

Diversité génétique et richesse allélique : concepts et application à des races bovines

J.-L. FOULLEY (1), L. OLLIVIER (2)

(1) INRA-SGQA - 78352 Jouy-en-Josas cedex (2) 8 impasse Calmette - 78350 Jouy-en-Josas

RESUME - La diversité génétique et sa caractérisation sont une préoccupation majeure des généticiens appliqués. Divers paramètres peuvent être utilisés à cet égard. A côté des critères classiques qui s'appuient sur les fréquences géniques et leur variabilité entre races, tels que les indices de différenciation de Wright (ou Nei) et les distances génétiques, on peut envisager des critères plus qualitatifs tels que le nombre total d'allèles présents ou présents exclusivement dans une race (allèles "privés"). Ces paramètres caractéristiques de la richesse allélique sont non seulement révélateurs de l'historique des races mais aussi déterminants dans leur évolution à long terme. L'objet de cette présentation est de faire le point sur les méthodes d'estimation de la richesse allélique prenant en compte la taille des échantillons qui conditionne le nombre d'allèles observés. On compare ici les méthodes basées sur le principe de la raréfaction à celles basées sur celui de l'extrapolation. On montre également comment ces quantités peuvent être utilisées pour calculer des indices de similarité/dissimilarité entre races. Ces concepts sont illustrés sur des données de polymorphisme de 9 microsatellites typés sur 18 races bovines européennes et africaines.

Genetic diversity and allelic richness: concepts and application to cattle breeds

J.-L. FOULLEY (1), L. OLLIVIER (2)

(1) INRA-SGQA - 78352 Jouy-en-Josas cedex

SUMMARY - Genetic diversity and its characterisation are of major concern to applied geneticists. Various parameters are available for that purpose. In addition to conventional criteria based on allelic frequencies and their between-breed variability, such as the differentiation indices of Wright (or Nei) and genetic distances, one can consider more qualitative criteria such as the number of alleles present or only present in one breed ("private" alleles). These parameters of allelic richness reflect the breeds' histories and are also indicative of their potential evolution in the long term. The purpose of this paper was to present an overview of the methods of estimation of allelic richness taking due account of sample size, by comparing the methods based on the principle of rarefaction and those based on extrapolation. It is also shown how indices of similarity/dissimilarity between breeds, measuring between-breed allelic diversity, may be derived from the number of alleles observed. These concepts are illustrated using 9 microsatellite polymorphisms in a set of 18 European and African cattle breeds.

INTRODUCTION

L'utilisation optimale de la diversité génétique des populations animales suppose que cette diversité soit caractérisée sous tous ses aspects. Divers paramètres peuvent être utilisés, notamment ceux basés sur les marqueurs moléculaires. A côté des critères classiques qui s'appuient sur les fréquences géniques et leur variabilité, tels que les indices de différenciation de Wright (ou Nei) et les distances génétiques, des critères plus qualitatifs tels que la richesse allélique ont également été proposés (voir la mise au point de Ollivier *et al.*, 2000). Cet article apporte une illustration sur un échantillon de races bovines.

1. MATERIEL ET METHODES

1.1. LES RACES ET LES MARQUEURS

Les 18 races bovines échantillonnées sont listées dans le tableau 1. Elles comprennent 10 races françaises et 8 races africaines. Ces dernières appartiennent à 2 sous-espèces, *Bos taurus taurus* (taurin) et *Bos taurus indicus* (zébu). Les marqueurs génétiques utilisés sont des microsatellites, connus pour être bien répartis sur le génome de chaque espèce et extrêmement polymorphes. Dans l'échantillon retenu pour cette étude, les 9 microsatellites considérés (INRA5, INRA16, INRA32, INRA35, INRA63, INRA72, INRAK, ETH152, ETH225) comportent en moyenne 12,4 allèles (allant de 8 à 17 selon le locus). Les nombres extrêmes d'individus typés par race selon le locus figurent dans le tableau 1. Pour plus de détails sur les données utilisées on peut consulter les articles de Moazami-Goudarzi *et al.*, (1997 et 2001) et de Souvenir Zafrindrajaona *et al.*, (1999).

