

Revue bibliographique

Méthodes d'analyse des données démographiques et généalogiques dans les populations d'animaux domestiques

Jacqueline VU TIEN KHANG

*I.N.R.A. Station d'Amélioration génétique des Animaux
Centre de Recherches de Toulouse
B.P. 12, F 31320 Castanet-Tolosan*

Résumé

Cette revue bibliographique présente les méthodes classiquement utilisées dans l'analyse des données démographiques et généalogiques se rapportant à des populations d'animaux domestiques.

L'analyse de la population des élevages repose sur l'étude de leur structure par âge, de leur durée de vie et de leur renouvellement. Une méthode permet de décrire la stratification fonctionnelle des élevages selon leurs pratiques d'enregistrement et la destination de leurs reproducteurs. La méthode de « description numérique de la structure d'une population » (ROBERTSON, 1953) évalue la concentration des origines des reproducteurs.

Les entrées et sorties d'animaux sont traitées dans les pyramides des âges et les tables de mortalité. La reproduction de la population est analysée à travers l'étude des carrières reproductives, des tables de fécondité, des taux de renouvellement et de réforme, et des intervalles de génération.

L'analyse génétique repose sur les coefficients de consanguinité et de relation (WRIGHT, 1922). WRIGHT et MC PHEE (1925) en proposent une méthode de calcul approchée.

L'évolution de la consanguinité observée dans une population permet d'estimer son effectif génétique (WRIGHT, 1931). La consanguinité peut être décomposée en trois termes : la consanguinité proche, la consanguinité à long terme et la consanguinité de lignée (ROBERTSON et ASKER, 1951a).

Le coefficient de relation entre un individu et une population comprend deux parts dues respectivement aux relations directes et collatérales : la première fournit une estimation de l'impact génétique de l'individu sur la population. WIENER (1953) introduit une méthode destinée à calculer les contributions des divers élevages au stock génétique d'une population.

Mots clés : structure de population, analyse démographique, analyse génétique, animaux domestiques.

Summary

Methods of analysis of demographic and genealogical data in the populations of domestic animals

This review deals with the most frequently used methods of demographic and genetic analysis of domestic animal populations from their demographic characteristics and pedigree registrations.

Analysis of the herd or flock population is based upon the study of their age structure, their duration and their replacement. A method describes functional stratification of the herds or flocks according to their registration practices and their importance as a source of breeding material. The method

of numerical description of breed structure (ROBERTSON, 1953) characterizes the concentration of the origins of sires.

Entries and removals of animals are taken into account in age distribution and life tables. Reproduction of the population is analysed through surveys of reproductive life, fecundity tables, replacement and wastage rates and generation length.

Genetic analysis is based upon inbreeding and relationship coefficients (WRIGHT, 1922). WRIGHT and Mc PHEE (1925) developed an approximate method of calculating these coefficients.

The increase in observed inbreeding in a population provides an estimate of its effective size (WRIGHT, 1931). Total inbreeding may be divided into three parts : current inbreeding, long term inbreeding and strain inbreeding (ROBERTSON and ASKER, 1951a).

The coefficient of relationship of an individual to a population includes two components respectively due to direct and collateral relationship. The direct relationship is an estimate of the genetic contribution of an individual to the population. WIENER (1953) gives a method for calculating contributions of herds or flocks to the genetic make-up of a breed.

Kew words : breed structure, demographic analysis, genetic analysis, domestic animals.

I. Introduction

L'analyse démographique permet de décrire la structure et la dynamique des populations considérées comme des ensembles renouvelés d'individus. L'analyse génétique s'intéresse à l'évolution du patrimoine génétique des populations. Or l'histoire des gènes est indissociable de celle des individus : c'est pourquoi démographie et génétique de populations sont liées et complémentaires. Ainsi, dans les publications consacrées à l'étude de populations d'animaux domestiques, les aspects démographiques et génétiques sont souvent étroitement imbriqués. Cependant, dans un souci de clarté, nous traiterons séparément les méthodes relevant de chacune des deux disciplines.

L'objet de cette revue bibliographique est de présenter quelques méthodes d'analyse démographique et génétique classiquement employées dans l'étude des populations d'animaux domestiques. Nous restreindrons le champ de cette présentation aux seules méthodes d'analyse descriptive de populations réelles, fondées sur l'exploitation de données démographiques et généalogiques. Les références bibliographiques citées constituent des exemples destinés à illustrer l'exposé des diverses méthodes.

II. Méthodes d'analyse démographique

L'objet de l'analyse démographique est de décrire l'état d'une population (c'est-à-dire son effectif, sa composition selon divers critères : sexe, âge, localisation géographique...) et les divers phénomènes (tels que la reproduction, la mortalité, les éliminations, les mouvements migratoires) qui déterminent son évolution dans le temps. L'existence d'une population est régie par le processus suivant :

- des individus naissent ou immigrent, entrant ainsi dans la population ;
- des individus meurent, émigrent ou sont réformés, sortant ainsi de la population ;
- un laps de temps s'écoule entre l'entrée et la sortie de chaque individu : cette durée de vie est jalonnée par la procréation d'un certain nombre de descendants qui contribuent au renouvellement de la population.

C'est à l'étude des populations considérées comme des ensembles renouvelés que s'attache l'analyse démographique. Celle-ci est fondée sur la collecte et le traitement de données statistiques provenant de registres, de recensements ou d'enquêtes spécifiques. Elle peut déboucher sur une recherche causale et, éventuellement, sur une prévision (HICKEY, 1960 ; TURNER et DOLLING, 1965 ; TACHER, 1975 a et b).

L'analyse démographique est dominée par la distinction fondamentale existant entre la vision transversale et la vision longitudinale. L'analyse transversale se rapporte aux manifestations d'un phénomène démographique durant une période donnée. Le qualificatif de « transversal » s'applique également à toute coupe pratiquée pour saisir, à un instant donné, l'état d'une population. L'analyse longitudinale se rapporte aux manifestations d'un phénomène dans une cohorte, c'est-à-dire dans un ensemble d'individus ayant vécu un même événement démographique (généralement la naissance ou l'entrée dans la vie reproductive) durant une période déterminée (une année civile, par exemple). L'analyse longitudinale des diverses cohortes impliquées dans une analyse transversale en constitue un complément essentiel. L'analyse longitudinale, qui repose sur une vision rétrospective, nécessite de reconstituer l'histoire démographique des individus de la cohorte considérée. Cette difficulté explique que l'approche transversale soit beaucoup plus fréquemment pratiquée que l'approche longitudinale.

Le découpage du temps est un problème inhérent à toute étude démographique. Le choix de l'unité de temps requiert une connaissance préalable des ordres de grandeur de certains paramètres relatifs à la durée de vie, au rythme de reproduction et au mouvement saisonnier des principaux phénomènes démographiques. Dans la plupart des définitions et des exemples traités, nous avons arbitrairement adopté l'année civile comme unité de temps. Mais elle pourrait être remplacée par toute autre unité, de durée inférieure ou supérieure.

Les méthodes de l'analyse démographique ont été initialement développées afin d'étudier des populations humaines (PRESSAT, 1973). Cependant, elles peuvent être utilement adaptées au cas des populations animales (BETTINI, 1968). Bien décrire la structure et le fonctionnement d'une population d'animaux domestiques constitue en effet une étape préalable à la gestion de son patrimoine génétique. L'analyse démographique fournit des éléments utiles à la gestion d'une petite population dont on cherche à minimiser la consanguinité et dans laquelle l'aménagement du plan d'accouplement doit être établi en fonction de certains paramètres relatifs à la durée de vie et à la reproduction (ROCHAMBEAU *et al.*, 1979). D'autre part, l'organisation de la sélection ne peut être raisonnée indépendamment des caractéristiques démographiques de la population considérée. Ainsi, l'estimation de la valeur génétique des reproducteurs testés sur descendance peut être influencée par la répartition de leurs produits dans les divers élevages. La réponse à la sélection par unité de temps dépend notamment de l'intervalle de génération et des capacités de reproduction de la population. L'évolution et la diffusion du progrès génétique dans une population est conditionnée par sa structure démographique. Ces quelques exemples illustrent l'intérêt que présente l'analyse démographique pour le généticien.

Dans les publications consacrées à l'analyse de populations domestiques, les aspects ayant trait à la démographie sont très divers. Aussi organiserons-nous la présentation des principaux concepts et méthodes autour de quelques points-clés tels que la collecte des données, la subdivision de la population en élevages, les entrées et sorties d'animaux, le renouvellement de la population. Il n'est pas question de faire ici un inventaire de l'ensemble des méthodes disponibles : cette présentation se limitera aux méthodes habituellement utilisées dans la description démographique des populations domestiques.

A. Collecte des données

La collecte des données est grandement facilitée lorsqu'existent des registres où sont consignés les principaux événements démographiques concernant la vie de chaque animal (tels que sa naissance, sa généalogie, les événements de sa carrière reproductive, sa mort ou sa réforme). Cependant, une critique des données est toujours utile. Les auteurs se posent en général deux sortes de questions.

1. Quelle est la fraction de la population totale appréhendée grâce à ces enregistrements ?

Par exemple, dans certains fichiers, seuls sont enregistrés les animaux dits « de renouvellement », c'est-à-dire les animaux conservés en vue de la reproduction : il est évident qu'une telle sélection de l'information a une incidence sur l'interprétation des résultats.

2. Quelle est la fiabilité des données ?

La fiabilité des données peut être vérifiée par des recoupements entre diverses sources d'information. Ainsi, les données généalogiques peuvent être confrontées aux données de polymorphisme sanguin : la proportion d'exclusions de parenté, interprétée en fonction du degré de polymorphisme de la population, permet d'évaluer le taux d'erreurs dans les enregistrements généalogiques (COTTERMAN, 1969 ; LANGANEY et NADOT, 1973 ; CHAKRABORTY *et al.*, 1974).

En l'absence d'enregistrements écrits, il est nécessaire de recourir à des enquêtes directes fondées sur des méthodes spécifiques (POIVEY *et al.*, 1981). Ces aspects ne seront pas développés, car les études auxquelles nous faisons référence se rapportent, pour la plupart, à des populations munies d'un « état civil ».

B. Description macroscopique de la population ; analyse de la population des élevages

Toute description démographique commence par situer la population étudiée dans l'espace et dans le temps : il s'agit de rassembler des données globales sur sa répartition géographique, sur l'évolution de ses effectifs au cours du temps et sur l'impact numérique d'éventuelles immigrations. Cette description macroscopique ne requérant aucune méthode spécifique, il n'est pas utile de la développer.

Une population d'animaux domestiques est généralement subdivisée en élevages. L'élevage représente une unité fondamentale du point de vue géographique, écologique et économique. Une analyse prenant en compte les élevages est indispensable à la compréhension de la structure et de la dynamique d'une population. Un premier niveau d'analyse consiste à envisager la population comme un ensemble d'élevages et à considérer les élevages comme des entités démographiques qui apparaissent, vivent un certain temps, puis disparaissent. Dans cette perspective, certaines méthodes classiques de l'analyse démographique, appliquées d'ordinaire à des ensembles d'individus, peuvent l'être à des ensembles d'élevages. C'est ainsi que les auteurs, après avoir étudié la distribution géographique des élevages et l'évolution de leur nombre au cours du temps, analysent leur durée de vie et leur renouvellement. Pour traiter ces aspects, ils ont recours à deux méthodes :

— la première consiste à établir la structure de la population des élevages en fonction de leur âge, l'âge d'un élevage étant défini comme le laps de temps écoulé depuis la date de son apparition (DONALD et AUERBACH, 1942 ; DONALD et EL ITRIBY, 1946 ; WIENER,

1954, 1961 ; YOUNG et PURSER, 1962 ; ÖZKÜTÜK et BICHARD, 1977 ; FOGARTY, 1978 a). Cette méthode relève d'une vision transversale ;

— la seconde consiste à étudier la distribution de la durée de vie au sein de diverses cohortes d'élevages, une cohorte étant ici définie comme l'ensemble des élevages apparus au cours d'une année (ou d'une campagne) donnée (DONALD et AUERBACH, 1942 ; DONALD et EL ITRIBY, 1946 ; STEWART, 1952 b ; WIENER, 1954, 1961 ; YOUNG et PURSER, 1962 ; MARTIN, 1975 a ; HERRON et PATTIE, 1977 a). Cette méthode, complémentaire de la précédente, relève de l'analyse longitudinale.

Le renouvellement des élevages est habituellement caractérisé par la proportion d'élevages qui apparaissent pendant une année donnée, paramètre équivalant à un taux brut de natalité, et par la proportion d'élevages qui disparaissent pendant la même période, paramètre équivalant à un taux brut de mortalité (DONALD et AUERBACH, 1942 ; DONALD et EL ITRIBY, 1946 ; YOUNG et PURSER, 1962 ; C.P. Mc PHEE, 1965 ; HERRON et PATTIE, 1977 a).

C. Les élevages considérés en tant qu'ensembles d'individus

Un élevage se caractérise par son effectif et sa composition en diverses catégories d'animaux (DONALD et AUERBACH, 1942 ; DONALD, 1945 ; DONALD et EL ITRIBY, 1945 ; STEWART, 1952 b ; WIENER et YAO, 1952 ; WIENER, 1953 ; BARKER, 1957 a ; YOUNG et PURSER, 1962 ; DAVEY et BARKER, 1963 ; MATASSINO *et al.*, 1965 ; LAUVERGNE *et al.*, 1973 ; DRAGANESCU, 1979 ; SCHOEMAN, 1979). Les auteurs recherchent classiquement les relations éventuelles entre la taille et la durée de vie des élevages (WIENER, 1954 ; HERRON et PATTIE, 1977 a ; FOGARTY, 1978 a).

Un élevage n'est pas un isolat : il échange généralement des reproducteurs avec d'autres élevages. Au cours d'une campagne de reproduction, un même mâle peut procréer des descendants dans plusieurs élevages. Ce phénomène est amplifié par l'insémination artificielle, qui augmente la capacité de procréation des mâles et permet de dissocier leur localisation de celle de leurs produits. L'intensité de la circulation des reproducteurs entre élevages est appréciée par le pourcentage de reproducteurs utilisés, à un moment donné, dans d'autres élevages que leur élevage d'origine (WIENER 1955 ; BARKER et DAVEY, 1960 ; WIENER, 1961 ; DAVEY et BARKER, 1963 ; LAUVERGNE *et al.*, 1973 ; HERRON et PATTIE, 1977 a ; ÖZKÜTÜK et BICHARD, 1977).

La circulation des reproducteurs est caractérisée par une matrice dont l'élément n_{ij} désigne le nombre de produits nés dans l'élevage i (ou dans l'ensemble géographique ou fonctionnel d'élevages i), d'un parent (généralement d'un père) lui-même issu de l'élevage j (ou de l'ensemble d'élevages j) : WIENER, 1955 ; C.P. Mc PHEE, 1965 ; LAUVERGNE *et al.*, 1973 ; ÖZKÜTÜK et BICHARD, 1977 ; FOGARTY, 1978 a. A partir d'un tel tableau, on peut éprouver par un test du chi 2 l'hypothèse d'indépendance entre l'origine des pères et celle de leurs produits, ou calculer le coefficient de corrélation entre ces origines (C.P. Mc PHEE, 1965). Dans certaines populations, l'analyse de la circulation des reproducteurs permet de mettre en évidence une stratification des élevages.

