

## INDICATEURS DE DIVERSITÉ GÉNÉTIQUE ANIMALE<sup>1</sup>

par Louis Ollivier<sup>2</sup> et Jean-Louis Foulley<sup>3</sup>

### RÉSUMÉ

La caractérisation des ressources génétiques animales est une condition indispensable à leur bonne gestion. À cet effet, plusieurs indicateurs de diversité génétique sont couramment utilisés. Nous présentons ici trois types d'indicateurs, qui sont les inventaires de races, les mesures de consanguinité et les marqueurs génétiques. Des inventaires de races au niveau mondial, appuyés sur des informations fournies par chaque pays, ont été organisés depuis le début des années 1980. Le système d'information en ligne mis en place par la FAO depuis 1993 répertorie aujourd'hui environ 14000 races de mammifères et d'oiseaux, réparties dans 182 pays. Les composantes nationale, régionale et interrégionale de cette diversité raciale représentent respectivement 1, 26 et 73 % de la diversité raciale totale. Les mesures de consanguinité, basées sur des pedigrees ou sur les effectifs de reproducteurs en service, offrent le moyen de suivre et de contrôler l'évolution de la diversité interne à chaque race, de manière à minimiser les effets dépressifs de la consanguinité sur les performances et l'érosion de la variabilité génétique des caractères sélectionnés. Les marqueurs génétiques offrent aujourd'hui des évaluations objectives de la diversité des génomes animaux de nature à compléter les indicateurs précédents. On évalue ainsi à la fois la diversité génétique *stricto sensu* (basée sur les hétérozygoties dans l'approche de Nei) et la diversité allélique (basée sur le nombre d'allèles différents dans chaque race et sur l'existence d'allèles spécifiques à certaines races). Ces deux types de diversité sont indépendants et méritent d'être tous les deux considérés dans le choix des priorités de conservation. L'article met l'accent sur des principes généraux d'analyse de diversité, communs aux indicateurs passés en revue, et sur leurs complémentarités.

### SUMMARY

*Characterising animal genetic resources is of utmost importance for their proper management. For that purpose, a number of indicators of genetic diversity are being currently employed. Here we present three types of indicators, which are breed inventories, inbreeding measurements and genetic markers. Breed inventories at world level, based on country reporting, have been implemented since the early 1980s. The on-line information system initiated by FAO in 1993 presently reports about 14000 breeds, mammalian and aviary, distributed over 182 countries. Breed diversity can be partitioned into national, regional and interregional components, representing respectively 1, 26 and 73 % of the present total diversity. Inbreeding measurements derived either from pedigrees or from the number of breeding animals, offer a mean of monitoring the internal diversity of each breed, in order to minimize deleterious effects of inbreeding depression as well as depletion of genetic variability of the traits being selected for. Genetic markers presently allow objective evaluation of animal genome diversities, usefully complementing the previous indicators. One may thus evaluate gene diversity (based on expected heterozygosities, as in the Nei approach) and allelic diversity (based on the number of different alleles in each breed and on the existence*

---

<sup>1</sup> Cet article est dédié à la mémoire de François Grosclaude, décédé le 3 octobre 2012. François Grosclaude a lancé les recherches en diversité génétique animale à l'Inra au début des années 1960 à partir de la caractérisation du polymorphisme biochimique (groupes sanguins, caséines). Il a animé ces recherches infatigablement tout au long de sa carrière, en tant que directeur du laboratoire de génétique biochimique à Jouy-en-Josas, puis comme chef du département de génétique animale, et ensuite comme directeur scientifique du secteur des productions animales de l'INRA.

<sup>2</sup> Membre de l'Académie d'Agriculture de France. E-mail: [louis.ollivier@free.fr](mailto:louis.ollivier@free.fr)

<sup>3</sup> Institut de Mathématiques et Modélisation de Montpellier, UMR-CNRS 5149 Université de Montpellier II, 34095 Montpellier Cedex 05, France. E-mail: [Jean-Louis.Foulley@math.cnrs.fr](mailto:Jean-Louis.Foulley@math.cnrs.fr)

*of alleles specific to some breeds). These two types of genetic diversity appear to be independent across breeds and they both deserve being considered in defining conservation priorities. General principles of diversity analysis, common to the indicators reviewed, are emphasized, as well as their complementarities.*

## 1. Introduction

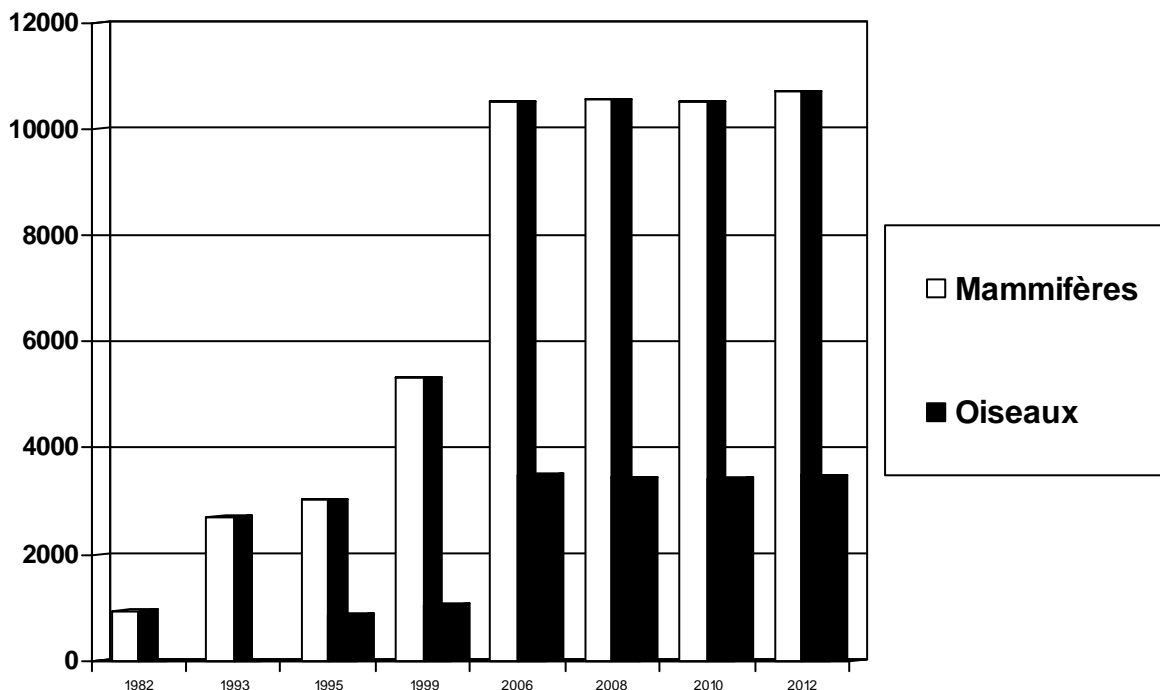
La diversité génétique peut se replacer dans le contexte plus général de la diversité biologique, telle qu'elle s'exprime, par exemple, en nombre d'espèces. Historiquement d'ailleurs, c'est la diversité des espèces qui a été le premier champ d'investigation des auteurs qui ont cherché, dès les années 1960, à définir des indicateurs de diversité biologique. Mais, s'agissant de diversité animale domestique, force est de reconnaître le faible nombre d'espèces que l'Homme a réussi à domestiquer. Les « 5 grands » sont le Bœuf, la Chèvre, le Mouton, le Porc et la Poule, dont les effectifs dans le monde sont de l'ordre du milliard d'individus pour les 4 premiers et approchent les 30 milliards pour la Poule. À ces cinq s'ajoute une trentaine d'espèces de mammifères et d'oiseaux (FAO, 2007 ; tableaux 6 et 7). C'est donc la diversité intra-spécifique que nous allons privilégier dans cet article.

