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.

Multilevel genetic and single-cell analyses reveal associations of SELP with high prolificacy and ovarian vascular endothelial states in sheep

Xiaorui Sun
State Key Laboratory of Animal Biotech Breeding, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing

Hao Li
State Key Laboratory of Animal Biotech Breeding, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing

Yanan Zhou
College of Veterinary Medicine, Northwest A&F University

Xiaoru Li
College of Veterinary Medicine, Northwest A&F University

Jun Liu
Hainan Institute of Northwest A&F University

Wentao Ma
Hainan Institute of Northwest A&F University

Zhangyuan Pan
panzhangyuan@caas.cn

State Key Laboratory of Animal Biotech Breeding, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing

Research Article

Keywords: SELP, Sheep, Prolificacy, Ovarian vascular endothelial cells, Single-cell RNA sequencing

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11714573/v1

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

Additional Declarations: No competing interests reported.

Multilevel genetic and single-cell analyses reveal associations of *SELP* with high prolificacy and ovarian vascular endothelial states in sheep

Xiaorui Sun [1,2,3,4,†], Hao Li [1,2,5,†], Yanan Zhou [1,2,3,4,†], Xiaorui Li [1,2,3,4], Jun Liu[3,4,*], Wentao Ma [3,4,*] Zhangyuan Pan [1,2,*]

1 State Key Laboratory of Animal Biotech Breeding, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing 100193, China

2 National Nanfan Research Institute (Sanya), Chinese Academy of Agricultural Sciences, Sanya 572024, China

3 College of Veterinary Medicine, Northwest A&F University, Yangling 712100, Shaanxi, China

4 Hainan Institute of Northwest A&F University, Sanya 572025, Hainan, China

5 Department of Animal Science, College of Agriculture, Yanbian University, Yanji 133002, China

Correspondence:

Corresponding authors: Jun Liu; Wentao Ma; Zhangyuan Pan.

E-mail addresses: liujun2013@nwsuaf.edu.cn (J.L.); mawentao@nwafu.edu.cn (W.M.); panzhangyuan@caas.cn (Z.P.).

†These authors contributed equally to this work and share first authorship.

Abstract

Background: Prolificacy in sheep (*Ovis aries*) is an important trait shaped by genetic background and the ovarian microenvironment. However, the vascular molecular processes probably involved in high prolificacy remain poorly understood. P-selectin is an important adhesion molecule that mediates interactions between vascular endothelial cells and circulating immune cells, but its association with sheep prolificacy and ovarian function remains unclear. The study aims to elucidate the relationship between *SELP* and high prolificacy in sheep.

Results: We used RNA-seq to compare ovarian *SELP* expression between high- and low-prolificacy sheep and found higher expression in the high-prolificacy group. Whole-genome sequencing and

genotype–phenotype analyses identified eight intronic variants associated with average litter size, with chr12:36538237 retained as a representative candidate. Protein sequence and structural analyses showed that SELP was conserved among ruminants and placed the missense variant chr12:36518637 T>G (p.Ser758Arg) near the predicted SELP–CD44 interaction interface. Using ovarian single-cell RNA sequencing, we found that *SELP* expression was mainly concentrated in blood vascular endothelial cells. Virtual knockout analysis further predicted perturbations in gene networks related to cell migration, leukocyte migration, cytokine and chemokine signaling, and the MAPK cascade. **Conclusions:** These findings support an association between *SELP* and sheep prolificacy and suggest that *SELP* may participate in the ovarian vascular–immune microenvironment.

Keywords: *SELP*, Sheep, Prolificacy, Ovarian vascular endothelial cells, Single-cell RNA sequencing

Background

With the development of the breeding industry, the supply of mutton is becoming increasingly tight. Reproductive efficiency has become a top priority. Fecundity is a trait that directly affects the reproductive efficiency of sheep, and recent studies support a complex genetic basis for variation in litter size[1, 2]. Previous studies have identified major reproductive genes, including *BMPR1B, BMP15, and GDF9,* whose genetic variants are associated with ovulation rate and litter size in sheep[3, 4]. Genome-wide association studies have further implicated multiple genomic regions in the genetic regulation of litter size in sheep [2, 5]. Recent sheep ovarian single-cell transcriptomic studies have resolved distinct ovarian cell populations and reproductive-state-associated transcriptional programs [6, 7]. Integrating population genetic, gene expression, and single-cell evidence may therefore provide a more comprehensive framework for evaluating the potential relationship between *SELP* and reproductive phenotypes.

Follicular growth and corpus luteum formation involve dynamic angiogenesis and tissue remodeling, while VEGF- and endocrine-related signaling contributes to ovarian endothelial function and

follicular vascularization [8, 9]. The periovulatory period is characterized by vascular remodeling, altered permeability, immune-cell recruitment, and tissue remodeling, giving the ovulatory microenvironment several inflammation-like features [10, 11]. In cattle (*Bos taurus*), P-selectin can mediate leukocyte adhesion, providing direct experimental evidence for P-selectin-dependent vascular–immune interactions in a ruminant species [12]. These observations do not, however, directly establish the function of *SELP* in the sheep ovary or its effects on litter size.

SELP (selectin P, P-selectin) encodes a cell adhesion molecule. P-selectin is stored in platelet α-granules and endothelial Weibel–Palade bodies and can translocate to the cell surface following stimulation [13, 14]. At the vascular surface, P-selectin participates in the initial capture and rolling of leukocytes and thereby contributes to leukocyte recruitment [15]. Fang C et al. performed preliminary transcriptomic and population genetic screening and found that *SELP* may be associated with sheep prolificacy [16, 17]. Accordingly, we hypothesize that the association between *SELP* and high prolificacy may have a cellular basis, involving ovarian vascular endothelial cells and the vascular–immune microenvironment. However, few studies have focused on the function of *SELP* in sheep. In this study, we explored the expression profile of *SELP* using ovarian single-cell transcriptome data. We also applied virtual perturbation analysis to investigate potentially linked molecular processes. Our results provide timely and essential data for clarifying the relationship between *SELP* and high prolificacy in sheep.

Results

SELP expression differs between high- and low-prolificacy sheep and among tissues

SELP expression was significantly higher in the six high-prolificacy ovarian samples than in the three low-prolificacy ovarian samples ($P<0.05$) (Fig. 1A). Exon-level expression profiles across the *SELP* transcript were also higher in the high-prolificacy group than in the low-prolificacy group (Fig. 1B). Genome-browser tracks showed stronger RNA-seq signals across the *SELP* region in the high-prolificacy ovarian samples, whereas lower signals were observed in the low-prolificacy samples and

in several other tissues (Fig. 1C). *SELP* expression differed among sheep tissues (Fig. 1D). Cross-species profiling detected *SELP* expression in multiple tissues from human, pig, sheep, and cattle, with differences in expression level and expression proportion among tissues and species (Fig. 1E).

DNA differences, linkage relationships and haplotype changes around the *SELP* gene

Population genetic analysis across the broader *SELP* interval identified 878 variant sites, of which 46 had $F_{ST} > 0.15$ (Table S2). Among the eight intronic variants evaluated in the genotype-phenotype analysis, F_{ST} values were 0.224615 for chr12:36536704; 0.208803 for chr12:36537867, chr12:36537958, and chr12:36538109; and 0.233333 for chr12:36538237, chr12:36538328, chr12:36538409, and chr12:36539559 (Fig. 2A; Table S2 SELP differentiated variants identified by population genetic analysis). The separate coding variant chr12:36518637 T>G had an F_{ST} value of 0.0359431 (Table S2). LD analysis of 370 SNPs in 42 matched individuals showed heterogeneous pairwise R^2 values and several local LD blocks across the *SELP* region (Fig. 2B). Haplotype analysis of 233 sheep using 14 SNPs spanning chr12:36546317-36549144 identified 14 haplotypes (H1-H14), with different haplotype compositions and frequencies across the analyzed populations (Fig. 2C, D; Table S3).

***SELP* candidate variants are associated with average litter size**

Allele-count comparisons at the eight displayed *SELP*-region sites showed different reference- and alternative-allele counts between the high- and low-prolificacy groups (Fig. 3A). In the genotype-phenotype analysis of 543 sheep, all eight intronic *SELP* candidate variants were associated with average litter size, with P values ranging from 0.00128557 to 0.00272971 and Benjamini-Hochberg-adjusted FDR values ranging from 0.00253213 to 0.00272971 (Table S1). At chr12:36538237, the wild-type homozygous, mutant heterozygous, and mutant homozygous genotypes were observed in 259, 232, and 52 sheep, respectively, with mean litter sizes of 2.329, 2.344, and 2.702. The overall P value was 0.001 and FDR = 0.002. Pairwise comparisons P values were 0.001 for mutant homozygous versus wild-type homozygous, 0.002 for mutant homozygous versus mutant

heterozygous, and 0.969 for mutant heterozygous versus wild-type homozygous (Fig. 3B). At chr12:36536704, the wild-type homozygous, mutant heterozygous, and mutant homozygous were observed in 234, 237, and 72 sheep, with mean litter sizes of 2.339, 2.319, and 2.650, respectively; the overall P value was 0.001 and the FDR was 0.002. The pairwise comparisons shown in Fig. 3B were different for mutant homozygous versus wild-type homozygous and mutant homozygous versus mutant heterozygous, whereas wild-type homozygous versus mutant heterozygous was not different.

