Pakistan J. Zool., pp 1-10, 2023. **DOI**: https://dx.doi.org/10.17582/journal.pjz/20220118110134

Polymorphisms of *ESR1*, *PGR* and *CYP19A1* Genes and their Association with Litter Size in Small-Tailed Han Sheep

Qingqing Liu[1,2], Zhilong Tian[1], Chunhuan Ren[2], Zijun Zhang[2] and Mingxing Chu[1]*

[1]Key Laboratory of Animal Genetics and Breeding and Reproduction of Ministry of Agriculture and Rural Affairs, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China
[2]College of Animal Science and Technology, Anhui Agricultural University, Hefei 230036, China.

Article Information
Received 18 January 2022
Revised 05 April 2022
Accepted 23 April 2022
Available online 08 September 2023 (early access)

Authors' Contribution
MC, ZT and QL, who performed the experimental analyses. ZT and QL analyzed the data. QL drafted the paper. MC contributed to revisions of the paper. ZZ and CR conducted the background investigation. All authors read and approved the final paper for publication.

Key words
Sheep, *ESR1, PGR, CYP19A1*, litter size, Small-tailed Han sheep

ABSTRACT

Litter size is one of the most important economic traits in sheep. This study aimed to explore the association of polymorphisms of *ESR1* gene (g.75521224T > A, g.75378892A>T), *PGR* gene (g.4187121T>C) and *CYP19A1* gene (g.56122588G> A) with litter size in sheep, searching for new molecular markers for litter size trait in sheep. Sequenom Mass ARRAY® SNP technology was used to detect polymorphisms at these loci of *ESR1, PGR* and *CYP19A1* genes in seven sheep breeds including polytocous sheep (Small-tailed Han sheep (STH), Hu sheep and Cele Black sheep) and monotocous sheep (Tan sheep, Sunite sheep, Suffolk sheep and Tibetan sheep), then the association between the loci and litter size in STH sheep was analyzed. Our results revealed that each locus in most sheep breeds contained three genotypes. For genotyping, we found that genotype and allele frequencies of three loci (g.75521224T> A, g.41871219T> C, g.56122588G > A) of the three genes were significantly different ($p < 0.05$) between the monotocous and polytocous groups, and we found that the polymorphism information content value in all sheep breeds ranged from 0 to 0.37, and most sheep breeds were under Hardy-Weinberg equilibrium (HWE) ($p > 0.05$). In addition, the association analysis showed that there was a significant association between the g.75378892A > T polymorphism and the litter size of the first, second and third parties ($p \leq 0.05$). In conclusion, the g.75378892A > T locus of *ESR1* gene may provide a valuable reference to litter size trait selection in STH sheep and the g.75521224T > A locus of *ESR1*gene, g.41871219T > C of *PGR* gene, g.56122588G > A of *CYP19A1* were not suitable for breeding of STH sheep. We speculate that the misstense mutation of *ESR1* may reduce the ability to bind to *NCOA1,* resulting in fecundity in the Small-tailed Han sheep. This study provides valuable genetic markers for sheep breeding.

Online First Article

INTRODUCTION

Most of the sheep belong to seasonal estrus single lamb breed, only a few sheep belong to multiple lamb breed and its distribution is limited by region, the contribution of the number of lambs to the economic benefits of sheep breeding can reach 74% to 96% (Notter, 2012). Reproductive character is an important economic character of sheep. At present, there have been many research reports on improving sheep fecundity. Up to now, many candidate genes affecting lambing number of sheep have been studied deeply, include bone morphogenetic protein receptor 1B (*BMPR1B*, *FecB*), bone morphogenetic protein 15 (*BMP15*, *FecF*), growth differentiation factor-9, (*GDF9, FecG*), estrogen receptor (*ESR*), progesterone receptor (*PGR*). These genes play crucial roles in the regulation of lambing traits of sheep (Pramod *et al*., 2013).

