



## OPEN ACCESS

## EDITED BY

Pedro Aponte,  
Universidad San Francisco de Quito, Ecuador

## REVIEWED BY

Regiane R. Santos,  
Schothorst Feed Research, Netherlands  
Ke Wang,  
Northwest A&F University, China

## \*CORRESPONDENCE

Yong Han  
✉ 2642900225@qq.com

RECEIVED 08 February 2025

ACCEPTED 22 July 2025

PUBLISHED 13 August 2025

## CITATION

Zhao Y, Han Y, Yuan C, Yang Y, Long Y and  
Xiao W (2025) Genome-wide selection signal  
analysis reveals copy number variation  
associated with litter size in Guizhou Black  
goat.

*Front. Vet. Sci.* 12:1573093.

doi: 10.3389/fvets.2025.1573093

## COPYRIGHT

© 2025 Zhao, Han, Yuan, Yang, Long and  
Xiao. This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited,  
in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Genome-wide selection signal analysis reveals copy number variation associated with litter size in Guizhou Black goat

Yanpin Zhao<sup>3</sup>, Yong Han<sup>2,1\*</sup>, Chao Yuan<sup>2</sup>, Yang Yang<sup>2</sup>,  
Yong Long<sup>3</sup> and Wen Xiao<sup>3</sup>

<sup>1</sup>Guizhou University of Traditional Chinese Medicine, Guiyang, China, <sup>2</sup>Institute of Animal Husbandry and Veterinary Sciences, Guizhou Academy of Agricultural Sciences, Guiyang, China, <sup>3</sup>College of Animal Science, Guizhou University, Guiyang, China

**Introduction:** Copy number variation (CNV) is a common form of genomic structural variation arising from genome sequence rearrangements, which primarily involves variations in the copy number of large genomic segments. This study focuses on performing selection signal analysis of CNVs to identify candidate loci and genes associated with litter size in Guizhou Black goat.

**Methods:** Sample selection: 30 Guizhou Black goats were divided into high-lambing and low-lambing groups based on their continuous lambing records. Genetic analysis: Phylogenetic tree construction and principal component analysis (PCA) were performed to assess genetic differences between the two groups. Selection signal indicators: Genetic differentiation index ( $F_{ST}$ ) and nucleotide diversity ( $\theta_{\pi}$ ) were used as indicators for identifying selection signals. CNV identification: 180 CNVs associated with litter size were identified through selection signal analysis. Functional annotation: Candidate genes within these CNVs were annotated, and further analysis was conducted using Gene Ontology (GO) and KEGG pathway enrichment methods.

**Results:** Genetic differentiation: Phylogenetic tree construction and PCA revealed significant differences in genetic structure between the high-lambing and low-lambing groups. CNV identification: 180 CNVs were found to be associated with litter size. Candidate genes: Functional annotation of the CNVs identified 49 candidate genes. Pathway enrichment: GO and KEGG pathway analysis revealed significant enrichment of these genes in several reproduction-related biological pathways, including the Hippo signaling pathway, steroid hormone biosynthesis, and retinol metabolism.

**Discussion:** This study provides novel insights into the genetic mechanisms underlying prolificacy in Guizhou Black goats. The identification of CNVs associated with litter size offers potential targets for genomic selection breeding, which could be used to improve reproductive traits in this breed. The enriched pathways, such as the Hippo signaling pathway, steroid hormone biosynthesis, and retinol metabolism, suggest that these biological processes may play a critical role in regulating fertility and reproductive efficiency in goats. The findings highlight the potential for utilizing CNVs as markers for improving reproductive performance in livestock breeding programs.

## KEYWORDS

CNV, goats, litter size, whole genome sequencing, select signal

## 1 Introduction

As economically significant livestock, goats contribute substantially to the global production of meat, milk, and hides. Increasing production volume alone is insufficient to meet the growing global demand for goat-derived products. Enhancing goat reproductive performance is a pressing concern, as litter size is a critical trait influencing reproductive efficiency and overall productivity. Consequently, enhancing litter size has become a primary objective in goat breeding programs. While factors such as maternal nutrition and dietary supplements (e.g., folic acid) play roles in determining litter size, genetic factors are also crucial. Recently, many studies have focused on identifying genetic variants and genes associated with litter size in goats. For instance, a genome-wide association study (GWAS) identified single nucleotide polymorphism (SNP) variants associated with litter size at birth (LSB) in Markhoz goats. Candidate genes annotated from these variants include GABRA5, AKAP13, and SV2B (1). Similarly, when analyzing SNPs associated with litter size in Chuankhong Black goats, Guo et al. utilized  $F_{ST}$  and  $\theta_{\pi}$  analyses to examine population selection signals. Their findings annotated 506 and 528 genes in high-litter and low-litter groups, respectively. Additionally, polymorphism analysis revealed a significant association between the g.89172108 T > G variant in the exon of the AMH gene and litter size, suggesting its potential as a marker for selecting high-fertility individuals (2). As a complex trait, litter size is influenced by multiple genetic factors. Focusing exclusively on SNPs may not provide a comprehensive understanding of their genetic basis. Copy number variation (CNV), an essential form of structural variation involving changes in the copy number of large genomic segments, can significantly impact gene expression and function. Increasing evidence indicates that CNVs play a vital role in animal reproduction. For example, a CNV-based GWAS in Canadian Holstein cattle identified two CNVRs linked to three reproductive traits: calf survival rate, interval from first service to conception, and non-return rate (3).

The Guizhou Black Goat is a local meat-type breed native to the southwestern region of Guizhou Province, China. It is characterized by a moderate body size, tender meat with a mild odor, and strong adaptability to cold climates and low-quality forage. However, its relatively low reproductive performance has become a limiting factor for the breed's development and widespread utilization. This study aims to identify copy number variations (CNVs) associated with lambing traits by analyzing whole-genome sequencing data of high- and low-litter-size Guizhou Black goat groups. We conducted selection signal analysis of CNVs between these groups to identify candidate genes potentially linked to litter size. These genes contribute to elucidating the genetic mechanisms underlying lambing traits and serve as valuable molecular markers for genetic improvement and enhanced reproductive performance in goats.

