How reliable are sheep pedigrees?

A.M. CRAWFORD, M.L. TATE, J.C. McEWAN[1], G. KUMARAMANICKAVEL, K.M. McEWAN[1], K.G. DODDS[1], P.A. SWARBRICK AND P. THOMPSON

AgResearch Molecular Biology Unit, Biochemistry Department, Otago University, PO Box 56, Dunedin.

ABSTRACT

The advent of highly polymorphic genetic markers in sheep has enabled, for the first time, a retrospective study of the pedigree records of New Zealand sheep flocks. Three research flocks and one stud flock have been analysed. The flocks were genotyped using a variety of microsatellite DNA markers and/or the polymorphic blood proteins transferrin, vitamin D binding protein, and ovine plasminogen antigen.

The first flocks analysed were the half-sib Booroola pedigrees from progeny tests of heterozygous Booroola rams and their daughters. Of 291 daughters of 9 rams, 7 were shown to have an incorrect sire assignment. The remaining flocks consisted of full sire/dam/offspring pedigrees. The stud flock showed 2 instances in 45 pedigrees where the sire was incorrectly assigned. Of 53 Inverdale lambs tested 5 were shown to have an incorrectly assigned parent and in 387 pedigrees from a progeny test flock at Woodlands only two pedigree errors were detected.

These error rates are lower than previously estimated from field observations which give confidence in the use of these pedigrees for genetic selection. Simple flock management techniques are described which help achieve the low error rates reported.

Keywords: Pedigree auditing, Parentage testing, Microsatellites, Protein polymorphisms.

INTRODUCTION

The rate of genetic progress made in selecting for desirable traits is reduced if the estimation procedures are dependent on knowing the animal's pedigree and these records contain errors. (Van Vleck, 1970; Geldermann *et al.*, 1986). It is therefore highly desirable to minimise recording errors in commercially recorded flocks. In research flocks the validity of parentage records are essential when testing for genetic linkage or segregation and independent verification of parentage is required. Parentage information is also used for genetic parameter estimates. With the availability of relatively cheap and informative protein markers in sheep as well as the development of highly polymorphic microsatellite markers, good estimates of the error rate in recording parentage information can now be made retrospectively.

This paper describes the use of genetic markers to audit the pedigree records of four sheep flocks; three research flocks and one stud flock. We examine what effect various error levels are likely to have on genetic progress and based on the results of the flock showing the lowest level of pedigree error suggest simple flock management techniques that will keep recording errors to a minimum.

MATERIALS AND METHODS

Sheep Flocks:

Four sheep flocks were used in this study: A Romney stud breeder kindly allowed us access to his flock in which a total of 45 pedigrees were sampled. Each pedigree comprised a sire, dam and their offspring. From a commercial farmer's flock in which the Inverdale gene was segregating a total of 53 pedigrees, similar to the stud pedigrees described above, were sampled. The Booroola flock consisted of half-sib pedigrees in which the ram and his daughters were sampled. Each ram had between 24 and 41 daughters, each sire group being lambed in separate paddocks. The largest number of samples came from a progeny test flock at Woodlands. In this flock a total of 387 sire/dam/offspring pedigrees were analysed.

Preparation of sheep DNA for microsatellite analysis:

The DNA from the Booroola and Inverdale flocks was purified from heparinised blood using the method of Montgomery and Sise (1990). This is a time consuming and expensive process that can only be justified when the pedigrees are undergoing extensive genotype analysis.

For the stud flock pedigrees and the parasite resistance pedigrees a very simple DNA preparation procedure was undertaken based on that of Mercier *et al.*, 1990. A heparinised blood sample was taken in a 15 ml vacutainer. The tube was then centrifuged at 1500 x g for 5 minutes. The white cell material at the interface of the red cell pellet and the plasma, the so called "buffy coat", was removed and frozen. A 2µl sample of this was then added to 50 µl of Taq polymerase reaction buffer (10mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1 mg/ml gelatin) in a 0.5 ml polypropylene tube and heat treated. The heat treatment consisted of heating the sample for 3 minutes at 95°C then cooling for 3 minutes at 55°C. This two step process was repeated three times. The tube

[1]AgResearch Invermay Agricultural Centre, PO Box 50034, Mosgiel, New Zealand.

was then briefly spun in a microfuge to pellet coagulated protein material. The supernatant then became the DNA stock, 2µl of this stock was used for each PCR reaction.