SELP shows conserved protein features across vertebrate species

Phylogenetic analysis placed sheep SELP with goat (*Capra hircus*), takin (*Budorcas taxicolor*), zebu cattle (*Bos indicus*), and cattle within the bovid group (Fig. 4A). The motif profiles showed similar motif composition and ordering among these bovid species, while greater variation was present in more distantly related species (Fig. 4A). Multiple-sequence alignment showed a series of conserved amino acid positions among sheep and the other bovids included in the analysis (Fig. 4B). TMHMM analysis produced a membrane-associated topology profile containing predicted extracellular, transmembrane, and cytoplasmic regions (Fig. 4C). In the subcellular localization profile, the highest confidence values were assigned to extracellular, extracellular or plasma-membrane, and plasma-membrane categories (Fig. 4D). SignalP analysis returned an N-terminal Sec/SPI signal-peptide profile for sheep SELP (Fig. 4E).

Ser758 is located near the predicted SELP-CD44 interface

The predicted SELP-CD44 complex contained several local contact regions between the two proteins (Fig. 5A, B). The residue corresponding to the missense variant chr12:36518637 T>G (p.Ser758Arg) was located near one of the predicted interface regions. In this region, SELP Ser758 was positioned near SELP Phe760 and Thr761 and CD44 Ser122 and Asn57; representative interatomic distances shown in the structure were approximately 1.9 Å and 2.8 Å (Fig. 5C). A second local interface region showed SELP Lys49 in spatial proximity to CD44 Ile172 and Ser174, with representative distances of approximately 2.7 Å and 2.9 Å (Fig. 5D).

Single-cell transcriptomic and virtual knockout analyses of *SELP* in ovarian blood vascular endothelial cells

Uniform Manifold Approximation and Projection (UMAP) visualization resolved the annotated ovarian single-cell dataset into multiple cell populations, including blood vascular endothelial cells, lymphatic endothelial cells, granulosa cells, macrophages, NK T cells, T cells, neutrophils, oocytes, smooth muscle cells, stromal cells, epithelial cells, and glial cells (Fig. 6A). Projection of *SELP* expression onto the UMAP showed that the strongest *SELP* signal was concentrated in blood vascular endothelial cells, with lower or sparse expression in most other annotated populations (Fig. 6B). Blood vascular endothelial cells were divided into 6,645 *SELP*-positive and 6,563 *SELP*-negative cells. *SELP* showed an average $\log_2$ fold change of 15.619 in the *SELP*-positive versus *SELP*-negative comparison, with an expression proportion of 1.000 in *SELP*-positive cells and close to 0 in *SELP*-negative cells. The heatmap showed differences between the two groups for representative genes including *CA4, CXCL9, EMCN, KDR, PECAM1, TNFAIP3, SELP, F5, IL6, BIRC3, ICAM1, NFKBIA, CEBPD, MAP3K8, VWF*, and *PDE4B* (Fig. 6C). The *SELP*-centered network included *NFKBIA, CEBPD, ICAM1, F5, CXCL9, IL6, TFPI2, INPP4B, CA4, SLC4A7, GADD45A, IER3, ADGRG6,* and other genes, with both positive and negative inferred network weights (Fig. 6D). The *SELP* virtual-knockout analysis produced 3,000 different Regulation candidate response genes, of which 77 met the FDR < 0.05 threshold. Marked responses included *SELP*, *LOC101117955, LOC101113728, LOC101117184, BUB1, INPP4B, NFKBIA, CEBPD, TFPI2*, and *SLC4A7.* The perturbation manifold showed displacement between WT and virtual-KO states for representative genes, including *CEBPD, NFKBIA, GADD45A, F5, CA4, CLDN11, IL6, IER3, TFPI2, ICAM1, SLC4A7, CXCL9, BIRC3, ADGRG6*, and *INPP4B* (Fig. 6E). Gene Ontology enrichment of the response genes included positive regulation of MAPK cascade, cell migration, neutrophil chemotaxis, myeloid leukocyte migration, leukocyte migration, inflammatory response,

response to cytokine, cytokine-mediated signaling pathway, and chemokine-mediated signaling pathway (Fig. 6F).

Discussion

SELP encodes P-selectin, a cell adhesion molecule mainly associated with platelets and vascular endothelial cells. P-selectin can move to the cell surface after stimulation and take part in the capture and rolling of leukocytes [13–15]. The ovary also undergoes marked vascular and immune changes during follicular growth, ovulation, and corpus luteum formation [8–11]. These known features make *SELP* a reasonable gene to examine in the ovarian vascular environment. In our study, *SELP* expression was higher in the high-prolificacy group than in the low-prolificacy group. *SELP* was also expressed in several sheep tissues and showed different expression levels across tissues and species. We therefore speculate tha*t SELP* may be associated with litter size. This finding provides a new clue to explore the molecular mechanism of high fecundity in sheep.

Litter size is a complex trait and is affected by many genes and biological processes [16–22]. In the *SELP* region, we detected population differentiation, local linkage disequilibrium, and different haplotype frequencies among sheep populations. Eight intronic variants were associated with average litter size in the 543 sheep used for genotype-phenotype analysis. Among these variants, chr12:36538237 was retained as a representative candidate. Sheep carrying the mutant homozygous genotype at this site had a higher mean litter size than sheep with the wild-type homozygous or mutant heterozygous genotype. Similar genotype differences were also observed at chr12:36536704. These findings show that genetic variation in the *SELP* region is related to reproductive phenotype in the population analyzed here. These variant sites may serve as candidate molecular markers for sheep breeding.

Most of the variants associated with litter size in our analysis were located in introns. Noncoding variants can affect gene expression through regulatory elements, transcription factor binding, chromatin state, RNA processing, or linkage with nearby functional variants [23–27]. Therefore,

chr12:36538237 may be a regulatory candidate or an association marker. This site lays a foundation for further functional verification of noncoding mutation in sheep.

Protein analysis showed that SELP was relatively conserved among ruminants and had the expected features of a membrane-associated protein. We also examined a separate coding variant, chr12:36518637 T>G, which causes p.Ser758Arg. Structural modeling placed Ser758 near a predicted SELP-CD44 interface. This result only provides structural context for the missense variant. It does not prove that p.Ser758Arg changes SELP-CD44 binding or SELP function. Importantly, chr12:36518637 and chr12:36538237 represent different types of evidence. The first is a coding variant used for structural analysis, whereas the second is an intronic variant associated with average litter size. The structural prediction offers a reference for exploring protein interaction changes caused by this amino acid mutation.

The single-cell results provided a clearer view of where *SELP* is expressed in the ovary. *SELP* signal was mainly concentrated in blood vascular endothelial cells. *SELP*-positive endothelial cells also showed different expression patterns from *SELP*-negative cells. Genes such as *NFKBIA, CEBPD, ICAM1, F5, CXCL9, IL6,* and *TFPI2* were among the representative genes that differed between the two groups. Previous studies show that ovarian vascular and immune cells interact during reproductive processes [28–32]. P-selectin is also well known for its role in leukocyte adhesion to vascular endothelium [33–35]. Therefore, the endothelial enrichment of *SELP* and the expression pattern of *SELP*-positive cells are consistent with a possible role in the ovarian vascular-immune environment. This result identifies ovarian vascular endothelial cells as the primary cell type for *SELP* function in ovary.

The virtual knockout analysis gave additional information about the *SELP*-related network. Seventy-seven genes were significantly perturbed at FDR < 0.05. Enriched biological processes included cell migration, leukocyte migration, neutrophil chemotaxis, inflammatory response, cytokine signaling, chemokine signaling, and positive regulation of the MAPK cascade. These results are in line with the

expression differences observed between *SELP*-positive and *SELP*-negative endothelial cells, and have important application value.

Conclusions

In this study, we found that *SELP* expression was significantly higher in high-prolificacy sheep than in low-prolificacy sheep ($P<0.05$). We identified eight intronic variants including chr12:36538237 that were associated with average litter size. Single-cell analysis showed that *SELP* was mainly expressed in ovarian blood vascular endothelial cells. Protein structure prediction and virtual knockout analysis further supported the potential function of *SELP* in ovarian vascular-immune regulation. We propose *SELP* as a candidate gene related to high prolificacy in sheep. These findings provide a new perspective for exploring the molecular mechanism of sheep prolificacy from the ovarian vascular microenvironment.

Methods

Animal and sample collection

This study used previously collected RNA-seq data of Small-tailed Han sheep from our laboratory. A total of nine ovarian RNA-seq samples were included for expression comparison, including six high-prolificacy sheep and three low-prolificacy sheep. Whole-genome sequencing (WGS) data for genotype-phenotype association analysis were obtained from 543 Small-tailed Han sheep with individual litter size records in our laboratory. Population differentiation analysis was performed using 233 sheep, including 101 individuals in the high-prolificacy group and 132 individuals in the low-prolificacy group.