## 2 Materials and methods

### 2.1 Sample collection and whole genome resequencing

The 30 Guizhou Black goats used in this study were all sourced from Maiping Town, Huaxi District, Guizhou Province. The selected

goats were healthy females aged 2–3 years, raised under uniform management conditions, and underwent annual disease screening. All experimental goats were fed a semi-grazing system, provided with the same basic diet (including silage corn, hay and concentrate) during the dry season, and allowed to eat natural pasture freely during the green season. Based on lambing records from the past 2 years, 15 high-litter-size individuals (average litter size  $\geq 2$ ) were selected to form the high-lambing group (HL), while 15 low-litter-size individuals (average litter size = 1) were chosen as the low-lambing group (LL) for whole-genome resequencing. Blood samples were collected using vacuum blood collection tubes between 8:00 and 9:00 in the morning. The studies involving animals were reviewed and approved by Guizhou University Subcommittee of Experimental Animal Ethics [eae-gzu-2024-e054]. Blood samples were collected for DNA extraction, and high-quality DNA samples were subsequently used to construct sequencing libraries. The DNBSEQ-T7 sequencer was employed to sequence, producing raw sequencing data with an average coverage depth of 33.3X.

### 2.2 CNV detection and annotation

To minimize errors caused by human factors during sequencing, the raw data were first filtered using Fastp (v0.23.4), and reads containing adapters were removed based on the following criteria. Paired reads were discarded if the proportion of “N” bases in a sequencing read exceeded 1% of the total bases. Additionally, paired reads were excluded if the number of low-quality bases ( $Q \leq 5$ ) in a sequencing read exceeded 50% of the total bases. Subsequently, BWA (v0.7.17) was used to align the clean reads to the reference genome (GCF\_001704415.2\_ARSL1.2\_genomic.chr), and SAMtools (v1.17) was employed to calculate basic metrics, such as sequencing depth and genome coverage, to prepare for subsequent variant detection. For structural variant (SV) detection, Manta was used to identify SVs for each individual. SVs of the same type were merged across individuals based on genomic position and variant length to remove redundancy. Paragraph genotyping was employed to genotype each individual using the deduplicated SV results. At the population level, further deduplication was performed based on positional overlap criteria (50% for deletions; 90% for insertions, inversions, and duplications) and genotype consistency (population genotype consistency  $\geq 0.95$ ).

Quality control was conducted on the deduplicated SVs using the following filtering criteria: variant length (ABS (INFO/SVLEN)  $\leq 10,000,000$ ), allelic frequency heterozygosity (INFO/ExcHet)  $\geq 0.05$ , missing rate (F\_MISSING)  $\leq 0.2$ , and minor allele frequency (INFO/MAF)  $> 0$ . The Plink software was used to calculate CNV frequencies across all individuals, and R (v4.2.0) was used for visualization. Custom Perl scripts were employed for CNV analyses specific to different groups. Finally, all CNVs were annotated using Annotvar.

### 2.3 Population structure analysis and selective feature detection

The experimental animals were divided into high-litter-size (HL,  $n = 15$ ) and low-litter-size (LL,  $n = 15$ ) groups based on their lambing records. Selection signature detection was then performed on these two groups. To understand the genetic structure of the HL and LL

groups, principal component analysis (PCA) was conducted using the GCTA software (4) with the make-grm and pca commands. The PCA results were visualized using R software.

Next, two methods, genetic differentiation index ( $F_{ST}$ ) and nucleotide diversity ( $\theta\pi$ ), were employed for selection signature analysis using VCFtools (v0.1.17). For  $F_{ST}$  analysis, a sliding window approach was applied with a 100 kb chromosome interval and a 10 kb step size. Subsequently, nucleotide diversity ( $\pi$ ) values were calculated separately for each group, and the relative difference in nucleotide diversity between the groups was determined as  $\log_2(\pi_{HL}/\pi_{LL})$ . Candidate loci were identified by selecting those present in the top 1% of windows from both the  $F_{ST}$  sliding window analysis and the  $\log_2(\pi_{HL}/\pi_{LL})$  sliding window analysis.

## 2.4 Functional annotation and pathway enrichment of candidate gene

The resulting data were filtered further after identifying candidate CNV loci within the selected regions based on the top 1% of  $F_{ST}$  sliding window analysis and  $\log_2(\pi_{HL}/\pi_{LL})$ . Goat gene annotation files were downloaded from the ENSEMBL database, and bedtools was used to annotate the candidate CNVs to their corresponding genes. Loci that were not enriched for genes or had negative  $F_{ST}$  values were excluded. Differential gene sets were then subjected to Gene Ontology (GO) and KEGG pathway enrichment analysis using the clusterProfiler package in R (v4.4.0). This analysis aimed to identify significantly enriched biological processes and pathways associated with the candidate CNVs.

## 3 Results

### 3.1 Identification and distribution of CNVs

Following variant detection and filtering, a total of 2,083 CNV sites were retained. All detected CNVs were deletions, and the identification statistics are presented in Figure 1. A length distribution analysis of all CNVs revealed that most loci fell within the 1–5 kb range. Functional annotation of CNV regions indicated that approximately 60.5% were located in intergenic regions, 23.8%

in intronic regions, and only 10.9% in exonic regions. The proportions of CNVs in downstream (0.4%) and upstream (0.3%) regions were relatively low. CNVs related to non-coding RNA were also uncommon, with ncRNA exons accounting for 1.9% and ncRNA introns for 1.2%. Additionally, CNVs were identified in 3' untranslated regions (3' UTR, 0.4%), 5' untranslated regions (5' UTR, 0.2%), splicing sites (0.2%), and overlapping upstream and downstream regions (0.05%).

