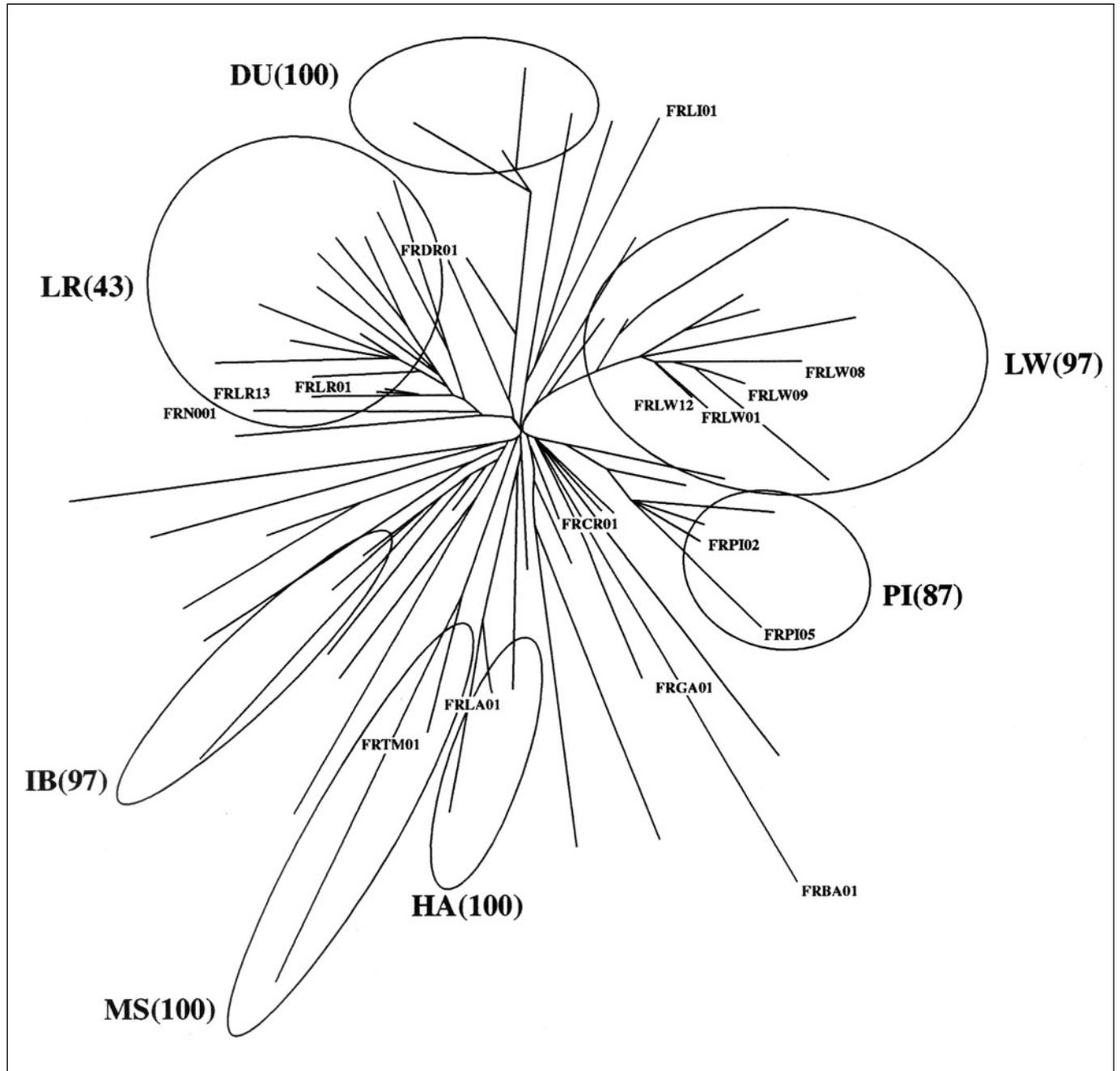
## 2.4. Arbres phylogénétiques

La figure 1 présente l'arbre des 70 races établi par la méthode de Jonction des Voisins, également utilisée par LAVAL et al (2000a et b). On peut voir que les races des catégories 2 et 3 du tableau 1 (races internationales et lignées commerciales de race pure) se regroupent autour de leur race de référence (*Large White*, *Landrace*, *Piétrain*, *Duroc* et *Hampshire*), alors que les races locales montrent plutôt une topologie dite "en étoile", confirmant celle montrée par les auteurs précédents sur un nombre de races plus

