SNP-based genetic diversity of South African commercial dairy and fibre goat breeds

S.F. Lashmar[1], C. Visser[1#] & E. van Marle-Köster[1]

[1]Department of Animal and Wildlife Sciences, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa

---

Abstract

Goats (*Capra aegagrus hircus*) have not been a prioritized livestock species with regards to molecular research. The genetic characterization of commercial South African (SA) goat breeds should contribute to improving the management of available animal genetic resources. The aim of this study was to investigate genetic diversity within and among SA commercial goat breeds utilizing the 50k goat beadchip. 88 goats originating from four breeds (dairy: British Alpine, Saanen, Toggenburg; fibre: Angora) were genotyped with the goat SNP50 beadchip. Average MAF values ranged from 0.25 for the Angora to 0.29 for the Saanen, with 46 983 and 50 368 polymorphic SNPs obtained for the respective breeds. Observed heterozygosity values ranged from 0.365 for the Angora to 0.431 for the Toggenburg breed. Linkage disequilibrium (LD) estimation revealed average r^2 values of 0.12 and 0.15 for dairy and fibre breeds, respectively. LD decay was shown to occur after a distance interval of 20-40kb and 40-60kb for dairy and Angora breeds, respectively. Principal component analysis (PCA) produced clusters corresponding to the different production types (dairy and fibre). The Angora, British Alpine and Saanen breeds showed high proportions of membership to respective inferred ancestral populations with ADMIXTURE (97%, 84% and 92%, respectively). The results obtained in this study indicated genetic uniformity within dairy and fibre goats due to production-specific trait selection. Sufficient levels of genetic variation was, however, observed to allow genetic progress for SA commercial goat breeds pending the improved management of these goat genetic resources.

Corresponding author: carina.visser@up.ac.za

Introduction

Goats in South Africa (SA) are one of the less recognised livestock species despite their contribution to agricultural production on different levels. In Africa, and also SA, goats have not been a priority with regards to research and development - ranked only fourth after cattle, sheep and poultry (Simela & Merkel, 2008). The relatively small physical size of the goat makes them easy to handle, therefore making goat production less labour intensive (Aziz, 2010; Devendra & Solaiman, 2010). Furthermore, goats are well adapted to diverse agro-ecological conditions and well suited to both commercial and small-holder farming systems. In SA the three primary goat commodities namely dairy, fibre and meat are produced by 40% of the total goat population consisting of primarily the recognised commercial goat breeds (Directorate: Marketing, 2013; Du Toit *et al*., 2013). The other 60% consists of goat populations that are not yet defined as breeds and are mostly found on small-holder farms or in rural households where they contribute to household food security (Directorate: Marketing, 2013), and also serve cultural and other socio-economic needs (Devendra & Solaiman, 2010).

The SA goat population consists of approximately 6.13 million animals and contributes an estimated 1.7% to the African goat population and 0.6% to the global goat population (FAOSTAT, 2013). The SA goat population is comprised of 60.8%, 24.6%, 14.3% and 0.3% of unimproved or communal meat, commercial meat, Angora and dairy goats, respectively (Du Toit *et al*., 2013). Goats of the respective sectors are in general outcompeted by other ruminant species, and therefore only serve niche markets (Olivier *et al*., 2005; Simela & Merkel, 2008; Visser & Van Marle-Köster, 2014).

Goat dairy products have been promoted as an alternative source of milk for individuals suffering allergies and gastro-intestinal ailments due to cow's milk (Haenlein, 2004), but in South Africa it is

most commonly marketed as goat's cheese for the affluent section of the population. The SA Angora goat is well known as a major role-player in global mohair production and is currently the largest in the world with a production of 2.3 million kg per year (Mohair South Africa, 2012; DAFF, 2013).

The United Nations (UN) has predicted a positive human population growth for SA of approximately 10 million people by the year 2050 (2015: ~53.5 million, 2050: ~63.4 million; United Nations, 2012). The SA urban population is furthermore projected to increase by 13% from 2015 to 2050, while the rural population is predicted to decrease by more than 4 million people during the same time period (United Nations, 2014). It is therefore expected that underutilized livestock species such as goats will have to contribute to fulfil the growing demands for protein.

Genetic research has been extensive for the SA Angora goat, and has included quantitative studies (Visser & Van Marle-Köster, 2009) and molecular research using microsatellites (Visser & Van Marle-Köster, 2009; Visser *et al*., 2010, 2011a, 2011b, 2013) and one SNP-based validation study (Lashmar *et al*., 2015). For SA dairy goats, genetic parameters for milk yield and composition have been estimated (Muller, 2005) and microsatellite markers have been utilized to investigate within and between breed diversity (Bosman *et al*., 2015). Molecular genetic research on SA meat and indigenous breeds has been limited to a few genetic diversity studies using microsatellite markers (Visser *et al*., 2004; Pieters, 2007, Pieters *et al*., 2009) and one SNP-based study by Mdladla *et al*. (2014).

The establishment and availability of molecular tools for goats such as a radiation hybrid map, a reference genome sequence and a 50K SNP chip have created positive prospects for genetic analyses of goats previously not possible. SNP-based approaches provide opportunities to improve on the limitations faced by traditional quantitative studies to accelerate genetic progress. High density SNP panels have been shown to be most useful for analyses of genetic diversity and population structure for goats (Kijas *et al*., 2013) and other livestock species (eg. Bovine HapMap Consortium, 2009; Kijas *et al*., 2009). In this study the aim was to investigate the genetic diversity and population structure of the commercial SA goat breeds, based on genotypes generated using the goat SNP50 beadchip.

