



Article Genome-Wide Analyses Reveal Genetic Convergence of Prolificacy between Goats and Sheep

Lin Tao¹, Xiaoyun He¹, Yanting Jiang², Yufang Liu^{1,3}, Yina Ouyang², Yezhen Shen⁴, Qionghua Hong^{2,*} and Mingxing Chu^{1,*}

- ¹ Key Laboratory of Animal Genetics, Breeding and Reproduction of Ministry of Agriculture and Rural Affairs, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China; 82101182357@caas.cn (L.T.); hexiaoyun@caas.cn (X.H.); yufangliu@hebeu.edu.cn (Y.L.)
- ² Yunnan Animal Science and Veterinary Institute, Kunming 650224, China; jiangyanting-2007@163.com (Y.J.); yinaouyang@163.com (Y.O.)
- ³ College of Life Science and Food Engineering, Hebei University of Engineering, Handan 056038, China
- ⁴ Annoroad Gene Technology Co., Ltd., Beijing 100176, China; yezhenshen@genome.cn
- * Correspondence: yxh7168@163.com (Q.H.); chumingxing@caas.cn (M.C.)

Abstract: The litter size of domestic goats and sheep is an economically important trait that shows variation within breeds. Strenuous efforts have been made to understand the genetic mechanisms underlying prolificacy in goats and sheep. However, there has been a paucity of research on the genetic convergence of prolificacy between goats and sheep, which likely arose because of similar natural and artificial selection forces. Here, we performed comparative genomic and transcriptomic analyses to identify the genetic convergence of prolificacy between goats and sheep. By combining genomic and transcriptomic data for the first time, we identified this genetic convergence in (1) positively selected genes (*CHST11* and *SDCCAG8*), (2) differentially expressed genes (*SERPINA14*, *RSAD2*, and *PPIG* at follicular phase, and *IGF1*, *GPRIN3*, *LIPG*, *SLC7A11*, and *CHST15* at luteal phase), and (3) biological pathways (genomic level: osteoclast differentiation, ErbB signaling pathway; transcriptomic level: the regulation of viral genome replication at follicular phase, and protein kinase B signaling and antigen processing and presentation at luteal phase). These results indicated the potential physiological convergence and enhanced our understanding of the overlapping genetic makeup underlying litter size in goats and sheep.

Keywords: litter size; genetic convergence; genome; RNA-seq; sheep; goats

1. Introduction

As ruminant livestock, sheep and goats are globally distributed and play important roles in providing meat, milk, and fur to humans. One primary objective of sheep and goat breeding programs is to increase the yield by phenotypic and genetic selection, and female fecundity is a basic and crucial key. Litter size (LS), the lambing number per parturition per ewe or the kidding number per parturition per doe, is an important component of goat and sheep productivity. Importantly and interestingly, variation in individual LS is common within breeds, such as 1–4 in Yunshang black goats, 1–5 Romanov sheep, and 1–6 in Finnsheep [1,2]. Breeds with exceptional prolificacy can provide genetic material to develop new lines/populations with ideal performance.

It is important to elucidate the genetic mechanisms of LS to efficiently utilize these prolific resources. Therefore, numerous studies have been separately performed to enhance our understanding of goat and sheep LS [3–7]. Compared with sheep, knowledge of hircine prolificacy is in its infancy. Thus, previous studies explored the association between hircine LS and homologous genetic markers first identified in sheep (such as *FecB* and *FecX*) [8,9], in which the genetic convergence in prolificacy between goats and sheep was expected in principle, at least at the single gene level. However, the approach using



Citation: Tao, L.; He, X.; Jiang, Y.; Liu, Y.; Ouyang, Y.; Shen, Y.; Hong, Q.; Chu, M. Genome-Wide Analyses Reveal Genetic Convergence of Prolificacy between Goats and Sheep. *Genes* 2021, *12*, 480. https://doi.org/ 10.3390/genes12040480

Academic Editors: Cord Drögemüller and Gesine Lühken

Received: 24 February 2021 Accepted: 23 March 2021 Published: 26 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). candidate genes/loci did not work well. Notwithstanding the convergent genomic signatures of domestication and high-altitude adaptation that have been explored in sheep and goats [10,11], thus far, there has been a paucity of research on the genetic convergence of prolificacy from a whole-genome perspective. We hypothesize that the genetic convergence of prolificacy occurred in goats and sheep, likely because they were exposed to similar natural and/or artificial selection forces.

Therefore, the objective of this study was to identify the genetic convergence of prolificacy between goats and sheep using whole-genome data by examining (1) positively selected genes (PSGs), (2) differentially expressed genes (DEGs), and (3) biological pathways. These results will enhance our knowledge of prolificacy and can be used to improve the molecular design breeding of both goats and sheep.

2. Materials and Methods

2.1. Animals, Phenotypes and Sample Collection

This study included two Yunshang black goat populations (Yunshang black goats used for DNA sequencing (YSD), n = 40; Yunshang black goats used ovarian mRNA sequencing (YSR), n = 20) and three sheep populations (Finnsheep, FIN, n = 19; Romanov, ROM, n = 19; and Small Tailed Han sheep, STH, n = 12); the YSD, FIN and ROM data were obtained from previous studies [1,2]. All animals were as unrelated as possible as determined by pedigree or farmers' knowledge. On the basis of average LS, the five populations were classified into a low-yield group (LG, control) and a high-yield group (HG, case) to represent extreme individuals within each breed (Table S1). For each population, a highly significant difference was observed in LS between the HG and LG (p-value < 0.01, Mann–Whitney U test, Table 1).