Blood protein and microsatellite analysis:

Three blood protein polymorphisms, transferrin, vitamin D binding protein, and ovine plasminogen antigen were used in the analysis. These methods have been previously described (Tate *et al.*,1992).

Three microsatellites were used in the analysis, MAF36 (Swarbrick *et al.*, 1991), MAF45 (Swarbrick *et al.*, 1992) and MAF209 (Buchanan and Crawford 1992). The method of microsatellite analysis is fully described in Crawford *et al.*, (1991). Details of the markers are shown in Table 1.

The Booroola pedigrees have been genotyped with a large number of other microsatellites as part of our search for a linked marker to the Booroola gene. The number of microsatellites used on each pedigree varied from 8 to 51, however, apart from using different primer pairs and varying the annealing temperature to suit the primers the same method was used for all microsatellites.

TABLE 1: Details of the Genetic Markers used for pedigree auditing

Marker	number of alleles	PIC*	Paternity** exclusion
MAF36	13	0.86	0.74
MAF45	12	0.84	0.72
MAF209	8	0.79	0.63
Transferin	5	0.67	0.48
Plasminogen Antigen	4	0.44	0.26
Vitamin D binding protein	2	0.29	0.15

* PIC – Polymorphism Information Content (Botstein et al., 1980)
**paternity exclusion was estimated using the method of Jamieson (1965)
This method assumes that the dam is known and has been genotyped.

RESULTS

Different genetic markers were used on the different flocks and as a result the chances of detecting errors in the pedigrees varied. The most extensively studied Booroola pedigree has been genotyped with a total of 51 microsatellites so that the chances of detecting an error in this pedigree were greater than 99.9%. The least studied Booroola pedigree was genotyped with 8 microsatellites and 2 protein markers. We estimated the chances of detecting an error in this pedigree to be 95% (see appendix). The other flocks were genotyped with one or more of the markers described in Table 1.

In the half-sib Booroola pedigrees the only relationship we could check was between sire and offspring as the dams were not sampled.

To make the flock error rates comparable we have expressed the error rate in two ways, either errors found per sire/offspring or dam/offspring relationship or, as is more usual, the proportion of lambs identified with errors in their pedigree. The second measurement is usually double the first as two relationships are checked for each lamb.

The errors detected per parentage relationship, shown in table 2, varied from a high of 5% in the Inverdale flock to less than 0.4% in the progeny test flock. The chances of detecting an error varied depending on the markers used for genotyping. By correcting for this variation we were able to assign to each flock its likely error rate. The Inverdale flock had the highest rate (15% of lambs with incorrect pedigrees). The Booroola and stud flocks had similar rates (5% of lambs with incorrect pedigrees) and the lowest error rate was in the progeny test flock in which less than 1% of the lambs had errors in their recorded parentage.

DISCUSSION

The Inverdale flock which had the highest error rate is kept on a farm near Tuatapere, Southland. The higher error rate probably reflects the level of intensity of the stock managment. This flock is lambed in much larger paddocks with significantly fewer shepherds per animal than the other flocks. It is pleasing however that the other three flocks which were either kept in sire groups for lambing (Booroola flock) or were monitored quite intensively during lambing all had rates of error that were markedly below the range of 7.5 to 29.7% reported by others who used observation, immunological or protein techniques to determine pedigree errors (Welch & Kilgour, 1971; Welch & Kilgour, 1972; Alexander