Materials and methods

Sampling of animals and phenotypic data

Ethics approval for this study was obtained from the Ethics Committee of the Faculty of Natural and Agricultural Sciences at the University of Pretoria (EC130618-060 & EC104-13). Whole blood samples of 88 goats originating from four SA commercial goat breeds (48 Angora, 20 Saanen, 14 British Alpine and 6 Toggenburg) were analysed. The 48 Angora goat samples included in this study have previously been described by Lashmar *et al*. (2015). These samples originated from four commercial herds located in the Karoo region of SA. The British Alpine and Toggenburg, as well as 15 of the Saanen samples originated from a single commercial herd located in the Western Cape province of SA. The remaining five Saanen samples originated from the University of Pretoria's experimental farm (Hatfield, Gauteng province, South Africa). All the dairy goats were selected on the premise of having sufficient milk production records. Based on available pedigree data, all animals were expected to be unrelated.

SNP genotyping and quality control

DNA extractions were performed at the Animal Breeding and Genetics laboratory of the Department of Animal and Wildlife Sciences at the University of Pretoria. The Qiagen DNeasy Blood and Tissue kit® (Qiagen – Whitehead Scientific [Pty] Ltd, Cape Town, South Africa, www.qiagen.com) was used to extract concentrated DNA following the manufacturer's protocol. The quantity and quality of genomic DNA for all samples were quantified using the Qubit® 2.0 fluorometer at the ARC Biotechnology Platform and, where necessary, was verified using the Nanodrop spectrophotometer (Nanodrop ND-1000). The average DNA concentration obtained over all samples was ~72 ng/μL and a 260:280 of ~1.9.

Genotyping was conducted at the ARC-Biotechnology Platform with the Illumina goat SNP50 beadchip (53 347 SNPs, inter-SNP spacing: ~40kb; Tosser-Klopp *et al.*, 2012). Following SNP genotyping, the processed beadchips were imaged using the Illumina iScan Reader after which the appropriate genotyping data was transferred to Illumina GenomeStudio 1.9.0 software for primary data analysis. PLINK (Purcell *et al.*, 2007) input files, were generated and used to perform quality control. Individuals missing more than 2% of genotypic data were removed (sample call rate<98%). SNPs that had a call rate below 98%, MAF below 5% or violated HWE (P<0.001) were removed from further analysis.

Adjacent pairwise analysis was performed in order to calculate LD estimates (average r^2 and D' values), using the EM algorithm (Excoffier & Slatkin, 1995) in SNP Variation Suite version 7 (SVS; www.goldenhelix.com). All LD estimates were calculated for mapped, autosomal SNP that passed quality control. LD estimates, as average r^2 and D' values, were calculated per chromosome for Angora and dairy breeds. LD estimates were also calculated for different mapping distance intervals namely <10kb, 10-20kb, 20-40kb, 40-60kb, 60-100bk, 100-200kb, 200-500kb and 500kb-1Mb. Summary statistics such as the mean expected (H_E) and observed (H_O) heterozygosity, average individual inbreeding coefficient (F_{IS}) as well as average MAF values were also calculated for breeds. Heterozygosity estimates were estimated across the entire set of SNPs (53 347) as well as across the polymorphic SNPs (MAF>5%), whereas the individual inbreeding coefficient values were estimated across independent SNPs. The --indep-pairwise (50 5 0.2) option in PLINK was used to remove SNPs in high linkage disequilibrium (LD). PLINK uses the EM algorithm (Excoffier & Slatkin, 1995) to calculate the r^2 values between all SNP pairs across all autosomes.

GCTA version 1.24 (Genome-wide Complex Trait Analysis; Yang *et al.*, 2011) was used to estimate SNP-based genetic relatedness between individuals. A genetic relationship matrix was first calculated, after which eigenvalues and eigenvectors were generated for principal component analysis (PCA). The eigenvectors for the first three principal components were then plotted against one another using Microsoft Excel (Microsoft Office, 2013).

ADMIXTURE version 1.23 (Alexander *et al.*, 2009) was used to investigate the population structure of the studied goat breeds, and was run at K-values of 2-5. The cross-validation (CV) procedure was

used in order to choose the optimal number of inferred clusters by identifying the K-value with the lowest cross-validation error estimate. Genesis version 0.2.3 (Buchmann, R. & Hazelhurst, S., University of the Witwatersrand, Johannesburg, SA, http://www.bioinf.wits.ac.za/software/genesis) was used to generate population structure bar plots for the appropriate K-values.

Results

Sample-based quality control resulted in the removal of one animal (British Alpine). Missing genotype rates for the dairy and fibre production types ranged from 0.004-0.017 and from 0.003-0.019, respectively. Average call rates of 99.5 and 99.6% were obtained for these respective production types. A total of 53 347 SNPs, including 1417 unmapped and 1986 X-chromosome linked SNPs, were considered before marker-based quality filtering. More than double the number of SNPs were removed for fibre animals compared to the dairy animals, with the number of polymorphic SNPs (MAF>5%) being the major contributor to this fairly large discrepancy (Table 1).

A larger number of SNPs were observed to be polymorphic in the dairy population in comparison with the Angora population (diary: 51 590, or ~97%; Angora: 46 983, or ~88%). The dairy population, furthermore, had a 10% higher SNP validation rate than the Angora population (92% and 82%, respectively).

Heterozygosity estimates ranged from 0.365 for the Angora breed to 0.431 for the Toggenburg breed. All of the individual breeds, except for the Angora, showed positive gene diversity ($H_E<H_O$). A loss of gene diversity ($H_E>H_O$) was observed for the dairy animals when heterozygosity was estimated across production types, indicating that diversity within breeds was higher than between breeds.