Tableau 1 : échantillons étudiés

Race	Individus typés
Bretonne Pie-Noire	31 -50
Charolais	43 -50
Holstein	49-50
Jersey	48-50
Limousin	45-47
Rouge des Prés ¹	30-31
Montbéliard	44-49
Normand	44-50
Parthenais	45-50
Vosgienne	45-50
Baoulé	36-37
Borgou	31-50
Lagunaire	44-51
N'Dama	28-30
Somba	42-50
Zébu Choa	32-36
Zébu M'Bororo	23-26
Zébu Peul	36-50

¹ connue précédemment sous l'appellation "Maine-Anjou"

1.2. DIVERSITE GENETIQUE

La diversité génétique globale d'un ensemble de races peut être évaluée par la fonction de diversité de Weitzman basée sur la matrice des distances génétiques entre les races prises deux à deux (voir Thaon d'Arnoldi *et al.*, 1998). Nous avons ici retenu la distance de Reynolds (Reynolds *et al.*, 1983). L'approche Weitzman permet d'analyser la diversité globale et d'estimer les contributions individuelles des races à la diversité totale.

1.3. RICHESSE ALLELIQUE D'UNE RACE

La richesse allélique d'une race, définie comme le nombre d'allèles présents à un locus donné, est connue pour dépendre de la taille de l'échantillon, puisque les chances de découvrir un nouvel allèle augmentent chaque fois qu'un nouvel individu est observé. Pour comparer la richesse allélique de deux races il est donc nécessaire de corriger le nombre d'allèles observé en fonction de la taille de l'échantillon. Deux procédures de correction sont possibles, basées soit sur le principe de "raréfaction" (El Mousadik et Petit, 1996), soit sur le principe d'"extrapolation" (Foulley et Ollivier, 2006a). La première méthode consiste à calculer le nombre d'allèles attendus dans un échantillon de taille donnée, inférieure à celle de l'échantillon dont on dispose. Ainsi en calculant ce nombre pour le plus petit échantillon peut-on comparer toutes les races sur la même base. La richesse allélique ainsi estimée est donc au plus égale au nombre d'allèles observés. Dans la méthode d'extrapolation au contraire, on ajoute au nombre d'allèles observés dans chaque race le nombre des allèles qu'on s'attend à voir manquer du fait de la taille limitée de l'échantillon.

1.4. DIVERSITE ALLELIQUE ENTRE RACES

La diversité allélique des races dépend de l'existence d'allèles spécifiques de chaque race. En considérant l'ensemble des races, on peut définir pour chaque race *i* le nombre d'allèles présents dans *i* et absents dans toutes les autres races, ce que l'on appelle la richesse allélique "privée" de *i*. On peut aussi définir une distance allélique (ou dissimilarité) entre deux races *i* et *j* par le nombre (*a*) des allèles présents dans *i* et absents dans *j*. La distance ainsi définie est généralement dissymétrique, puisque le nombre (*b*) des allèles présents dans *j* et absents dans *i* peut être différent de *a*. Une distance symétrique est obtenue en considérant le rapport $(a+b)/(a+b+c)$, où *c* représente le nombre d'allèles communs aux deux races *i* et *j*. Comme détaillé dans Foulley et Ollivier (2006b), les richesses alléliques privées de chaque race et les distances entre races peuvent être estimées en appliquant la raréfaction ou l'extrapolation pour tenir compte des tailles d'échantillon variables. A partir d'une matrice symétrique de distances alléliques ainsi obtenue, l'approche décrite en 1.2. peut être utilisée pour analyser la diversité allélique.

2. RESULTATS

2.1. DIVERSITE GENETIQUE ET RICHESSE ALLELIQUE

Les contributions individuelles de chaque race à la diversité génétique totale et les richesses alléliques estimées sont présentées dans le tableau 2. On note une très grande variation des contributions raciales individuelles, qui vont de 0,3 (zébu M'Bororo) à 17,8 % (Lagunaire). Les richesses alléliques varient relativement moins entre races. Dans l'ensemble, si globalement les races africaines échantillonnées diffèrent assez peu de l'échantillon européen dans leurs contributions individuelles à la diversité génétique et leurs richesses alléliques (voir les deux dernières lignes du tableau 2), on note cependant une plus grande richesse allélique des zébus, qui se retrouve également dans la race Borgou qui est un croisement zébu x taurin. Les mêmes différences se retrouvent dans les hétérozygoties, mais avec une différence légèrement en faveur des races européennes. Comme attendu, la richesse

allélique estimée par raréfaction est notablement inférieure à celle obtenue par extrapolation, mais elles sont fortement corrélées ($r=0,96$).