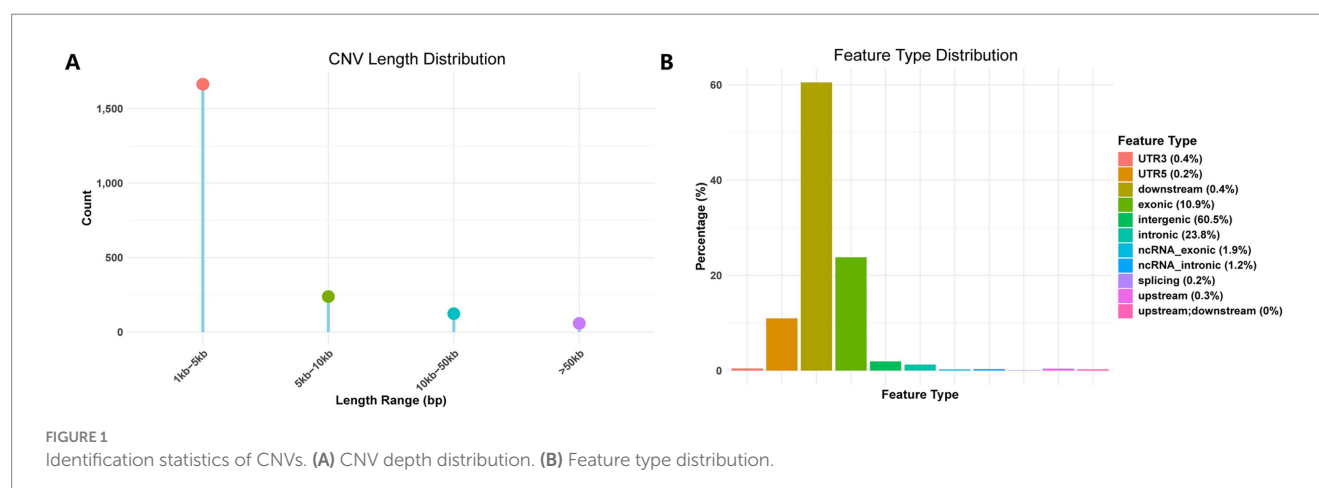
We analyzed the chromosomal distribution of CNVs and the number of CNVs at different frequencies in the HL and LL groups, and the results are shown in Figure 2. For statistical evaluation of the CNV distribution, we performed a chi-square test for the CNV distribution of each chromosome in both groups with a Bonferroni correction for multiple testing. The results, summarized in Supplementary Tables S1, S2, indicated no significant differences in CNV distribution at the chromosomal level between the HL and LL groups. Furthermore, the frequency statistics of CNVs showed no significant differences between the two groups. It is worth noting that, the results of Fisher's exact test showed that the frequencies of deletion CNVs were significantly different between the HL and LL groups in the frequency ranges of 0–0.1 and 0.2–0.3.

### 3.2 Analysis of group structure

The results of the principal component analysis (PCA) are presented in Figure 3. Despite some overlap between individuals from the HL and LL groups, the two groups display a distinct differentiation trend. This observation is corroborated by the phylogenetic tree in Figure 3, where all individuals are predominantly clustered into two separate branches, this also indicates that there is a clear differentiation between the two groups.

### 3.3 Selective signal analysis of HL and LL

The results of the selection signal analysis are presented in Figure 4. A total of 180 CNV loci were selected based on the relative difference in  $\pi$  values,  $\log_2(\pi_{HL}/\pi_{LL})$ , and the top 1% of  $F_{ST}$ . After annotation, 49 candidate genes (Table 1) were identified.



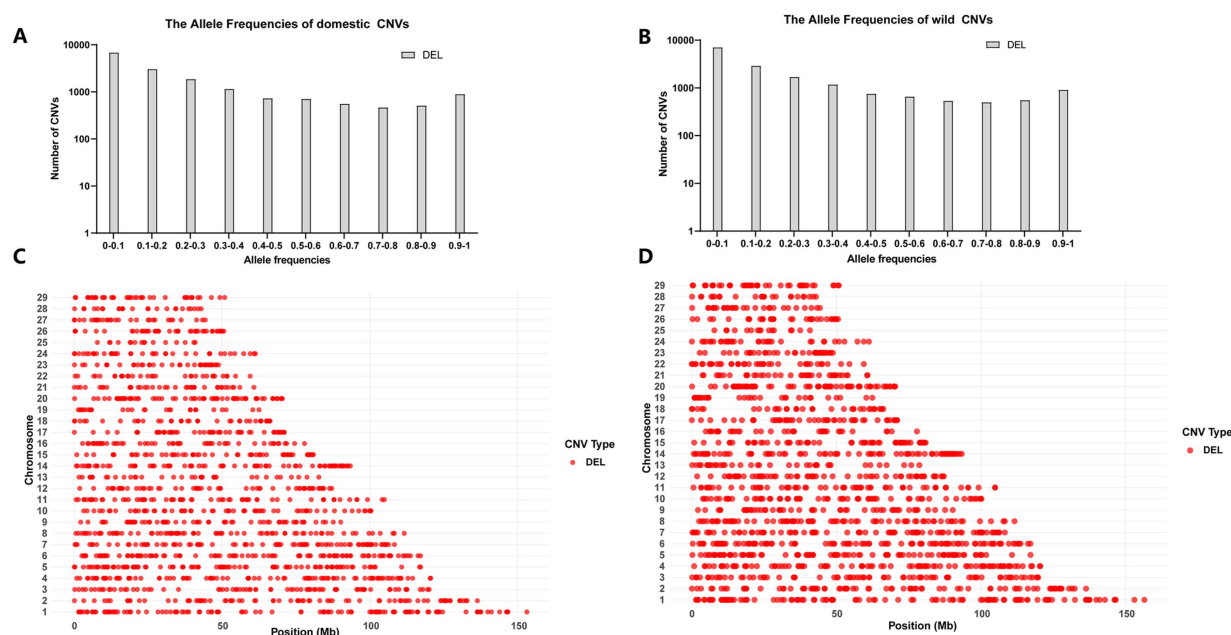


FIGURE 2

Distribution statistics of CNVs. (A) The number of CNVs at different frequencies in the HL group. (B) The number of CNVs at different frequencies in the LL group. (C) The distribution of chromosome CNVs in the HL group. (D) The distribution of chromosome CNVs in the LL group.

