

Discriminating among cattle breeds using genetic markers

S. C. BLOTT†, J. L. WILLIAMS & C. S. HALEY*

Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, U.K.

Genetic markers provide a potentially powerful means of identifying the breed of individual animals. In this study diallelic and microsatellite loci were compared for their efficiency in discriminating among cattle breeds. Data were simulated for seven European cattle breeds using allele frequencies estimated at 20 microsatellite and 30 diallelic markers. Animals were assigned to the breed for which their genotype had the highest probability, and the power of the method assessed by estimating the error rate or proportion of animals misclassified. The number of markers required for discriminating among pure, or both pure and crossbreed, animals was investigated using either randomly sampled markers or markers selected on individual error rate. The relationship between individual marker variability and discriminatory power was also investigated. Microsatellite markers were found to be more powerful than diallelic markers for distinguishing among the breeds. The most discriminatory markers were those with the highest average heterozygosity and observed number of alleles. The number of markers needed to achieve a particular error rate could be reduced by selecting markers with the lowest individual error rates. Discrimination among both crossbreeds and pure breeds required approximately three times as many markers as discrimination among pure breeds alone.

Keywords: breed identification, cattle, genetic markers.

Introduction

Livestock populations have been subject to a variety of evolutionary forces during their histories. The cumulative effects of genetic drift, caused by founder effects and small population size, together with natural and artificial selection has led to the formation of distinct breeds. Studies of cattle breed relationships, based on genetic markers, have found that breeds are significantly differentiated at the genetic level (MacHugh, 1996; Blott *et al.*, 1998). Genetic markers could therefore provide a potentially powerful way of identifying the breed to which an individual animal belongs, when pedigree information is missing or suspect.

Livestock genetic conservation is currently based on the maintenance of pure-breeding populations (Hall & Bradley, 1995) and genetic markers can provide a method of identifying when introgression or crossbreeding has taken place (Bradley *et al.*, 1994). A test of breed identity would also be valuable for the validation of livestock products. This application may

become increasingly important as breed names become more widely used as a 'brand' name for livestock products (e.g. 'Aberdeen Angus' beef, 'Ayrshire' and 'Jersey and Guernsey' milk). Protection of the brand name may require that the products can be validated by a DNA test.

Different types of genetic marker will vary in their information content. Microsatellite loci have been identified as particularly useful markers for discriminating among populations. Bowcock *et al.* (1994) and MacHugh (1996) found that individuals from the same populations clustered together when a phylogenetic tree was constructed from microsatellite marker genotypes. New genotyping technologies have recently been developed (e.g. microarray chips for typing single nucleotide polymorphisms (SNPs)) which allow large numbers of samples to be genotyped using semiautomated assays (Delahunty *et al.*, 1996). Single nucleotide polymorphisms are usually diallelic (Delahunty *et al.*, 1996), and although this means that they are inherently less informative than markers with multiple alleles such as microsatellites, they have several advantages for forensic applications (Kwok *et al.*, 1994).

The main objective of this paper was to compare the two types of marker, diallelic and microsatellite markers, for their efficiency in distinguishing among cattle

*Correspondence. E-mail: chris.haley@bbsrc.ac.uk

†Present address: Department of Genetics, Faculty of Veterinary Medicine, University of Liège, 20 Boulevard de Colonster, 4000-Liège, Belgium.

breeds. The analysis was based on published estimates of allele frequencies at microsatellite loci (MacHugh *et al.*, 1997) and diallelic blood type loci (Blott *et al.*, 1998), in seven European cattle breeds. Bowcock *et al.* (1994) suggested that measures of marker variability, such as heterozygosity and F_{ST} , may be indicators of the power of individual loci to discriminate among populations. If these measures are correlated with discriminatory power, then markers could be selected on the basis of them. The relationship between marker variability and discriminatory power is investigated in this paper. Commercial tests of breed may also require that crossbreed animals can be identified, because a number of cattle breed meat marketing schemes (e.g. Aberdeen Angus marketing in the UK, and the 'Hereford Prime' scheme in New Zealand) not only accept pure-bred animals but also offspring of pure-bred bulls. The number of markers required to identify crossbreed animals was investigated by simulating data for crossbreed populations and assessing the efficiency of microsatellite markers at distinguishing among them.