1. Stratification des élevages

Dans de nombreuses populations d'animaux domestiques, il existe une stratification fonctionnelle des élevages fondée sur la reconnaissance, par les éleveurs, de la supériorité (réelle ou imaginée) de certains troupeaux (LUSH, 1946 ; ROBERTSON et ASKER, 1951 a). La

place de chaque élevage dans la hiérarchie pyramidale dépend de son renom. Plus un élevage est réputé, plus les reproducteurs qu'il produit sont demandés. STEWART (1952 b, 1955) classe les élevages en plusieurs catégories, selon leurs pratiques d'enregistrement des animaux et la destination des mâles qu'ils produisent (fig. 1). BARKER (1957 b) assortit cette classification d'une nomenclature qui sera reprise par de nombreux auteurs (BARKER et DAVEY, 1960 ; DAVEY et BARKER, 1963 ; C.P. Mc PHEE, 1965 ; MARTIN, 1975 b ; HERRON et PATTIE, 1977 a ; FOGARTY, 1978 b ; HERRON, 1978 ; SCHOEMAN, 1979).

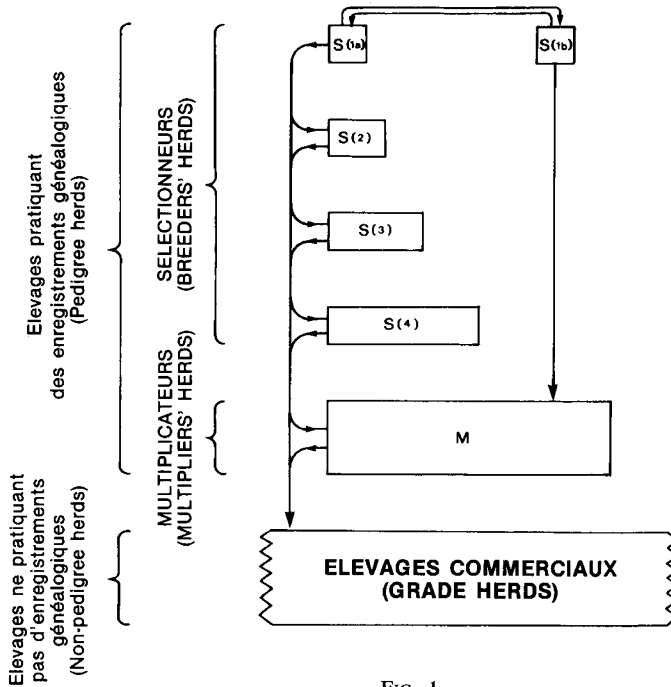


FIG. 1

Représentation schématique de la structure pyramidale d'une population.
Diagrammatic representation of hierarchical breed structure.

Les flèches symbolisent les migrations de reproducteurs.
Les notations sont définies dans le paragraphe II, C, 1.

Arrows indicate movement of sires.
Notations are defined in paragraph II, C, 1.

Les élevages se distribuent en deux grands ensembles selon qu'ils pratiquent ou non des enregistrements généalogiques. Les élevages ne pratiquant pas d'enregistrements généalogiques sont appelés « élevages commerciaux ».

Les élevages pratiquant des enregistrements généalogiques comprennent :

- les élevages qui produisent des mâles utilisés dans d'autres élevages pratiquant eux-mêmes des enregistrements généalogiques : ce sont les sélectionneurs (notés S) ;
- les autres, dont les mâles peuvent être utilisés soit dans leur élevage d'origine, soit dans des troupeaux commerciaux : ce sont les multiplicateurs (notés M).

Les sélectionneurs se répartissent en plusieurs strates, suivant leur importance en tant que source de matériel génétique. On peut caractériser chaque strate par les propriétés suivantes : les élevages d'une strate donnée n'utilisent que des mâles issus de cette même strate ou de strates supérieures ; inversement, ils ne fournissent jamais de mâles aux strates supérieures. On est conduit à subdiviser la strate supérieure S (1) en deux classes :

— S (1a), qui comprend les élevages fournissant des reproducteurs aux autres strates de sélectionneurs ;

— S (1b), qui comprend les élevages dont les mâles ne sont jamais utilisés dans les autres strates de sélectionneurs, mais uniquement dans S (1) et, éventuellement, chez des multiplicateurs (fig. 1).

WIENER (1953) propose un autre critère de subdivision des sélectionneurs. Il distingue :

— ceux dont les mâles procréent des fils de renouvellement chez d'autres sélectionneurs ;

— ceux dont les mâles ne procréent pas de fils, mais seulement des filles de renouvellement chez les autres sélectionneurs.

Les multiplicateurs (M) peuvent être classés en quatre catégories, en fonction de leurs pratiques d'enregistrement et, secondairement, de la destination des mâles qu'ils produisent :

— M (H) comprend les troupeaux enregistrant des mâles qui ne sont utilisés que dans leur troupeau d'origine ou, éventuellement, dans des élevages commerciaux ;

— M (N) comprend les troupeaux enregistrant des mâles qui ne sont jamais utilisés dans un élevage pratiquant des enregistrements généalogiques ;

— M (S) comprend les troupeaux qui n'enregistrent pas de mâles, mais uniquement des femelles ;

— M (M) comprend les troupeaux qui n'enregistrent que des mâles, et jamais de femelles.

Les mâles utilisés chez les multiplicateurs proviennent des sélectionneurs ou, exceptionnellement, de l'élevage multiplicateur lui-même si celui-ci appartient à la catégorie M (H). Les mâles issus des multiplicateurs sont utilisés dans les troupeaux commerciaux, exception faite de certains mâles de la catégorie M (H) utilisés dans leur élevage d'origine.

Un autre type de structure des élevages, principalement développé depuis 1967 en Nouvelle-Zélande et en Australie, est celui correspondant à la sélection de groupe avec noyau ouvert* (EVANS, 1976 ; RAE, 1976 ; SHEPHERD, 1976). Dans cette structure, les élevages se répartissent en deux catégories : le noyau et la base. Le noyau de sélection fournit les reproducteurs mâles à l'ensemble de la population (noyau et base). Les femelles de renouvellement sont produites à la fois par le noyau et par la base. Elles peuvent migrer du noyau vers la base, mais aussi de la base vers le noyau. De ce fait, les gènes circulent dans les deux sens entre le noyau et la base. Il existe donc une « remontée » des gènes de la base vers le noyau : c'est ce qui distingue fondamentalement une structure à noyau ouvert d'une structure hiérarchique traditionnelle.

* En anglais : « group breeding sheme », ou « co-operative breeding scheme », ou « open nucleus breeding scheme ».

Plusieurs auteurs ont étudié, d'un point de vue théorique, l'évolution et la diffusion du progrès génétique dans une population possédant une structure hiérarchique traditionnelle (BICHARD, 1971 ; ELSEN et MOCQUOT, 1974 ; HILL, 1974 ; ELSEN, 1980) ou une structure à noyau ouvert (JACKSON et TURNER, 1972 ; JAMES, 1976, 1977).

La situation génétique d'une population dépend souvent de la politique d'accouplements pratiquée par un assez petit nombre d'éleveurs qui dominent la population. Dans de nombreuses études descriptives, cette situation est analysée en termes qualitatifs plutôt que quantitatifs. C'est pourquoi ROBERTSON (1953) propose une description numérique de la domination d'une population par un petit nombre d'élevages.

2. Description numérique de la structure d'une population

ROBERTSON (1953) remarque que la concentration des origines des reproducteurs sur un petit nombre d'élevages apparaît clairement lorsqu'on examine les lignées d'ascendance paternelle de pedigrees pris au hasard : le nombre d'élevages rencontrés se réduit au fur et à mesure qu'on remonte ces lignées. Le problème est de définir un paramètre caractérisant cette concentration des origines des pères, grands-pères paternels, etc... d'un groupe d'animaux. ROBERTSON définit le paramètre C_s (s comme « sire ») comme la probabilité pour que deux animaux pris au hasard dans l'échantillon étudié aient des pères issus du même élevage. Soient N le nombre d'élevages ayant fourni les pères de l'échantillon étudié

et n_i le nombre d'animaux dont les pères sont issus de l'élevage i . Parmi les $\sum_{i=1}^N n_i$ animaux de l'échantillon, le nombre de paires distinctes est égal à :

$$\frac{\left(\sum_{i=1}^N n_i \right) \left(\sum_{i=1}^N n_i - 1 \right)}{2}$$

Dans cet ensemble de paires distinctes, celles qui comportent des animaux ayant des pères issus du même élevage i sont au nombre de :

$$\frac{n_i (n_i - 1)}{2}$$

La probabilité pour que deux animaux pris au hasard aient des pères issus du même élevage est donc :

$$C_s = \frac{\sum_{i=1}^N \left(\frac{n_i (n_i - 1)}{2} \right)}{\frac{\left(\sum_{i=1}^N n_i \right) \left(\sum_{i=1}^N n_i - 1 \right)}{2}} = \frac{\sum_{i=1}^N \left(n_i (n_i - 1) \right)}{\left(\sum_{i=1}^N n_i \right) \left(\sum_{i=1}^N n_i - 1 \right)}$$

De même, on peut définir C_{ss} (respectivement C_{sss}) comme la probabilité pour que deux individus pris au hasard aient des grands-pères paternels (respectivement des arrière-grands-pères paternels) issus du même élevage. L'indice (s , ss , sss , etc.) caractérise la génération considérée dans la lignée d'ascendance paternelle.

YULE (1944) avait rencontré un problème analogue lors du traitement statistique d'un style littéraire : quelques mots apparaissent très fréquemment alors que de nombreux autres n'apparaissent qu'une fois. Il a défini une « caractéristique » de leur distribution tout à fait similaire au paramètre C défini par ROBERTSON.

ROBERTSON s'intéresse à l'inverse de C_s , noté H_s , qu'il nomme « nombre efficace de troupeaux fournissant des pères », par analogie avec l'effectif génétique (« effective number ») défini par WRIGHT (1931). Le concept de nombre efficace appliqué ici aux élevages est l'équivalent formel de celui concernant les individus. H_s peut se définir de la façon suivante : la probabilité pour que deux animaux pris au hasard dans l'échantillon étudié aient des pères issus du même élevage est la même que dans un échantillon idéal de référence où les pères des individus seraient répartis de façon équiprobable dans H_s élevages.

L'estimation de H présente des inconvénients sur le plan statistique : l'inverse de l'estimateur de C n'est pas un estimateur centré de H . Lorsqu'on désire savoir s'il existe des différences statistiquement significatives entre populations, il est préférable de travailler sur C plutôt que sur H .

La méthode de ROBERTSON, nommée « description numérique de la structure d'une population » a été fréquemment appliquée (BARKER, 1957 b ; BARKER et DAVEY, 1960 ; YOUNG et PURSER, 1962 ; DAVEY et BARKER, 1963 ; C.P. Mc PHEE, 1965 ; MARTIN, 1975 b ; HERRON et PATTIE, 1977 a ; ÖZKÜTÜK et BICHARD, 1977 ; FOGARTY, 1978 b ; SCHOEMAN, 1979 ; TOLL et BARKER, 1979).

D. Traitement des données relatives aux entrées et sorties d'animaux

1. Une vision transversale : la pyramide des âges

La pyramide des âges est un double histogramme qui fournit une représentation instantanée de la structure de la population par sexe et par âge. Cette représentation est classiquement utilisée (CANNON et HANSEN, 1939 ; PETTIT, 1940 ; DONALD et EL ITRIBY, 1947 ; HICKEY, 1960 ; BETTINI *et al.*, 1962 ; MATASSINO *et al.*, 1965 ; JOAKIMSEN, 1969 ; GIDEL, 1972 ; LAUVERGNE *et al.*, 1973 ; BRZOZOWSKI, 1974 ; KNOTHE, 1976 ; LANGLOIS, 1976 ; VERA *et al.*, 1978).

2. Une vision longitudinale : la table de mortalité

Une table de mortalité (ou de réforme) décrit la survenue, en fonction de l'âge, des sorties (morts, réformes, ventes...) dans une cohorte d'animaux. A partir d'un nombre arbitraire d'individus présents à l'âge zéro (ce nombre, appelé « racine de la table », étant généralement une puissance de 10), la table fournit, pour la suite des anniversaires x :

— le nombre de survivants S_x à chaque anniversaire x (la représentation graphique de S_x en fonction de x étant la courbe de survie) ;

* Le terme de « mortalité » est pris ici dans un sens large : par mortalité, nous entendons non seulement la mortalité au sens strict, mais aussi l'ensemble des phénomènes entraînant des sorties d'animaux (réforme, vente).

— le nombre de sorties $d(x, x+1)$ entre deux anniversaires successifs x et $x+1$:

$$d(x, x+1) = S_x - S_{x+1}$$

— le quotient de mortalité (ou de réforme) q_x à chaque âge x , égal à la probabilité, pour un individu présent dans la population à l'anniversaire x , de sortir avant l'anniversaire suivant $x+1$:

$$q_x = \frac{d(x, x+1)}{S_x} = \frac{S_x - S_{x+1}}{S_x}$$

A partir d'une telle table de mortalité, on peut calculer l'espérance de vie e_x à l'âge x . En admettant que les individus disparus entre les anniversaires x et $x+1$ sortent de la population en moyenne à l'âge $(x+0,5)$, il est aisé de calculer e_x :

$$e_x = \frac{1}{2} + \frac{S_{x+1} + S_{x+2} + \dots}{S_x}$$

La construction d'une table complète de mortalité nécessite de suivre une cohorte depuis son apparition jusqu'à son extinction totale (COMFORT, 1958, 1959 et 1962). La difficulté d'une telle entreprise explique que de nombreux auteurs aient recours à des synthèses transversales : sans l'annoncer explicitement, ils établissent en fait une table de mortalité du moment, c'est-à-dire une table résultant de l'attribution à une cohorte fictive des quotients de mortalité calculés sur les diverses cohortes réelles soumises au risque de mortalité pendant un laps de temps donné (TURNER *et al.*, 1959 ; HICKEY, 1960 ; MATASSINO *et al.*, 1965 ; KNOTHE, 1976). Pour construire une telle table, ils font parcourir tous les âges de la vie à une cohorte fictive, en lui faisant subir à chaque âge le risque de mortalité calculé pendant la période étudiée : autrement dit, ils reportent sur une cohorte fictive une succession de risques qui ont tous été observés sur des cohortes différentes au cours d'une période fixée. La table de mortalité du moment permet de calculer les espérances de vie inhérentes aux conditions de l'époque considérée (HICKEY, 1960).

D'autres auteurs assimilent la structure instantanée par âge d'une population à une courbe de survie, et l'âge moyen de la population à l'espérance de vie (BUCHANAN SMITH et ROBISON, 1931 ; CANNON et HANSEN, 1939 ; PETTIT 1940 ; DONALD et EL ITRIBY, 1947 ; ANDRUS *et al.*, 1970). Cette pratique repose sur l'hypothèse implicite de stationnarité de la population considérée (une population stationnaire étant une population dans laquelle la table de mortalité est identique dans toutes les cohortes et le nombre annuel d'entrées, constant).

E. Étude de la reproduction de la population

Une population résulte de la coexistence de diverses cohortes. Cet ensemble subit un processus permanent de renouvellement : chaque année, pendant que les cohortes déjà existantes sont réduites par la mortalité, une nouvelle cohorte apparaît, issue d'animaux reproducteurs appartenant à des cohortes antérieures. Des paramètres analytiques ont été définis afin de caractériser les diverses composantes biologiques de l'aptitude d'un animal à créer une descendance viable (TURNER et DOLLING, 1965 ; DESVIGNES, 1968 ; LEGAULT, 1978). D'un point de vue démographique, il est essentiel, pour bien décrire le processus global de renouvellement d'une population, de reconstituer la carrière reproductive de chaque cohorte.