restreint. On voit aussi que les races et lignées de race pure françaises se regroupent logiquement autour des bouquets *Large White*, *Landrace* et *Piétrain*. Les lignées synthétiques *DRB*, *Laconie* et *Tia Meslan* se regroupent avec respectivement *Duroc*, *Hampshire* et *Meishan*, en accord aussi avec ce que l'on sait de leur composition raciale d'origine.

## 2.5. Analyse de la diversité

La méthode de Weitzman (WEITZMAN, 1993 ; THAON d'ARNOLDI et al, 1998) permet d'analyser la diversité glo-



**Figure 1** - Arbre phylogénétique des 70 races de l'étude européenne, basé sur les distances Reynolds sur microsatellites, et établi selon l'algorithme NJ, ou « jonction des voisins », décrit par HARTL et CLARK (1997, pages 370-371)

Les 16 races et lignées françaises sont identifiées à leur place dans l'arbre (voir code race du tableau 2)

Les chiffres indiqués entre parenthèses expriment la fiabilité (entre 0 et 100) à accorder à chacun des nœuds correspondants :

DU Duroc, HA Hampshire, IB porc ibérique, LR Landrace, LW Large White, MS Meishan, PI Piétrain

(valeurs calculées sur l'arbre des 59 races de SANCRISTOBAL et al, 2002)

bale observée et d'estimer les contributions à cette diversité globale de chacune des races étudiées ou de groupes de races définis. Quand on considère les trois catégories de races définies au tableau 1, une telle analyse conduit à conclure que plus de la moitié de la diversité totale est fournie par les 29 races locales (voir tableau 5). Selon le même principe, on établit aussi que les 16 races et lignées françaises contribuent pour près de 25 % à la diversité européenne totale (résultat non montré). La diversité globale peut également être analysée à l'intérieur de groupes définis, tels que par exemple les différents pays. Une telle analyse pour la France montre une distribution de la diversité entre les trois catégories assez voisine de celle observée sur l'ensemble du continent européen (voir plus loin le tableau 6). Des résultats similaires (non montrés) sont obtenus avec les marqueurs AFLP, qui donnent cependant une part un peu moindre aux races locales dans la diversité totale.

**Tableau 5** - Contribution des groupes à la diversité des microsatellites

Groupe	Nombre de races	Contribution à la diversité raciale <sup>1</sup> (%)	Hétérozygotie moyenne par race (%)
Races locales	29	<b>55,9</b>	54,1
Races internationales	18	15,4	<b>58,7</b>
Lignées commerciales	21	28,7	54,9
Total (ou moyenne)	68	100,0	55,6

<sup>1</sup> En utilisant la fonction de diversité de WEITZMAN (1993) appliquée aux distances de Reynolds  
En gras contribution la plus importante

### 3. DISCUSSION

On trouve dans la littérature de nombreuses études de diversité consacrées aux espèces animales d'élevage, dont le porc. Cependant la plupart de ces études ont été jusqu'à présent assez limitées, autant pour ce qui est du nombre de races échantillonnées que de celui des marqueurs génétiques utilisés (VAN ZEVEREN et al, 1995 ; LI et al, 2000 ; MARTINEZ et al 2000 ; LAVAL et al, 2000a et b ; OLLIVIER et al, 2001b ; SUN et al, 2002). Les actions menées sous l'égide de l'Union Européenne entre 1995 et 2000 (voir OLLIVIER et al, 2003) ont permis de recueillir un ensemble de données probablement unique au monde à ce jour, tant par le nombre des marqueurs utilisés (près de 200 locus différents au total) que par celui des races étudiées (71 en incluant la race italienne *Mora Romagnola* qui n'a été typée que pour un nombre limité de microsatellites sur un mélange d'ADN). Au total, près d'un demi-million de génotypes ont été ainsi générés. Notons que cette information est du domaine public et notamment disponible sur le serveur <http://www.project.roslin.ac.uk/pigbiodiv/>.

