



## DIVERSITE ET RELATIONS GENETIQUES ENTRE LES RACES DE CHEVAUX ELEVEES EN FRANCE : CONSEQUENCES PRATIQUES POUR LA GESTION DES RACES

Par :

- G. Leroy, L. Callède, E. Verrier, J.C. Meriaux, A. Ricard, C. Danchin-Burge, X. Rognon
- AgroParisTech, UMR Génétique et diversité animales, 16 rue Claude Bernard, 75231 Paris Cedex 05

### Résumé

La diversité génétique des races équinées élevées en France a été analysée à partir de génotypages sur 11 marqueurs microsatellites. L'étude a porté sur 1679 animaux appartenant à 33 races reconnues, dont 19 races d'origine française ou implantées et sélectionnées en France depuis plus d'un siècle.

De façon concordante, différentes méthodes d'analyse permettent d'identifier quatre groupes clairement différenciés sur le plan génétique : chevaux de sport et de course, chevaux de trait, poneys, races nordiques. Au sein de chacun de ces groupes, les races montrent une différenciation plus ou moins marquée, qui peut être liée à leur histoire ou leur origine géographique.

Nos résultats suggèrent une importance prépondérante des origines Pur-Sang au sein de la race Selle Français, tandis qu'une fragmentation en noyaux plus ou moins fermés est mise en évidence au sein de la race Arabe. Plusieurs races méritent de voir se renforcer les actions de conservation en leur faveur, l'originalité de certaines d'entre elles, comme le Boulonnais, devant conduire à exclure le recours à des reproducteurs issus d'autres races.

**Mots-clés :** *Equus caballus*, races équinées, microsatellite, diversité génétique, conservation

### Summary

The genetic diversity and structure of horse breeds raised in France were investigated using 11 microsatellite markers and 1679 animals, belonging to 33 recognized breeds. Nineteen of the breeds were of French origin or have been raised in France for at least a century.

According to different approaches, breeds could be classified into four genetic differentiated groups: warm-blooded, draught, Nordic and pony breeds. Within each group, breeds were more or less differentiated, which could be linked to their history and geographical origin.

Our results suggest the predominance of Thoroughbred in Selle-Français breed, while Arab horses seem to be raised in more or less closed clusters. Conservation measures should be increased in some of the breeds. In order to maintain the genetic diversity in some of those breeds, such as Boulonnais breed, crosses with other breeds should be avoided.

**Key-words:** *Equus caballus*, horse breeds, microsatellite, genetic variation, conservation

## Introduction

Qu'il s'agisse des pratiques de sélection, des flux de reproducteurs entre pays, d'une évolution des débouchés ou même de phénomènes de modes, de multiples facteurs peuvent influencer la variabilité génétique au sein d'une espèce domestique. De ce fait, la structure génétique de l'espèce équine a connu depuis plusieurs dizaines d'années d'importantes évolutions, tant au sein des races qu'entre les races. Les évolutions démographiques des effectifs raciaux constituent la conséquence la plus visible de ces bouleversements, avec, comme effet le plus dramatique, la disparition de certaines races. Selon la FAO, pas moins de quinze races françaises auraient disparu au cours des cinquante dernières années, et huit seraient considérées comme menacées. La mise en place de mesures d'aide à la conservation semble, cependant, avoir permis de freiner la chute des effectifs au sein des races menacées, même si la situation reste toujours inquiétante pour certaines d'entre elles, et nécessite un suivi attentif (Danvy et Heydemann, 2005). La diminution de la variabilité génétique à l'intérieur des races, liée à la sur-utilisation de certains reproducteurs, à des pratiques d'accouplement entre proches apparentés, ou plus simplement à un nombre de reproducteurs limités au sein d'une population, constitue un autre problème. Les conséquences peuvent être des phénomènes de dépression de consanguinité, l'émergence d'affections héréditaires, ainsi qu'un appauvrissement de la base de sélection à l'intérieur de la race. Enfin, l'existence de flux de gènes, passés ou actuels, entre populations, peut amener les éleveurs et les gestionnaires à s'interroger sur l'originalité génétique de leur cheptel, ainsi que sur la quantification des gènes apportés par d'autres populations, notamment dans le cas des races composites. La nature des problèmes posés nécessite la mise en place d'outils permettant de quantifier la diversité génétique, de manière inter- et intra- raciale.