TABLE 2

Flock (markers used)	Number of relationships checked	Number of errors detected	Percent errors detected	Chance of error detection*	Error rate per relationship	Error rate per pedigree
Booroola (multiple microsatellites)	291	7	2.4	>95%	2.5%	N/A
Stud (MAF36, MAF209)	97	2	2.1	83%	2.5%	5%
Inverdale (MAF45)	102	5	5.0	66%	7.5%	15%
Parasite Flock A (protein markers)	383	1	0.35	85%	0.4%	0.8%
Parasite Flock B (MAF36)	390	1	0.31	74%	0.4%	0.8%

*Determined for the Booroola pedigrees by assuming the loci were independent and using the estimate of parental exclusion outlined in the appendix. Determined for all other flocks by comparing each sire and dam pair with all progeny of other parents. The average proportion which could be excluded was used to estimate the chance of error detection.

et al., 1983; Sabo 1980; Iovenko, 1984; Marzanov & Lyutskanov, 1989; Wang & Foot, 1990).

The effect of pedigree errors on the rate of genetic progress under selection depends on their importance in determining the breeding value of the individual. At one extreme, if selection is based solely on an individual own value, pedigree mistakes have no effect. Another situation commonly encountered in the sheep industry is sire referencing where all information is dependent on pedigrees. In this case Van Vleck (1970) showed the reduction in genetic progress was largely independent of the heritability of the trait and mainly dependent on the percentage of progeny misidentified and the average number of progeny tested per sire. For example if 0.8 of the progeny are correctly identified then the reduction in genetic progress is about 10% if 20 progeny are tested and about 6% if 100 progeny are tested. These drop to 5% and 2% respectively if p=0.9.

Presently, within flock breeding values in sheep are commonly estimated on a combination of the records from the individual and its close relatives. Long *et al.*, (1990) simulated the effects of pedigree errors in a swine population when using best linear unbiased prediction with ancestoral or contemporary information. They found reduced genetic gains of 12.4% and 3.2% for traits with heritabilities of 0.13 and 0.50 respectively when when 20% pedigree errors were introduced into the data.

Efficient use of information from different types of relatives requires that the heritability be well estimated, but Van Vleck (1970a) shows that paternal half-sib heritability estimates are reduced by approximately $1-p^2$ when only a proportion of the progeny are assigned to the correct sire. The use of these biased estimates is another way in which pedigree errors can impede genetic progress

On research farms more extreme situations can occur when different selection lines or breeds are being grazed and lambed together as part of experimental comparisons. For example, the mis-identification and subsequent use of a ram between high and control fecundity selection lines could have dramatic impacts on the estimates of realised heritability of the trait. Also a single paternity error can invalidate or severely compromise gene mapping studies dependent on genetic segregation. Paternity verification is therefore an essential component of this work.

The very low rate of error (0.8%) found in the progeny test flock grazed at Woodlands Research Station is a credit to the people involved. It is even more exceptional given that from this flock, 800 ewes with a lamb drop of 1.8 lambs born per ewe lambing and weaned 147 lambs per 100 ewes mated, were lambed and recorded by one shepherd. While the shepherd was highly experienced, the major factor was the simple stock and record management methods he used to minimise ewe mis-mothering and the labour required for recording. These procedures were based on personal observation and experience by the shepherd and the results of research work on lambing behaviour in ewes. Much of this latter work has been summarised in clear simple language in Kilgour (1982) and Kilgour *et al.*, (1982) and is essential reading by any person intending to record parentage at lambing.