L'énorme quantité d'information collectée soulève de nombreuses difficultés d'interprétation. Partant de l'information essentielle, qui est constituée des génotypes individuels à chaque locus pour chacune des races étudiées, des fréquences alléliques sont estimées pour résumer l'informa-

tion entre races, au prix d'une certaine perte d'information sur la variabilité intra-race. La première question qui se pose est celle du choix d'une mesure de distance. On sait qu'un très grand nombre de distances ont été proposées dans la littérature et que le choix de la distance la plus adéquate selon les marqueurs et les populations a donné lieu également à une abondante littérature (LAVAL et al, 2002). Les distances génétiques de Reynolds et de Nei standard ont des valeurs attendues facilement calculables sous le modèle de dérive (fluctuation aléatoire des fréquences alléliques dans une population d'effectif limité) d'une part et sous le modèle d'équilibre entre dérive et mutation d'autre part.

Même si l'on peut supposer que la dérive est un des facteurs ayant eu une grande influence sur la différenciation des races porcines européennes, il est néanmoins évident que les migrations entre populations ont également joué un rôle dans la constitution des races actuelles, mais elles sont malheureusement difficiles à quantifier.

D'autre part, la sélection a joué un rôle important dans l'évolution des races. A travers des marqueurs neutres, son effet le plus facilement observable est la réduction de l'effectif génétique ( $N$ ) des populations. En supposant le modèle de dérive, la distance de Reynolds ( $D_R$ ) est en effet approximativement égale à  $t/2N$ , où  $t$  est le nombre de générations de divergence entre les deux populations. Si on connaît  $t$ , la distance de Reynolds fournit une indication sur les effectifs génétiques  $N$ . Alternativement, connaissant  $N$ , on peut estimer  $t$ . Par contre, si un marqueur est associé à un gène sélectionné et ainsi indirectement sélectionné, les écarts de fréquences peuvent dépasser ce qui est attendu sous l'hypothèse de dérive pure. On sous-estimera alors  $N$  et/ou on surestimera  $t$ .

Compte tenu des temps de divergence considérés, on peut admettre que les mutations ont eu peu d'influence sur l'évolution des fréquences alléliques, malgré le taux élevé de mutation des microsatellites. De toute façon, les mécanismes de mutation touchant les microsatellites ne permettent pas d'appliquer les relations classiques de Nei prédisant une linéarité entre distance de Nei et temps de divergence.

L'exploitation des distances calculées requiert aussi l'attention. Le nombre même des distances calculées, qui est 8252 dans cette étude en totalisant les deux distances calculées pour les deux types de marqueur, appelle des schémas d'analyse aptes à résumer au mieux le volume d'information disponible. Essentiellement trois types d'analyse sont envisageables : (1) comparaisons entre types de distance et entre types de marqueur, (2) établissement d'arbres phylogénétiques et (3) analyse de la diversité raciale. Ces analyses ont été mises en œuvre dans cette étude et les enseignements qui s'en dégagent sont résumés ci-après.

(1) Les **comparaisons entre la distance  $D_R$  et la distance  $D_S$**  montrent que les distances observées entre la race *Meishan* et les races européennes ne peuvent être expliquées par la dérive génétique et que l'effet des mutations doit être invoqué (SANCRISTOBAL et al,

2002). Les différences observées entre les **distances Reynolds microsatellites et Reynolds AFLP** (voir plus haut) résultent des particularités des marqueurs AFLP, qui sont moins polymorphes et donc moins informatifs. De ce fait, leur analyse requiert aussi des précautions particulières (SANCRISTOBAL et al, 2002 ; OLLIVIER et FOULLEY, 2004).