Materials and methods

Data

The analysis was based on the comparison of seven European (*Bos taurus*) breeds; Aberdeen Angus (AA), Charolais (CH), Holstein-Friesian (FR), Hereford (HF), Jersey (JS), Kerry (KY) and Simmental (SM). Blood type data (described in Blott *et al.*, 1998) were provided by the Roslin Cattle Blood Typing Service. The allele frequencies at 30 red cell antigen loci were assumed to represent frequencies at a typical set of diallelic loci in European cattle breeds. These blood type allele frequencies are available on the Cattle Diversity Database (<http://www.ri.bbsrc.ac.uk/cdiv-www/>). Allele frequencies at 20 microsatellite markers were published by MacHugh *et al.* (1997) (ftp site: [acer.gen.tcd.ie/pub/cow_microsat/](ftp://acer.gen.tcd.ie/pub/cow_microsat/)). Details of the microsatellite data collection and genotyping protocols are described in MacHugh *et al.* (1997).

Data simulation

The observed allele frequencies in the seven breeds were used to simulate new genotypes. First, microsatellite and diallelic marker genotypes were generated for the seven pure breeds. Second, microsatellite marker genotypes were generated for five of these breeds (Aberdeen Angus, Charolais, Friesian, Hereford and Simmental) and four crosses, namely between the Friesian and (i) Aberdeen Angus (FR × AA), (ii) Charolais (FR × CH), (iii) Hereford (FR × HF) or (iv) Simmental (FR × SM).

These represent typical crosses that might be made by farmers when using terminal beef sires on nonelite dairy cows. Genotypes were simulated for the pure breed animals assuming Hardy–Weinberg equilibrium (random mating within breeds). Genotypes for the crossbreed animals were generated by sampling the first allele from one parental breed and the second allele from the other parental breed. The number of animals (genotypes) simulated for each pure breed was the same number as in the original data sets; this ranged from 52 Kerry to 7778 Holstein-Friesian for the blood typing data (Blott *et al.*, 1998), and 33 Aberdeen Angus to 40 Friesian and Kerry for the microsatellite data (MacHugh *et al.*, 1997). The same numbers of animals as in the original pure-breed samples were simulated in order to reflect the degree of error in the original sampling process. Fifty animals were generated for each of the crossbreed populations. The simulation was repeated 200 times for each data set.

Allocation of an individual to a breed

Breed identification can be treated as a decision problem. Given the genotype of an animal a decision must be made as to which breed the animal comes from. Each individual (x) can be considered as a vector of genotypes at m loci, having a probability density function $f_k(x)$ in breed k and $f_j(x)$ in breed j . An allocation rule is required that partitions the sample space, so that an individual is assigned to a particular breed according to the region in which its genotype falls. Two regions of error will be associated with the allocation rule. The probability of allocating an individual to breed j when it actually came from breed k and the probability of allocating the individual to breed k when it actually came from breed j . The two types of error may not be equally serious and this can be accounted for by specifying two costs caused by misclassification. An allocation rule is derived by minimizing the expected overall cost (Krzanowski & Marriott, 1995). The incidence of individuals may also differ between the two populations; some breeds are more numerous than others, so prior probabilities q_j , q_k of drawing an individual from breed j or breed k , respectively, can be specified.

If all costs are assumed to be equal then the rule can be simplified (Krzanowski & Marriott, 1995), so that the individual (x) is assigned to breed k rather than breed j if

$$q_j f_j(x) < q_k f_k(x), \quad j \neq k.$$

In other words, an individual is assigned to the breed for which its genotype posterior probability is highest. Within each breed the probabilities of the genotypes can

be estimated from a multinomial distribution, the parameters of each distribution being the allele frequencies in each breed. Estimates of the gene frequencies, obtained by sampling individuals from each breed, can be used to construct the allocation rule.