Production type	SNP<98% Call Rate	SNP<5% MAF	SNP<HWE (p<0.001)	Total SNP removed	SNP validated for downstream analysis
Dairy	2211	1757	628	4596	49 227
Fibre	2279	6364	1681	9588	43 759

**Table 1** Results following marker-based quality control for the separate production types

Table 2 Gene diversity and individual inbreeding estimates for individual breeds and production types

Production Type	Breed	Polymorphic loci	H_E*	H_O*	PH_E**	PH_O**	Individual inbreeding coefficient (F_{is})***
Dairy	British Alpine	47 859	0.355	0.385	0.392	0.424	-0.088
	Saanen	50 368	0.373	0.378	0.393	0.399	-0.015
	Toggenburg	47 561	0.339	0.385	0.379	0.431	-0.135
	Total	51 590	0.398	0.382	0.410	0.393	0.024
Fibre	Angora	46 983	0.333	0.324	0.371	0.365	0.009
Total		**52 177**	**0.405**	**0.350**	**0.414**	**0.357**	**0.078**

* Estimated across all 53 347 SNPs (before quality control), ** Estimated across polymorphic loci (MAF>5%), *** Estimated across independent SNPs (after LD filtering)

Individual inbreeding estimates supported the gene diversity that was observed per production type, and showed positive F_{IS} values for the Angora and dairy production group that showed a loss of heterozygosity. When considering the four individual breeds, the Angora breed was shown to be the most inbred even though F_{IS} values were insignificantly low positive whereas the Toggenburg was shown to be the least inbred. Comparison between the two production groups, however, indicated a higher average F_{IS} value for the combined dairy animals. The animal with the highest individual

inbreeding (0.019) was from the Saanen population, whereas the animal with the lowest individual inbreeding (-0.220) was from the Toggenburg population. SNPs were partitioned into 10% MAF intervals ranging from 0-50% for each individual breed (Figure 1).
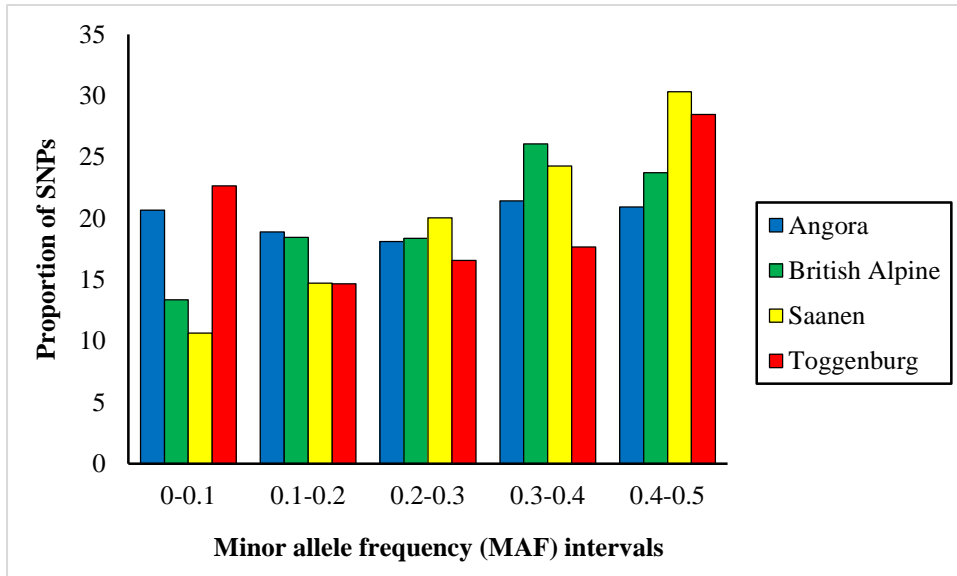
Figure 1 MAF distribution for each breed

The average MAF for the breeds varied little, with a range from 0.25 for the Angora goats to 0.29 for Saanen goats. In Figure 1 it is indicated that the Angora breed had a relatively high proportion of SNPs in the lowest MAF interval compared to the proportions of SNPs observed in the higher MAF intervals. The proportions of SNPs for the British Alpine breed were found to be higher with increasing MAF intervals and showed a decrease at the last interval. For the Saanen breed, linear growth in proportion of SNPs was observed and no real pattern was observed for the Toggenburg breed.

LD, r^2 and D', was estimated on a per-chromosome basis and are summarized in Table 3. Samples were separated into dairy and fibre breeds in order to draw comparisons between the statuses of linkage for different production types.

Table 3 A summary of LD estimates per chromosome for the two production groups

| Production type	Average number of SNPs per chromosome	Average r^2	Average D'
Dairy	1589	0.12	0.481
Fibre	1442	0.15	0.563

The results for the production types are in agreement with regards to the chromosomes with the lowest and highest number of SNPs (CHI 25 and 1, respectively). The LD estimates for the Angora breed indicated higher r^2 and D' values compared to the dairy breeds. LD estimates were estimated for increasing distance intervals in order to observe the effect of increasing distance between consecutive SNPs on these estimates, and the results are reported in Table 4.

Table 4 Mean LD between mapped, autosomal SNPs for different map distances

| Distance Intervals (kb)	Dairy			Fibre		
	SNP pairs	Average r^2	Average D'	SNP pairs	Average r^2	Average D'
<10	222	0.092	0.454	155	0.161	0.574
10-20	303	0.111	0.458	220	0.147	0.553
20-40	21 826	0.117	0.473	15 081	0.154	0.569
40-60	18 313	0.114	0.472	12 784	0.155	0.567
60-100	7 655	0.113	0.466	8 570	0.148	0.560
100-200	1 535	0.111	0.463	3 710	0.126	0.531
200-500	61	0.147	0.473	454	0.111	0.508
500-1Mb	-	-	-	14	0.085	0.633

Table 4 indicates that the majority of SNP pairs had a distance of between 20-40kb between consecutive SNPs. r^2 estimates peaked at the 40-60kb distance interval for the Angora breed, whereas these estimates peaked at the 20-40kb distance interval for the dairy production group. D' estimates peaked at the 20-40kb distance interval for both production groups.