Both *ESR1* and *PGR* belong to the steroid nuclear receptor superfamily (Osz *et al*., 2012). *ERα* in sheep is encoded by the gene *ESR1*, which has the function of regulating transcription proteins, and thus regulates the expression of target genes by binding to specific effector elements on target genes (Paterni *et al*., 2014). Adult female mutant mice lack *ERα* in neonatal neurons and cannot express estrous cycles or negative feedback (Cheong *et al*., 2014). In addition, a recent study showed that the *ESR1* variation revealed by SNP analysis is related to the egg-laying traits of quail (Wu *et al*., 2015) and some

* Corresponding author: mxchu@263.net
0030-9923/2023/0001-0001 $ 9.00/0

Copyright 2023 by the authors. Licensee Zoological Society of Pakistan.
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/).

studies have shown that *ESR1* is associates with in obesity, ovarian and breast disease in women (Guclu-Geyik *et al*., 2020; Turner *et al*., 2020; Tan *et al*., 2020). Progesterone receptor (*PGR*), also known as *NR3C3*, is a member of the steroid hormone nuclear receptor family (Arck *et al*., 2007; Tsai and O'malley, 1994). Previous studies have shown that *PGR* is highly expressed in ovary and uterus of Small-tailed Han (STH) sheep and can be used as a candidate gene for multiple lambs in sheep (Tian *et al*., 2018). In addition, *CYP19A1* many candidates first exons and nine different transcription start sites are known, which can carry out complex and tissue-specific regulation (Bouchoucha *et al*., 2014; Grumbach and Auchus, 1999). Aromatase have been described, which are products encoded by *CYP19A1* gene. In vertebrates, it can catalyze synthetases of different forms of sex hormones, which are also key enzymes in the process of rate limiting estrogen synthesis. The level of its activity and the high and low levels of estrogen in the synthesis reaction are widely distributed. It is involved in the production, processing, and transportation of protein in various tissues and cells of animals. It can not only act on the reproductive system, but also affect the occurrence of animal tumors (Balthazart and Ball., 1998; Ghosh *et al*., 2009). Therefore, the research of *CYP19A1* is crucial for sheep reproduction. At present, *CYP19A1* gene has been mostly reported in mice, chicken, human and other animals, but there are few studies on sheep breeding and reproduction (Zhang and Xu, 2021; Jin *et al*., 2020; Szaflik *et al*., 2020).

Small-tailed Han sheep (STH) is famous with the early maturity, good meat quality, genetic stability and perennial estrus (Jiao *et al*., 2020), so that, the research on litter size has been a hot topic. In this study, major genes or genetic markers related to litter size of sheep were found. To identify DNA markers for selecting elite sheep at an early stage through marker-assisted selection (MAS), Sequenom MassARRAY®SNP assay and detection of single nucleotide polymorphisms (SNPs) in four loci of *ESR1*, *PGR* and *CYP19A1* genes in seven sheep breeds by DNA sequencing. The association of identified SNPs with the litter size performance traits of something was explored and the correlation analysis of different sheep population and STH sheep litter size at this locus can provide the basis for sheep breeding.

MATERIALS AND METHODS

Ethics

All experimental procedures involving animals used in this study were approved by the Science Research Department (in charge of animal welfare issues) of the Institute of Animal Science, Chinese Academy of Agricultural Sciences (IAS-CAAS; Beijing, China). Ethical approval on animal survival was given by the animal ethics committee of IAS-CAAS (No. IAS2020-63, 28 April 2020).

Animals, sample collection and DNA extraction

Among this study, the seven sheep breeds were selected and divided into polytocous groups (Cele Black sheep; Hu sheep; and STH sheep) and monotocous groups (Prairie Tibetan sheep; Suffolk sheep; Sunite sheep; and Tan sheep) (Table I). Then, Jugular vein blood samples (10 mL blood per ewe) were collected using citrate glucose as an anticoagulant. Genomic DNA was extracted by the phenol-chloroform method, dissolved in ddH_2O and stored at -20℃. At the same time, the lambing order and lambing number of STH sheep were collected.