Derry letter 1	Average Litter S		
Population -	Low-Yield Group	High-Yield Group	<i>p</i> -value -
YSD	1.32 ± 0.19 (<i>n</i> = 20)	3.00 ± 0.38 (<i>n</i> = 20)	$5.16 imes10^{-8}$
YSR	$1.75 \pm 0.26 \ (n = 10)$	$3.30 \pm 0.35 \ (n = 10)$	$1.21 imes 10^{-4}$
FIN	$1.66 \pm 0.33 \ (n = 9)$	$3.59 \pm 0.41 \ (n = 10)$	$2.76 imes10^{-4}$
ROM	$1.53 \pm 0.25 \ (n = 9)$	3.13 ± 0.17 (<i>n</i> = 10)	$2.01 imes10^{-4}$
STH	$1.00 \pm 0.00 \ (n = 6)$	$2.89 \pm 0.17 \ (n = 6)$	$2.22 imes 10^{-3}$

Table 1. Phenotypes of does and ewes used in this study.

¹ YSD and YSR indicate Yunshang black goats used for DNA sequencing and ovarian mRNA sequencing, respectively. FIN: Finnsheep; ROM: Romanov sheep; STH: Small Tailed Han sheep. ² Mann–Whitney *U* test. The number in the parentheses is sample size.

To detect PSGs for LS, 20 YSD (LS \leq 1.5), 9 FIN (LS \leq 2), and 9 ROM (LS \leq 1.8) were labeled as controls, and 20 YSD (LS > 2.3), 10 FIN (LS > 3.28), and 10 ROM (LS \geq 3) were labeled as cases. Likewise, to identify DEGs for LS, 10 YSR (LS \leq 2) and 6 STH (LS = 1) were selected as controls, and 10 YSR (LS \geq 3) and 6 STH (LS > 2.66) were selected as cases. For ovary collection, estrus synchronization with controlled internal drug release (CIDR, progesterone 300 mg) for 16 and 12 days was separately performed for YSR and STH. Finally, 10 YSR (5 cases and 5 controls) and 6 STH (3 cases and 3 controls) were slaughtered within 45–48 h of CIDR removal for ovary harvest at follicular phase (FP), and the remaining 10 YSR and 6 STH were slaughtered on days 11 and 9 for ovary harvest at luteal phase (LP).

2.2. Whole-Genome Single Nucleotide Polymorphism (SNP)

Genomic DNA was extracted from the jugular venous blood of YSD and ear marginal tissues of FIN and ROM following a standard phenol–chloroform method. To obtain genomic SNPs of YSD, whole-genome sequencing of a single library was performed using the Illumina NovaSeq 6000 system (Illumina Inc., San Diego, CA, USA), and 150-bp paired-end reads were generated. To obtain clean reads, the following unqualified raw reads were removed: (1) adaptor-polluted reads, (2) reads with over 50% bases with a Phred score <19, and (3) reads with over 5% N bases. Clean reads were mapped to the ASR1 reference genome using a Burrows–Wheeler Aligner v0.7.9a [12], with an average depth of ~20-fold coverage, and SNPs were subsequently called using a Genome Analysis Toolkit v3.3 [13]. FIN and ROM were genotyped by the Ovine Infinium HD BeadChip, which comprises 606,006 SNPs. To obtain eligible SNPs and samples, the following quality control criteria were implemented in Plink v1.90 [14]: (1) call rate >90%, (2) SNPs on autosomes, (3) minor allele frequency >1%, (4) missing genotype frequency <10%, and (5) Hardy–Weinberg disequilibrium *p*-value > 10^{-6} .

2.3. Population Structure

To explore the genetic relationship of individuals from YSD, FIN and ROM, Plink command "–indep-pairwise 50 5 0.2" was run to generate independent effective SNP sets for principal component analysis [14].

2.4. Selective Sweeps

To detect PSGs for LS, the fixation index (Fst), which is a measure of population differentiation, was calculated in VCFtools v0.1.14 [15] using Weir and Cockeram's method [16] to compare LG and HG for YSD (Fst(YSD)), FIN (Fst(FIN)) and ROM (Fst(ROM)). For YSD, a 100-kb window with a 50-kb step size was used. For sheep, Fst(FIN) and Fst(ROM) were estimated SNP by SNP, and their root mean square (Fst_RMS) was then calculated with the following formula:

$$Fst_RMS = \sqrt{\frac{1}{2}} (Fst(FIN)^2 + Fst(ROM)^2)$$

To calculate Fst_RMS, only SNPs shared by FIN and ROM and those with nonnegative Fst values were used. The results were visualized using R package qqman [17].

To identify loci with extremely high Fst values that drive reproductive differences of LG and HG, we ranked the regions/loci in descending order on the basis of Fst(YSD) and Fst_RMS, respectively. The regions with extremely high Fst(YSD) (0.05, top 0.8%) were treated as candidates for goats, and the SNPs with extremely high Fst_RMS (0.3, top 0.7%) were selected as candidates for sheep. Finally, PSGs wholly or partially located at these regions or covering candidate SNPs were detected by ANNOVAR v2020-06-07 [18].

To explore the association of three HS3ST2 variations and reproductive performance, an additional 544 2–5-year-old female Yunshang black goats were genotyped using a MassARRAY[®] SNP system. The primer sequences are in Table S2. Chi-squared test (all expected frequencies \geq 5) or Fisher's exact test were employed to detect associations between the genotype and LS of Yunshang black goats.

2.5. Ovarian RNA-Seq

The total RNA of hircine ovaries for LG and HG at FP and LP (five biological replicates) was extracted and sequenced using the Illumina NovaSeq 6000 platform (Illumina Inc., San Diego, CA, USA). After filtering out the adapter reads and unqualified reads, clean reads were mapped to the ASR1 reference genome using HISAT2 [19], followed by the assembly and calculation of read count and fragments per kilobase per million values in StringTie [20], and identification of DEGs with DESeq2 [21]. In the two comparisons of LG vs. HG, genes were accepted as being differentially expressed when (1) *q*-value (*p*-value adjusted by the Benjamini–Hochberg method) <0.5, and $|Log_2FoldChang| \ge 0.24$.

