Genet. Res., Camb. (1977), **29**, *pp.* 239–253
Printed in Great Britain

Genetic distance between the Australian Merino and the Poll Dorset Sheep

BY CLYDE MANWELL AND C. M. ANN BAKER
Department of Zoology, University of Adelaide, South Australia, 5001

(*Received* 17 *January* 1977)

SUMMARY

Merinos (189) and Poll Dorsets (106) were compared for genetic variation at 30 loci for blood proteins. 'Malic enzyme' and NADH diaphorase I polymorphism occur in Merinos but not in Poll Dorsets, whereas both breeds are polymorphic for haemoglobin, erythrocyte X-protein, serum esterase, catalase and transferrin, although the breeds differ in the presence or absence of certain rare transferrin variants. Poll Dorsets but not Merinos have genetic variation of erythrocyte pyruvate kinase at low gene frequencies; Merinos but not Poll Dorsets have genetic variation of two erythrocyte Gly-Leu peptidases, glucosephosphate isomerase, superoxide dismutase, and NADH diaphorase II, all at low gene frequencies ($P < 0{\cdot}05$).

Using Masatoshi Nei's (1972, 1976) standard genetic distance, over all 30 loci, to calculate the time of divergence of the two breeds, we obtained $t = 69700$ years – whereas sheep are believed to have been domesticated by man for not more than 11000 years, and ancestors of the British breeds and of the Merino have probably not been separated for more than 2000 years. Several possible explanations for this discrepancy are discussed, including the accuracy of the coefficient of codon change (α), changes in population size and structure, selection, and additional hybridization in the ancestry of the Merino.

1. INTRODUCTION

One of the major accomplishments of theoretical population genetics is the development of methods for the estimation of the degree of divergence between populations. In particular, the method of Nei (1972, 1975, 1976) has a linear relationship between the *standard genetic distance*, D, and the *time of divergence*, t:

$$D = -\log_e \frac{J_{XY}}{(J_X J_Y)^{\frac{1}{2}}} = 2\alpha t,$$

where α is the rate of codon change (gene substitution) per locus per year. J_{XY} is the probability of gene identity between the two populations X and Y, i.e. the probability that, for each locus, if one gene is sampled from the gene pool of population X, and another gene is sampled from the gene pool of population Y, then the two genes will be identical,

$$J_{XY} = \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{k} p_{Xij}\, p_{Yji},$$

with from 1 to k alleles at each of n loci, p being the gene frequency. J_X is the probability of gene identity within population X, i.e. the probability of homozygosity,

$$J_X = \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{k} (p_{Xij})^2.$$

and J_Y is the same for population Y.

Although the choice of measures of genetic distance is still a controversial topic in genetic theory, Nei's model has a number of advantages besides linearity with respect to time, e.g. the model is robust to variations in the size of the diverging population, and the coefficient of codon change, α, has been calculated independently, using data on amino acid substitutions in proteins and, thus, the model can be used both at low levels of genetic divergence, i.e. between populations within a species, and at large levels of genetic divergence, i.e. higher taxonomic categories (Nei, 1975, 1976).

Though the model was originally derived under the assumption of neutrality, the expression $2\alpha t$ can be changed to accommodate simple patterns of selection (Nei, 1975, 1976). Although it remains to be determined just how sensitive the model is to different types of selection at different loci, it is reasonable to assume that the effects of weak and variable selection would tend to cancel out.

For example, Nei's genetic distance statistic has given reasonable estimates for the time of divergence of the three major races of man (Nei & Roychoudhury, 1974) although it is believed that some of the polymorphisms used to calculate the genetic distance statistic show reduced variances in gene frequency as a consequence of selection (Lewontin & Krakauer (1973); see, however, Ewens & Feldmann (1976) for a critique on Lewontin and Krakauer's analysis).

The use of Nei's standard genetic distance to estimate the times of separation of some species of *Drosophila* on five of the Hawaiian Islands results in similar dates to those arrived independently from geological information and allows the 'cautious suggestion' (Carson, 1976) that: 'accumulation of electrophoretically detected protein differences over geologic time may possibly serve as a useful clock mechanism, permitting details of the past evolutionary history of a species to be inferred from the biochemical state of the living forms'.

The limitations in studies on genetic distance may be more experimental than theoretical. Where data from a large number of different workers are used, such as in the comparison of the three major races of man, it is inevitable that different proteins have been subjected to different intensities of study. For example, inspection of tables 6.1 and 6.5 in Nei (1975) reveals that within the most surveyed race, Caucasoid, the proportion of polymorphic loci and the average heterozygosity increases in the loci studied in both Caucasoids and Negroids, and increases still further in the loci studied in all three races.

Furthermore, when only small numbers of protein loci have been studied, it has been necessary to include data on blood groups as well. For example, Ananthakrishnan's (1973) pioneering use of genetic-distance estimates on some sheep breeds used literature data on only four proteins, all reported in the literature

because they were polymorphic, and five blood-group antigens. Ananthakrishnan (1973) was properly sceptical of the results when the closest genetic distance was between the Southdown and the Navajo – two very different breeds of sheep. Although examples of strong association between protein variants and blood groups are known (e.g. for sheep: Rendel & Stormont 1964; Hope 1966), some blood-group antigens are complex carbohydrate moieties and the genetic determination is mediated through changes in the substrate specificity, or the presence or absence, of glycosyl transferases and other enzymes. That caution is necessary in pooling data on protein polymorphisms and on blood groups is suggested by the fact that the gene frequency distribution of blood group loci is significantly different from that for electrophoretic protein loci (Yamazaki & Maruyama, 1974).