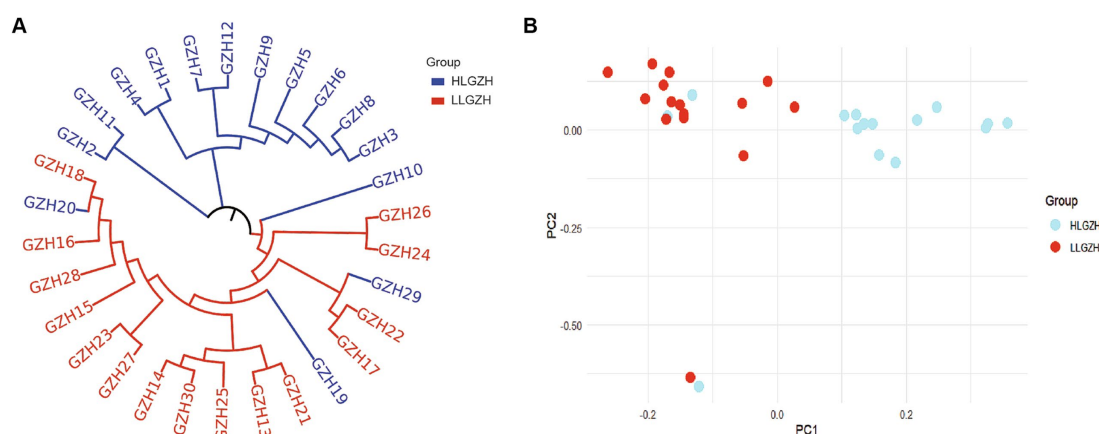


FIGURE 3

(A) Phylogenetic tree. (B) Principal component analysis.

### 3.4 GO and KEGG enrichment analysis of candidate genes

The primary significantly enriched pathways identified by GO and KEGG analyses are presented in Figure 5. A total of 59 KEGG pathways were enriched, including 14 significant pathways ( $p < 0.05$ ), primarily associated with metabolism, hormones, signaling, and chemical carcinogenesis. The most significantly enriched pathway was ascorbate and aldate metabolism (chx00053). In the GO enrichment analysis, 532 pathways were identified, of which 134 were significant ( $p < 0.05$ ). Among these, 453 biological processes (BP), 24 cellular components (CC), and 55 molecular functions (MF) were enriched, primarily related to apoptosis, stress response, lipid metabolism, development and regeneration, signal transduction, protein interactions, gene

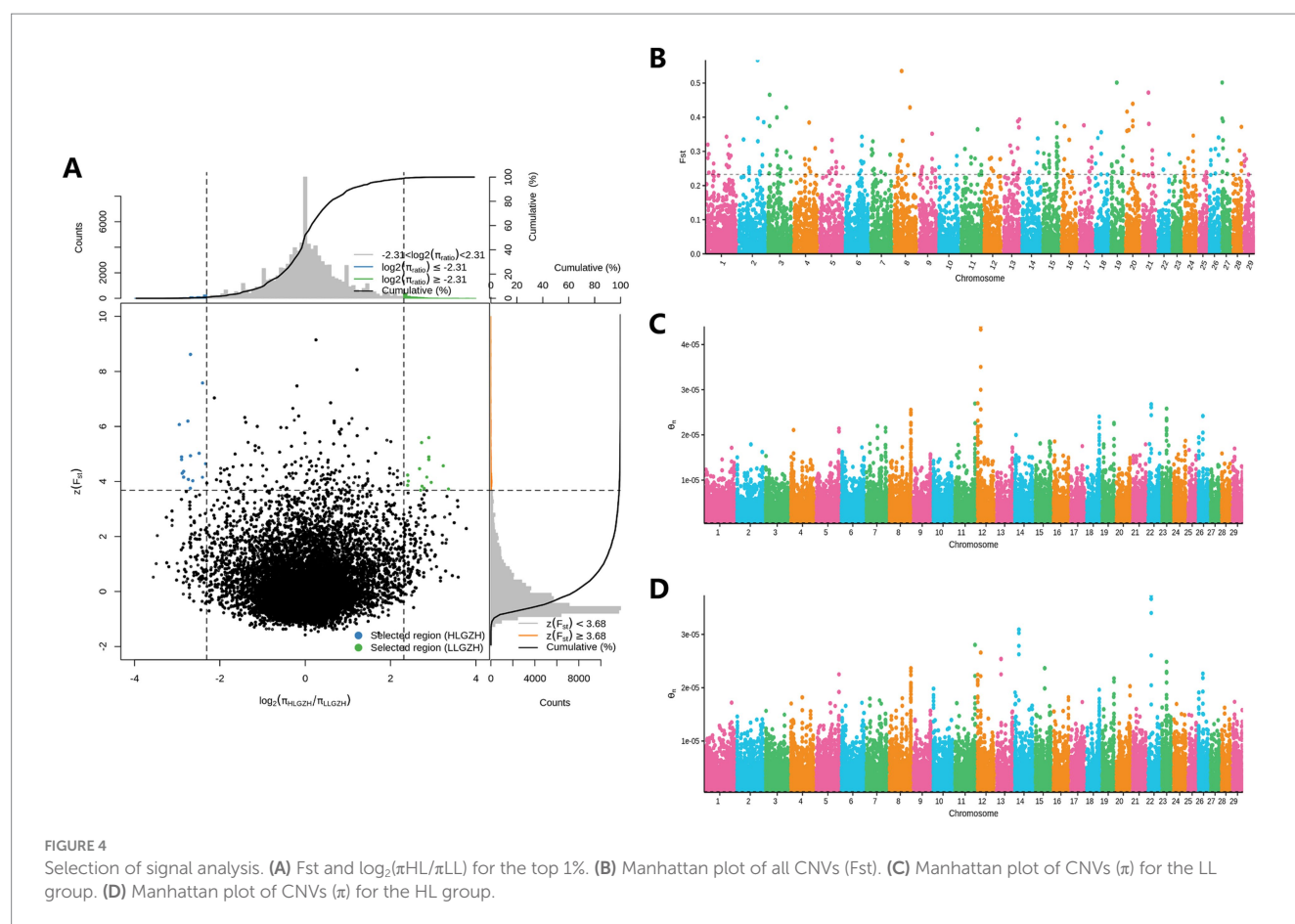
expression, and repair. The most significantly enriched biological process was positive regulation of cell death (GO:0010942), the most significantly enriched cellular component was connexin complex (GO:0005922), and the most significantly enriched molecular function was glucuronosyltransferase activity (GO:0015020).