La question est de définir ce qu'est la diversité génétique animale et de savoir comment on la mesure. En général, on s'intéressera à la diversité en termes d'éléments différents dans un ensemble, comme des *racés* d'une espèce dans une région donnée, ou encore des *gènes* dans une race donnée, ces éléments étant caractérisés, soit par leur abondance (ou fréquence) soit plus simplement par leur incidence (présence/absence), la diversité étant dans ce cas mesurée par le nombre d'éléments différents observés, ce qu'on appelle aussi une « richesse ». Quel que soit l'indice de diversité retenu, il sera généralement possible de décomposer la diversité en une diversité *intra*, par exemple intra-race ou intra-région, et une diversité *entre*, et ces deux composantes devront être prises en compte conjointement pour quantifier la diversité *totale*. Le principe de cette décomposition, établi par MacArthur *et al.* (1966) dans une étude sur la diversité des espèces, s'énonce comme suit. La diversité *entre* deux échantillons (appelée  $\beta$  en écologie) est la diversité combinée des deux échantillons (qui est aussi la diversité *totale*, appelée  $\gamma$ ) moins la diversité moyenne des deux échantillons (qui est la diversité *intra*, appelée  $\alpha$ ), de telle sorte que  $\alpha + \beta = \gamma$ . Ce principe se généralise facilement à un nombre quelconque d'échantillons. Dans ce cadre général s'insèrent les divers indicateurs de diversité génétique animale que nous allons passer en revue dans cet article.

## 2. Les inventaires de races

Un premier niveau d'évaluation de la diversité est à l'évidence la race. Depuis le *Dictionnaire des races de bétail* de Mason (1951), les inventaires de races rassemblent des informations visant à couvrir l'ensemble des races de chaque espèce. Les inventaires sont aujourd'hui réalisés et mis en base de données, à des échelles nationale, régionale ou mondiale. Le recueil des informations au niveau mondial a été lancé par la Fédération Européenne de Zootechnie au début des années 1980, et poursuivi ensuite à partir de 1993 sous l'égide de la FAO, dans le prolongement de la Conférence de Rio sur la Biodiversité de 1992. La FAO a mis en place progressivement un système d'information sur la Diversité Animale Domestique (appelé DAD), consultable en ligne sur <http://dad.fao.org/>. Les races animales domestiques peuvent être classées selon une typologie évolutive, qui distingue quatre catégories de populations, des populations sauvages, des populations primaires (ou traditionnelles), des races standardisées et des lignées sélectionnées (Lauvergne, 1982). Ces deux dernières catégories constituent la majeure partie de l'ensemble. Au total, environ 14 000 races sont aujourd'hui répertoriées à travers le monde par le système DAD. Ce total recouvre 37 espèces de mammifères et d'oiseaux exploitées par l'Homme. La Figure 1 indique l'évolution du nombre des races recensées au niveau mondial au cours des 30 dernières années. On peut voir que, depuis la mise en place du système DAD par la FAO en 1993, ce nombre a été multiplié par 11 pour les races de Mammifères et par 4 pour les races aviaires. On note aussi un doublement du nombre total de races inventoriées entre 1999 et 2006, ce qui correspond au lancement par la FAO en 1999 de la préparation d'un rapport sur l'état des ressources génétiques animales dans le monde, qui fut publié en 2007 (FAO, 2007).

FIGURE 1



Évolution du nombre des races de mammifères et d'oiseaux recensées dans le monde de 1982 à 2012

Sources :

- 1982 : Maijala *et al.*, 1984
- 1993-2012 : FAO, 2012a

Cette richesse raciale est décomposable en une diversité nationale (intra-pays), une diversité régionale (intra-région du globe), et une diversité interrégionale (entre les 7 régions du globe définies par la FAO), dont le total représente le nombre de *races différentes* répertoriées dans le monde. Cette décomposition prend en compte l'existence de races dites transfrontalières, qui sont communes à plusieurs pays, à l'exemple de la race bovine Charolaise, que l'on trouve représentée dans 69 pays, ou encore de la race bovine Holstein, qui est quasi-universellement distribuée puisque son existence est signalée dans 160 pays (selon DAD). Le principe de décomposition décrit dans l'introduction doit ici être étendu au cas de deux niveaux de hiérarchie, soient la région du globe et le pays intra-région. La composante intra-pays ( $\alpha$ ) est le nombre moyen de races par pays, et la composante entre pays ( $\beta$ ) se répartit en une composante entre pays intra-région ( $\beta_1$ ) et une composante entre régions ( $\beta_2$ ), de telle sorte que  $\alpha + \beta_1 + \beta_2 = \gamma$ , où  $\gamma$  représente le nombre total de races. La procédure est décrite en détail par Veech *et al.* (2002 ; voir notamment la figure 4 de l'article). Sur l'ensemble des mammifères et oiseaux répertoriés dans DAD (FAO, 2012a), les composantes nationale ( $\alpha$ ), régionale ( $\beta_1$ ) et interrégionale ( $\beta_2$ ) représentent respectivement 1 %, 26 % et 73 % de la diversité raciale mondiale. Il vaut d'être noté que des décompositions similaires s'obtiennent pour chacune des grandes espèces traitée séparément (voir Tableau). On peut donc dire que la diversité en nombre de races animales domestiques s'exprime pour près des trois quarts sous la forme d'une diversité entre les grandes régions du globe.