Depuis près d'une vingtaine d'années, le développement de marqueurs moléculaires permet d'analyser le polymorphisme génétique de manière directe avec d'autant plus de précision, que le nombre de marqueurs employés et d'individus analysés est important. En comparaison des analyses généalogiques, une telle approche a notamment l'avantage de permettre l'analyse des relations génétiques entre races (Rognon *et al.*, 2005), et éventuellement de quantifier l'importance de flux de gènes d'une population vers une autre. A l'aide des approches de Weitzman (1992), il est aussi possible de mesurer la contribution des races à la variabilité génétique existante au sein de l'espèce. Un tel indicateur peut être utile au décisionnaire confronté au choix des races devant bénéficier en priorité de programmes de conservation.

Nous avons réalisé une étude de la diversité génétique présente chez les principales races équines élevées en France, à l'aide de marqueurs microsatellites. Nous présenterons ensuite les recommandations pratiques qui peuvent être envisagées à partir de ces résultats.

## 1. Populations analysées et information utilisée

L'échantillonnage des individus a été effectué de manière à disposer d'une image la plus représentative et la plus exhaustive des races équinées élevées en France. Pour cela, les génotypages de 1679 chevaux appartenant à 39 populations ont été employés dans le cadre de l'étude. Ces 39 populations incluaient 31 races reconnues, 1 échantillon de chevaux de Przewalski, et 7 sous-populations résultant de la séparation des individus Anglo-Arabe et Selle Français, respectivement en quatre et trois groupes. Cette séparation était faite à partir des règlements de 2005. Si cette différenciation n'est plus entièrement d'actualité, elle est néanmoins apparue pertinente dans notre étude, puisque basée sur la proportion d'origines étrangères selon des analyses généalogiques. Alors que les sous-populations codées AA6 et AA9 sont considérées comme Anglo-Arabe descendant uniquement d'ascendants Pur-Sang et Arabe, les individus appartenant aux groupes AA5 et AA10 possèdent des ancêtres d'autres origines, la proportion d'origine Arabe étant plus haute dans les groupes AA5 et AA6 que dans les autres. Les SFA97 correspondent à la section A, fermée aux apports extérieurs directs, alors que les SFB98, issus de la section B, possèdent un parent appartenant à une autre race. Les individus SF8 (communément appelés AQPS pour Autre Que Pur Sang) ont une forte proportion d'origine pur-sang et peuvent donc être employés en tant que facteurs d'Anglo-Arabe.

Au sein de chacune des populations, entre 23 et 50 animaux, nés entre 1996 et 2005, ont été génotypés. Afin d'avoir des échantillons les plus représentatifs des populations, les individus ont été sélectionnés de manière à n'avoir aucun parent en commun. Les génotypages ont été effectués par le laboratoire Labogena, sur la base d'un panel de 11 marqueurs microsatellites, employé dans le cadre du protocole d'identification génétique et de contrôle de filiation (pour plus de détails sur les méthodes employées, voir Leroy *et al.*, 2008).

## 2. Variabilité génétique intra- raciale

Le taux d'hétérozygotie attendu selon les hypothèses d'Hardy-Weinberg et non biaisé ( $H_e$ ) correspond à l'un des indicateurs les plus communément employés pour mesurer la variabilité génétique existante au sein d'une population (Rognon *et al.*, 2005). Dans le cas des 39 populations étudiées, la valeur moyenne de  $H$  est 0,70, pour une plage de variation allant de 0,43 en cheval Frison, à 0,79 pour le Poney français de Selle (Tableau 1). Dans le cas du Frison, le manque de variabilité est expliqué par d'importants goulets d'étranglement liés à des effectifs faibles (Ducro *et al.*, 2006) alors que la forte hétérozygotie au sein du Poney français de selle est probablement liée à son statut de race composite, ouverte à de nombreux croisements. On notera aussi une diversité relativement faible au sein de la race Poitevine, avec un taux d'hétérozygotie de 0,57.