## 4 Discussion

### 4.1 Distribution of CNV

We conducted a statistical analysis of the distribution and types of CNVs between the HL and LL populations. The results indicated that only deletion CNVs showed a significant difference, suggesting





that the deletion copy number may be associated with lambing traits. Previous studies have demonstrated that CNV deletions are linked to animal reproductive performance. For example, Zhang Ruiqian (5) conducted a comparative analysis of CNVs in goats with different litter size trait groups. They found that in the deletion CNV analysis of high-yield and low-yield groups, 173 genes were unique to the high-yield group, indicating that the missing CNV may be related to the high litter size of goats. Similarly, Ting Sun et al. (6) found that a CNV deletion in SPAG16 was associated with reproductive ability (semen motility) in bulls.

## 4.2 Population structure analysis and selection signal analysis

The phylogenetic tree and principal component analysis results revealed clear genetic differentiation between the HL and LL groups, indicating structural genetic differences between the high-yield and low-yield groups, warranting further investigation. We reviewed the relevant literature and compiled statistics for the 49 selected genes (Table 1). Among them, 16 genes were classified as unknown, with no relevant studies identified. Of the remaining 33 genes, 16 have been identified in the literature as being associated with animal reproduction, while the others are linked to diseases, organ development, and cell differentiation. Notably, TRNAC-GCA was selected as a candidate gene for fertility in Katahdin ewes through whole-genome selection signal analysis (7), while STXP6 was

identified as a candidate gene in a meta-analysis of key SNPs related to pig litter size. Additionally, several candidate genes have been implicated in the regulation of ovulation-related processes in animals. For instance, GJA4 (8) supports oocyte maturation, while the roles of LINGO1 (9) and PTGER3 (10) appear more complex and may involve regulatory or inhibitory functions during follicular development and angiogenesis. Half of the 33 known candidate genes have been implicated in reproduction, further supporting the accuracy of the results in this study.

## 4.3 GO and KEGG enrichment

During the enrichment analysis, we identified three GO pathways directly related to placental development: “placenta development,” “embryonic placenta morphogenesis,” and “cell differentiation involved in embryonic placenta development.” Previous studies have shown that placental traits in sheep are associated with litter size (11), and alterations in litter size can lead to a reduction in placental weight (12). Abnormal placental development also can result in the death of some embryos, thereby reducing litter size. Furthermore, the placenta secretes various hormones that regulate pregnancy maintenance and fetal development, which influence litter size. For example, the placenta secretes progesterone (13), and research has shown that progesterone plays a role in determining bovine oocyte quality and embryo development (14). Oocyte quality and embryo development are closely linked to litter size. The two most significant GO terms in

TABLE 1 Candidate gene functions (References (8–10, 21–48) are all listed in the table in top and bottom order).

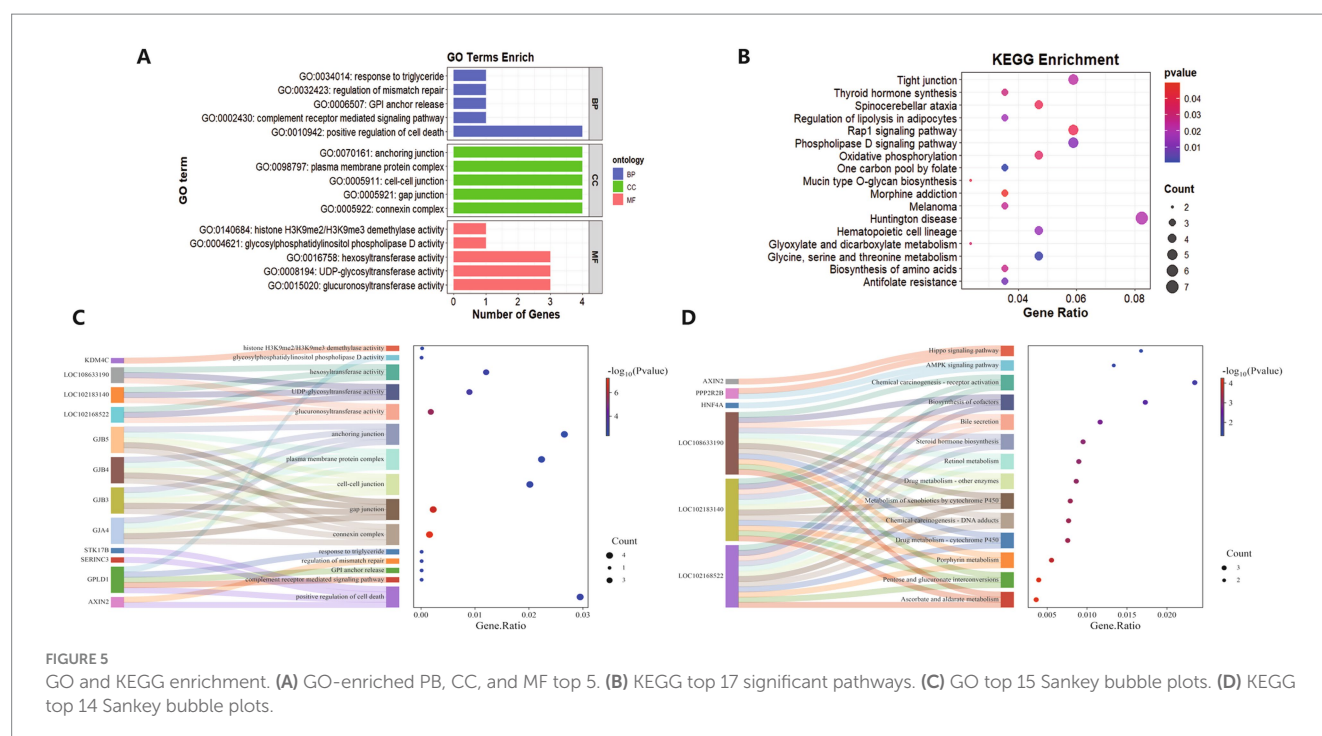
Gene		Functionality	Reference
GJA4	Reproduction	Maturation and developmental competence of bovine oocytes	(8)
LINGO1		Bovine oocyte development	(9)
PTGER3		Promotes the onset of ovulation in monkeys	(10)
ZNF131		Inhibits estrogen signaling	(21)
MIR186		Regulates Follicle Stimulating Hormone (FSH) Expression	(22)
GJB3		Gene expression affects mouse embryo survival	(23)
HNF4A		Human Embryonic Stem Cell Differentiation	(24)
KDM4C		Mouse Embryonic Stem Cell Identity	(25)
PCDH15		Mouse Embryonic Development	(26)
SOX5		Involvement in mouse gonadal development	(27)
AXIN2		Human Embryonic Stem Cell Self-Renewal	(28)
STXBP6		Candidate Genes for Littering Traits in Pigs	(29)
TRNAC-GCA		Candidate genes for fertility in sheep	(7)
STK17B		Causes severe defects in porcine sperm acrosomes	(30)
CEP112		Coordinate spermatogenesis	(5)
DNAH7		Male reproductive disorders (sperm motility)	(32)
ALDH5A1	Disease	Good predictor of ovarian cancer	(33)
DGKE		Causes kidney related diseases	(34)
DLGAP3		Brain disorders	(35)
GPLD1		New therapeutic target for chronic diseases	(36)
PPP2R2B		Inhibitors of Triple Negative Breast Cancer (TNBC)	(37)
SERINC3		Limits HIV-1 Infectivity	(38)
TRNAG-CCC		Involved in inflammatory process of mastitis	(39)
TTPAL		Promotes gastric tumorigenesis	(40)
UTRN		Potential Biomarkers for Breast Cancer	(41)
ZNF438		Polycystic Ovary Syndrome (PCOS) Candidate Genes	(42)
ZRANB2		Involved in Adriamycin Resistance in Breast Cancer Cells	(43)
BET1	Other	Regulates protein transport between the endoplasmic reticulum and the Golgi apparatus	(44)
GJB4, GJB5		Mouse Epididymis Development	(45)
KIAA0319		Human neuroepithelial cell development	(46)
PKIG		Osteoblast differentiation	(47)
R3HDML		Satellite cell proliferation and differentiation	(48)
C19H17orf67, LOC102168522, LOC102179380, LOC102180243, LOC102180519, LOC102180784, LOC102183140, LOC102187692, LOC102188962, LOC102189228, LOC102189502, LOC102189778, LOC106501965, LOC108633190, LOC108638409, SMIM12			NA