Tableau : Analyse de la diversité des races animales domestiques recensées dans 182 pays et 7 régions du globe (situation 2012)

## MESURE ET ÉVOLUTION DE LA DIVERSITÉ GÉNÉTIQUE DES PLANTES CULTIVÉES ET DES ANIMAUX DOMESTIQUES

(Source : FAO, 2012a)

ESPÈCE	Bœuf	Cheval	Chèvre	Mouton	Porc	Poule	Toutes espèces
Nombre de races	3080	1442	1200	2416	1293	2371	14194
Nombre de races locales	1004	618	557	1106	556	1269	6581
Nombre de races régionales	99	62	47	128	28	48	504
Nombre de races internationales	107	63	38	101	30	106	549
Nombre de races différentes * ( $\gamma$ )	1210	743	642	1335	614	1423	7634
Diversité intra pays ( $\alpha$ )	17	8	7	13	7	13	78
Diversité intra région ( $\beta_1$ )	423	198	165	332	178	326	1950
Diversité entre régions ( $\beta_2$ )	770	537	471	990	429	1084	5606
$\alpha/\gamma$	0,01	0,01	0,01	0,01	0,01	0,01	0,01
$\beta_1/\gamma$	0,35	0,27	0,26	0,25	0,29	0,23	0,26
$\beta_2/\gamma$	0,64	0,72	0,73	0,74	0,70	0,76	0,73

\*Ce nombre est la somme des nombres de races locales, régionales et internationales

### 3. La caractérisation des races

Les inventaires comportent le recueil d'informations sur chaque race, pour une *caractérisation phénotypique* dont le but est d'évaluer les différences entre races et de déceler, si possible, la part d'origine génétique de ces différences (FAO, 2012b). Les caractères considérés sont d'abord les caractères visibles et généralement qualitatifs, comme la couleur de la robe ou du plumage, et de nombreux caractères morphologiques. Il s'agit souvent de caractères dont le déterminisme génétique est assez simple et qui sont peu influencés par le milieu. À l'opposé, les performances zootechniques sont des caractères quantitatifs très dépendants du milieu. L'estimation des différences génétiques entre races pour ces caractères requiert la mise en œuvre de protocoles expérimentaux rigoureux, dont le coût devient vite rédhibitoire à mesure que le nombre des races à évaluer augmente et que leur dispersion à travers le monde s'élargit.

La diversité en nombre de races reste malgré tout un indicateur de diversité génétique, puisque l'isolement géographique et reproductif des races implique de la dérive génétique et que s'y ajoutent les changements génétiques induits par la sélection des éleveurs pour des objectifs différents selon les régions, ou qui résultent de l'adaptation des races à leur environnement de production et au milieu naturel propre à chaque région. Mais si on mesure la diversité génétique sur la base du nombre de races, on s'expose à la surévaluer dans les régions ayant une longue tradition de sélection organisée pouvant aboutir à distinguer des races très apparentées, et, inversement à la sous-estimer dans les pays à sélection peu structurée, où les races peuvent recouvrir une diversité de sous-types non encore identifiés (voir FAO, 2007, 5<sup>e</sup> partie, section A). Les bases génétiques de la diversité raciale devront être précisées, comme nous le verrons plus loin, par l'étude des marqueurs génétiques.

### 4. Parenté, consanguinité et démographie

La diversité génétique intra-race conditionne l'efficacité de tout programme d'amélioration génétique, puisque la variabilité génétique de la population est la matière première qu'exploite le sélectionneur. Le paramètre de base est le *coefficient de consanguinité*  $f$  d'un individu, défini, selon Malécot (1948), comme la probabilité pour que deux gènes du même locus soient identiques, *i. e.* dérivent par descendance mendélienne d'un même gène ancêtre. Comme l'un des gènes de l'individu provient de son père et l'autre de sa mère,  $f$  est aussi le *coefficient de parenté* de ses parents. Rappelons la définition classique et le mode de calcul de  $f$ , pour un individu dont les parents ont un ancêtre commun A, lui-même consanguin à hauteur de  $f_A$ , et que  $n_1$  et  $n_2$  générations séparent respectivement des parents de l'individu considéré, soit

## MESURE ET ÉVOLUTION DE LA DIVERSITÉ GÉNÉTIQUE DES PLANTES CULTIVÉES ET DES ANIMAUX DOMESTIQUES

---

$f = (1 + f_A)(0,5)^{n_1+n_2+1}$ . En totalisant les  $f$  obtenus pour chacun des ancêtres communs des parents de l'individu on obtient sa consanguinité.

La connaissance des pedigrees permet de calculer à tout instant la consanguinité moyenne de la population et ainsi de suivre son évolution au cours du temps. Des logiciels bien adaptés existent pour réaliser ces calculs (Boichard, 2007 ; Colleau, 2002). En plus de la consanguinité classique, ces programmes de calcul permettent des analyses plus poussées en évaluant des caractéristiques telles que l'*intervalle de génération*, sur la base des dates de naissance, les *probabilités d'origine des gènes* permettant de déduire les contributions relatives de différents ancêtres à la population présente, le nombre d'ancêtres de chaque individu, et un *nombre efficace d'ancêtres fondateurs* dont la population présente dérive.

Une belle illustration de ces concepts nous est fournie par le pedigree du cheval *Pur-Sang* analysé par une équipe de chercheurs irlandais (Cunningham et al. 2001). Le *Pur-Sang* est sans doute la race animale la plus ancienne pour ce qui est de la tenue de pedigree, puisque son pedigree remonte à 1791 et que, de plus, la race est considérée comme pratiquement fermée depuis cette date. Le pedigree étudié par les irlandais comporte près de un million de chevaux et peut être considéré comme complet, ce que tend à confirmer par ailleurs la relation quasi-linéaire entre les coefficients de parenté d'un échantillon de quelques 200 chevaux contemporains pris deux à deux et la proportion d'allèles qu'ils ont en commun. Le nombre d'ancêtres fondateurs des chevaux actuels est de 158, mais leurs contributions très inégales font que le nombre efficace n'est que de 28, ce qui montre la base génétique exceptionnellement étroite de cette race. Sur l'ensemble des races domestiques, les méthodes modernes de sélection tendent à accélérer le mouvement général d'accroissement de consanguinité en fonction du temps, un mouvement qui est considéré comme néfaste, du double point de vue de la réduction potentielle de la diversité sélectionnable et des effets dépressifs de la consanguinité sur les performances. Un historique des méthodes de gestion de ce risque à l'aide des pedigrees pour plusieurs espèces animales domestiques a été présenté par Colleau et al. (2007).

