ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12776

Genetic diversity and population structure in multiple Chinese goat populations using a SNP panel

H. Berihulay*†, Y. Li*†, X. Liu*†, G. Gebreselassie*†, R. Islam*†, W. Liu*†, L. Jiang*† and Y. Ma*†

*Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China. †The Key Laboratory for Farm Animal Genetic Resources and Utilization of Ministry of Agriculture of China, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China.

Summary

Information about genetic diversity and population structure among goat breeds is essential for genetic improvement, understanding of environmental adaptation as well as utilization and conservation of goat breeds. Here, we measured genetic diversity and population structure in multiple Chinese goat populations, namely, Nanjiang, Qinggeda, Arbas Cashmere, Jining Grey, Luoping Yellow and Guangfeng goats. A total of 193 individuals were genotyped for about 47 401 autosomal single nucleotide polymorphisms (SNPs). We found a high proportion of informative SNPs, ranging from 69.5% in the Luoping Yellow to 93.9% in the Jining Grey goat breeds with an average mean of 84.7%. Diversity, as measured by expected heterozygosity, ranged from 0.371 in Luoping Yellow to 0.405 in Jining Grey goat populations. The average estimated pair-wise genetic differentiation (F_{ST}) among the populations was 8.6%, ranging from 0.2% to 16% and indicating low to moderate genetic differentiation. Principal component analysis, genetic structure and phylogenetic tree analysis revealed a clustering of six Chinese goat populations according to geographic distribution. The results from this study can contribute valuable genetic information and can properly assist with within-breed diversity, which provides a good opportunity for sustainable utilization of and maintenance of genetic resource improvements in the Chinese goat populations.

Keywords China, genetic structure, genetic variation, goat breeds, SNP analysis

Introduction

Archaeological and genetic evolution evidence suggests that the domestic goat was one of the first ruminant animals to be domesticated around 10 000 years ago (Zeder 2000) at the dawn of the Neolithic period in the Fertile Crescent. Goats were domesticated from bezoars (*Capra aegagrus*) in western Asia (Naderi *et al.* 2008). For the past few decades, these domestic breeds have represented an important animal genetic resource and have aroused great attention for the development of sustainable breed improvement strategies. Goat is the most special, geographically widespread livestock species for different weather conditions (Aziz 2010) and is an important agricultural production resource used for meat, milk, skins and fiber for the livelihood of low-input production systems (Devendra 2015).

The biological diversity of the local goat population serves as the basis for the existence and development of raw material for millions of farmers wanting to improve their breeds to adapt to the changing environments and changing demands (Devendra 2015). Hence, consideration of the genetic diversity and conservation of diverse goat populations is an imperative aspect of facing future challenges, such as climate change, emerging diseases and food security for an increasing human population, and is important for the conservation and utilization of the indigenous goat breeds (Groeneveld *et al.* 2010).

To properly maintain indigenous goat populations, the Chinese government has gradually cultivated ecological adaptation (Song *et al.* 2016) and systematic selection (Guan *et al.* 2016). China is home to over 69 goat breeds, including 58 local breeds, eight improved breeds and three introduced breeds, and many of them are listed in the national goat breed directory (Zhang *et al.* 2012; Guan *et al.* 2016; Ma *et al.* 2017). However, genetic diversity of goats and the relationship among the goat populations have not yet been well studied using

Address for correspondence

L. Jiang and Y. Ma, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China.
E-mails: jianglin@caas.cn (L. Jiang) and mayuehui@caas.cn (Y. Ma)

Accepted for publication 16 January 2019

single nucleotide polymorphism (SNP) technology. Most previous studies that were focused on genetic diversity and structures of Chinese goat populations used low-density microsatellite markers (Di *et al.* 2011; Wei *et al.* 2014) and mitochondrial DNA (Wang *et al.* 2008). Nevertheless, today the analysis of SNP markers is becoming a standard approach for diversity analysis and genome-wide studies. The availability of SNPs allows for comprehensive investigation and provides superior power for the examination of genetic variation and the relationship across many diverse livestock genomes that is impossible to attain with other types of marker panels (Lin *et al.* 2013; Lashmar *et al.* 2016; Onzima *et al.* 2018). SNPs represent one of the more interesting approaches for genotyping because they are abundant in the genome, genetically stable and amenable for conducting genome-wide studies (Brito *et al.* 2017) and whole genome re-sequencing analysis (Song *et al.* 2016). The importance of SNP array technology in measuring genetic diversity and divergence within and among populations has been demonstrated in several studies (Kijas *et al.* 2012; Grasso *et al.* 2014).

A detailed understanding of genome-wide SNP panels will provide an opportunity to apply them to genome-based association studies and complex traits in the future. Despite the many SNPs identified in the International Goat Genome Consortium project, few have been validated in Chinese goat populations. Breed characterization requires basic knowledge of genetic variations that can be effectively measured within and between populations. Therefore, in this study we investigated the genetic diversity and population structure of six Chinese goat populations using a SNP panel generated from the Illumina goat SNP50 BeadChip.