1. Etude de la carrière reproductive d'une cohorte d'animaux

Ce type d'étude est le plus souvent réalisé sur des femelles, mais rien n'empêche d'utiliser les mêmes méthodes sur des mâles (LANGLOIS, 1976). Cependant, on peut se heurter à une légère difficulté dans le cas particulier d'inséminations artificielles réalisées avec du sperme congelé : un tel procédé dissocie l'aptitude d'un mâle à procréer de son état de vie ou de mort. On peut lever cette difficulté en définissant la carrière reproductive d'un mâle comme le laps de temps pendant lequel il procréé des descendants, sa carrière se prolongeant éventuellement au-delà de sa mort ou de sa réforme.

L'analyse de la carrière d'un groupe d'animaux reproducteurs se situe nécessairement dans une perspective longitudinale : il s'agit de suivre la carrière reproductive d'une cohorte d'animaux, c'est-à-dire d'un ensemble d'animaux nés (ou entrés dans la vie reproductive) pendant une période donnée. Pour chaque animal de la cohorte considérée, on enregistre les naissances de ses produits. Selon les cas, on prend en compte la totalité des produits nés ou la seule fraction des descendants « utiles » (au sens génétique du terme), c'est-à-dire des descendants qui deviendront à leur tour des reproducteurs.

A partir d'un tel ensemble de données, on peut étudier la distribution de l'âge au début de la carrière reproductive (ROBERTSON et ASKER, 1951b ; ANDRUS *et al.*, 1970 ; MATASSINO *et al.*, 1970 ; BRZOWSKI, 1974 ; NARDONE *et al.*, 1975 ; LANGLOIS, 1976 ; LEGAULT, 1978), de l'âge des parents à la naissance des produits de rangs successifs (ANDRUS *et al.*, 1970), des intervalles entre mises bas (ROBERTSON et ASKER, 1951b ; FISTEAG *et al.*, 1968 ; MATASSINO *et al.*, 1970 ; ANDRUS *et al.*, 1970 ; LANDAIS *et al.*, 1980), de la durée de la carrière reproductive (ROBERTSON et ASKER, 1951b ; MATASSINO *et al.*, 1970 ; LANGLOIS, 1976), du nombre total de mises bas par mère (ROBERTSON et ASKER, 1951b ; MATASSINO *et al.*, 1970 ; NARDONE 1970 ; NARDONE *et al.*, 1975), du nombre total de produits par géniteur (ROBERTSON et ASKER, 1951 b ; FISTEAG *et al.*, 1968 ; MATASSINO *et al.*, 1970 ; LANGLOIS, 1976). En pratique, ces calculs ne sont généralement pas réalisés sur des cohortes, mais sur des ensembles composites d'animaux appartenant à diverses cohortes : les auteurs supposent implicitement que les paramètres démographiques caractérisant la carrière reproductive ne varient pas sensiblement au fil des cohortes successives.

La survenue des naissances dans une cohorte d'animaux (mâles ou femelles) peut être décrite par ce qu'on appelle classiquement en démographie humaine une table brute de fécondité. Sur la base d'un nombre arbitraire d'animaux (généralement une puissance de 10) non soumis à la mortalité, une telle table donne, pour la suite des anniversaires x :

- le nombre de naissances $n(x, x+1)$ entre deux anniversaires successifs x et $x+1$;
- la descendance atteinte à chaque anniversaire x :

$$D_x = \sum_{\xi=1}^{x-1} n(\xi, \xi+1)$$

La descendance finale, réduite aux individus du même sexe que les géniteurs considérés, correspond au taux brut de reproduction. Le nombre de naissances $n(x, x+1)$ s'identifie au taux de fécondité à l'âge x (également appelé « naissances réduites ») égal, par définition, au rapport du nombre de naissances survenues chez les animaux de la classe d'âge x , à l'effectif moyen des animaux ayant cet âge.

La construction d'une table brute de fécondité exige un recul assez important dans le temps, puisqu'il faut retracer l'histoire complète de la cohorte considérée. C'est pourquoi,

ici encore, les auteurs ont recours à des synthèses transversales : ils établissent des tables de fécondité du moment en juxtaposant dans une cohorte fictive les taux de fécondité (au sens précédemment défini) observés sur les diverses cohortes en âge de procréer pendant une période fixée (HICKEY, 1960 ; TURNER et DOLLING, 1965). LANGLOIS (1976) calcule des taux de fécondité par âge sur des ensembles d'animaux morts ou réformés pendant une période déterminée : il ne s'agit pas d'une vraie table de fécondité puisque les animaux considérés appartiennent à différentes cohortes ; il ne s'agit pas non plus d'une table du moment au sens strict, puisque la période d'observation est relativement longue.

2. Taux de renouvellement ; taux de réforme

Ces notions s'appliquent à la population des reproducteurs. Son renouvellement est assuré par l'entrée de nouveaux animaux en reproduction et par le départ concomitant d'animaux plus anciens. Le taux de renouvellement est défini comme le rapport du nombre d'entrées pendant une unité de temps (année civile, campagne) à l'effectif moyen de la population pendant cette même période. En pratique, le critère d'entrée d'un animal peut varier selon les situations et les auteurs : par exemple, pour une femelle, ce critère peut être la mise à la reproduction ou la première mise bas. Le taux de réforme est défini symétriquement comme le rapport du nombre de sorties pendant une unité de temps à l'effectif moyen de la population pendant cette même période (l'écart entre taux de renouvellement et taux de réforme représentant le taux d'accroissement de la population). Le taux de réforme défini ici est distinct du quotient de réforme introduit dans le paragraphe II, D, 2. En effet, un taux est le rapport du nombre d'événements (renouvelables, comme les mises bas, ou non renouvelables, comme les réformes) survenus dans une population durant une période donnée, à l'effectif moyen de la population durant cette même période. Un quotient est la probabilité de survenue d'un événement non renouvelable dans une cohorte, entre deux anniversaires fixés : la notion de quotient est liée à celle de table à extinction dont elle constitue un élément (PRESSAT, 1979).

Les taux de renouvellement et de réforme sont classiquement utilisés en zootechnie (BUCHANAN et ROBISON, 1931 ; PETTIT, 1940 ; BARKER, 1957a ; BETTINI *et al.*, 1962 ; LAUVERGNE *et al.*, 1973 ; KEANE, 1977 ; KOTZE *et al.*, 1977 ; MARKUSHIN, 1978). Plusieurs auteurs assimilent l'inverse du taux de renouvellement à la durée moyenne de la carrière reproductive, faisant implicitement l'hypothèse qu'ils s'adressent à une population stationnaire (BUCHANAN et ROBISON, 1931 ; PETTIT, 1940).

3. Intervalle de génération

• Définitions de l'intervalle de génération

Dans la plupart des populations d'intérêt zootechnique, les générations ne se succèdent pas de façon séparée dans le temps : elles sont dites « chevauchantes » ou « superposées » (BETTINI et MATASSINO, 1961 ; BETTINI, 1968). Dans ces conditions, l'intervalle de génération est habituellement défini comme l'âge moyen des parents à la naissance de leurs produits. Cette définition, très générale, peut être précisée de diverses manières.

Une première distinction peut être opérée selon le sexe des parents et des produits ; ainsi, on peut calculer l'intervalle de génération sur les quatre voies suivantes : père-fils, père-fille, mère-fils et mère-fille.

Une seconde distinction peut être opérée suivant que l'on considère l'ensemble des produits nés ou les seuls produits « utiles » (c'est-à-dire conservés pour le renouvellement) : les

résultats peuvent différer sensiblement selon le critère adopté. Définir l'intervalle de génération comme l'âge moyen des parents à la naissance de leurs produits « utiles » revient à le définir comme le laps de temps moyen s'écoulant entre la reproduction des parents et celle de leurs produits.

Une troisième distinction est inhérente au choix de la cohorte de référence. En effet, il existe deux modes de calcul de l'intervalle de génération. Le premier consiste à calculer l'âge moyen d'une cohorte de parents à la naissance de ses produits. Le second consiste à calculer l'âge moyen des parents d'une cohorte de produits : c'est cette dernière méthode qui, pour des raisons de commodité, est la plus fréquemment employée. Dans une population stationnaire, ces deux modes de calcul sont équivalents ; l'intervalle de génération est uniquement fonction de la « table de mortalité » (cf. § II, D, 2) et de la « table brute de fécondité » (cf. § II, E, 1).

• Méthodes d'estimation de l'intervalle de génération

Une méthode très utilisée dans les publications les plus anciennes (DICKSON et LUSH, 1933 ; LUSH *et al.*, 1936 ; YODER et LUSH, 1937 ; WILLHAM, 1937 ; LUSH et ANDERSON, 1939 ; CARTER, 1940 ; STONAKER, 1943 ; FLETCHER, 1945 ; KUTSAL, 1954 ; YOUNG et PURSER, 1962) consiste à tracer, à partir d'un échantillon représentatif d'une population, des lignées aléatoires d'ascendance, c'est-à-dire des lignées dans lesquelles, à chaque génération, le choix entre le côté paternel et le côté maternel se fait au hasard et de façon équiprobable. Ces lignées remontent jusqu'aux fondateurs, c'est-à-dire jusqu'aux animaux d'ascendance inconnue. Les auteurs estiment l'intervalle de génération en calculant la moyenne, sur toutes les lignées considérées, de l'écart entre les dates de naissance des fondateurs et des animaux de l'échantillon, divisé par le nombre de générations les séparant. On peut faire plusieurs remarques à propos de cette méthode d'estimation. Le paramètre ainsi estimé est en fait l'intervalle de génération moyen observé sur l'ensemble des arbres généalogiques des individus de l'échantillon étudié. Il ne tient pas compte du sexe des individus. Il ne concerne que des individus « utiles » puisque, du fait de la méthode utilisée, tous les individus considérés sont des géniteurs. En outre, cette méthode repose sur deux hypothèses implicites. La première est celle de la constance de l'intervalle de génération au fil des générations successives. La seconde est que les individus apparaissant dans les lignées à un rang d'ascendance donné sont représentatifs de l'ensemble de leurs contemporains « utiles » : cette seconde hypothèse est rarement vérifiée dans le cas de populations soumises à une sélection.

La deuxième méthode d'estimation de l'intervalle de génération, largement utilisée depuis les années cinquante (LERNER et HAZEL, 1947 ; ROBERTSON et ASKER, 1951a ; STEWART, 1952a ; BARKER, 1957b ; BARKER et DAVEY, 1960 ; YOUNG et PURSER, 1962 ; DAVEY et BARKER, 1963 ; JOAKIMSEN, 1969 ; LAUVERGNE *et al.*, 1973 ; MARTIN, 1975a ; LANGLOIS, 1976 ; HERRON et PATTIE, 1977 b ; ÖZKÜTÜK et BICHARD, 1977 ; KOTZE *et al.*, 1977 ; HERRON, 1978 ; TOLL et BARKER, 1979) consiste à calculer l'âge moyen des parents d'un échantillon de produits nés pendant une période déterminée (année civile, campagne). Les auteurs distinguent quatre voies : père-fils, père-fille, mère-fils, mère-fille. Ils calculent l'intervalle moyen de génération en faisant la moyenne arithmétique des intervalles calculés sur ces quatre voies. Un tel intervalle moyen ne constitue une estimation de l'âge moyen des parents à la naissance de leurs produits que dans les cas où les deux sexes sont également représentés parmi les produits. Cette condition n'est généralement pas vérifiée lorsqu'on s'intéresse aux seuls animaux de renouvellement. Néanmoins, l'intervalle moyen ainsi estimé à partir des produits utiles présente un intérêt du point de vue génétique car il correspond au paramètre figurant dans l'espérance mathématique du progrès génétique réalisé par unité de temps.

III. Méthodes d'analyse génétique

WRIGHT (1923a) propose des concepts et des méthodes destinés à analyser l'histoire génétique d'une population à partir de données généalogiques. Dans les années vingt, WRIGHT et Mc PHEE publient une série d'articles consacrés à l'analyse mendélienne de plusieurs rameaux de la race bovine Shorthorn, dans lesquels ils mettent en œuvre ces concepts et ces méthodes (WRIGHT, 1923b ; McPHEE et WRIGHT, 1925 ; Mc PHEE et WRIGHT, 1926). A leur suite, de nombreux auteurs ont recouru à ces méthodes, pour retracer l'histoire des gènes à partir de celle des individus et décrire la structure génétique des populations étudiées. La lecture des publications relatives à ces questions, classées en ordre chronologique de 1923 à nos jours, suggère plusieurs remarques.

Les publications les plus anciennes présentent encore, pour le lecteur contemporain, un intérêt autre que purement historique. Ceci tient au fait que les concepts fondamentaux sur lesquels reposent les méthodes utilisées n'ont pas changé : ce sont ceux de la génétique mendélienne.

Si les concepts de base n'ont pas évolué, les méthodes se sont néanmoins affinées au cours du temps. En outre, le développement de l'informatique, en permettant l'exécution de certains calculs impossibles à réaliser à la main, a ouvert de nouvelles voies d'investigation.

Les diverses méthodes proposées tentent d'apporter des réponses aux questions suivantes : quelle a été, dans la population considérée, l'évolution de la consanguinité et de l'apparentement ? Existe-t-il une tendance à la subdivision en lignées ? Quelle est l'influence génétique exercée par divers ancêtres ? Quel est le « poids génétique » des divers élevages ?

A. Mesure de la consanguinité

En 1922, WRIGHT définit le coefficient F de consanguinité d'un individu comme le coefficient de corrélation entre les deux gamètes dont il est issu. Plus tard, HALDANE et MOSHINSKY (1939), COTTERMAN (1940) et MALECOT (1948) préférèrent le considérer comme la probabilité pour que deux gènes homologues de l'individu considéré soient « identiques », c'est-à-dire dérivent par descendance mendélienne d'un même gène ancêtre (les mutations étant exclues).

Dès 1923, WRIGHT présente, dans « Journal of Heredity », un article consacré à l'analyse mendélienne des races pures d'animaux domestiques, dans lequel il traite notamment de la mesure de la consanguinité. Il montre comment le coefficient F permet de quantifier et de traduire en termes de génétique mendélienne un concept dont les sélectionneurs étaient familiers depuis longtemps. WRIGHT insiste sur le caractère relatif du coefficient de consanguinité. En effet, ce coefficient mesure la réduction relative d'hétérozygotie de l'individu (ou du groupe d'individus) considéré, par rapport au stock génétique ancestral, qui dépend lui-même de la date à laquelle remontent les pedigrees. En outre, WRIGHT souligne que cette mesure de la diminution relative du degré d'hétérozygotie ne vaut que pour des gènes non soumis à la sélection.

WRIGHT expose, de façon détaillée et à l'aide d'exemples, le mode de calcul du coefficient F :

$$F_A = \sum \binom{1}{-}^{n+n'+1} (1 + F_c)$$

Le coefficient de consanguinité F_A de l'individu A est égal à une somme de termes de la forme $\left(\frac{1}{2}\right)^{n+n'+1} (1+F_c)$. Chacun de ces termes se rapporte à une chaîne de parenté (« tie ») partant du père de l'individu A, remontant par n générations jusqu'à l'ancêtre C (ayant lui-même un coefficient de consanguinité égal à F_c) et redescendant par n' générations jusqu'à la mère de A. Les nombres n et n' peuvent prendre la valeur zéro : par convention, on admet que tout individu est son propre ancêtre. Chaque chaîne ne doit pas passer plus d'une fois par le même individu, ceci afin de ne conserver que des probabilités relatives à des événements incompatibles entre eux. La somme comprend autant de termes qu'il existe de chaînes distinctes, un même ancêtre pouvant éventuellement être au sommet de plusieurs chaînes.

WRIGHT (1923a) montre la supériorité du coefficient F par rapport aux autres méthodes d'évaluation de la consanguinité (PEARL, 1917) : cette supériorité tient au fait que le coefficient F repose sur une conception mendélienne de l'hérédité.