En l'absence de pedigrees bien tenus, la consanguinité ne peut évidemment pas être directement mesurée. Le nombre des reproducteurs en service devient alors le meilleur indicateur de diversité. Le critère le plus utilisé est la *taille efficace* de la population, qui a été définie par Wright (1931) comme  $N_e = 4MF/(M + F)$ , expression qui combine le nombre des reproducteurs mâles ( $M$ ) et femelles ( $F$ ) de la population. On voit en faisant  $M=F$  que  $N_e$  équivaut à une population constituée de  $N_e/2$  reproducteurs de chaque sexe. L'évolution attendue de la consanguinité moyenne en fonction de  $N_e$  et du numéro de génération  $g$  s'écrit  $f_g = 1 - e^{-g/2N_e}$ , d'où un accroissement de consanguinité approximatif par génération  $\Delta f = 1/2N_e$ .

Les prédictions de consanguinité basées sur la taille efficace supposent cependant que les reproducteurs soient accouplés au hasard et qu'il n'y ait aucune sélection. Ces prédictions sous-estiment en général la consanguinité réelle des populations sélectionnées, puisque la sélection entraîne un supplément de consanguinité (Robertson, 1961). Notons aussi que, à nombre de reproducteurs constant, des paramètres démographiques comme l'âge des reproducteurs à la naissance de leur premier descendant et le taux de leur renouvellement exercent des effets sur l'accroissement de consanguinité. Par ailleurs, les mutations et les migrations peuvent jouer en sens inverse en freinant les augmentations de consanguinité. Les paramètres de génétique des populations à prendre en considération sont alors  $4N_e\mu$  et  $4N_em$ ,  $\mu$  et  $m$  étant les taux respectifs de mutation et de migration par génération (voir Hartl et Clark, 1997). Pour ce qui est des mutations, leur rôle dans le maintien de la variabilité génétique dans un contexte de sélection naturelle au sein de populations de petite taille est maintenant bien établi (Houle et al. 1996), et il n'est pas exclu qu'elles puissent jouer un rôle similaire dans un contexte de sélection artificielle (Hill et Zhang, 2009). Quant aux migrations, les races animales n'étant généralement pas strictement fermées, les échanges entre races jouent un rôle voisin de celui des migrations dans les populations naturelles, dont les effets dans le maintien de la diversité sont connus pour être significatifs, même pour des taux de migration  $m$  modérés. Slatkin (1985) a montré la relation inverse qui existe entre  $N_em$  et la proportion des allèles présents dans une seule race sur l'ensemble des allèles présents, que depuis Neel (1973) on appelle des allèles « privés ». Sur cette base, les valeurs assez faibles des richesses alléliques privées (voir plus loin § 5.2) indiqueraient des valeurs non négligeables de  $N_em$  et donc un rôle significatif des migrations dans le maintien de la diversité.

L'intérêt de ce paramètre de taille efficace est aussi de fournir une base objective pour évaluer le risque d'extinction d'une race, qu'on sait être lié à sa consanguinité et aux effets dépressifs de celle-ci, et

pour choisir les races justiciables en priorité d'un programme de conservation. Une taille critique peut ainsi être définie par exemple sur la base d'une consanguinité à ne pas dépasser au bout d'un temps  $t$ . A titre d'exemple, 3 étalons suffisent pour assurer une consanguinité ne dépassant pas 25 % au bout de 50 ans, alors qu'il faut 36 coqs pour obtenir le même résultat, compte tenu des longueurs de génération de ces espèces (Ollivier, 1998). Dans le même ordre d'idées, des règles fondées sur des valeurs minimum de  $N_e$  ont été édictées comme signal d'alarme sur un danger de mise en péril voire d'extinction de certaines espèces. Un exemple en est fourni par la règle des 50/500 (Franklin, 1980 ; Jamieson et Allendorf, 2012). Notons cependant que l'évaluation du risque d'extinction d'une race doit prendre en compte d'autres facteurs que sa diversité génétique, notamment sa démographie et les conditions socio-économiques de son élevage. C'est un sujet que nous ne pouvons détailler ici et pour lequel nous renvoyons à une étude de Gandini *et al.* (2004), qui analyse l'évolution de la taille de 110 races bovines européennes et les facteurs susceptibles d'influencer cette évolution, comme par exemple la taille des troupeaux. Notons que l'un des objectifs des inventaires de race décrits précédemment est précisément de recueillir le maximum d'information sur ces facteurs pour une évaluation pertinente des risques d'extinction.

## 5. Les marqueurs génétiques: d'autres inventaires

Face à la difficulté évoquée plus haut d'estimer la part génétique des différences phénotypiques entre races, les marqueurs génétiques offrent l'avantage d'une estimation objective de la diversité des génomes, totalement indépendante des conditions de milieu. Les *groupes sanguins* et les *polymorphismes biochimiques* des animaux domestiques sont étudiés depuis le début des années 60 et ils ont dès cette époque été utilisés pour comparer des races et étudier les relations entre elles (Grosclaude *et al.*, 1990). Ce n'est cependant qu'à partir des années 1990, avec le développement de la cartographie génétique et des *marqueurs moléculaires* tels que microsatellites et AFLP, que des marqueurs ont pu être *localisés avec précision* sur les chromosomes, et en *nombre suffisant* pour réaliser une couverture adéquate des génomes et assurer ainsi la représentativité des marqueurs choisis. Enfin, une étape importante a aussi été franchie avec la mise au point de *procédures automatiques* permettant le recueil à moindre coût d'une vaste information, tant en nombre de marqueurs qu'en nombre de populations examinées. En France, la première application de ces techniques est l'étude des races bovines de Moazami-Goudarzi *et al.* (1997) utilisant les microsatellites. La mise en pratique du marquage moléculaire à grande échelle a favorisé le lancement de nombreux projets internationaux, notamment des consortiums financés par l'Union Européenne concernant les principales espèces animales domestiques (voir des références à ce sujet dans Ollivier et Foulley, 2009).

La pertinence des marqueurs moléculaires dans l'évaluation de la diversité a été mise en cause du fait de la neutralité supposée de certains types d'entre eux (microsatellites par exemple), leur diversité pouvant ne pas refléter celle des gènes d'intérêt zootechnique. Mais nous savons par ailleurs que des gènes neutres peuvent être affectés par la sélection qui s'exerce sur des gènes voisins, ce qu'on appelle l'*effet d'entraînement* (Maynard-Smith et Haig, 1974). De cette manière la diversité génétique aux locus sélectionnés « entraînerait » une diversité corrélative de gènes neutres voisins. Ce phénomène peut être vu comme la situation exploitée en sens inverse dans la sélection assistée par marqueurs, où la sélection sur des marqueurs neutres affecte les locus de caractère quantitatifs (QTL) voisins. Le recensement des QTL dans les espèces animales montre en effet que des marqueurs liés à des caractères zootechniques sont aujourd'hui identifiés sur toute l'étendue des génomes (Hu *et al.*, 2013). Une couverture adéquate du génome par un ensemble de marqueurs doit donc, en théorie, assurer une corrélation de la diversité-marqueur avec au moins quelques performances.