La comparaison des taux d'hétérozygotie attendus et observés ( $H_o$ ), évaluée selon la formule  $F_{is} = 1 - (H_o / H_e)$ , permet de mettre en évidence d'éventuelles sous-structures à l'intérieur des races, ainsi que des pratiques d'accouplements entre proches apparentés, lorsque les valeurs de  $F_{is}$  sont fortement positives. En ce qui concerne la race Arabe, avec une valeur de  $F_{is}$  de 0,08, les résultats obtenus permettent de confirmer les résultats des analyses généalogiques de Moureaux *et al.* (1996) : quoique l'hétérozygotie attendue soit plutôt élevée au sein de la race (0,72), les accouplements entre proches apparentés, liés à une fragmentation de la race en noyaux plus ou moins fermés, a pour conséquence une hétérozygotie observée bien plus faible (0,66). A terme, des pertes de variabilité importantes pourraient avoir lieu dans la race, ce qui nous amène à recommander d'augmenter les échanges de reproducteurs.

Tableau 1 : Paramètres génétiques des 39 populations étudiées (*He* : hétérozygotie attendue, *Pex* : probabilité d'extinction, *CW/CB/D* : contributions à la diversité intra-population/inter-population/agrégée, *CP* : potentiel de cryoconservation)

*Table 1: Genetic parameters within the 39 populations studied (He: expected heterozygosity, Pex: extinction probability, CW/CB/D: within-breed/between-breed/aggregate contributions to diversity, CP: cryoconservation potential)*

Race / Breed	Code	<i>He</i>	<i>Fis</i>	<i>Pex</i>	<i>CW</i>	<i>CB</i>	<i>D</i>	<i>CP</i>
Anglo-Arabe	AA10	0,71	-0,01					
	AA5	0,73	0,03	0,11	0,18	0,85	0,24	0,10
	AA6	0,73	0,03					
	AA9	0,69	-0,01					
Appaloosa	APPAL	0,77	0,06					
Arabe-Barbe	AB	0,76	0,03					
Arabe	AR	0,72	0,08	0,03	0,21	10,90	1,18	0,35
Ardennais	ARD	0,64	0,03	0,08	-0,46	1,33	-0,30	0,10
Auxois	AUX	0,65	0,05	0,57	-0,32	3,14	-0,01	1,79
Barbe	BA	0,74	0					
Boulonnais	BOUL	0,62	0,03	0,24	-0,60	12,35	0,57	2,95
Breton	BR	0,66	-0,02	0,02	-0,24	5,57	0,29	0,12
Camargue	CAM	0,73	0,07	0,12	0,27	7,99	0,97	0,97
Connemara	CO	0,75	0,03					
Cob Normand	COBND	0,72	-0,01	0,21	0,24	2,42	0,44	0,52
Comtois	COMT	0,69	0,03	0,02	-0,01	3,63	0,32	0,06
Fjord	FJ	0,67	-0,03					
Friesian	FRI	0,43	0					
Haflinger	HAF	0,65	0,05					
Islandais	IS	0,7	0,03					
Landais	LAND	0,75	0,05	0,74	0,48	3,99	0,79	2,95
Lusitanien	LUS	0,74	0,04					
Merens	MER	0,7	-0,01	0,15	-0,02	10,41	0,96	1,53
New Forrest	NF	0,76	0,03					
Percheron	PER	0,68	-0,01	0,07	-0,10	4,60	0,33	0,34
Poney Français de selle	PFS	0,79	0	0,14	0,78	1,93	0,89	0,27
Poitevin	POIT	0,57	-0,02	0,38	-1,00	12,60	0,23	4,83
Pottok	POT	0,77	-0,03	0,15	0,58	1,33	0,64	0,20
Pure Race Espagnole	PRE	0,7	0,11					
Przewalsky	PRW	0,59	0,05					
Pur Sang	PS	0,69	-0,01	0,04	0,01	6,17	0,57	0,22
Quarterhorse	QH	0,73	0,01					
Selle Français	SF8	0,71	-0,03					
	SFA97	0,74	0,01	0,03	0,34	1,33	0,43	0,04
	SFB98	0,75	0					
Shetland	SHE	0,69	0,06					
Trait du Nord	TDN	0,64	-0,08	0,85	-0,41	1,93	-0,20	1,64
Trotteur Français	TF	0,7	0,01	0,02	0,02	7,51	0,70	0,18
Poney Welsh	WAB	0,76	0,03					