The probability of each simulated individual's genotype occurring in each breed was calculated over all markers. Assuming Hardy–Weinberg equilibrium within the pure breeds, genotype probabilities at a single locus are p_{jx}^2 for the homozygotes and $2p_{jx}p_{jy}$ for the heterozygotes. The genotype probabilities in the crossbreed animals are $p_{jx}p_{ky}$ for the homozygotes and $p_{jx}p_{ky} + p_{jy}p_{kx}$ for the heterozygotes. p_{jx} and p_{kx} are the frequencies of allele x in breeds j and k , and p_{jy} and p_{ky} are the frequencies of allele y in breeds j and k , where $x \neq y$. The genotype probability over all m markers is given by. If equal prior probabilities are assigned to the breeds then the individual is simply allocated to the breed for which this probability is highest.

Prior probabilities

When sampling has been carried out separately from each population (the situation when animals are deliberately sampled from each breed), rather than from a mixture of the breeds (which might be the case for data collected over a given time period by a commercial genotyping service, for example), there are no simple estimates of the prior probabilities available from the data (Krzanowski & Marriott, 1995). Unless extra information is available, equal prior probabilities must be assumed. The analysis presented here assumes equal prior probabilities for the breeds, but in a commercial test it might be desirable to incorporate prior probabilities. Additional information that might be used to provide prior probabilities is the population distribution or census sizes of breeds.

Estimation of the error rate

The efficiency of the different markers for discriminating among the breeds was measured by calculating the proportion of animals that were misclassified with each set of markers. The overall proportion of animals misclassified (error rate) was calculated as the total number of animals allocated to the wrong breeds divided by the total number of animals to be allocated.

Two types of error were determined: type I error, or the proportion of individuals of one breed that are allocated to another breed (really breed A but allocated to breed B), and type II error, or the proportion of individuals that are allocated to one breed but are really of another breed (allocated to breed A but really a

member of breed B). For any individual breed the two types of error may not be equal; however, across all breeds being compared the two error rates are equal to one another. Both types of error were calculated for each individual breed. Error rates were determined at each replicate of the simulation and the mean error rate over all 200 replicates calculated.

Random sampling of markers

The effect of increasing the number of markers, on the accuracy of breed identification, was investigated by randomly sampling markers with replacement from the original set of markers. For the microsatellite loci the number of markers sampled ranged between one and 60, and for the diallelic markers the numbers sampled ranged between one and 250. This random sampling of markers assumes that the original markers are a random set of markers of their type, and are representative of the distribution of allele frequencies at these loci.

Selection of microsatellite markers

The efficiency of individual microsatellite markers for discriminating among breeds was investigated by estimating the error rate for each marker separately. Markers were then selected on their individual error rates. The best markers were combined, those with lowest error rates being added first, and the resulting combined error rates were compared with those obtained when markers were randomly sampled. Marker selection was carried out for the microsatellite markers only, as they were more variable in numbers of alleles and average heterozygosity than the diallelic markers, allowing any relationship between these measures and error rate to be more clearly seen.

The relationship between individual marker variability and error rate was investigated by comparing the error rate against average expected heterozygosity (calculated as $1 - \sum p^2$), number of alleles observed (averaged over all breeds) and F_{ST} for each individual marker. F_{ST} was calculated as $\sigma^2/[p(1-p)]$, where p is the mean allele frequency over all the breeds and σ^2 the variance of the allele frequency among the breeds. This value was then averaged over all alleles at a locus.