Likewise, the total RNA of 12 ovine ovaries for LG and HG at FP and LP (three biological replicates) were subjected to the same processing, except they were mapped to the Oar_v3.1 reference genome.

2.6. Identification of Genetic Convergence

To reveal the potential genetic convergence in goat and sheep prolificacy, we focused on (1) PSGs, (2) DEGs, and (3) biological pathways at the level of gene set. Given the availability of PSGs in HG using the number of segregating sites by length (nSL), a haplotype-based statistic with test power to both soft and hard sweeps [2,22], convergent PSGs were identified as the intersection of ovine candidates identified by Fst RMS and hircine candidates shared by Fst(YSD) and nSL (≥ 6) of HG. Convergent DEGs at LP and FP were defined as these genes that were shared by goats and sheep, but excluded their expression trend. The number of genes shared by Fst(YSD) and nSL of HG is limited and nonideal for gene set analysis; therefore the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed for the PSGs on the basis of Fst(YSD) and Fst_RMS using Metascape [23] to obtain the related gene set which was composed of functionally associated PSGs. Terms with *p*-value < 0.01 were considered to be significantly different. Similarly, DEGs were also used to identify the connected gene set for prolificacy between goats and sheep. Therefore, convergent biological pathways were identified as the mutual GO (including only biological processes) and KEGG Ontology (ko) terms.

3. Results and Discussion

3.1. Population Structure

The principal component analysis (Figure S1) demonstrated a clear split between FIN and ROM along the first principal component (PC), indicating that they were independent populations. However, they were northern European short-tailed breeds with common ancestry components [24]. Therefore, they are thought to be useful for exploring the genetic mechanisms underlying LS as single and separate herds. For both goats and sheep, the cases and controls within each breed were not split by the first two PCs, which was consistent with the circumstances of dogs in a previous report on behavior [25]. These findings indicate that population structure was not linked to LS of each breed.

3.2. Convergent PSGs Linked to LS

To reveal the signatures of selection for goat prolificacy, a set of 350 PSGs was identified by Fst(YSD), of which *HS3ST2* on chromosome 25 was found to have the strongest selection signature (Figure 1 and Table S3). The variation g.20423876C>T at *HS3ST2* was significantly associated with LS (Table S4). Coding heparan sulfate-glucosamine 3-sulfotransferase 2, *HS3ST2* was reported to be linked to circadian rhythm regulation in the pineal gland of rats [26]. *HS3ST5*, an *HS3ST2* paralog, was associated with reproductive seasonality traits in sheep [27]. Given the association of seasonality-associated *DIO3* and LS in pigs and goats, and the empirical concurrence of higher LS and year-round estrus in goats [2,28], these results indicate that *HS3ST2* may be involved in female reproductive health, and seasonal breeding in goats.



Figure 1. Manhattan plots for fixation index (Fst)(YSD) and root mean square of fixation index (Fst_RMS). The red lines indicate the cut-offs of Fst(YSD) (0.05, top 0.8%) and Fst_RMS (0.3, top 0.7%).

Eight overlapping PSGs were revealed on the basis of Fst(YSD) and nSL of HG in goats: CHST11, SDCCAG8, ATF6, FAM13A, HSP90B1, IBSP, IFNLR1, and STAB2 (Figure 2). Of these, two genes (CHST11 and SDCCAG8) were also under selection in sheep for which 162 PSGs were identified by Fst_RMS (Figures 1 and 2). This number was significantly larger than expected (p-value = 0.05), suggesting that CHST11 and SDCCAG8 are significantly convergent in goat and sheep prolificacy. CHST11 (also known as C4ST-1), which encodes carbohydrate sulfotransferase 11, is a target gene of transforming growth factor β (TGFβ)-like growth factors (such as TGFβ, BMP2, and Activin) [29]. Chst11 was found to regulate the proliferation and apoptosis of embryo chondrocytes in mice [30]. The mRNA level of CHST11 increases upon the interaction of BMP4 with a phospho-Smad3 binding site in the CHST11 promoter [31]. Furthermore, CHST11 was linked to ovarian cancer [32], endometrial carcinomas [33], and porcine farrowing interval [34]. However, the CHST11 transcripts were not differentially expressed in goat ovary (Table S5). Another PSG, SDCCAG8, is a risk gene for Bardet–Biedl syndrome, which is primarily characterized by postaxial polydactyly, hypogonadotropic hypogonadism, and/or genitourinary malformations, and secondly characterized by polycystic ovary syndrome [35]. SDCCAG8 is a candidate gene for hind limb and rib cage malformations in mice [36,37] and teat number in pigs [38]. Interestingly, BMPR1B, a major gene for sheep LS, plays important roles in both chondrogenesis and embryogenesis [39,40], indicating the pleiotropism of candidate genes for bone formation and reproduction. This was, to some extent, supported by the "osteoclast differentiation" pathway, which was enriched by some PSGs for LS (Table 2). Moreover, there is a transcription factor binding site at SDCCAG8 for SMAD5, a member of the TGF β signaling pathway, and strong differentiation signals have been observed in SMAD5 of Shaanbei White Cashmere goats with different LS [7]; this may contribute to the differential expression of SDCCAG8 at both LP and FP (Table S5). Taken together, these findings indicate that CHST11 and SDCCAG8 regulate LS via TGFβ and BMP signaling in goats and sheep.



Figure 2. Venn diagrams of genetic convergence of prolificacy between goats and sheep. The top panels denote the convergence at the single gene level (positively selected genes and differentially expressed genes of ovaries), whereas the convergence at the gene set level is represented as Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Ontology (ko) terms in the bottom panels; only biological processes (BP) were considered in GO analysis.