Ovarian RNA-seq processing and *SELP* expression analysis

Total RNA from ovarian tissue had been extracted using TRIzol reagent, and RNA quantity, purity, and integrity were evaluated before library construction. Only RNA samples that passed the quality-control criteria of the original dataset were used for sequencing and downstream analysis. Poly(A)-enriched RNA libraries were sequenced in paired-end mode on the BGISEQ-500 platform [36],

following the library-preparation procedure used for the source dataset. Raw reads were processed to remove adaptor sequences and low-quality bases, and clean reads were aligned to the sheep reference genome ARS-UI_Ramb_v2.0 (GCF_016772045.1) [37]. Gene-level abundance values were obtained from the processed expression matrix, and normalized *SELP* expression was extracted for subsequent analyses. *SELP* expression between the high- and low-prolificacy groups was compared in R v4.5.3, with $P < 0.05$ considered statistically significant. Expression distributions were visualized using ggplot2, and genome-browser tracks were examined to compare transcriptional signals across the *SELP* region. These bulk RNA-seq analyses were performed at the tissue level and were kept analytically separate from the ovarian single-cell transcriptomic analyses described below.

Cross-tissue and cross-species expression profiling

Public sheep transcriptomic datasets were integrated to compare *SELP* expression among ovary, blood, spleen, lung, liver, skin, and other tissues. Tissue expression profiles from sheep, cattle, pig (*Sus scrofa*), and human (*Homo sapiens*) were further used to evaluate the cross-species distribution of *SELP* expression [38–40]. Differences in *SELP* expression among tissues were evaluated by one-way analysis of variance, with $P < 0.05$ considered statistically significant. Genome-browser visualization was used to display *SELP* gene structure, GC content, and transcriptional signals across representative tissues.

Population genetic differentiation at the *SELP* region

Genomic coordinates were defined according to the sheep reference genome ARS-UI_Ramb_v2.0 (GCF_016772045.1), in which *SELP* spans chr12:36517739–36552490. The *SELP* gene together with 10-kb flanking regions on both sides (chr12:36507739–36562490) was used as the primary interval for focused candidate-variant screening. SNPs were extracted from a whole-genome variant call format (VCF) generated from the laboratory WGS dataset after the quality-control procedure applied in the original genomic analysis. Site-level fixation index (F_{ST}) values between the 101 high-prolificacy and 132 low-prolificacy sheep were calculated using VCFtools v0.1.16. Variants with F_{ST}

> 0.15 were prioritized as highly differentiated candidate sites, while additional representative regional variants were retained for comparison. Variants were annotated according to their genomic positions and classified as upstream, downstream, intronic, exonic, or other variants. The prioritized variant chr12:36538237 was annotated as intronic, whereas chr12:36518637 T>G represented a separate coding missense candidate used for structural analysis. To characterize the broader differentiation landscape around *SELP*, F_{ST} results were additionally examined across a region extending approximately 50 kb upstream and downstream of the gene. The variant-containing portion of this broader interval extended from chr12:36467743 to chr12:36602413 and contained 878 variant sites, of which 46 had $F_{ST} > 0.15$.

Linkage disequilibrium and haplotype analysis

Local linkage disequilibrium analysis was performed using the *SELP*-region matched dataset retained after sample matching and quality control. This dataset contained 42 individuals and 370 SNPs, and pairwise linkage disequilibrium was quantified as R^2. The resulting pairwise R^2 matrix was used to visualize local LD patterns surrounding the prioritized *SELP* variants. This matched LD dataset was analytically distinct from the 233-sheep population used for F_{ST} estimation. Haplotype analysis was conducted separately using genotype data from 233 sheep. Fourteen SNPs recorded in the finalized *SELP* candidate-block file, spanning chr12:36546317–36549144, were used to characterize the local haplotype segment. Haplotype composition and frequencies were summarized across the analyzed sheep populations.

Genotype–phenotype association analysis

Genotype–phenotype association analyses were performed in 543 sheep with usable genotype data and longitudinal litter-size records. For each ewe, average litter size was calculated as the arithmetic mean of litter size across all available valid parities. Eight prioritized intronic *SELP* variants were evaluated by comparing average litter size among genotype groups. The overall genotype effect at each variant was evaluated using one-way ANOVA, with N = 543 for each successfully analyzed

site. Benjamini–Hochberg correction was applied across the eight omnibus tests to control the false discovery rate. For prioritized regions, Tukey's honestly significant difference (HSD) test was used for pairwise comparisons among genotype groups, and these adjusted pairwise P values were distinguished from the omnibus ANOVA P values and across-site FDR values. All eight variants were retained in the reporting, with chr12:36538237 treated as one representative prioritized candidate rather than as the only associated region.

Evolutionary and structural analysis of SELP

SELP protein sequences from sheep and representative vertebrate species were retrieved from the NCBI Protein database [41]. Multiple sequence alignment and phylogenetic analysis were performed using MEGA11 [42], and a Neighbor-Joining tree was constructed with 1,000 bootstrap replicates to assess branch support. Conserved protein motifs were identified using MEME Suite and compared across species to assess conservation of motif composition and arrangement. Transmembrane topology of sheep SELP was evaluated using TMHMM 2.0, and the N-terminal signal peptide was predicted using SignalP 6.0 [43]. Subcellular localization was described on the basis of established P-selectin biology and was not treated as an independent computational prediction. The three-dimensional SELP–CD44 complex was predicted using AlphaFold 3 [44], and structural visualization and interface inspection were performed in PyMOL v3.1 [45]. Residue pairs with minimum interatomic distances of ≤4.0 Å were defined as close contacts for interface characterization. The coding candidate chr12:36518637 T>G (p.Ser758Arg) was mapped onto the SELP structure to examine its local spatial context relative to the predicted SELP–CD44 interface.

Ovarian single-cell transcriptomics and *SELP v*irtual knockout

Single-cell analyses were performed using a previously annotated sheep ovary Seurat object in R v4.5.3, with the existing cell_type metadata retained for cell-type identities and the RNA assay used for expression analysis [46, 47]. UMAP visualization was based on the existing low-dimensional representation and was used to display major ovarian cell populations and the distribution of *SELP*

expression. Particular attention was given to Blood vascular endothelial cells, Lymphatic endothelial cells, Macrophages, NK T cells, and Granulosa cells. Blood vascular endothelial cells were classified as *SELP*-positive when *SELP* expression was detected and as *SELP*-negative when no *SELP* expression was detected. The final comparison included 6,645 *SELP*-positive and 6,563 *SELP*-negative endothelial cells. Differential expression between the two groups was evaluated using Seurat FindMarkers, and representative genes were prioritized according to adjusted P values, average $\log_2$ fold change, and expression proportions.

***SELP* virtual knockout and functional enrichment**

SELP virtual knockout was performed in Blood vascular endothelial cells using scTenifoldKnk v1.0.3 [48]. Genes with FDR < 0.05 in the scTenifoldKnk output were considered perturbed. Gene Ontology Biological Process enrichment analysis of perturbed genes was performed using clusterProfiler [49], with Benjamini–Hochberg correction for multiple testing.

Abbreviations

ANOVA, analysis of variance; FDR, false discovery rate; FST, fixation index; LD, linkage disequilibrium; MAPK, mitogen-activated protein kinase; NCBI, National Center for Biotechnology Information; RNA-seq, RNA sequencing; SNP, single-nucleotide polymorphism; UMAP, Uniform Manifold Approximation and Projection; VCF, Variant Call Format; WGS, whole-genome sequencing; R^2, squared correlation coefficient.

Declarations

Ethics approval and consent to participate

All animal procedures were reviewed and approved by the Animal Welfare and Ethics Committee of the Institute of Animal Science, Chinese Academy of Agricultural Sciences (approval nos. IAS2024-129 and IAS2024-130). The sheep were institution-owned research animals maintained by the Institute of Animal Science, Chinese Academy of Agricultural Sciences and were used exclusively

for research purposes. The animals were not privately or client owned. Their use in this study was authorized in writing by all responsible institutional representatives.

Availability of data and materials

The RNA-seq data, whole-genome sequencing (WGS) data, and genotype–phenotype data from 543 Small-tailed Han sheep used in this study were obtained from existing datasets generated and maintained by our laboratory. Public multi-tissue expression data for sheep were obtained from NCBI and EBI, while data for pig, cattle, and human were obtained from PigGTEx, CattleGTEx, and the GTEx Portal, respectively.

Competing interests

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.

Funding

This work was supported by the National Key R&D Program of China (grant numbers 2023YFF1001800 and 2022YFF1000100) and the Xinjiang Uygur Autonomous Region Major Science and Technology Special Project, ‘Breeding and Multiplication of New Varieties/Lines of Wide-adaptability Meat Sheep in Xinjiang’ (grant number 2024A02004).

Authors' contributions

XS, HL, YZ, ZP, and WM: study concept, design, and data interpretation. XS, XL, and YZ: data acquisition and analysis. XS: experiment and manuscript writing. XS, HL, ZP, JL, and WM: preparation and manuscript editing, ZP, JL, and WM: resources. All authors contributed to the article and approved the submitted version.