(2) **L'arbre « phylogénétique »** obtenu (fig. 1) ne fait que résumer, visuellement, les 2065 distances Reynolds calculées entre nos 70 races. C'est un arbre de classification. Il montre que la structure de la population porcine européenne, vue à travers les microsatellites, reflète assez bien le fond génétique et l'histoire de nos races. Mais cet arbre a aussi l'avantage de mettre en lumière des topologies inattendues, comme la position d'un *Landrace* scandinave dans le groupe *Large White*. La longueur des branches traduit par ailleurs une diversité dont les causes peuvent être, soit un effectif génétique restreint ayant entraîné une perte d'allèles (3,4 allèles observés en moyenne chez la race *Basque*), soit l'existence d'allèles originaux (la race *Meishan*), soit plus généralement des différences de fréquences alléliques (voir ci-dessous). Les rééchantillonnages effectués montrent également que l'ensemble des *Large White* européens est regroupé d'une manière plus fiable que celui des *Landrace* (fiabilités respectives 97 et 43), ce qui pourrait traduire une plus grande homogénéité génétique des premiers, que tendent à confirmer les distances internes aux deux groupes.

(3) Enfin **l'analyse de la diversité raciale** est probablement l'information la plus utile qui puisse être dérivée d'un ensemble de distances dans un contexte de conservation. La question centrale est ici : quelle diversité appréhendons-nous avec les marqueurs génétiques ? Il faut se garder de conclure de la neutralité des marqueurs *per se* au caractère neutre de la diversité qui est mesurée. On sait en effet que des gènes neutres voient leur comportement influencé par la sélection appliquée à des gènes voisins, à travers ce qu'on appelle l'effet d'entraînement. Ce phénomène génétique est en fait à la base de la sélection assistée par marqueur (voir, par exemple, OLLIVIER, 2002). Il faut donc s'attendre à une corrélation entre la diversité des marqueurs et celle des gènes de caractères quantitatifs (QTL), surtout si les marqueurs sont proches de gènes sélectionnés, bien que nous manquions d'information sur l'importance de ces associations. La diversité des marqueurs nous fournit ainsi une précieuse information pour guider les politiques de conservation. On peut aussi combiner les deux diversités, entre races et intra-race par diverses méthodes, comme celle proposée par exemple par OLLIVIER et FOULLEY (2002). Le tableau 6 montre ainsi que si le porc *Basque* mérite une attention particulière pour la conservation de la variabilité entre races (voir  $D_2$ ), une lignée synthétique comme la *Tia Meslan* est prioritaire pour maximiser la variabilité génétique d'une population qui combinerait les races et lignées françaises existantes (voir  $D_1$ ). Ceci n'est pas a priori surprenant, sachant que cette lignée inclut des gènes de deux races

chinoises, *Meishan* et *Jiaying*, mais montre quand même que cette lignée a probablement gardé, depuis sa création, une bonne part de sa diversité initiale.

Il faut préciser que les chiffres de diversité du tableau 5 sont basés sur des distances entre populations. En fait, la fonction de diversité de Weitzman est une somme de distances et la diversité ainsi mesurée est le « reflet » d'un ensemble de distances entre des populations plus ou moins éloignées les unes des autres (ou « originales »). Dans des études de phylogénie entre espèces, les temps de divergence et les tailles de populations sont très grands. Les distances utilisées (Nei) mesurent les nouveautés apportées par les mutations et permettent donc d'appréhender une très large diversité génétique. Notre étude se place dans un cadre différent constitué de populations d'une même espèce et d'une même zone géographique. Les temps de divergence sont beaucoup plus courts, et les tailles de populations généralement plus réduites. Pour l'essentiel toutes les populations partagent les mêmes allèles (voir tableau 2). Une grande distance entre deux races peut être due à la présence d'allèles originaux (ou « privés ») dans l'une des deux races - on parlera alors de *diversité allélique* - ou à des fortes différences de fréquences alléliques pour des allèles communs, et on parlera alors de *diversité structurelle*. Ainsi, à l'extrême, une population homozygote (consanguine) ne possédant que des allèles présents dans les autres populations se trouvera très éloignée des autres populations, et contribuera ainsi

**Tableau 6** - Contributions relatives (%) des 16 races françaises à la diversité entre races (CE) et intra-race (CI) et à la diversité globale des microsatellites ( $D_1 = 0,23 \text{ CE} + 0,77 \text{ CI}$  et  $D_2 = 0,83 \text{ CE} + 0,17 \text{ CI}$ )