### **3. Variabilité génétique inter- raciale**

#### **3.1. Une structuration de l'espèce en quatre groupes**

La méthode la plus classique permettant de visualiser les relations génétiques entre les races passe par le calcul de distances génétiques sur la base des fréquences alléliques. A partir des matrices de distances ainsi calculées, des arbres de relation génétique (ou phénogrammes) sont généralement tracés. L'utilisation de ces arbres pose néanmoins problème dans le cas des races domestiques, issues d'une succession d'isolements et de croisements, ne correspondant pas à des relations dichotomiques. En conséquence, la représentation sous la forme de réseaux, prenant en compte toutes les relations possibles, apparaît plus satisfaisante pour représenter les liens existant entre les races. Les relations entre les 39 populations sont représentées sur la Figure I, qui semble structurer ces dernières en quatre groupes : chevaux de sport et de course, chevaux de trait, poneys, races nordiques. Une telle typologie, guère éloignée de la classification actuelle employée par les Haras Nationaux, semble regrouper les races d'abord en fonction de leur utilisation, et dans une moindre mesure, selon leurs origines géographiques.

Deux races apparaissent cependant isolées de ces quatre groupes. Le cheval de Przewalski (PRW), par son statut et son histoire, est logiquement porteur d'un pool génétique assez différent des races domestiques, ce qui explique son éloignement sur la Figure I. L'isolement du cheval Frison (FRI) doit en revanche être mise sur le compte de la faible hétérozygotie de la race : les distances génétiques sont en effet biaisées lorsqu'elles sont établies sur des races ayant subi, comme le Frison, une dérive génétique importante. En réalité, de par son histoire et son utilisation à l'origine, il convient plutôt de rattacher cette race au groupe des chevaux de trait.

L'utilisation d'une approche complémentaire d'analyse des génotypes va nous permettre d'affiner ces résultats. Pritchard et Wen (2003) a en effet développé une méthode bayésienne, qui, à partir de données de génotypages, permet d'affecter les génotypes d'individus ou de populations à un nombre choisi K de groupes (ou clusters). L'utilisation de cette approche pour des valeurs de clusters de K allant de 2 à 5 (Figure II) semble confirmer et affiner les résultats obtenus avec les distances génétiques. En supposant l'existence de deux populations au sein du jeu de données, les chevaux de trait, incluant le Frison (FRI), sont clairement séparés des chevaux de sport et de course, les autres races ayant des appartenances intermédiaires entre ces deux groupes. Lorsque le nombre de clusters augmente, un certain nombre de populations semblent plus ou moins se distinguer. Les races nordiques/primitives, incluant le cheval de Przewalski (PRW), sont très clairement différenciées. Les races de Poneys forment un groupe moins homogène. Enfin, les races de chevaux de sang semblent former un gradient, avec des races qui ségrègent entre les influences Pur-Sang et Arabo-ibériques.

Figure I : Réseau Neighbor-Net, des 39 populations équinées, basé sur la distance de Reynolds  
 Figure I: Neighbor-Net for the 39 horse populations, based on Reynolds distance.