Studies on domesticated species are of value in that there are independent estimates of the time of domestication and of breed formation (e.g. for sheep: Ryder, 1964, 1973) and there are data on the type of selection to which man has exposed them.

2. MATERIALS AND METHODS

(i) *Sheep*

Merinos were from the Roseworthy Agricultural College's Index and Visual Selection flocks, derived basically from the 'Bungaree' strain of Merino and representing a high quality of South Australian 'strong wool' Merino. A total of 189 Merinos were sampled, 146 from the Index flock and 32 from the Visual flock; and, despite limitations imposed by the availability of certain reagents, most of these animals were surveyed for most of the 30 loci, details being placed in Table 1. One advantage of the choice of these Roseworthy Merinos is that the flocks were also sampled close to 10 years ago for two of the polymorphic protein loci, transferrin and haemoglobin (Mayo *et al.* 1970), allowing a check on gene frequencies.

The pure-bred Poll Dorsets are also kept at the Roseworthy Agricultural College and this means that the comparison of the two breeds is done with animals kept under the same general environment in recent times. However, the Poll Dorset flock is not entirely closed as on occasion rams from other Poll Dorset flocks have been added to the Roseworthy flock.

(ii) *Electrophoresis*

Erythrocytes were washed three times with 0·9 % NaCl, equilibrated with carbon monoxide, and lysed with the 1 % digitonin, pH 8, glycyl–glycine buffer as described earlier (Baker & Manwell, 1976*a*); in general, 1 vol. of cells was lysed with 2 vols. of extractant, and freezing and thawing were avoided.

The electrophoretic procedures were the same as those used on organ extracts (Baker & Manwell, 1976*a*). It was found that the acrylamide polymer gradient gels, 'Gradipore', were especially useful in studying erythrocytes, for the low levels of certain enzymes still allowed staining of zones within 1 h, except for phosphoglucomutase, which required overnight incubation. Two special enzyme localization procedures which give satisfactory results on sheep erythrocyte extracts are detailed:

Table 1. *Comparison of 30 protein loci in Merinos and Poll Dorset sheep*

Protein (Enzyme Commission no. where appropriate)	Merino (no. of sheep surveyed)	Poll Dorset (no. of sheep surveyed)	Statistical significance of breed difference (see Materials and Methods)
Lactate dehydrogenase (EC 1.1.1.27), H chain	Monomorphic (140)	Monomorphic (102)	—
Malate dehydrogenase (EC 1.1.1.37), supernatant	Monomorphic (179)	Monomorphic (102)	—
'Malic enzyme' (EC 1.1.1.40); major isozyme	Polymorphic (154) $p_F = 0{\cdot}662$ $p_S = 0{\cdot}338$	Monomorphic (102) $p_F = 1{\cdot}00$ $p_S = 0$	*** $P = 2{\cdot}0 \times 10^{-24}$
Glucose-6-phosphate dehydrogenase (EC 1.1.1.49)	Monomorphic (155)	Monomorphic (102)	—
NADH diaphorase I of r.b.c.s (EC 1.6.4.3 or 1.6.99.3)	Polymorphic (143) $p_F = 0{\cdot}766$ $p_S = 0{\cdot}234$	Monomorphic (102) $p_F = 1{\cdot}00$ $p_S = 0$	*** $P = 1{\cdot}6 \times 10^{-16}$
NADH diaphorase II of r.b.c.s (EC 1.6.4.3 or 1.6.99.3)	'Monomorphic' (143) $p_F = 0{\cdot}997$ $p_S = 0{\cdot}003$	Monomorphic (102) $p_F = 1{\cdot}00$ $p_S = 0$	N.S. $P = 1$
Catalase (EC 1.11.1.6)	Polymorphic (117) $p_F = 0{\cdot}483$ $p_S = 0{\cdot}517$	Polymorphic (82) $p_F = 0{\cdot}598$ $p_S = 0{\cdot}402$	* $\chi^2_{(1)} = 4{\cdot}64P \sim 0{\cdot}03$
Superoxide dismutase (EC 1.15.1.1)	'Monomorphic' (149) $p_A = 0{\cdot}003$ $p_C = 0{\cdot}997$	Monomorphic (103) $p_A = 0$ $p_C = 1{\cdot}00$	N.S. $P = 1$
Pyruvate kinase (EC 2.7.1.40)	Monomorphic (86) $p_F = 0$ $p_S = 1{\cdot}00$	'Monomorphic' (66) $p_F = 0{\cdot}015$ $p_S = 0{\cdot}985$	N.S. $P = 0{\cdot}51$
Phosphoglucomutase I (EC 2.7.5.1)	Monomorphic (77)	Monomorphic (23)	—
Lee's (1966) H esterase of serum (EC 3.1.1.2?), H, active on α-naphthyl acetate; h is not	Polymorphic (189) $p_H = 0{\cdot}152$ $p_h = 0{\cdot}848$	Polymorphic (106) $p_H = 0{\cdot}425$ $p_h = 0{\cdot}575$	*** $\chi^2_{(1)} = 40{\cdot}7P < 0{\cdot}001$
Serum cholinesterase (EC 3.1.1.8); active only on acetylthiocholine	Monomorphic (140)	Monomorphic (106)	—
Acid phosphatase (EC 3.1.3.2)	Monomorphic (119)	Monomorphic (80)	—
Gly-Leu peptidase of r.b.c.s (EC 3.4.3.2?), major (slow) locus	'Monomorphic' (135) $p_F = 0{\cdot}952$ $p_S = 0{\cdot}048$	Monomorphic (104) $p_F = 1{\cdot}00$ $p_S = 0$	*** $P = 0{\cdot}00062$
Gly-Leu peptidase of r.b.c.s (EC 3.4.3.2?), trace (fast) locus	'Monomorphic' (135) $p_F = 0{\cdot}004$ $p_S = 0{\cdot}996$	Monomorphic (104) $p_F = 0$ $p_S = 1{\cdot}00$	N.S. $P = 1$
Gly-Leu peptidase of serum (not an arylamidase) (EC 3.4.3.2?)	Monomorphic (52)	Monomorphic (28)	—