Genotyping

First, single-base extended primers for g.75521224T>A, g.75378892A> T in *ESR1*, g.7491179C>T in *PGR* and g.56122588G>A in *CYP19A1* were selected for genotyping in 768 samples from STH, Tan, Sunite, Suffolk, Hu, Cele Black and Prairie Tibetan sheep. Genotyping was performed using a MassARRAY®SNP analysis (Johansen *et al*., 2013). The polymerase chain reactions system and temperature were described in detail in previous study (Zhang *et al*., 2019). The DNA required quantity of each sample was 20μL, and the DNA concentration was 40-80 ng/μL.

Table I. Information of seven sheep breeds selected for genotyping.

Breed	Number	District
Small-tailed Han sheep	384	Yuncheng, Shandong Province, China
Hu sheep	83	Xuzhou, Jiangsu Province, China
Cele black sheep	68	Cele, Hetian, Xinjiang Uigur Autonomous Region, China
Prairie Tibetan sheep	80	Dangxiong, Tibet Autonomous Region, China
Sunite sheep	70	Wulatezhongqi, Bayannaoer, Inner Mongolia Autonomous Region, China
Tan sheep	23	Yanchi, Ningxia Hui Autonomous Region, China
Suffolk sheep	60	Shunyi, Beijing, China

Online First Article

Statistical analysis

Genotype frequencies, allele frequencies, heterozygosity (*He*), the number of effective alleles (*Ne*), p-values, and polymorphism information content (PIC) were calculated using the data obtained from genotyping results. Sheep populations with $p > 0.05$ (chi-square test) were considered to conform to the Hardy-Weinberg equilibrium (HWE). Multiple comparisons of means were performed using the least significant difference method. The applied model was expressed as follows: $y_{ijn} = \mu + P_i + G_j + IPG + e_{ijn}$, where y_{ijn} is the phenotypic value of litter size; μ is the population mean; P_i is the fixed effect of the ith parity (i = 1, 2, 3); G_j is the fixed effect of the jth genotype (j=1, 2, 3); IPG is the interaction effect of parity and genotype; and e_{ijn} is the random residual.

Protein structure prediction

The secondary structure of *ESR1* with and without missense mutants was analyzed using SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html). The STRING database (v.11.0) was applied to collect and integrate these functional interactions by consolidating protein–protein association data for massive organisms.

RESULTS

Polymorphisms of the ESR1, PGR *and* CYP19A1 *genes*

The genotyping results of the g.75521224T > A, g.75378892A > T, g.7491179C >T, and g.56122588G > A loci of *ESR1*, *PGR* and *CYP19A1* are shown in Figure 1. For example, for g.75521224T > A, there are three genotypes: wild homozygous TT, heterozygous AT, and mutant homozygous AA.

Population genetic analysis of SNPs in ESR1, PGR *and* CYP19A1

Four SNPs (g.75378892A >T, g.75521224T > A) in *ESR1,* (g.60495375A >G) in *PGR* and (g.56122588G>A) in *CYP19A1* were detected in monotocous and polytocous sheep breeds (Table II). The genotype frequency and allele frequency of g.75521224T > A locus were significantly different between polytocous and monotocous sheep ($p < 0.05$); and the genotype frequency and allele frequency of g.7491179C > T locus in *PGR* and g.56122588G > A locus in *CYP19A1* were extremely significantly different between polytocous and monotocous sheep ($p < 0.01$) (Table I).

The genotype frequency, allele frequency, *He*, *Ne,* PIC and $\chi 2$ test results from population genetic analysis for four SNPs in the seven sheep breeds are listed in Table III. The results showed that two (g.75378892A > T, g.75521224T > A) loci in *ESR1* were at moderate polymorphism (0.25 < PIC < 0.5) in STH, Suffolk, Sunite, Cele Black, Hu, Tan, and Prairie Tibetan sheep; g.7491179C > T in *PGR* locus had low polymorphism (PIC < 0.25) in seven sheep populations; g.56122588G > A locus in *CYP19A1* was at moderate polymorphism (0.25 < PIC < 0.5) in Suffolk, Cele Black, Hu, Tan, and Prairie Tibetan sheep except for STH sheep, Sunite (PIC < 0.25). In addition, the $\chi 2$ test revealed that g.75521224T > A was in Hardy-Weinberg equilibrium (HWE) ($p < 0.05$) except for STH sheep and Suffolk sheep; the g.75378892A > T was in HWE in the seven sheep breeds ($p > 0.05$); the g.7491179C > T was in HWE in the six sheep breeds ($p > 0.05$) except for Suffolk sheep; g.56122588G > A locus was under HWE ($p >0.05$) except for Hu sheep.

