Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Deep genome sequencing provides potential novel insights into plateau adaptations in domestication of goats to extreme environments

Dehong Tian
Chinese Academy of Sciences
Baicheng Zhou
General Station of Animal Husbandry of Qinghai province
Buying Han
Chinese Academy of Sciences
Sijia Liu
Chinese Academy of Sciences
Fei Tian
Chinese Academy of Sciences
Delin Qi
Qinghai University
Kai Zhao (zk1101@yeah.net)
Chinese Academy of Sciences

Research Article

Keywords: cashmere goat, defensive, whole-genome resequencing, olfactory perception, adaptability

Posted Date: May 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-453887/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Background: Environmental conditions and extensive rearing impose a myriad of selective pressures on domesticated animals living at high altitude. These pressures have selected for the characteristics of strong adaptability for life in the alpine pastoral area. While it is accepted that various natural and anthropogenic factors, such as environmental pressures, human migration, and socioeconomic practices, have shaped the genomic structure of goats, little is known regarding the resulting genetic adaptations. The goat (Capra hircus) have become well adapted to certain extreme environments (e.g., plateaus and deserts). Through whole-genome sequencing, it was hypothesized here that the immune system and olfactory perception genetic regulatory pathways may contribute to the defense related adaptations of the goat population. As such, CNGA4 and Camk 2b have been identified as key candidate genes that drive these adaptations during goat domestication, and they may enhance immunity and defenses against external adverse factors. This study provides a new insight into the unique regional defense adaptation of Qaidam cashmere goats during the generations long domestication process, in order to provide a theoretical basis for the adaptation of cashmere goats to a commercial environment.

Results: More than 3,000 candidate genes have been found that may have been selected for through both natural and artificial processes during domestication. These candidate genes require validation with additional samples in the future. novel immune and olfactory perception genetic regulatory pathways were identified, which may correspond to defensive adaptations to the extreme environments.

Conclusion: CNGA4 and Camk 2b may enhance immunity and defense against external adverse factors during domestication.

Background

Goats are one of the most important livestock resources of local breeders that was domesticated approximately 10,000 years ago[1], The domestication of goats transformed natural reproduction into a controlled process, which has turned out to be a critical development in the history of modern human civilization. Nevertheless, the origins of caprine domestication are controversial[2]. A great diversity of goat breeds has been created through selective breeding, adaptation, isolation, and genetic drift. Domesticated goats retain many of their natural characteristics, and are an important genetic resource. There are numerous local goat breeds in China, which have a strong fitness and foraging capability under a wide range of habitats, including the dry, cold, and harsh Qinghai-Tibet Plateau and the warm and humid lands of South China[3].

Artificial selection during domestication and production-oriented breeding has greatly shaped the level of genomic variability in goats[4]. High-throughput whole genome sequencing can be used to characterize adaptive genetic variations in harsh or extreme environments[5,6]. Wang et al. detect selection signals in 8 goat breeds by analyzing the sequences of 5 domestic goats and 3 introduced breeds[4], Benjelloun et al. found a number of candidate genes in which mutations allowed for the adaptation of metabolic

pathways or biological processes associated with local environmental conditions by screening 3 indigenous goat breeds[7]. Selective sweep analyses identified multiple regions in which genetic variations have been described in duck[8], pandas[9], pig[10], sheep[11], dog[12] and goat[13]. These populations have evolved numerous adaptations, and several genomic studies have been conducted to identify the genes that drive these adaptations. High-altitude environments exert strong selective pressures, and human and animal populations have evolved in convergent ways to cope with a chronica lack of oxygen[14]. Different environmental adaptations are thought to be the result of favorable mutations and selection pressures. This provides interesting opportunities to identify the selective signatures associated with adaptation genes. However, little is known about the many diverse genetic mechanisms by which have allowed goats to successfully adapt to extreme environments, such as the olfactory sense and the immune system.

To elucidate the potential genes or variations in signal characteristics and phenotypes of goats adapted to local extreme climatic conditions, whole-genome sequencing was performed on samples from 10 indigenous Chinese Qinghai-Tibetan Plateau goat breeds (altitude >4,000 m above sea level). The resulting data was compared with previously published genome sequences of 9 low altitude goat breeds. The aim of this study was to develop a basic understanding of the genetic variations associated with the acclimation characteristics of goats that are influenced by both artificial and natural selection. Taken together, this information will provide a new perspective for the research on the mechanisms of acclimation and adaptation in Chinese goat breeds.

Materials And Methods

Ethic statement

This study followed the recommendations of the “Regulations for the Management of Affairs Concerning Experimental Animals” (Ministry of Science and Technology, China, revised in June 2004). The study were approved by the Animal Care and Use Committees of the Northwest Institute of Plateau Biology, Chinese Academy of Sciences. The animals are not harmed during sample collection. The study is in accordance with the ARRIVE Guidelines.

Animals and whole-genome sequencing

Ear tissue samples were collected from 10 Qaidam Cashmere goats (CMDG) in the Qaidam area, Haixi prefecture, Qinghai Province (altitude>3,000 m above sea level). For a population sample from a contrasting climate, ear tissue samples were collected from 9 Chengdu Brown goats (CDBG) of the Chengdu plain (altitude<1000 m above sea level). Total genomic DNA was extracted from samples and a minimum of 3 μg genomic DNA was used to construct paired-end libraries with an insert size of 500 bp using the Paired-End DNA Sample Prep kit (Illumina Inc., San Diego, CA, USA). These libraries were sequenced using the novaseq6000 (Illumina Inc., San Diego, CA, USA) NGS platform at Gene denovo company (Guangzhou, China).

Short insertions and deletions (InDels) and Structural Variants (SVs)

The Burrows-Wheeler Aligner (BWA) was used to align the clean reads from each sample against the reference genome (https://www.ncbi.nlm.nih.gov/assembly/GCF_001704415.1) with the default parameters. Variant calling was performed for all samples using the GATK's Unified Genotyper. Both SNPs and InDels were filtered using GATK's Variant Filtration with proper standards (-Window 4□-filter "QD < 2.0 || FS > 60.0 || MQ < 40.0 "□-G_filter "GQ < 20"). Those exhibiting segregation distortion or sequencing errors were discarded. In order to determine the physical positions of each SNP, the software tool ANNOVAR[15], was used to align and annotate SNPs or InDels. Structural Variation (SV) types include translocations, inversions and insertion events, which were determined the breakdancer software.

Phylogenetic and population structure analysis

Next, a phylogenetic tree was constructed using the neighbor joining method in the PHYLIP software program[16] (version 3.69). The bootstrapped confidence interval was based on 1000 replicates. The population subdivision pattern was preliminarily classified in the principal component analysis by the GCTA program[17]. The admixture model-based software Admixture[18] V1.3.0 was further used to estimate population structure (Q matrix). The tested K was set from 1 to 10, and the optimal K was determined as that which resulted in the lowest cross-validation error. The kinship matrix (K matrix) was calculated using the method of Loiselle[19] et al by the software of SPAGeDi[20] V1.5.