Race <sup>1</sup>	CE	CI	D1	D2
FRCR01	4,10	<b>1,45</b>	2,06	3,65
FRBA01	<b>14,41</b>	-2,03	1,75	<b>11,62</b>
FRGA01	8,05	-0,50	1,46	6,60
FRLI01	11,38	-1,10	1,77	9,26
FRNO01	8,27	-0,43	1,57	6,79
FRLW01	1,74	0,53	0,81	1,53
FRLW12	1,20	0,67	0,79	1,11
FRLR01	3,65	0,57	1,28	3,13
FRPI02	3,46	0,22	0,96	2,91
FRDR01	4,83	0,52	1,51	4,10
FRLA01	6,75	-0,17	1,42	5,57
FRLR13	5,20	0,21	1,35	4,35
FRLW08	5,45	-0,39	0,95	4,46
FRLW09	2,72	-0,01	0,62	2,26
FRPI05	5,98	-0,55	0,95	4,87
FRTM01	9,83	1,01	<b>3,04</b>	8,33

<sup>1</sup> Code race : voir tableau 2

En gras contribution la plus forte

En italique contribution la plus faible

fortement à la diversité. Dans l'interprétation de notre analyse de diversité, il est difficile de faire la part qui revient aux deux diversités, allélique et structurelle, qui sont d'ailleurs en partie confondues.

Notons bien que seule la diversité structurelle est considérée au tableau 5, mais que l'originalité allélique de chaque race est cependant évoquée au tableau 2 avec le décompte des allèles « privés ». Ce critère est cependant peu discriminant, car les valeurs sont faibles en moyenne, et il dépend par ailleurs de la taille des échantillons examinés dans chaque race, puisque le nombre d'allèles recensés dans une race ne peut qu'augmenter avec la taille de l'échantillon. Une méthode palliant cette difficulté a été proposée par PETIT et al (1998) pour évaluer la contribution de chaque race à la richesse allélique de l'ensemble. Il est intéressant d'observer que, dans l'exemple des 12 populations d'arganier marocain traité par ces auteurs, la population qui contribue le plus à la diversité structurelle et allélique ne possède aucun allèle privé et que celle qui a le plus grand nombre d'allèles privés (2 sur 12 locus) ne vient qu'en 3<sup>ème</sup> position pour le critère de richesse allélique (tableaux 2 et 3 aux pages 847 et 850 de PETIT et al, 1998). Notre étude révèle un paradoxe similaire, puisque le porc *Créole* qui a le plus grand nombre d'allèles, tant observés que privés (tableau 2), contribue assez faiblement à la diversité d'ensemble des races françaises (tableau 6). Des analyses complémentaires restent donc à faire sur cet aspect.

## CONCLUSIONS

Cette étude vient compléter l'étude pilote de LAVAL et al (2000a et b) et confirme la faisabilité d'une évaluation à grande échelle de la diversité génétique - puisqu'ici elle a été réalisée à l'échelle du continent européen - selon les recommandations de la FAO (BARKER et al, 1998). Notons aussi que le projet financé par l'Union Européenne a permis de constituer un stock d'ADN, noyau d'un futur « centre de ressources biologiques » et disponible, sous certaines conditions d'accès définies, pour des recherches sur d'autres marqueurs ou d'autres matériels génétiques.

Une information de grand intérêt pour la gestion de la diversité génétique porcine a ainsi été rassemblée. Une caractérisation plus détaillée est cependant nécessaire si on veut conserver cette diversité dans les meilleures conditions (par exemple, LABROUE et al, 2000 a et b pour les races locales françaises). La difficulté est de combiner la diversité avec le « mérite génétique perçu » de chaque race (PIYASATIAN et KINGHORN, 2003) ou, plus généralement en termes économiques, le « caractère distinctif » et l'« utilité » (WEITZMAN, 1998). Le poids relatif à accorder à la diversité des races et à leur variabilité interne de manière à définir un objectif global de diversité est une autre difficulté (OLLIVIER et FOULLEY, 2002).