Materials and methods

Sample collection

Samples were collected from a total of 193 goats representing the Nanjiang (NJ), Qinggeda (QGD), Arbas Cashmere (AC), Jining Grey (JN), Luoping Yellow (LP) and Guangfeng (GF) goat populations, living mainly in five Chinese provinces (Fig. 1). Summary of breed name, sample code and sample size of the goat populations are depicted in Table S1. To minimize the likelihood of studying related individuals, samples were collected from different areas in the provinces to capture a representative sample of within-breed genetic diversity. A total of 10 ml of blood was drawn from the jugular vein into a tube. DNA extraction was conducted at the Institute of Animal Science, Chinese Academy of Agricultural Sciences laboratory using Promega kits, following the manufacturer's protocol.

Genotyping, quality control and data pruning

All DNA samples were genotyped using the Illumina goat SNP 50K BeadChip by GeneSeek/Neogen, featuring over 53 347 SNPs. The high-quality SNPs were obtained using PLINK v.1.07 (Purcell *et al.* 2007) with the parameters –mind, 0.1; –geno, 0.1; –maf, 0.05; and –hwe, 1e-5 and then were used for further analysis.

Data analysis

Genetic diversity analysis

PLINK v1.07 (Purcell *et al.* 2007) was used for the estimation of genetic diversity measures: inbreeding coefficient (F_{IS}), average relatedness (PI_HAT), proportion of polymorphic SNPs (P_N), minor allele frequency (MAF), observed (H_O) and expected (H_E) heterozygosity across all the populations. H_o and H_e were calculated using the --hardy flag, and the inbreeding coefficient (F_{IS}) and MAF were calculated using the commands --het and --frq respectively. Average relatedness was estimated as the proportion of identity-by-descent between individual pairs using the same software.

Population genetic differentiation, structure and principal component analysis

The analysis of molecular variance (AMOVA), pairwise genetic difference (F_{ST}) and Reynolds' genetic distance (Reynolds *et al.* 1983) were computed using ARLEQUIN 3.5.2 (Excoffier & Lischer 2010) with 1000 permutations for a different hierarchical level of population genetic data analysis. A neighbor-joining tree was constructed from allele-sharing distances using POWERMARKER (Liu & Muse 2005), and the tree was visualized using FIGTREE 1.4.2. (http://tree.bio.ed.ac.uk/software/figtree/). A Bayesian-model-based clustering method was used to elucidate the population genetic structure using STRUCTURE software version 2.3.4 (Falush *et al.* 2007) with 20 000 Markov-chain Monte Carlo runs after a burn-in period of 10 000 interactions and five independent runs for each value of K ($K = 2$–6). The server STRUCTURE HARVESTER (Earl 2012) was used to display the STRUCTURE results. The optimal K value was computed using the procedures of Evanno *et al.* (2005) based on ΔK. Principal component analysis (PCA) was estimated from the genetic relationship matrix of the first three eigenvalues and eigenvectors in R package version 3.4 (Zheng *et al.* 2012).

Results

Genetic variants

In this study, 53 347 SNPs were considered before marker-based quality control. SNP marker quality control excluded 6642 SNPs across populations (Table 1). Significant deviations for MAF values less than 0.05 caused 3839 SNPs to be excluded, and 1719 and 1715 SNPs were removed because of Hardy-Weinberg equilibrium ($P < 1e\text{-}5$) and call rates less than 0.95 respectively. Based on SNPs with MAF