the enrichment analysis were “connexin complex” and “gap junction,” both of which are critical components of cell-to-cell communication and closely linked to reproductive function. Studies on gap junction proteins in the female reproductive organs of mice have shown that certain connexins (such as CX43, CX37, and CX26) are vital for female reproductive function (15).

In the KEGG enrichment analysis, we identified several pathways directly related to animal reproduction, such as the Hippo Signaling Pathway, Steroid Hormone Biosynthesis, and Retinol Metabolism. Studies have shown that the Hippo signaling pathway plays a critical role in ovarian development, follicular development, and oocyte

maturation (16). Steroid hormones, including estrogen, progesterone, and testosterone, are directly involved in regulating the reproductive system and serve as core pathways for reproduction, controlling follicular development, ovulation, and pregnancy maintenance (17). Additionally, retinol (vitamin A) plays a vital role in the reproductive system, particularly in regulating embryonic development (18), follicular maturation, and spermatogenesis (19).

Notably, when we created Sankey bubble plots for the GO and KEGG enrichment results, we identified three genes common to both plots: LOC102168522, LOC102183140, and LOC108633190. A review of the relevant literature revealed that, while the overexpression of



LOC102168552 has been shown to promote the growth of mammary epithelial cells (20), no studies have reported on the functional roles of the other two genes. We hypothesize that these three genes may be associated with lambing performance in Guizhou Black goats; however, definitive conclusions can only be drawn through further in-depth investigation of these genes.

## 5 Conclusion

In this study, we conducted a genome-wide selection signal analysis using the genetic differentiation index ( $F_{st}$ ) and nucleotide diversity ( $\theta_{\pi}$ ) to identify CNVs associated with lambing performance in Guizhou Black goats. A total of 2083 CNVs were detected in the HL and LL groups, and 180 CNVs were selected as potentially associated with lambing performance. Furthermore, we identified 49 candidate genes, including TRNAC-GCA, STXPB6, and GJA4, which may be linked to lambing traits.

## Data availability statement

The original sequence data of Guizhou Black Goat used in this paper have been deposited in the National Genome Science Data Center (Nucleic Acids Research 2022) and are publicly accessible at the following URL: <https://bigd.big.ac.cn/gsa/browse/CRA019130>.

## Ethics statement

The studies involving animals were reviewed and approved by Guizhou University Subcommittee of Experimental Animal Ethics [eae-gzu-2024-e055].

## Author contributions

YZ: Conceptualization, Data curation, Writing – original draft, Writing – review & editing. YH: Investigation, Methodology, Writing – original draft, Writing – review & editing. CY: Supervision, Validation, Writing – review & editing. YY: Project administration, Software, Writing – review & editing. YL: Methodology, Project administration, Writing – review & editing. WX: Validation, Formal analysis, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The sources of funding for this study were Innovative Utilization of Livestock and Poultry Germplasm Resources and Construction of Basic Science and Technology Platform (Qiannongke Technology Innovation [2023] 04) and Major Special Projects of the Guizhou Province Department of Science and Technology provided funding for this work (Qianke Service Enterprise [2020] 4009).