The largest principal components (PCA1, PCA2 and PCA3) were plotted against one another and the resulting plots can be seen in Figure 2.
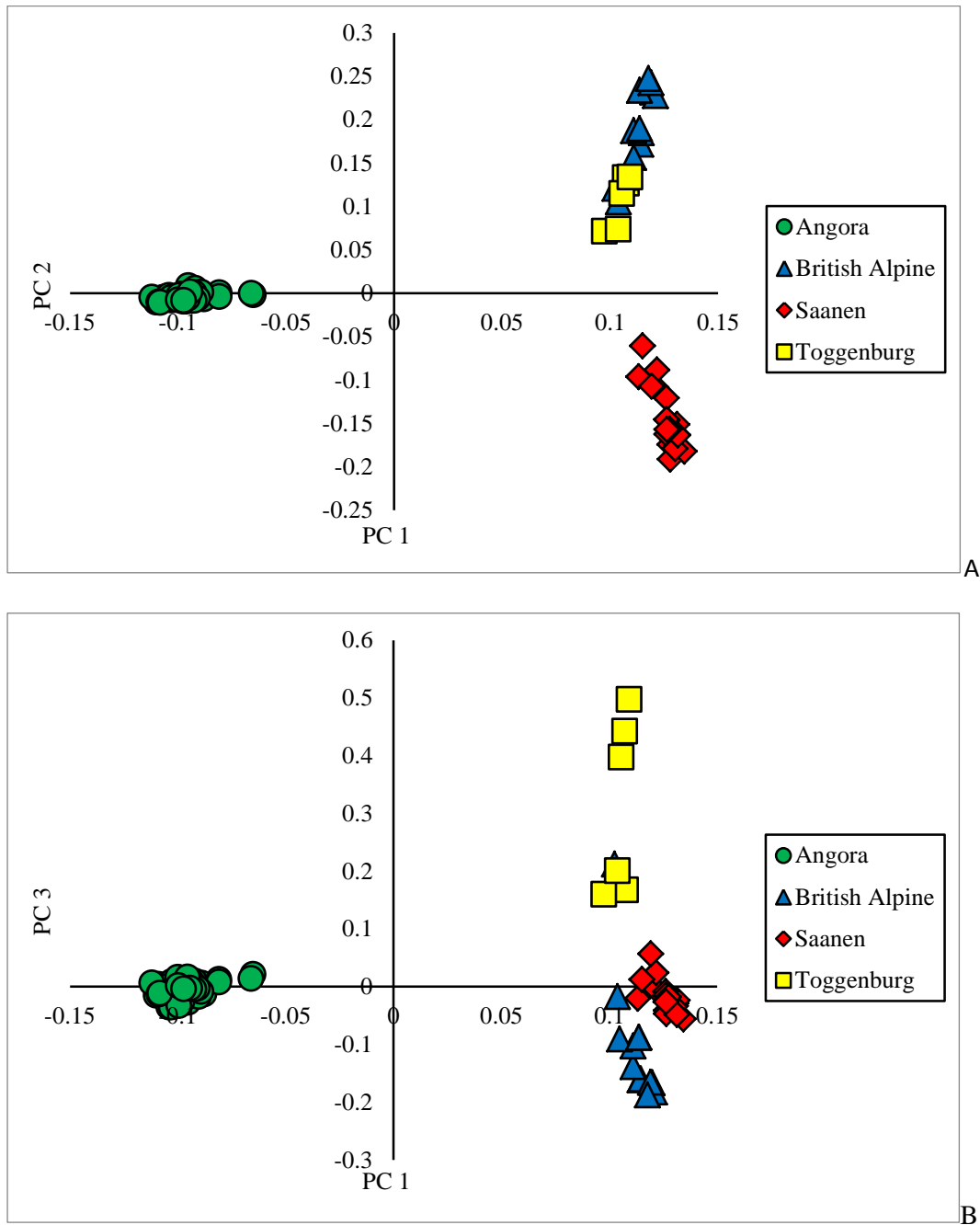
Figure 2 The genetic relationships among the 87 goats as seen when plotting **A**) the first and second principal components (PCA1 and PCA2, left) and **B**) the first and third principal components (PCA1 and PCA3, right) against one another

Figure 2a illustrates that the animals were clustered according to their appropriate production types (dairy and fibre). The Angora breed (green dots) formed the most defined cluster. The dairy animals formed two separate clusters in close proximity to one another, with the Saanen animals forming the first cluster and the British Alpine and Toggenburg animals combining into a second cluster. Figure 2b illustrates that by incorporating the third largest principal component (PCA3), the Angora animals maintained tight clustering whereas higher levels of dispersion were indicated in the dairy goat clusters. In Figure 2b the Toggenburg animals (yellow squares) tended to deviate from the British Alpine animals (blue triangles), however, these animals still did not form a restricted cluster.