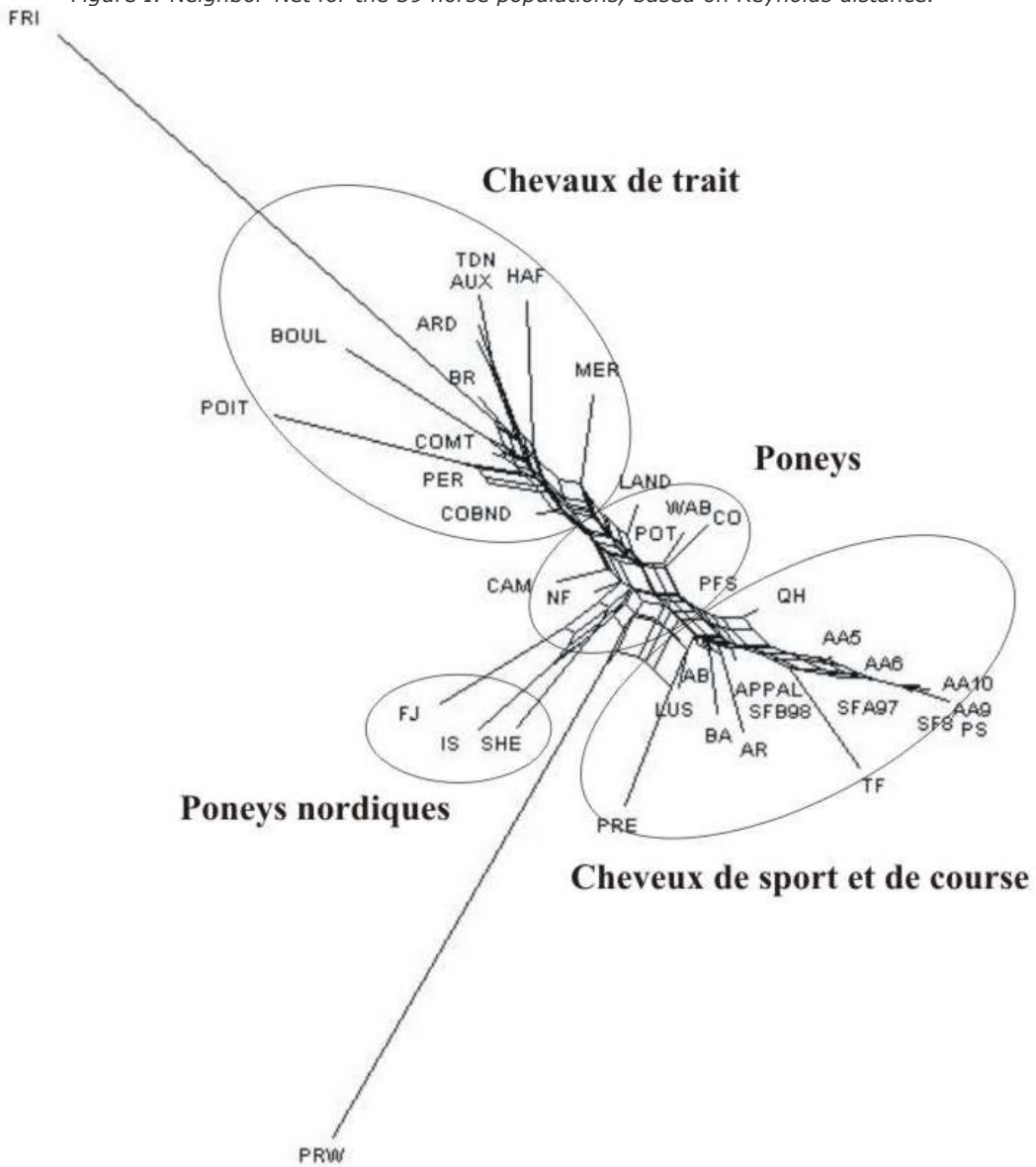
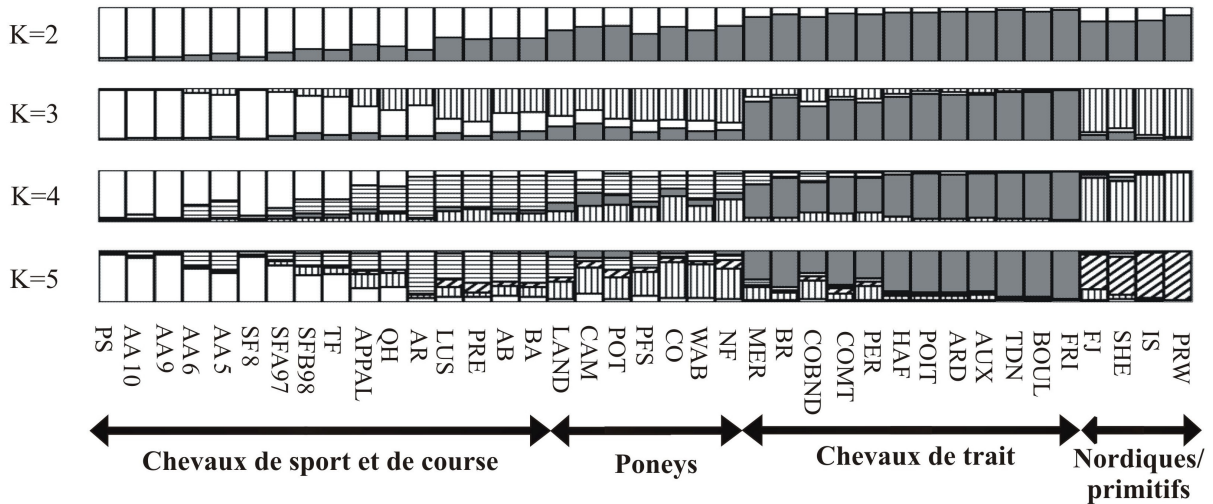