 doi: 10.1111/age.12776

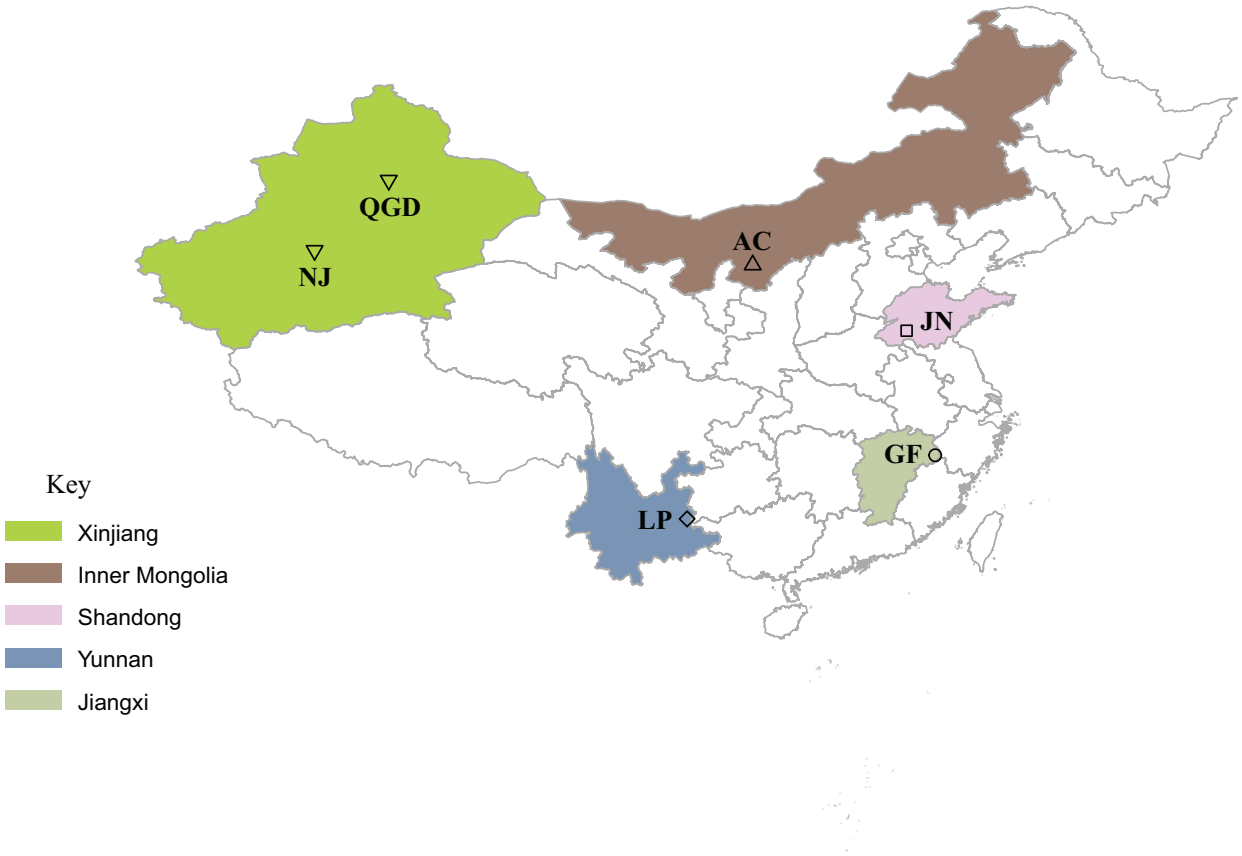

Figure 1 Geographic distribution and provinces of the six goat populations in China.

values less than 0.05, the LP breed had the highest number of SNPs excluded ($n = 12\,466$) and the JN breed had the least excluded ($n = 3037$). A common subset of 47 401 SNPs (88.85%) from 193 individuals was used for downstream analysis.

The total distribution of MAF values across all the populations is shown in Fig. 2. Among the six Chinese goat breeds, QGD (28.8%) and LP (20.2%) showed the highest and the lowest proportion of SNPs (MAF ≥ 0.3) respectively. The LP breed had a higher proportion of SNPs in the lowest MAF interval compared to the proportion of SNPs in the higher MAF intervals. The proportion of SNPs for the NJ, QGD, AC, JN and GF goat populations were shown to be higher as the MAF interval increased except in the last interval, for which the percentage decreased.

Genetic diversity analysis

The results of the genetic diversity both after and before quality control in all the populations are summarized in Tables 2 & S2 respectively. Based on the percentage of identity-by-descent, the average relatedness between individual pairs was estimated at 0.3%, 1.6%, 2.7%, 3.4%, 5.5% and 6.4% for the QGD, NJ, JN, AC, GF and LP goat breeds respectively with an overall mean of 3.3%. The average MAF was lowest for the LP (0.234 ± 0.16) and

Table 1 Single nucleotide polymorphism chip marker-based quality filtering results.

		Excluded SNPs				
Population	*n*	SNP CR < 0.95	MAF < 0.05	HWE (*P* < 1e-5)	Total	Remaining SNPs
Nanjiang	23	1726	6495	47	7085	46 262
Qinggeda	24	1608	4476	3	4931	48 416
Aarbas Cashmere	59	1806	7113	37	7759	45 588
Jining Grey	39	2863	3037	89	4797	48 550
Luoping Yellow	24	1591	12 466	7	12 876	40 471
Guangfeng	24	1684	7725	18	8235	45 112
Merged	193	1715	3839	1719	5946	47 401

CR, call rate; HWE, Hardy-Weinberg equilibrium; MAF, minor allele frequency; *n*, number of individuals.

highest for the JN (0.314 ± 0.12) breeds with an average mean of 0.279 ± 0.14 across the populations. With MAF less than 0.05, the proportion of polymorphic loci (P_N) was found to be highest in JN (93.9%) and lowest in LP (69.5%) with an overall mean of 84.7% across the populations.

The inbreeding coefficient was positive for the GF ($F_{IS} = 0.062$), QGD ($F_{IS} = 0.046$) and NJ ($F_{IS} = 0.014$) breeds with an overall mean of 0.016 across all the populations. The expected heterozygosity varied from 0.371 (LP) to 0.405 (JN) with an overall mean of 0.387, whereas the observed heterozygosity ranged from 0.306 in LP to 0.402 in JN with a mean of 0.362 across all populations. The expected heterozygosity was to some extent greater than the observed heterozygosity ($H_E > H_O$) both before and after quality control.