Results

Comparison of diallelic markers with microsatellite markers

Figure 1 shows the comparison between microsatellite and diallelic markers when used to discriminate among seven pure breeds. Approximately 10 microsatellite

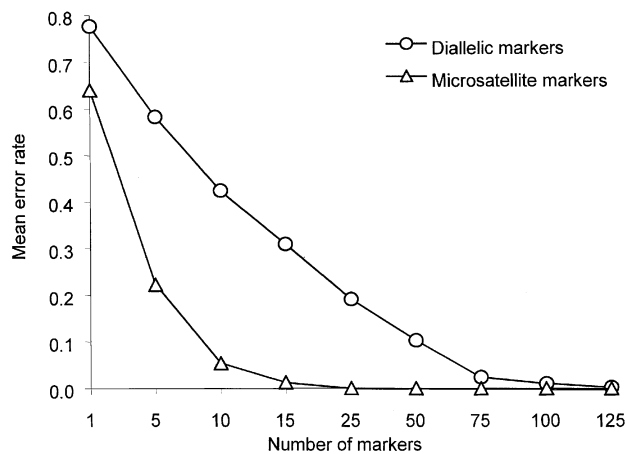


Fig. 1 Efficiency of microsatellite markers compared with diallelic markers for discriminating among seven European cattle breeds (markers were randomly sampled).

markers were required to achieve a 5% mean error rate compared with about 65 diallelic markers.

Selection of individual microsatellite markers

Table 1 shows the average heterozygosity, number of alleles observed, F_{ST} and mean error rate for each individual microsatellite marker (mean values across all breeds). The relationships between average heterozygosity, average number of alleles observed, F_{ST} and error rate for each of the markers are shown in Fig. 2. The markers that are most efficient at distinguishing among breeds (those that have the lowest error rate) are markers with the highest heterozygosities, and the greatest number of observed alleles. There is no clear relationship between F_{ST} and error rate.

Figure 3 illustrates the difference between using randomly sampled markers compared with selecting markers on their individual error rates. It can be seen that, using microsatellite markers to discriminate among the seven pure breeds, error rates of less than 5% can be achieved with approximately five to six selected markers compared with about 10 randomly sampled markers.

Discrimination among pure and crossbred animals

First crosses between breeds are closely related to both parental breeds, and this makes the problem of discrimination more difficult. Table 2 shows both the type I and type II errors when the best five selected microsatellite markers (lowest individual error rates) were used to discriminate among five pure breed (Aberdeen Angus, Charolais, Friesian, Hereford, Simmental) and four crossbred populations (Friesian crossed with Aberdeen

Table 1 Average heterozygosity, number of alleles observed, F_{ST} and error rate for each microsatellite marker. Markers are ranked by error rate

Marker	Average heterozygosity	No. alleles	F_{ST}	Error
BM2113	0.73	6	0.164	0.66
ETH131	0.76	8	0.133	0.67
BTMICROS	0.63	5	0.234	0.70
HEL1	0.64	5	0.240	0.72
HBB	0.67	6	0.131	0.75
ETH152	0.51	4	0.308	0.76
ETH225	0.65	5	0.201	0.77
HEL5	0.69	6	0.174	0.78
BoLA DRP1	0.79	8	0.064	0.79
ILSTS001	0.44	4	0.295	0.80
HIS-H1	0.40	3	0.218	0.83
ILSTS005	0.51	3	0.224	0.84
OCAM	0.33	3	0.359	0.85
TGLA116	0.53	3	0.222	0.85
RBP3	0.46	3	0.205	0.86
TGLA48	0.51	3	0.214	0.87
ILSTS014	0.33	3	0.166	0.87
RASA	0.52	3	0.204	0.87
PRL	0.20	2	0.132	0.90
BoLA DP2B	0.27	2	0.078	0.95

Angus, Charolais, Hereford or Simmental). The highest type I errors are seen in the first-cross Friesians, i.e. a fairly high proportion of these animals (30–40%) are mistaken for another breed, usually one of the parental breeds or another of the Friesian crosses. Type II error is highest for the pure-bred Friesian where 36% of animals allocated to Friesian are from other breeds; these misallocated animals are most likely to be first-cross Friesians. Type II error is also high among the crossbred Friesian populations (20–30%), either because animals from the parental breeds are wrongly allocated to them or the Friesian crosses are confused with one another.