	Table 2. Pathway	ys shared by	positively	y selected g	enes of sheep	and goats.
--	------------------	--------------	------------	--------------	---------------	------------

	N T	—lg(<i>p</i>)		Gene		
lerm	Name	Sheep	Goat	Sheep	Goat	
hsa04012	ErbB signaling pathway	3.92	2.18	CBL PIK3CA MAPK10 KCNH8 NRG4	CAMK2B GRB2 MAPK10 SRC AKT3	
ko04380	Osteoclast differentiation	3.27	2.89	MAPK14 NFATC2 PIK3CA MAPK10 TGFB2	FCGR2A FCGR2B FCGR3A FCGR3B GRB2 MAPK10 TEC AKT3	
hsa04380	Osteoclast differentiation	3.14	2.72	MAPK14 NFATC2 PIK3CA MAPK10 TGFB2	FCGR2A FCGR2B FCGR3A FCGR3B GRB2 MAPK10 TEC AKT3 RHBDF2	
ko04012	ErbB signaling pathway	3.01	2.34	CBL PIK3CA MAPK10 NRG4	CAMK2B GRB2 MAPK10 SRC AKT3	
hsa04926	Relaxin signaling pathway	2.29	2.13	MAPK14 PIK3CA MAPK10 CREB5	ADCY3 GNG5 GRB2 MMP9 MAPK10 SRC AKT3	

3.3. Convergent DEGs Associated with LS

For goats, compared with LG, 2207 (1126 up- and 1081 down-regulated) and 1665 DEGs (826 up- and 839 down-regulated) were identified in HG at LP and FP, respectively (Figure 3 and Tables S6 and S7). For sheep, 104 and 39 DEGs were identified in HG compared with LG at LP and FP, respectively (Figure 3 and Tables S8 and S9). In total, eight overlapping DEGs were shared by goats and sheep, including three (SERPINA14, RSAD2, and PPIG) at FP and five (IGF1, GPRIN3, LIPG, SLC7A11, and CHST15) at LP (Figures 2 and 4). Of these, *IGF1* and *SLC7A11* were identified as DEGs in the transcriptomic analysis of ovaries from pigs with high and low LS [41]. Additionally, SLC7A11 is differentially expressed in ovaries among sheep breeds with different fecundity [42]. IGF1 is a candidate gene for LS in goats and sheep [43–45], suggesting the convergence at the levels of DNA and change in mRNA expression.

Differentially expressed genes of ovaries



Figure 3. Volcano plots of differentially expressed genes in the ovaries of Yunshang black goats (**A**,**B**) and Small Tail Han sheep (**C**,**D**). (**A**,**C**) indicate luteal phases, and (**B**,**D**) denote follicular phases. Compared with the low-yield groups, points in red, blue and green are the down-regulated, up-regulated, and non-differentially expressed genes in high-yield groups, respectively. The dashed lines are thresholds of 1.30 (-lg(0.05)).



Figure 4. Differentially expressed genes of ovaries shared by sheep and goats. FP and LP denote follicular and luteal phases, respectively.

Previous studies reported the differential expression of LIPG mRNA or protein in the theca interna during the transition from small to large antral bovine follicles [46], in the placenta of Berkshire pigs with different LS [47], and in the human follicular fluid of poor ovarian responders compared with a control group [48]. Additionally, a genome-wide association study on Large White pigs identified *LIPG* as a candidate gene for piglet uniformity measured as the coefficient of variation of birth weights within one litter [49]. These findings highlight the potential roles in follicular development and pregnancy. Secreted from uterine endometrium and induced by progesterone, uterine serpin (SERPINA14) plays crucial roles in the pregnancy of farm animals, such as nutrient supply to the conceptus, immunoregulation, and maternal recognition of pregnancy [50–52]. *RSAD2* is a marker for dendritic cell maturation [53] and early pregnancy diagnosis in dairy cows [54]. Moreover, compared with healthy follicles, *RSAD2* was up-regulated in early atresia and progressive atresia groups in pigs [55]. The functions of *GPRIN3*, *CHST15*, and *PPIG* in reproduction by

multiple processes, such as follicular development, immunoregulation, and pregnancy maintenance.

3.4. Convergent Biological Pathways

At the single gene level, it is possible to infer convergent signals from overlapping DEGs or PSGs between goats and sheep. At the gene set level, GO and KEGG pathway analyses for PSGs and DEGs were implemented to identify the overlapping biological pathways that include a series of functionally connected genes. For goat PSGs, 192 GO and 30 ko terms were revealed, as expected, six acrosome reaction-related GO terms were detected: binding of sperm to zona pellucida (GO:0007339), regulation of acrosome reaction (GO:0060046), single fertilization (GO:0007338), regulation of fertilization (GO:0080154), sperm–egg recognition (GO:0035036), and fertilization (GO:0009566) (Table S10). Additionally, some endocrine resistance ko terms were identified: endocrine resistance (ko01522), estrogen signaling pathway (ko04915 and hsa04915), relaxin signaling pathway (hsa04926), and fluid shear stress and atherosclerosis (hsa05418) (Table S11). For sheep PSGs, we identified 132 GO and 58 ko terms, of which osteoclast differentiation (ko04380 and hsa04380), ErbB signaling pathway (hsa04012 and ko04012), and relaxin signaling pathway (hsa04926) were shared by goats (Figure 2 and Table 3).

Table 3. Terms shared by differentially expressed genes of sheep and goats.