Table 1 (*cont.*)

Protein (Enzyme Commission no. where appropriate)	Merino (no. of sheep surveyed)	Poll Dorset (no. of sheep surveyed)	Statistical significance of breed difference (see Materials and Methods)
Carbonic anhydrase II (EC 4.2.1.1)	Monomorphic (152) (all 'slow')	Monomorphic (106) (all 'slow')	—
Glucosephosphate isomerase (EC 5.3.1.9)	'Monomorphic' (158) $p_A = 0{\cdot}006$ $p_C = 0{\cdot}994$	Monomorphic (104) $p_A = 0$ $p_C = 1{\cdot}00$	N.S. $P = 0{\cdot}3$
Haemoglobin, α chain locus	Monomorphic (188)	Monomorphic (106)	—
Haemoglobin, β chain locus	Polymorphic (188) $p_A = 0{\cdot}371$ $p_B = 0{\cdot}629$	Polymorphic (106) $p_A = 0{\cdot}198$ $p_B = 0{\cdot}802$	*** $\chi^2_{(1)} = 17{\cdot}8 P < 0{\cdot}001$
Tucker *et al.*'s (1967) erythrocyte X-protein	Polymorphic (164) $p_+ = 0{\cdot}426$ $p_0 = 0{\cdot}574$	Polymorphic (102) $p_+ = 0{\cdot}131$ $p_0 = 0{\cdot}869$	*** $\chi^2_{(1)} = 37{\cdot}3 P < 0{\cdot}001$
Erythrocyte 'Y-protein' (migrates anodally to X)	Variable (164) (see text)	Variable (102)	—
Erythrocyte protein (migrates near 2nd zone of X)	Monomorphic (164)	Monomorphic (102)	—
Serum prealbumin	Variable (178)	Variable (106)	—
Serum albumin	Monomorphic (178)	Monomorphic (106)	—
Serum postalbumin	Variable (178)	Monomorphic (106)	—
Caeruloplasmin [*o*-dianisidine ferro-oxidase] (recently assigned EC 1.16.3.1)	Monomorphic (178)	Monomorphic (106)	—
Unidentified major serum protein, migrating between transferrin and slow α_2	Variable (178)	Variable (106)	
Slow α_2 macroglobulin	Monomorphic (178)	Monomorphic (106)	—
Transferrin	Polymorphic (178) $p_A = 0{\cdot}219$ $p_A^- = 0$ $p_B = 0{\cdot}233$ $p_C = 0{\cdot}006$ $p_D = 0{\cdot}506$ $p_D^- = 0{\cdot}028$ $p_E = 0{\cdot}008$	Polymorphic (106) $p_A = 0{\cdot}118$ $p_A^- = 0{\cdot}009$ $p_B = 0{\cdot}278$ $p_C = 0{\cdot}080$ $p_D = 0{\cdot}514$ $p_D^- = 0$ $p_E = 0$	*** $\chi^2_{(1)} = 8{\cdot}5 P \sim 0{\cdot}006$ $P = 0{\cdot}53$ $\chi^2_{(1)} = 1{\cdot}21 P \sim 0{\cdot}28$ $P = 0{\cdot}000036$ $\chi^2_{(1)} = 0{\cdot}012 P \sim 0{\cdot}9$ $P = 0{\cdot}0094$ $P = 0{\cdot}3$

(iii) *Carbonic anhydrase*

Previous workers have usually studied this enzyme either by detecting the change in pH which occurs upon hydration of carbon dioxide and dissociation into H^+ and HCO_3^-, by detecting a minor protein zone after staining with Amido Black 10B or other total protein stain, or by using its weak esterase activity on naphthyl acetate esters (e.g. Tucker, Suzuki & Stormont, 1967). We have found

that a modification of the conventional photometric assay method using hydrolysis of *p*-nitrophenyl acetate is quick and reliable: 10 mg of *p*-nitrophenyl acetate (Sigma Chemical Company) are dissolved in 0·3 ml of acetone and added to 5 ml of a pH 6·0 potassium phosphate buffer, $\mu = 0{\cdot}01$. The gel is first rinsed in 50 ml of this pH 6 buffer for 15 min and blotted dry; then the aqueous solution of *p*-nitrophenyl acetate is soaked into a filter-paper overlay and placed on top of the sliced gel. Usually, yellow zones of carbonic anhydrase activity show up in 15 min. On occasion it is necessary to wait 30 min and then, if no zones are visible, to raise the pH to 8 or so by rinsing the gel with 10 ml of the concentrated Aronsson-Grönwall buffer described previously (Baker & Manwell, 1976*a*) in order to increase the proportion of the yellow-coloured *p*-nitrophenolate ion. Although the initial incubation can be made at pH 7, in our experience the faster development of yellow zones by carbonic anhydrase at this pH is offset by the instability of the reagent, which gives a yellow background. With the correct change in pH, the result is bright yellow zones on a white or very pale yellow background.

