Contributed paper

A GENOMIC COMPARISON OF AUSTRALIAN, NEW ZEALAND AND NORWEGIAN DAIRY GOAT POPULATIONS

K.G. Dodds[1], J.H. Jakobsen[2], M. Wheeler[3], T. Blichfeldt[2], R. Brauning[1], H.J. Baird[1] and J.C. McEwan[1]

[1] AgResearch, Invermay Agricultural Centre, Mosgiel, New Zealand
[2] The Norwegian Association of Sheep and Goat Breeders, Box 104, 1431 Ås, Norway
[3] AgResearch, Ruakura Agricultural Centre, Mosgiel, New Zealand

SUMMARY

Six dairy goat populations that had been genotyped using genotyping-by-sequencing (GBS) were compared. The populations were an industry data set from Norway along with four herds from New Zealand (NZ) and one from Australia. The Norwegian population was found to be the most genetically diverged from the others. One of the NZ populations was also quite distinct while the other NZ populations appear to be genetically similar to each other and closer to the Australian population than the other NZ population. It may be useful to combine these three NZ populations and the Australian population to provide better genomic evaluation.

INTRODUCTION

AgResearch have been providing genotyping services for several dairy goat enterprises with clients in New Zealand, Australia, and Norway. A genotyping-by-sequencing (GBS) platform has been used to provide a medium-density (~60k) SNP profile. In most cases the genetic background of the populations is poorly recorded. This study characterises and compares these populations based on genotypes from the common GBS platform used. One outcome of a comparison is that it would inform the likely usefulness of combining populations for genomic evaluation.

MATERIALS AND METHODS

Animals. The animals used in this study were from herds in the Norwegian Association of Sheep and Goat Breeders ('Norway', www.nsg.no, Norway), Meredith Dairy ('Aus1', Victoria, Australia), Northland ('NZ1', Northland, New Zealand) and three other New Zealand herds ('NZ2', 'NZ3', 'NZ4'). The Norway population descends from the Norwegian Landrace breed with some recent infusion of French Alpine (Ådnøy 2014). NZ1 is primarily Saanen (Wheeler *et al*. 2018) while Aus1 is a composite of Saanen, Toggenberg and British Alpine (Wheeler *et al*. 2018) with similar likely origins for NZ2, NZ3, NZ4. To approximately balance numbers across the groups, younger animals were removed from some groups (those with birth years from 2018, 2014 and 2017 for Norway, Aus1 and NZ1 respectively).

GBS genotypes. The animals were genotyped by genotyping-by-sequencing (GBS) using the methods described by Dodds *et al.* (2015) and Wheeler *et al*. (2018). Prior to this study, sequence reads from a set of 5,395 goats, that were available at the time, from a range of sources (including 3,702 from Aus1, 1,458 from NZ1 and 201 from NZ3 but none from Norway) were used to detect variants. The variants were discovered using UNEAK (Lu *et al.* 2013) on the adapter-trimmed sequences and without using a reference genome. These variants were placed into a catalogue which was used to allow counts of reference and alternate alleles for each variant in any GBS'd sample using TagDigger (Clark and Sacks, 2016).

Only SNPs that mapped on to the autosomal chromosomes of the goat reference assembly (ARS1, https://www.ncbi.nlm.nih.gov/assembly/GCF_001704415.1) were retained. In addition, SNPs with a raw (not adjusted for read depth) Hardy-Weinberg disequilibrium value < 0.05 (Dodds *et al.* 2015) or a depth adjusted Hardy-Weinberg test p-value < 10^{-100} (Dodds *et al*. 2018a) were

removed. Animals that were genotyped multiple times but had inconsistent genotypes (n=10) or that had a mean read depth < 0.3 (n=191) were removed from the initial set of 8,340 goats.

Population structure. A genomic relationship matrix (GRM) was calculated using the method of Dodds *et al.* (2015) which accounts for the read depth in a genotype call. The overall allele frequencies, calculated on the total number of reads for each allele, were used in these calculations. The GRM was then used to perform a principal component analysis. The mean relatedness by group pair was also calculated and plotted using the heatmap function in R (R core team, 2020), which also performs hierarchical clustering. The fixation index (F_{st}) between groups was calculated by the depth-adjusted method of Dodds *et al*. (2018b) using KGD software (www.github.com/AgResearch/KGD) with default settings.

RESULTS AND DISCUSSION

The SNP catalogue contained 60,225 SNPs. After filtering 51,680 SNPs and 8539 animals remained. These SNPs had a 73.9% call rate and mean read depth of 2.71.

The first two principal components are shown in Figure 1. The first component explains 69.3% of the variance and separates out the Norwegian goats from the others. This is consistent with the Norwegian group being genetically isolated for over 1000 years until 2007. The effect of the recent use of semen from French Alpine goats in Norway can be seen by the small cluster (on the left of the main Norway cluster).