Analysis of Nucleotide diversity and Selective Sweep Regions

To ensure accuracy of the analysis, missing SNPs at a rate of greater than 20% were excluded from further analysis. Multiple indices related to population genetics, including Nucleotide Diversity[21] (π) and Fst, were calculated with Pop Genome using the sliding window approach. The window size was set to 100 kb, and the step size was set to 10kb. Selective sweep regions were selected according to the interception of 2 indices, which were Fst and π ratio with a top 5% threshold level. All related graphs were drawn by R scripts. Candidate genes within the sweep regions were extracted for further analysis.

Results And Discussion

Whole-Genome Sequencing and Genetic Variation

Genome sequencing yielded a total of 438.46 billion reads of raw paired-end reads with a length of 150-bp from 19 goats in the present study. Stringent quality filtering yielded 430.32 billion reads or 64,548 Gb of genome data (approximately 147× effective sequence depth). Over 98% of the total clean reads were mapped to the annotated goat reference genome with a coverage of ~99.66% (Supplementary Table S1), These data are indicative of high quality sequences for the subsequent analyses. A total of 22.29 million high quality SNPs were obtained, relative to the published reference goat genome. Most of the SNPs (62.98%) in the individual genome of goat were located in intergenic regions. The remaining SNPs were located upstream (0.55%) and downstream (0.57%) of open reading frame, in introns (34.51%) and only 0.76% of the SNPs were located in exonic regions (0.63%) (Supplementary Table S2). A total of 60,259

non-synonymous SNPs and 84,818 synonymous SNPs were located within exons, forming a nonsynonymous/synonymous ratio of 0.71 was closed to the ration of Mouflon sheep[22] (0.689) (Supplementary Table S2). Next, the 18.58 million SNPs were divided into 3 categories: 9.86 million SNPs shared between CDMG and CDBG, as well as 8.57 million and 0.15 million SNPs unique to CDMG and CDBG (Fig.1), respectively. Next, 18.58 million SNPs were measured for genome-wide variation and the frequency spectrum so as to accurately detect genomic footprints left by selection. The numbers of SNPs in the exonic, intronic, and intergenic regions of CDBG were lower than was observed in CDMG. Furthermore, higher genomic diversity was observed in CDMG relative to CDBG (Supplementary Table S3). These data indicate that the high altitude goat breed underwent intense natural and artificial selection in extreme weather conditions.

The sequence diversity (π) was estimated to be 1.35×10^{-3} in CDBG and 2.05×10^{-3} in CDMG, which were slightly lower than the native sheep[22] (2.28×10^{-3}). Moreover, the goats exhibited a nucleotide diversity that was higher than estimates for human populations[23] (1.0×10^{-3}), and lower than that of Chinese cattle breeds[24] (Leiqiong cattle,4.2×10^{-3}). The level of genetic differentiation between them, Fst, was estimated to be 0.03, which is lower than that between Tibetan and Hu sheep[25] (about 0.249). The high proportion of shared SNPs and the low Fst revealed a small genetic divergence between plateau and plain goats. However, the principal component analysis (PCA) was significantly separated along the first principal component between plateau and plain goats (Fig.2), which accounts for the greatest proportion of genetic variability. The neighbor-joining tree also grouped the plateau and plain goats into two distinct groups.

Short insertions and deletions (InDels) and Structural Variants (SVs)

As efforts turn to whole genome sequencing as an approach for identifying adaptive variation, it will become even more critical to detect InDels fully and accurately[26]. Furthermore, InDels and SVs are important types of genetic variations that contribute significantly to genomic diversity[27]. A total of 861,995 insertions and 1,114,290 deletions were identified. Within all InDels, 1,129,583 were observed in both domestic goat, 388,539 were specific to CDBG and 741,044 were specific to CDMG. The length of InDels ranged from 0 to 10 bp, with the large majority consisting of single nucleotide deletions (Fig.3a). Observation of a high number of unique genetic variations suggests a benefit to the animals in protecting genetic diversity.

A total of 31,483 reliable structure variants were detected. Among these, 4,941 were insertions, 16,785 were deletions, 292 were inversions, 7,888 were chromosomal translocations, and 1,577 were tandem duplications. Within these SVs, 1,366 were specific to CDBG and the other 11,915 were specific to CDMG (Fig.3b). No significant differences were observed in the length of these types of SVs. Apparently, more variation was retained under selection pressure for CDMG in this environment. This phenomenon could be the result of the life-sustaining adaptions to high altitude, low temperature, and low oxygen conditions. The discovery range of InDels overlapped with that of structural variation. Thus, breed-specific variants

could be useful for further studies of breed characterization, while overlapping variants may reflect closer relations between breeds[28].

Signatures of Positive Selection in goat Genome

Genome variations were compared to both nucleotide diversity and allele frequency in response to natural and artificial selection. The ratio of the genetic diversity in CDBG versus CDMG goats (πCDBG/πCDMG) across the goat genome along the autosome using the sliding window method was measured. The 5% of windows with a Fst>=0, π ratio (CDBG compared to CDMG=0.775), in comparison to CDBG, CDMG was observed to possess lower levels of polymorphisms, suggesting that these genomic regions were strongly selected for in the extreme environment. These selected regions also exhibited significant differences (Tajima.D= - 0.313). This result quantitatively reflects that the extreme climatic conditions had a strong impact on the emergence of adaptive genetic changes in highland goats. In addition, although the overall goat genome is similar, two divergent clusters were observed that may be the result of adaptations to the extreme environmental conditions in which they lived (Fig.2). Moreover, the combination of Fst values and π ratios were more reliable indicators of goat genome for the selected genomic regions from the plateau and high-altitude regions. A total of 1277, were identified as high leves of differentiation regions (Supplementary Table S4). Using the top 5% of Fst>=0 and π ratio>=0.78 cutoffs for the CDBG/CDMG (Fig.4b). A total of 3184 overlapping selective candidate genes were observed to be shared between the Fst values and π ration values (Fig.4a). The combined analysis of these methods provided further support that the candidate genes were in fact a result directly related to the treatment.