Population genetic differentiation, AMOVA, and structure

Pairwise genetic differentiation (F_{ST}) and Reynolds' genetic distance were measured using 47 401 SNPs (Table S3). The global F_{ST} was 8.6% and ranged from 0.02% to 0.16% ($P < 0.05$), indicating moderate genetic differentiation among the populations. This finding was further confirmed by the AMOVA (Table 3), which showed 90.2% of the variation to be intra-population and 8.6% ($P < 0.001$) to be inter-population.

A PCA is often used to envisage individual relationships within and among populations. The goat populations were grouped according to their geographic locations and breed origins. Our PCA results revealed that two breeds were clearly differentiated in the first and the second eigenvectors, which accounted for 11.21% and 8.64% of the total variation respectively, suggesting that a separate genetic structure exists between the AC and LP goat populations (Fig. 3). The NJ goat breed clustered closely with QGD goats, and the JN breed clustered closely with GF population.

To determine the geographic relationship between the individual animals, a neighbor-joining tree was constructed using genome-wide allele sharing. All individuals were grouped together and four different clusters were created (Fig. 4), which is in concordance with the results of the PCA analysis. The first cluster contained the GF and JN populations, QGD and NJ goats were grouped together in the second cluster, and the AC and LP goat breeds were seen as genetically distinct as the third and fourth clusters respectively.

The STRUCTURE analysis (Fig. 5) revealed a result similar to the PCA and neighbor-joining tree analysis. At $K = 2$, the AC goat population was the only breed that formed an independent genetic structure. Similarly, at $K = 3$ and 4, the AC and LP goat populations displayed a separate cluster, whereas the NJ, QGD, and JN presented as having similar genetic construction. Based on the ΔK value, $K = 3$ was the most optimal number for the inferred genetic structure of the populations (Fig. S1).

Discussion

Genetic diversity of Chinese goat populations

Knowledge about genetic diversity and population structure among goat populations is essential for genetic improvement, understanding of environmental adaptation as well as utilization and conservation of goat breeds. Although the 50K SNP panel was developed from sequence data for goat breeds, such as those used in this study (NJ, QGD, AC, JN, LP and GF; https://www.animalgenome.org/repository/pub/CAAS2018.1121/), all breeds showed sustained high levels of genetic variability, which could be due to the greater genetic diversity of the animals' ancestors (Wei *et al.* 2014). The JN breed demonstrated the highest level of genetic diversity compared to the others, which might be attributed to the absence of strong artificial selection pressures, a high gene flow and

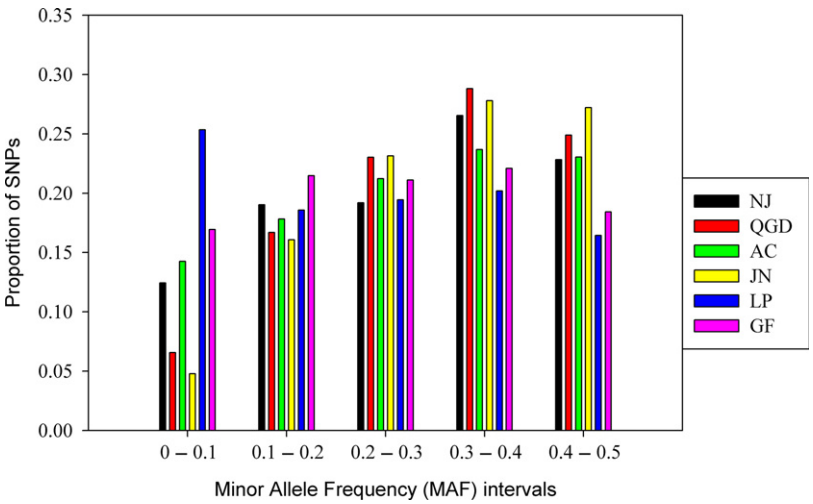

Figure 2 Distribution of minor allele frequency across the six Chinese goat populations.

Table 2 Genetic diversity, showing the name of the breed, sample size (n), proportion of polymorphic SNPs (P_N), average relatedness (PI_HAT), minor allele frequency, expected (H_E) and observed (H_O) heterozygosity and inbreeding coefficient (F_{IS}) for the six Chinese goat populations.

Breed name	n	% P_N	PI-HAT	MAF	H_E	H_O	F_{IS}
Nanjiang	23	86.1	0.016	0.284 ± 0.14	0.384 ± 0.15	0.374 ± 0.16	−0.014
Qinggeda	24	90.8	0.003	0.307 ± 0.13	0.397 ± 0.12	0.396 ± 0.14	0.046
Arbas Cashmere	59	84.6	0.034	0.276 ± 0.14	0.374 ± 0.15	0.365 ± 0.15	−0.013
Jining Grey	39	93.9	0.027	0.314 ± 0.12	0.405 ± 0.24	0.402 ± 0.12	−0.002
Luoping Yellow	24	69.5	0.064	0.234 ± 0.16	0.371 ± 0.17	0.306 ± 0.18	0.014
Guangfeng	24	83.1	0.055	0.258 ± 0.14	0.402 ± 0.13	0.320 ± 0.16	0.062
Overall	193	84.7	0.033	0.279 ± 0.14	0.387 ± 0.16	0.362 ± 0.15	0.016

Table 3 Analysis of molecular variance among Chinese goat populations.