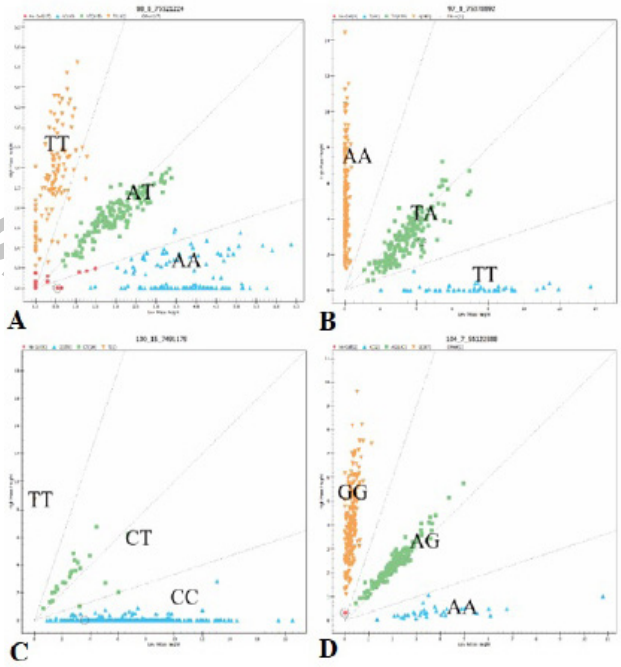

Fig. 1. Scatter plot of genotypes of *ESR1* (A, B), *PGR* (C) and *CYP19A1* (D).

Association between four loci in ESR1, PGR *and* CYP19A1 *with litter size in STH sheep*

The results revealed that the g.75378892A > T locus in *ESR1* was significantly correlated with litter size in STH sheep and that the litter size of ewes with AA genotype was higher than that of ewes with the AT and TT genotype ($p < 0.05$). However, g.75521224T > A, g.7491179C > T, g. 56122588G > A loci were not significantly correlated with litter size in STH sheep ($p >0.05$) (Table IV).

Online First

Table II. Genotype and allele frequencies of four loci in sheep with different litter size characteristics.

Gene	Locus	Characteristics of litter size	Genotype frequency			x^2 test (P value) value	Allele frequency		x^2 Test (P-value)
			AA	(P-value)	TT		A	T	
ESR1	g.75521224T>A	Polytocous sheep	0.31	0.40	0.29	0.04	0.51	0.49	0.01
		Monotocous sheep	0.40	0.38	0.22		0.59	0.41	
			AA	TA	TT		A	T	
ESR1	g.75378892A>T	Polytocous sheep	0.47	0.43	0.10	0.53	0.68	0.32	0.53
		Monotocous sheep	0.45	0.45	0.10		0.67	0.33	
			CC	CT	TT		C	T	
PGR	g.7491179C>T	Polytocous sheep	0.83	0.16	0.01	0.00	0.91	0.09	0.00
		Monotocous sheep	0.92	0.08	0.00		0.96	0.04	
			AA	AG	GG		A	G	
CYP19A1	g.56122588G>A	Polytocous sheep	0.04	0.26	0.70	0.00	0.19	0.81	0.00
		Monotocous sheep	0.10	0.44	0.46		0.32	0.68	

$p < 0.05$ denotes significant differences, and $p < 0.01$ denotes extremely significant difference 111.

First Article

Bioinformatic analysis of ESR1, PGR *and* CYP19A1

Bioinformatics analysis showed that g.75378892A > T was A missense mutation, leading to the mutation of isoleucine at 36 to phenylalanine, which may cause some changes in the function of this secreted factor in follicles. The results were identified of alpha helix, extended strand, random coil, beta turn was changed, and many proteins interact (Figs. 2 and 3).