GO and pathway enrichment analysis of selection signals

It is reasonable to expect that specific genes and biological pathways be involved in adaptation to extreme environments. To gain insights into the molecular functions of the candidate genes in the selected region and to provide clues for further screening validation, a gene ontology enrichment analysis was conducted. The analysis identified significant enrichment of 1867 specific candidate genes embedded in the selective regions in 175 GO terms ($P < 0.05$, Supplementary Table S5). These genes were significantly enriched in detection of stimulus involved in sensory perception (GO:0050906), neurological system process (GO:0050877), positive regulation of phospholipase activity (GO:0010518), detection of stimulus (GO:0051606). These pathways reflect the likely adaptations of cashmere goats to the low temperatures on the plateau. Also enriched were cardiac conduction (GO:0061337), pulmonary valve morphogenesis (GO:0003184), positive regulation of potassium ion transmembrane transporter activity (GO:1901018). These pathways are associated with adaptations to activity under the hypoxic environmental conditions associated with life at high altitude. Moreover, some immune activities were also significantly enriched. For example, positive regulation of T cell mediated immunity (GO:0002711), regulation of humoral immune response (GO:0002920), this would probably have been the result of a rapid response to new pathogens associated with frequent contact with humans during domestication[29]. Multiple candidate genes involved in production traits. For example, coat color genes included *OSTM1* (ncbi_102189757), *KITLG* (ncbi_100860807), *POMC* (ncbi_102182514), *APC* (ncbi_102184724), *CDK5*

(ncbi_102178040), body size included *BMP2K* (ncbi_102186290), *FOSL2* (ncbi_102171522), *IGF2* (ncbi_100861277), *BMP1* (ncbi_102179476), *ACSL3* (ncbi_102185694), *POMC* (ncbi_102182514), *CKM* (ncbi_102170262), *MYF5* (ncbi_100861252), cashmere traits included *PHF14* (ncbi_102173711), *FGF5* (ncbi_102171226), *KRT74* (ncbi_102176789). In addition, some specific selective genes that were expected to be mutated to allow for environmental adaptation were also identified. These included genes involved in adaptability to hypoxia, such as *GNB1* (ncbi_102187070), *HIF1A* (ncbi_100861391), *RHOG* (ncbi_102191156), *PPARA* (ncbi_100861393), *TMTC2* (ncbi_102190760), *EPO* (ncbi_102180175), *TH* (ncbi_102179849), *PPARA* (ncbi_100861393), *ACER1* (ncbi_102179093), *HSP70.1* (ncbi_100860849). These findings substantially enriched the catalog of genetic variation and provided new insights into the unique characteristics of Chinese native goat breeds.

Different genes coordinate with each other to perform their biological functions, and pathway-based analysis is helpful to further understand the biological functions of genes. Selected gene were annotated using KEGG pathways. The analysis identified 21 genes that were significantly enriched (P□0.05) (Fig.5, Supplementary Table S6). Many of these pathways that were identified are involved in inflammatory responses in the lung tissue and immune system. For example some of the enriched pathways included Asthma pathway, Pertussis pathway, Antigen processing and presentation pathway. This may be because in the hypoxic environment, lung tissue and pulmonary blood vessels are damaged, so that the pathway of organ inflammatory response is in a state of inhibition, and by activating a large number of immune-related biological pathways, the inflammatory response caused by blood vessels in the remodeling process is reduced[30]. This is probably a defensive adaptation mechanism of the cardiovascular system to the environment of high altitude hypoxia. The immune regulatory cytokine interleukin 10 (IL-10) is an important regulatory cytokine with an overall immunosuppressive effect, which may represent an evolutionary adaptation that permits reactivation of the IL-10 gene under appropriate physiological or pathological conditions[31]. There are also candidate gene clusters associated with neural and automatic immune responses. Immune related genes are expected to be a main target for selection because they play a key role in immunity and defense, for example interleukin 1A (*IL1A*), interleukin 1B (*IL1B*), interleukin 10 (*IL10*), interleukin 9 (*IL9*), interleukin 4 (*IL4*), interleukin 13 (*IL13*), interleukin 18 (*IL18*)[32,33,34]. This reflects the effect of natural selection pressure on plateau goats in the harsh habitat of high altitude areas. There may be general patterns of recent evolutionary adaptation of the human immune system to particular regions and that these adaptations can produce differences in disease susceptibility[35]. Taken together, these results suggest that regulatory variations affecting immune-related genes, particularly those associated with viruses, may confer a selective advantage on the populations.

In addition to the environmental challenges posed by low oxygen and high UV radiation, smell plays a key role in the survival of animals through a series of core behaviors, such as sexual selection, alarm signals, foraging, and defensive awareness[36]. As the data presented here indicated, the Olfactory transduction signaling pathway was highly enriched (Q< 0.05) (Fig.5a). The ability of prey to innately recognize the odor of a potential predator provides a strong selective advantage. These neural mechanisms enable chemical surveillance against other species, interpretation of the cues, and initiation of a defensive

response [37,38]. Signals that induce fearful behavior, secreted by one individual and received by a second individual of the same species, in which they initiate a specific reaction, are cues transmitted between species that selectively disadvantage the sender and advantage the recipient[39]. Pheromones are chemical signals between two members of the same species which are known to be found mainly in animal communication. Although these models have not provided insight into the organization of the neural response in the receiving animals[40], subsets of sensory neurons are thought to be genetically determined and modulate innate behavior.

However, the evolution of olfactory perception during goat domestication has previously received little attention. Olfactory perception acuity is less important for survival of domesticated goats living in the predictable farm environment than those living in the wild. The three genes enriched in the olfactory perception signaling pathway were Cyclic nucleotide gated channel alpha 4(*CNGA4*), Adenylate cyclase 3 (*ADCY3*), cGMP-dependent 3',5'-cyclic phosphodiesterase (*PDE2A*). Heteromultimeric cyclic nucleotide—gated (CNG) channels play a central role in the transduction of odorant signals and subsequent adaptation, Cyclic nucleotide gated channel alpha 4 (*CNGA4*)subunit rapid olfactory adaptation allows an animal to continuously assess changes in the odor environment and the intensity of odor[41]. The *CNGA4* subunit in Olfactory sensory neurons (OSNs) was a functionally important mechanism for sensory adaptation of the olfactory system *in vivo*[42]. This volatile "alarm signal" produces and releases defensive chemicals that play an important role in animals adapting to their natural lives, as originally described by Darwinian natural selection in extreme environments. In this study, the CNGA4 gene revealed a higher level of diversity in CDMG (Fig.6a, 6b). Further screening of this candidate gene for non-synonymous mutations in exon, which may be representative of putative functional variants, revealed 4 missense SNPs (exon3: c.G178A/p.D60N, exon4: c.T835C/p.F279L, exon4: c.G893A/p.R298Q and exon6: c.C1489A:p.L497M) (Fig.6c) in the domestic goats. These results suggest that variations may be an important adaptation of the olfactory defense system.

Moreover, *ADCY3* gene encodes the adenylate cyclase involved in olfactory signal transduction. Animals carrying functional knockouts of *ADCY3* are unable to smell, as olfactory signaling across the main olfactory epithelium is mediated via receptor stimulation of type 3 adenylyl cyclase[43]. An indispensable role of *ADCY3* in postnatal maturation of OSNs and the maintenance of olfactory ciliary ultrastructure in mice through adulthood[44]. The *PDE2A* gene was expressed in olfactory sensory neurons of olfactory epithelium[45]. As a marker for a unique set of olfactory sensory neurons, *PDE2A* is now defined as the “necklace olfactory”[46]. In the present study, 2 remarkable selection signals were detected in the *ADCY3* (Fst=0 and π ratio=0.87) and *PDE2A* (Fst=0 and π ratio=0.79). These observations suggest that these selection signals may result in an enhancement of the olfactory perception defense system.