References

1. Zhong T, Hou D, Zhao Q, Zhan S, Wang L, Li L, et al. Comparative whole-genome resequencing to uncover selection signatures linked to litter size in Hu Sheep and five other breeds. BMC genomics. 2024;25:480. https://doi.org/10.1186/s12864-024-10396-x.

2. Bao J, Xiong J, Huang J, Yang P, Shang M, Zhang L. Genetic Diversity, Selection Signatures, and Genome-Wide Association Study Identify Candidate Genes Related to Litter Size in Hu Sheep. International journal of molecular sciences. 2024;25:9397. https://doi.org/10.3390/ijms25179397.

3. Zhang Y, Wang H, Li T, Zhang N, Chen J, Yang H, et al. Association of BMP15 and GDF9 Gene Polymorphisms with Litter Size in Hu Sheep. Genes. 2025;16:168. https://doi.org/10.3390/genes16020168.

4. Ji X, Cao Z, Hao Q, He M, Cang M, Yu H, et al. Effects of New Mutations in BMPRIB, GDF9, BMP15, LEPR, and B4GALNT2 Genes on Litter Size in Sheep. Veterinary sciences. 2023;10:258. https://doi.org/10.3390/vetsci10040258.

5. Gholizadeh M, Esmaeili-Fard SM. Meta-analysis of genome-wide association studies for litter size in sheep. Theriogenology. 2022;180:103–12. https://doi.org/10.1016/j.theriogenology.2021.12.025.

6. Zhao B, Dai K, Li S, Zhang G, Lv W, Liu W, et al. Integrative single-cell transcriptomics of sheep ovarian development reveals dynamic transcriptional programs relevant to reproductive traits. iScience. 2025;28:112422. https://doi.org/10.1016/j.isci.2025.112422.

7. Ge T, Wen Y, Li B, Huang X, Jiang S, Zhang E. Single-cell sequencing reveals the reproductive variations between primiparous and multiparous Hu ewes. Journal of animal science and biotechnology. 2023;14:144. https://doi.org/10.1186/s40104-023-00941-1.

8. Lund M, Pearson AC, Sage MAG, Duffy DM. Luteinizing hormone receptor promotes angiogenesis in ovarian endothelial cells of Macaca fascicularis and Homo sapiens†. Biology of reproduction. 2023;108:258–68. https://doi.org/10.1093/biolre/ioac189.

9. Guzmán A, Hernández-Coronado CG, Gutiérrez CG, Rosales-Torres AM. The vascular endothelial growth factor (VEGF) system as a key regulator of ovarian follicle angiogenesis and growth. Molecular reproduction and development. 2023;90:201–17. https://doi.org/10.1002/mrd.23683.

10. Bazzano MV, Köninger A, Solano ME. Beyond defence: Immune architects of ovarian health and disease. Seminars in immunopathology. 2024;46:11. https://doi.org/10.1007/s00281-024-01021-w.

11. Duffy DM, Ko C, Jo M, Brannstrom M, Curry TE. Ovulation: Parallels With Inflammatory Processes. Endocrine reviews. 2019;40:369–416. https://doi.org/10.1210/er.2018-00075.

12. Chen X, Cheng Z, Werling D, Pollott GE, Salavati M, Johnson KF, et al. Bovine P-selectin mediates leukocyte adhesion and is highly polymorphic in dairy breeds. Research in veterinary science. 2016;108:85–92. https://doi.org/10.1016/j.rvsc.2016.08.004.

13. Hellen N, Mashanov GI, Conte IL, le Trionnaire S, Babich V, Knipe L, et al. P-selectin mobility undergoes a sol-gel transition as it diffuses from exocytosis sites into the cell membrane. Nature communications. 2022;13:3031. https://doi.org/10.1038/s41467-022-30669-x.

14. Naß J, Terglane J, Gerke V. Weibel Palade Bodies: Unique Secretory Organelles of Endothelial Cells that Control Blood Vessel Homeostasis. Frontiers in cell and developmental biology. 2021;9:813995. https://doi.org/10.3389/fcell.2021.813995.

15. Escopy S, Chaikof EL. Targeting the P-selectin/PSGL-1 pathway: discovery of disease-modifying therapeutics for disorders of thromboinflammation. Blood vessels, thrombosis & hemostasis. 2024;1:100015. https://doi.org/10.1016/j.bvth.2024.100015.

16. Fang C, Druet T, Cao H, Liu W, Chen Q, Farnir F. Whole genome sequences of 297 Duolang sheep for litter size. Scientific data. 2025;12:1086. https://doi.org/10.1038/s41597-025-05448-0.

17. Li H, He Y, Du X, Zhang S, Nie S, Zhang X, et al. Multi-level transcriptomics reveals FecB-associated and non-FecB ovarian regulatory programs underlying sheep prolificacy. Journal of Integrative Agriculture. 2026;:S2095311926002248. https://doi.org/10.1016/j.jia.2026.05.041.

18. Liu B, Liu S, Li P, Tulafu H, Ma Y, Huang J, et al. Whole-genome resequencing uncovers population structure and candidate molecular markers for litter size in hetian sheep. BMC genomics. 2025;26:1005. https://doi.org/10.1186/s12864-025-12168-7.

19. Tian D, Liu S, Tian F, Ding N, Li X, Zhao K. Comparative transcriptome of reproductive axis in Chinese indigenous sheep with different FecB genotypes and prolificacies. Animal reproduction science. 2020;223:106624. https://doi.org/10.1016/j.anireprosci.2020.106624.

20. Yao X, Li F, Wei Z, Ei-Samahy MA, Feng X, Yang F, et al. Integrative Genome-Wide DNA Methylome and Transcriptome Analysis of Ovaries from Hu Sheep with High and Low Prolific. Frontiers in cell and developmental biology. 2022;10:820558. https://doi.org/10.3389/fcell.2022.820558.

21. Ren Z, He X, Wang X, Chu M. Polymorphisms of the HRG, FETUB, and GUCY1A1 genes and their association with litter size in sheep. Archives animal breeding. 2024;67:153–61. https://doi.org/10.5194/aab-67-153-2024.

22. Cui W, Wang H, Li J, Lv D, Xu J, Liu M, et al. Sheep litter size heredity basis using genome-wide selective analysis. Reproduction in domestic animals = Zuchthygiene. 2024;59:e14689. https://doi.org/10.1111/rda.14689.

23. Kim-Hellmuth S, Aguet F, Oliva M, Muñoz-Aguirre M, Kasela S, Wucher V, et al. Cell type-specific genetic regulation of gene expression across human tissues. Science (New York, NY). 2020;369:eaaz8528. https://doi.org/10.1126/science.aaz8528.

24. ENCODE PC, Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature. 2020;583:699–710. https://doi.org/10.1038/s41586-020-2493-4.

25. Stolze LK, Conklin AC, Whalen MB, López Rodríguez M, Õunap K, Selvarajan I, et al. Systems Genetics in Human Endothelial Cells Identifies Non-coding Variants Modifying Enhancers, Expression, and Complex Disease Traits. American journal of human genetics. 2020;106:748–63. https://doi.org/10.1016/j.ajhg.2020.04.008.

26. Alsheikh AJ, Wollenhaupt S, King EA, Reeb J, Ghosh S, Stolzenburg LR, et al. The landscape of GWAS validation; systematic review identifying 309 validated non-coding variants across 130 human diseases. BMC medical genomics. 2022;15:74. https://doi.org/10.1186/s12920-022-01216-w.

27. French JD, Edwards SL. The Role of Noncoding Variants in Heritable Disease. Trends in genetics : TIG. 2020;36:880–91. https://doi.org/10.1016/j.tig.2020.07.004.

28. Jones ASK, Hannum DF, Machlin JH, Tan A, Ma Q, Ulrich ND, et al. Cellular atlas of the human ovary using morphologically guided spatial transcriptomics and single-cell sequencing. Science advances. 2024;10:eadm7506. https://doi.org/10.1126/sciadv.adm7506.

29. Zhao B, Luo H, Guan D, Fu X, Wang F, Zhang G, et al. Comparative Single-Cell Transcriptomic Atlas Reveals the Genetic Regulation of Reproductive Traits. Advanced science (Weinheim, Baden-Wurttemberg, Germany). 2026;13:e17633. https://doi.org/10.1002/advs.202517633.

30. Pan B, Qin J, Du K, Zhang L, Jia G, Ye J, et al. Integrated ultrasensitive metabolomics and single-cell transcriptomics identify crucial regulators of sheep oocyte maturation and early embryo development in vitro. Journal of advanced research. 2025;73:147–60. https://doi.org/10.1016/j.jare.2024.08.040.

31. Guzmán A, Hughes CHK, Murphy BD. Liver receptor homolog 1 (LRH-1) regulates follicle vasculature during ovulation in mice. Reproduction (Cambridge, England). 2022;164:283–9. https://doi.org/10.1530/REP-22-0135.

32. Zhang C, Yan C, Wei W, Dong T, Jitjumnong J, Yang H, et al. Single-Cell RNA Sequencing Reveals the Role of Macrophage-Granulosa Cell Crosstalk in Regulating Granulosa Cells Function in Sheep. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2026;40:e71648. https://doi.org/10.1096/fj.202502486RR.