Briefly, mis-mothering is primarily a function of the number of ewes lambing close together in both time and space, the "mothering" ability of the ewe and whether the ewe has been disturbed, lambed or removed from her lambing site. Problems are accentuated if the ewe has twins and triplets, is lambing for the first time, and the lambing paddock has uneven topography. In this progeny test flock the ewes were set stocked during the lambing period in mobs of 20 to 30 ewes for periods of 10 or more days on flat sheltered 1.0 to 1.5 ha paddocks. First parity ewes were mixed with experienced ewes. Lambs born to ewes lambing overnight were marked with spray-raddle first thing in the morning and the ewe and number of lambs noted in the lambing book. Tagging was done later after any shepherding problems were attended to. Care was taken not to disturb ewes from their lambing site for 24 hours and to assist any ewes with malpresentations on their site. A second lambing round was done in the afternoon and lambs were either tagged or marked at this time. While in excess of 100 of the 800 ewes could lamb on a single day, on average only 1 or 2 ewes would lamb in a particular paddock between rounds and these were generally widely spaced, aiding accurate recording and the bonding of the ewe to her lambs.

Techniques for resolution of suspect pedigree identification included visual examination of the ewe for signs of having recently given birth and counting of umbilical cords in the placenta. The sheep used were a Coopworth strain that has been heavily selected for ease of lambing, mothering ability and lack of "flightiness" to shepherds. Accurate recording and checking of the information entered was also important in reducing errors and appropriate methods have been clearly outlined by Dalton (1982). In particular, checking ewe tag numbers mated to a particular ram, both prior to and after joining, was essential to detect misdrafting and "fence jumping". Similarly, daily completion of both progeny and dam order lists during lambing detected duplicate or non-existent identifications at a time when correction was possible.

The development of retrospective testing allows the performance of stock managers to be quantified so that allowances can be made if errors are detected. Although microsatellites are more informative than protein markers, at this stage in their development, it is cheaper and easier to use the protein markers to check pedigree records. The 3 blood protein markers used in this study can all be read from the one electrophoresis run and will detect approximately 80% of the pedigree errors, enough to provide a good estimate of the accuracy of the records i.e. parentage auditing as distinct from verification. It would cost in the region of $10 per animal.

If a very valuable animal requires its parentage information verified and perhaps only one parent is available, then genotyping with a panel of highly polymorphic microsatellites is the only way the information can be adequately checked. This testing is currently quite expensive, probably costing between $100 and $150 to analyse one sample with five microsatellites. Economies of scale should reduce this cost by at least 50% over the next few years. Microsatellite testing would only need to be undertaken after the protein markers mentioned above failed to exclude the assignment.

Now that the technology is available research flock managers should seriously consider getting their pedigree records audited to identify potential problems and in the case of genetic selection flocks we would recommend that all sires should have their parentage information verified.

ACKNOWLEDGEMENTS

The authors wish to thank the Romney breeder for allowing us access to his stud flock and pedigree records; Anne Schmack and Gordon Greer for collection of the blood samples; Kevin Knowler, Roger Wheeler and the other staff at Woodlands for the animal management and collection of records and samples from progeny test flock. Mr A.J. Gray for access to the Inverdale flock and the staff of the AgResearch Molecular Biology Unit for providing the Booroola and Inverdale flock DNA samples.

APPENDIX

Single parent – offspring exclusion for the Booroola pedigrees:

Suppose a locus has *v* alleles, and the the frequency of the *i*th allele (Ai) is pi. Then the progeny that can be excluded and the probability of this occuring are listed below.

Parent		Excluded Progeny	
Genotype	Probability	Genotypes	Probability
A_iA_i	p_i^2	AjAj,AjAk	$(1-p_i)^2$
A_iA_j	p_ip_j	A_kA_k,A_kA_l	$(1-p_i-p_j)^2$

The exclusion probability is:

$$\sum_{i=1}^{v} p_i^2(1-p_i)^2 + \sum_{i=1}^{v}\sum_{j\neq i} p_ip_j(1-p_i-p_j)^2$$

$$= \sum_{i=1}^{v}\sum_{j\neq i} p_ip_j\,(1 + 2p_ip_j + p_i^2 - 3p_i)$$

REFERENCES

Alexander, G.; Stevens, D.; Mottershead, B. (1983) Problems in the accurate recording of lambing data. *Australian Journal Experimental Animal Husbandry* **23**:361-368

Botstein D.; White R.L.; Skolnick M.; Davis R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**: 314-331.