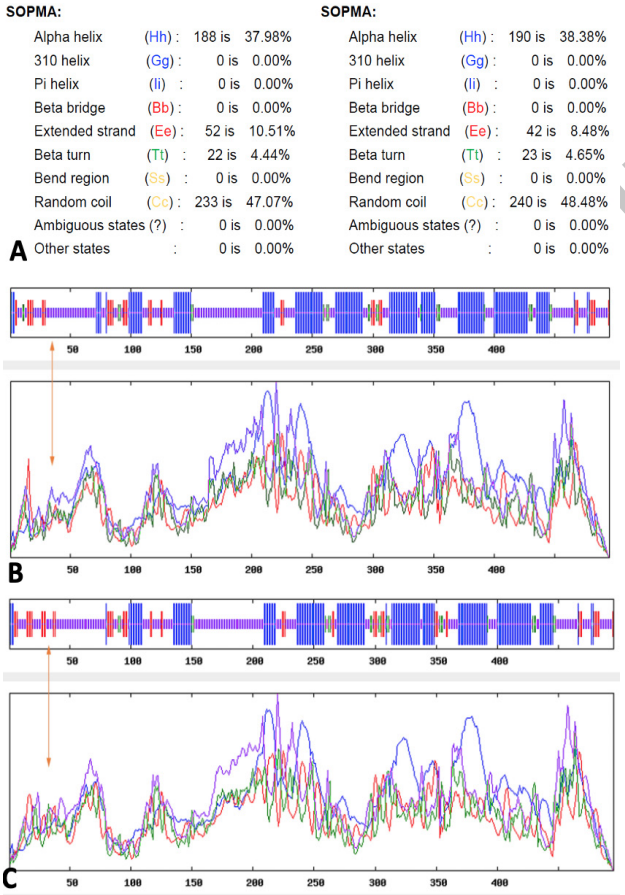

Fig. 2. Secondary protein structure for the *ESR1* product before (A, C) and after (B, D) the mutation at g.75378892A>T based on its amino acid sequence.

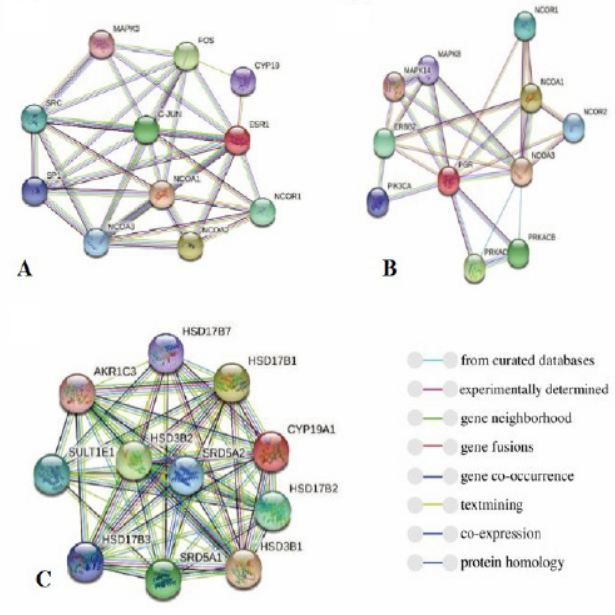

Fig. 3. The prediction of the protein interaction networks of *ESR1* (A), *PGR* (B), and *CYP19A1* (C), based on the STRING database.

Table III. Population genetic analysis of four loci in seven sheep breeds.