Source of variation	DF	SS	VC	% of variations
Among population	5	255296.058	700.33400	8.61
Among individual within a population	187	1408043.310	97.21158	1.20
Within individuals	193	1415697.500	7335.22021	90.19
Total	385	3079036.868	8132.76579	100

DF, degree of freedom; SS, sums of squares; VC, variance component.

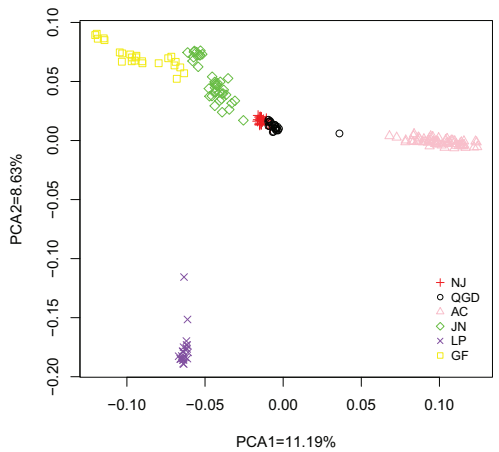

Figure 3 Principal component analysis showing genetic relationship results from a dataset of 39 552 SNPs in six Chinese indigenous goat breeds: NJ, Nanjiang; QGD, Qinggeda; AC, Arbas Cashmere; JN, Jining Grey; LP, Luoping Yellow; GF, Guangfeng. The first PCA accounts for 11.21% and the second PCA accounts for 8.64% of the variation.

heterogeneous characteristics of communal smallholder production systems in the local area. On the other hand, the lowest value was observed for the LP breed, which is consistent with its geographic history, as the breed is geographically distributed in Yunnan province, located in Southwest China, which is surrounded by the Qinling Mountains and Huaihe River Line (Wei *et al.* 2014); as a result, high inbreeding due to a small breeding stock population could have contributed to the low genetic diversity. We also observed that the LP breed showed a lower proportion of polymorphic loci (69.5%), which could be explained by the small size of the population and relatively high inbreeding. A relatively lower number of informative SNPs indicates populations with a narrow genetic base and individuality of the breeds (Mdladla *et al.* 2016). However, populations with a larger sample size would be needed to obtain a better estimate of inbreeding measures (Brito *et al.* 2017).

The genetic diversity values (such as observed and expected heterozygosity) observed in the six goat populations studied are in concordance with the reported estimated values for Spanish goats, $H_E = 0.29–0.42$ and $H_O = 0.28–0.41$ (Manunza *et al.* 2016); South African goats, $H_E = 0.41$ and $H_O = 0.42$ (Mdladla *et al.* 2016); and Italian goats, $H_E = 0.37–0.41$ and $H_O = 0.36–0.41$ (Nicoloso *et al.* 2015) based on a SNP panel. However, our results were lower than the previously reported average heterozygosity of 0.972 (Wang *et al.* 2008), 0.806 (Zhang *et al.* 2012) and 0.6433 (Wei *et al.* 2014) for multiple Chinese goat breeds based on microsatellite markers.

Monitoring and controlling inbreeding is important for limiting the potential impact of deleterious alleles, inbreeding depression and loss of variance. In this study, we obtained a positive inbreeding coefficient for the GF, QGD and LP goat breeds, indicating no deliberate inbreeding in the populations. Overall, the proportion of polymorphic loci across all study populations was high and a deficiency in heterozygosity was observed ($H_O < H_E$). The H_E values were within almost a fraction of one standard error both before and after quality control, which reflects that the population substructure has been maintained within the Chinese goat breeds, which could be due to the Wahlund effect rather than inbreeding (Oliveira *et al.* 2010).

Population genetic structure

The neighbor-joining tree (Fig. 4) revealed the individual animals clustering into four major groups on the basis of the degree of the genetic relationships and geographic separation. The NJ and QGD populations were grouped together, suggesting a high genetic relationship between these populations, which could be due to possible high gene flow and short domestication history, as they are raised in nearby geographic regions. The tree also revealed that the GF breed was located in the same clade as the JN breed, indicating that these breeds may share common ancestry.