Nous distinguerons dans ce qui suit les deux types de diversité que les marqueurs génétiques peuvent mesurer, soit la diversité génétique *stricto sensu* et la diversité allélique. La première, qui est la plus classiquement utilisée, se rattache à la catégorie générale des diversités basées sur l'abondance (évoquée dans l'introduction), alors que la seconde peut être vue comme une diversité d'incidence.

### 5.1 Diversité génétique *stricto sensu*

La méthode classique de mesure de la diversité génétique des populations est celle des indices de fixation ( $F$ ) proposée par Wright en 1943. Wright définit ainsi un  $F_{ST}$  qui mesure le degré de différenciation entre des sous-populations ( $S$ ) résultant du fractionnement d'une population totale ( $T$ ). Cet indice de différenciation, qui était vu par Wright comme une généralisation du concept de consanguinité, peut aussi se

rattacher au concept de diversité selon l'approche de Nei (1973), la diversité étant assimilée à l'hétérozygotie moyenne ( $H$ ) attendue à plusieurs locus sous l'hypothèse de l'équilibre de Hardy-Weinberg.  $H$  est donc une quantité mesurable connaissant les fréquences alléliques  $p_i$ , puisque  $H = 1 - \sum_i p_i^2$ . Nei montre que la diversité totale ( $H_T$ ) se décompose en une diversité intra-sous-population ( $H_S$ ) et une diversité entre sous-populations ( $D_{ST}$ ). On a  $H_T = H_S + D_{ST}$ , et Nei définit ainsi un coefficient de différenciation génique entre les sous-populations tel que  $G_{ST} = D_{ST}/H_T = 1 - H_S/H_T$ , équivalent au  $F_{ST}$  de Wright. La confrontation des valeurs observées et espérées de cet indice sous certaines hypothèses génétiques permet de tester la vraisemblance de ces hypothèses : par exemple neutralité vs sélection (Lewontin et Krakauer, 1973 ; Gautier et al., 2010).

Une autre méthode de mesure de la différenciation entre populations (ou races) est basée sur les distances génétiques de ces populations entre elles deux à deux et sur le principe selon lequel l'ajout d'une population  $i$  à un ensemble ( $R$ ) de populations doit augmenter la diversité ( $V$ ) de cet ensemble d'une quantité au moins égale à la distance entre la population  $i$  et l'ensemble  $R$ , soit  $d(i, R)$ . L'application de ce principe conduit à définir la diversité  $V$  d'un ensemble  $R$  de populations comme le maximum, sur toutes les populations de l'ensemble, de la distance entre une population  $i$  donnée et l'ensemble  $R$  privé de cette population, soit  $d(i, R \setminus i)$ , plus la diversité de l'ensemble  $R$  privé de cette population. La diversité de  $R$  s'écrit alors :  $V(R) = \max_{i \in S} [d(i, R \setminus i) + V(R \setminus i)]$  (Weitzman, 1992). La méthode, à l'origine proposée pour l'étude de la diversité des espèces, est facilement transposable aux races animales, comme cela a été montré par Thaon d'Arnoldi et al. (1998) avec une application à un ensemble de races bovines. L'utilisation de cette fonction de diversité dans la gestion de la diversité génétique animale est décrite en détail par Foulley et Ollivier (2007), qui référencent également les nombreuses applications auxquelles la méthode a donné lieu dans ce secteur.

L'établissement d'une fonction de diversité ouvre la voie à des analyses plus fines en vue de mesurer les *contributions individuelles* de races animales à la diversité d'un ensemble de races, puisque la fonction de diversité est applicable à n'importe quel sous-ensemble. Weitzman (1993) a sans doute été le premier à recommander l'usage de ces contributions pour définir des politiques rationnelles de conservation. C'est ainsi qu'en utilisant la fonction  $V(R)$  définie plus haut la contribution d'une race  $i$  quelconque à la diversité d'un ensemble de races  $R$  peut s'écrire  $CB_i = 1 - V(R \setminus i)/V(R)$ , une quantité qui mesure en quelque sorte l'originalité génétique de la race  $i$  par rapport à l'ensemble. Petit et al. (1998) ont proposé une approche similaire pour les diversités de Nei (1973), qui consiste à remplacer  $V$  dans la formule précédente par les quantités  $H_T$ ,  $H_S$ , ou  $D_{ST}$  définies plus haut, de manière à définir des contributions  $C_T$ ,  $C_S$ , ou  $C_D$ , selon que l'intérêt se porte sur la diversité totale ou sur ses composantes.

## 5.2 Diversité (richesse) allélique

Le nombre d'allèles par locus est un indicateur de diversité d'un grand intérêt en sélection. Alors que l'hétérozygotie est liée à la réponse immédiate à la sélection, la réponse à long-terme dépend plutôt du nombre des allèles par locus (Barker, 2001). La richesse allélique est aussi un critère utile dans la pratique de la conservation puisque la maximisation assistée par marqueurs de la richesse allélique se révèle efficace dans la conservation d'un maximum d'allèles aussi bien neutres que non neutres (Bataillon et al., 1996). Mais le nombre d'allèles observé dans une race tend à augmenter avec la taille de l'échantillon examiné, ce qui oblige à tenir compte de celle-ci dans les comparaisons entre races. C'est un problème que l'on rencontre aussi en écologie dans l'évaluation du nombre d'espèces. La correction par *raréfaction*, proposée par El Mousadik et Petit (1996), consiste à définir la richesse allélique à partir du nombre d'allèles manquants attendus par rapport au nombre total d'allèles observés, dans l'hypothèse où l'échantillon serait d'une taille inférieure à celle de l'échantillon réel. On choisit généralement comme taille de référence ( $g$ ) à un locus donné le nombre minimum de gènes examinés par race sur l'ensemble des races examinées à ce locus. Un allèle  $k$ , observé  $N_k$  fois parmi les  $N$  gènes de la race examinée, a une probabilité de manquer dans un échantillon de taille  $g < N$  qui est égale à  $P_k^{(g)} = C(N - N_k, g)/C(N, g)$ , expression dans laquelle  $C(N, g)$  représente le nombre de combinaisons de  $N$  objets pris  $g$  à  $g$  (ou  $N - N_k$  objets dans le cas du numérateur).

L'allèle  $k$  considéré a donc une probabilité  $1 - P_k^{(g)}$  d'être présent dans un échantillon de taille  $g$ . En sommant

sur les  $K$  allèles observés dans une race on obtient sa richesse allélique  $r^{(g)} = \sum_{k=1}^K (1 - P_k^{(g)}) = K - \sum_{k=1}^K P_k^{(g)}$ . Une

autre méthode de correction consiste à ajouter au nombre d'allèles observés le nombre d'allèles qu'on s'attend à voir manquer du fait de la taille limitée de l'échantillon. C'est la méthode d'*extrapolation* proposée par Foulley et Ollivier (2006a), que ces auteurs ont appliquée et comparée à la méthode de raréfaction sur des échantillons de l'espèce porcine et de l'arganier marocain.