The most probable number of inferred populations was chosen as three based on cross-validation error estimation and the resulting population structure bar plots for K=2-4 can be seen in Figure 3.
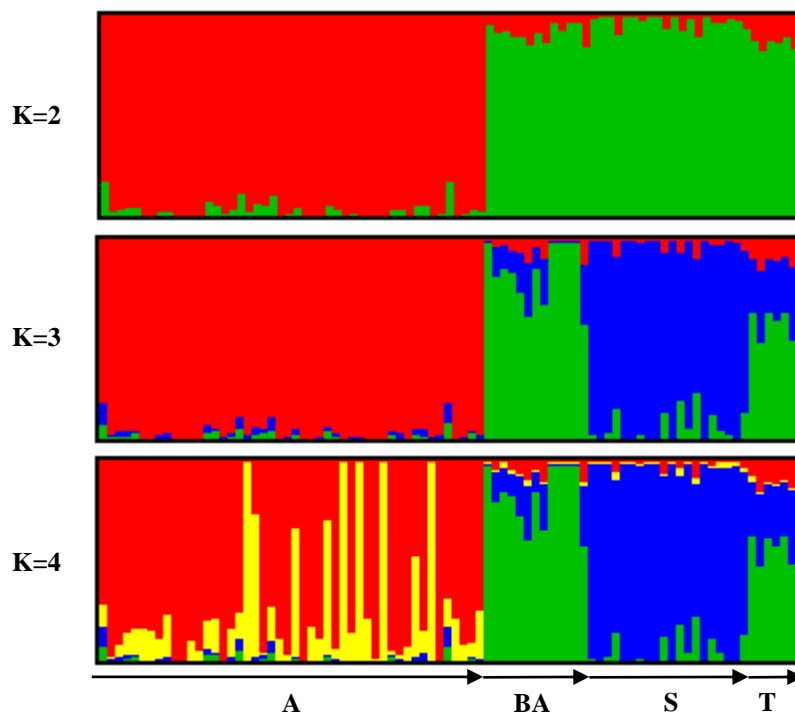
Figure 3 Population structure plots showing the proportions of ancestral populations for each individual (A = Angora, BA = British Alpine, S = Saanen and T = Toggenburg) for K=2 to K=4

ADMIXTURE results illustrated distinct ancestral backgrounds for the two production types at K=2. The incorporation of an additional ancestral population (K=3), divided the dairy production group and illustrated two separate ancestral populations – the Saanen animals predominantly originating from

the first, and the other two breeds originating predominantly from the second. The Toggenburg breed had the highest level of admixture, while the lowest levels of admixture can be observed for the Angora and meat breeds. When four ancestral populations were assumed (K=4), different ancestries for the subpopulations, or different herds, that were sampled for the Angora animals were illustrated.

Discussion

Genetic characterization and investigating the genetic diversity and admixture of commercial goat breeds in SA, and across Africa, will assist in the sustainable management of genetic resources and conservation of unique, adaptable characteristics of the local animals. Genetic variation, amongst other factors such as selection intensity, accuracy of selection and trait heritability; directly influences genetic progress, and therefore the genetic characterization of commercial goat breeds serves as an essential starting point (Bourdon, 2000).

Heterozygosity values obtained in the current study were comparable with values obtained by Mdladla *et al*. (2014) for indigenous SA goats (mean H_E: 0.42±0.09, mean H_O: 0.38±0.09). Similar to the latter authors, a slight loss of heterozygosity ($H_O < H_E$) was also observed in the current study across dairy and fibre production types. Positive heterozygosity estimates for the individual dairy breeds indicated high gene diversity within these breeds, or at least between the animals sampled in this specific study. Due to limited sample size, results for specifically the Toggenburg breed should be interpreted with caution. Due to the similarity between H_E and H_O values for the Angora, it is reasonable to assume that there is still sufficient genetic diversity within this breed to allow the genetic improvement of these important fibre-producing goats in SA. Ascertainment bias influences diversity estimates such as H_E that are dependent on allele frequency (Clark *et al*., 2005), and the presence of non-discovery breeds in the current study (Angora and Toggenburg) was expected to influence this parameter. A microsatellite-based study by Bosman *et al*. (2015) observed an insignificantly small loss of heterozygosity for SA Saanen and British Alpine breeds. The

heterozygosity estimates obtained in this study were higher than values previously reported for Angora (H_E: 0.233, H_O: 0.285) goats in a SNP-based study by Kijas *et al*. (2013).

The number of polymorphic SNPs for the Angora (46 983) was higher than the 33 597 polymorphic SNPs obtained by Kijas *et al*. (2013), and furthermore the number of validated SNPs for this breed (43 759) was lower than the number obtained by Roldán *et al*. (2014) (47 359) for Angora x Creole backcrossed goats. Other non-discovery breeds have been found to have comparable numbers of polymorphic loci (Jinlan: 45 648, Skopelos: 50 908, Spanish Florida: 49 156) to the Angora breed (Tosser-Klopp *et al.*, 2014; Zidi *et al*., 2014). The level of polymorphicity obtained for the dairy breeds (British Alpine: 47 859, Saanen: 50 368, Toggenburg: 47 561) were higher than obtained for Canadian populations (British Alpine: 45 338, Saanen: 45 010, Toggenburg: 33 761; Brito *et al*., 2014), but slightly lower than obtained in the goat SNP50 beadchip development study (British Alpine: 51 339, Saanen: 51 689; Tosser-Klopp *et al*., 2014). The relatively large discrepancy between the numbers of polymorphic loci for SA goat breeds, such as the Toggenburg, compared to these breeds originating from other countries (eg. Brito *et al.* (2013): Canada) is possibly attributable to large sample size differences or the fact that these are different subpopulations of the same breed that are geographically isolated from one another.

The distribution of SNPs within different MAF intervals, as well as the average MAF values obtained per breed, highlighted the marker-based differences between the breeds included in the study. The high proportion of SNPs with low MAF values (<10%) observed for the Angora breed compared to the dairy breeds, may indicate a higher proportion of fixed alleles within this populations due the association of SNPs with specific traits under intensive selection. The SA dairy goat population has generally not been subjected to intensive artificial selection practices (Bosman *et al*., 2015), and therefore higher mean MAF values were expected. The higher MAF values obtained for the Saanen and British Alpine breeds compared to the Toggenberg breed can be attributed to possible ascertainment bias. The Toggenburg breed also had a smaller sample size and all sampled animals originated from a single herd, and the lower MAF might therefore not be an accurate representation of all SA Toggenburg animals.