The detection of carbonic anhydrase by hydrolysis of *p*-nitrophenyl acetate was checked by two independent methods: yellow zones did not appear when the sliced gel was first equilibrated with the specific carbonic anhydrase inhibitor Diamox (= acetazolamide) at 10^{-4} M; and both the two zones in human erythrocytes and the single zone (carbonic anhydrase II) in sheep erythrocytes were shown to have molecular weights in the range 32000–34000 by retardation electrophoresis (Manwell, 1977).

Activity of carbonic anhydrase is much greater on *p*-nitrophenyl acetate than on α- or β-naphthyl acetate; however, none of these compounds are absolutely specific for this enzyme, for a number of esterases will hydrolyse all three of these esters, e.g. certain of the liver and kidney esterases in sheep. This is not a problem in the present study for these other esterases do not occur in erythrocytes and they are not inhibited by acetazolamide. In studies on erythrocyte extracts the *p*-nitrophenyl acetate method has the advantage that it will detect carbonic anhydrase variants that overlap haemoglobin, variants that would be missed in total protein staining. One minor disadvantage of the *p*-nitrophenyl acetate technique is that the *p*-nitrophenolate ion does not form a precipitate and thus the yellow zones, originally very sharp and only 2–3 mm wide, become diffuse after 1–2 h.

(iv) *Acid phosphatase*

We found the 4-methylumbelliferyl phosphate method recommended by McDermid, Agar & Chai (1975) useful and, in agreement with those workers, we found no genetic variation in sheep. We did observe a spurious 'polymorphism' when older samples were used because of the development at different rates of a faster migrating satellite zone. Using 2 mg of 4-methylumbelliferyl phosphate (Sigma) dissolved in 5 ml of pH 5·7 $\mu = 0{\cdot}03$ sodium acetate buffer to which 0·2 ml of 1 M-$MgCl_2$ was added, the sliced gel is covered with a double piece of filter paper soaked in this solution. For gels at alkaline pH's at 15 min preincubation

with only dilute pH 5·7 acetate buffer is necessary. When viewed with long-wavelength ultraviolet light, bright zones of bluish fluorescence denote acid phosphatase. For erythrocyte extracts the zones show up within 1 h; if not, the gel is rinsed quickly with 10 ml of pH 8·7 Aronsson and Grönwall buffer, which enhances the fluorescence.

(v) *Analysis of data*

Although some data in the literature on protein electrophoresis are presented in a dichotomous classification, monomorphic versus polymorphic, it is necessary to recognize that occasionally one observes variation which is not readily genetically interpretable; such systems are placed into a third category, 'variable' (Manwell & Baker, 1975). 'Variable' systems are fairly common among certain serum proteins, e.g. serum prealbumins, and some esterases, where the variation is physiological rather than genetic. However, it is possible that some 'variable' systems represent the subtle effects of a 'hidden' genetic polymorphism involving the substitution of similar amino acids so that the primary structures do not differ by an ionizable group. Such 'variable' cases should be lumped with the monomorphic loci for genetic distance studies as α is determined using amino acid sequence data representing only mutations which result in a net integral difference in charge on the protein and, thus, α is calibrated for the easily separated electrophoretic variants (Nei, 1972, 1975).

Each difference in allele frequency between the two breeds was tested for statistical significance; the contingency χ^2 was calculated when none of the cells of data had expectations fewer than 5; otherwise R. A. Fisher's exact test was used (Bailey, 1959).

3. RESULTS

Data on the genetic variability of 30 protein loci in Merinos and in Poll Dorsets are summarized in Table 1.

The most striking difference between Merinos and Poll Dorsets is that the former have polymorphism with high gene frequencies for the major 'malic enzyme' [NADP-dependent malate dehydrogenase (decarboxylating)] and the NADH diaphorase I of erythrocytes, whereas the latter are monomorphic at both these loci.

In addition, Merinos but not Poll Dorsets have low-frequency variants for the minor and major erythrocyte peptidases using Gly-Leu as a substrate, for NADH diaphorase II, glucosephosphate isomerase and superoxide dismutase, although only the difference involving the major Gly-Leu peptidase was statistically significant. Merinos with the rare variants of NADH diaphorase II, glucosephosphate isomerase and superoxide dismutase had three major zones in each case; the pattern of a single hybrid zone in heterozygotes is typical for the last two enzymes (Manwell & Baker, 1970). There is a lack of published data on the pattern of variation of NADH diaphorases, especially the more slowly migrating forms of higher molecular weight; however, a similar NADH diaphorase in erythrocytes of sipunculid worms has a pattern of genetic variation of three zones in

heterozygotes (Manwell, unpublished). The fast-moving diaphorase I of sheep erythrocytes has a low molecular weight (25000–31000) and the absence of hybrid zones in heterozygotes for this diaphorase is in agreement with Brewer, Eaton, Knutsen & Beck (1967).

Poll Dorsets, but not Merinos, have low-frequency variation of erythrocyte pyruvate kinase. As the enzyme is localized by a negative method, the removal of UV-induced NADH fluorescence, it is difficult to ascertain whether or not there is a hybrid zone.

Both breeds share, though with different gene frequencies, polymorphism for haemoglobin, the erythrocyte X-protein of Tucker *et al.* (1967), the haloxon differentiable major serum esterase of Lee (1964), transferrin and catalase.