B. Mesure de la parenté

1. Coefficient de relation

WRIGHT (1922) définit le coefficient de relation * R_{AB} entre deux individus A et B comme le coefficient de corrélation entre ces deux individus, pour un caractère dont le déterminisme est entièrement génétique, autosomal, sans dominance et sans épistasie. WRIGHT donne le mode de calcul de ce coefficient R_{AB} :

$$R_{AB} = \frac{\sum \left[\left(\frac{1}{2}\right)^{n+n'} (1+F_c) \right]}{\sqrt{(1+F_A) (1+F_B)}}$$

F_A , F_B et F_c désignent respectivement les coefficients de consanguinité des individus A, B et C. Le numérateur est composé d'une somme de termes de la forme $\left(\frac{1}{2}\right)^{n+n'} (1+F_c)$.

Chaque terme se rapporte à une chaîne de parenté reliant les individus A et B par l'ancêtre C, n et n' désignant les nombres respectifs de générations qui séparent l'ancêtre commun C des individus A et B. Chaque chaîne ne doit pas passer plus d'une fois par le même individu, pour la raison déjà évoquée lors de la présentation du calcul du coefficient de consanguinité (§ III, A).

2. Coefficient de parenté

MALECOT (1948), adoptant une approche probabiliste, définit le coefficient de parenté* Φ_{AB} entre deux individus A et B comme la probabilité d'identité entre deux gènes homologues pris au hasard l'un chez A, l'autre chez B. Il découle de cette définition que le coefficient de consanguinité d'un individu est égal au coefficient de parenté entre son père et sa mère.

* En anglais : « coefficient of relationship ».

* Le terme « coefficient de parenté » a été traduit en anglais de diverses manières : « coefficient of kinship » (MALECOT), « coefficient of parentage » (KEMPTHORNE, 1957), « coefficient of coancestry » (FALCONER, 1960).

Entre le coefficient de relation R_{AB} de WRIGHT et le coefficient de parenté Φ_{AB} de MALECOT, existe la relation suivante :

$$R_{AB} = \frac{2 \Phi_{AB}}{\sqrt{(1 + F_A) (1 + F_B)}}$$

où F_A et F_B désignent respectivement les coefficients de consanguinité des individus A et B. Dans le cas particulier où les individus A et B ne sont pas consanguins, le coefficient de relation R_{AB} est simplement le double du coefficient de parenté Φ_{AB} .

Lorsque les pedigrees s'étendent sur de nombreuses générations, le calcul manuel des coefficients de consanguinité et de relation devient long et laborieux, voire impraticable, en raison du foisonnement des lignées d'ascendance à explorer. C'est pourquoi WRIGHT et Mc PHEE (1925) proposent une méthode approchée, fondée sur l'échantillonnage des lignées d'ascendance.

C. Méthodes approchées de calcul des coefficients de consanguinité et de relation

1. - Présentation des méthodes approchées de WRIGHT et Mc PHEE (1925) et de ROBERTSON et MASON (1954)

Pour des raisons évidentes, les ancêtres communs sont, dans les populations d'animaux domestiques, beaucoup plus souvent des mâles que des femelles. Le choix de lignées entièrement mâles ou femelles est donc à rejeter à cause du biais systématique qu'il entraînerait. Le système de lignées alternant régulièrement mâles et femelles donne des résultats satisfaisants, mais la théorie statistique ne peut, dans ce cas, s'appliquer simplement pour calculer l'erreur. WRIGHT et Mc PHEE optent donc pour l'extraction purement aléatoire de lignées d'ascendance, le choix entre ascendant mâle ou femelle s'opérant à chaque génération par tirage au sort équiprobable.

Pour estimer le coefficient moyen de consanguinité d'un ensemble d'individus, la méthode la plus simple consiste à tracer, pour chaque individu A, deux lignées aléatoires d'ascendance, l'une à partir de son père, l'autre à partir de sa mère. Puis on recherche dans ces deux lignées l'ancêtre commun C le plus proche. S'il n'y en a pas, le coefficient est nul. S'il y en a un, situé à n générations du père et n' de la mère, sa contribution au coefficient de consanguinité de l'individu A est égale à $\left(\frac{1}{2}\right)^{n+n'+1} (1 + F_C) (1 + F_C)$ où F_C désigne son propre coefficient de consanguinité. Comme le père de A possède 2^n ancêtres de rang n et que sa mère en possède $2^{n'}$ de rang n', la paire de lignées considérée n'est qu'une des $2^{n+n'}$ paires possibles : pour obtenir une estimation de la consanguinité, il faut multiplier la contribution de l'ancêtre commun C par $2^{n+n'}$. L'estimation du coefficient F_A est donc simplement égale à $\frac{1}{2} (1 + F_C)$. Si on néglige provisoirement la consanguinité des ancêtres communs, l'estimation de la consanguinité d'un individu ne peut prendre que deux valeurs :

- 0 en l'absence d'ancêtre commun dans les deux lignées d'ascendance considérées ;
- $\frac{1}{2}$ s'il existe au moins un ancêtre commun dans ces deux lignées.

L'estimateur de la consanguinité moyenne d'un groupe de M individus est donc égal à $\frac{1}{2} \frac{m}{M}$, où m désigne le nombre d'individus (parmi les M considérés) pour lesquels l'examen

des deux lignées aléatoires d'ascendance montre l'existence d'au moins un ancêtre commun. Il est évident qu'une telle méthode ne présente aucun intérêt pour estimer la consanguinité d'un individu en particulier. En revanche, elle donne de bons résultats quand il s'agit d'estimer la consanguinité moyenne d'un groupe d'individus.

Pour estimer le coefficient de consanguinité d'un individu particulier, WRIGHT et Mc PHEE proposent une méthode qui consiste à reconstituer intégralement l'arbre généalogique de l'individu considéré jusqu'aux 2^n ancêtres de rang n (par exemple jusqu'aux 32 ancêtres de rang 5), puis à tracer des lignées aléatoires d'ascendance, telles qu'elles ont été définies précédemment, à partir de ces 2^{n-1} (soit 16 si $n=5$) ancêtres paternels et 2^{n-1} ancêtres maternels. Il faut alors traiter, non plus une, mais $2^{n-1} \times 2^{n-1}$ (soit $16 \times 16 = 256$ si $n=5$) paires de lignées. Cette méthode peut servir à estimer les coefficients F_C de consanguinité des ancêtres communs qui, dans un premier temps, avaient été négligés lors de l'estimation de la consanguinité moyenne d'un groupe d'individus. Les auteurs présentent de façon détaillée le mode de calcul du coefficient et de l'erreur.

Enfin, ils généralisent la méthode d'estimation de la consanguinité moyenne d'un groupe d'individus par la prise en compte de pedigrees à 4 (ou même à 8) lignées d'ascendance partant des 4 grands-parents (ou des 8 arrière-grands-parents). Pour chaque pedigree, il faut traiter, non plus une, mais $2 \times 2 = 4$ (ou $4 \times 4 = 16$) paires de lignées.

L'estimation du coefficient de relation R_{AB} de deux individus A et B se fait selon le même principe que celle du coefficient de consanguinité. Au lieu d'explorer tout l'arbre généalogique de chacun des deux individus A et B, on considère seulement une ou deux lignées aléatoires d'ascendance à partir de chacun d'eux.

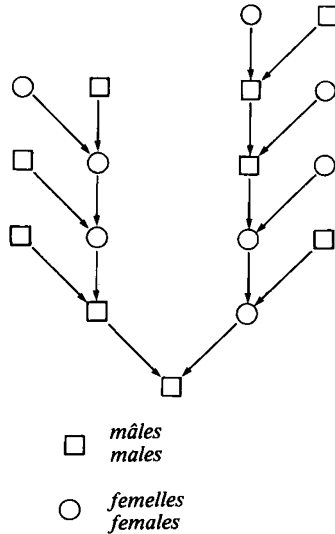


FIG. 2

Exemple de pedigree à deux doubles lignées.
Example of a two-double lines pedigree.

En 1954, ROBERTSON et MASON proposent une variante de la méthode d'estimation du coefficient moyen de consanguinité d'un échantillon d'animaux. A partir de chaque animal, ils tracent deux lignées aléatoires d'ascendance, au sens précédemment défini par WRIGHT

et Mc PHEE (1925). En outre, ils enregistrent, à chaque génération, le numéro de l'autre parent (le père ou la mère suivant les cas), constituant ainsi un pedigree dit « à deux doubles lignées » dont un exemple est fourni par la figure 2. La constitution d'un tel « pedigree à deux doubles lignées » ne nécessite pratiquement aucun travail supplémentaire par rapport à celle d'un pedigree à deux lignées simples : en effet, lorsqu'on recherche dans un fichier généalogique le numéro du père d'un animal, on obtient simultanément celui de sa mère, et réciproquement. Considérons un pedigree « à deux doubles lignées » d'un individu A et, dans ce pedigree, un ancêtre commun situé à n générations du père et n' générations de la mère de A. Sa contribution au coefficient de consanguinité de A est égale, par définition, à $\left(\frac{1}{2}\right)^{n+n'+1}$ si l'on suppose nulle sa propre consanguinité. Or un pedigree « à deux doubles lignées » ne comprend que 2 des 2^n ascendants de rang n du père de A et 2 des $2^{n'}$ ascendants de rang n' de la mère de A. Pour obtenir une estimation de la contribution des ancêtres de rangs respectifs n et n' au coefficient de consanguinité de A, il faut donc multiplier $\left(\frac{1}{2}\right)^{n+n'+1}$ par le facteur $\frac{2^n \times 2^{n'}}{2 \times 2}$. On obtient la valeur $\frac{1}{8}$. Ainsi, en appliquant un raisonnement analogue à celui de WRIGHT et Mc PHEE, on montre que le nombre de chaînes de parenté distinctes observées sur les pedigrees « à deux doubles lignées » d'un échantillon de M individus, divisé par 8 M, fournit une estimation du coefficient moyen de consanguinité de cet échantillon (à condition de négliger la consanguinité des ancêtres communs).

En 1976, COOK et HARTL généralisent les résultats de WRIGHT et Mc PHEE (1925) en calculant la variance de l'estimateur du coefficient de consanguinité dans le cas de pedigrees à 2^n lignées aléatoires d'ascendance (n du côté paternel et n du côté maternel).

CILLIERS (1976) propose une généralisation de la méthode approchée de WRIGHT et Mc PHEE : il suggère de reconstituer intégralement l'arbre généalogique de l'individu considéré jusqu'aux 2^n ascendants d'un rang n fixé, puis de tracer un nombre arbitraire de lignées aléatoires d'ascendance de chaque côté (paternel et maternel) du pedigree.

2. Comparaison entre méthode approchée (« short method ») et méthode complète (« long method »)

WRIGHT et Mc PHEE (1925) appliquent leur méthode approchée, fondée sur des pedigrees réduits à quatre lignées d'ascendance, pour estimer la consanguinité moyenne de 64 femelles Shorthorn de la famille « Duchess » sélectionnée par Thomas Bates pendant la première moitié du XIX^e siècle. Le coefficient moyen approché ainsi obtenu est de $42,2 \text{ p.}100 \pm 1,1 \text{ p.}100$, alors que le coefficient calculé à partir des pedigrees complets est de $40,9 \text{ p.}100$: les auteurs concluent que, dans un tel cas, le considérable gain de temps dû à la méthode approchée compense le léger écart observé entre la valeur estimée et la valeur exacte de la consanguinité.

FOWLER (1932), dans l'étude qu'il consacre à la race bovine Ayrshire, compare avantages et inconvénients des deux méthodes. Il conclut que la méthode approchée est bien adaptée lorsqu'on désire estimer la consanguinité moyenne d'un groupe d'individus dont les pedigrees s'étendent sur un assez grand nombre de générations (supérieur ou égal à 6). En revanche, il est évident que, dans le cas d'une information généalogique plus restreinte (remontant, par exemple, jusqu'aux arrière-grands-parents), il vaut mieux utiliser la méthode complète. La méthode complète s'impose également, même dans le cas d'importants pedigrees, lorsqu'on souhaite analyser dans le détail la façon dont les accouplements ont été réalisés.

3. Intérêt des méthodes approchées

La méthode approchée de WRIGHT et Mc PHEE (1925) ainsi que celle de ROBERTSON et MASON (1954) ont été très largement utilisées (tableau 1). On aurait pu penser que l'avènement de l'informatique allait considérablement diminuer l'intérêt de ces méthodes et les faire tomber en désuétude. Or ce n'est pas le cas. Certes, les ordinateurs permettent de stocker de grands fichiers généalogiques et de réaliser rapidement les calculs complexes de consanguinité. Mais dans certaines études rétrospectives, exploitant de très importants fichiers non informatisés, la saisie exhaustive de l'information généalogique peut s'avérer trop longue et trop onéreuse. Dans de tels cas, les méthodes approchées, fondées sur un échantillonnage aléatoire des lignées, conservent tout leur intérêt.

D. Calcul de l'effectif génétique à partir de la consanguinité

Considérons une population idéale comprenant, à chaque génération, N individus diploïdes à sexes confondus. Les $2N$ gamètes nécessaires à la constitution des N zygotes de la génération G_n sont tirés au hasard dans un pool comprenant un nombre infini de gamètes qui proviennent, de façon équiprobable, des N individus de la génération précédente G_{n-1} (l'autofécondation n'étant pas exclue).

Soit F_n le coefficient de consanguinité d'un individu quelconque de la génération G_n . Soit H_n l'hétérozygotie dans la génération G_n .

$$H_n = (1 - F_n) H_0 \quad (1)$$

D'où :

$$F_n = 1 - \frac{H_n}{H_0} \quad (2)$$

Le rapport $\frac{H_n}{H_0}$ est, par définition, « l'hétérozygotie relative » dans la génération G_n .

Le coefficient F_n s'exprime en fonction de F_{n-1} et de N (WRIGHT, 1931) :

$$F_n = \left(1 - \frac{1}{2N}\right) F_{n-1} + \frac{1}{2N} \quad (3)$$

Si l'on substitue à F_n et F_{n-1} leurs expressions en fonction de l'hétérozygotie relative (2), la relation (3) s'écrit :

$$\frac{H_{n-1} - H_n}{H_{n-1}} = \frac{1}{2N} \quad (4)$$

Le rapport $\frac{H_{n-1} - H_n}{H_{n-1}}$ représente la diminution relative d'hétérozygotie entre deux générations successives : celle-ci est égale à $\frac{1}{2N}$.

L'effectif génétique (dans le sens d'« inbreeding effective number ») N_e d'une population réelle quelconque est égal à l'effectif de la population idéale de référence (telle qu'elle vient d'être définie) qui aurait la même diminution relative d'hétérozygotie entre deux générations successives.