Gene	Locus	Breed	Genotype frequency			Allele frequency		PIC	He	Ne	x^2 Test (P-value)
ESR1	g.75521224T>A		AA	AT	TT	A	T				
		Small-tailed Han sheep	0.37	0.39	0.24	0.57	0.43	0.37	0.49	0.97	0.00
		Suffolk sheep	0.71	0.21	0.08	0.81	0.19	0.26	0.30	1.43	0.04
		Sunite sheep	0.13	0.43	0.44	0.35	0.65	0.35	0.45	1.82	0.64
		Cele Black sheep	0.18	0.42	0.4	0.39	0.61	0.36	0.48	1.91	0.36
		Hu sheep	0.13	0.41	0.46	0.33	0.67	0.35	0.45	1.8	0.54
		Tan sheep	0.19	0.48	0.33	0.43	0.57	0.37	0.49	1.96	0.90
		Prairie Tibetan sheep	0.49	0.44	0.07	0.71	0.29	0.33	0.41	1.70	0.52
ESR1	g.75378892A>T		AA	TA	TT	A	T				
		Small-tailed Han sheep	0.48	0.42	0.10	0.69	0.31	0.34	0.41	1.75	0.46
		Suffolk sheep	0.50	0.40	0.10	0.70	0.30	0.33	0.42	1.72	0.71
		Sunite sheep	0.43	0.49	0.08	0.68	0.34	0.34	0.44	1.78	0.40
		Cele Black sheep	0.49	0.46	0.05	0.72	0.28	0.32	0.4	1.68	0.35
		Hu sheep	0.41	0.48	0.11	0.65	0.35	0.35	0.46	1.83	0.66
		Tan sheep	0.41	0.41	0.18	0.61	0.39	0.36	0.47	1.90	0.52
		Prairie Tibetan sheep	0.44	0.46	0.10	0.67	0.33	0.34	0.44	1.79	0.69
PGR	g.7491179C>T		CC	CT	TT	C	T				
		Small-tailed Han sheep	0.78	0.20	0.02	0.88	0.12	0.19	0.21	1.27	0.94
		Suffolk sheep	0.95	0.03	0.02	0.97	0.03	0.07	0.07	1.07	0.00
		Sunite sheep	0.84	0.16	000	0.92	0.08	0.14	0.15	1.17	0.48
		Cele Black sheep	0.97	0.03	0.00	0.99	0.01	0.03	0.03	1.03	0.90
		Hu sheep	0.88	0.12	0.00	0.94	0.06	0.11	0.11	1.13	0.71
		Tan sheep	0.91	0.09	0.00	0.96	0.04	0.08	0.09	1.09	0.83
		Prairie Tibetan sheep	0.96	0.04	0.00	0.98	0.02	0.04	0.04	1.04	0.86
YP19A1	g.56122588G>A		AA	AG	GG	A	G				
		Small-tailed Han sheep	0.04	0.25	0.71	0.17	0.83	0.24	0.28	1.38	0.14
		Suffolk sheep	0.15	0.42	0.43	0.36	0.64	0.35	0.46	1.85	0.47
		Sunite sheep	0.04	0.26	0.70	0.17	0.83	0.24	0.28	1.39	0.44
		Cele Black sheep	0.12	0.37	0.51	0.31	0.69	0.33	0.42	1.74	0.32
		Hu sheep	0.19	0.81	0.00	0.60	0.40	0.37	0.48	1.93	0.00
		Tan sheep	0.22	0.61	0.17	0.53	0.47	0.37	0.50	2.00	0.29
		Prairie Tibetan sheep	0.09	0.56	0.35	0.37	0.63	0.36	0.47	0.87	0.06

Abbreviations: *He*, heterozygosity; PIC, polymorphic information content; *Ne*, number of effective alleles; $p > 0.05$ indicates the locus was under Hardy-Weinberg equilibrium.

Online First Article

DISCUSSION

Relationship between ESR1 *polymorphism and reproductive performance of animals*

Combined with estrogen, *ESRα* plays an important role in the growth and development of embryonic mammary glands and female follicles during the reproductive cycle (Li *et al*., 2015). Knockout of *ESR1* gene would lead to disorders of LH hormone regulation and ovulation failure of ovaries in mice (Hewitt and Korach, 2003). In addition, many studies have confirmed that *ESR* gene is one of the major genes regulating high litter size in pigs (Findlay *et al*., 2001). Wang *et al*. (2015) used genome-wide association analysis to find that the T669C mutation of *ESR1* gene

Table IV. Analysis of different loci and litter size in STH sheep.