Interestingly, KEGG enrichment analysis showed that oocyte meiosis pathways were also involved in the adaptations to the environment. Calcium/calmodulin-dependent protein kinaseII (Camk 2b) might be related the hypoxia-induced factors-1 (*HIF-1*) pathway, or may possibly even function through the HIF-1 signal pathway[47]. The HIF pathway is a known complex hypoxic regulatory system involved in

homeostasis, embryogenesis, and development, and can stimulate hundreds of down-regulated genes in response to hypoxia. In addition, the anaphase-promoting complex subunit 13 (*ANAPC13*) gene in this pathway is specifically involved in oocyte maturation[48]. In hypoxic environment, through the transmission of multiple collective signals, cellular metabolism may be converted from aerobic to glycolytic metabolism, which is conducive to the maturation of granulosa cells in the process of follicle maturation. This would allow for the cells to adapt their energy requirements using pathways with reduced oxygen requirements in hypoxic environments. Therefore, after a period of Darwinian natural selection, the relevant adaptation mechanism of a species may be the result of interactions between multiple genes. This indicates the complexity of selecting important adaptively mutated genes in signaling regions, and the associated regulatory pathways of these genes.

Some of the enriched signaling pathways were not believed to be related to the adaptations of indigenous goats to domestication in the present study, such as the Riboflavin metabolism pathway, Lipoic acid metabolism pathway, Linoleic acid metabolism pathway of metabolism system and prolactin signaling pathway of endocrine system. However, these pathways are involved in the basic regulation of physiological metabolic and endocrine processes. Therefore, it can be speculated that the selected genes associated with these pathways could be involved in responses resulting from domestication, possibly providing new clues to the functions of such signaling pathways.

Conclusion

To our knowledge, more than 3,000 candidate genes have been found that may have been selected for through both natural and artificial processes during domestication. These candidate genes require validation with additional samples in the future. In addition, novel immune and olfactory perception genetic regulatory pathways were identified, which may correspond to defensive adaptations to the extreme environments in which Chinese native goat breeds were forced to survive during the process of domestication. CNGA4 and Camk 2b may enhance immunity and defense against external adverse factors during domestication. Furthermore, these results shed new light on the field of evolution and provide a novel basis for the identification of new genes affecting goat domestication and evolution.

Declarations

Acknowledgements

We thank the editor and anonymous reviewers for their constructive comments. We greatly appreciated our collaborators, for their assistance in sample collection.

Funding

This research was supported by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP), Grant No.2019 QZKK0501. Strategic Priority Research Program of Chinese Academy of Sciences,

Pan-Third Pole Environment Study for a Green Silk Road (Pan-TPE) (4-1-3-3 XDA2004010303) . Dr Zhao supported by K.C. Wong Education Foundation.

Authors' contributions

DQ and KZ designed the experiments and revised the manuscript. BZ conducted the RNA-seq and qRT-PCR analyses. DT analyzed the data and wrote the paper. FT conducted WGCNA analysis. SL and BH prepared the DNA samples for SNP identification and genotyping and analyzed the data. All authors read and approved the final manuscript.

Ethics approval and consent to participate

All animal experiments were approved by the Animal Care and Use Committees of the Northwest Institute of Plateau Biology, Chinese Academy of Sciences. We obtained written permission from the owner of the sheep farm to take samples.

Consent for publication

Not Applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The raw reads produced in this study were deposited in the NCBI SRA with the accession number SRA SUB9339333 under Bio-project PRJNA723729.

References

1. Zeder, Melinda, A., Hesse & Brian. The Initial Domestication of Goats (Capra hircus) in the Zagros Mountains 10,000 Years Ago. *Science (80-.).* (2000).
2. Naderi, S. *et al.* The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 17659–17664 (2008).
3. Liu, Y. P., Cao, S. X., Chen, S. Y., Yao, Y. G. & Liu, T. Z. Genetic diversity of Chinese domestic goat based on the mitochondrial DNA sequence variation. *J. Anim. Breed. Genet.* **126**, 80–89 (2015).
4. Wang, X. *et al.* Whole-genome sequencing of eight goat populations for the detection of selection signatures underlying production and adaptive traits. *Rep* **6**, 38932 (2016).
5. Simonson, T. S., Yang, Y., Huff, C. D., Yun, H. & Ge, R. L. Genetic Evidence for High-Altitude Adaptation in Tibet. *Science (80-.).* **329**, 72–75 (2010).
6. Scheinfeldt, L. B. *et al.* Genetic adaptation to high altitude in the Ethiopian highlands. *Genome Biol.* **13**, R1 (2012).

7. Badr, B. *et al.* Characterizing neutral genomic diversity and selection signatures in indigenous populations of Moroccan goats (Capra hircus) using WGS data. *Front. Genet.* **6**, 107 (2015).
8. Zhang, Z. *et al.* Whole-genome resequencing reveals signatures of selection and timing of duck domestication. *Gigascience* 4 (2018).
9. Zhao, S. *et al.* Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. *Nat. Genet.* **45**, 67-U99 (2013).
10. Strong signatures of selection in the domestic pig genome. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 19529–19536 (2012).
11. Lv, F. H., Agha, S., Kantanen, J., Colli, L. & Marsan, P. A. Adaptations to Climate-Mediated Selective Pressures in Sheep. *Mol. Biol. Evol.* **31**, 3324–3343 (2014).
12. Liu, Y. H. *et al.* Whole-genome sequencing of African dogs provides insights into adaptations against tropical parasites. *Mol. Biol. Evol.* **35**, (2017).
13. Yang *et al.* Reference genome of wild goat (capra aegagrus) and sequencing of goat breeds provide insight into genic basis of goat domestication. *BMC Genomics* (2015).
14. Witt, K. E. & Huerta-Sánchez, E. Convergent evolution in human and domesticate adaptation to high-altitude environments. *Philos. Trans. R. Soc. B Biol. Sci.* **374**, (2019).
15. Kai, W., Li, M. & Hakon, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* e164 (2010).
16. Felsenstein, J. PHYLIP-Phylogeny Inference Package (Version 3.2). *Cladistics-the Int. J. Willi Hennig Soc.* **5**, 164–166 (1989).
17. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
18. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
19. Loiselle, B. A., Sork, V. L., Nason, J. & Graham, C. Spatial genetic structure of a tropical understory shrub, PSYCHOTRIA OFFICINALIS (RuBIACEAE). *Am. J. Bot.* **82**, (1995).
20. Hardy, O. J. & Vekemans, X. SPAGeDI: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* **2**, 618–620 (2002).
21. Nei, M. & Li, W. H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci.* **76**, 5269–5273 (1979).
22. Ji, Y. *et al.* Whole-Genome Sequencing of Native Sheep Provides Insights into Rapid Adaptations to Extreme Environments. *Mol. Biol. Evol.* **33**, 2576–2592 (2016).
23. Consortium, H. M. Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. *Science (80-.).* **324**, 528 (2009).
24. Mei, C. *et al.* Genetic Architecture and Selection of Chinese Cattle Revealed by Whole Genome Resequencing. *Mol. Biol. Evol.* msx322 (2017).