There are differences between Merinos and Poll Dorsets in the presence or absence of certain rare transferrin types, including two variants similar to but distinct from Tf A and Tf D, temporarily termed Tf A^- and Tf D^- respectively. There are many rare transferrin alleles in sheep (e.g. Fésüs & Rasmussen, 1971) and, since serum samples cannot be brought into Australia (quarantine laws), we do not know if Tf D^- and Tf A^- are among the rare variants found by others.

Although some sheep are polymorphic for serum albumin (e.g. Stratil & Glasnák, 1974) the Roseworthy Merinos and Poll Dorsets were monomorphic for serum albumin in all buffer systems. Similarly, we did not find the low-frequency carbonic anhydrase F variant described by Tucker *et al.* (1967); presumably the sheep surveyed by us are monomorphic for the S variant. We observed variability of the serum pre-albumins in several starch-gel electrophoretic systems but we could not, using the system recommended by Efremov, Vaskov & Hrisoho (1970), obtain genetically interpretable results; however, we did find, too late to use in the survey, that marked prealbumin variation can be detected when sheep sera are electrophoresed in pH 4·7, 0·01 ionic-strength sodium acetate buffer, using acrylamide polymer gradient gels, providing they are pre-electrophoresed for 24 h to bring the pH sufficiently close to 4·7.

4. DISCUSSION

(i) *Comparison of our data with earlier studies on the Roseworthy Merinos*

Combining the data on the low-frequency variant Tf D^- with Tf D, it is remarkable how similar the gene frequencies are for the transferrin polymorphism in the Roseworthy Merinos surveyed approximately 10 years ago by Mayo *et al.* (1970) and by us in 1975 (see Table 2).

For haemoglobin, Mayo and colleagues (1970) reported a gene frequency for Hb A of 0·256 in the Roseworthy Visual Merino flock and of 0·333 in the Roseworthy Index Merino flock; we observed slightly higher frequencies of Hb A in both flocks – 0·281 for 32 sheep from the Visual flock and 0·390 for 146 from the Index flock. However, even the larger difference between ourselves and Mayo *et al.* (1970) is not statistically significant, $\chi^2 = 2{\cdot}66$ (1 D.F.). In the haemoglobin

studies we observed one individual with nearly 50% Hb C; when that individual was bled again 3 months later no Hb C was found; as the first time the individual had a very active glucose-6-phosphate dehydrogenase zone, it is reasonable to suppose that it had switched haemoglobin synthesis in response to increased erythropoiesis and its true haemoglobin type was Hb A homozygous; only one other individual had a trace of Hb C and that was in a Hb AB heterozygote. Mayo and colleagues (1970) did not report any Hb C.

A further survey of the transferrins and haemoglobins of the Roseworthy Merinos has also shown constancy of the transferrins and a slight increase in Hb A (Mr Tom Mann, personal communication). It is thus clear that for the two polymorphisms where 'historical' data are available that neither sheep, nor shepherd, nor electrophoresis have changed more than slightly.

(ii) *Genetic distance and the time of divergence for the Merino and the Poll Dorset*

If all differences in gene frequency for the 30 loci are used, J_X for the Merino is 0·8906, J_Y for the Poll Dorset is 0·9267, and J_{XY} between the two breeds is 0·8959, yielding a standard genetic distance $D = 0{\cdot}01393$, in turn corresponding to a time of divergence of 69700 years, using $\alpha = 10^{-7}$ electrophoretically detectable codon changes per locus per year.

If only the statistically significant breed differences are used (Table 1), D is 0·01389, corresponding to 69500 years. Thus, at this level of divergence the few statistically insignificant rare variants have only a marginal effect on the results.

As archaeological estimates indicate that the domestication of sheep took place at the earliest 10870 ± 300 years ago, and possibly later (Bökönyi, 1976), and as the time of divergence of the ancestors of the fine-wooled Merino and the various modern British breeds took place closer to 2000 years ago (Ryder, 1973), there is a discrepancy to be explained.

(iii) *Accuracy of the calculation of the time of genetic divergence?*

The present data are among the better data available for calculations of genetic distance in that: (1) Approximately similar electrophoretic effort was devoted to each protein. (2) We have satisfied Nei's (1975, 1976) suggestion that, as a consequence of the large variance of gene frequencies for alleles at polymorphic loci, approximately 30 loci should be examined. (3) The number of individuals surveyed from each population is greater than the minimal number of 30 suggested by Nei (1975, 1976). (4) Most of the differences in gene frequency between the two breeds are highly statistically significant (Table 1) and, thus, the estimate of the time of divergence is not appreciably altered when only the statistically significant differences are used in calculating genetic distance. (5) Gene frequencies for two of the polymorphic protein systems studied earlier by Mayo and colleagues (1970) on the same flocks of Roseworthy Merinos are in agreement with our independent sample. Thus, the measure of genetic distance itself is reasonably accurate.

However, estimation of the time of divergence depends also upon the coefficient of codon change, α, for electrophoretically detectable protein polymorphisms. Here there is of necessity greater uncertainty, for α has been calculated using the not inconsiderable but still limited data on amino acid sequences of proteins. As Nei (1976, p. 758) points out, the magnitude of the variance of α is 'not known at the present time'. As the better-known phylogenetic assemblages of homologous proteins, e.g. the haemoglobins, the pyridine-adenine dinucleotide-linked dehydrogenases, and the peptidases, better known at the level of primary structure, are included in the present study, and as some of the primary structure data has been obtained on proteins from sheep or related artiodactyls, α is partly standardized for studies on sheep or other mammals. To explain the large time of divergence between the two breeds of sheep purely on the basis of uncertainty in α requires postulating a true α some 30×10^{-7}. It is unlikely that the error could be anywhere near that large and, as mentioned in the introduction of this paper, $\alpha = 10^{-7}$ has been used in other calculations of the time of divergence which are in reasonable agreement with independent estimates.