Gene	Locus	Genotype	1st parity litter size(N)	2nd parity litter size(N)	3rd parity litter size(N)
ESR1	g.75521224T>A	AA	1.86±0.05(139)	2.22±0.09(79)	2.38±0.16(29)
		AT	1.88±0.06(147)	2.18±0.08(87)	2.37±0.14(38)
		TT	1.87±0.07(87)	2.04±0.10(56)	2.31±0.18(21)
ESR1	g.75378892A>T	AA	1.93±0.05(184) a	2.15±0.07(106) a	2.46±0.13(41) a
		TA	1.86±0.05(162) ab	2.14±0.08(99) ab	2.33±0.28(40) ab
		TT	1.68±0.10(38) b	1.83±0.16(23) b	2.08±0.17(9) b
PGR	g.7491179C>T	CC	1.85±0.04(285)	2.09±0.06(160)	2.53±0.11(61)
		CT	1.95±0.08(74)	2.12±0.10(57)	2.32±0.17(25)
		TT	2.20±0.30(5)	2.33±0.44(4)	2.38±0.41(3)
CYP19A1	g.56122588G>A	AA	1.71±0.25(7)	2.00±0.34(5)	2.09±0.59(4)
		AG	1.91±0.06(114)	2.14±0.09(69)	2.24±0.17(25)
		GG	1.85±0.04(261)	2.11±0.06(153)	2.18±0.10(64)

Note that different small letters in the same group mean significant difference ($p < 0.05$).

was closely related to the litter size of pigs and found that *ESR1* gene was involved in the process of pig domestication. Studies of STH sheep and pigs have shown that *ESR* gene is closely related to lamb multiplication (Bi *et al*., 2005; Rahman *et al*., 2021). That is, more and more evidences indicate that mutations in the *ESR1* gene locus are related to animal reproduction. We found that the *ESR1* gene g.75378892A > T and g.75521224T > A loci were moderately polymorphic in seven sheep populations (0.25 < PIC < 0.5). Except that g.75521224T > A locus were not HWE in STH sheep and Suffolk sheep, the remaining SNPs showed HWE in different sheep breeds. This could be due to an imbalance in the site due to natural selection or artificial intervention, or it could be due to the small number of sheep involved in this study. Association analysis showed that the polymorphism of *ESR1* in g.75378892A > T was significantly associated with different litter size of STH sheep ($p < 0.05$). Litter size of AA type ewes was significantly higher than that of TT type ewes ($p < 0.05$). Therefore, in the process of sheep breeding, AA homozygous individuals can be selected and retained to improve the fertility of sheep. In addition, bioinformatics analysis showed mutations in *ESR1* in g.75378892A>T sites, resulting in amino acid changes, which we speculate is the main reason for the decline in lambing in STH sheep.

Relationship between PGR *polymorphism and reproductive performance of animals*

Studies have shown that *PGR* are low expressed in all major physiological systems of the animal body, but high expressed in the central nervous system and female reproductive tract (Pelch *et al*., 2011). Song (2017) obtained the polymorphism information of *PGR* gene through DNA mixed pool sequencing and flight mass spectrometry typing. According to the correlation analysis of *PGR* with Rex rabbit reproductive performance, it was found that the three SNPs of *PGR* were significantly correlated with the live litter size, weaning number and litter weight at 21 days of age. There is a significant association between the *PGR* rs660149 G variant of the Malays and the susceptibility to preterm birth (Langmia *et al*., 2015). The expression of *PGR* related genes affects piglet size, such as *IHH, NR2F2* and *BMP2* (Chen *et al*., 2016). In this study, it was found that the *PGR* gene g.7491179C > T was low polymorphism (PIC < 0.25) in all seven sheep breeds and had no significant correlation with litter size of STH sheep. Previous studies have shown that mutations in the *PGR* gene are associated with mammalian reproduction, which is inconsistent with this study and may be a species difference.