TABLEAU 1

Méthodes de calcul des coefficients de consanguinité et de parenté utilisées dans quelques publications de 1923 à 1981

Methods of calculating inbreeding and relationship coefficients used in some papers from 1923 to 1981

Année Year	Auteurs Authors	Espèce Species	Race Breed	Pays Land	Méthodes utilisées Methods used					
					1	2	3	4	5	6
1923 b	WRIGHT	B	Shorthorn (Famille Duchess)	G.B.	+					
1925	WRIGHT et H.C Mc PHEE	B	Shorthorn	G.B.				+	+	
1926	H.C. Mc PHEE et WRIGHT	B	Dairy Shorthorn	G.B.				+		
1927	CALDER	E	Clydesdale	G.B.	+					
1931	BROCKELBANK et WINTERS	B	Shorthorn	U.S.A.		+(6)				
1932	FOWLER	B	Ayrshire	G.B.	+			+		
1933	DICKSON et LUSH	O	Rambouillet	U.S.A.				+		
1936	LUSH <i>et al.</i>	B	Holstein	U.S.A.				+	+	
1937	YODER et LUSH	B	Brown Swiss	U.S.A.				+		
1937	WILLHAM	B	Hereford	U.S.A.	+			+		
1939	LUSH et ANDERSON	P	Poland China	U.S.A.				+		
1943	STONAKER	B	Aberdeen Angus	U.S.A.				+		
1945	FLETCHER	E	Quarter Horse	U.S.A.					+	
1947	LENER et HAZEL	V	Californienne	U.S.A.	+					
1951 a	ROBERTSON et ASKER	B	Frisonne	G.B.						+
1952 b	STEWART	B	Jersey	Nlle-Zélande						+
1954	KUTSAL	B	Brune des Alpes	Suisse		+(5)				
1954	ROBERTSON et MASON	B	Rouge danoise	Danemark						+
1954	STEWART	B	Jersey	Nlle-Zélande						+
1955	STEWART	B	Ayrshire	Nlle-Zélande						+
1956	CLAYTON	B	Shorthorn	G.B.						+
1957 b	BARKER	B	Jersey	Australie						+
1957	SYRSTAD	B	Telemark et Red Poll	Norvège					+	
1959	BEKEDAM	P	Races hollandaises	Pays-Bas		+(4)				
1960	BARKER et DAVEY	B	Poll Hereford	Australie					+	
1961	LANGLET et GRAVERT	B	Schwarzbunt	R.F.A.		+(5 et 7)				
1962	YOUNG et PURSER	O	Border Leicester	G.B.						+
1963	BOLDUAN	B	Schwarzbunt	R.D.A.		+(5)				
1963	DAVEY et BARKER	B	Hereford	Australie					+	
1963	WATSON	B	Welsh Black	G.B.		+(3)				
1965	ALLAIRE et HENDERSON	B	Holstein	U.S.A.		+(2)				
1965	FLADE	B	Brandebourgeoise	R.D.A.		+(5)				
1965	C.P. Mc PHEE	P	Large White	Australie					+	
1966	SCHWARK et KITTNER	B	Schwarzbunt	R.D.A.		+(5)				
1967	FIEDLER et HIMMEL	B	Fleckvieh	R.D.A.		+(4)				
1967	O'CONNOR et WILLIS	B	Frisonne	G.B.						+
1968	LINDEMANN	B	Schwarzbunt	R.D.A.		+(3 et 5)				
1971	JONSSON	P	Landrace	Danemark					+	
1971	LINDSTRÖM	B	Finnoise Ayrshire	Finlande		+(7)				
1973	HANSET	P	Piétrain	Belgique	+	+(5)				
1975	LEDERER <i>et al.</i>	B	Fleckvieh	R.F.A.		+(3)				
1975 c	MARTIN	O	Galway	Irlande	+					
1977 b	HERRON et PATTIE	B	Illawarra Shorthorn	Australie						+
1977	ÖZKÜTÜK et BICHARD	B	Hereford	G.B.		+(4)		+		
1978	BOWMAN <i>et al.</i>	B	Frisonne	G.B.					+	
1978 b	FOGARTY	O	Dorset Horn	Australie				+		
1978	HERRON	B	Sahiwal	Australie						+
1978	IKIN et HERRON	C	British Alpine	Australie						+
1978	SMITH <i>et al.</i>	P	Large White, Welsh et Landrace	G.B.		+(4 et 5)				
1979	HODGES <i>et al.</i>	B	Holstein	U.S.A.		+(4)				
1979	TOLL et BARKER	B	Poll Hereford	Australie		+(4)				+
1980	MAYRHOFER et WOLLINGER	E	Lipizzan	Autriche	+					
1981	MAHON et CUNNINGHAM	E	Pur sang	Irlande		+(5)			+	+

LEGENDE DU TABLEAU 1

Espèces

B = bovins, *C* = caprins, *E* = équins, *O* = ovins, *P* = porcins, *V* = poule

Méthodes

1 pedigree complet

2 pedigree sur *n* générations (*n* figure entre parenthèses)

3 pedigree à 2 lignées

4 pedigree à 4 lignées } (méthode de WRIGHT et Mc PHEE, 1925)

5 pedigree à 8 lignées

6 pedigree à 2 doubles lignées (méthode de ROBERTSON et MASON, 1954)

Species

B = cattle, *C* = goat, *E* = horse, *O* = sheep, *P* = swine, *V* = poultry

Methods

1 complete pedigree

2 *n*-generations pedigree (*n* is bracketed)

3 2-lines pedigree

4 4-lines pedigree } (WRIGHT and Mc PHEE, 1925)

5 8-lines pedigree

6 2-double lines pedigree (ROBERTSON and MASON, 1954)

Considérons maintenant une population présentant les mêmes caractéristiques que la population idéale précédemment décrite, mais où les sexes sont séparés : chaque génération comprend N_m mâles et N_f femelles. WRIGHT (1931) a établi la formule donnant la diminution relative d'hétérozygotie entre deux générations successives G_{n-1} et G_n :

$$\frac{H_{n-1} - H_n}{H_{n-1}} = \frac{1}{8 N_m} + \frac{1}{8 N_f} \quad (5)$$

Si N_m est petit devant N_f (ce qui est généralement le cas dans les populations d'animaux domestiques) on peut écrire :

$$\frac{H_{n-1} - H_n}{H_{n-1}} \approx \frac{1}{8 N_m} \quad (6)$$

Par définition, le « nombre efficace de reproducteurs » (« effective number of males ») N_m d'une population réelle quelconque est égal au nombre de mâles de la population idéale à sexes séparés qui aurait la même diminution relative d'hétérozygotie entre deux générations successives.

Plusieurs auteurs utilisent les résultats relatifs à l'augmentation moyenne de consanguinité par génération pour estimer le nombre efficace de reproducteurs (LUSH, 1932 ; DICKSON et LUSH, 1933 ; YODER et LUSH, 1937 ; LUSH et ANDERSON, 1939 ; LERNER et HAZEL, 1947 ; YOUNG et PURSER, 1962 ; TOLL et BARKER, 1979). Il est à noter que cette méthode d'estimation suppose la stationnarité de la population pendant la période considérée.

Le nombre efficace de reproducteurs de chaque population réelle traitée représente le nombre de reproducteurs de la population idéale équivalente du point de vue de l'augmentation de consanguinité : il constitue un élément de comparaison entre populations (LUSH, 1932). De plus, il peut permettre de prévoir l'évolution de la consanguinité, moyennant l'hypothèse de la persistance de la stationnarité.

E. Décomposition de la consanguinité

1. Décomposition en une somme de contributions dues à divers ancêtres

A la suite de WRIGHT (1923a), plusieurs auteurs tels que FOWLER (1932), STONAKER (1943), WATSON (1963) ou HANSET (1973), décomposent la consanguinité totale d'un individu (ou d'un groupe d'individus) en une somme de contributions dues aux divers ancêtres communs situés au sommet des chaînes de parenté. Ils définissent la contribution d'un ancêtre commun C comme la somme des termes de la forme $\left(\frac{1}{2}\right)^{n+n'+1} (1+F_c)$.

Il faut toutefois remarquer que, dans la contribution de l'ancêtre C ainsi définie, intervient la propre consanguinité de C qui résulte elle-même d'ancêtres communs à son père et à sa mère appartenant à des générations antérieures. La méthode décrite conduit à surestimer les contributions des ancêtres communs les plus proches aux dépens des plus éloignés. Si cette décomposition de la consanguinité est discutable, elle n'en a pas moins le mérite de permettre un repérage des ancêtres jouant un rôle-clé dans la consanguinité de la population.

Certains auteurs (FOWLER, 1932 ; ROBERTSON et MASON, 1954 ; MARTIN, 1975c) classent les ancêtres communs en fonction du pourcentage de chaînes de parenté dont ils constituent le sommet. Il s'agit là d'une approche tout à fait différente, qui permet simplement de repérer les ancêtres sur lesquels confluent les plus grands nombres de chaînes de parenté. Comme le souligne FOWLER (1932) dans son étude sur les bovins Ayrshire de Grande-Bretagne, il n'existe pas de relation automatique entre la proportion de chaînes de parenté au sommet desquelles se trouve un individu et sa contribution à la consanguinité (au sens défini précédemment).

2. Décomposition de la consanguinité en consanguinité « proche », consanguinité « à long terme » et consanguinité « de lignée »

H.C. Mc PHEE et WRIGHT (1925) établissent la relation existant, dans une population panmictique, entre le coefficient moyen de consanguinité F et le coefficient de relation moyen R entre individus :

$$\text{D'où :} \quad R = \frac{2 F}{1 + F} \quad (7)$$

$$F = \frac{R}{2 - R} \quad (8)$$

Dans une population quelconque, la comparaison entre le F observé et le F calculé à partir de R permet de déterminer s'il existe une tendance à la subdivision en lignées : l'écart entre F observé et F calculé est d'autant plus grand que cette tendance est plus affirmée (DICKSON et LUSH, 1933 ; LUSH *et al.*, 1936 ; WILLHAM, 1937 ; YODER et LUSH, 1937 ; LUSH et ANDERSON, 1939 ; CARTER, 1940 ; STONAKER, 1943 ; FLETCHER, 1945 ; LERNER et HAZEL, 1947).

LUSH (1946) définit un « index de subdivision » égal au rapport entre le F observé et le F calculé à partir de R. Cet index, égal à 1 dans une population panmictique, augmente avec le degré de subdivision de la population en lignées.

ROBERTSON et ASKER (1951a) distinguent plusieurs composantes dans la consanguinité totale d'une population. La consanguinité proche (« current inbreeding ») résulte d'accouplements entre proches parents (parent-enfant, germains ou demi-germains). La consanguinité restante (« non-current inbreeding ») comprend :

— la consanguinité à long terme, due à des ancêtres ayant exercé une grande influence génétique sur la population considérée ;

— la consanguinité de lignée, résultant de la subdivision de la population en lignées.

Une fois mise à part la consanguinité proche, la consanguinité à long terme F_1 est définie comme la part de la consanguinité restante qui serait observée dans le cas d'accouplements au hasard. La relation (8) établie par H.C. Mc PHEE et WRIGHT en 1925 permet d'exprimer F_1 en fonction du coefficient de relation moyen R de la population :

$$F_1 = \frac{R}{2 - R} \quad (9)$$

La consanguinité de lignée se calcule par différence entre « non-current inbreeding » et consanguinité à long terme. Cette méthode de décomposition de la consanguinité totale d'une population est utilisée par de nombreux auteurs (tableau 2). Il semble cependant qu'elle n'établisse pas une partition rigoureuse de la consanguinité. En effet, la consanguinité à long terme, telle qu'elle est estimée à partir du coefficient moyen de relation R , peut englober une certaine part de consanguinité proche : dans une population quelconque, il n'est pas exclu que des accouplements panmictiques se produisent entre apparentés proches.

WRIGHT (1969) introduit une approche de la subdivision hiérarchique d'une population en lignées, donnant ainsi un prolongement théorique à la pratique de la décomposition de la consanguinité. Il considère une population totale T subdivisée en sous-populations S , sans faire d'hypothèses particulières sur leur degré d'isolement ni sur leur mode de reproduction (panmictique ou non). Il appelle F_{IT} le coefficient de corrélation gamétique entre les gamètes s'unissant pour engendrer les individus I de la population totale T , F_{IS} la moyenne sur toutes les subdivisions du coefficient de corrélation gamétique entre les gamètes s'unissant pour engendrer les individus I d'une subdivision S et F_{ST} le coefficient de corrélation gamétique entre des gamètes pris au hasard dans les subdivisions S de la population totale T . Utilisant la formule de WAHLUND (1928), il établit la relation :

$$F_{ST} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}} \quad (10)$$

Le coefficient F_{ST} représente le rapport de la variance de la fréquence d'un gène autosomal dans les diverses subdivisions à la valeur maximale de cette variance (dans le cas où les sous-populations sont complètement isolées et génétiquement fixées). Le coefficient F_{ST} est toujours positif ; le coefficient F_{IS} , habituellement positif, peut être négatif si les unions consanguines sont systématiquement évitées dans les sous-populations.

WRIGHT (1977) applique cette décomposition de la consanguinité à diverses populations réelles étudiées par lui-même ou d'autres auteurs (Mc PHEE et WRIGHT, 1925 et 1926 ; CALDER, 1927 ; DICKSON et LUSH, 1933 ; SCIUCHETTI, 1935 ; LUSH *et al.*, 1936 ; ROTTENSTEN, 1937 ; WILLHAM, 1937 ; YODER et LUSH, 1937 ; LUSH et ANDERSON, 1939 ; CARTER, 1940 ; STONAKER, 1943 ; STEELE, 1944). Le coefficient F_{IT} s'identifie au coefficient de consanguinité totale et le coefficient F_{ST} (nommé par WRIGHT « consanguinité

TABLEAU 2

Principales méthodes d'analyse génétique d'une population utilisées dans quelques publications de 1923 à 1979

Most frequently used methods of genetic analysis of a breed in some papers from 1923 to 1979

Année Year	Auteurs Authors	Espèce Species	Race Breed	Pays Land	Méthodes utilisées Methods used				
					1	2	3	4	5
1923 b	WRIGHT	B	Shorthorn (Famille Duchess)	G.B.	+				
1925	H.C. Mc PHEE et WRIGHT	B	Shorthorn	G.B.	+				
1926	H.C. Mc PHEE et WRIGHT	B	Dairy Shorthorn	G.B.	+				
1932	FOWLER	B	Ayrshire	G.B.	+				
1933	DICKSON et LUSH	O	Rambouillet	U.S.A.	+		+	+	
1936	LUSH <i>et al.</i>	B	Holstein	U.S.A.	+		+		
1937	YODER et LUSH	B	Brown Swiss	U.S.A.	+		+		
1937	WILLHAM	B	Hereford	U.S.A.	+		+	+	
1939	LUSH et ANDERSON	P	Poland China	U.S.A.	+		+	+	
1940	CARTER	O	Hampshire	U.S.A.	+		+		
1943	STONAKER	B	Aberdeen Angus	U.S.A.	+		+	+	
1945	FLETCHER	E	Quarter Horse	U.S.A.	+		+		
1951 a	ROBERTSON et ASKER	B	Frisonne	G.B.	+	+	+	+	(+)
1952 b	STEWART	B	Jersey	Nlle-Zélande				+	
1953	WIENER	B	Ayrshire	G.B.					+
1954	KUTSAL	B	Brune des Alpes	Suisse	+		+		
1954	ROBERTSON et MASON	B	Rouge danoise	Danemark	+		+		
1954	STEWART	B	Jersey	Nlle-Zélande	+	+			
1955	STEWART	B	Ayrshire	Nlle-Zélande			+	+	+
1957 b)	BARKER	B	Jersey	Australie	+	+	+	+	+
1959 }		B	Telemark et Red Poll	Norvège	+	+	+		
1957	SYRSTAD	B	Telemark et Red Poll	Norvège	+	+	+		
1958	JAMES et Mc BRIDE	V	Australorps	Australie			+		
1960	BARKER et DAVEY	B	Poll Hereford	Australie	+	+	+	+	+
1962	YOUNG et PURSER	O	Border Leicester	G.B.	+	+	+		
1963	DAVEY et BARKER	B	Hereford	Australie	+		+	+	+
1963	WATSON	B	Welsh Black	G.B.	+				+
1965	C.P. Mc PHEE	P	Large White	Australie	+	+			+
1966	SCHWARK et KITTNER	B	Schwarzbunt	R.D.A.	+		+		
1967	O'CONNOR et WILLIS	B	Frisonne	G.B.	+	+	+	+	
1971	LINDSTRÖM et MAIJALA	B	Finnoise Ayrshire	Finlande	+		+	+	
1975 b	MARTIN	O	Galway	Irlande	+	+			+
et c		B	Illawarra Shorthorn	Australie	+	+	+		+
1977 b	HERRON et PATTIE	B	Hereford	G.B.	+	+	+		+
1977	ÖZKÜTÜK et BICHARD	B	Hereford	G.B.	+	+	+		+
1978	BOWMAN <i>et al.</i>	B	Frisonne	G.B.	+	+	+	+	
1978 b	FOGARTY	O	Dorset Horn	Australie	+	+	+		+
1978	HERRON	B	Sahiwal	Australie	+	+	+	+	+
1978	IKIN et HERRON	C	British Alpine	Australie	+	+			+
1979	SCHOEMAN	O	Karakul	Sud Ouest Africain					+
1979	TOLL et BARKER	B	Poll Hereford	Australie	+	+	+	+	+

