DNA-Matching In Deer

Mike Tate, Robert McDonald, Jamie Ward, and Ken Dodds

Abstract

New DNA technologies offer the potential to identify parents of a deer calf from among hundreds, or even thousands of possible parents. This paper examines, in theory and in practice, the power of the 13 marker DNA parentage test for deer developed by the ***Genomnz*** laboratory (AgResearch New Zealand). Theoretical calculations and case studies show that the test can match all calves to their parents in a typical deer mating group (e.g. 30-50 breeding animals). However, the case studies also show that DNA testing must be applied carefully and should not be viewed as a panacea for poor record keeping. In particular, blood samples need to be collected from all possible parents, requiring accurate recording and DNA sampling of mating and calving groups. Our calculations show the test is most powerful in "out-bred" groups of red deer and red deer hybrids. The test is substantially less powerful in Canadian elk and in "line-bred" herds, where the sires used are closely related. Finally, the power of the test is also greatly reduced when samples are not available from both parents (e.g. when only sires and calves are sampled).

Introduction

Genetic improvement has the potential to increase the productivity of deer and reduce losses from disease. Pedigree recording is an integral part of any genetic improvement programme because this family data enables the farmer to separate the genetic component of animal production from environmental factors (e.g. differences in feeding and drenching).

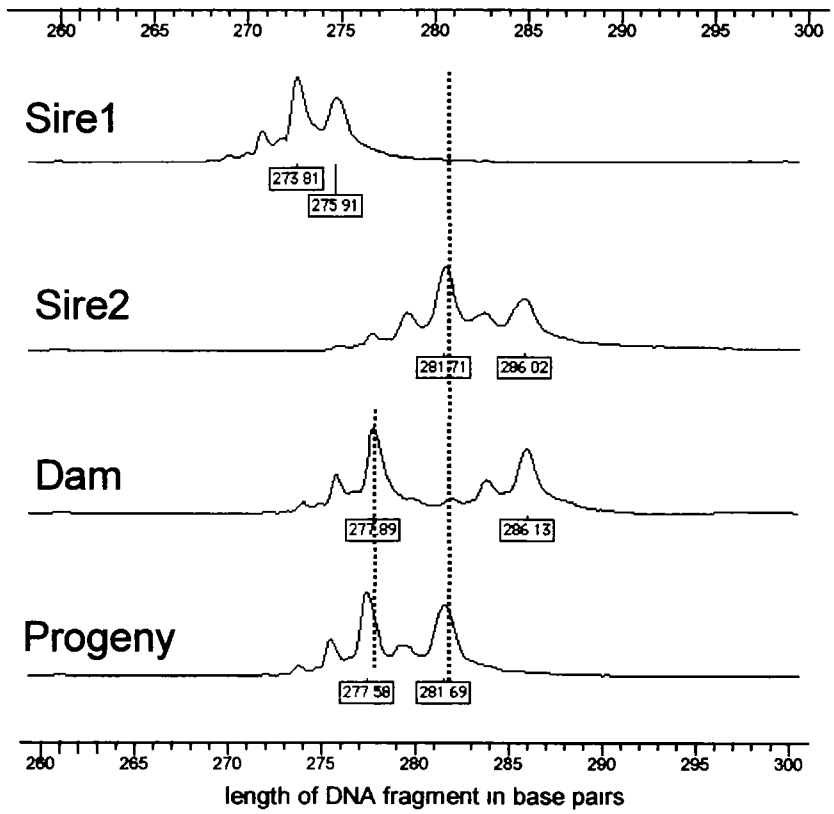
The primary sources of pedigree records on stud deer farms are the stag and hind locations during mating and observations of hind-calf association after calving. Advances in blood typing and, more recently, DNA analysis have provided new tests to check pedigrees recorded on farms using such traditional methods. Unfortunately, DNA-based tests have tended to identify a high level of errors in farm records (Tate *et al* 1992). Observational studies indicate that at least some of these errors may arise from the fact that hinds, particularly first calvers and those who have lost a calf, may suckle calves other than their own.