Sur le plan français, l'échantillon étudié est certainement restreint par rapport aux 30 races ou lignées répertoriées pour la France dans la base génétique de Hanovre (OLLIVIER et al, 2001a). L'étude confirme cependant l'originalité génétique de nos races locales et particulièrement des porcs *Basque* et *Limousin* (voir tableau 6). Notons aussi l'intérêt de l'introduction de populations très différentes, telles que les races chinoises importées en France en 1979 (à la base de lignées synthétiques telles que la *Tia Meslan*), pour créer une diversité génétique additionnelle. Signalons enfin la possibilité de prendre en compte les distances calculées dans cette étude pour optimiser des plans de croisement, compte tenu de la corrélation attendue entre les distances génétiques et les effets d'hétérosis.

## REMERCIEMENTS

Ce travail a bénéficié du soutien de plusieurs programmes de l'Union Européenne (contrats BIO2-CT94-3044, RESGEN-CT95-012 et BIO4-CT98-0188), ainsi que de financements complémentaires du Ministère de l'Agriculture et de l'INRA, que nous remercions très sincèrement.

L'échantillonnage des races étrangères, les typages d'ADN et le traitement des données ont bénéficié de la contribution des participants étrangers aux contrats mentionnés plus haut, à savoir : A. ARCHIBALD et C.S. HALEY (Roslin, Royaume-Uni), L. PEELMAN (Gand, Belgique), H. GELDERMANN (Stuttgart, Allemagne), M. FREDHOLM (Copenhague, Danemark), M. GROENEN (Wageningen, Pays-Bas), L. ANDERSSON (Uppsala, Suède), G. PLASTOW (Abingdon, Royaume-Uni), K. HAMMOND (FAO, Rome), P. GLODEK (Göttingen, Allemagne), R. DAVOLI (Reggio Emilia, Italie), G. GANDINI (Milan, Italie), J.V. DELGADO (Cordoue, Espagne), E. FIMLAND (Oslo, Norvège), L. ALDERSON (Shrewsbury, GB) et M. RAMOS (Vila Real, Portugal).

L'échantillonnage des races françaises a bénéficié du concours de B. COUDURIER et C. DÉSOUTÉS-SAWADOGO (Agence de la Sélection porcine), P. CHÉREL (France Hybrides), C. BAZIN (Pen Ar Lan), C. GUILLAUME et C. LÉCOUR (SCAAPAG), L. MAIGNEL et R. GUEBLEZ (LGPC/ITP), avec L. KERNALEGUEN (ADN), C. GASNIER (Gène +) et B. LE ROSSIGNOL (Nucléus), et E. DESPRES et J. J. DELATTE pour la collecte des échantillons de porc Créole en Guadeloupe et en Haïti respectivement.

S. DERBAN (étudiant à l'Ecole Nationale de Statistiques Administratives et Economiques de Paris), A. NEAU et X. ROGNON (INRA, Jouy-en-Josas), et A. MARTINEZ et J. L. VEGA-PLA (Cordoue, Espagne) doivent également être remerciés pour l'aide apportée dans l'analyse des données. Merci enfin à H. LAGANT, S. NUGIER, M. L. LE PAIH (INRA-SGQA, Jouy-en-Josas) et M. WEBER (INRA, Communication, Jouy-en-Josas) pour leur aide dans la gestion informatique des données et la préparation du manuscrit.