33. Beckman JD, DaSilva A, Aronovich E, Nguyen A, Nguyen J, Hargis G, et al. JAK-STAT inhibition reduces endothelial prothrombotic activation and leukocyte-endothelial proadhesive interactions. Journal of thrombosis and haemostasis : JTH. 2023;21:1366–80. https://doi.org/10.1016/j.jtha.2023.01.027.

34. O'Dwyer M, Kirkham-McCarthy L, Cerreto M, Foà R, Natoni A. PSGL-1 decorated with sialyl Lewis(a/x) promotes high affinity binding of myeloma cells to P-selectin but is dispensable for E-selectin engagement. Scientific reports. 2024;14:1756. https://doi.org/10.1038/s41598-024-52212-2.

35. Mylvaganam S, Riedl M, Vega A, Collins RF, Jaqaman K, Grinstein S, et al. Stabilization of Endothelial Receptor Arrays by a Polarized Spectrin Cytoskeleton Facilitates Rolling and Adhesion of Leukocytes. Cell reports. 2020;31:107798. https://doi.org/10.1016/j.celrep.2020.107798.

36. Zhu F-Y, Chen M-X, Ye N-H, Qiao W-M, Gao B, Law W-K, et al. Comparative performance of the BGISEQ-500 and Illumina HiSeq4000 sequencing platforms for transcriptome analysis in plants. Plant methods. 2018;14:69. https://doi.org/10.1186/s13007-018-0337-0.

37. Davenport KM, Bickhart DM, Worley K, Murali SC, Salavati M, Clark EL, et al. An improved ovine reference genome assembly to facilitate in-depth functional annotation of the sheep genome. GigaScience. 2022;11:giab096. https://doi.org/10.1093/gigascience/giab096.

38. Teng J, Gao Y, Yin H, Bai Z, Liu S, Zeng H, et al. A compendium of genetic regulatory effects across pig tissues. Nature genetics. 2024;56:112–23. https://doi.org/10.1038/s41588-023-01585-7.

460 39. Liu S, Gao Y, Canela-Xandri O, Wang S, Yu Y, Cai W, et al. A multi-tissue atlas of regulatory
461 variants in cattle. Nature genetics. 2022;54:1438–47. https://doi.org/10.1038/s41588-022-01153-5.

462 40. GTEx C. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science
463 (New York, NY). 2020;369:1318–30. https://doi.org/10.1126/science.aaz1776.

464 41. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence
465 (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic
466 acids research. 2016;44:D733–45. https://doi.org/10.1093/nar/gkv1189.

467 42. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version
468 11. Molecular biology and evolution. 2021;38:3022–7. https://doi.org/10.1093/molbev/msab120.

469 43. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, et al.
470 SignalP 6.0 predicts all five types of signal peptides using protein language models. Nature
471 biotechnology. 2022;40:1023–5. https://doi.org/10.1038/s41587-021-01156-3.

472 44. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction
473 of biomolecular interactions with AlphaFold 3. Nature. 2024;630:493–500.
474 https://doi.org/10.1038/s41586-024-07487-w.

475 45. Yuan S, Chan HCS, Hu Z. Using PyMOL as a platform for computational drug design. WIREs
476 Comput Mol Sci. 2017;7:e1298. https://doi.org/10.1002/wcms.1298.

477 46. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WMR, et al. Comprehensive
478 Integration of Single-Cell Data. Cell. 2019;177:1888-1902.e21.
479 https://doi.org/10.1016/j.cell.2019.05.031.

480 47. Hao Y, Hao S, Andersen-Nissen E, Mauck WMR, Zheng S, Butler A, et al. Integrated analysis of
481 multimodal single-cell data. Cell. 2021;184:3573-3587.e29.
482 https://doi.org/10.1016/j.cell.2021.04.048.

483 48. Osorio D, Zhong Y, Li G, Xu Q, Yang Y, Tian Y, et al. scTenifoldKnk: An efficient virtual
484 knockout tool for gene function predictions via single-cell gene regulatory network perturbation.
485 Patterns (New York, NY). 2022;3:100434. https://doi.org/10.1016/j.patter.2022.100434.

486 49. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool
487 for interpreting omics data. Innovation (Cambridge (Mass)). 2021;2:100141.
488 https://doi.org/10.1016/j.xinn.2021.100141.

Figure legends

Figure 1. Expression characteristics of *SELP* in sheep.

(A) Comparison of *SELP* expression levels between high- and low-prolificacy sheep groups. Box plots show the distribution of *SELP* expression levels in the two groups, and individual points represent independent samples. **(B)** Genomic structure and regional expression profiles of the *SELP* region. The exon–intron organization, transcript structure, and transcriptional signals across the *SELP* genomic region are shown. **(C)** Genome browser visualization of the *SELP* genomic region. RNA-seq signals from ovarian and other tissues are displayed across the genomic region surrounding the *SELP* region. **(D)** *SELP* expression profiles across different tissues. Expression levels in multiple tissues, including blood-associated and endothelial-related tissues, are shown for comparison. **(E)** Cross-species expression profiles of *SELP*. Bubble plots show *SELP* expression levels and expression proportions across tissues from sheep, cattle, pigs, and humans.

Figure 2. Genetic variation, linkage disequilibrium structure, and haplotype distribution of the *SELP* candidate region.

(A) Population differentiation and functional annotation of genetic variants within the *SELP* region and its flanking region. Each point represents an individual SNP, with colors indicating genomic annotations, including upstream, downstream, exon, and intron regions. The *SELP* genomic region is indicated by the gray background. Representative variants are labeled according to their genomic positions, including the coding variant chr12:36518637 T>G. **(B)** Linkage disequilibrium (LD) structure of SNPs within the *SELP* candidate region. The triangular heatmap displays pairwise LD values (R^2) among SNPs. The upper track indicates the genomic positions of the analyzed SNPs, and black outlines denote local LD blocks. **(C)** Haplotype network of the *SELP* candidate region. Each node represents a haplotype (H1–H14), and node size corresponds to haplotype frequency. Colored sectors within each node represent the proportional contribution of different breeds or populations. Lines connecting nodes indicate inferred relationships among haplotypes. **(D)** Distribution of major

haplotype frequencies across breeds or populations. Rows represent haplotypes and columns represent breeds or populations. Values indicate haplotype frequencies, with color gradients representing differences in frequency. The locations of the candidate variants chr12:36518637 T>G (p.Ser758Arg) and chr12:36538237 are indicated in the corresponding analyses.

Figure 3. Association analysis of *SELP* candidate SNPs with litter-size traits in sheep.

(A) Allele distributions of *SELP* candidate SNPs in high- and low-prolificacy sheep groups. Bar plots show the allele counts of representative candidate variants in the two groups. **(B)** Association between genotypes of prioritized *SELP* candidate SNPs and average litter size. Genotype-based comparisons were performed for chr12:36538237 and chr12:36536704. Average litter size was compared among different genotype groups for each variant. Overall genotype effects were assessed using one-way ANOVA. Benjamini–Hochberg correction was applied across the eight SNP-level tests to control the false discovery rate. Pairwise genotype comparisons for the prioritized loci were performed using Tukey's honestly significant difference (HSD) test.

Figure 4. Evolutionary conservation and structural characteristics of ovine SELP protein.

(A) Phylogenetic relationships and conserved motif distribution of SELP proteins from representative vertebrate species. Ovine SELP and homologous proteins from other species are shown in the phylogenetic tree, together with their conserved motifs. **(B)** Multiple sequence alignment of SELP proteins from representative vertebrate species. Conserved amino acid residues and sequence regions are indicated. **(C)** Predicted transmembrane topology of ovine SELP based on TMHMM analysis. The predicted extracellular, transmembrane, and cytoplasmic regions are shown. **(D)** Predicted subcellular localization of ovine SELP. The predicted cellular and extracellular localization categories are displayed. **(E)** Signal peptide and structural region prediction of ovine SELP. SignalP analysis was used to identify the N-terminal signal peptide, and the predicted structural regions are shown.

Figure 5. Structural localization of the candidate Ser758 variant in the predicted SELP–CD44 complex.