Term	Name	$-\lg(p)$		Gene		
		Goat	Sheep	Goat	Sheep	
Protein kinase B signaling	GO:0043491	2.99	2.58	TCF7L2 IGFBP5 RPS6KB2 MYOC LOC102175889	IGF1 KIT LOX TRK2 PKHD1	
Antigen processing and presentation	ko04612	2.08	3.86	PDIA3 CREB1 CALR NFYC	HLA-A HLA-BLA-C HLA-F	
Regulation of viral genome replication	GO:0045069	2.25	6.15	RSAD2 VAPA	IFIT1 MX1 ISG15 RSAD2	

For goat DEGs, as expected, the phosphatidylinositol 3-kinase (PI3K)-Akt signaling pathway (ko04151), oocyte meiosis (ko04114), progesterone-mediated oocyte maturation (ko04914), and estrogen signaling pathway (ko04915) were in the top 10 ko terms (Figure S2), which reemphasizes the significance of some biological processes in LS. Consequently, goats shared three terms with sheep: protein kinase B signaling (GO:0043491) and antigen processing and presentation (ko04612) at LP, and the regulation of viral genome replication (GO:0045069) at FP (Table 3 and Figure 2). An implication of these convergent biological pathways is that there may be differences in relaxin concentration during pregnancy, immunity, and disease resistance that contribute to LS in goats and sheep. Moreover, it is helpful to pinpoint the potential causal genes for LS, such as *ADCY3*, which is necessary for male fertility [56]; *TGFB2*, which participates in ovine fecundity [57]; and *IGFBP5*, which is associated with LS in pigs [58]. Therefore, the convergence of biological pathways from the perspective of gene sets and single genes, indicates potential physiological convergence and enhances our understanding of the genetic basis underlying goat and sheep LS.

From a genomic and transcriptomic standpoint, we attempted to provide a genomewide landscape to reveal the genetic convergence of prolificacy between goats and sheep. The results may indicate the physiological convergence of goats and sheep; however, more research is needed to characterize reproductive hormones and biological factors. However, these results may differ among other populations/breeds because of the sample size and/or genetic backgrounds in this study. Therefore, it would be useful to integrate an abundance of multi-omics data, such as genomics, transcriptomics, proteomics, and metabolomics, for additional populations or reproductive organs in a large population.

4. Conclusions

By combining genomic and transcriptomic data, this study identified the genetic convergence of prolificacy between goats and sheep for the first time on the basis of (1) PSGs (*CHST11* and *SDCCAG8*), (2) DEGs (*SERPINA14*, *RSAD2*, and *PPIG* at FP, and *IGF1*, *GPRIN3*, *LIPG*, *SLC7A11*, and *CHST15* at LP), and (3) biological pathways (PSG: osteoclast differentiation, ErbB signaling pathway and relaxin signaling pathway; DEG: regulation of viral genome replication at FP, and protein kinase B signaling and antigen processing and presentation at LP). These results indicated potential physiological convergence and enhanced our understanding of the overlapping genetic makeup underlying LS between goats and sheep.

Supplementary Materials: The following are available online at https://www.mdpi.com/2073-442 5/12/4/480/s1, Figure S1: Principal component analysis of goats and sheep, Figure S2: Top 10 KEGG terms of differentially expressed genes from luteal phase (LP) and follicular phase (FP) in Yunshang black goats, Table S1: Phenotypes of animals involved in this study, Table S2: Primer sequences of *HS3ST2*, Table S3: Positively selected genes identified by Fst, Table S4: Association between loci on *HS3ST2* and litter size of Yunshang black goats, Table S5: Ovarian mRNA expression of *CHST11* and *SDCCAG8* in Yunshang black goats, Table S6: Differentially expressed mRNAs of Yunshang black goat ovaries at luteal phase, Table S7: Differentially expressed mRNAs of Small Tailed Han sheep ovaries at follicular phase, Table S9: Differentially expressed mRNAs of Small Tailed Han sheep ovaries at luteal phase, Table S9: Differentially expressed mRNAs of Small Tailed Han sheep ovaries at follicular phase, Table S9: Differentially expressed mRNAs of Small Tailed Han sheep ovaries at follicular phase, Table S9: Differentially expressed mRNAs of Small Tailed Han sheep ovaries at follicular phase, Table S9: Differentially expressed mRNAs of Small Tailed Han sheep ovaries at follicular phase, Table S10: Functional category of positively selected genes in Yunshang black goats, Table S11: Terms related to acrosome reaction and endocrine resistance.

Author Contributions: Conceptualization, M.C. and Q.H.; methodology, L.T.; software, L.T. and Y.S.; validation, L.T., X.H. and Y.L.; formal analysis, L.T.; investigation, L.T., Y.J. and Y.O.; resources, M.C. and Q.H.; data curation, L.T. and Y.S.; writing—original draft preparation, L.T.; writing—review and editing, X.H., Q.H. and M.C.; visualization, L.T.; supervision, M.C. and Q.H.; project administration, M.C.; funding acquisition, M.C. and Q.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the National Natural Science Foundation of China (31772580), the Earmarked Fund for China Agriculture Research System (CARS-38), the Agricultural Science and Technology Innovation Program of China (CAAS-ZDRW202106 and ASTIP-IAS13), the Science and Technology Innovation Talents Program of Yunnan Province (2018HC017), the Basic Research Foundation Key Project of Yunnan Province (202001AS070002). The APC was funded by 31772580.

Institutional Review Board Statement: STH, YSD and YSR involved in this study were approved by the Science Research Department (in charge of animal welfare issues) of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (IAS-CAAS) (Beijing, China) and ethical approval was given by the Animal Ethics Committee of the IAS-CAAS (No. IASCAAS-AE-03 for STH, and No. IAS2019-63 for YSD and YSR).

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw sequencing data of 40 Yunshang black goats (YSD) have been deposited in the NCBI SRA database under accession code PRJNA611688. Genotype and phenotype datasets for Finnsheep and Romanov sheep were available at the following link: https://www.animalgenome.org/repository/pub/CAAS2018.0302/ (accessed on 24 September 2020). The RNA-seq data reported in this article are available upon request for research purpose.

Acknowledgments: We are grateful to all volunteers who assisted in sampling. We thank Liwen Bianji, Edanz Group China (www.liwenbianji.cn/ac, accessed on 10 March 2021), for editing the English text of a draft of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest. Yezhen Shen is employee of Annoroad Gene Technology Co. Ltd., Beijing, China. This article has no conflict of interest with the company.