Figure II : Affectation des 39 populations équinées à K clusters, selon des approches bayésiennes  
 Figure II: Cluster assignment of the 39 horse populations to K clusters, using Bayesian approaches



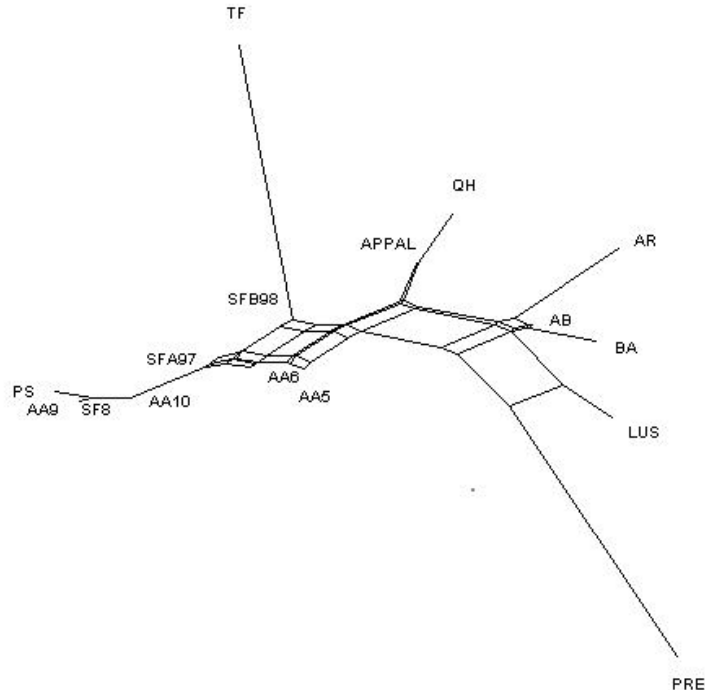


### 3.2. Races de chevaux de sport et de course

L'utilisation d'un réseau Neighbor-Net (Figure III) à partir des races de chevaux de sport et de course confirme bien les deux influences prépondérantes des races Pur-sang et Arabe au sein du groupe.

Figure III : Réseau Neighbor-Net, des 16 populations de chevaux de sport et de course, basé sur la distance de Reynolds

Figure III: Neighbor-Net for the 16 race/riding horse populations, based on Reynolds distance.



Deux races semblent s'isoler plus particulièrement au sein de ce réseau : le Trotteur français et le Pure race espagnole, quoique ce dernier montre une certaine proximité avec le Lusitanien (LUS), sans doute du fait d'origines communes.

Les différentes sous-populations de Selle Français et d'Anglo-Arabe se retrouvent très proches du Pur-Sang. Parmi ces sous-populations, les SFB98, AA6 et AA5 semblent plus influencés par le pool génétique Arabe, ce qui est, dans le cas des Anglo-Arabe, cohérent avec le classement effectué sur les bases généalogiques. Ce résultat permet de mettre en évidence l'importance de l'influence du Pur-Sang au sein du Selle Français. Il n'apparaît en conséquence donc pas justifié, d'un point de vue génétique, d'accroître cette influence par l'utilisation de reproducteurs Pur-Sang au sein du Selle Français.