## Acknowledgments

Thanks to everyone who helped with this article.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fvets.2025.1573093/full#supplementary-material>

## References

- Mahmoudi P, Rashidi A, Nazari-Ghadikolaei A, Rostamzadeh J, Razmkabir M, Huson HJ. Genome-wide association study reveals novel candidate genes for litter size in Markhoz goats. *Front Vet Sci.* (2022) 9:1045589. doi: 10.3389/fvets.2022.1045589
- Guo C, Ye J, Liu J, Li Z, Deng M, Guo Y, et al. Whole-genome sequencing identified candidate genes associated with high and low litter size in Chuanshong black goats. *Front Vet Sci.* (2024) 11:1420164. doi: 10.3389/fvets.2024.1420164
- Rojas DOH, Chud T, Oliveira GJ, Hermisdorff IC, Narayana SG, Rochus CM, et al. Genome-wide association analyses reveal copy number variant regions associated with reproduction and disease traits in Canadian Holstein cattle. *J Dairy Sci.* (2024) 107:7052–63. doi: 10.3168/jds.2023-24295
- Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis[J]. *Am J Hum Genet.* (2011) 88:76–82. doi: 10.1016/j.ajhg.2010.11.011
- Zhang RQ, Wang JJ, Zhang T, Zhai HL, Shen W. Copy-number variation in goat genome sequence: a comparative analysis of the different litter size trait groups. *Gene.* (2019) 696:40–6. doi: 10.1016/j.gene.2019.02.027
- Sun T, Pei S, Liu Y, Hanif Q, Xu H, Chen N, et al. Whole genome sequencing of simmental cattle for SNP and CNV discovery. *BMC Genomics.* (2023) 24:179. doi: 10.1186/s12864-023-09248-x
- Sanchez-Ramos R, Trujano-Chavez MZ, Gallegos-Sanchez J, Becerril-Perez CM, Cadena-Villegas S, Cortez-Romero C. Detection of candidate genes associated with fecundity through genome-wide selection signatures of Katahdin ewes. *Animals.* (2023) 13:272. doi: 10.3390/ani13020272
- Hao T, Xu X, Hao H, Du W, Pang Y, Zhao S, et al. Melatonin improves the maturation and developmental ability of bovine oocytes by up-regulating GJA4 to enhance gap junction intercellular communication. *Reprod Fertil Dev.* (2021) 33:760–71. doi: 10.1071/RD21145
- Janowski D, Salilew-Wondim D, Torner H, Tesfaye D, Ghanem N, Tomek W, et al. Incidence of apoptosis and transcript abundance in bovine follicular cells is associated with the quality of the enclosed oocyte. *Theriogenology.* (2012) 78:656–669.e5. doi: 10.1016/j.theriogenology.2012.03.012
- Trau HA, Davis JS, Duffy DM. Angiogenesis in the primate ovulatory follicle is stimulated by luteinizing hormone via prostaglandin E2. *Biol Reprod.* (2015) 92:15. doi: 10.1095/biolreprod.114.123711
- Ocak S, Emsen E, Koycegiz F, Kutluca M, Onder H. Comparison of placental traits and their relation to litter size and parity weight in sheep. *J Anim Sci.* (2009) 87:3196–201. doi: 10.2527/jas.2009-1913
- Mesa H, Safranski TJ, Johnson RK, Lamberson WR. Correlated response in placental efficiency in swine selected for an index of components of litter size 1,2. *J Anim Sci.* (2003) 81:74–9. doi: 10.2527/2003.81174x
- Tuckey RC. Progesterone synthesis by the human placenta. *Placenta.* (2005) 26:273–81. doi: 10.1016/j.placenta.2004.06.012
- Loneragan P. Influence of progesterone on oocyte quality and embryo development in cows. *Theriogenology.* (2011) 76:1594–601. doi: 10.1016/j.theriogenology.2011.06.012
- Winterhager E, Kidder GM. Gap junction connexins in female reproductive organs: implications for women's reproductive health. *Hum Reprod Update.* (2015) 21:340–52. doi: 10.1093/humupd/dmv007
- Clark KL, George JW, Przygrodzka E, Plewes MR, Hua G, Wang C, et al. Hippo signaling in the ovary: emerging roles in development, fertility, and disease. *Endocr Rev.* (2022) 43:1074–96. doi: 10.1210/edrev/bnac013
- Drummond AE. The role of steroids in follicular growth. *Reprod Biol Endocrinol.* (2006) 4:16. doi: 10.1186/1477-7827-4-16
- Clagett-Dame M, Deluca HF. The role of vitamin A in mammalian reproduction and embryonic development. *Annu Rev Nutr.* (2002) 22:347–81. doi: 10.1146/annurev.nutr.22.010402.102745E
- Hogarth CA, Griswold MD. The key role of vitamin A in spermatogenesis. *J Clin Invest.* (2010) 120:956–62. doi: 10.1172/JCI41303
- Xuan R, Zhao X, Li Q, Zhao Y, Wang Y, Du S, et al. Characterization of long noncoding RNA in nonlactating goat mammary glands reveals their regulatory role in mammary cell involution and remodeling. *Int J Biol Macromol.* (2022) 222:2158–75. doi: 10.1016/j.ijbiomac.2022.09.291
- Oh Y, Chung KC. Zinc finger protein 131 inhibits estrogen signaling by suppressing estrogen receptor  $\alpha$  homo-dimerization. *Biochem Biophys Res Commun.* (2013) 430:400–5. doi: 10.1016/j.bbrc.2012.11.031
- Han DX, Xiao Y, Wang CJ, Jiang H, Gao Y, Yuan B, et al. Regulation of FSH expression by differentially expressed miR-186-5p in rat anterior adenohypophyseal cells. *PLoS One.* (2018) 13:e0194300. doi: 10.1371/journal.pone.0194300
- Plum A, Hallas G, Willecke K. Expression of the mouse gap junction gene Gjb3 is regulated by distinct mechanisms in embryonic stem cells and keratinocytes. *Genomics.* (2002) 79:24–30. doi: 10.1006/geno.2001.6671
- Delaforest A, Nagaoka M, Si-Tayeb K, Noto FK, Konopka G, Battle MA, et al. HNF4a is essential for specification of hepatic progenitors from human pluripotent stem cells. *Development.* (2011) 138:4143–53. doi: 10.1242/dev.062547
- Das PP, Shao Z, Beyaz S, Apostolou E, Pinello L, De Los AA, et al. Distinct and combinatorial functions of Jmjd2b/Kdm4b and Jmjd2c/Kdm4c in mouse embryonic stem cell identity. *Mol Cell.* (2014) 53:32–48. doi: 10.1016/j.molcel.2013.11.011
- Murcia CL, Woychik RP. Expression of Pcdh15 in the inner ear, nervous system and various epithelia of the developing embryo. *Mech Dev.* (2001) 105:163–6. doi: 10.1016/s0925-4773(01)00388-4
- Schartl M, Schories S, Wakamatsu Y, Nagao Y, Hashimoto H, Bertin C, et al. Sox5 is involved in germ-cell regulation and sex determination in medaka following co-option of nested transposable elements. *BMC Biol.* (2018) 16:16. doi: 10.1186/s12915-018-0485-8
- Kim H, Wu J, Ye S, Tai CI, Zhou X, Yan H, et al. Modulation of  $\beta$ -catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal. *Nat Commun.* (2013) 4:2403. doi: 10.1038/ncomms3403
- Sell-Kubiak E, Dobrzanski J, Derks M, Lopes MS, Szwaczkowski T. Meta-analysis of SNPs determining litter traits in pigs. *Genes.* (2022) 13:1730. doi: 10.3390/genes13101730
- Sironen A, Uimari P, Nagy S, Paku S, Andersson M, Vilkkij J. Knobbed acrosome defect is associated with a region containing the genes STK17b and HECW2 on porcine chromosome 15. *BMC Genomics.* (2010) 11:699. doi: 10.1186/1471-2164-11-699
- Zhang X, Huang G, Jiang T, Meng L, Li T, Zhang G, et al. CEP112 coordinates translational regulation of essential fertility genes during spermiogenesis through phase separation in humans and mice. *Nat Commun.* (2024) 15:8465. doi: 10.1038/s41467-024-52705-8
- Whitfield M, Thomas L, Bequignon E, Schmitt A, Stouvenel L, Montantin G, et al. Mutations in DNAH17, encoding a sperm-specific axonemal outer dynein arm heavy chain, cause isolated male infertility due to Asthenozoospermia. *Am J Hum Genet.* (2019) 105:198–212. doi: 10.1016/j.ajhg.2019.04.015
- Shen T, Wang H, Tang B, Zhu G, Wang X. The impact of RNA binding proteins and the associated long non-coding RNAs in the TCA cycle on cancer pathogenesis. *RNA Biol.* (2023) 20:223–34. doi: 10.1080/15476286.2023.2216562
- Ozaltin F, Li B, Rauhauser A, An SW, Soylemezoglu O, Gonul II, et al. DGKE variants cause a glomerular microangiopathy that mimics membranoproliferative GN. *J Am Soc Nephrol.* (2013) 24:377–84. doi: 10.1681/ASN.2012090903
- Rasmussen AH, Rasmussen HB, Silahatoglu A. The DLGAP family: neuronal expression, function and role in brain disorders. *Mol Brain.* (2017) 10:43. doi: 10.1186/s13041-017-0324-9
- Cao J, Zhou A, Zhou Z, Liu H, Jia S. The role of GPLD1 in chronic diseases. *J Cell Physiol.* (2023) 238:1407–15. doi: 10.1002/jcp.31041