 doi: 10.1111/age.12776

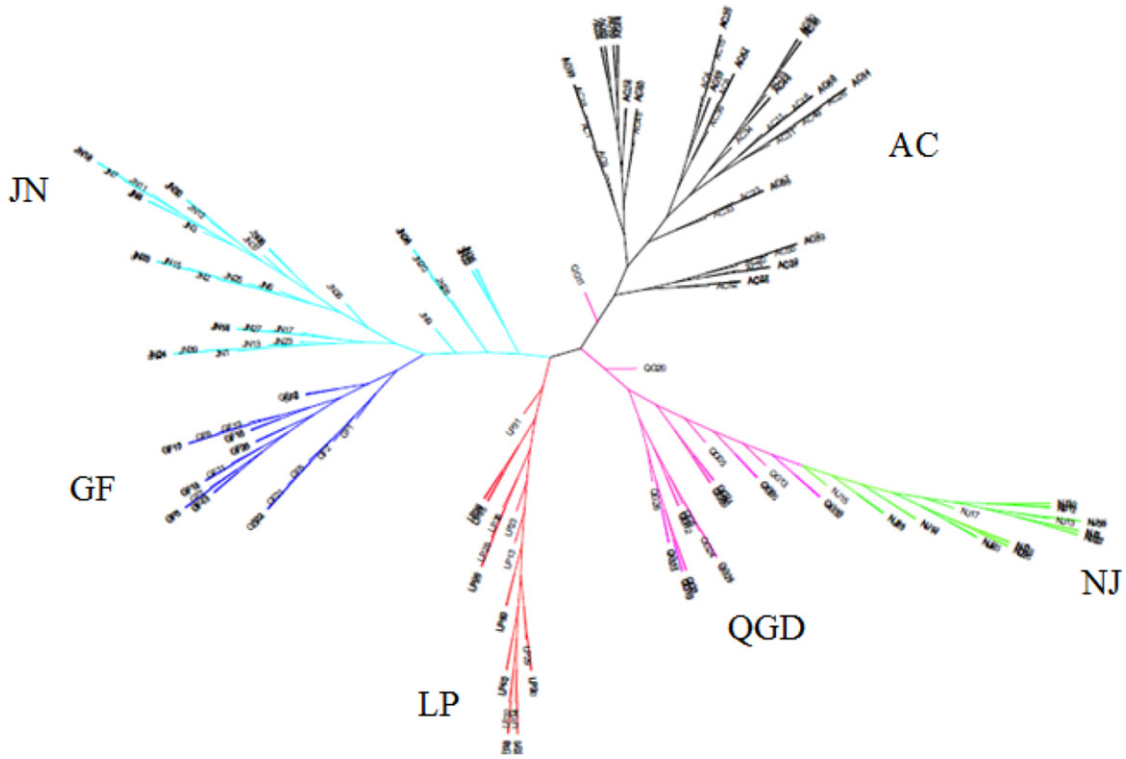

Figure 4 Phylogenetic relationship constructed using a neighbor-joining tree from a dataset of 39 552 SNPs in six Chinese indigenous goat breeds: NJ, Nanjiang; QGD, Qinggeda; AC, Arbas Cashmere; JN, Jining Grey; LP, Luoping Yellow; GF, Guangfeng.

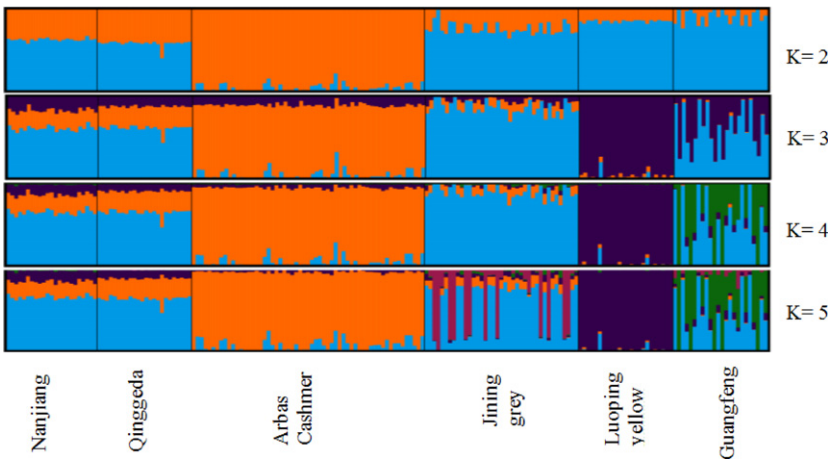

Figure 5 Population structure plots showing the proportions of ancestral populations of the six Chinese goat populations for $K = 2$ to $K = 5$.

Our findings were consistent with those in a previous study of different Chinese goat breeds based on microsatellite markers (Di *et al.* 2011). The observation was further supported by the PCA analysis results, which separated the populations according to their geographic isolation (Fig. 3). The results of this study are also in line with the clustering and geographic patterns described in previous studies based on microsatellite markers in different Chinese goat populations (Zhang *et al.* 2012) and on the SNP BeadChip in Italian goats (Nicoloso *et al.* 2015).

Furthermore, the STRUCTURE analyses allowed for the assignment of individuals to groups based on their genetic similarities and geographic isolation, thereby providing information about the number of ancestral populations underlying the observed genetic diversity. Certainly, the representative clusters in LP were also present in AC breeds with a low ancestry coefficient. A similar pattern was observed in LP, which contained some features of the AC cluster. The two clusters formed by LP are suggestive of genetic substructuring resulting from using LP animals belonging to divergent selection lines. At most, the effect of geographic isolation on the local scale was seen in the LP goat breed, which is traditionally reared in the Yunnan province located in southwest China (Fig. 1), a region that

has a special geographic structure due to the Hengduan Mountains located at the middle of Qinghai-Tibet Plateau and Yunnan-Guizhou Plateau, which has restricted livestock gene flow between northern and southwestern China (Wei *et al.* 2014). In addition, the AC goat breed is one of the most well-known Chinese goats for the cashmere trait, which has experienced strong artificial selection (Guan *et al.* 2016; Liu *et al.* 2018). Our findings are in agreement with previous research reported to have used microsatellite analysis in goat populations (Zhang *et al.* 2012) and based on the SNP panel in sheep (Kijas *et al.* 2012) and cattle (Edea *et al.* 2015).