### 3.3. Races de chevaux de trait

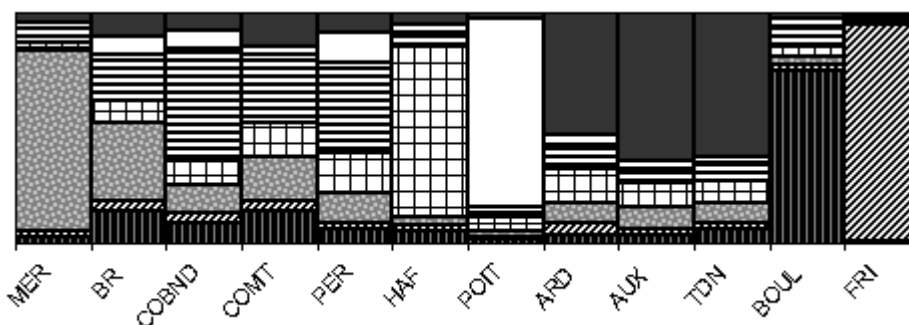
Le cluster spécifique des chevaux de trait apparaît très clairement séparé des races de chevaux de sport et de course. Il inclut l'intégralité des races classées en cheval de trait par la nomenclature des Haras Nationaux, ainsi que trois autres races, classées en cheval de sang (Frison (FRI), Merens (MER)) et Poney (Haflinger (HAF)). Ce résultat est probablement lié au fait que les trois races ont, par le passé, été employées en tant que cheval de trait, et ont pu avoir été sujettes à des croisements avec des races de chevaux lourds.

Il est également intéressant de constater l'absence de lien entre ce groupe et celui incluant la race Arabe (y compris en augmentant la valeur de K (Leroy *et al.*, 2008)).

Ce résultat contredit donc les hypothèses historiques supposant une influence significative du sang arabe sur certaines races de chevaux lourds.

L'analyse spécifique du groupe des chevaux de trait (Figure IV) permet d'individualiser quatre races : le cheval Frison (FRI), le Boulonnais (BOUL), le Poitevin (POIT), et le Haflinger (HAF). En revanche, les races Ardennaise (ARD), Auxois (AUX) et Trait du Nord (TDN) montrent une similarité génétique importante, ce qui est en lien avec une origine géographique et une histoire relativement communes, et apparaissent difficiles à distinguer génétiquement.

Figure IV : Affectation des 12 races de chevaux de trait à 7 clusters  
Figure IV: Cluster assignment of the 12 draught horse breeds to 7 clusters



### 3.4. Races de poneys et races nordiques

Le cluster incluant les races de poney constitue un groupe intermédiaire entre les chevaux de trait et les chevaux de sport et de course. Ce groupe inclut de plus le Camarguais (CAM), classé en cheval de sang, mais morphologiquement proche des poneys (Quittet et Blanc, 1974). Les races de poneys nordiques et le cheval de Przewalski (PRW) constituent un groupe distinct des poneys, mais semble cependant partager une certaine similarité génétique avec ces derniers.

## 4. Approche de conservation

Dans un contexte de conservation, il est rarement possible, pour des raisons économiques, d'affecter à toutes les populations d'une espèce des moyens nécessaires à leur préservation. Aussi est-il parfois nécessaire d'établir des priorités et de choisir les populations pour lesquelles des moyens doivent être engagés en priorité. Sur la base d'indicateurs de variabilité (hétérozygotie, distances génétiques...), il est possible d'évaluer la contribution d'une race à la variabilité intra-raciale ( $CW$ ) et inter-raciale ( $CB$ ) de l'espèce (Ollivier et Foulley, 2005). Ces deux valeurs, pondérées par l'indice de fixation  $F_{st}$  (qui correspond au pourcentage de la diversité génétique expliquée par les différences entre races), peuvent être additionnées en un indice de contribution à la diversité agrégée  $D$ . Appliquée aux 19 races d'origine française, ou élevées en France depuis au moins cent ans, cette méthode montre que la race ayant la contribution la plus élevée à la diversité agrégée est la race Arabe (AR). Parmi les races à petit effectif, les races Landais (LAND) et Mérens (MER) montrent aussi des valeurs de  $D$  relativement importantes.