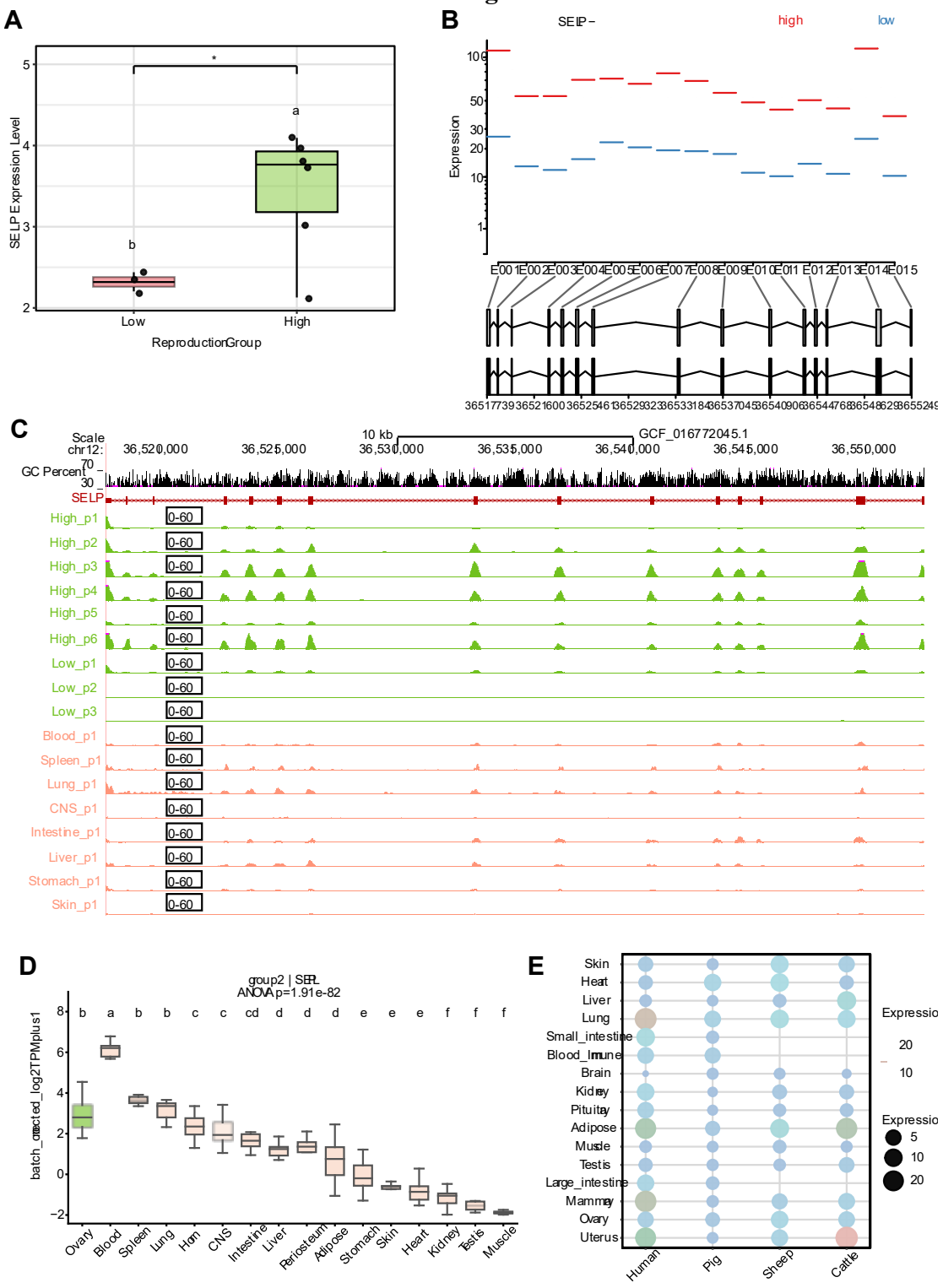
(A) Overall three-dimensional structure of the predicted SELP–CD44 complex. SELP is shown in orange and CD44 is shown in cyan. **(B)** Local views of the predicted SELP–CD44 interaction interface. Two representative regions are indicated by red boxes and shown at higher magnification in the lower panels. **(C)** Structural position of the candidate Ser758 residue and neighboring residues in the predicted SELP–CD44 complex. Ser758 is highlighted together with neighboring SELP residues Phe760 and Thr761 and CD44 residues Ser122 and Asn57. Dashed lines indicate the measured distances between selected atoms. **(D)** Representative local region of the predicted SELP–CD44 interface showing the spatial positions of SELP Lys49 and CD44 Ile172 and Ser174. Dashed lines indicate the measured atomic distances.

Figure 6. Single-cell transcriptomic analysis and virtual knockout of *SELP* in ovarian blood vascular endothelial cells.

(A) UMAP visualization of the ovarian single-cell transcriptomic dataset. Major cell populations are indicated according to cell-type annotation. **(B)** Distribution of *SELP* expression across ovarian cell populations. The expression pattern of *SELP* is shown across the annotated cell populations. **(C)** Heatmap showing representative gene expression profiles in *SELP*-negative and *SELP*-positive blood vascular endothelial cells. Expression values were displayed using a unified gene-wise normalization scale and consistent color range. Representative genes include *SELP*, *F5, IL6, CXCL9, ICAM1, NFKBIA,* and *CEBPD.* **(D)** *SELP*-centered regulatory network inferred from the virtual knockout analysis. Representative genes associated with the *SELP*-centered network include *NFKBIA, CEBPD, ICAM1, F5, IL6, CXCL9, TFPI2, INPP4B,* and *CA4*. Edges represent inferred network relationships. **(E)** Normalized perturbation manifold following *SELP* virtual knockout. Blue open circles represent the wild-type (WT) state, and red filled circles represent the *SELP* virtual knockout state. Connecting lines indicate the displacement between the two

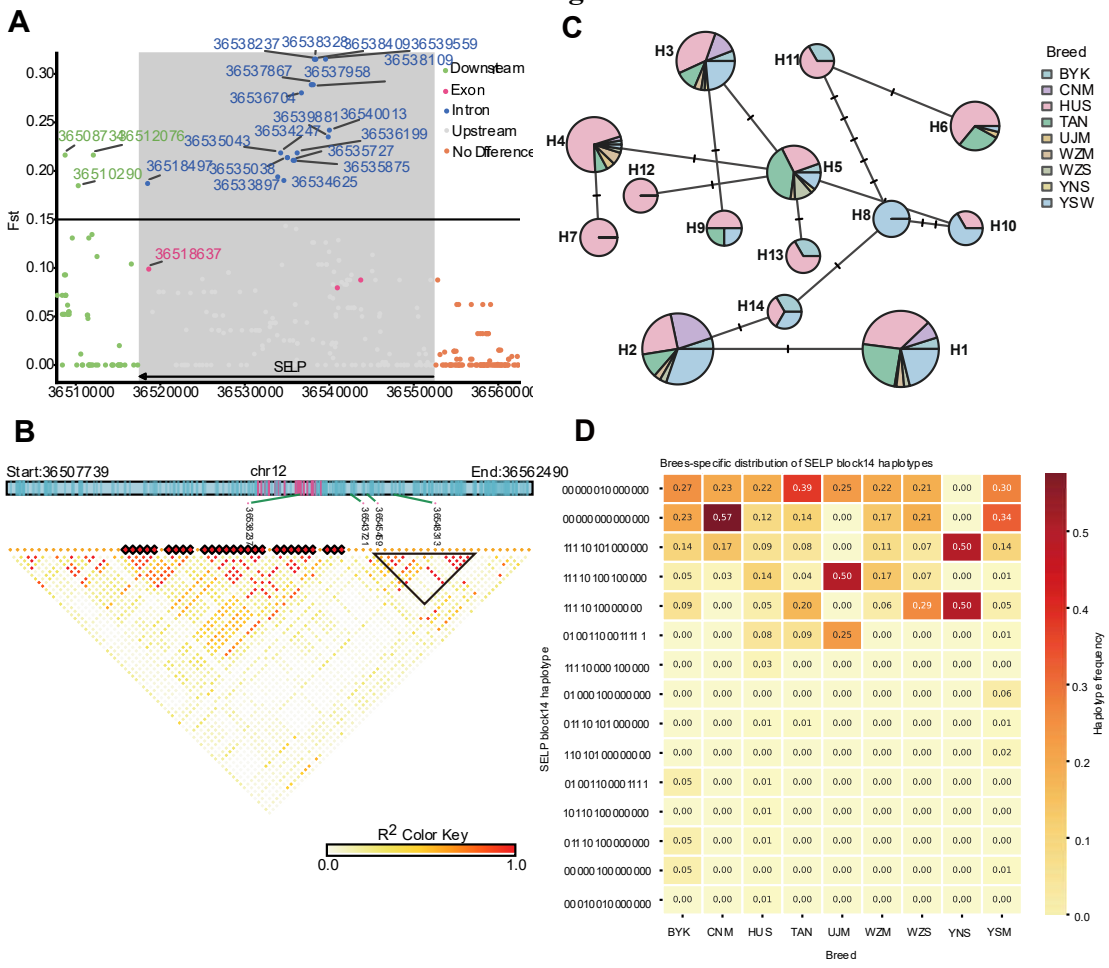
states. Vector directions were derived from the original manifold displacement, whereas vector lengths were normalized for visualization. **(F)** Functional enrichment analysis of genes associated with the *SELP* virtual knockout response. Enriched biological processes and pathways are displayed according to their statistical significance and enrichment values.