Tableau 2 : diversité génétique et richesses alléliques des différentes races

Race	Diversité génétique (%)	Richesse allélique (nb d'allèles/locus)	
		Raréfaction	Extrapolation
Bretonne Pie-Noire	2,80	5,41	9,89
Charolais	5,30	5,86	10,15
Holstein	7,68	5,11	9,42
Jersey	9,62	4,30	8,62
Limousin	5,89	4,73	9,29
Rouge des Prés	5,12	4,89	9,42
Montbéliard	2,36	4,75	9,12
Normand	6,95	5,57	10,03
Parthenais	4,03	5,87	9,99
Vosgienne	4,05	5,03	9,40
Baoulé	5,19	3,98	8,69
Borgou	4,45	6,58	10,58
Lagunaire	17,81	3,27	7,50
N'Dama	7,37	4,65	9,39
Somba	1,64	5,08	9,38
Zébu Choa	7,01	6,79	10,36
Zébu M'Bororo	0,29	6,22	10,25
Zébu Peul	2,43	6,43	10,29
Moyenne Europe	5,38	5,15	9,53
Moyenne Afrique	5,78	5,37	9,56

2.2. RICHESSE ALLELIQUE PRIVÉE

Tableau 3 : richesses alléliques privées des différentes races

Race	Allèles privés observés	Allèles privés estimés	
		Raréfaction	Extrapolation
Bretonne Pie-Noire	1	1,12	0,94
Charolais	2	1,46	1,55
Holstein	0	0,13	0,00
Jersey	0	0,25	0,00
Limousin	0	0,04	0,00
Rouge des Prés	0	0,25	0,00
Montbéliard	0	0,31	0,00
Normand	0	0,47	0,00
Parthenais	1	0,57	0,54
Vosgienne	0	0,04	0,00
Baoulé	0	0,00	0,00
Borgou	3	3,44	2,43
Lagunaire	3	1,65	1,82
N'Dama	0	0,00	0,00
Somba	3	2,55	2,40
Zébu Choa	2	2,38	1,23
Zébu M'Bororo	0	0,86	0,00
Zébu Peul	1	1,96	0,61
Total Europe	4	4,63	3,10
Total Afrique	12	12,85	8,50

Le tableau 3 donne les nombres totaux d'allèles spécifiques de chaque race (allèles privés) sur l'ensemble des 9 locus étudiés, soit basés sur les nombres directement observés, soit corrigés pour les tailles d'échantillon par les deux méthodes décrites plus haut. Des différences notables de richesse allélique privée sont à remarquer. L'échantillon des races africaines se distingue par une plus grande diversité allélique, que reflète le nombre total des allèles privés dans les deux ensembles. On voit également que les richesses alléliques privées obtenues par les deux méthodes sont voisines ($r=0,90$). Les valeurs raréfiées sont en moyenne

supérieures, ce qui résulte en partie des situations où la richesse privée estimée excède la valeur observée, ce qui ne se produit jamais avec l'extrapolation. L'originalité allélique des races africaines est également confirmée par le nombre total d'allèles spécifiques des races africaines (32) comparé à celui des races européennes (14).

2.3. DIVERSITE GENETIQUE ET DIVERSITE ALLELIQUE

Le tableau 4 permet de comparer la répartition entre races de la diversité génétique (reprise du tableau 2) à la diversité allélique, en appliquant la fonction de diversité de Weitzman aux matrices de distance génétique et allélique correspondantes. La répartition des deux diversités entre les 18 races montre un certain parallélisme, avec cependant des exceptions, comme celle de la race Somba, qui contribue notablement à la diversité allélique malgré sa faible part dans la diversité génétique.

Tableau 4 : analyse de la diversité allélique

Race	Diversité	Diversité allélique (%)	
	génétique (%)	Raréfaction	Extrapolation
Bretonne Pie-Noire	2,80	4,16	4,51
Charolais	5,30	5,47	4,41
Holstein	7,68	4,95	5,58
Jersey	9,62	6,46	7,78
Limousin	5,89	5,11	5,05
Rouge des Prés	5,12	6,27	6,27
Montbéliard	2,36	5,10	5,30
Normand	6,95	4,66	4,52
Parthenais	4,03	4,83	4,55
Vosgienne	4,05	4,18	4,63
Baoulé	5,19	7,47	7,92
Borgou	4,45	5,90	5,29
Lagunaire	17,81	9,74	8,86
N'Dama	7,37	6,47	6,74
Somba	1,64	6,84	6,99
Zébu Choa	7,01	3,10	3,05
Zébu M'Bororo	0,29	3,83	4,74
Zébu Peul	2,43	5,12	3,80
Moyenne Europe	5,38	5,15	5,26
Moyenne Afrique	5,78	6,06	5,92

Le tableau 5 résume les relations qui existent entre les 3 genres de diversité analysés dans ce travail. Sur l'échantillon étudié, il apparaît donc clairement que les diversités génétiques et alléliques ne doivent pas être confondues. La distinction est particulièrement claire pour la richesse allélique privée, qui est pratiquement indépendante de la diversité génétique classique.