The extent of LD across the genome determines the SNP density necessary for the accurate estimation of GEBV for genomic selection as well as the value of predicting QTL effects in GWAS studies (Qanbari *et al.*, 2010; García-Gámez *et al.*, 2012). LD is influenced by factors such as population history, breeding systems and the pattern of geographic subdivision (Ardlie *et al.*, 2002), and estimates were therefore expected to differ between the sampled breeds. The r^2 measure is known to be a more robust parameter for bi-allelic marker analysis, and therefore requires a smaller sample size for accurate LD estimate estimation than D' (Ardlie *et al.*, 2002; Khatkar *et al.*, 2008; Lipkin *et al.*, 2009). Due to the fact that breed populations had relatively small sample sizes, only average r^2 will be discussed further.

Average r^2 values of 0.12 and 0.15 were obtained for the dairy and fibre breeds, respectively. Average r^2 values for the dairy breeds were slightly lower than average r^2 values of 0.14, 0.15 and 0.24 found for British Alpine, Saanen and Toggenburg breeds, respectively by Brito *et al.* (2014). Average r^2 values were also lower than an average value of 0.17 found by Carillier *et al.* (2013) across a French multi-breed (Alpine and Saanen) population. Angora animals (0.15) also had comparable r^2 values to the studies mentioned before. LD estimates for the dairy and fibre goats in this study were lower than obtained for meat goats in previous studies (Boer: 0.29; Brito *et al.*, 2014). SNP densities of 1 SNP/~59kb and 1 SNP/~226kb have been estimated for Angora goats before and after LD filtering, respectively (Lashmar *et al.*, 2015), and were found to be higher than estimates obtained for British Alpine and Saanen breeds, but lower than obtained for Boer and Toggenburg breeds by Brito *et al.* (2014).

Research has suggested an r^2 value of 0.2 to be sufficient for the application of genomic selection (Meuwissen *et al.*, 2001; Calus, 2010), implying that genomic selection is not feasible for SA goats given the resources currently available. LD has been shown to decay with increasing inter-marker distances, and therefore denser SNP genotyping arrays are necessary to incorporate sufficient association between SNPs for the application of GS in SA Angora goats and dairy breeds.

PCA and ADMIXTURE results were generally in agreement and classified the animals according to the appropriate production types that the respective breeds belong to. It was interesting to note that the inference of K=4 did not assign a separate ancestral population for the Toggenburg breed as might

have been expected, and this was a verification of the fact that the Toggenburg animals sampled were an admixture of the ancestral populations inferred for the Saanen (30%) and British Alpine (58%) animals, respectively. The proportion of membership observed for the Angora, British Alpine and Saanen breeds to their respective clusters (97%, 84% and 92%, respectively) were comparable with a range of 81-97% obtained by Makina *et al*. (2014) for six SA cattle breeds.

For the Angora breed, well-defined PCA clustering and high proportions of shared ancestry within the breed were expected for various reasons. The Angora is the sole mohair-producing goat breed and therefore represents a unique goat genetic resource in SA. Shared SNP genotypes associated with mohair traits will therefore inevitably result in close genetic relationships between animals within this breed. The results obtained were in concordance with observations made by Kijas *et al*. (2013) based on goat SNP genotyping data. The latter authors observed that Australian cashmere goats (fibre-producing) were genetically distinct from Rangeland (indigenous) and Boer (meat) goats, attributing the unique genetic profile of these goats to genes contributing to fibre production and morphology. The fact that the different herds sampled separated at K=4 onwards, indicated the existence of genetically isolated herds.

Even though the dairy breeds clustered together, presumably because all these breeds are selected for milk production and quality traits, there was clearly more hybridization between these breeds. A higher level of diversity was expected for dairy breeds as there is a tendency in SA to keep mixed herds (Bosman *et al*., 2015). With 92% shared co-ancestry within the Saanen breed, the animals sampled from this breed were clearly more purebred. There have been reports of crossbreeding between Toggenburg and SA "milch" goats that originated from the Saanen breed, for improved milk production (www.milkgoats.co.za/milkgoats_society/); and this would explain the 30% co-ancestry between these breeds. According to the South African Milk Breeders' Society, dark variations of Toggenburg animals were mated in an attempt to imitate and develop the British Alpine breed in SA during the early 1900's. Purebred British Alpine goats were later (between 1924 and 1934) imported into the country from England (www.milkgoats.co.za/milkgoats_society/). Crossbreeding between dark Alpine-like Toggenburg variants and pure British Alpine animals would explain the gene flow between these breeds. The afore-mentioned results are in partial agreement with microsatellite-based

results obtained by Bosman *et al.* (2015) which revealed a crossbred population using STRUCTURE (Pritchard *et al.*, 2000; Falush *et al.*, 2003) analysis, existing of animals with membership from all three the dairy breeds sampled in this study. The latter authors did, however, also observe separate populations for the British Alpine and Toggenburg breeds; but it should be kept in mind that larger sample sizes with animals sampled from several geographically separated herds, as opposed to one, were studied.

Conclusion

Improved production of commercial goat breeds occurring in SA, in comparison with unimproved indigenous breeds, is a result of artificial selection practices and has been achieved at the risk of losing genetic diversity. Selection for production-specific traits in dairy and fibre-producing goats has led to genetic uniformity observed within the respective production types. Model-based genetic diversity analyses (PCA and ADMIXTURE) revealed the dairy and fibre production types to be distinct from one another and furthermore to originate from separate ancestral populations. The dairy group seemed to be the most genetically complex production type and this might be a reflection of unscientific and unverifiable history records, with regards to the importation of individual breeds into SA, as well as the presence of mixed herds and lack of performance recording. The status of LD in the sampled individuals generally deemed necessary higher SNP densities for applications such as genomic selection in SA commercial goat breeds. Genetic variation is, however, sufficient to allow progress to be made for goats in SA pending the simultaneous improvement of performance recording and other management issues.