Abbreviations

LS	litter size
LG	low-yield group
HG	high-yield group
FP	follicular phase
LP	luteal phase
PSGs	positively selected genes
DEGs	differentially expressed genes
GO	gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
ko	KEGG ontology
YSD	Yunshang black goats for DNA sequencing
YSR	Yunshang black goats for mRNA sequencing
FIN	Finnsheep
ROM	Romanov sheep
STH	Small Tailed Han sheep
Fst_RMS	root mean square of fixation index (Fst)
nSL	number of segregating sites by length

References

- Xu, S.S.; Gao, L.; Xie, X.L.; Ren, Y.L.; Shen, Z.Q.; Wang, F.; Shen, M.; Eyþórsdóttir, E.; Hallsson, J.H.; Kiseleva, T.; et al. Genomewide association analyses highlight the potential for different genetic mechanisms for litter size among sheep breeds. *Front. Genet.* 2018, 9, 118. [CrossRef]
- Tao, L.; He, X.Y.; Jiang, Y.T.; Lan, R.; Li, M.; Li, Z.M.; Yang, W.F.; Hong, Q.H.; Chu, M.X. Combined approaches to reveal genes associated with litter size in Yunshang black goats. *Anim. Genet.* 2020, *51*, 924–934. [CrossRef] [PubMed]
- 3. Yao, Y.; Pan, Z.; Di, R.; Liu, Q.; Hu, W.; Guo, X.; He, X.; Gan, S.; Wang, X.; Chu, M. Whole genome sequencing reveals the effects of recent artificial selection on litter size of Bamei mutton sheep. *Animals* **2021**, *11*, 157. [CrossRef] [PubMed]
- 4. Wang, J.J.; Zhang, T.; Chen, Q.M.; Zhang, R.Q.; Li, L.; Cheng, S.F.; Shen, W.; Lei, C.Z. Genomic signatures of selection associated with litter size trait in Jining Gray goat. *Front. Genet.* **2020**, *11*, 286. [CrossRef] [PubMed]
- Demars, J.; Fabre, S.; Sarry, J.; Rossetti, R.; Gilbert, H.; Persani, L.; Tosser-Klopp, G.; Mulsant, P.; Nowak, Z.; Drobik, W. Genomewide association studies identify two novel *BMP15* mutations responsible for an atypical hyperprolificacy phenotype in sheep. *PLoS Genet.* 2013, 9, e1003482. [CrossRef]
- Lai, F.N.; Zhai, H.L.; Cheng, M.; Ma, J.Y.; Cheng, S.F.; Ge, W.; Zhang, G.L.; Wang, J.J.; Zhang, R.Q.; Wang, X. Whole-genome scanning for the litter size trait associated genes and SNPs under selection in dairy goat (*Capra hircus*). Sci. Rep. 2016, 6, 38096. [CrossRef]
- 7. Wang, K.; Liu, X.; Qi, T.; Hui, Y.; Yan, H.; Qu, L.; Lan, X.; Pan, C. Whole-genome sequencing to identify candidate genes for litter size and to uncover the variant function in goats (*Capra hircus*). *Genomics* **2021**, *113*, 142–150. [CrossRef]
- 8. Hua, G.H.; Chen, S.L.; Ai, J.T.; Yang, L.G. None of polymorphism of ovine fecundity major genes *FecB* and *FecX* was tested in goat. *Anim. Reprod. Sci.* 2008, 108, 279–286. [CrossRef]
- Mulsant, P.; Lecerf, F.; Fabre, S.; Schibler, L.; Monget, P.; Lanneluc, I.; Pisselet, C.; Riquet, J.; Monniaux, D.; Callebaut, I.; et al. Mutation in bone morphogenetic protein receptor-IB is associated with increased ovulation rate in Booroola Mérino ewes. *Proc. Natl. Acad. Sci. USA* 2001, *98*, 5104–5109. [CrossRef]
- 10. Alberto, F.J.; Boyer, F.; Orozco-terWengel, P.; Streeter, I.; Servin, B.; de Villemereuil, P.; Benjelloun, B.; Librado, P.; Biscarini, F.; Colli, L.; et al. Convergent genomic signatures of domestication in sheep and goats. *Nat. Commun.* **2018**, *9*, 813. [CrossRef]
- 11. Wu, D.-D.; Yang, C.-P.; Wang, M.-S.; Dong, K.-Z.; Yan, D.-W.; Hao, Z.-Q.; Fan, S.-Q.; Chu, S.-Z.; Shen, Q.-S.; Jiang, L.-P.; et al. Convergent genomic signatures of high-altitude adaptation among domestic mammals. *Nat. Sci. Rev.* 2019, 7, 952–963. [CrossRef]
- 12. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009, 25, 1754–1760. [CrossRef] [PubMed]
- Mckenna, A.; Hanna, M.G.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.M.; Garimella, K.; Altshuler, D.S.; Gabriel, S.; Daly, M. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010, 20, 1297–1303. [CrossRef] [PubMed]
- 14. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]
- 15. Danecek, P.; Auton, A.; Abecasis, G.R.; Albers, C.A.; Banks, E.; Depristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef] [PubMed]
- 16. Weir, B.S.; Cockerham, C.C. Estimating F-statistics for the analysis of population structure. *Evolution* **1984**, *38*, 1358–1370. [CrossRef]