25. Wang, W., Zhang, X., Zhou, X., Zhang, Y. & Jiang, Z. Deep Genome Resequencing Reveals Artificial and Natural Selection for Visual Deterioration, Plateau Adaptability and High Prolificacy in Chinese Domestic Sheep. *Front. Genet.* **10**, (2013).
26. Mullaney, J. M., Mills, R. E., Pittard, W. S. & Devine, S. E. Small insertions and deletions (INDELs) in human genomes. *Hum. Mol. Genet.* **19**, R131-6 (2010).
27. Kim, J. *et al.* KoVariome: Korean National Standard Reference Variome database of whole genomes with comprehensive SNV, indel, CNV, and SV analyses. *Sci. Rep.* **8**, 5677 (2018).
28. Stafuzza, N. B., Zerlotini, A., Lobo, F. P., Yamagishi, M. & Silva, M. V. G. B. _Da_ B. T.-A. M. & T. S. Single nucleotide variants and InDels identified from whole-genome re-sequencing of Guzerat, Gyr, Girolando and Holstein cattle breeds. in (2018).
29. Wang, K. *et al.* Genome-wide variation within and between wild and domestic yak. *Mol. Ecol. Resour.* **14**, 794–801 (2014).
30. Madjdpour, C., Jewell, U. R., Kneller, S., Ziegler, U. & Beck-Schimmer, B. Decreased alveolar oxygen induces lung inflammation. *AJP Lung Cell. Mol. Physiol.* **284**, L360-7 (2003).
31. Im, S. H., Hueber, A., Monticelli, S., Kang, K. H. & Rao, A. Chromatin-level Regulation of the IL10 Gene in T Cells. *J. Biol. Chem.* **279**, 46818–46825 (2004).
32. Quach, H. *et al.* Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. *Cell* **167**, 643-656.e17 (2016).
33. Schmitt, E., Klein, M. & Bopp, T. Th9 cells, new players in adaptive immunity. *Trends Immunol.* **35**, 61–68 (2014).
34. Heeb, L., Egholm, C. & Boyman, O. Evolution and function of interleukin-4 receptor signaling in adaptive immunity and neutrophils. *Genes Immun.* 1–7 (2020).
35. Souëf, P. N. L., Goldblatt, J. & Lynch, N. R. Evolutionary adaptation of inflammatory immune responses in human beings. *Lancet* **356**, 242–244 (2000).
36. Dielenberg, R. A. & Mcgregor, I. S. Defensive behavior in rats towards predatory odors: a review. *Neurosci Biobehav Rev* **25**, 597–609 (2001).
37. Takahashi, L. K., Nakashima, B. R., Hong, H. & Watanabe, K. The smell of danger: a behavioral and neural analysis of predator odor-induced fear. *Neurosci. Biobehav. Rev.* **29**, 1157–1167 (2005).
38. Vyas, A., Kim, S. K., Giacomini, N., Boothroyd, J. C. & Sapolsky, R. M. Behavioral changes induced by Toxoplasma infection of rodents are highly specific to aversion of cat odors. *Proc. Natl. Acad. Sci.* **104**, 6442–6447 (2007).
39. Pasteels, J. & Daloze, D. Pheromones and Animal Behaviour: Communication by Smell and Taste. *Chemoecology* **13**, (2003).
40. Stowe, M. K., Turlings, T. C., Loughrin, J. H., Lewis, W. J. & Tumlinson, J. H. The chemistry of eavesdropping, alarm, and deceit. *Proc. Natl. Acad. Sci.* **92**, 23–28 (1995).
41. Munger & S., D. Central Role of the CNGA4 Channel Subunit in Ca2+-Calmodulin-Dependent Odor Adaptation. *Science (80-.).* **294**, 2172–2175 (2001).

42. Kelliher, K. R., Ziesmann, J., Munger, S. D., Reed, R. R. & Zufall, F. Importance of the CNGA4 channel gene for odor discrimination and adaptation in behaving mice. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 4299 (2003).
43. Trinh, K. & Storm, D. R. Detection of Odorants through the Main Olfactory Epithelium and Vomeronasal Organ of Mice. *Nutr. Rev.* (2004).
44. Zhe, Z. *et al.* Deletion of Type 3 Adenylyl Cyclase Perturbs the Postnatal Maturation of Olfactory Sensory Neurons and Olfactory Cilium Ultrastructure in Mice. *Front. Cell. Neurosci.* **11**, (2017).
45. Kensaku, M., Hiroyuki, M. & Kimiya, N. Possible functional role of olfactory subsystems in monitoring inhalation and exhalation. *Front. Neuroanat.* **8**, 107 (2014).
46. Luo, M. The necklace olfactory system in mammals. *J. Neurogenet.* **22**, 229–238 (2008).
47. Fang, M., Feng, C., Zhao, Y. X. & Liu, X. Y. Camk2b protects neurons from homocysteine-induced apoptosis with the involvement of HIF-1α signal pathway. *Int. J. Clin. Exp. Med.* **7**, 1659–1668 (2014).
48. Yin, X. Y., Cheng, G. H., Guo, H. Y., Wang, Q. & Li, Y. J. Single cell transcriptome profiling revealed differences in gene expression during oocyte maturation in Haimen white goats. *Genet. Mol. Res. GMR* **16**, (2017).