Acknowledgments

References

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Research 19, 1655-1664.

Ardlie, K.G., Kruglyak, L., Seielstad, M., 2002. Patterns of linkage disequilibrium in the human genome. Nature Reviews Genetics 3, 299-309.

Aziz, M.A., 2010. Present status of the world goat populations and their productivity. Lohmann Information 45, 42-52.

Bosman, L., van Marle-Köster, E., Visser, C., 2015. Genetic diversity of South African dairy goats for genetic management and improvement. Small Ruminant Research 123, 224-231.

Bourdon, R.M., 2000. Understanding Animal Breeding, second ed. Prentice-Hall Inc., New Jersey.

Bovine HapMap Consortium, 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science 324, 528-532.

Brito, L.F., Jafarikia, M., Grossi, D.A., Maignel, L., Sargolzaei, M., Schenkel, F.S., 2014. Characterization of linkage disequilibrium and consistency of gametic phase in Canadian goats. In: *Proceedings of 10th World Congress of Genetics Applied to Livestock Production.* Vancouver, BC, Canada, 2014. Poster presentation, 882.

Buchmann, R., Hazelhurst, S., 2014. Genesis Manual. Available at URL: http://www.bioinf.wits.ac.za/software/genesis/Genesis.pdf.

Calus, M.P.L., 2010. Genomic breeding value prediction: methods and procedures. Animal 4, 157-164.

Campbell, Q.P., 2003. The origin and description of southern Africa's indigenous goats. South African Journal of Animal Science 4, 18-22.

Carillier, C., Larroque, H., Palhière, I., Clément, V., Rupp, R., Robert-Granié, C., 2013. A first step towards genomic selection in multi-breed French dairy goat population. Journal of Dairy Science 96, 1-12.

Casey, N.H., Webb, E.C., 2010. Managing goat production for meat quality. Small Ruminant Research, DOI:10.1016/j.smallrumres.2009.12.047.

Clark, A.G., Hubisz, M.J., Bustamante, C.D., Williamson, S.H., Nielsen, R., 2005. Ascertainment bias in studies of human genome-wide polymorphism. Genome Research 15, 1496-1502.

DAFF (Department of Agriculture, Forestry, Fisheries), 2013. Trends in the Agricultural Sector, pp. 68–69.

Devendra, C., Solaiman, S.G., 2010. Perspectives on goats and global production. In: Goat science and production. Ed. Solaiman, S.G. Wiley-Blackwell, Ames, Iowa, USA. pp. 3-20.

Directorate: Marketing, 2013. A profile of the South African goat market value chain. Department of Agriculture Forestry and Fisheries, Pretoria, South Africa.

Du Toit, C.J.L., Van Niekerk, W.A., Meissner, H.H., 2013. Direct greenhouse gas emissions of the South African small stock sectors. South African Journal of Animal Science 43, 240-361.

Excoffier, L., Slatkin, M., 1995. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Molecular Biology and Evolution 12, 921-927.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of populationstructure: extensions to linked loci and correlated allele frequencies.Genetics 164, 1567-1587.

FAOSTAT, 2013. http://faostat3.fao.org/home/index.html#DOWNLOAD accessed on 26 August 2014.

García-Gámez, E., Sahana, G., Gutiérrez-Gil, B., Arranz, J.J., 2012. Linkage disequilibrium and inbreeding estimation in Spanish Churra sheep. BMC Genetics 13, 43.

Haenlein, G.F.W., 2004. Goat milk in human nutrition. Small Ruminant Research 51, 155-163.

Khatkar, M.S., Nicholas, F.W., Collins, A.R., Zenger, K.R., Cavanagh, J.A.L., Barris, W., Schnabel, R.D., Taylor, J.F., Raadsma, H.W., 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics 9, 187.

Kijas, J.W., Townley, D., Dalrymple, B.P., Heaton, M.P., Maddox, J.F., McGrath, A., Wilson, P., Ingersoll, R.G., McCulloch, R., McWilliam, S., Tang, D., McEwan, J., Cockett, N., Oddy, V.H., Nicholas, F.W. & Raadsma, H., 2009. A Genome Wide Survey of SNP Variation Reveals the Genetic Structure of Sheep Breeds. PLoS One 4, e4668.

Kijas, J.W., Ortiz, J.S., McCulloch, R., James, A., Brice, B., Swain, B., Tosser-Klopp, G., International Goat Genome Consortium, 2013. Genetic diversity and investigation of polledness in divergent goat populations using 52 088 SNPs. Animal Genetics 44, 325-335.

Lashmar, S.F., Visser, C., Van Marle-Köster, E., 2015. Validation of the 50k Illumina goat SNP chip in the South African Angora goat (Short communication). South African Journal of Animal Science 45, 56-59.

Lipkin, E., Straus, K., Stein, R.T., Bagnato, A., Schiavini, F., Fontanesi, L., Russo, V., Medugorac, I., Foerster, M., Sölkner, J., Dolezal, M., Medrano, J.F., Friedmann, A., Soller, M., 2009. Extensive long range and non-syntenic linkage disequilibrium in livestock populations: deconstruction of a conundrum. Genetics 181, 691-699.

Makina, S.O., Muchadeyi, F.C., van Marle-Köster, E., MacNeil, M.D., Maiwashe, A., 2014. Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel. Frontiers in Genetics 5, 1-7.

Mdladla, K., Dzomba, E.F., Muchadeyi, F.C., 2014. Genetic diversity, population structure and relationships of South African indigenous goat populations using genome-wide SNP markers. In *Programme and Abstracts of the SA Genetics Society (SAGS) & SA Society of Bioinformatics (SASBi) Joint Congress.* Kwalata Game Ranch, Hammanskraal, Gauteng, SA, 23-26 September 2014. Oral session, S4-9.