- 17. Turner, S.D. qqman: An R package for visualizing GWAS results using QQ and manhattan plots. bioRxiv 2014. [CrossRef]
- 18. Wang, K.; Li, M.Y.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010, *38*, e164. [CrossRef]
- 19. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef]
- 20. Kovaka, S.; Zimin, A.V.; Pertea, G.M.; Razaghi, R.; Salzberg, S.L.; Pertea, M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* 2019, 20, 278. [CrossRef]
- 21. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef] [PubMed]
- Ferrer-Admetlla, A.; Liang, M.; Korneliussen, T.; Nielsen, R. On detecting incomplete soft or hard selective sweeps using haplotype structure. *Mol. Biol. Evol.* 2014, 31, 1275–1291. [CrossRef] [PubMed]
- 23. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A.H.; Tanaseichuk, O.; Benner, C.; Chanda, S.K. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **2019**, *10*, 1523. [CrossRef] [PubMed]
- Rochus, C.M.; Jonas, E.; Johansson, A.M. Population structure of five native sheep breeds of Sweden estimated with high density SNP genotypes. *BMC Genet.* 2020, 21, 27. [CrossRef] [PubMed]
- Cao, X.; Liu, W.-P.; Cheng, L.-G.; Li, H.-J.; Wu, H.; Liu, Y.-H.; Chen, C.; Xiao, X.; Li, M.; Wang, G.-D.; et al. Whole genome analyses reveal significant convergence in obsessive-compulsive disorder between humans and dogs. *Sci. Bull.* 2021, 66, 187–196. [CrossRef]
- Kuberan, B.; Lech, M.; Borjigin, J.; Rosenberg, R.D. Light-induced 3-O-sulfotransferase expression alters pineal heparan sulfate fine structure: A surprising link to circadian rhythm. *J. Biol. Chem.* 2004, 279, 5053–5054. [CrossRef] [PubMed]
- Martinez-Royo, A.; Alabart, J.L.; Sarto, P.; Serrano, M.; Lahoz, B.; Folch, J.; Calvo, J.H. Genome-wide association studies for reproductive seasonality traits in Rasa Aragonesa sheep breed. *Theriogenology* 2017, 99, 21–29. [CrossRef]
- 28. Coster, A.; Madsen, O.; Heuven, H.C.; Dibbits, B.; Groenen, M.A.; van Arendonk, J.A.; Bovenhuis, H. The imprinted gene DIO3 is a candidate gene for litter size in pigs. *PLoS ONE* **2012**, *7*, e31825. [CrossRef]
- Klüppel, M.; Vallis, K.A.; Wrana, J.L. A high-throughput induction gene trap approach defines C4ST as a target of BMP signaling. Mech. Dev. 2002, 118, 77–89. [CrossRef]
- Klüppel, M.; Wight, T.N.; Chan, C.; Hinek, A.; Wrana, J.L. Maintenance of chondroitin sulfation balance by chondroitin-4-sulfotransferase 1 is required for chondrocyte development and growth factor signaling during cartilage morphogenesis. *Development* 2005, 132, 3989–4003. [CrossRef]
- 31. Bhattacharyya, S.; Feferman, L.; Tobacman, J.K. Regulation of chondroitin-4-sulfotransferase (CHST11) expression by opposing effects of arylsulfatase B on BMP4 and Wnt9A. *Biochim. Biophys. Acta* 2015, *1849*, 342–352. [CrossRef] [PubMed]
- Oliveira-Ferrer, L.; Heßling, A.; Trillsch, F.; Mahner, S.; Milde-Langosch, K. Prognostic impact of chondroitin-4-sulfotransferase CHST11 in ovarian cancer. *Tumor Biol.* 2015, 36, 9023–9030. [CrossRef] [PubMed]
- 33. Farkas, S.A.; Sorbe, B.G.; Nilsson, T.K. Epigenetic changes as prognostic predictors in endometrial carcinomas. *Epigenetics* **2017**, 12, 19–26. [CrossRef]
- 34. Wu, P.; Wang, K.; Zhou, J.; Chen, D.; Yang, Q.; Yang, X.; Liu, Y.; Feng, B.; Jiang, A.; Shen, L.; et al. GWAS on imputed wholegenome resequencing from genotyping-by-sequencing data for farrowing interval of different parities in pigs. *Front. Genet.* **2019**, 10, 1012. [CrossRef]
- Forsyth, R.; Gunay-Aygun, M. Bardet-Biedl syndrome overview. In *GeneReviews*(®); Adam, M.P., Ardinger, H.H., Pagon, R.A., Wallace, S.E., Bean, L.J.H., Stephens, K., Amemiya, A., Eds.; University of Washington: Seattle, WA, USA, 1993.
- Weihbrecht, K.; Goar, W.A.; Carter, C.S.; Sheffield, V.C.; Seo, S. Genotypic and phenotypic characterization of the Sdccag8^{Tn(sb-Tyr)2161B.CA1C2Ove} mouse model. PLoS ONE 2018, 13, e0192755. [CrossRef] [PubMed]
- Airik, R.; Schueler, M.; Airik, M.; Cho, J.; Ulanowicz, K.A.; Porath, J.D.; Hurd, T.W.; Bekker-Jensen, S.; Schrøder, J.M.; Andersen, J.S.; et al. SDCCAG8 interacts with RAB effector proteins RABEP2 and ERC1 and is required for Hedgehog signaling. *PLoS ONE* 2016, *11*, e0156081. [CrossRef]
- Zhuang, Z.; Ding, R.; Peng, L.; Wu, J.; Ye, Y.; Zhou, S.; Wang, X.; Quan, J.; Zheng, E.; Cai, G.; et al. Genome-wide association analyses identify known and novel loci for teat number in Duroc pigs using single-locus and multi-locus models. *BMC Genom.* 2020, 21, 344. [CrossRef] [PubMed]
- 39. Wu, M.; Chen, G.; Li, Y.-P. TGF-β and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. *Bone Res.* **2016**, *4*, 16009. [CrossRef]
- Abdoli, R.; Zamani, P.; Mirhoseini, S.Z.; Ghavi Hossein-Zadeh, N.; Nadri, S. A review on prolificacy genes in sheep. *Reprod. Domest. Anim.* 2016, *51*, 631–637. [CrossRef] [PubMed]
- 41. Zhang, X.; Huang, L.; Wu, T.; Feng, Y.; Ding, Y.; Ye, P.; Yin, Z. Transcriptomic analysis of ovaries from pigs with high and low litter size. *PLoS ONE* **2015**, *10*, e0139514. [CrossRef] [PubMed]
- 42. Pokharel, K.; Peippo, J.; Honkatukia, M.; Seppala, A.; Rautiainen, J.; Ghanem, N.; Hamama, T.M.; Crowe, M.A.; Andersson, M.; Li, M.H.; et al. Integrated ovarian mRNA and miRNA transcriptome profiling characterizes the genetic basis of prolificacy traits in sheep (*Ovis aries*). *BMC Genom.* **2018**, *19*, 104. [CrossRef] [PubMed]
- 43. He, J.; Zhang, B.; Chu, M.; Wang, P.; Feng, T.; Cao, G.; Di, R.; Fang, L.; Huang, D.; Tang, Q. Polymorphism of insulin-like growth factor 1 gene and its association with litter size in Small Tail Han sheep. *Mol. Biol. Rep.* **2012**, *39*, 9801–9807. [CrossRef] [PubMed]