De la richesse allélique on passe à la diversité allélique en définissant une *diversité allélique intra-race* et une *diversité allélique totale*, selon le principe exposé dans l'Introduction et appliqué plus loin à la diversité des races. Si les échantillons raciaux sont de taille égale, la diversité totale est le nombre total d'allèles observés sur l'ensemble des races et la diversité intra-race est la moyenne des nombres d'allèles observés dans chaque race. La *diversité allélique entre races* s'obtient par différence entre les deux composantes ainsi définies et on en déduit un coefficient de différenciation allélique, analogue au  $G_{ST}$  de Nei. Les nombres d'allèles observés devront être corrigés par les méthodes indiquées plus haut en cas de taille inégale des échantillons raciaux.

Les contributions individuelles des races à la diversité allélique reposent sur l'existence d'allèles spécifiques à certaines races, que depuis Neel (1973) on appelle des allèles « privés » (voir plus haut). La contribution d'une race à la diversité allélique d'un ensemble de races se définit comme le nombre d'allèles privés par locus observés dans la race. Elle se décompose, comme pour la diversité génétique *stricto sensu*, en une contribution à la diversité allélique intra-race et une contribution à la diversité allélique entre races. Dans le cas le plus général d'échantillons raciaux de taille inégale, la contribution de chaque race à la diversité allélique se définit comme l'espérance du nombre d'allèles présents dans cette race et absents dans toutes les autres (Kalinowski, 2004). Sur la base de la probabilité d'absence  $P_k^{(g)}$  définie plus haut pour un allèle  $k$  donné, la contribution de la race  $i$  est donc  $(1 - P_i^{(g)}) \prod_{j \neq i} P_j^{(g)}$  pour cet allèle, quantité qu'il suffit de sommer sur les allèles de la race pour obtenir sa contribution au titre du locus considéré. La contribution totale est la moyenne des contributions sur l'ensemble des locus considérés.

Cette méthodologie d'analyse de la diversité allélique a maintenant été appliquée dans plusieurs espèces animales et végétales (références dans Ollivier et Foulley, 2013). Les études réalisées jusqu'à présent montrent que diversité génétique *stricto sensu* et diversité allélique sont deux formes différentes de diversité, quasi-indépendantes entre elles. La Figure 2 illustre cette situation dans une analyse de microsatellites sur 10 races bovines (Foulley et Ollivier, 2006b). Les exemples abondent de races représentant une part importante de diversité sans qu'elles aient le moindre allèle spécifique, tout autant que de situations inverses. Aucun de ces indicateurs de diversité ne doit donc être négligé quand on définit des priorités de conservation.



FIGURE 2

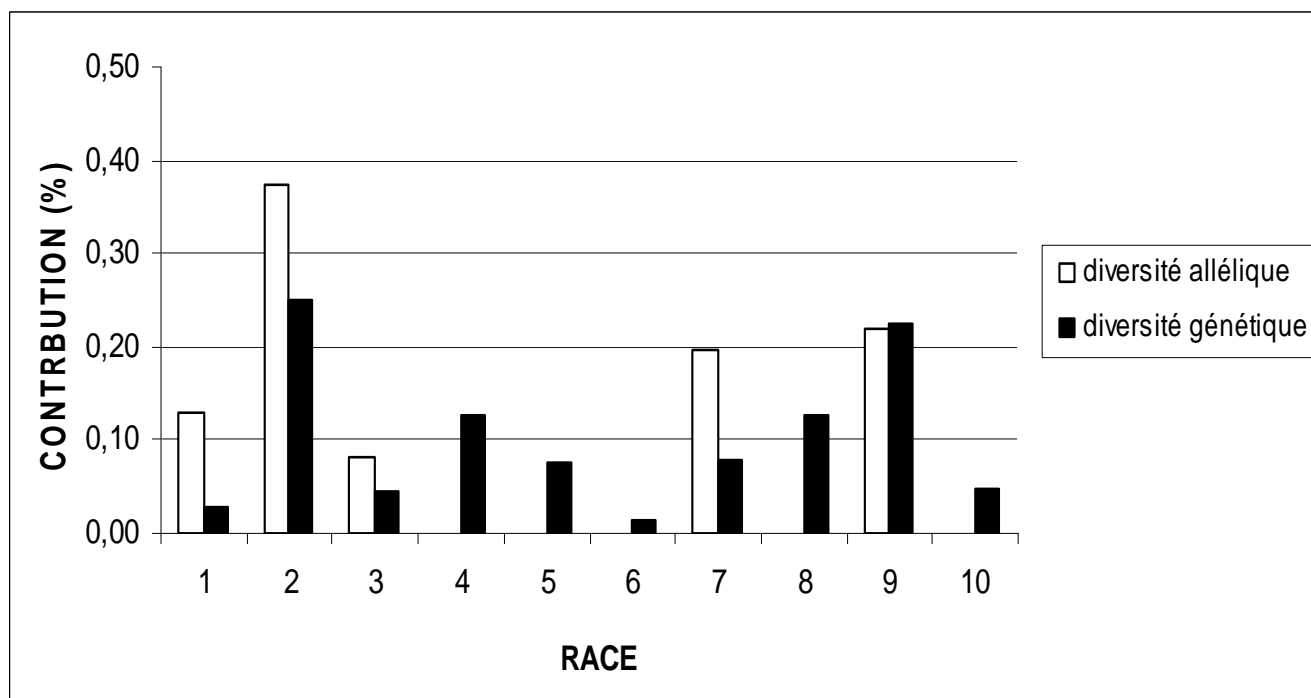


Figure 2 Analyse de la diversité génétique de 9 locus de microsatellites sur 10 races bovines françaises : contributions raciales à la diversité exprimées relativement à la diversité totale (données analysées dans Foulley et Ollivier, 2006b)

Voir texte pour la définition des contributions à la diversité allélique et à la diversité génétique *stricto sensu*

Les races sont : 1 Bretonne Pie Noire ; 2 Charolais ; 3 Holstein ; 4 Jersey ; 5 Limousin ; 6 Rouge des Prés (anciennement Maine-Anjou) ; 7 Montbéliard ; 8 Normand ; 9 Parthenais ; 10 Vosgienne.