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819-1829.

Mohair South Africa, 2012. Mohair Review, http://www.mohair.co.za/wp-content/uploads/2013/11/Mohair Review 2012.pdf (accessed 15.10.14).

Muller, C.J.C., 2005. Genetic parameter estimation and breeding plans for the South African dairy goat herd. Ph.D. Thesis. Department of Animal Sciences, Faculty of Agricultural and Forestry Sciences, University of Stellenbosch. Stellenbosch.

Olivier, J.J., Cloete, S.W.P., Schoeman, S.J., Muller, C.J.C., 2005. Performance testing and recording in meat and dairy goats. Small Ruminant Research 60, 83-93.

Pieters, A., 2007. Genetic characterization of commercial goat populations in South Africa. MSc Thesis. Department of Animal and Wildlife Sciences, Faculty of Natural and Agricultural Sciences, University of Pretoria. Pretoria.

Pieters, A., Van Marle-Köster, E., Visser, C., Kotze, A., 2009. South African developed meat type goats: a forgotten animal genetic resource? Animal Genetic Researc 44, 33-43.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotye data. Genetics 155, 945-959.

Purcell, S., Neale, B., Todd-Brown, K., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics 81, 559-575.

Qanbari, S., Pimentel, E.C., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A.R., Simianer, H., 2010. The pattern of linkage disequilibrium in German Holstein cattle. Animal Genetics 41, 346–356.

Roldán, D.L., Debenedetti, S., Cano, E.M., Taddeo, H.R., Poli, M.A., 2014. Preliminary refined localization of QTL for fleece traits in five goat chromosomes using SNP markers in a backcross population. In:

Proceedings of 10th World Congress of Genetics Applied to Livestock Production. Vancouver, BC, Canada, 2014. Poster presentation, 885.

Simela, L., Merkel, R., 2008. The contribution of chevon from Africa to global meat production. Meat Science 80, 101-109.

South African Milch Goat Breeders' Society, 2013. Breeds. Available at URL: http://www.milkgoats.co.za/milkgoat_society/index.php?p=59 accessed on 15 September 2014.

Tosser-Klopp, G., Bardou, P., Cabau, C., International Goat Genome Consortium, 2012. Goat genome assembly, availability of an international 50 K SNP chip and RH panel: an update of the International Goat Genome Consortium projects. In: *International Plant & Animal Genome Conference*, January 2012, San Diego, CA.

Tosser-Klopp, G., Bardou, P., Bouchez, O., Cabau, C., Crooijmans, R., Dong, Y., Donnadieu-Tonon, C., Eggen, A., Heuven, H.C.M., Jamli, S., Jiken, A.J., Klopp, C., Lawley, C.T., McEwan, J., Martin, P., Moreno, C.R., Mulsant, P., Nabihoudine, I., Pailhoux, E., Palhiere, E., Rupp, R., Sarry, J., Sayre, B.L., Tircazes, A., Wang, J., Wang, W., Zhang, W., International Goat Genome Consortium, 2014. Design and Characterization of a 52K SNP Chip for Goats. PLoS One 9, e86227.

United Nations (UN), 2012. World Population Prospects: The 2012 Revision. Department of Economic and Social Affairs, Population Division. Available at URL: http://populationpyramid.net/south-africa/2050/#.

United Nations (UN), 2014. World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352). Department of Economic and Social Affairs, Population Division.

Visser, C., Hefer, C.A, Van Marle-Köster, E., Kotze, A., 2004. Genetic variation of three commercial and three indigenous goat populations in South Africa. South African Journal of Animal Science 34, 24-27.

Visser, C., Van Marle-Köster, E., 2009. Genetic variation of the reference population for quantitative loci research in South African Angora goats. Animal Genetic Resources Information 45, 113-119.

Visser, C., Crooijmans, R.P.M.A., Van Marle-Köster, E., 2010. A genetic linkage map for the South African Angora goat. Small Ruminant Research 93, 171–179.

Visser, C., Crooijmans, R.P.M.A., Bovenhuis, H., Van Marle-Köster, E., 2011a. QTL for mohair traits in South African Angora goats. Small Ruminant Research 100, 8-14.

Visser, C., Van Marle-Köster, E., Friedrich, H., 2011b. Parentage verification of South African Angora goats, using microsatellite markers. South African Journal of Animal Science 41, 3, 250–255.

Visser, C., Van Marle-Köster, E., Snyman, M.A., Bovenhuis, H., Crooijmans, R.P.M.A., 2013. Quantitative trait loci associated with pre-weaning growth in South African Angora goats. Small Ruminant Research 112, 15-20.

Visser, C., Van Marle-Köster, E., 2014. Strategies for the genetic improvement of South African Angora goats. Small Ruminant Research, http://dx.doi.org/10.1016/j.smallrumres.2014.01.012.

Webb, E.C., Casey, N.H., Simela, L., 2005. Goat meat quality. Small Ruminant Research 60, 153-166.

Yang, J., Hong Lee, S., Goddard, M.E., Visscher, P.M., 2011. GCTA: A Tool for Genome-wide Complex Trait Analysis. The American Journal of Human Genetics 88, 76-82.

Zidi, A., Abo-Shady, H., Molina, A., Menéndez-Buxadera, A., Sánchez-Rodríguez, M., Díaz, C., Carabaño, M.J., Serradilla, J.M., 2014. Genome wide association for heat stress tolerance/susceptibility in Florida dairy goats. In: *Proceedings of 10th World Congress of Genetics Applied to Livestock Production*. Vancouver, BC, Canada, 2014. Oral presentation, 340.