LEGENDE DU TABLEAU 2

Espèces

B = bovins, *C* = caprins, *E* = équins, *O* = ovins, *P* = porcins, *V* = poule

Méthodes

1 coefficients de consanguinité et de « relationship » (WRIGHT, 1922)

2 décomposition de la consanguinité (ROBERTSON et ASKER, 1951 a)

3 coefficient de relation directe (DICKSON et LUSH, 1933)

4 étude des fondateurs (DICKSON et LUSH, 1933)

5 contributions génétiques des élevages (WIENER, 1953)

Species

B = cattle, *C* = goat, *E* = horse, *O* = sheep, *P* = swine, *V* = poultry

Methods

1 coefficients of inbreeding and relationship (WRIGHT, 1922)

2 decomposition of the total inbreeding (ROBERTSON and ASKER, 1951 a)

3 coefficient of direct relationship (DICKSON and LUSH, 1933)

4 study of the founders (DICKSON and LUSH, 1933)

5 genetic contributions of the herds or flocks (WIENER, 1953)

cumulative ») au coefficient F_I de consanguinité « à long terme » qui peut être calculé à partir du coefficient moyen de relation R de la population grâce à la formule (9). Le coefficient F_{IS} est calculé à partir des coefficients F_{IT} et F_{ST} grâce à la relation (10) :

$$F_{IS} = \frac{F_{IT} - F_{ST}}{1 - F_{ST}} \quad (11)$$

Alors que le coefficient F_{ST} mesure la consanguinité accumulée dans la population totale depuis l'origine considérée, le coefficient F_{IS} évalue l'importance des unions consanguines actuelles. WRIGHT appelle le coefficient F_{IS} « current inbreeding », bien qu'il ne s'identifie pas au coefficient ainsi nommé par ROBERTSON et ASKER (1951a) et par de nombreux auteurs à leur suite (tableau 2).

F. Coefficient de relation entre un individu et l'ensemble d'une population

Le coefficient de relation d'un ancêtre avec l'ensemble de la population permet de mesurer son influence génétique (WRIGHT, 1923b). Dans une étude consacrée à la population ovine Rambouillet d'Amérique, DICKSON et LUSH (1933) mettent l'accent sur un point important concernant l'interprétation du coefficient de relation d'un animal donné avec l'ensemble de la population. Ce coefficient dépend :

- de l'importance de la descendance laissée dans la population par l'animal considéré ;
- de ses relations collatérales avec la population, résultant de l'existence de descendants issus de ses propres ancêtres.

Ceci conduit les auteurs à distinguer deux composantes dans le coefficient de relation d'un individu I avec l'ensemble d'une population. La première (relation directe) comprend la somme des termes relatifs aux chaînes de parenté dont l'individu I est lui-même le sommet. La seconde (relation collatérale) comprend la somme des termes relatifs aux autres chaînes.

A la suite de DICKSON et LUSH (1933), de nombreux auteurs utilisent la méthode approchée de WRIGHT et Mc PHEE (1925) pour estimer le coefficient de relation directe

d'un ancêtre particulier avec l'ensemble d'une population (tableau 2). Ils estiment le coefficient de relation directe d'un animal par la proportion de lignées d'ascendance dans lesquelles il apparaît, négligeant implicitement la consanguinité de cet animal et celle des individus de l'échantillon représentatif de la population.

De façon générale, les auteurs assimilent le coefficient de relation directe d'un animal à sa contribution au stock génétique de la population. Si, par exemple, un animal possède un coefficient de 0,10, ils concluent que 10 % des gènes de la population proviennent de cet animal (en l'absence de sélection). D'un point de vue strictement théorique, cette assimilation nous paraît abusive. Rappelons l'expression du coefficient de relation directe entre un animal A et un de ses descendants D :

$$R = \frac{\left(\frac{1}{2}\right)^n (1 + F_A)}{(1 + F_A) (1 + F_D)}$$

F_A et F_D désignent respectivement les coefficients de consanguinité des animaux A et D ; n désigne le nombre de générations séparant A et D. La somme figurant au numérateur comprend autant de termes qu'il existe de chaînes d'ascendance reliant D à A. La proportion P de gènes de l'animal D issus de son ancêtre A est égale à $\left(\frac{1}{2}\right)^n$. Il n'y a donc égalité entre R et P que dans le cas particulier où les individus A et D ne sont pas consanguins. Néanmoins, d'un point de vue pratique, l'estimation de R par la méthode approchée de WRIGHT et Mc PHEE peut être assimilée à celle de P, puisqu'elle néglige implicitement les coefficients de consanguinité F_A et F_D .

G. Étude des fondateurs

Les fondateurs sont, par définition, les individus situés au sommet des lignées d'ascendance. DICKSON et LUSCH (1933), dans leur étude consacrée à la population ovine Rambouillet d'Amérique, observent une concentration des lignées d'ascendance sur certains fondateurs : ce sont ceux qui ont joué un rôle important dans l'histoire génétique de la population. Appliquant la méthode approchée précédemment décrite, DICKSON et LUSH estiment la contribution génétique d'un fondateur par la proportion de lignées dont il est le sommet.

Les auteurs regroupent les fondateurs en fonction de leur date de naissance afin d'évaluer les contributions génétiques des vagues successives d'importation d'animaux. D'autre part, le regroupement des fondateurs par élevage d'origine leur permet d'apprécier les contributions génétiques respectives des divers troupeaux.

A la suite de DICKSON et LUSH, de nombreux auteurs ont utilisé cette méthode d'étude des fondateurs, notamment pour mesurer l'impact génétique des importations d'animaux (tableau 2).

H. Contributions génétiques des divers élevages

L'analyse démographique permet de mettre en évidence, quand elles existent, la stratification des élevages et la concentration des origines des reproducteurs (cf. § II, C, 1 et II, C, 2). On est naturellement conduit à poursuivre, en termes génétiques, cette analyse purement démographique. Dans cette perspective, il paraît intéressant de détecter les élevages

qui ont contribué le plus fortement à la formation du stock génétique d'une population et de quantifier leurs contributions par un calcul de pourcentages de gènes.

ROBERTSON et ASKER (1951 a) apprécient l'importance d'un élevage par la proportion de lignées d'ascendance traversant cet élevage. Ils notent que cette proportion n'a pas de signification génétique précise mais qu'elle fournit une évaluation grossière du rôle de l'élevage considéré dans l'histoire de la population étudiée.

Dans une étude consacrée à la structure de la population bovine Ayrshire de Grande-Bretagne, WIENER (1953) propose une méthode destinée à mesurer précisément la contribution d'un élevage, ou d'un groupe d'élevages, à la constitution du stock génétique d'une population. Il choisit un échantillon représentatif d'animaux et reconstitue leurs pedigrees complets sur 4 générations (c'est-à-dire jusqu'aux arrière-arrière-grands-parents), en enregistrant, pour chaque ascendant rencontré, l'élevage dont il est issu.

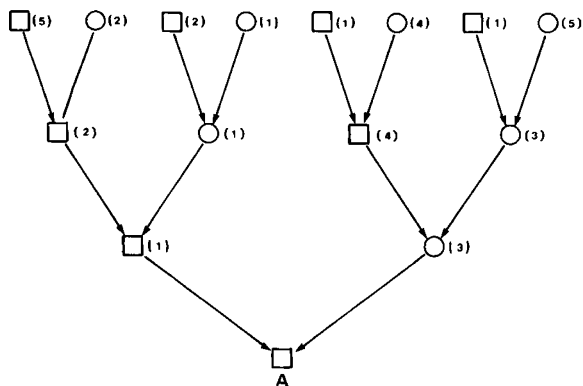
Afin de repérer les élevages importants, il envisage deux méthodes. La première consiste à compter le nombre d'apparitions de chaque élevage dans l'ensemble des pedigrees considérés. BARKER (1957 b) proposera des variantes simplifiées de cette méthode, consistant à réduire le champ d'investigation aux élevages ayant fourni les mâles ou, plus encore, aux seuls élevages apparaissant dans les lignées paternelles (« sires-of-sires lines »). La deuxième méthode de WIENER (1953) est fondée sur le nombre d'apparitions de chaque élevage dans les pedigrees, modifié par un système de pondération qui fait intervenir la génération dans laquelle se situe chaque apparition : ainsi, une apparition d'un élevage dans la génération des parents compte 2 fois plus que dans celle des grands-parents, 4 fois plus que dans celle des arrière-grands-parents, et 8 fois plus que dans celle des arrière-arrière-grands-parents.

WIENER souligne que ces méthodes ne fournissent pas une estimation de la contribution génétique de chaque élevage considéré, car les apparitions d'un même élevage dans une même lignée d'ascendance ne sont pas indépendantes les unes des autres. C'est pourquoi il définit la contribution génétique d'un élevage dans la génération n comme le pourcentage d'apparitions de cet élevage dans la génération n qui sont indépendantes d'apparitions dans les générations m ($m < n$). Par convention, les générations sont numérotées de 0 pour les individus de l'échantillon étudié, à 4 pour leurs arrière-arrière-grands-parents. Si, par exemple, un élevage apparaît dans un pedigree au niveau du père, les autres apparitions de ce même élevage au niveau des ascendants du père ne seront pas prises en compte. La contribution génétique totale d'un élevage est égale à la somme des contributions calculées dans chacune des 4 générations d'ascendants.

La même définition vaut pour un groupe d'élevages. Il est à noter que la contribution d'un groupe de n élevages n'est pas la somme des contributions des n élevages, mais lui est en général inférieure, en raison des migrations d'animaux entre ces élevages.

Pour illustrer la méthode de WIENER, considérons un pedigree réduit à trois générations d'ascendants (dans un but de simplification). A côté de chaque ascendant, figure le numéro de l'élevage dont il est issu (fig. 3). Calculons, à titre d'exemple, la contribution génétique de l'élevage 1. L'élevage 1 apparaît au niveau du père de l'animal A : dans la génération parentale, sa contribution est donc égale à $\frac{1}{2}$. Les deux apparitions de l'élevage 1 au niveau d'ascendants (grand-mère et arrière-grand-mère) du côté paternel ne sont pas prises en compte. D'autre part, l'élevage 1 apparaît deux fois dans la génération des arrière-grands-parents : dans cette génération, sa contribution est donc égale à $2 \times \frac{1}{8} = \frac{1}{4}$.

La contribution génétique totale de l'élevage 1 est égale à $\frac{1}{2} + \frac{1}{4} = \frac{3}{4}$. De même, on peut calculer les contributions des quatre autres élevages apparaissant dans le pedigree considéré (fig. 3).



<i>Elevage Herd (or flock)</i>	<i>Contribution au stock génétique de A Contribution to genetic make up of A</i>
1	$\frac{3}{4}$
2	$\frac{3}{8}$
3	$\frac{1}{2}$
4	$\frac{1}{4}$
5	$\frac{1}{4}$

FIG. 3

Calcul des contributions génétiques des élevages par la méthode de WIENER : un exemple (les numéros entre parenthèses désignent les élevages dont sont issus les divers ascendants).

Computation of genetic contributions of herds (or flocks) by WIENER's method : an example (the figures in brackets indicate the herds -or flocks- in which ancestors have been bred).

On remarque immédiatement que la somme des contributions génétiques des divers élevages n'est pas égale à 1 (dans l'exemple traité, elle est égale à 2,125). Ceci est, à notre avis, le point faible de la méthode de WIENER qui a pourtant connu un grand succès auprès de nombreux auteurs jusqu'à la période actuelle (tableau 2). En fait, la méthode de WIENER permet de déterminer le pourcentage des gènes d'un échantillon d'animaux qui

ont « transité » par un élevage donné. Un même gène transite, en général, par plusieurs élevages distincts. C'est pourquoi la somme des contributions génétiques, telles qu'elles sont définies par WIENER, n'est pas égale à 1. Selon la population considérée, cette somme peut prendre des valeurs très différentes : elle est d'autant plus élevée que la circulation des reproducteurs dans les élevages est plus intense. De ce fait, il est difficile de comparer des contributions d'élevages appartenant à des populations distinctes.

La méthode de WIENER (1953) pour évaluer la contribution génétique d'un élevage se distingue de celle de DICKSON et LUSH (1933) évoquée dans le paragraphe précédent. D'après la méthode de WIENER, la contribution d'un élevage représente la probabilité pour qu'un gène autosomal quelconque de l'échantillon traité ait transité par cet élevage au cours des quatre dernières générations. Le fait, pour un gène, d'avoir transité par un élevage donné n'est pas exclusif du fait d'avoir transité par un autre élevage : aussi la somme des contributions génétiques des divers élevages est-elle généralement supérieure à 1. D'après la méthode de DICKSON et LUSH, la contribution d'un élevage représente la probabilité pour qu'un gène autosomal quelconque de l'échantillon traité soit issu d'un fondateur appartenant à cet élevage. Les événements relatifs aux divers élevages ayant fourni des fondateurs constituent un système complet d'événements : aussi la somme des contributions génétiques des divers élevages est-elle égale à 1.

IV. Conclusion

L'analyse démographique des populations d'animaux domestiques a conduit à l'invention de quelques méthodes spécifiques : la méthode dite de « description numérique de la structure d'une population », proposée par ROBERTSON en 1953, en est un exemple. Mais la plupart des méthodes utilisées en démographie animale dérivent directement de méthodes classiques de démographie humaine.

Toutefois, la construction des tables n'est généralement pas fondée sur une analyse longitudinale nécessitant la reconstitution de l'histoire d'une cohorte, mais sur une série d'observations réalisées, pendant un laps de temps donné, sur un échantillon d'animaux appartenant à diverses cohortes : sans le mentionner explicitement, les auteurs établissent en fait des « tables du moment » (§ II, D, 2 et II, E, 1). L'imprécision relative à la définition des échantillons d'animaux se répercute sur le vocabulaire : on remarque, par exemple, une confusion fréquente entre les taux et les quotients (§ II, E, 2).

En outre, la validité de certains calculs est conditionnée par l'hypothèse sous-jacente de stationnarité : l'assimilation de la structure par âge à une table de survie, de l'âge moyen observé à une espérance de vie (§ II, D, 2) et de l'inverse du taux de renouvellement à la durée moyenne de la carrière reproductive (§ II, E, 2) deviennent des pratiques abusives en l'absence de stationnarité. Or, cette hypothèse est rarement vérifiée par les auteurs.

Si les auteurs ont fréquemment recours aux synthèses transversales que sont les tables du moment et à l'hypothèse implicite de stationnarité, c'est pour faire l'économie d'une véritable analyse longitudinale. Or l'analyse longitudinale est irremplaçable quand il s'agit de décrire l'évolution démographique d'une population, de détecter ses tendances et de faire des prévisions. Elle nécessite une observation continue des phénomènes démographiques pendant un laps de temps suffisamment long. Sa mise en œuvre dans les populations animales ne devrait pas se heurter à des difficultés majeures, car la durée de vie y est plus courte que dans les populations humaines.