Figures

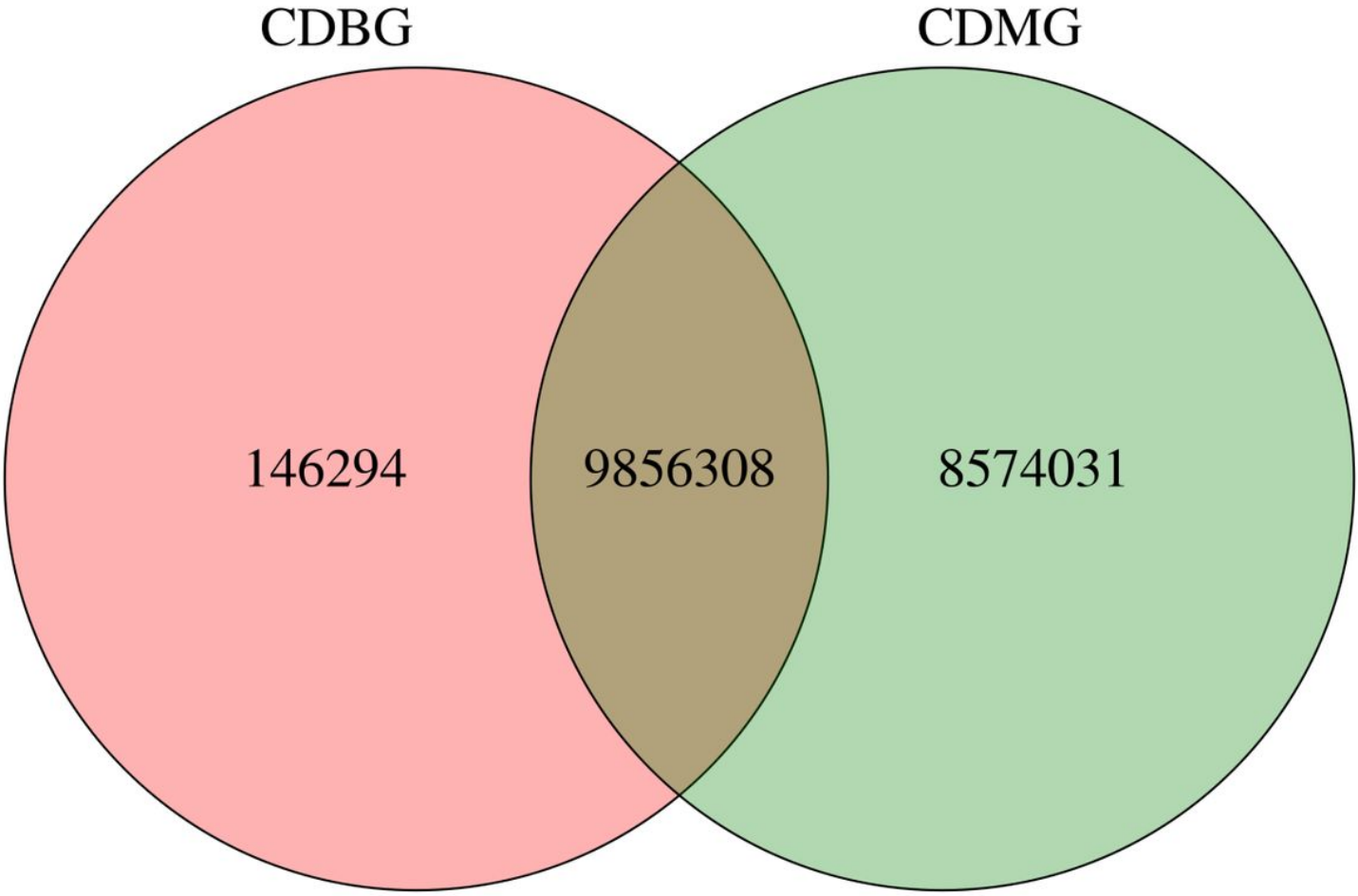

Figure 1

Number of SNPs for the two populations are indicated in the Venn diagram.

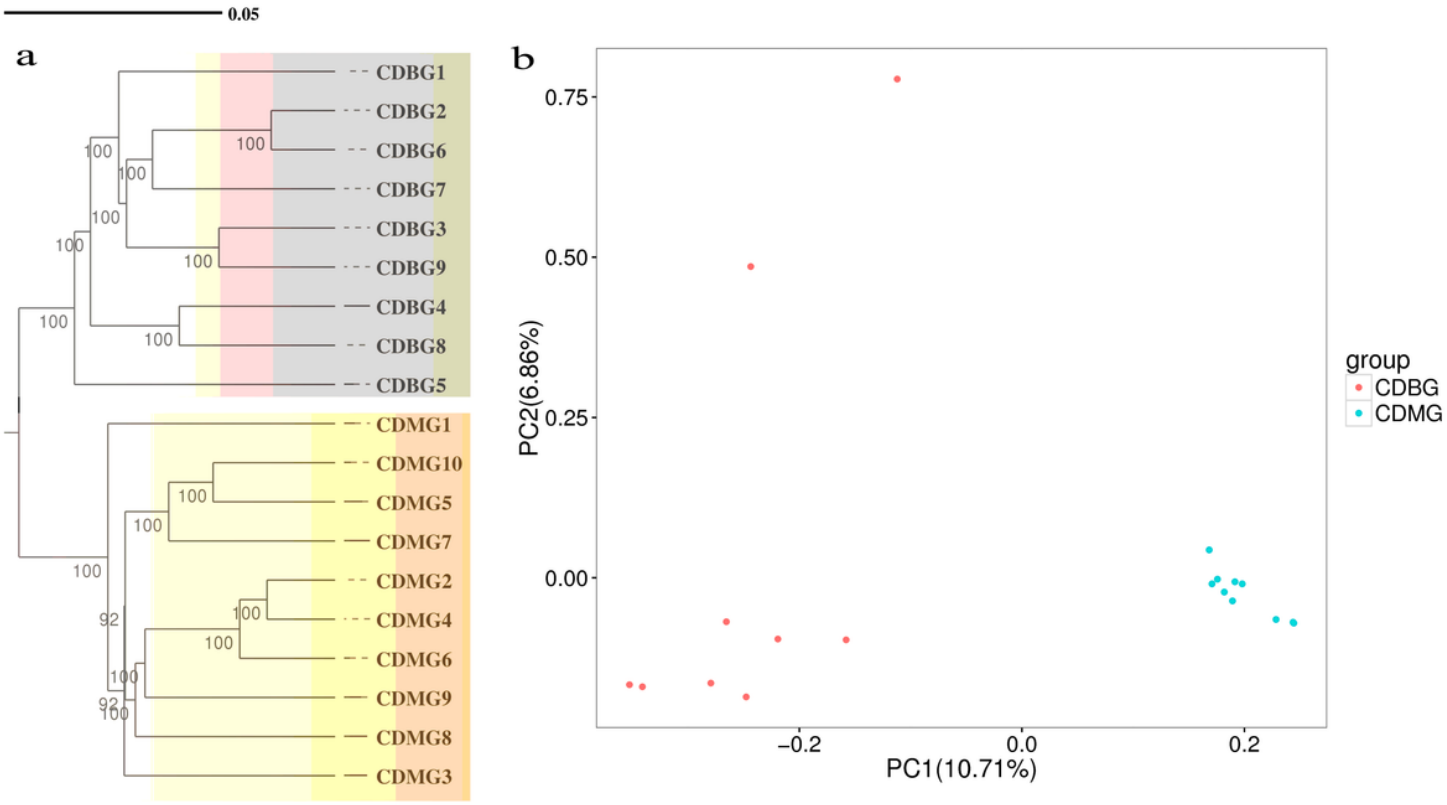

Figure 2

Genome-wide relationship between CDMG and CDBG. (a) Neighbour-joining tree for the 19 goats. The number adjacent to each branch is the bootstrap value. (b) The PCA of the 19 goats. CDMG and CDBG are separated along the first principal component (PC1).