Figure 3 illustrates the use of randomly sampled markers against selected markers for discriminating among pure and crossbred animals. About 30 randomly sampled markers were sufficient to achieve error rates of 5%, compared with the 10 randomly sampled markers required to discriminate with the same error rate among the seven pure breeds. The selection of markers was restricted by the fact that a total of only 20 microsatellite markers were available. The curve asymptotes at an error rate of around 12% for seven or more markers. It is possible that given markers that were all equally as discriminatory as the best five then error rates of 5% or less could be achieved with no more than 10–15 markers, at least for the crossbreeds chosen in this example.

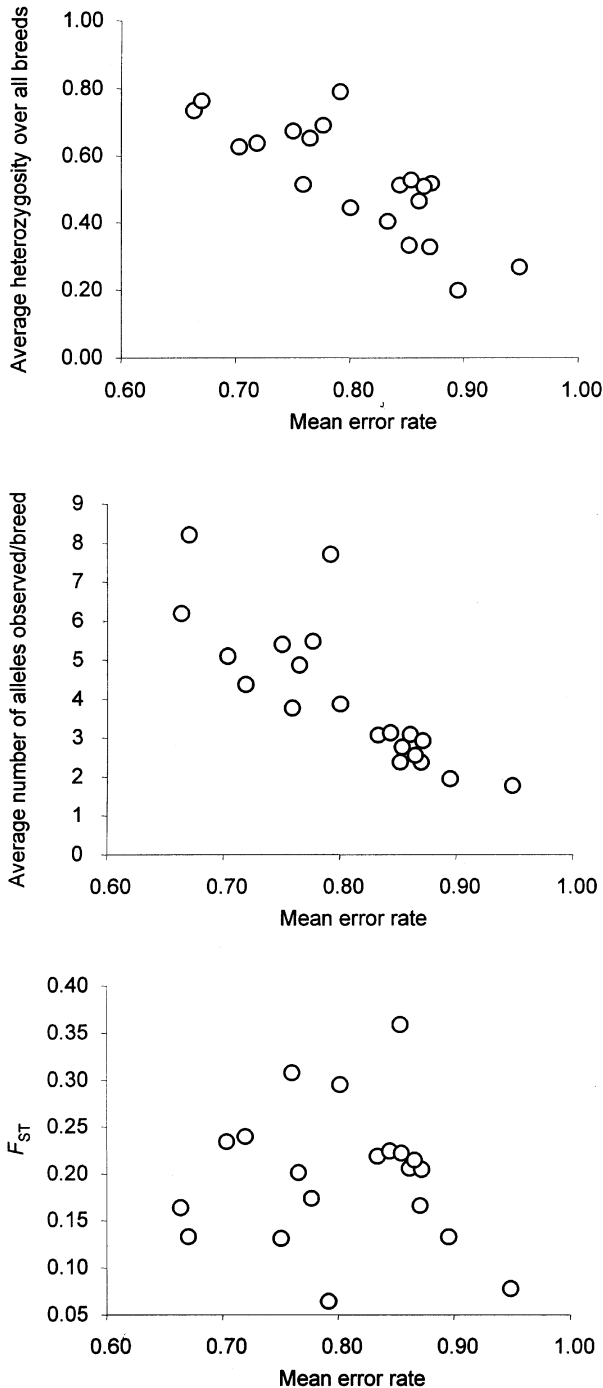


Fig. 2 Three measures of marker variability: (a) average heterozygosity; (b) average number of alleles observed; and (c) F_{ST} , plotted against mean error rate for each of 20 microsatellite markers.