- 44. Thomas, N.; Venkatachalapathy, T.; Aravindakshan, T.; Raghavan, K.C. Molecular cloning, SNP detection and association analysis of 5' flanking region of the goat IGF1 gene with prolificacy. *Anim. Reprod. Sci.* **2016**, *167*, 8–15. [CrossRef] [PubMed]
- Wang, P.-Q.; Tan, Y.; Zhang, B.-Y.; Chu, M.-X.; Deng, L.-M.; Fan, Q.; Liu, C.-X. DNA polymorphisms of 5'-flanking region of insulin-like growth factor 1 gene and their association with reproduction traits in goats. *Agric. Sci. China* 2011, 10, 1609–1617. [CrossRef]
- 46. Hatzirodos, N.; Hummitzsch, K.; Irving-Rodgers, H.F.; Rodgers, R.J. Transcriptome profiling of the theca interna in transition from small to large antral ovarian follicles. *PLoS ONE* **2014**, *9*, e97489. [CrossRef]
- 47. Kwon, S.G.; Hwang, J.H.; Park, D.H.; Kim, T.W.; Kang, D.G.; Kang, K.H.; Kim, I.-S.; Park, H.C.; Na, C.-S.; Ha, J. Identification of differentially expressed genes associated with litter size in Berkshire pig placenta. *PLoS ONE* **2016**, *11*, e0153311. [CrossRef]
- 48. Oh, J.W.; Kim, S.K.; Cho, K.C.; Kim, M.S.; Suh, C.S.; Lee, J.R.; Kim, K.P. Proteomic analysis of human follicular fluid in poor ovarian responders during in vitro fertilization. *Proteomics* 2017, *17*, 1600333. [CrossRef]
- 49. Wang, Y.; Ding, X.; Tan, Z.; Ning, C.; Xing, K.; Yang, T.; Pan, Y.; Sun, D.; Wang, C. Genome-wide association study of piglet uniformity and farrowing interval. *Front. Genet.* **2017**, *8*, 194. [CrossRef]
- 50. Padua, M.B.; Hansen, P.J. Evolution and function of the uterine serpins (SERPINA14). *Am. J. Reprod. Immunol.* **2010**, *64*, 265–274. [CrossRef]
- Ulbrich, S.E.; Frohlich, T.; Schulke, K.; Englberger, E.; Waldschmitt, N.; Arnold, G.J.; Reichenbach, H.D.; Reichenbach, M.; Wolf, E.; Meyer, H.H.D.; et al. Evidence for estrogen-dependent uterine serpin (SERPINA14) expression during estrus in the bovine endometrial glandular epithelium and lumen. *Biol. Reprod.* 2009, *81*, 795–805. [CrossRef]
- 52. Rohit, K.; Ramteke, P.W.; Amar, N.; Singh, S.P. Role of candidate genes in regulation of embryonic survival and maternal recognition of pregnancy in farm animals. *Vet. World* 2013, *6*, 280–284. [CrossRef]
- 53. Jang, J.-S.; Lee, J.-H.; Jung, N.-C.; Choi, S.-Y.; Park, S.-Y.; Yoo, J.-Y.; Song, J.-Y.; Seo, H.G.; Lee, H.S.; Lim, D.-S. *Rsad2* is necessary for mouse dendritic cell maturation via the IRF7-mediated signaling pathway. *Cell Death Dis.* **2018**, *9*, 823. [CrossRef] [PubMed]
- 54. Cheng, L.; Xiang, M.; Hu, X.; Yu, J.; Xia, Y.; Tao, B.; Zhao, S. Duplex quantitative polymerase chain reaction of *ISG15* and *RSAD2* increases accuracy of early pregnancy diagnosis in dairy cows. *Ann. Anim. Sci.* **2019**, *19*, 383–401. [CrossRef]
- 55. Zhang, J.; Liu, Y.; Yao, W.; Li, Q.; Liu, H.; Pan, Z. Initiation of follicular atresia: Gene networks during early atresia in pig ovaries. *Reproduction* **2018**, 156, 23–33. [CrossRef]
- 56. Livera, G.; Xie, F.; Garcia, M.A.; Jaiswal, B.; Chen, J.; Law, E.; Storm, D.R.; Conti, M. Inactivation of the mouse adenylyl cyclase 3 gene disrupts male fertility and spermatozoon function. *Mol. Endocrinol.* **2005**, *19*, 1277–1290. [CrossRef]
- Zhang, Y.; Li, F.; Feng, X.; Yang, H.; Zhu, A.; Pang, J.; Han, L.; Zhang, T.; Yao, X.; Wang, F. Genome-wide analysis of DNA methylation profiles on sheep ovaries associated with prolificacy using whole-genome Bisulfite sequencing. *BMC Genom.* 2017, 18, 759. [CrossRef]
- 58. Sironen, A.I.; Uimari, P.; Serenius, T.; Mote, B.; Rothschild, M.; Vilkki, J. Effect of polymorphisms in candidate genes on reproduction traits in Finnish pig populations. *J. Anim. Sci.* **2010**, *88*, 821–827. [CrossRef] [PubMed]