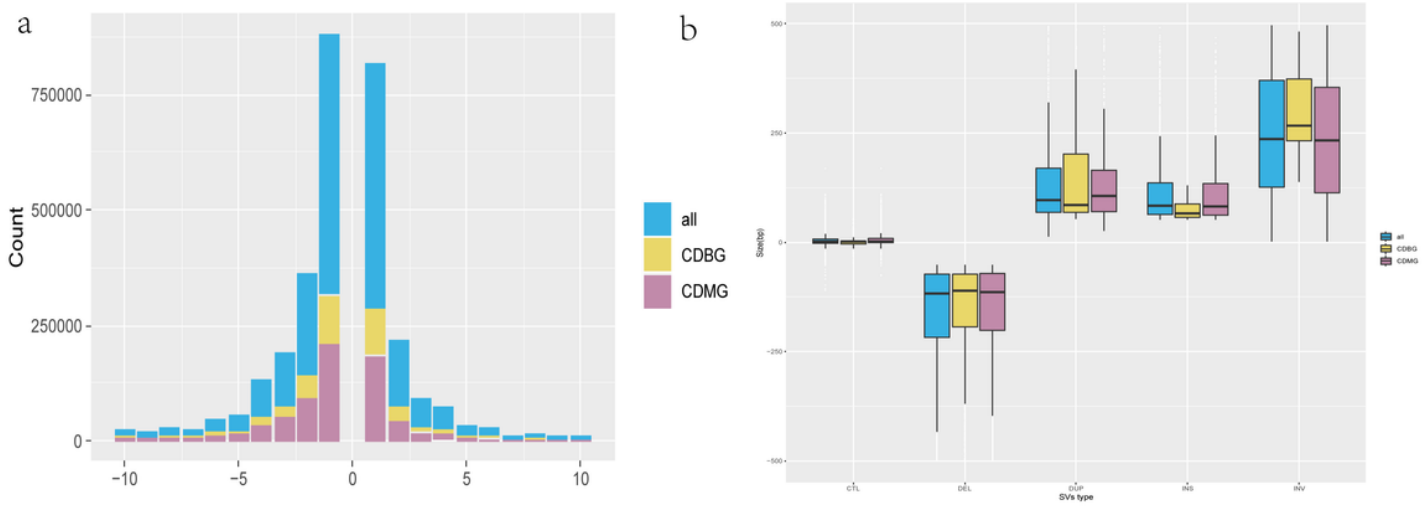

Figure 3

An overview of the variation in yak genome. (a) The relationship between the number of InDels and their size. The CDMG InDels are shown in purple, the CDBG InDels in yellow and the InDels present in both

populations in blue. (b) The size distribution of different types of SVs.

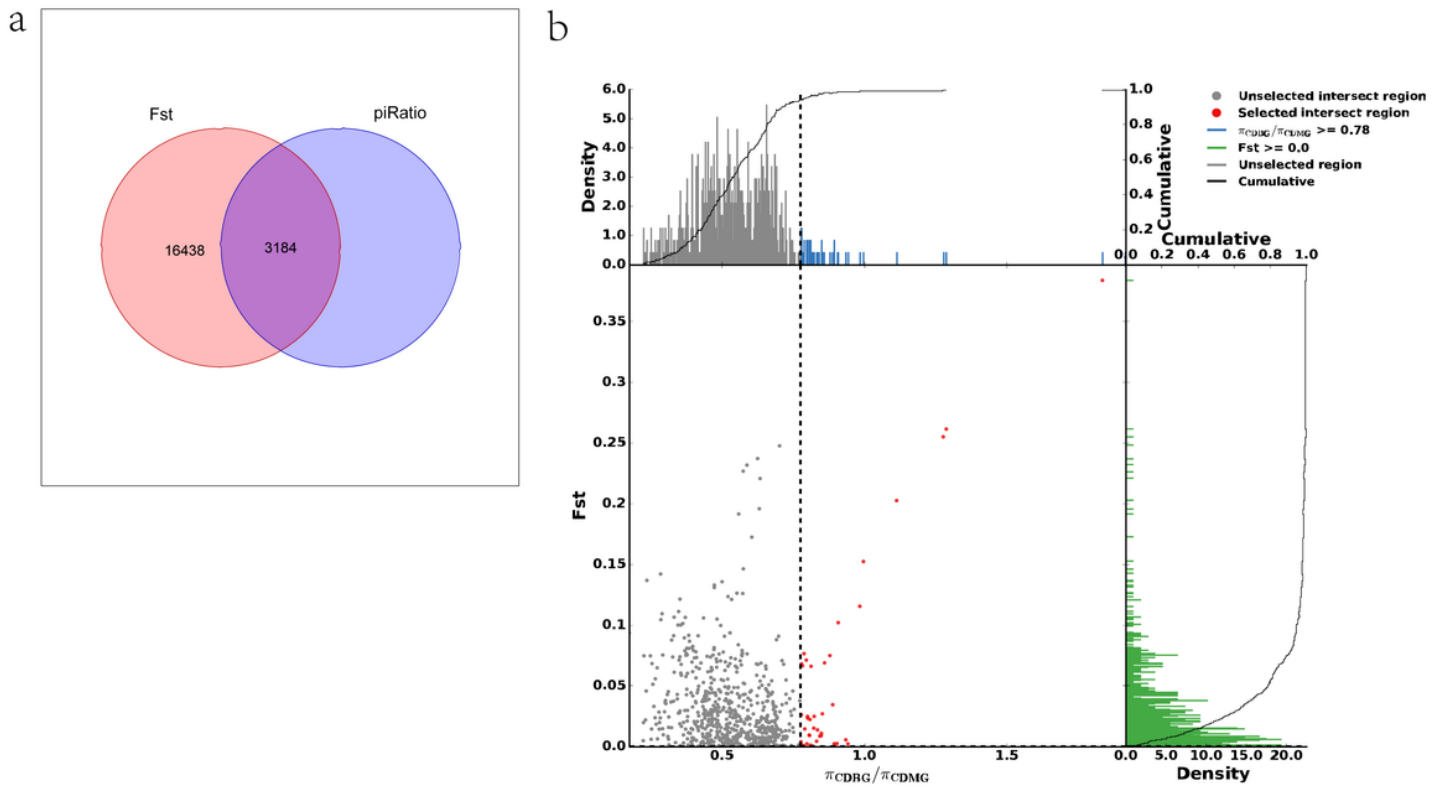

Figure 4

An overview of the variation in yak genome. (a) The relationship between the number of InDels and their size. The CDMG InDels are shown in purple, the CDBG InDels in yellow and the InDels present in both populations in blue. (b) The size distribution of different types of SVs.