Discussion

The results presented show that microsatellite markers are more powerful than diallelic markers for distin-

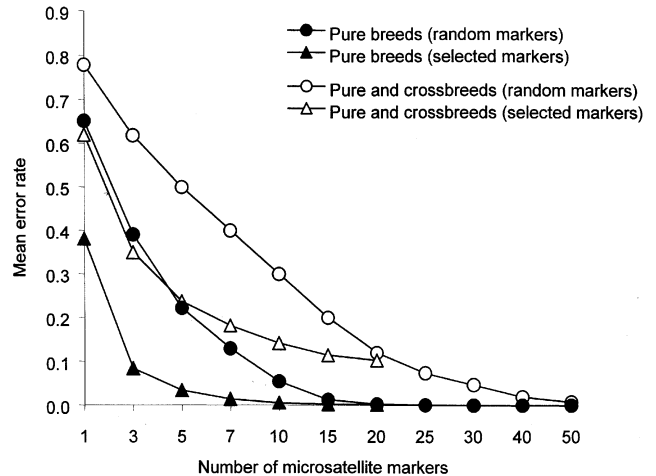


Fig. 3 Comparison between randomly sampled microsatellite markers and markers selected on individual error rates, when used to discriminate among either seven pure breeds or five pure breeds of cattle (Aberdeen Angus, Charolais, Friesian, Hereford and Simmental) and four Friesian crosses.

guishing among cattle breeds, at least when the distribution of diallelic marker frequencies is similar to that observed at blood type loci. Error rates below 5% can be achieved with 11–18 randomly sampled microsatellites against 65–100 randomly sampled diallelic markers. If the most discriminatory markers are selected then the number of markers required to achieve the same error rates can be reduced, by about half. The most powerful markers for breed discrimination are those with high average heterozygosities, and with greater mean numbers of observed alleles (across all breeds). The between-population variance in gene frequencies, estimated by F_{ST} , does not appear to have any clear relationship with the error rate for individual microsatellite markers. This suggests that the concern expressed by Bowcock *et al.* (1994), that markers with high diversity (high average heterozygosity across breeds) will be less informative for distinguishing among populations, is unfounded. Markers could be preselected on their average heterozygosity and number of alleles, using published information, in order to reduce the amount of genotyping required to develop a breed identification test.

The number of breeds being compared also affects the error rates obtained with a given set of markers. The larger the number of breeds being compared, the more markers are required, particularly if the breeds are closely related. When only the seven pure breeds are compared the error rate is less than 2% using 15 markers. However, if crossbreed populations are also considered, then more than 40 markers are required to achieve the same error rate (Fig. 3). The precise number of markers required to achieve a particular error rate

Table 2(a) Proportion of true breed (rows) allocated to each breed category

True breed	Allocated breed									Type I error
	AA	CH	FR	HF	SM	FR × AA	FR × CH	FR × HF	FR × SM	
AA	0.920	0.004	0.002	0.007	0.001	0.058	0.003	0.004	0.001	0.080
CH	0.006	0.854	0.005	0.009	0.012	0.010	0.084	0.006	0.014	0.146
FR	0.003	0.003	0.808	0.001	0.006	0.081	0.049	0.026	0.024	0.192
HF	0.010	0.009	0.001	0.904	0.001	0.006	0.004	0.064	0.001	0.096
SM	0.000	0.010	0.003	0.001	0.918	0.003	0.008	0.002	0.054	0.082
FR × AA	0.096	0.008	0.086	0.006	0.004	0.701	0.046	0.045	0.009	0.299
FR × CH	0.007	0.115	0.093	0.004	0.016	0.063	0.590	0.043	0.068	0.410
FR × HF	0.010	0.012	0.056	0.079	0.003	0.073	0.060	0.696	0.013	0.304
FR × SM	0.003	0.013	0.114	0.001	0.121	0.029	0.055	0.014	0.650	0.350

The proportion of animals allocated to breeds other than their true breed is the type I error. For example, in row one 92% of true AA were allocated to AA, 8% of true AA were allocated to other breeds (including 6% to FR × AA). Results shown were obtained by using the five best microsatellite markers, selected for lowest individual error rates, to discriminate among five pure European cattle breeds and four Friesian crosses.