(iv) *Changes in population size and breeding structure?*

Although Nei's formula was derived under the assumption of constant population size, it has subsequently been shown to be 'quite robust' to founder effect and changes in population size, even fluctuations of two orders of magnitude (Nei, 1975, 1976). Genetic distance, however, is amplified if the population passes through a tight bottleneck, but this requires extremely small breeding numbers, e.g. 2–10, for several generations (Chakraborty & Nei, 1977). If over a long period of time the population size increases there will be a decrease in the rate of gene substitution by drift and, thus, the estimated time of divergence will be less than the true time (Nei, 1976). The overall expansion of the sheep populations in Australia from which we sampled should result in an underestimate of the time of divergence.

However, that effect is likely to be more than offset by the breeding structure. The effective population size will be greatly reduced by the hierarchical structure of 'studs' within the breeds – although the Roseworthy Merinos have been a completely closed flock since 1948. The use of relatively small numbers of sires would, of course, be expected to accelerate drift. The formula due to Wright (1931) can be used to estimate the difference,

$$N_e = \frac{4N_f N_m}{N_f + N_m},$$

where N_e is the effective population size, N_f is the number of females and N_m is the number of males. If an extreme value is taken, e.g. one ram to 100 ewes, N_e is close to 1/25 of the total breeding population. In other words, drift would be accelerated approximately 25-fold. As in some cases rams are used for a period shorter than the length of life of the average ewe, the actual difference between N_e and $N_f + N_m$ will be less.

Furthermore, depending on the intensity of selection and the heritability of the selected characters, N_e decreases yet further as a consequence of the variability in the selected characters among the different family groups within the selected population (Robertson, 1961). Thus, the hierarchical breeding system has the potential to accelerate drift – and thereby increase the expected genetic distance.

Although the time interval is relatively small, and thus its utility as a check is only marginal, our data on transferrin gene frequencies are extremely close to those obtained by Mayo and colleagues (1970), who surveyed the same large flock of Merinos approximately 10 years before we did (Table 2). As discussed previously, the haemoglobin gene frequencies are also similar. For few other populations in the protein polymorphism literature are such independent checks available. The methods of breeding and the changes in population size have not brought about large changes in two of the marker gene loci in the last ten years; the frequency of Hb A has increased in both the visual and the index flocks, although not to the point of being statistically significant.

Table 2. *Comparison of transferrin gene frequency data obtained on Roseworthy Merinos*

	Tf^A	Tf^B	Tf^C	Tf^D (includes D⁻)	Tf^E	No. of sheep
Roseworthy Index Merinos surveyed by Mayo *et al.* (1970)	0·250	0·211	0	0·525	0·013	334
Roseworthy Index Merinos surveyed by us	0·229	0·240	0	0·524	0·003	146
Roseworthy Visual Merinos surveyed by Mayo *et al.* (1970)	0·184	0·212	0·007	0·565	0·032	345
Roseworthy Visual Merinos surveyed by us	0·17	0·20	0·03	0·56	0·03	32

(v) *Migration?*

Migration of genes between the two diverging populations will result in an underestimate of the true time of divergence (Nei, 1976). Many breeds of sheep evolved slowly with repeated crossing with similar, or sometimes not-so-similar, breeds; therefore the pattern of evolution is reticulate, especially before breed type is stabilized. There is evidence in the history of these breeds to consider gene migration seriously, if probably of only minor significance (Austin, 1944; Carter, 1964; Ryder, 1964).

(vi) *Selection?*

As Nei (1976) has shown, if differential selection is sufficiently strong in relation to genetic drift, then the standard genetic distance is increased. This requires sufficiently long periods of *different* selection pressures on the diverging populations. The Roseworthy Merinos have been quite strongly selected for wool production (Mayo *et al.* 1970) and the Poll Dorset has been selected for crossing in fat lamb production (Close, 1971). However, the Roseworthy Merinos are derived wholly from the 'Bungaree' strain of Merino, which also has relatively good carcass

characteristics, and Merinos are increasingly being selected also for meat production (Dolling, 1970).

Evidence for selection has been presented for polymorphism of three NADP-dependent dehydrogenases and three esterases in the various organs of sheep (Baker & Manwell, 1976*a*). However, of these enzymes only the major 'malic enzyme' polymorphism occurs also in erythrocytes and thus is part of the present breed comparison.

Most searches for associations between production characters and biochemical or blood-group polymorphisms have resulted in only weak and variable correlations (Manwell & Baker, 1970; Mayo *et al.* 1970). Such changes in gene frequency of the marker protein loci may well be the result of 'hitch-hiking' to separate but linked loci actually determining the selected production properties.

Neglecting all other factors, if the time of divergence of ancestors of British breeds and the Merino was close to 2000 years ago (Ryder, 1973) and yet the calculated genetic distance is equivalent to a time closer to 70000 years, one can estimate that the observed results are 97% selection and 3% drift. However, we would wish to be cautious in such suggestions for the present, as data are needed on other strains of the Merino and on other sheep selected for different purposes.