## REFERENCES BIBLIOGRAPHIQUES

- BARKER J.S.F., 2002. In : Animal Breeding and Animal Genetic Resources, FAL, Braunschweig, Allemagne, 15-21.
- BARKER J.S.F., HILL W.G., BRADLEY D. et al, 1998. Measurement of domestic animal diversity (MoDAD) : original working group report. FAO, Rome, 55p.
- BELKHIR K. BORSA P. GOUDET J CHIKHI L. BONHOMME F., 1998. Belkhir Biosoft, Université de Montpellier II.
- DERBAN S., FOULLEY J.L., OLLIVIER L., 2002. WeitzPro. Calculation of the Weitzman function of diversity. Windows et Unix.
- FELSENSTEIN J., 2000. PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genome Science, University of Washington, Seattle.
- GROENEN M.A.M., JOOSTEN R., BOSCHER M.Y. et al, 2003. Arch. Zootec., 52, 145-155.
- HARTL D.L., CLARK A.G., 1997. Principles of Population Genetics (3<sup>rd</sup> edition). Sinauer, Sunderland, Massachusetts, 542p.
- LABROUE F., GOUMY S., GRUAND J., MOUROT J., NEELZ V., LEGAULT C., 2000a. Journées Rech. Porcine en France, 32, 403-411
- LABROUE F., GUILLOUET P., MARSAC H., BOISSEAU C., LUQUET M., ARRAYET J., MARTINAT-BOTTÉ F., TERQUI M., 2000b. Journées Rech. Porcine en France, 32, 413-418.
- LAVAL G., IANNUCCCELLI N., LEGAULT C. et al, 2000a. Genet. Sel. Evol., 32,187-203.
- LAVAL G., IANNUCCCELLI N., LEGAULT C. et al, 2000b. Journées Rech. Porcine en France, 32, 397-402.
- LAVAL G., SANCRISTOBAL M., CHEVALET C., 2002. Genet. Sel. Evol., 34, 481-507.
- LI K., CHEN Y., MORAN C., FAN B., ZHAO S., PENG Z., 2000. Anim. Genet., 91, 322-325.
- MARTINEZ A.M., DELGADO J.V., RODERO A., VEGA-PLA J.L., 2000. Anim. Genet., 31, 295-301.
- NEI M., 1972. Am. Nat., 106, 283-292.
- OLLIVIER L., 2002. Eléments de génétique quantitative (2<sup>ème</sup> édition). INRA-Editions, Paris.
- OLLIVIER L., WREDE J., DISTL O., 2001a. In : Pig genetic resources in Europe. EAAP publication n°104, 5-14.
- OLLIVIER L., CARITEZ J.L., FOULLEY J.L., et al, 2001b. In : Pig genetic resources in Europe. EAAP publication n°104, 87-97.
- OLLIVIER L., FOULLEY J.L., 2002. Book of Abstracts of the 53<sup>rd</sup> Annual Meeting of the European Association for Animal Production. Wageningen Academic Publishers, Wageningen, Pays-Bas, G2.3, 2.
- OLLIVIER L., FOULLEY J.L., 2004. In : Kolloquium "Nutztierzüchtung im Wandel der Zeit" (sous presse).
- OLLIVIER L., AMIGUES Y., BOSCHER M.Y., 2003. Arch. Zootec., 52, 137-144.
- OLLIVIER L., ALDERSON L., GANDINI G. et al, 2004. En préparation.
- PETIT R.J., EL MOUSADIK A., PONS O., 1998. Conservation Biology, 12, 844-855.
- PITEL F., RIQUET J., 2000. In : INRA Prod. Anim., 2000, numéro hors série "Génétique moléculaire: principes et applications aux populations animales", 45-53.
- PIYASATIAN N., KINGHORN B.P., 2003. J. Anim. Breed. Genet., 120, 137-149.
- PLASTOW G., SIGGENS K., BAGGA M. et al, 2003. Arch. Zootec., 52, 157-164.
- PLASTOW G., HALEY C.S., SANCRISTOBAL M., 2004. En préparation.
- RAYMOND M., ROUSSET F., 1995. J. Heredity, 86, 248-249.
- REYNOLDS J., WEIR B.S., COCKERHAM C.C., 1983. Genetics, 105, 767-779.
- SANCRISTOBAL M., CHEVALET C., HALEY C.S. et al, 2002. World Congr. Genet. Appl. Livest. Prod., 33, 525-528.
- SANCRISTOBAL M., CHEVALET C., HALEY C.S. et al, 2004. En préparation.
- SUN F., ZHANG Y., WANG Z., YANG S., 2002. World Congr. Genet. Appl. Livest. Prod., 33, 529-532.
- THAON D'ARNOLDI C., FOULLEY J.L., OLLIVIER L., 1998. Genet. Sel. Evol., 30, 149-161.
- VAN ZEVEREN A., PEELMAN L., VAN DE WEGHE A., BOUQUET Y., 1995. J. Anim. Breed. Genet., 112, 191-204.
- WEIR B.S., COCKERHAM C.C., 1984. Evolution, 38, 1358-1370.
- WEITZMAN M.L., 1993. Quart. J. Econ., 108, 157-183.
- WEITZMAN M.L., 1998. Econometrica, 66, 1279-1298.