Genetic differentiation and AMOVA

In this study, the findings for genetic distance and genetic differentiation reflected the natural geographic separation of the populations; for instance, when the geographic distance between populations was small, genetic differentiation was low, but increased geographical distance went along with increased genetic differentiation among the populations. Therefore, isolation by geographic separation has played a major role in shaping the genetic differentiation among Chinese goat populations. A similar investigation was reported in multiple Chinese goat breeds (Wei *et al.* 2014) based on microsatellite markers. The global F_{ST} value of 8.6% in this study was similar to the F_{ST} value of 8.9% found in South African goat breeds in a study based on a SNP panel (Visser *et al.* 2016) but higher than the F_{ST} values of 6.3% (Zhang *et al.* 2012) and 6.2% (Di *et al.* 2011) for different Chinese goat breeds in studies based on microsatellite markers. Our results were further confirmed by AMOVA, which showed that 90.2% ($P < 0.05$) of the variation was within populations and 8.6% ($P < 0.001$) among populations. The findings of this study was consistent with the values of 87.86% and 11.86% reported for South African, French and Argentinian goat populations (Visser *et al.* 2016) and values of 89.83% among and 7.49% within individual variations for South African and Italian goat breeds (Nicoloso *et al.* 2015) using the SNP panel analysis.

Conclusion

Single nucleotide polymorphisms are quite reliable markers for use in studies of genetic diversity and variability of populations and to determine whether individuals belong to the claimed populations. Here, we observed that the genetic structures of the studied breeds were explained mainly by their geographic origin. In general, low to moderate levels of genetic variability and high within-breed diversity was maintained in the studied goat populations and thereby provide a good opportunity for sustainable utilization and improvements in the conservation of animal genetic resources. Information from this study may also assist in the reduction of unnecessary inbreeding or gene flow among populations. Further studies will be helpful to confirm and refine our results through the inclusion of more samples and other tools such as whole genome re-sequencing, high-density SNP genotyping and phenotypic data.

Acknowledgements

This study was financially supported by the National Natural Science Foundation China (31472064, 31601910, U160 3232), the Special Fund for Agro-scientific Research in the Public Interest (201303059) and the earmarked fund for Modern Agro-industrial Technology Research System (CARS-40-01).

Conflict of interests

The authors confirmed that there is no conflict of interests.

References

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

Brito L.F., Kijas J.W., Ventura R.V., Sargolzaei M., Porto-Neto L.R., Cánovas A., Feng Z., Jafarikia M. & Schenkel F.S. (2017) Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics* **18**, 229.

Devendra C. (2015) Dynamics of goat meat production in extensive systems in Asia: improvement of productivity and transformation of livelihoods. *Agrotechnology* **4**, 131.

Di R., Vahidi S.M., Ma Y.H., He X.H., Zhao Q.J., Han J.L., Guan W.J., Chu M.X., Sun W. & Pu Y.P. (2011) Microsatellite analysis revealed genetic diversity and population structure among Chinese Cashmere goats. *Animal Genetics* **42**, 428–31.

Earl D.A. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359–61.

Edea Z., Bhuiyan M., Dessie T., Rothschild M., Dadi H. & Kim K. (2015) Genome-wide genetic diversity, population structure and admixture analysis in African and Asian cattle breeds. *Animal* **9**, 218–26.

Evanno G., Regnaut S. & Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611–20.

Excoffier L. & Lischer H.E. (2010) ARLEQUIN suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564–7.

Falush D., Stephens M. & Pritchard J.K. (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7**, 574–8.

Grasso A.N., Goldberg V., Navajas E.A., Iriarte W., Gimeno D., Aguilar I., Medrano J.F., Rincón G. & Ciappesoni G. (2014) Genomic variation and population structure detected by single nucleotide polymorphism arrays in Corriedale, Merino and Creole sheep. *Genetics and Molecular Biology* **37**, 389–95.

Groeneveld L., Lenstra J., Eding H., Toro M., Scherf B., Pilling D., Negrini R., Finlay E., Jianlin H. & Groeneveld E. (2010) Genetic diversity in farm animals – a review. *Animal Genetics* **41**, 6–31.

 doi: 10.1111/age.12776

Guan D., Luo N., Tan X., Zhao Z., Huang Y., Na R., Zhang J. & Zhao Y. (2016) Scanning of selection signature provides a glimpse into important economic traits in goats (*Capra hircus*). *Scientific Reports* **6**, 36372.

Kijas J.W., Lenstra J.A., Hayes B., Boitard S., Neto L.R.P., San Cristobal M., Servin B., Mcculloch R., Whan V. & Gietzen K. (2012) Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology* **10**, E1001258.