En matière d'analyse génétique des populations, la situation est différente : de nombreux concepts et méthodes ont été inventés par des généticiens soucieux de décrire l'évolution du stock génétique de populations animales soumises à une sélection artificielle. C'est initialement afin d'analyser le patrimoine génétique de populations animales domestiques que WRIGHT a introduit des concepts et des méthodes qui sont devenus des éléments classiques et fondamentaux de la génétique de populations. Bien qu'elles aient été mises au point sur des populations d'animaux domestiques, ces méthodes présentent un intérêt plus général : leur champ d'application s'étend à toute population diploïde à reproduction sexuée, sans autofécondation.

Une méthode apparemment plus spécifique comme la méthode d'évaluation des contributions des divers élevages à la constitution du stock génétique d'une population animale (WIENER, 1953), répond en fait à une question plus large : dans une population structurée en groupes (définis selon un critère géographique ou autre), quelle est, à un instant donné, la proportion de gènes issus de chacun des groupes ? Bien que la méthode proposée par WIENER en 1953 ne soit pas entièrement satisfaisante (§ III, H), elle est encore très largement utilisée.

Du reste, il est frappant de constater que la plupart des méthodes utilisées pour analyser l'histoire génétique des populations d'animaux domestiques à partir de leurs pedigrees ont été mises au point et fixées il y a trente ans ou plus. Il semble pourtant que certaines de ces méthodes soient susceptibles d'être perfectionnées et que le champ des investigations dans ce domaine demeure ouvert. Ainsi, le concept de probabilité d'origine des gènes pourrait donner lieu à d'intéressants développements. La contribution d'un fondateur à la formation du patrimoine génétique d'un individu s'exprime par la probabilité pour qu'un gène autosomal désigné au hasard chez l'individu considéré soit la copie d'un gène appartenant à ce fondateur. Chaque individu de la population peut donc être caractérisé par un vecteur de n probabilités d'origine des gènes relatives aux n fondateurs de la population (la somme de ces probabilités étant égale à 1). Le calcul sur ordinateur de ces probabilités est simple et rapide. En outre, la matrice des probabilités ainsi obtenues se prête particulièrement bien à des traitements d'analyse des données tels que l'analyse factorielle des correspondances. Cette approche possible de la structure d'une population n'a encore guère été développée sur le plan théorique. Elle pourrait cependant permettre de décrire la structuration d'une population en lignées, prolongeant ainsi les recherches concernant la décomposition de la consanguinité (§ III, E, 2).

L'examen comparé des méthodes démographiques et génétiques conduit aux réflexions suivantes. En démographie, l'unité d'observation est l'individu, dont la durée de vie est bornée par sa naissance et sa mort. En génétique, l'unité de raisonnement est le gène, qui peut être transmis d'un parent à son enfant avec une probabilité égale à $\frac{1}{2}$. Les gènes

survivent aux individus, mais leur histoire est conditionnée par celle des individus qui les portent et les transmettent. La description de la structure génétique d'une population à un moment donné (en termes de probabilités d'origine des gènes, ou de coefficients de parenté et de consanguinité) constitue, en quelque sorte, un résumé de l'histoire démographique des générations qui ont précédé. Par exemple, il est aisé de mettre en évidence, par un calcul de probabilités d'origine des gènes, la forte contribution d'un fondateur à la formation du stock génétique d'une population. Un tel phénomène s'explique par le fait que ce fondateur a eu des descendants qui ont eux-mêmes procréé de nombreux produits etc. L'analyse démographique ne fournit pas de méthode simple permettant d'appréhender ce genre d'événement. En revanche, le calcul des contributions génétiques en donne une synthèse chiffrée.

Le concept de gène permet de réaliser une synthèse entre des caractéristiques démographiques observées sur des générations successives. Ainsi, les méthodes démographiques destinées à analyser la structuration d'une population en élevages trouvent un prolongement dans les méthodes génétiques évaluant les contributions des divers élevages. Les approches démographique et génétique ne sont pas interchangeables mais complémentaires, les résultats des calculs génétiques pouvant être interprétés à la lumière des caractéristiques démographiques.

Remerciements

L'auteur remercie les deux lecteurs mandatés par la revue ainsi que MM. C. CHEVALET, L. OLLIVIER et B. POUJARDIEU pour leurs critiques et suggestions lors de la lecture du manuscrit.

Reçu le 31 mars 1982.

Accepté le 17 janvier 1983.

Références bibliographiques

- ALLAIRE F.R., HENDERSON C.R., 1965. Inbreeding within an artificially bred dairy cattle population. *J. Dairy Sci.*, **48**, 1366-1371.
- ANDRUS D.F., FREEMAN A.E., EASTWOOD B.R., 1970. Age distribution and herd life expectancy in Iowa dairy herds. *J. Dairy Sci.*, **53**, 764-771.
- BARKER J.S.F., 1957 a. Expansion and distribution of the pedigree Jersey breed in Australia. *Aust. J. agric. Res.*, **8**, 547-560.
- BARKER J.S.F., 1957 b. The breed structure and genetic analysis of the pedigree cattle breeds in Australia. I. The Jersey. *Aust. J. agric. Res.*, **8**, 561-586.
- BARKER J.S.F., 1959. The breed structure and genetic analysis of the pedigree cattle breeds in Australia. I. The Jersey : a correction. *Aust. J. agric. Res.*, **10**, 769-770.
- BARKER J.S.F., DAVEY G. P., 1960. The breed structure and genetic analysis of the pedigree cattle breeds in Australia. II. The Poll Hereford. *Aust. J. agric. Res.*, **11**, 1072-1100.
- BEKEDAM M., 1959. Inbreeding in the Dutch Swine breeds. *Z. Tierz. Züchtungsbiol.*, **73**, 218-220.
- BETTINI T.M., 1968. *Elementi di demografia zootecnica*. Portici, Naples : Produzione Animale, 316 p.
- BETTINI T.M., MATASSINO D., 1961. Il ciclo di sostituzione dei genitori, la successione delle generazioni, e l'età media dei genitori alla nascita dei figli. *Genet. agrar.*, **13**, 321-358.
- BETTINI T.M., MATASSINO D., MAFFINI P., 1962. Indagine sulla struttura di un grande allevamento bovino da latte e sulla dinamica del suo equilibrio numerico in relazione ai fattori che influenzano la mortalità e gli scarti alle diverse età. *Ann. Fac. Sci. agrar. Univ. Stud.*, Napoli Portici. Serie Terza., Vol. XXVII, 13-118.
- RICHARD M., 1971. Dissemination of genetic improvement through a livestock industry. *Anim. Prod.*, **13**, 401-411.
- BOLDUAN G., 1963. Sind Inzucht und Verwandtschaft der Besamungsbullen angestiegen ? *Tierzucht*, **17**, 160-163.
- BOWMAN J.C., BUTLER E.A., TUNCEL E., 1978. Coefficients of inbreeding and degree of relationship for the British Friesian herd. *Anim. Prod.*, **27**, 269-276.
- BROCKELBANK E.E., WINTERS L.M., 1931. A study of the methods of breeding the best Shorthorns. *J. Hered.*, **22**, 245-249.
- BRZOWOSKI P., 1974. The age structure of cows in large herds of Polish Red- and- White Lowland cattle (polonais). *Acta agrar. silvestria, Ser.zootech*, **14**, 23-31.

- BUCHANAN SMITH A.D., ROBISON O.J., 1931. The average ages of cows and bulls in six breeds of cattle. *J. agric. Sci.*, **21**, 136-149.
- CALDER A., 1927. The role of inbreeding in the development of the Clydesdale breed of horses. *Proc. r. Soc., Edinburgh*, **47**, 118-140.
- CANNON C.Y., HANSEN E.N., 1939. Expectation of life in dairy cows. *J. Dairy Sci.*, **22**, 1025-1032.
- CARTER R.C., 1940. A genetic history of Hampshire sheep. *J. Hered.*, **31**, 89-93.
- CHAKRABORTY R., SHAW M., SCHULL W.J., 1974. Exclusion of paternity : the current state of the art. *Am. J. hum. Genet.*, **26**, 477-488.
- CILLIERS B., 1976. A generalisation of an approximate method to calculate inbreeding coefficients. *S. afr. J. anim. Sci.*, **6**, 191-193.
- CLAYTON G.A., 1956. Aspects of breed structure in pedigree British Shorthorn cattle. *Proc. bri. Soc. anim. Prod.* 1956, 107-110.
- COMFORT A., 1958. The longevity and mortality of Thoroughbred mares. *J. Gerontol.*, **13**, 342-350.
- COMFORT A., 1959. The longevity and mortality of Thoroughbred stallions. *J. Gerontol.*, **14**, 9-10.
- COMFORT A., 1962. A life table for Arabian mares. *J. Gerontol.*, **17**, 14.
- COOK R.D., HARTL D.L., 1976. Estimation of inbreeding by random walks in pedigrees. *Theor. appl. genet*, **48**, 171-178.
- COTTERMAN C.W., 1969. Factor-union phenotype system. In Computer Applications in Genetics, ed. MORTON N.E., University of Hawai Press, Honolulu, 1-18.
- COTTERMAN C.W., 1969. Factor-union phenotype system. In Computer Applications in Genetics, ed. MORTON N.E., University of Hawai Press, Honolulu, 1-18.
- DAVEY G., P., BARKER J.S.F., 1963. The breed structure and genetic analysis of the pedigree cattle breeds in Australia. III. The Hereford. *Aust. J. agric. Res.*, **14**, 93-118.
- DESIGNES A., 1968. Proposition de définitions de critères zootechniques en élevage ovin. *FEZ*, Dublin.
- DICKSON W.F., LUSH J.L., 1933. Inbreeding and the genetic history of the Rambouillet sheep in America. *J. Hered.*, **24**, 19-33.
- DONALD H.P., 1945. The growth and distribution of the pedigree Red Poll cattle population in England. *Emp. J. exp. Agric.* **13**, 169-183.
- DONALD H.P., AUERBACH C., 1942. The duration and extent of pedigree breeding herds of Large White pigs as a factor in breed improvement. *Emp. J. exp. Agric.*, **10**, 232-237.
- DONALD H.P., EL ITRIBY A.A., 1945. Herd size and its genetical significance in pedigree cattle breeding. *J. agric. Sci.*, **35**, 84-94.
- DONALD H.P., EL ITRIBY A.A., 1946. The duration of pedigree herds in three breeds of cattle in relation to selective breeding. *J. agric. Sci.*, **36**, 100-110.
- DONALD H.P., EL ITRIBY A.A., 1947. The age distribution of pedigree cattle and factors affecting it. *Emp. J. exp. Agric.*, **15**, 119-131.
- DRAGANESCU C., 1979. Some aspects of breeding structure of local breeds of sheep (roumain). *Lucr. Stiint. Inst. Agron.*, **22**, 15-18.
- ELSEN J.M., 1980. Diffusion du progrès génétique dans les populations avec générations imbriquées : quelques propriétés d'un modèle de prévision. *Ann. Génét. Sél. anim.*, **12**, 49-80.
- ELSEN J.M., MOCQUOT J.C., 1974. Méthode de prévision de l'évolution du niveau génétique d'une population soumise à une opération de sélection et dont les générations se chevauchent. *Bull. tech. Dép. Génét. anim. (I.N.R.A.)*, **17**, 30-54.
- EVANS B., 1976. The Coolah high fertility Merino flock. A co-operative group breeding scheme. In *Sheep Breeding*, Int. sheep breeding cong. Muresk and Perth, Butterworths, Australia (2nd ed.), 269-273.
- FALCONER D.S., 1960. *Introduction to Quantitative Genetics*. Oliver and Boyd, Londres, 365 p.
- FIEDLER H., HIMMEL V., 1967. Die Inzuchtverhältnisse beim Deutschen Fleckvieh, untersucht an den Besamungsbullen und Bullenmüttern. *Arch. Tierz.*, **10**, 25-29.
- FISTEAG I., LUNGULESCU G., NITA A., 1968. Productive potential of Romanian Simmental cattle (roumain). *Lucr. stiint. Inst. agron. Timosoara*, Ser. Med. vet., **11**, 523-532.
- FLADE D., 1965. Über genetische Veränderungen in der brandenburgischen Rinderzucht und die Verwendung ingezüchteter Bullen. *Arch. Tierz.*, **8**, 293-304.
- FLETCHER J.L., 1945. A genetic analysis of the American Quarter horse. *J. Hered.*, **36**, 346-352.

- FOGARTY N.M., 1978 a. Pedigree Dorset Horn sheep in Australia. I. Breed expansion and other vital statistics. *Aust. J. exp. Agric. anim. Husb.*, **18**, 165-172.
- FOGARTY N.M., 1978 b. Pedigree Dorset Horn sheep in Australia. II. Breed structure and genetic analysis. *Aust. J. exp. Agric. anim. Husb.*, **18**, 173-182.
- FOWLER A.B., 1932. The Ayrshire breed of cattle : a genetic study. *J. Dairy Res.*, **4**, 11-27.
- GIDEL R., 1972. Étude sur la composition moyenne des troupeaux de bovins de Haute-Volta et de Côte d'Ivoire en fonction de l'âge et du sexe. *Rev. Elev. Méd. vét. Pays trop.*, **25**, 453-550.
- HALDANE J.B.S., MOSHINSKY P., 1939. Inbreeding in mendelian populations with special reference to human cousin marriage. *Ann. Eugen.*, **9**, 321-340.
- HANSET R., 1973. Consanguinité et parenté chez le porc de Piétrain. *Ann. Génét. Sél. anim.*, **5**, 177-188.
- HERRON N.D., 1978. The breed structure and genetic analysis of the pedigree Sahiwal breed in Australia. *Aust. J. agric. Res.*, **29**, 381-393.
- HERRON N.D., PATTIE W.A., 1977 a. Studies of the Australian Illawarra Shorthorn breed of dairy cattle. I. Breed structure. *Aust. J. agric. Res.*, **28**, 1107-1117.
- HERRON N.D., PATTIE W.A., 1977 b. Studies of the Australian Illawarra Shorthorn breed of dairy cattle. II. Genetic analysis. *Aust. J. agric. Res.*, **28**, 1119-1132.
- HICKEY F., 1960. Death and reproductive rates of sheep in relation to flock culling and selection. *N. Z. J. agric. Res.*, **3**, 332-344.
- HILL W.G., 1974. Prediction and evaluation of response to selection with overlapping generations. *Anim. Prod.*, **18**, 117-139.
- HODGES J., TANNEN L., Mc GILLIVRAY B.J., HILEY P.G., ELLIS S., 1979. Inbreeding levels and their effect on milk fat and calving interval in Holstein-Friesian cows. *Can. J. anim. Sci.*, **59**, 153-158.
- IKIN T.L., HERRON N.D., 1978. Genetic analysis of the British Alpin goat herd in Australia. *Proc. aust. Soc. anim. Prod.*, **12**, 200.
- JACKSON N., TURNER H.N., 1972. Optimal structure for a co-operative nucleus breeding system. *Proc. aust. Soc. anim. Prod.*, **9**, 55-64.
- JAMES J.W., 1976. The theory behind breeding schemes. In *Sheep Breeding*, Int. sheep breeding cong. Muresk and Perth, Butterworths, Australia (2nd ed.), 205-213.
- JAMES J.W., 1977. Open nucleus breeding systems. *Anim. Prod.*, **24**, 287-305.
- JAMES J.W., Mc BRIDE G., 1958. The spread of genes by natural and artificial selection in a closed poultry flock. *J. Genet.*, **56**, 55-62.
- JOAKIMSEN Ö., 1969. Generation interval in Norwegian sheep. *Acta Agric. scand.*, **19**, 175-177.
- JONSSON P., 1971. Population parameters estimates of the Danish Landrace pig. *Acta Agric. scand.*, **21**, 11-16.
- KEANE M., 1977. Herd age and variation in cow disposals. *J. agric. Econ. rural Sociol.*, **6**, 213-220.
- KEMPTHORNE O., 1957. *An Introduction to Genetic Statistics*. John Wiley and Sons, New York, 545 p.
- KNOTHE A., 1976. Age structure of Polish Merino flocks (polonais). *Acta agrar. silv.*, Ser. zootech., **16**, 31-36.
- KOTZE F. de K., NEL J.W., SCHOEMAN S.J., 1977. The generation interval of the South African Mut-ton Merino (Afrikander). *S. afr. J. anim. Sci.*, **7**, 47.
- KUTSAL A., 1954. Untersuchungen über die Inzucht-und Verwandtschaftsverhältnisse und das Generationintervall beim schweizerischen Braunvieh. *Z. Tierz. Züchtungsbiol.*, **63**, 265-294.
- LANDAIS E., POIVEY J.P., SEITZ J.L., 1980. Recherches sur la reproduction du cheptel taurin sédentaire du Nord de la Côte-d'Ivoire : utilisation des intervalles entre vélages ; aspects méthodologiques et premiers résultats. *Rev. Elev. Méd. vét. Pays trop.*, **33**, 193-204.
- LANGANEY A., NADOT R., 1973. Détection automatique des anomalies de transmission génétique. *Ann. Genet.*, **16**, 247-253.
- LANGLET J., GRAVERT H.O., 1961. Inzucht-und Verwandtschaftsverhältnisse bei Bullen des schwarz-bunten Rindes in Schleswig-Holstein, die in der Künstlichen Besamung und im natürlichen Deckbetrieb eingesetzt wurden. *Züchtungskunde*, **33**, 373-379.
- LANGLOIS B., 1976. Estimation de quelques paramètres démographiques du pur sang anglais en France. *Ann. Génét. Sél. anim.*, **8**, 315-329.