When blood protein or DNA-based tests reveal that a pedigree recorded on the farm is incorrect, the farmer's immediate request is to use the biochemical test to identify the correct parents. However, finding the correct pedigree for an excluded calf among a large number of possible parents is a far more difficult problem than simply checking a single putative pedigree. In a typical mating group, the possible parents of a calf may include any combination of 2 stags and up to 50 hinds (i.e. 100 possible combinations per calf). Until now, a molecular test

which excludes 95 out of 100 incorrect pedigrees has been considered a perfectly adequate tool for parentage verification or exclusion. However, faced with this particular herd situation such a test would, on average, present the farmer with 5 possible parental combinations per calf! This produces an extremely unsatisfactory situation: a molecular test has told a farmer that a recorded pedigree is wrong - but the test is not powerful enough to identify the correct pedigree.

Recent developments in molecular genetics and gene mapping in deer have the potential to rectify this situation. The growing number of new DNA markers make it possible *match* pedigrees rather than simply verify or exclude farm recorded pedigrees. This paper evaluates the new DNA tests developed by the AgResearch technology development unit; ***Genomnz.*** We calculate the potential of these new tests for matching a pedigree from a very large number of alternatives and discuss the implementation of this test in a number of case studies.

Figure 1: Laboratory Results



The Genomnz DNA Parentage Test

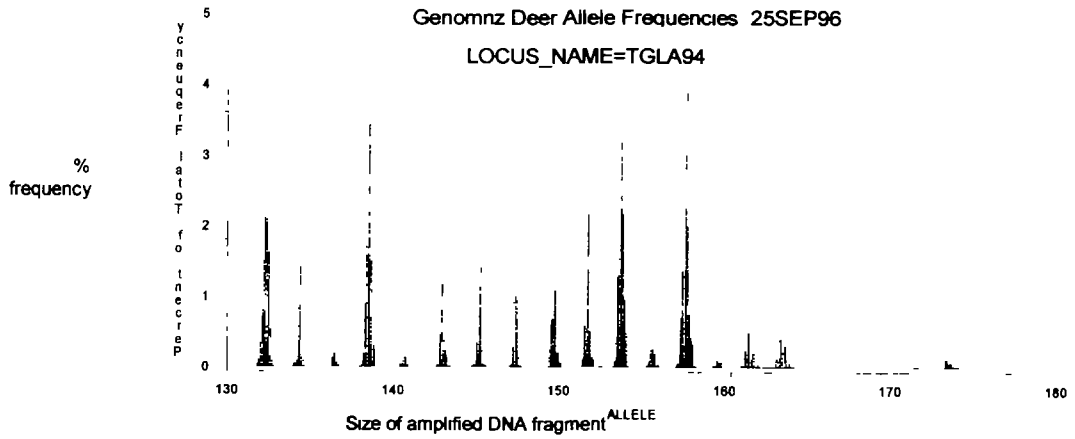
The ***Genomnz*** DNA profile for each animal is generated by polymerase chain reaction (PCR) amplification of thirteen regions of deer DNA which contain simple sequence repeats or microsatellite DNA. These regions or markers have been chosen because they are highly polymorphic (have many different forms) in the deer population. The different forms are often referred to as "types". Figure 1 shows the DNA from four individual deer analysed for one DNA marker. The data were generated by PCR amplification of the DNA region followed by

measurement of the length of the amplified fragments on an ABI 377 automated DNA analyser.

The measurement on the horizontal axis is the length of the amplified DNA fragment, measured in base pairs (bp). The peaks indicate the presence of amplified DNA. The tip of the peak is labelled (below) with its size. The minor peaks are a consistent feature of this particular marker system and are useful in the positive identification and labelling of the marker peaks by a computer algorithm. In all cases, the computer size calling is also checked manually.

Each deer in Figure 1 has two inherited DNA peaks. Geneticists call these two inherited types "alleles". For each marker in the test, one allele is inherited from the animal's sire and the other from its dam. The dam in Figure 1 was observed to suckle the progeny and, in farm records, this dam was recorded as the calf's mother. The DNA results are consistent with this observation because the dam and calf share the 277bp peak. As the 277bp allele comes from the dam, the calf's 281bp allele must come from its sire. From the mating records, it is known that there are two possible sires. Sire 1 does not contain the 281bp peak (or indeed any peak in common with the calf) and this excludes it from being a parent of the calf. Sire 2, however, shares the 281bp allele with the calf and therefore this stag could have sired the calf.