Tableau 5 : corrélations entre diversités génétique et allélique (1^{ère} ligne : raréfaction ; 2^e ligne : extrapolation)

	Diversité	Richesse allélique
	allélique	privée
Diversité génétique	0,61	-0,06
	0,55	0,11
Diversité allélique	1	0,03
	1	0,06

3. DISCUSSION

L'intérêt des généticiens pour la richesse allélique a été argumenté il y a déjà longtemps par des études théoriques, dont celles de James (1971), soulignant l'importance du nombre d'allèles dans une perspective d'amélioration génétique à long terme, et de Nei *et al.*, (1975), qui ont

montré le lien entre la richesse allélique et l'histoire évolutive des populations. Malgré cela, les applications en agriculture ont jusqu'à présent été assez limitées. L'attention des généticiens des plantes fut attirée par le travail de El Mousadik et Petit (1996) sur l'arganier marocain. La question du lien possible entre diversité génétique classique et diversité allélique a d'abord été examinée par Barker (2001), qui a montré la quasi indépendance des deux diversités en ré-analysant les mêmes données de l'arganier. Notre étude confirme donc une situation qui semble assez générale, puisqu'elle a été trouvée aussi sur un échantillon de races porcines européennes (Foulley et Ollivier, 2006a). Il faut aussi souligner que richesse et diversité allélique sont deux concepts distincts. La possession d'allèles nombreux par une race ne garantit pas nécessairement une grande originalité allélique. Dans notre échantillon, la corrélation entre richesse (tableau 2) et diversité allélique (tableau 3) est positive mais pas très élevée ($r=0,52$ et $0,12$ respectivement avec raréfaction et extrapolation).

Les deux méthodes utilisées ici pour apprécier l'originalité allélique des races prennent en compte la taille des échantillons aussi bien que le nombre de copies des allèles "privés", ce qui constitue une amélioration notable par rapport au simple comptage des allèles, généralement retenu dans ce domaine. Le tableau 3 montre bien que les 3 races africaines qui possèdent 3 allèles privés, Borgou, Lagunaire et Somba, ne peuvent pas être considérées comme équivalentes du point de vue de leur originalité allélique. Ces races se classent en effet dans l'ordre de leur nombre de copies d'allèles privés, 7 6 et 3 respectivement pour Borgou, Somba et Lagunaire. Rappelons ici l'intérêt de l'extrapolation lorsque les tailles d'échantillon sont très variables, Dans ce cas, le risque existe de baser la méthode de raréfaction sur des faibles tailles d'échantillon rendant la méthode peu sensible à la présence d'allèles rares (Foulley et Ollivier, 2006a). L'extrapolation peut alors conduire à un nombre d'allèles privés supérieur à la raréfaction (voir le cas de la race Lagunaire au tableau 3).

Les mêmes avantages s'appliquent aux calculs de distance basés sur la diversité allélique. Notons ici que la distance allélique "naturelle" est représentée par le nombre des allèles présents dans une race et absents dans une autre, soient les distances a et b définies en 1.4. A ces distances correspond une fonction de diversité qui est simplement le nombre des allèles différents dans un ensemble de races donné, comme le souligne Weitzman (1998). Ce dernier fait à ce sujet une comparaison éclairante avec la diversité des livres dans les bibliothèques. Le nombre de livres différents dans un ensemble de bibliothèques, qui mesure la diversité de cet ensemble, est ainsi analogue au nombre d'allèles d'un ensemble de races. Et la perte de diversité liée à la disparition d'une race (ou d'une bibliothèque) est mesurée par le nombre d'allèles (ou de livres) qui lui sont spécifiques. Des priorités de conservation basées sur la diversité allélique peuvent ainsi être établies en suivant une approche parallèle à celle appliquée dans la conservation de la diversité génétique. Il faut noter que la distance symétrique $(a+b)/(a+b+c)$ conduit à une fonction de diversité différente de la précédente comme le montre le tableau 5. La signification pratique de cette diversité allélique reste encore à éclaircir.