(vii) *Additional hybridization in the ancestry of the Merino?*

Argument over descent of domesticated plants and animals from single species ('monophyletic theory') or from species hybrids ('polyphyletic theory') has continued since the time of Darwin. The unusual nature of many of the common protein polymorphisms of domesticated animals, the multiple amino acid substitution differences between variants, is most easily explained by the polyphyletic theory (Baker & Manwell, 1976*b*; Jope, 1976; Manwell & Baker, 1976). If one breed receives a differential input of genes from additional species hybridization, then one would expect the time calculated from the standard genetic distance to be greater than the time of domestication. Our estimate of 69700 years is not far from the 75000 years now estimated as the time of the beginning of the last ice age (Hammond, 1976). Speciation in the genus *Ovis* is believed to have occurred in the Pleistocene ice ages (Geist, 1971). On the basis of historical accounts Spanish workers consider it possible that the ancestors of the Merino received genes from a species distinct from that of other European breeds (Asensio Llorente, 1974). This other species of sheep might have been the source of genes for fine wool.

While it would be desirable to make the discussion of a possible polyphyletic origin of Merino sheep more specific by identifying the species which is implicated, the scarcity of data on wild ovicaprines makes this at present impossible; however, one ovicaprine from northern Africa has been eliminated (Manwell & Baker, 1977).

(viii) *Partitioning of genetic distance*

In the sheep studied by us genetic variation falls into two distinct categories: polymorphism with variants at high gene frequencies ($0{\cdot}2 < P < 0{\cdot}5$) and variation with low gene frequencies ($P < 0{\cdot}05$). As Nei's J_{XY} is based on summing the

products of gene frequencies over all sampled loci, it can be partitioned arbitrarily into two components, one representing high gene frequency variation and the other representing low gene frequency variation.

The authors are aware that the procedure is not legitimate in the usual way that genetic distance is calculated, which is over *all* sampled loci. Nevertheless, let us consider a possible rationale for such partitioning and see what happens:

Without entering into the current controversy over 'selectionist' versus 'neutralist' population genetics of electrophoretic variants, it is not unreasonable to suspect that polymorphisms with high gene frequencies represent selection. It is also not unreasonable to suggest that low-frequency variants represent in many cases relatively recent mutations which have given rise to essentially neutral alleles. Thus, changes in the gene frequency of the latter are more likely to represent drift – which is a basic postulate in the usual way genetic distance is converted to time of divergence (Nei, 1972, 1975, 1976). That these low-gene-frequency variants may well have arisen *after* divergence between Merino and British breeds is suggested by the fact that *none* of these low gene frequency variants occur in *both* breeds (Table 1).

If a hypothetical genetic distance is calculated using only the 24 loci without high-gene-frequency polymorphism, i.e. the differences between the breeds for the rare transferrin alleles, pyruvate kinase, erythrocyte Gly-Leu peptidases, glucosephosphate isomerase, superoxide dismutase, and NADH diaphorase II, then a value of 0·000294 is obtained, which corresponds to a time of divergence of 1470 years.

This figure is sufficiently close to Ryder's (1973) independent estimate of 2000 years for the time of divergence of the ancestors of the Merino and the British breeds of sheep to suggest that the approach warrants further study with other populations of sheep, and other domesticated species, where independent historical or archaeological checks on the probable time of divergence are available.

The authors are grateful to the Vice-Chancellor of the University of Adelaide, Professor G. M. Badger, for provision of funds for research equipment, and to the staff and students of Roseworthy Agricultural College, in particular, Mr Tom Mann, and to Mr Greg Powell of the University of Adelaide, for assistance in catching and bleeding sheep. We thank Dr G. Poll, Wellcome Research Laboratories, Berkhamsted, for the gift of special chemicals, and the members of the Spanish delegation at the International Society for Animal Blood Group Research Conference, 13–17 July 1976, for information about Merino sheep. We thank Dr Masatoshi Nei and the editor and referees of *Genetical Research* for helpful comments on the manuscript.

REFERENCES

ANANTHAKRISHNAN, R. (1973). A study of gene differences between some breeds of sheep. *Animal Blood Groups and Biochemical Genetics* **4**, 141–146.

ASENSIO LLORENTE, J. (1974). Aportaciones al estudio Historico-Evolutivo de las razas ovinas en España mediante los Polimorfismos Bioquimicos. Doctoral dissertation, Faculty of Veterinary Medicine, Zaragoza, Spain.

AUSTIN, H. B. (1944). *The Merino: Past, Present and Probable.* Sydney: Grahame Book Company.

BAILEY, N. T. J. (1959). *Statistical Methods in Biology*, 5th (1969) impression. London: English Universities Press.

BAKER, C. M. A. & MANWELL, C. (1976*a*). Heterozygosity of the sheep: polymorphism of 'malic enzyme', isocitrate dehydrogenase, catalase and esterase. *Australian Journal of Biological Science* (in the Press).

BAKER, C. M. A. & MANWELL, C. (1976*b*). Contributions to the study of animals at the molecular level with particular reference to the domestic fowl. *Philosophical Transactions of the Royal Society* (*London*) B **275**, 109–113.

BÖKÖNYI, S. (1976). Development of early stock rearing in the Near East. *Nature* **264**, 19–23.

BREWER, G. J., EATON, J. W., KNUTSEN, C. S. & BECK, C. C. (1967). A starch-gel electrophoretic method for the study of diaphorase isozymes and preliminary results with sheep and human erythrocytes. *Biochemical and Biophysical Research Communications* **29**, 198–204.

CARSON, H. L. (1976). Inference of the time of origin of some *Drosophila* species. *Nature* **259**, 395–396.