Figure 2: Frequency of DNA types in New Zealand farmed deer for marker GNZ-DY157



Marker Variability

The power of a DNA parentage test depends on how many different markers are assayed and how variable or polymorphic they are in the population. The more markers and more different types observed per marker, the more likely an incorrect parent is to be detected and excluded. The 13 markers used by the ***Genomnz*** test contain between 6 and 17 types when run in New Zealand farmed deer. Figure 2 summarises the length of the amplified DNA fragments in over 1000 deer from 29 deer studs for one of the 13 markers. Seventeen distinct fragment sizes (DNA types) occur in the population. Each DNA type is approximately two DNA base pairs greater in size than the previous one, apart from the largest which is separated from the others by 10 DNA base pairs. The height of the peaks in Figure 2 relates to their frequency in the

population. The spread around the peaks results from the slight measurement error in the ABI 377 DNA analysis system.

Theoretical Power of the Genomnz DNA Parentage Test

The power of a parentage test is usually summarised using a statistic called the probability of exclusion (Dodds *et al.* 1996). This describes the chance that an animal drawn at random from a population will be excluded from an incorrect pedigree drawn from the same population. When the probability of exclusion becomes very large (e.g. 0.9998) it is easier to present the converse: the chance that the profile of a calf will *match* a set of unrelated parents in a particular parentage situation (e.g. 2 in 10,000). The chance of a match with an unrelated animal depends on the power of the test and the exact situation or problem encountered.

Table 1: Power of the *Genomnz* DNA Parentage Test in various test situations

Type of test	Example	Odds that an unrelated animal will match (mean within stud value)	
		Red Deer and hybrids (n>1000, 36 studs)	Canadian Elk (n>200, 7 studs)
Identity	Semen or tissue sample available *identify the animal it came from*	<1 in 1 million	<1 in 1000
Parent-calf	Sire-dam mating pairs known *find the calf*	<1 in 10 000	<1 in 100
Paternity	Dam-calf match known. *find the sire*	<1 in 1000	<4 in 100
Maternity	Sire-calf match known. *find the dam*	<1 in 1000	<4 in 100
Sire only	No dam DNA *find the sire for a calf*	~2 in 100	<2 in 10
Dam only	No sire DNA *find the dam for a calf*	~2 in 100	<2 in 10

Table 1 gives the odds of a DNA match with an unrelated animal, calculated for the various situations encountered in deer. In all cases, the red deer and hybrid probability is the average figure found within a stud. The calculations show that the chance that two randomly selected animals will have the exactly the same profile is very small (less than 1 in 1 million). The DNA profile can therefore be used to identify an animal or any product produced from an animal (antler, tissue, semen). To date, practical applications of this "identity test" have included identification of semen straws which had become muddled in a large AI programme and matching of tissue remains in a paddock to a carcass found elsewhere in a forensic case.

The chance of a parentage match with an unrelated animal is greater than an exact match of profiles. This is because genetic segregation and recombination can produce a wide range of progeny profiles from one set of parents. The table shows that the best situation for parentage testing is where parent pairs are known and both parents are sampled. This knowledge restricts the types possible for the calf and therefore the chance that an unrelated calf will match. The most difficult situation for parentage testing is where only one parent is tested. In

this case, a wide range of sire (or dam) profiles will match a calf thereby increasing the chance that an unrelated animal will (mistakenly) match (Dodds *et al* 1998).

One of the major features of the analysis of farmed deer was the observation that elk imported from Canada were typically less variable for the ***Genomnz*** markers than red deer or red deer hybrids. Table 1 compares the average within-stud power of the test in red deer to that in Canadian elk. While the test remains useful in imported elk, chance matches between a calf and an unrelated parent and/or matches with multiple parents will be more common, especially if samples are not available from both parents.

Power of the Genomnz DNA Parentage Test in Practice

The theoretical analysis (above) suggests the ***Genomnz*** test should be able to identify a correct parentage among hundreds or in some cases thousands of alternatives when samples are available from all calves and every possible sire and dam. In practice however, there are situations in which the power of test may be significantly lower: (1) if the animals tested are closely related; and (2) if all markers are not scored accurately in all animals.

As animals become more closely related it is more difficult to distinguish them genetically. This is not generally a major issue in "out-bred" herds, however it may become a significant issue where animals have been inbred or "line-bred" for several generations. In such situations, it may be prudent to request that some animals from the herd (particularly stags) are sampled and profiled to help evaluate the power of the test.