On note sur les races bovines analysées qu'une grande diversité allélique ne va pas nécessairement de pair avec une grande diversité génétique. Quelques exemples illustrent bien cette situation. C'est ainsi que la race Lagunaire, contrairement à la race Jersey, associe bien les deux diversités. A l'inverse, la race Borgou (en fait un croisement taurin x zébu) qui possède beaucoup d'allèles uniques contribue peu à la diversité génétique de l'ensemble. Ceci semble indiquer qu'une grande richesse allélique pourrait refléter le caractère "hétérogène" de la race. Les causes génétiques possibles de la diversité allélique sont cependant diverses, puisque les facteurs classiques de l'évolution génétique, dérive, sélection, mutation et migrations peuvent aussi influencer la richesse allélique. Il n'est d'ailleurs pas exclu que ces facteurs puissent agir différemment selon les régions du génome, ce qui motiverait une étude locus par locus.

Soulignons pour terminer le faible nombre de marqueurs considérés dans cette étude. Il faut donc se garder de tirer des conclusions générales relatives aux races examinées, et retenir surtout le caractère illustratif de notre analyse.

CONCLUSION

La conclusion principale de cette étude est l'indépendance des deux types de diversité examinés sur l'échantillon de races bovines considéré. Cela montre l'utilité de distinguer une diversité génétique qu'on peut appeler "structurelle", basée sur les différences de *fréquence* des allèles d'une race à une autre, de la diversité allélique, qui repose sur la nature des allèles et leur caractère unique en fonction de la race. On note en effet qu'une grande contribution à la diversité génétique ne s'accompagne pas toujours d'une grande diversité allélique (exemple de la race Jersey) et que, à l'inverse, une race possédant beaucoup d'allèles uniques peut ne pas contribuer notablement à la diversité génétique (exemple de la race Borgou). Plus généralement, dans une perspective de conservation de la biodiversité, rappelons que la richesse allélique n'est qu'un critère parmi beaucoup d'autres, comme cela a été souligné par plusieurs auteurs (El Mousadik et Petit, 1996 ; Barker, 2001 ; Foulley et Ollivier, 2006a). Tous ces critères ne sont pas forcément équivalents et aucun d'eux ne doit être négligé.

Nous remercions Denis Laloë (INRA-SGQA, Jouy-en-Josas) de nous avoir procuré les données ayant servi à cette étude

- Barker J.S.F., 2001.** *Can. J. For. Res.*, 31, 588-595
El Mousadik A., Petit R.J., 1996. *Theor. Appl. Genet.*, 92, 832-836
Foulley J.-L., Ollivier L., 2006a. *Livest. Sci.*, 101, 150-158
Foulley J.-L., Ollivier L., 2006b. 8th *World Cong. Genet. Appl. Livest. Prod.*, communication n°33-09
James J.W., 1971. *Genet. Res.*, 16, 241-250
Moazami-Goudarzi K., Laloë D., Furet J.-P., Grosclaude F., 1997. *Anim. Genet.*, 28,338-345
Moazami-Goudarzi K., Blemsaga D.M.A., Ceriotti G., Laloë D., Fagbohoun F., Kouagou N.T., Sidibé I., Codja V., Crimella C., Grosclaude F., Touré S.M., 2001. *Rev. Elev. Med. Vet. pays Trop.*, 55, 155-162
Nei M., Maryama T., Chakraborty R., 1975. *Evolution*, 29, 1-10
Ollivier L., Chevalet C., Foulley J.-L., 2000. In : "Génétique moléculaire : principes et applications aux populations animales", INRA Productions Animales, no spécial, 247-252
Reynolds J., Weir B.S., Cockerham C.C., 1983. *Genetics*, 105, 767-779
Souvenir Zafrindrajaona P., Zeuh V., Moazami-Goudarzi K., Laloë D., Bourzat D., Idriss A., Grosclaude F., 1999. *Rev. Elev. Med. Vet. pays Trop.*, 54, 1-10
Thaon d'Arnoldi C., Foulley J.-L., Ollivier L., 1998. *Genet. Sel. Evol.*, 30, 149-161
Weitzman M.L., 1998. *Econometrica*, 66, 1279-1298