Figure 1


A
SELP Expression Level
ReproductionGroup
Low
High
B
SELP –
high
low
Expression
C
Scale
chr12:
10 kb
GCF_016772045.1
GC Percent
SELP
High_p1
High_p2
High_p3
High_p4
High_p5
High_p6
Low_p1
Low_p2
Low_p3
Blood_p1
Spleen_p1
Lung_p1
CNS_p1
Intestine_p1
Liver_p1
Stomach_p1
Skin_p1
D
group2 | SELP
ANOVA p=1.91e-82
batch_corrected_log2TPMplus1
Ovary
Blood
Spleen
Lung
Horn
CNS
Intestine
Liver
Periosteum
Adipose
Stomach
Skin
Heart
Kidney
Testis
Muscle
E
Skin
Heart
Liver
Lung
Small_intestine
Blood_Immune
Brain
Kidney
Pituitary
Adipose
Muscle
Testis
Large_intestine
Mammary
Ovary
Uterus
Human
Pig
Sheep
Cattle

571

Figure 2

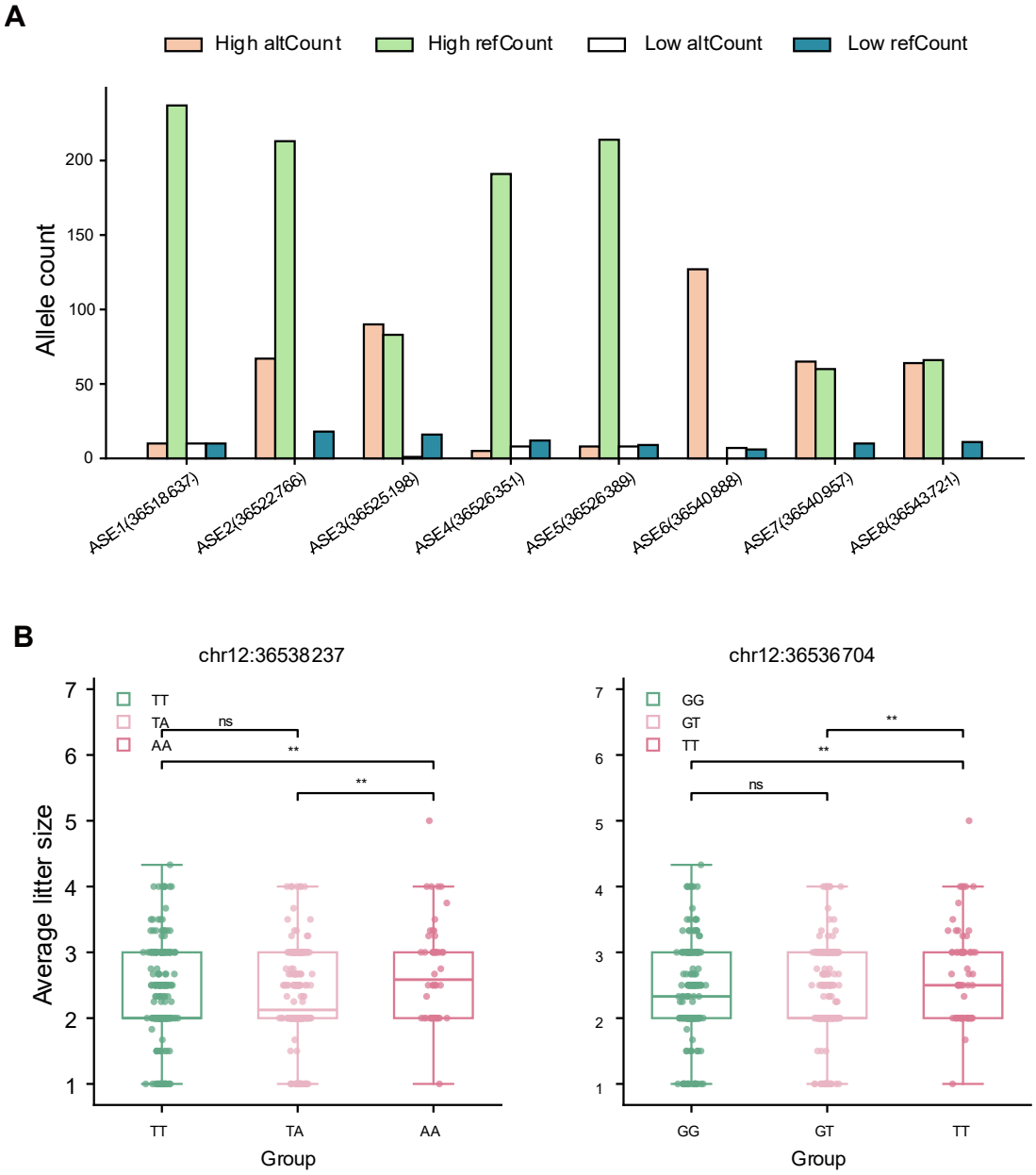


572

573

574

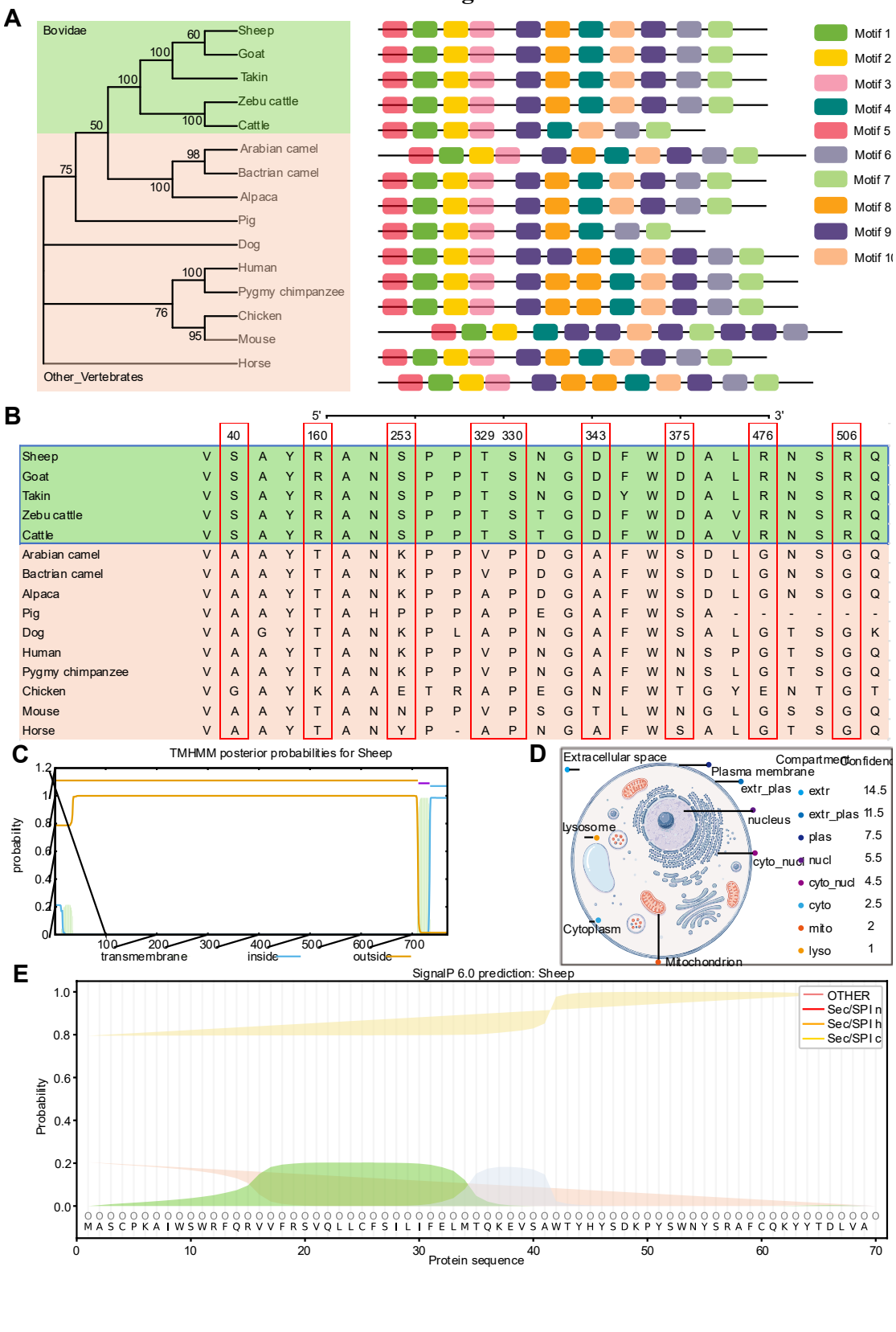
Figure 3



575

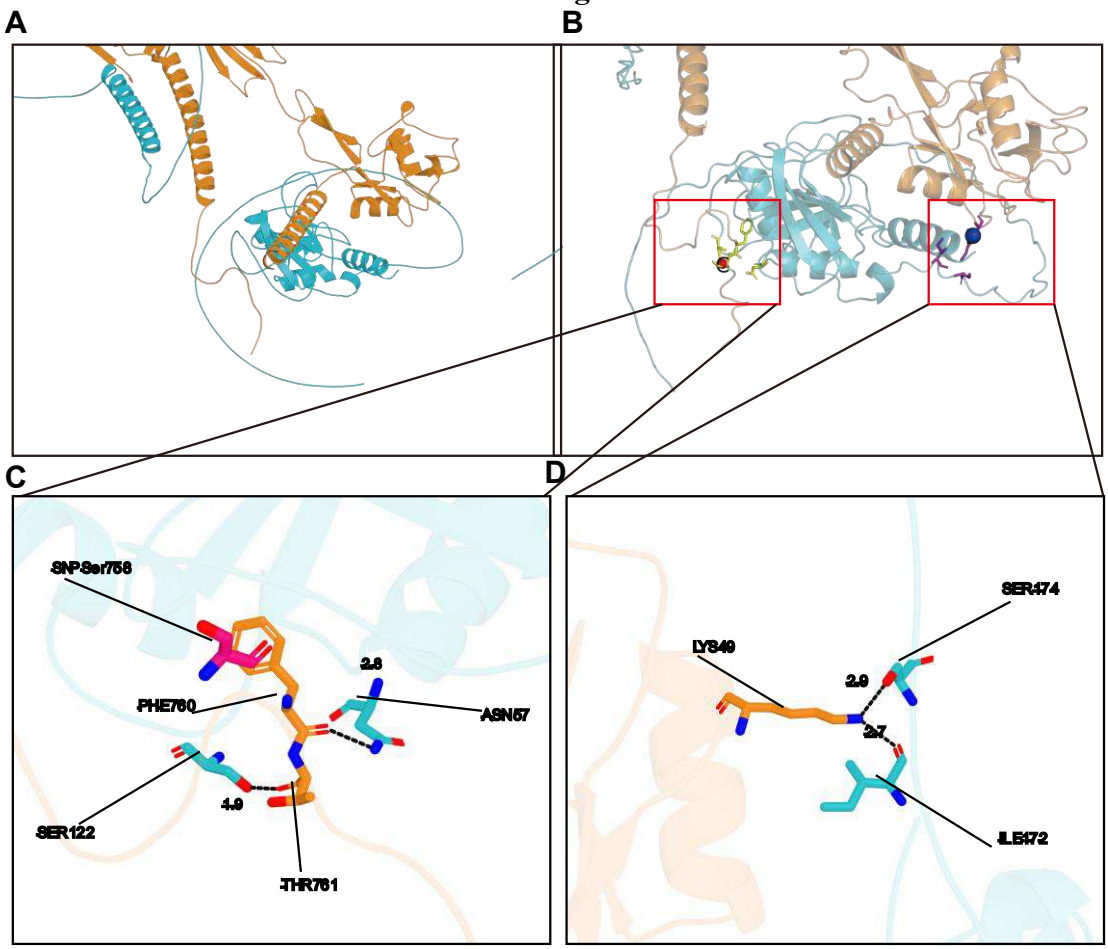
576

Figure 4


A
Bovidae
Sheep
Goat
Takin
Zebu cattle
Cattle
Arabian camel
Bactrian camel
Alpaca
Pig
Dog
Human
Pygmy chimpanzee
Chicken
Mouse
Horse
Other_Vertebrates
Motif 1
Motif 2
Motif 3
Motif 4
Motif 5
Motif 6
Motif 7
Motif 8
Motif 9
B
5'
3'
40
160
253
329 330
343
375
476
506
C
TMHMM posterior probabilities for Sheep
probability
transmembrane
inside
outside
D
Extracellular space
Plasma membrane
extr_plas
nucleus
lysosome
cyto_nucl
Cytoplasm
Mitochondrion
Compartment
Confidence
extr 14.5
extr_plas 11.5
plas 7.5
nucl 5.5
cyto_nucl 4.5
cyto 2.5
mito 2
lyso 1
E