## 6. Conclusion

Les indicateurs de diversité génétique que nous avons passés en revue nous disent, chacun à leur façon, quelque chose qui a trait à la diversité génétique, mais aucun d'eux n'est tout à fait exempt de critiques. Le flou qui entoure le concept de *race*, notamment selon la région du globe considérée, enlève beaucoup de valeur au critère du nombre de races. L'utilité des *pedigrees* pour suivre la diversité génétique intra-race est très dépendante de la qualité et de l'exhaustivité des informations recueillies. Les *marqueurs moléculaires* sont réputés neutres et leurs relations avec les gènes d'intérêt zootechnique restent encore à évaluer avec précision. Il s'ensuit qu'il n'est pas facile de se prononcer sur l'état de la diversité génétique animale à un moment donné, et qu'il est encore plus difficile d'avoir une vue rétrospective de son évolution dans le temps, même si des *pedigrees* bien tenus sur de longues périodes offrent un suivi des diversités internes aux races.

Un intérêt pour la diversité génétique animale domestique n'est en fait apparu qu'assez tardivement, et avec un décalage notable par rapport au domaine végétal. C'est à partir des années 1960 que s'est manifestée une inquiétude sur le maintien de la diversité génétique animale face à une tendance générale vers l'abandon des races les moins productives et, au sein des grandes races, l'utilisation intensive de l'insémination artificielle. Nous avons vu qu'il a fallu cependant attendre une bonne vingtaine d'années pour que des actions concrètes soient prises pour mieux connaître et mieux gérer les ressources génétiques animales. Un

aspect important a aussi été le renforcement de la coopération internationale, dont témoigne l'énorme travail d'évaluation de l'état de ces ressources à travers le monde (FAO, 2007).

La caractérisation moléculaire des ressources est aujourd'hui reconnue comme un outil essentiel dans leur gestion. Les marqueurs moléculaires offrent à l'évidence un moyen d'objectiver le concept de diversité, et, à ce titre, ils contribuent souvent à dissiper le flou entourant l'idée de race et à préciser les liens existant entre diverses races. Ils permettent même d'inférer une structure raciale latente dans des populations n'offrant aucune apparence de subdivision en races (Pritchard et al., 2000). Ces marqueurs pourront aussi à l'avenir servir à suivre l'évolution de la diversité génétique animale dans le temps, ce qui jusqu'à présent n'a pas été possible, faute de disposer du recul nécessaire dans la constitution de stocks de matériel biologique (sources d'ADN), contrairement au secteur végétal (Gallais, 2013). De ce point de vue, les centres de ressources biologiques génomiques qui se mettent en place, à côté des ressources reproductives des cryobanques (Danchin-Burge, 2012), devraient contribuer utilement à la coordination des échantillonnages et à la conservation sur le long terme des matériels récoltés. Sur un plan plus général, l'étude des marqueurs moléculaires offre des perspectives pour une meilleure connaissance des génomes animaux et pour élucider les bases génétiques des divergences adaptatives des races animales à travers le monde.

### RÉFÉRENCES BIBLIOGRAPHIQUES

- (1) BARKER, J.S.F., 2001. – Conservation and management of genetic diversity: a domestic animal perspective. *Can. J. For. Res.*, **31**, 588-595.
- (2) BATAILLON T.M., DAVID J.L., SCHOEN D.J., 1996. – Neutral genetic markers and conservation genetics : simulated germplasm collections. *Genetics*, **144**, 409-417.
- (3) BOICHARD D., 2007. – Pedig : a Fortran package for pedigree analysis suited to large populations. ([http://www-sgqa.jouy.inra.fr/rubrique.php3?id\\_rubrique=5](http://www-sgqa.jouy.inra.fr/rubrique.php3?id_rubrique=5))
- (4) COLLEAU J.-J., 2002. – An indirect approach to the extensive calculation of relationship coefficients. *Genet. Sel. Evol.*, **34**, 409-421.
- (5) COLLEAU J.-J., REGALDO D., PALHIÈRE I., TRIBOUT T., MOUREAUX S., FRITZ S., BARBAT A., DE PREAUMONT H., TUAL K., MATTALIA S., 2007. – Gestion optimisée de la variabilité génétique dans les populations sélectionnées. Académie d'Agriculture de France, Séance du 7 mars 2007. ([http://www.academieagriculture.fr/mediatheque/seances/2007/0702007/20070307communication1\\_integral.pdf](http://www.academieagriculture.fr/mediatheque/seances/2007/0702007/20070307communication1_integral.pdf))
- (6) CUNNINGHAM E.P., DOOLEY J.J., SPLAN R.K., BRADLEY D.G., 2001. – Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. *Anim. Genet.*, **32**, 360-364.
- (7) DANCHIN-BURGE C., 2012. – Gestion durable de la variabilité génétique des ruminants domestiques : approches *in situ* et *ex situ*. Thèse AgroParisTech.
- (8) EL MOUSADIK A., PETIT R.J., 1996. – High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* L. Skeels) endemic to Morocco. *Theoret. Appl. Genet.*, **92**, 832-836.
- (9) FAO, 2007. – The State of the World's Animal Genetic Resources for Food and Agriculture. Rischkowsky B. and Pilling D. (Eds), FAO, Rome, 511 p.
- (10) FAO, 2012a. – Status and trends of animal genetic resources. Document distribué à la 7e séance du groupe de travail technique intergouvernemental sur les ressources génétiques animales pour l'alimentation et l'agriculture, Rome, 24-26 octobre 2012, 40 p.
- (11) FAO, 2012b. – Phenotypic characterization of animal genetic resources. FAO Animal Production and Health guidelines No 11. Rome, 142 p.
- (12) FOULLEY J.-L., OLLIVIER L., 2006a. – Estimating allelic richness and its diversity. *Livest. Sci.*, **101**, 150-158.
- (13) FOULLEY J.-L., OLLIVIER L., 2006b. – Diversité génétique et richesse allélique : concepts et application à des races bovines. Journées 3R, Paris, 4 p.