Table 2(b) Each column shows proportion of animals assigned to each breed category according to their true breed origin

True breed	Allocated breed									Type II error
	AA	CH	FR	HF	SM	FR × AA	FR × CH	FR × HF	FR × SM	
AA	0.825	0.004	0.001	0.006	0.000	0.039	0.003	0.003	0.000	0.000
CH	0.006	0.788	0.003	0.009	0.010	0.008	0.071	0.005	0.012	0.012
FR	0.003	0.003	0.644	0.001	0.006	0.066	0.045	0.024	0.024	0.024
HF	0.010	0.008	0.001	0.856	0.001	0.004	0.003	0.050	0.001	0.001
SM	0.000	0.009	0.002	0.001	0.806	0.003	0.007	0.002	0.048	0.048
FR × AA	0.130	0.010	0.085	0.009	0.005	0.713	0.053	0.052	0.011	0.011
FR × CH	0.010	0.148	0.093	0.006	0.020	0.064	0.686	0.050	0.084	0.084
FR × HF	0.013	0.015	0.055	0.109	0.004	0.075	0.069	0.799	0.016	0.016
FR × SM	0.004	0.017	0.114	0.002	0.147	0.030	0.064	0.016	0.804	0.804
Type II error	0.175	0.212	0.356	0.144	0.194	0.287	0.314	0.201	0.196	0.196

The proportion of animals allocated to a breed, which were really of another breed, is the type II error. For example, in column one 82.5% of animals allocated to AA were true AA, 17.5% of animals allocated to AA were from other breeds (13% being FR × AA).

will depend on the number and type of populations that are to be compared.

The most powerful markers for distinguishing among populations would be those that are fixed for different alleles in different breeds (breed-specific alleles). European cattle breeds have, traditionally, been distinguished by phenotypic differences in coat colour, presence or absence of horns (polling) and traits such as double muscling. The genes controlling some of these traits have recently been identified (Brenneman *et al.*, 1996; Joerg *et al.*, 1996; Grobet *et al.*, 1997), and it has been suggested that these markers may be useful in a DNA test for breed identification (Georges & Andersson, 1996).

Allocation of individual animals to breeds using the method described in this paper depends on the estimation of multilocus genotype probabilities. These prob-

abilities are subject to sampling fluctuation, because the allele frequencies used to calculate them are estimated from samples drawn from the different breeds. When reporting the results of a test on an individual animal it would be desirable to be able to attach some level of confidence to its allocation. Chakraborty *et al.* (1993) have described how standard errors and confidence intervals for multilocus genotype probabilities can be evaluated, and give simple approximations to the sampling variance. Alternatively, bootstrapping could be used to estimate the confidence interval, by resampling the observations on which the allele frequency estimates are based. A new multilocus genotype probability would be generated at each bootstrap replicate, and 95% confidence intervals could be constructed around the probability of the genotype occurring in each breed.

Monte Carlo methods, such as bootstrapping, could also be used to approximate the distribution of genotypes within each breed making it possible to test whether an animal's genotype is extreme for the breed. Animals falling into the bottom or top 2.5% of the distribution, for example, could be rejected as members of the breed. Gotz & Thaller (1998) have suggested that this enables the hypothesis that the animal belongs to a known population vs. all other possible populations to be tested. In reality, it may be necessary to compare the genotype distribution for all possible breeds of interest. Enough markers should then be used to give the best separation of the genotype probability distributions, as the more the distributions overlap the greater will be the error rates.

Current genotyping technologies make a DNA test of breed possible. Microsatellite markers have been shown to be more powerful than diallelic markers for this type of test, and would be the marker of choice at the moment. There are some practical disadvantages to the use of microsatellites, however. Genotyping error can be high and it is difficult to compare genotypes typed in different laboratories. Developments in automated genotyping methods may ultimately favour the use of diallelic markers, because these technologies will allow many hundreds of loci to be scanned in a reliable and cost-effective manner.