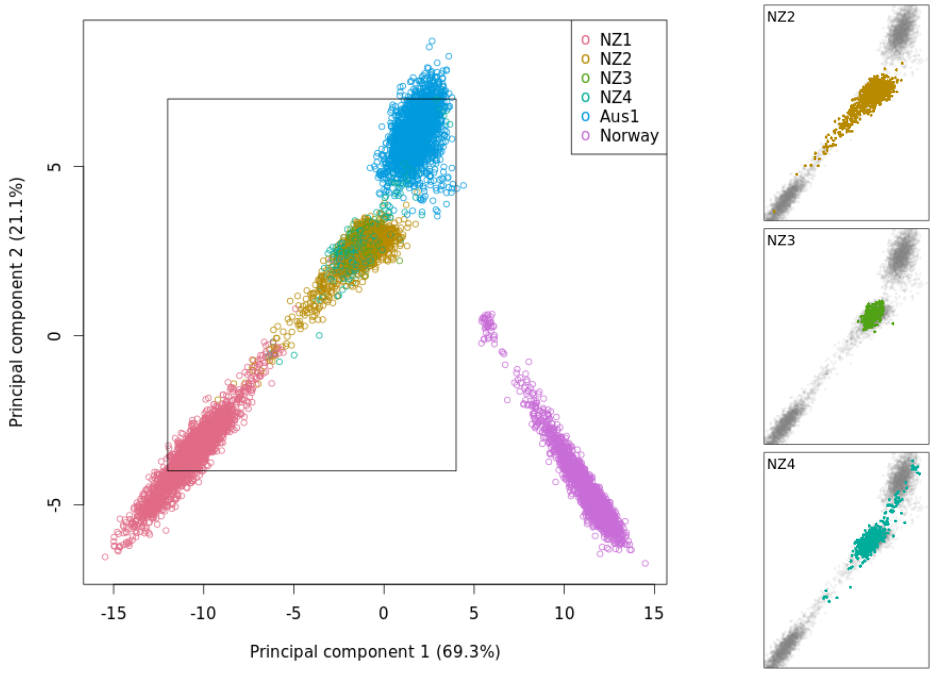

Figure 1. Principal components plot of the GRM coloured by population. Subplots on the right show three of the populations plotted over the others for the region in the box

The Australian and New Zealand populations fall into three partially overlapping groups (Figure 1). These groups almost form a linear arrangement. NZ1 forms a group almost on its own at one end. This population is from Northland where climatic & environmental conditions have necessitated a region-specific breeding programme. Aus1 forms a group at the other end with the

other three NZ populations in between. Some of NZ4 overlaps with Aus1. Both NZ2 and NZ4 have animals near or just overlapping NZ1, while NZ3 is tightly clustered within the central group of NZ populations. The studies of Brito *et al*. (2017) and Oget *et al*. (2019), which also include some alpine dairy goat breeds but using a 50k SNP chip, suggest other populations that could be investigated using a common genotyping platform.

The clustering based on mean population pair relatedness (Figure 2) also indicates that NZ4 and NZ2 are genetically similar, although it clusters Aus1 with this pair before adding NZ3. The Norway population is added last and appears to have the highest within population similarity. This could be due to the isolation of the Norwegian from other goat populations but may also be partly due to Norway being the most outlying population, a minority (~25%) of the data analysed and that SNPs were ascertained in non-Norwegian goats.

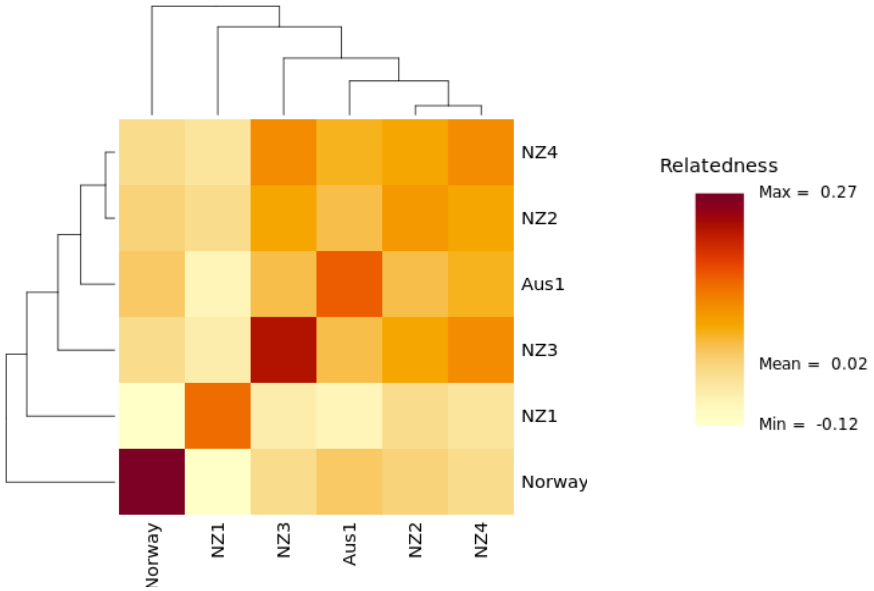

Figure 2. Heatmap plot of the mean GRM relatedness within and between populations