Calculé à partir du produit entre la contribution à la diversité inter-raciale  $CB$  et la probabilité d'extinction  $P_{ex}$  (établie à partir des effectifs de reproducteurs), le potentiel de cryoconservation  $CP$  permet d'étudier plus précisément les races à petits effectifs. A partir de ce critère, les efforts de conservation pourraient être portés sur le Poitevin (POIT), le Boulonnais (BOUL) et le Landais (LAND).



## Conclusion

L'utilisation des marqueurs moléculaires nous a permis de mettre en évidence les proximités génétiques ainsi que les spécificités génétiques que peuvent présenter les races élevées en France. Il convient en revanche de préciser que dans la majorité des cas, il ne nous était pas possible d'affecter des chevaux à leur race supposée sur la base de marqueurs moléculaires. Ceci était lié d'une part au nombre limité de marqueurs, d'autre part, au fait qu'une partie des populations étudiées, sont des races composites, sujettes à des croisements fréquents. L'ajout de marqueurs supplémentaires devrait cependant permettre une telle application, au moins dans le cas des populations élevées en race pure.

Les marqueurs nous ont permis de classer les races dans des groupes relativement similaires à la nomenclature actuellement employée par les haras nationaux. Il convient cependant de noter les quelques différences décrites précédemment. Un tel résultat ne doit pas pour autant constituer une remise en cause, la nomenclature des Haras Nationaux devant être basée sur les utilisations actuelles et non sur les origines passées.

L'utilisation des approches de conservation nous a aussi permis de mettre en évidence, parmi les races d'effectif limité, lesquelles montraient une certaine originalité génétique (Boulonnais) et mériteraient un renforcement des actions de soutien en leur faveur. Afin de maintenir cette originalité, il conviendrait d'éviter d'avoir recours à des reproducteurs issus d'autres races, tant que la variabilité génétique intra- raciale reste substantielle. Bien que les approches que nous avons employées au cours de cette étude puissent constituer une aide précieuse pour le décisionnaire, il convient aussi pour ce dernier de prendre en compte d'autres sources de données (Fadlaoui 2006) telles que les valeurs socio-économiques et environnementales que peuvent avoir les races, ainsi que la motivation des éleveurs concernés.

## Bibliographie

DANVY S. et HEYDEMANN P., Le point sur les ressources génétiques équinnes en France. In 31<sup>ème</sup> journée de la recherche équine, Les Haras Nationaux, Le Pin-au-Haras, 2005, 135-146.

DUCRO B.J., BOVENHUIS H., NEUTEBOOM M. et HELLINGA I., Genetic diversity in the Dutch Friesian horse. In Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, 13-18 August 2006, Belo Horizonte, Brazil.

FADLAOUI A. Modélisation bioéconomique de la conservation des ressources génétiques animales. Mémoire de thèse. Université Catholique de Louvain, 2006.

LEROY G., CALLEDE L., VERRIER E., MERIAUX J.C., RICARD A., DANCHIN-BURGE C. et ROGNON X., Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism. *Genet. Sel. Evol.* 41, (2009), 5

MOUREAUX S., VERRIER E., RICARD A. et MERIAUX J.C. Genetic variability within French race and riding horse breeds from genealogical data and blood marker polymorphisms. *Genet. Sel. Evol.* 28, 1996, 83-102.

OLLIVIER L. et FOULLEY J.L. Aggregate diversity: New approach combining within- and between-breed genetic diversity. *Livest. Prod. Sci.* 95, 2005, 247-254.

PRITCHARD J.K. et WEN W. Documentation for STRUCTURE software: Version 2, University of Chicago, 2003, <http://pritch.bsd.uchicago.edu>.

QUITTET E. et BLANC H. Races chevalines en France, 2nde édition. La Maison Rustique, Paris, 1974.

ROGNON X., MERIAUX J.C. et VERRIER E. Caractérisation génétique des races à l'aide des marqueurs moléculaires. In 31<sup>ème</sup> journée de la recherche équine, Les Haras Nationaux, Le Pin-au-Haras, 2005, 147-160.

WEITZMAN M.L., On diversity. *Quart. J. Econom.* 107, 1992, 363-405.