Acknowledgements

We are grateful to David MacHugh for making the microsatellite data available. We would also like to thank Peter Keightley and Robin Thompson for their helpful comments. S.C.B. received financial support from the Rare Breeds Survival Trust, C.S.H. and J.L.W. acknowledge support from MAFF and BBSRC.

References

- BLOTT, S. C., WILLIAMS, J. L. AND HALEY, C. S. 1998. Genetic relationships among European cattle breeds. *Anim. Genet.*, **29**, 273–282.
- BOWCOCK, A. M., RUIZ-LINARES, A., TOMFOHRDE, J., MINCH, E., KIDD, J. R. AND CAVALLI-SFORZA, L. L. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, **368**, 455–457.
- BRADLEY, D. G., MACHUGH, D. E., LOFTUS, R. T., SOW, R. S., HOSTE, C. H. AND CUNNINGHAM, E. P. 1994. Zebu–taurine variation in Y chromosomal DNA: a sensitive assay for genetic introgression in West African trypanotolerant cattle populations. *Anim. Genet.*, **25**, 7–12.
- BRENNEMAN, R. A., DAVIS, S. K., SANDERS, J. O., BURNS, B. M., WHEELER, T. C., TURNER, J. W. AND TAYLOR, J. F. 1996. The polled locus maps to BTA1 in a *Bos indicus* × *Bos taurus* cross. *J. Hered.*, **87**, 156–161.
- CHAKRABORTY, R., SRINIVASAN, M. R. AND DAIGER, S. P. 1993. Evaluation of standard error and confidence interval of estimated multilocus genotype probabilities and their implications in DNA forensics. *Am. J. Hum. Genet.*, **52**, 60–70.
- DELAHUNTY, C., ANKENER, W., DENG, Q., ENG, J. AND NICKERSON, D. 1996. Testing the feasibility of DNA typing for human identification by PCR and oligonucleotide ligation assay. *Am. J. Hum. Genet.*, **58**, 1239–1246.
- GEORGES, M. AND ANDERSSON, L. 1996. Livestock genomics comes of age. *Genome Res.*, **6**, 907–921.
- GOTZ, K. U. AND THALLER, G. 1998. Assignment of individuals to populations using microsatellites. *J. Anim. Breed. Genet.*, **115**, 53–61.
- GROBET, L., ROYO MARTIN, J. L., PONCELET, D., PIROTTIN, D., BROUWERS, B., RIQUET, J. ET AL. 1997. A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. *Nature Genet.*, **17**, 71–74.
- HALL, S. J. G. AND BRADLEY, D. G. 1995. Conserving livestock breed biodiversity. *Trends Ecol. Evol.*, **10**, 267–270.
- JOERG, H., FRIES, H. R., MEIJERINK, E. AND STRANZINGER, G. F. 1996. Red coat color in Holstein cattle is associated with a deletion in the MSHR gene. *Mamm. Genome*, **7**, 317–318.
- KRZANOWSKI, W. J. AND MARRIOTT, F. H. C. 1995. *Multivariate Analysis. Part II. Classification, Covariance Structures and Repeated Measurements. Kendall's Library of Statistics.* Edward Arnold, London.
- KWOK, P.-Y., CARLSON, C., YAGER, T., ANKENER, W. AND NICKERSON, D. 1994. Comparative analysis of human DNA variations by fluorescence-based sequencing of PCR products. *Genomics*, **23**, 138–149.
- MACHUGH, D. E. 1996. *Molecular Biogeography and Genetic Structure of Domesticated Cattle.* PhD Thesis, University of Dublin, Ireland.
- MACHUGH, D. E., SHRIVER, M. D., LOFTUS, R. T., CUNNINGHAM, E. P. AND BRADLEY, D. G. 1997. Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics*, **146**, 1071–1086.