Buchanan F.C.; Crawford A.M. (1992) Ovine dinucleotide repeat at the MAF209 locus. *Animal Genetics* **23**: 183.

Crawford A.M.; Buchanan, F.C.; Swarbrick, P.A. (1991) The use of dinucleotide repeats or microsatellites as genetic markers in domestic animals. *Proceedings of the New Zealand society of Animal Production* **51**: 79-83.

Dalton D.C. (1982) Performance Recording *In* Sheep Production: Volume One, Breeding and Reproduction G.A. Wickam and M.F. McDonald Eds. New Zealand Institute of Agricultural Science pp 199-210

Geldermann, H.; Pieper, U.; Weber, W.E. (1986) Effect of mis-identification on the estimation of breeding value and heritability in cattle. *Journal of Animal Science* **63**: 1759-1768

Iovenko, V.N. (1984) The effect of parentage errors on the effectiveness of selection *in* sheep Nauchno-Tekhnicheskii byulleten', Ukrainskii Nauchno-Issledovatel'skii Institut Zhivotnovodstva Stepnykh Raionov "Askaniya-Nova" No. 2 50-53

Kilgour, R.; Alexander, G.; Stevens, D. (1982) Lambing sites are important. *Flock and Herd* **7**: 26-28

Kilgour R. (1982) Better lambing procedures. *Proceedings of the Ruakura Farmers Conference* 9-13

Long,T.E.; Johnson, R.K.; Keele, J.W. (1990) Effects of errors in pedigree on three methods of estimating breeding value for litter size, backfat and average daily gain in swine. *Journal of Animal Science* **68**: 4067-4078

Jamieson A. (1965) The genetics of Transferrins in cattle. Heredity 20, 419-441.

Marzanov, N.S.; Lyutskanov, P.I. (1989) The use of blood groups in the selection of sheep. *In* 3 Shk.-semin. po genet. i selektsii zhivotnykh 12-19

Mercier B.; Gaucher C.; Feugeas O.; Mazurier C. (1990) Direct PCR from whole blood without DNA extraction. *Nucleic Acids Research* **18**: 5908.

Montgomery G.W.; Sise J.A. (1990) Extraction of DNA from sheep white blood cells. *New Zealand Journal of Agricultural Research* **33**: 437-441.

Sabo, I. (1980) Blood group studies and parentage tests in an experimental flock with recorded mating and group lambing. Thesis, Justus-Leibig-Universitat Geissen, German Democratic Republic 151pp

Swarbrick P.A.; Buchanan F.C.; Crawford A.M. (1991) Ovine dinucleotide repeat polymorphism at the MAF36 locus. *Animal Genetics* **22**: 377-378.

Swarbrick P.A.; Schmack A.E.; Crawford A.M. (1992) MAF45, a highly polymorphic marker for the pseudoautosomal region of the sheep genome, is not linked to the $FecX^I$ (Inverdale) gene. *Genomics* **13**: 849-851.

Tate M.L.; Manly H.C.; Dodds K.G.; Montgomery G.W. (1992) Genetic linkage analysis between protein polymorphisms and the Fec^B major gene in sheep. *Animal Genetics* **23**: 417-424.

Van Vleck, L.D. (1970) Misidentification and sire evaluation. *Journal of Dairy Science* **53**: 1697-1702

Van Vleck, L.D. (1970a) Misidentification in estimating the paternal sib correlation. *Journal of Dairy Science* **53**: 1469-1478

Wang, S.; Foote, W.C. (1990) Protein polymorphism in sheep pedigree testing. *Theriogenology* **34**: 1079-1085

Welch, R.A.S.; Kilgour, R. (1971) Mismothering in Romney sheep. *Proceedings of the New Zealand society of Animal Production* **31**: 41

Welch, R.A.S.; Kilgour, R. (1972) Survey of lambing practices in stud sheep flocks in New Zealand. *Proceedings of the New Zealand Society of Animal Production* **32**: 115-122