As NZ2, NZ3 and NZ4 appeared to be similar (Figure 1), they were treated as a single population ('NZ234') for the F_{st} analysis. The mean F_{st} for the resulting four populations was 0.066 while the pairwise values are shown in Table 1. These have an expected value, under the null hypothesis of no differentiation, of 1/(mean number of alleles in comparison) ≈ 0.0004 (negligible), as it was estimated using methods in Dodds *et al*. (2018b) that 10,000 alleles were seen (averaged over SNPs) in the 8539 individuals. The F_{st} results were broadly consistent with the relatedness results. For example, the highest F_{st} pair were Norway with NZ1 (F_{st}=0.061, Table 1) and this pair were the most distantly related (Figure 2) while the lowest F_{st} pair were Aus1 with NZ234 (F_{st}=0.018, Table 1) and these groups were the most closely related (Figures 1 and 2).

The SNP minor allele frequencies (MAFs) were also calculated for each of the F_{st} populations and the numbers of SNPs that had no variation (MAF=0) in each group are shown in Table 1. The two most divergent populations (Norway and NZ1) had the highest numbers of MAF=0 SNPs. The high number of non-polymorphic SNPs for Norway could be expected as no animals from Norway were included in the SNP detection process. Even so, there are only 4% of all SNPs used that were

not polymorphic in the Norway population.

Table 1. Mean F_{st} values between pairs of populations

	Number	F_{st} between population pairs				Number of SNPs
Population	of animals	Aus1	Norway	NZ1	NZ234	with MAF=0
Aus1	2199		0.044	0.047	0.018	358
Norway	2107			0.061	0.042	1967
NZ1	2357				0.032	890
NZ234*	1876					122

* NZ234 has 1249, 199 and 428 animals from NZ2, NZ3 and NZ4, respectively

CONCLUSIONS

These results indicate that there is some genetic differentiation between the populations of dairy goats investigated. The Norwegian population appears to be the most divergent, although the introduction of French Alpine into that population has reduced the amount of differentiation. Three New Zealand populations (NZ2, NZ3, NZ4) appear to be quite similar and it is likely that a combined genetic or genomic evaluation of those populations would be useful. If such an evaluation should be widened, it is Aus1 (rather than NZ1) that is most likely to be of benefit. There is unlikely to be much predictive power for genomic evaluations between the Norway population and the other populations studied here.

ACKNOWLEDGEMENTS

Financed by the Norwegian Research Council and the Norwegian Association of Sheep and Goat Breeders through the research project 296561. We are grateful to the breeders and their breeding organisations for access to their data. Some of the methods used in this study were developed by the "Genomics for Production & Security in a Biological Economy C10X1306" project, which was funded by the Ministry of Business, Innovation and Employment (NZ).

REFERENCES

Ådnøy T. (2014) *Small Rumin. Res.* **122**:4.

Brito L.F., Kijas J.W., Ventura R.V., Sargolzaei M., Porto-Neto L.R., Cánovas A., Feng Z., Jafarikia M. and Schenkel F.S. (2017) *BMC Genomics* **18**:229.

Clark L.V. and Sacks E.J. (2016) *Source Code Biol. Med.* **11**:1.

Dodds K.G., McEwan J.C., Brauning R., Anderson R.A., Van Stijn T.C., Kristjánsson T. and Clarke S.M. (2015) *BMC Genomics* **16**:1047.

Dodds K., McEwan J., Bilton T., Brauning R. and Clarke S. (2018a) *Australasian Applied Stat. Conf.*:38 (abstract).

Dodds K.G., Symonds J.E., Brauning R., McEwan J.C. and Clarke S.M. (2018b) *New Zealand Molecular Mapping Workshop (Mapnet)*(abstract).

Lu F., Lipka A.E., Glaubitz J., Elshire R., Cherney J.H., Casler M.D., Buckler E.S. and Costich D.E. (2013) *PLoS Genet.* **9**:e1003215.

Oget C., Servin B. and Palhière I. (2019) *Anim. Genet.* **50**:54.

R Core Team (2020) https://www.R-project.org/.

Wheeler M., Dodds K.G., Brauning R., Clarke S.M., Van Stijn T.C., Foote B.J., Foote J.F., Cameron A.W.N. and McEwan J.C. (2018) *Proc. 11th World Congress Genet. Appl. Livest. Prod.*:358.