SignalP 6.0 prediction: Sheep
Probability
Protein sequence
OTHER
Sec/SPI n
Sec/SPI h
Sec/SPI c

580

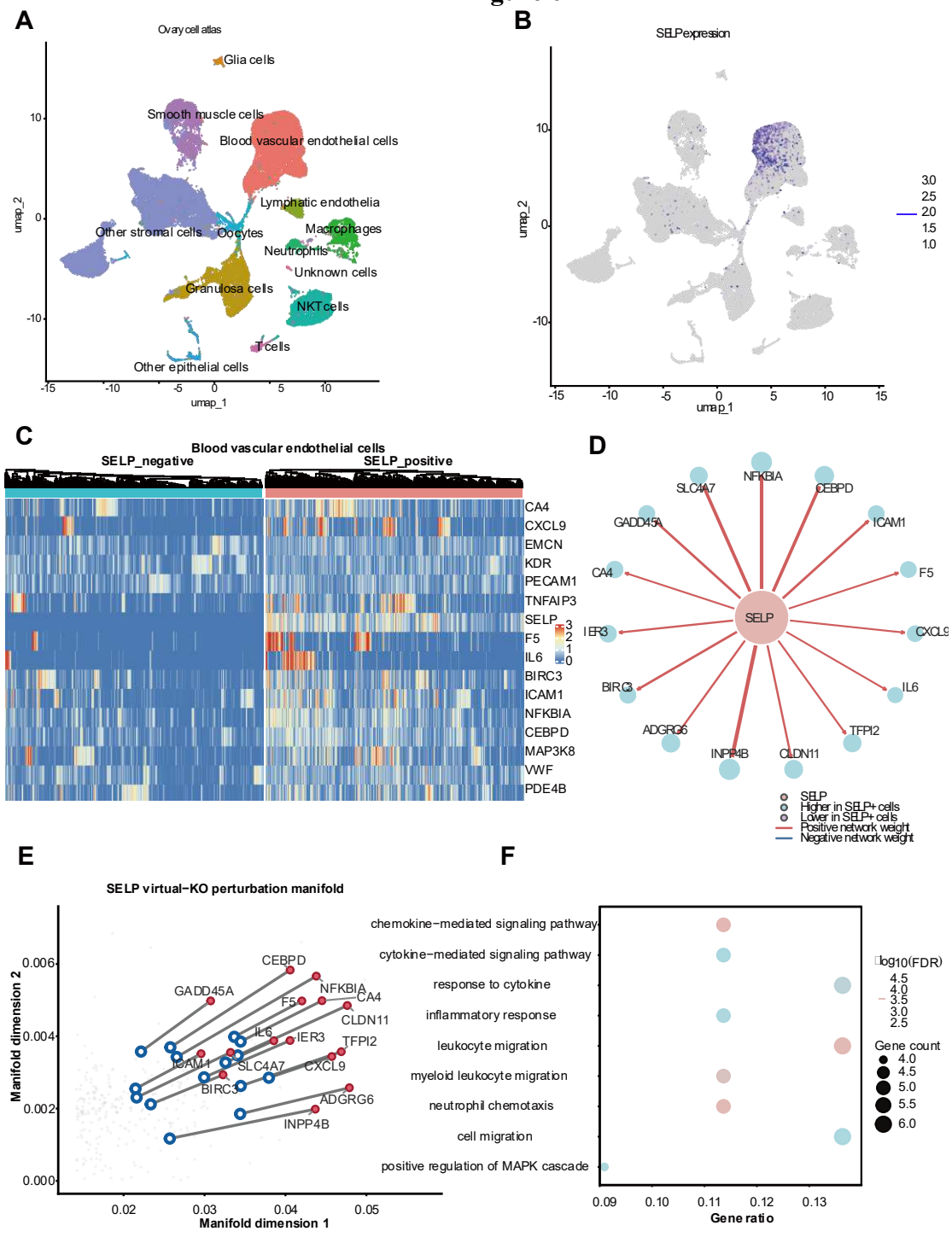
Figure 5



581

582

Figure 6


A
Ovary cell atlas
Glia cells
Smooth muscle cells
Blood vascular endothelial cells
Lymphatic endothelia
Other stromal cells
Oocytes
Macrophages
Neutrophils
Unknown cells
Granulosa cells
NKTcells
Tcells
Other epithelial cells
umap_1
umap_2
B
SELPexpression
C
Blood vascular endothelial cells
SELP_negative
SELP_positive
CA4
CXCL9
EMCN
KDR
PECAM1
TNFAIP3
SELP
F5
IL6
BIRC3
ICAM1
NFKBIA
CEBPD
MAP3K8
VWF
PDE4B
D
SLC4A7
NFKBIA
CEBPD
GADD45A
ICAM1
CA4
F5
SELP
IER3
CXCL9
BIRC3
IL6
ADGRG6
TFPI2
INPP4B
CLDN11
SELP
Higher in SELP+ cells
Lower in SELP+ cells
Positive network weight
Negative network weight
E
SELP virtual−KO perturbation manifold
CEBPD
GADD45A
NFKBIA
F5
CA4
CLDN11
IL6
IER3
TFPI2
ICAM1
SLC4A7
CXCL9
BIRC3
ADGRG6
INPP4B
Manifold dimension 1
Manifold dimension 2
F
chemokine−mediated signaling pathway
cytokine−mediated signaling pathway
response to cytokine
inflammatory response
leukocyte migration
myeloid leukocyte migration
neutrophil chemotaxis
cell migration
positive regulation of MAPK cascade
Gene ratio
log10(FDR)
Gene count

586 **Table S1 SELP candidate SNP**

587 **Table S2 SELP differentiated variants identified by population genetic analysis**

588 **Table S3 SELP LD block composit**