CARTER, H. B. (1964). *His Majesty's Spanish Flock: Sir Joseph Banks and Merinos of George III of England*. Sydney: Angus and Robertson.

CHAKRABORTY, R. & NEI, M. (1977). Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution* (in the Press).

CLOSE, R. L. (1971). Genetical studies of sheep. Honours thesis, Department of Zoology, University of Adelaide, South Australia.

DOLLING, C. H. S. (1970). *Breeding Merinos*. Adelaide: Rigby.

EFREMOV, G., VASKOV, B. & HRISOHO, R. (1970). Inherited variations in the prealbumins of sheep sera. *Proceedings of the XIth European Conference on Animal Blood Groups and Biochemical Polymorphisms*, Warsaw, 2–6 July 1968, pp. 505–511.

EWENS, W. J. & FELDMAN, M. W. (1976). The theoretical assessment of selective neutrality. In *Population Genetics and Ecology* (ed. by S. Karlin and E. Nevo), pp. 303–337. New York: Academic Press.

FÉSÜS, L. & RASMUSEN, B. A. (1971). The distribution of transferrin and hemoglobin types in families of Suffolk and Targhee sheep. *Animal Blood Groups and Biochemical Genetics* **2**, 39–43.

GEIST, B. (1971). *Mountain Sheep: A Study in Behavior and Evolution*. Chicago: University of Chicago Press.

HAMMOND, A. L. (1976). Paleoclimate: ice age earth was cool and dry. *Science* **191**, 455.

HOPE, R. M. (1966). Association between serum alkaline phosphatase variants and the R-O-i blood group system in the Australian Merino. *Australian Journal of Biological Science* **19**, 1171–1174.

JOPE, E. M. (1976). The evolution of plants and animals under domestication: the contribution of studies at the molecular level. *Philosophical Transactions of the Royal Society* (*London*) B **275**, 99–116.

LEE, R. M. (1964). Di-(2-chloroethyl)aryl phosphates: a study of their reaction with B-esterases, and of the genetic control of their hydrolysis in sheep. *Biochemical Pharmacology* **13**, 1551–1556.

LEWONTIN, R. C. & KRAKAUER, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**, 175–195.

MCDERMID, E. M., AGAR, N. S. & CHAI, C. K. (1975). Electrophoretic variation of red cell enzyme systems in farm animals. *Animal Blood Groups and Biochemical Genetics* **6**, 127–174.

MANWELL, C. (1977). A simplified electrophoretic system for determining molecular weight of proteins. *Biochemical Journal* (in the Press).

MANWELL, C. & BAKER, C. M. A. (1970). *Molecular Biology and the Origin of Species*. London: Sidgwick and Jackson.

MANWELL, C. & BAKER, C. M. A. (1974). Molecular genetics of avian proteins. XIII. Protein polymorphism in three species of Australian passerines. *Australian Journal of Biological Science* **28**, 545–557.

MANWELL, C. & BAKER, C. M. A. (1976). Protein polymorphisms in domesticated species: evidence for hybrid origin? In *Population Genetics and Ecology* (ed. S. Karlin and E. Nevo), pp. 105–140. New York: Academic Press.

MANWELL, C. & BAKER, C. M. A. (1977). *Ammotragus lervia*: barbary sheep or barbary goat? *Comparative Biochemistry and Physiology* (in the Press).

MAYO, O., COOPER, D. W., BRADY, R. E. & HOOPER, C. W. (1970). Response to partial selection on clean fleece weight in South Australian strong-wool Merino sheep. II. Associations between production characters, fertility, and three genetic polymorphisms. *Australian Journal of Agricultural Research* **21**, 541–547.

NEI, M. (1972). Genetic distance between populations. *American Naturalist* **106**, 283–292.

NEI, M. (1975). *Molecular Population Genetics and Evolution*. Amsterdam: North-Holland.

NEI, M. (1976). Mathematical models of speciation and genetic distance. In *Population Genetics and Ecology* (ed. S. Karlin and E. Nevo), pp. 723–766. New York: Academic Press.

NEI, M. & ROYCHOUDHURY, A. K. (1974). Genic variation within and between the three major races of man, Caucasoids, Negroids and Mongoloids. *American Journal of Human Genetics* **26**, 421–443.

RENDEL, J. & STORMONT, C. (1964). Variants of ovine alkaline serum phosphatases and their association with the R-O blood groups. *Proceedings of the Society for Experimental Biology and Medicine* **115**, 853–856.

ROBERTSON, A. (1961). Inbreeding in artificial selection programmes. *Genetical Research* **2**, 189–194.

RYDER, M. L. (1964). The history of the sheep breeds in Britain. *Agricultural History Review* **12**, 1–12, 65–82.

RYDER, M. L. (1973). The use of the skin and coat in studies of changes following domestication. In *Domestikationsforschung und Geschichte der Haustiere* (ed. J. Matolcsi), pp. 163–168. Budapest: Akadémiai Kiadó.

STRATIL, A. & GLASNÁK, V. (1974). Al^v – a new allele at the albumin locus in sheep. *Animal Blood Groups and Biochemical Genetics* **5**, 193–194.

TUCKER, E. M., SUZUKI, Y. & STORMONT, C. (1967). Three new phenotypic systems in the blood of sheep. *Vox Sanguinis* **13**, 246–262.

WRIGHT, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.

YAMAZAKI, T. & MARUYAMA, T. (1974). Evidence that enzyme polymorphisms are selectively neutral, but blood group polymorphisms are not. *Science* **183**, 1091–1092.