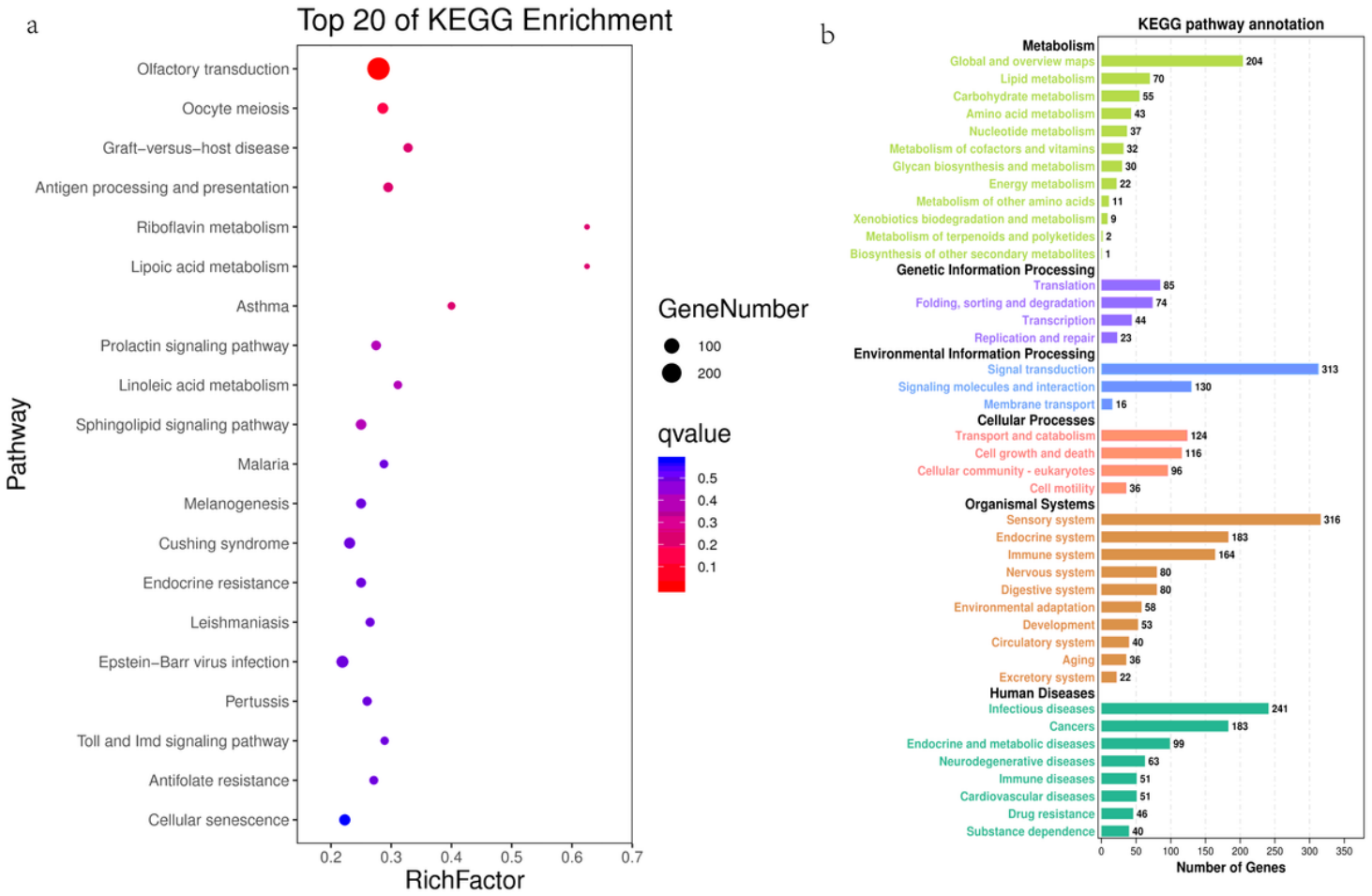

Figure 5

KEGG enrichment analysis of selected genes. (a) Enriched selected genes on twenty KEGG pathways. (b) The number of genes annotated on different KEGG pathways.

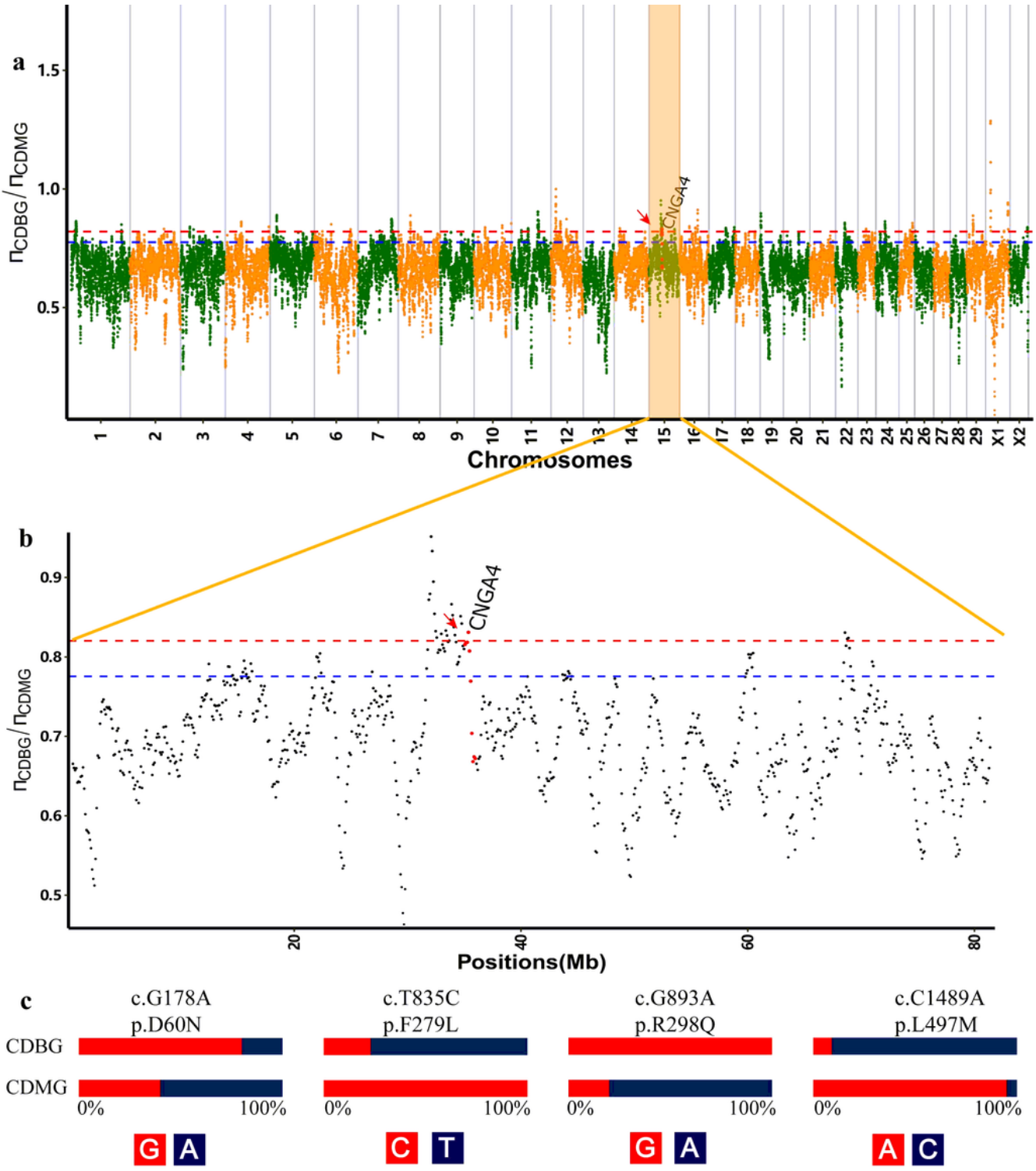

Figure 6

Signatures of the selective sweep of the CNGA4 gene region in CDMG. (a) π ration values across a chromosome 15 region harboring the CNGA4 gene. (b) π ration values around the CNGA4 gene region in chromosome 15. (c) The allele frequencies of four SNPs within the CNGA4 gene between CDMG and CDBG.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfileschecklist.doc
- SupplementaryTable.xlsx