MESURE ET ÉVOLUTION DE LA DIVERSITÉ GÉNÉTIQUE DES PLANTES CULTIVÉES ET DES ANIMAUX DOMESTIQUES

---

- (14) FOULLEY J.-L., OLLIVIER L., 2007. – Mesure de la diversité génétique : l'apport de Martin Weitzman. Académie d'Agriculture de France, Séance du 7 mars 2007. ([http://www.academie-agriculture.fr/mediatheque/seances/2007/07-03-2007/20070307communication2\\_integral.pdf](http://www.academie-agriculture.fr/mediatheque/seances/2007/07-03-2007/20070307communication2_integral.pdf))
- (15) FRANKLIN I.R., 1980. – Evolutionary change in small populations. In Conservation Biology: an Evolutionary–Ecological Perspective (Soulé M.E. and Wilcox B.A., eds), pp. 135–150, Sinauer Associates, Sunderland, Massachusetts, USA.
- (16) GALLAIS A., 2013. – Évolution de la diversité génétique des variétés de plantes cultivées. Académie d'Agriculture de France, Séance du 23 janvier 2013 ([http://www.academie-agriculture.fr/mediatheque/seances/2013/20130123\\_resume2.pdf](http://www.academie-agriculture.fr/mediatheque/seances/2013/20130123_resume2.pdf))
- (17) GANDINI G.C., OLLIVIER, L., DANELL, B., DISTL, O., GEORGOUDIS, A., GROENEVELD, E., MARTYNIUK, E., VAN ARENDONK, J.A.M., AND WOOLLIAMS, J., 2004. – Criteria to assess the degree of endangerment of livestock breeds in Europe. *Livest. Prod. Sci.*, **91**, 173-182.
- (18) GAUTIER M., HOCKING T. D., FOULLEY J.-L., 2010. – A Bayesian outlier criterion to detect SNPs under selection in large data sets. *PLoS ONE*, **5**, Issue 8, e11913.
- (19) GROSCLAUDE F., AUPETIT R.Y., LEFEBVRE J., MERIAUX J.C., 1990. – Essai d'analyse des relations génétiques entre les races bovines françaises à l'aide du polymorphisme biochimique. *Génét. Sél. Evol.*, **22**, 317-338.
- (20) HARTL D.L., CLARK A.G., 1997. – Principles of Population Genetics (3e ed). Sinauer Associates, Sunderland, Massachusetts, USA, 542 p.
- (21) HILL W.G., ZHANG X.S., 2009. – Maintaining genetic variation in fitness. In: van der Werf, J., Graser, H.-U., Frankham, R. and Gondro, C. (Eds) *Adaptation and Fitness in Animal Populations*, Springer, Berlin, pp. 59-81.
- (22) HOULE D., MORIKAWA B., LYNCH M., 1996. – Comparing mutational variabilities. *Genetics*, **143**, 1467-1483.
- (23) HU Z.-L., PARK C. A., WU X.-L., REECY J. M., 2013. – Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Res.*, **41**, Database issue D871–D879.
- (24) JAMIESON I.G., ALLENDORF F.W., 2012. – How does the 50/500 rule apply to MVPs? *Trends Ecol. Evol.*, **27**, 578-584.
- (25) KALINOWSKI S.T., 2004. – Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conserv. Genet.*, **5**, 539-543.
- (26) LAUVERGNE J.J., 1982. – Genetica en poblaciones animales después de la domestication: consecuencias para la conservacion de las razas. *Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production*, **6**, 77-87 .
- (27) LEWONTIN R.C., KRAKAUER J., 1973. – Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, **74**, 175-195.
- (28) MACARTHUR R., RECHER H., CODY M., 1966. – On the relation between habitat and species diversity. *Am. Nat.*, **100**, 319-332.
- (29) MAIJALA K., CHEREKAEV E.V., DEVILLARD J.M., REKLEWSKI Z., ROGNONI G., SIMON D.L., STEANE D.E., 1984. – Conservation of animal genetic resources in Europe. Final report of an EAAP working group. *Livest. Prod. Sci.* **11**, 3-22.
- (30) MALÉCOT G., 1948. – Les mathématiques de l'hérédité. Masson, Paris, 64 p.
- (31) MASON I.L., 1951. – A World Dictionary of Livestock Breeds Types and Varieties. Commonwealth Agricultural Bureaux, Slough, G.B., 270 p.
- (32) MAYNARD SMITH, J., HAIG, J., 1974. – The hitch-hiking effect of a favourable gene. *Genet. Res.*, **23**, 23-35.
- (33) MOAZAMI-GOUDARZI K., LALOË D., FURET J.P., GROSCLAUDE F., 1997. – Analysis of genetic relationships between 10 cattle breeds with 17 microsatellites. *Anim. Genet.*, **28**, 338-345.
- (34) NEEL J.V., 1973. – “Private” genetic variants and the frequency of mutation among South American Indians. *Proc. Natl. Acad. Sci. USA*, **70**, 3311-3315.
- (35) NEI M., 1973. – Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA*, **70**, 3321-3323.

MESURE ET ÉVOLUTION DE LA DIVERSITÉ GÉNÉTIQUE DES PLANTES CULTIVÉES ET DES ANIMAUX DOMESTIQUES

---

- (36) OLLIVIER L., 1998. – Results and prospects of reproductive technologies for (sustainable) breeding and conservation in farm animal species other than cattle. *Acta Agric. Scand., Suppl.* **29**, 115-120.
- (37) OLLIVIER L., FOULLEY J.-L., 2009. – Managing genetic diversity, fitness and adaptation of farm animal genetic resources. In: van der Werf, J., Graser, H.-U., Frankham, R. and Gondro, C. (Eds) *Adaptation and Fitness in Animal Populations*, Springer, Berlin, pp. 201-227.
- (38) OLLIVIER L., FOULLEY J.-L., 2013. – A note on the partitioning of allelic diversity *Conserv. Genet.* DOI 10.1007/s10592-013-0508-5.
- (39) PETIT R.J., EL MOUSADIK A., PONS O., 1998. – Identifying populations for conservation on the basis of genetic markers. *Conserv. Biol.*, **12**, 844-855.
- (40) PRITCHARD J. K., STEPHENS M., DONNELLY P., 2000. – Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- (41) ROBERTSON A., 1961. – Inbreeding in artificial selection programmes. *Genet. Res.*, **2**, 189-194.
- (42) SLATKIN M., 1985. Rare alleles as indicators of gene flow. *Evolution*, **39**, 53-65.
- (43) THAON D'ARNOLDI C., FOULLEY J.-L., OLLIVIER L., 1998. – An overview of the Weitzman approach to diversity. *Genet. Sel. Evol.*, **30**, 149-161.
- (44) VEECH J.A., SUMMERVILLE K.S., CRIST T.O., GERIN J.C., 2002. – The additive partitioning of species diversity: recent revival of an old idea. *Oikos*, **99**, 3-9.
- (45) VEECH J.A., CRIST T.O., 2010. – Toward a unified view of diversity partitioning. *Ecology*, 1988-1992.
- (46) WEITZMAN M.L., 1992. – On diversity. *Quart. J.Econ.*, **107**, 363-405.
- (47) WEITZMAN M.L., 1993. – What to preserve? An application of diversity theory to crane conservation. *Quart. J. Econ.*, **108**, 157-183.
- (48) WRIGHT S., 1931. – Evolution in Mendelian populations. *Genetics*, **16**, 97-159.
- (49) WRIGHT S., 1943. – Isolation by distance. *Genetics*, **28**, 114-138.