Lashmar S.F., Visser C. & Van Marle-Köster E. (2016) SNP-based genetic diversity of South African commercial dairy and fiber goat breeds. *Small Ruminant Research* **136**, 65–71.

Lin B.Z., Kato T., Kaneda M., Matsumoto H., Sasazaki S. & Mannen H. (2013) Genetic diversity and structure in Asian native goat analyzed by newly developed SNP markers. *Animal Science Journal* **84**, 579–84.

Liu K. & Muse S.V. (2005) POWERMARKER: an integrated analysis environment for genetic marker analysis. *Bioinformatics* **21**, 2128–9.

Liu Y., Wang L., Li X. *et al.* (2018) High-throughput sequencing of hair follicle development-related micrornas in cashmere goat at various fetal periods. *Saudi Journal of Biological Sciences* **7**, 1494–508.

Ma L., Qin Q., Yang Q., Zhang M., Zhao H., Pan C., Lei C., Chen H. & Lan X. (2017) Associations of six SNPS of POU1F1-PROP1-PITX1-SIX3 pathway genes with growth traits in two Chinese indigenous goat breeds. *Annals of Animal Science* **17**, 399–411.

Manunza A., Noce A., Serradilla J.M., Goyache F., Martínez A., Capote J., Delgado J.V., Jordana J., Muñoz E. & Molina A. (2016) A genome-wide perspective about the diversity and demographic history of seven Spanish goat breeds. *Genetics Selection Evolution* **48**, 52.

Mdladla K., Dzomba E., Huson H. & Muchadeyi F. (2016) Population genomic structure and linkage disequilibrium analysis of South African goat breeds using genome-wide SNP data. *Animal Genetics* **47**, 471–82.

Naderi S., Rezaei H.-R., Pompanon F. *et al.* (2008) The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 17659–64.

Nicoloso L., Bomba L., Colli L., Negrini R., Milanesi M., Mazza R., Sechi T., Frattini S., Talenti A. & Coizet B. (2015) Genetic diversity of Italian goat breeds assessed with a medium-density SNP chip. *Genetics Selection Evolution* **47**, 62.

Oliveira J., Ribeiro M., Rocha L., Gomes-Filho M., Delgado J., Martinez A., Menezes M., Bettencourt C. & Gama L. (2010) Genetic relationships between two homologous goat breeds from Portugal and Brazil assessed by microsatellite markers. *Small Ruminant Research* **93**, 79–87.

Onzima R., Upadhyay M., Mukiibi R., Kanis E., Groenen M. & Crooijmans R. (2018) Genome-wide population structure and admixture analysis reveals weak differentiation among Ugandan goat breeds. *Animal Genetics* **49**, 59–70.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D., Maller J., Sklar P., De Bakker P.I. & Daly M.J. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–75.

Reynolds J., Weir B.S. & Cockerham C.C. (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**, 767–79.

Song S., Yao N., Yang M., Liu X., Dong K., Zhao Q., Pu Y., He X., Guan W. & Yang N. (2016) Exome sequencing reveals genetic differentiation due to high-altitude adaptation in the Tibetan cashmere goat (*Capra hircus*). *BMC Genomics* **17**, 122.

Visser C., Lashmar S.F., Van Marle-Köster E., Poli M.A. & Allain D. (2016) Genetic diversity and population structure in South African, French and Argentinian Angora goats from genome-wide SNP data. *PLoS One* **11**, E0154353.

Wang J., Chen Y., Wang X. & Yang Z. (2008) The genetic diversity of seven indigenous Chinese goat breeds. *Small Ruminant Research* **74**, 231–7.

Wei C., Lu J., Xu L., Liu G., Wang Z., Zhao F., Zhang L., Han X., Du L. & Liu C. (2014) Genetic structure of Chinese indigenous goats and the special geographical structure in the Southwest China as a geographic barrier driving the fragmentation of a large population. *PLoS One* **9**, E94435.

Zeder M.A. (2000) The initial domestication of goats (*Capra hircus*) in the Zagros Mountains 10 000 years ago. *Science* **287**, 2254–7.

Zhang L., Yang Q., Li X., Zhou R. & Li L. (2012) Genetic diversity of five goat breeds in China based on microsatellite markers. *African Journal of Biotechnology* **11**, 11671–7.

Zheng X., Levine D., Shen J., Gogarten S.M., Laurie C. & Weir B.S. (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–8.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Distribution of (a) ΔK and (b) L(K) for 193 individuals of six Chinese goat populations with the modal value indicating the true K or the exact inference of genetic structure ($K = 3$).

Table S1 Summary of breed names, sample codes, sample size, geographic distribution of the six Chinese indigenous goat populations.

Table S2 Genetic diversity, showing the names of the breeds, sample size, minor allele frequency, expected (H_E) and observed (H_O) heterozygosity and inbreeding coefficient (F_{IS}) for the six goat populations before quality control.

Table S3 Mean value of pair-wise genetic differentiation (F_{ST}) between the six goat populations (below diagonal) and Reynolds' genetic distance (above diagonal).