- LAUVERGNE J.J., BOYAZOGLU J.G., CARTA R., CASU S., 1973. Caractéristiques démographiques de la race ovine sarde. *Ann. Génét. Sél. anim.*, **5**, 53-72.
- LEDERER J., BOGNER H., GOTTSCHALK A., AVERDUNK G., 1975. Inzuchtverhältnisse in der Fleckvieh-KB. Population und deren Auswirkung auf die Milchleistung. *Bayer. Landwirtschaft. Jahrb.*, **52**, 3-13.
- LEGAULT C., 1978. Analyse des composantes de la productivité numérique des truies. *Ann. Zootech.*, **27**, 457-470.
- LERNER I.M. HAZEL L.N., 1947. Population genetics of a poultry flock under artificial selection. *Genetics*, **32**, 325-339.
- LINDEMANN E., 1968. Linienzugehörigkeit, Inzucht- und Verwandtschaftsverhältnisse des Deutschen Schwarzbunten Rindes im Bezirk Rostock : ein Beitrag zur Linienzucht. *Arch. Tierz.*, **11**, 99-122.
- LINDSTRÖM U., MAIJALA K., 1971. Development of the coefficient of inbreeding and relationship in the Finnish Ayrshire breed. *Ann. Génét. Sél. anim.*, **3**, 101.
- LUSH J.L., 1932. The amount and kind of inbreeding which has occurred in the development of breeds of livestock. *Proc 6th Int. Genetics Cong.*, **2**, 123-126.
- LUSH J.L., 1946. Chance as a cause of changes in gene frequency within pure breeds of livestock. *Am. Nat.*, **80**, 318-342.
- LUSH J.L., ANDERSON A.L., 1939. A genetic history of Poland-China swine. I. Early breed history : the « Hot Blood » versus the « Big Type ». *J. Hered.*, **30**, 149-156.
- LUSH J.L., HOLBERT J.C., WILLHAM O.S., 1936. Genetic history of the Holstein-Friesian cattle in the United States. *J. Hered.*, **27**, 61-72.
- MAHON G.A.T., CUNNINGHAM E.P., 1981. Inbreeding and infertility in the Thoroughbred mare. *32nd Annual Meeting of the European Association for Animal Production*. Commission on Animal genetics.
- MALECOT G., 1948. *Les mathématiques de l'hérédité*. Masson, Paris.
- MARKUSHIN A.P., 1978. Replacement rate of generations in breeding flocks (russe). *Ovtsevodstvo*, **10**, 25-27.
- MARTIN I., 1975 a. A genetic analysis of the Galway sheep breed. 1. Some aspects of population dynamics of the pedigree and non-pedigree Galway sheep breed. *Ir. J. agric. Res.*, **14**, 245-253.
- MARTIN I., 1975 b. A genetic analysis of the Galway sheep breed. 2. The breed structure of the pedigree and non-pedigree Galway sheep breed. *Ir. J. agric. Res.*, **14**, 255-267.
- MARTIN I., 1975 c. A genetic analysis of the Galway sheep breed. 3. Level of inbreeding in the pedigree Galway sheep breed. *Ir. J. agric. Res.*, **14**, 269-274.
- MATASSINO D., BORDI A., NARDONE A., 1970. Alcune statistiche vitali nel coniglio nostrale migliorato. *Prod. anim.*, **9**, 21-73.
- MATASSINO D., LIBEROTTI G., GATTI L., 1965. Aspetti demografici di un allevamento bovino di razza Chianina. *Prod. anim.*, **4**, 333-378.
- MAYRHOFFER G., WOLLINGER F., 1980. Der Anstieg des Abstammungskoeffizienten nach Malécot im österreichischen Lipizzaner. *Wiener Tierärztliche Monatsschrift*, **67**, 243-244.
- Mc PHEE C.P., 1965. Inbreeding, migration and structure of the pedigree Large White pig population in Australia. *Aust. J. exp. Agric. anim. Husb.*, **5**, 270-278.
- Mc PHEE H.C., WRIGHT S., 1925. Mendelian analysis of the pure breeds of livestock. III. The Shorthorns. *J. Hered.*, **16**, 205-215.
- Mc PHEE H.C., WRIGHT S., 1926. Mendelian analysis of the pure breeds of livestock. IV. The British Dairy Shorthorns. *J. Hered.*, **17**, 397-401.
- NARDONE A., 1970. Cause di eliminazione e di morte nelle vacche controllate per la produzione lattea nel quadriennio. *Prod. anim.*, **9**, 1-20.
- NARDONE A., FABBRI G., RAMUNNO L., 1975. La mortalità e la eliminabilità nell' intervallo postparto in bovina da latte nel periodo 1969-1976. *Prod. anim.*, **14**, 111-144.
- O' CONNOR L.K., WILLIS M.B., 1967. The effect of artificial insemination on the breed structure of british Friesian cattle. *Anim. Prod.*, **9**, 287-293.
- ÖZKÜTÜK K., BICHARD M., 1977. Studies of pedigree Hereford cattle breeding. 1. Herdbook analyses. *Anim. Prod.*, **24**, 1-14.
- PEARL R., 1917. Studies on inbreeding. *Am. Nat.*, **51**, 545-559 ; **51**, 636-639.
- PETTIT G.N., 1940. Wastage, length of productive life, replacement and depreciation of dairy cows. *J. agric. Sci.*, **30**, 485-497.

- POIVEY J.P., LANDAIS E., SEITZ J.L., KOUYATE M., 1981. Détermination de l'âge des bovins par l'examen de la dentition. Méthodologie et principaux résultats acquis en milieu villageois dans le Nord de la Côte-d'Ivoire. *Rev. Elev. Méd. vét. Pays trop.*, **34**, 55-62.
- PRESSAT R., 1973. *L'analyse démographique*. P.U.F., Paris, 321 p.
- PRESSAT R., 1979. *Dictionnaire de démographie*. P.U.F., Paris, 295 p.
- RAE A.L., 1976. The development of co-operative breeding schemes in New Zealand. In *Sheep Breeding*, Int. sheep breeding cong. Muresk and Perth, Butterworths, Australia (2nd ed.), 275-285.
- ROBERTSON A., 1953. A numerical description of breed structure. *J. agric. Sci. (Cambridge)*, **43**, 334-336.
- ROBERTSON A., ASKER A.A., 1951 a. The genetic history and breed structure of British Friesian cattle. *Emp. J. exp. Agric.*, **19**, 113-130.
- ROBERTSON A., ASKER A.A., 1951 b. The expansion of a breed of dairy cattle. *Emp. J. exp. Agric.*, **19**, 295-301.
- ROBERTSON A., MASON I.L., 1954. A genetic analysis of the Red Danish breed of cattle. *Acta Agric. scand.*, **4**, 257-265.
- ROCHAMBEAU H. de, CHEVALET C., MALAFOSSE A., 1979. Le contrôle de la consanguinité dans les petites populations. *Bull. tech. Dép. Génét. anim.*, **31**, INRA, Paris, 122 p.
- ROTTENSTEN K., 1937. Inbreeding in Danish Landrace swine. *Nord. Jordbrugsforskning*, 3-4, 94-114.
- SCHOEMAN S.J., 1979. An analysis of the breed structure of the Karakul stud industry in South West Africa. *Agroanimalia*, **11**, 55-59.
- SCHWARK H.J., KITTNER M., 1966. Genealogischer Aufbau, Inzucht und Verwandtschaft bei schwarz-bunten Rindern im Bezirk Magdeburg : ein Beitrag zur Linienzucht. *Arch. Tierz.*, **9**, 107-129.
- SCIUCHETTI A., 1935. Beitrag zur genetischen Analyse der Schweizischen Braunrasse. Arch. Julius Klaus-Stiftung Vererbungsforsch. *Sozialanthropol. Rassenhyg.*, **10**, 85-99.
- SHEPHERD J.H., 1976. The Australian Merino society nucleus breeding scheme. In *Sheep Breeding*, Int. sheep breeding cong. Muresk and Perth, Butterworths, Australia (2nd ed.), 235-246.
- SMITH C., JORDAN C.H.C., STEANE D.E., SWEENEY M.B., 1978. A note on inbreeding and genetic relationship among British tested pigs. *Anim. Prod.*, **27**, 125-128.
- STEELE D., 1944. A genetic analysis of recent Thoroughbreds, Standardbreds and American Saddle horses. *Ky agric. Exp. Stn. Bull.* 462.
- STEWART A., 1952 a. Age distribution in New Zealand pedigree cattle. *Proc. N. Z. Soc. anim. Prod.*, **11**, 79-85.
- STEWART A., 1952 b. The expansion and structure of the New Zealand pedigree Jersey breed. *Proc. N. Z. Soc. anim. Prod.*, **12**, 32-39.
- STEWART A., 1954. Levels of inbreeding in New Zealand pedigree Jersey cattle. *Proc. N. Z. Soc. anim. Prod.*, **14**, 96-102.
- STEWART A., 1955. Expansion and structure of the New Zealand pedigree Ayrshire breed, 1910 to 1950. *N. Z. J. Sci. Technol.*, **36**, 493-505.
- STONAKER H.H., 1943. The breeding structure of the Aberdeen-Angus breed. *J. Hered.*, **34**, 322-328.
- SYRSTAD O., 1957. Inbreeding and relationship within Telemark cattle and Red Polled cattle of Norway. *Meld. Nor. Landbrukshoegsk.*, **36**, 1-11.
- TACHER G., 1975 a. Notes sur la démographie bovine au Sahel. I. Représentation et projection théoriques du troupeau par un modèle mathématique. *Rev. Elev. Méd. vét. Pays trop.*, **28**, 547-569.
- TACHER G., 1975 b. Notes sur la démographie bovine au Sahel. II. Représentation et projection théoriques du troupeau par un modèle mathématique après une catastrophe. *Rev. Elev. Méd. vét. Pays trop.*, **28**, 571-595.
- TOLL G.L., BARKER J.S.F., 1979. Genetic history of the pedigree Poll Hereford breed in Australia : effect of the importation ban. *Aust. J. agric. Res.*, **30**, 767-777.
- TURNER H.N., DOLLING C.H.S., SHEAFFE P.H.G., 1959. Vital statistics for an experimental flock of Merino sheep. I. Death rates in adult sheep, in relation to method of selection, age and sex. *Aust. J. agric. Res.*, **10**, 581-590.
- TURNER H.N., DOLLING C.H.S., 1965. Vital statistics for an experimental flock of Merino sheep. II. The influence of age on reproductive performance. *Aust. J. agric. Res.*, **16**, 699-712.
- VERA A., APARICIO F., ECHEVARRIA A., 1978. Estudio de la longevidad y la duración de la vida útil en rebaños de ovejas navarras. *Arch. Zootec.*, **27**, 133-143.
- WAHLUND S., 1928. Zusammensetzung von Population und Korrelationserscheinung vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*, **11**, 65-105.

- WATSON J.H., 1963. Inbreeding arising through artificial insemination and natural service in Welsh Black cattle. *Anim. Prod.*, **5**, 209-214.
- WIENER G., 1953. Breed structure in the pedigree Ayrshire cattle population in Great Britain. *J. agric. Sci. (Cambridge)*, **43**, 123-130.
- WIENER G., 1954. Flock size and flock duration of lowland breeds of sheep. *Proc. br. Soc. anim. Prod.*, 1954, 59-62.
- WIENER G., 1955. Migration of pedigree Ayrshire cattle in Great Britain. *J. agric. Sci.*, **45**, 476-480.
- WIENER G., 1961. Population dynamics in 14 lowland breeds of sheep in Great Britain. *J. agric. Sci.*, **57**, 21-28.
- WIENER G., YAO T.S., 1952. Growth of the pedigree Ayrshire cattle population in Great Britain. *Emp. J. exp. Agric.*, **20**, 195-208.
- WILLHAM O.S., 1937. A genetic history of Hereford cattle in the United States. *J. Hered.*, **28**, 283-294.
- WRIGHT S., 1922. Coefficients of inbreeding and relationship. *Am. Nat.*, **56**, 330-338.
- WRIGHT S., 1923 a. Mendelian analysis of the pure breeds of livestock. I. The measurement of inbreeding and relationship. *J. Hered.*, **14**, 339-348.
- WRIGHT S., 1923 b. Mendelian analysis of the pure breeds of livestock. II. The Duchess family of Shorthorns as bred by Thomas Bates. *J. Hered.*, **14**, 405-422.
- WRIGHT S., 1931. Evolution in mendelian populations. *Genetics*, **16**, 97-159.
- WRIGHT S., 1969. Evolution and the genetics of populations. Vol. 2. *The theory of gene frequencies*. The University of Chicago Press, Chicago and London, p. 294-295.
- WRIGHT S. 1977. Evolution and the genetics of populations. Vol. 3. *Experimental results and evolutionary deductions*. The University of Chicago Press, Chicago and London, p. 538-555.
- WRIGHT S., Mc PHEE H.C., 1925. An approximate method of calculating coefficients of inbreeding and relationship from livestock pedigrees. *J. agric. Sci. (Cambridge)*, **31**, 377-383.
- YODER D.M., LUSH J.L., 1937. A genetic history of the Brown Swiss cattle in the United States. *J. Hered.*, **28**, 154-160.
- YOUNG G.B., PURSER A.F., 1962. Breed structure and genetic analysis of Border Leicester sheep. *Anim. Prod.*, **4**, 379-389.
- YULE G.U., 1944. *A statistical study of literary vocabulary*. Cambridge University Press.