When DNA profiles are generated for very large numbers of individual animals, some types or alleles may be incorrectly labelled at some markers and a low level of errors may occur (1-2%). Such errors can easily be detected by careful rechecking of one or more markers in each excluded pedigree. This sort of human error is relatively simple to detect and rectify. However, there are rare situations where an "incompatible" DNA marker type is detected in the correct pedigree. This can be caused either by a DNA mutation in of one of the animals in the pedigree, or by the occasional occurrence of DNA which, while present in the animal, does not amplify in the test. The former, new mutations, are very rare and occur only in one out of every 10 000-100 000 matings in the sort of markers used by ***Genomnz***. The latter, known as null alleles, are screened for in the parentage test algorithm and have been occasionally detected in only three markers of the 13 marker set.

To allow for these anomalies and other rare possibilities, results of the ***Genomnz*** DNA Parentage Test are reported in three categories.

Match The DNA profile of the calf is entirely consistent with the profiles of the sire and dam according to the rules of genetic inheritance.

Qualified A single allele found in the progeny is not consistent with the alleles found in the parents at that marker. All other markers are consistent and the mismatch could have been caused by a mutation or a non-amplifying allele.

Excluded Two or more alleles are incompatible between the calf and parents.

Animals which are "Excluded" cannot be the parents. Animals which "Qualify" are reported but are only considered as correct parents if an exact "Match" cannot be found among all the possible parents.

Case studies

The ***Genomnz*** test was used in the 1996 and 1997 seasons to match 210 calves to parents in six "whole herd matching" situations. These are listed in Table 2 below as case studies 1 to 6. The most important parameter in ranking the difficulty of these problems is the total number of pedigree combinations tested. Case studies 3 and 4 involved the use of embryo transfer which reduced the number of combinations because only a few parents were used and the parent (mating) pairs were already known.

Only two of the case studies involved cases where calves matched multiple parents. These were case study 2, which involved a group of closely related elk in which the power of the test is reduced (see above) and case study 6, a very large mix and match problem where key research animals became mixed with other calves before tagging. In a problem of this size some chance matches are to be expected. In case study 6, only the parents which were used to produce the research animals were tested. The parentage test successfully identified 66 calves which could not have been from these parents.

Table 2 Summary of DNA Matching Case Studies, 1996-97

Case Study	Number of Combinations			Progeny matched	Multiple matches	No matches
	Sire x Dams	x Calves	Total Comb.	to a single parent pair		
#1	2 x 14	x 12	336	12		
#2	4 x 14	x 14	784	9	1	4*
#3	(6 ,16)	x 56	896	56		
#4	(5 ,16)	x 36	1296	36		
#5	(11 ,39)	x 36	1404	36		
#6	(6+,95+)	x 151	14,356	70	15	66*
			19,072	210		

*see text

While a successful outcome was achieved in all cases, in some instances this took many months and several testing runs. Invariably, this was due to the first testing run identifying calves which did not match any of the putative pedigrees. Case study 2 is still in this phase and further possible parents are due to be sent for testing. In each case, when new parents and or new samples were supplied a unique solution was found. This has highlighted the importance of ensuring samples are collected from *all* the possible parents involved. Ideally this should occur before mating, or as soon after as practical.

Conclusions

For the first time, the ***Genomnz*** DNA Parentage Test provides the potential to match pedigrees using DNA alone. The test offers an alternative to traditional matching but it is not

a panacea for poor record keeping. While the test performs well in some situations, in other situations it may not be able to uniquely identify pedigrees for all calves:

The Best Situation

- red deer or red deer hybrids
- unrelated stags
- mating pairs recorded (e.g. which hinds mated to which stags)
- samples taken from all parents
- calving groups recorded and relatively small (e.g. 30)
- samples from all calves

The Worst Situation

- Canadian elk
- Closely related sires (e.g. a sire and his sons)
- Line bred herd
- exact mating pairs unknown (e.g. test all possible sires with all dams)
- samples unavailable from some parents
- very large or unknown calving groups
- samples unavailable from some calves

In the first scenario the chance of uniquely matching all pedigrees is very good. In the latter the test can still be useful but the expected outcome would be that some calves may match two or more sets of parents.

References

Dodds, K , R. McDonald and M Tate (1998). Probability of parentage assignment using genetic markers *Proceedings of the 6th World Congress on Genetics as Applied to Animal Production* **25**:637-640

Dodds, K., M. Tate, J. McEwan and A Crawford (1996). Exclusion probabilities for pedigree testing in farm animals. *Theoretical and Applied Genetics* **92**: 966-975.

Tate, M. L , F C. Buchanan and A. M Crawford (1990) Parentage testing in farmed red deer *Proceedings of a Deer Course for Veterinarians* 7 177-182