37. Li Z, Li Y, Wang X, Yang Q. PPP2R2B downregulation is associated with immune evasion and predicts poor clinical outcomes in triple-negative breast cancer. *Cancer Cell Int.* (2021) 21:13. doi: 10.1186/s12935-020-01707-9
38. Usami Y, Wu Y, Gottlinger HG. SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. *Nature.* (2015) 526:218–23. doi: 10.1038/nature15400
39. Silva AA, Silva DA, Silva FF, Costa CN, Silva HT, Lopes PS, et al. GWAS and gene networks for milk-related traits from test-day multiple lactations in Portuguese Holstein cattle. *J Appl Genet.* (2020) 61:465–76. doi: 10.1007/s13353-020-00567-3
40. Liu W, Gou H, Wang X, Li X, Hu X, Su H, et al. TTPAL promotes gastric tumorigenesis by directly targeting NNMT to activate PI3K/AKT signaling. *Oncogene.* (2021) 40:6666–79. doi: 10.1038/s41388-021-01838-x
41. Li H, Zhang W, Liu Y, Cai Z, Lan A, Shu D, et al. UTRN as a potential biomarker in breast cancer: a comprehensive bioinformatics and in vitro study. *Sci Rep.* (2024) 14:7702. doi: 10.1038/s41598-024-58124-5
42. Wu X, Wang CC, Cao Y, Li J, Li Z, Ma H, et al. Novel genetic risk and metabolic signatures of insulin signaling and Androgenesis in the anovulation of polycystic ovary syndrome. *Engineering.* (2023) 23:103–11. doi: 10.1016/j.eng.2022.08.013
43. Tanaka I, Chakraborty A, Saulnier O, Benoit-Pilven C, Vacher S, Labiod D, et al. ZRANB2 and SYF2-mediated splicing programs converging on ECT2 are involved in breast cancer cell resistance to doxorubicin. *Nucleic Acids Res.* (2020) 48:2676–93. doi: 10.1093/nar/gkz1213
44. Newman AP, Graf J, Mancini P, Rossi G, Lian JP, Ferro-Novick S. SEC22 and SLY2 are identical. *Mol Cell Biol.* (1992) 12:3663–4. doi: 10.1128/mcb.12.8.3663-3664.1992
45. Cyr DG, Adam C, Dufresne J, Gregory M. Regulation of the gap junction interplay during postnatal development in the rat epididymis. *Cell Tissue Res.* (2024) 398:191–206. doi: 10.1007/s00441-024-03919-1
46. Paniagua S, Cakir B, Hu Y, Kiral FR, Tanaka Y, Xiang Y, et al. Dyslexia associated gene KIAA0319 regulates cell cycle during human neuroepithelial cell development. *Front Cell Dev Biol.* (2022) 10:967147. doi: 10.3389/fcell.2022.967147
47. Chen X, Hausman BS, Luo G, Zhou G, Murakami S, Rubin J, et al. Protein kinase inhibitor  $\gamma$  reciprocally regulates osteoblast and adipocyte differentiation by downregulating leukemia inhibitory factor. *Stem Cells.* (2013) 31:2789–99. doi: 10.1002/stem.1524
48. Sakamoto K, Furuichi Y, Yamamoto M, Takahashi M, Akimoto Y, Ishikawa T, et al. R3hdm1 regulates satellite cell proliferation and differentiation. *EMBO Rep.* (2019) 20:e47957. doi: 10.15252/embr.201947957