Relationship between CYP19A1 *polymorphism and reproductive performance of animals*

The aromatase encoded by the *CYP19A1* gene is an enzyme for estrogen synthesis, a monooxygenase, which can catalyze many reactions related to steroids and catalyze the conversion of androgens to estrogen (Bershteĭn, 1997). Cytochrome P450 aromatase is involved in the development of follicles at different stages and has a regulatory role (Bao and Garverick, 1998) and plays an important role in the development of sheep gonadal (Kwon *et al*., 2001). The development of testicular cells (Hess and Roser, 2004) is also very important. Previous studies on the *CYP19A1* gene have mainly focused on human cancer-related diseases (Wang *et al*., 2021; Akçurin *et al*., 2016) and the reproductive regulation of other species, and there has been little research in the field of

Online First Article

sheep reproduction. Moreover, Zhang and Xu (2021) used Yulin mixture to increase the expression of *CYP19A1* in mice, thereby regulating the proliferation and apoptosis of granulosa cells. Taking the STH sheep population with different fecundity as the research object, it is found that the *CYP19A1* gene is mainly expressed in the ovary and hypothalamus tissues. The expression level of the *CYP19A1* gene in the ovary tissues of the multi-lamb STH sheep is higher than that of the single-lamb and Sunite sheep ($p < 0.05$) (Tian *et al*., 2019). Vega *et al*. (2018) found that the rs718446508T > C and rs41651668T > C sites of the *CYP19A1* gene were significantly related to bovine oocyte production, survival, embryonic development, pregnancy and other reproductive traits ($p < 0.05$). El-Bayomi *et al*. (2018) found that the *CYP19A1* gene c.135T > C, c.559G>A, c.1285C > T, c.1394A > G sites were significantly related to the buffalo interestrous period ($p < 0.05$). The results of the correlation analysis showed that there was no significant correlation between the different genotypes of *CYP19A1* gene g.56122588G>A and the litter size in different parities of STH sheep ($p > 0.05$). It may be because natural selection or artificial selection has a greater impact on the distribution of the site, or it may be related to the number of sheep breeds we have selected. In addition, it may be that the site is in the 5'UTR region and does not directly participate in the encoding of the protein, thereby affecting its function.

Protein secondary structure prediction and interaction

Interestingly, through the construction of protein network interaction maps, interactions between the *ESR1*, *PGR* and *NCOA1* genes were found. The *NCOA1* gene, known as the steroid receptor co-activator protein 1 (*SRC1*) gene, is an estrogen receptor, progesterone receptor assistive activator (Oñate *et al*., 1995) and is considered a "switch" of reproductive hormones (Seitz et al., 2008) Breeding in chickens (Gholami *et al*., 2014) and sheep (Yuan *et al*., 2019; Xu *et al*., 2018) has been reported in many studies and has achieved more desirable results. However, the results of this study showed that the mutation of A>T site in g.75378892A resulted in the decrease of litter size of small-tailed Han sheep. It could be speculated that the mutation of this site reduced or blocked the recognition of *ESR1* and *NCOA1* genes, thus affecting the reproductive rate. In addition, the *PGR* site is synonymous mutation, which does not cause the secondary and tertiary structure mutation of the protein, so it cannot interact with *NCOA1*, so it has no significant influence on lambing of small-tailed Han sheep. *CYP19A1* also interacts with many proteins, and whether it affects reproduction needs to be further explored.

CONCLUSION

In this study, we first performed a population genetic analysis of *ESR1*, *PGR* and *CYP19A1* SNPs. Through association analysis, we found a key gene locus, namely g.75378892A > T in *ESR1,* this locus could influence each parity litter size in Small-tailed Han sheep ($p < 0.05$). Therefore, our studies could be useful in marker-assisted selection of the litter size in Small-tailed Han sheep.

ACKNPWLEDGEMENTS

This research was funded by National Natural Science Foundation of China (32172704), China Agriculture Research System of MOF and MARA (CARS-38), Agricultural Science and Technology Innovation Program of China (CAAS-ZDRW202106 and ASTIP-IAS13).

Statement of conflict of interest

The authors have declared no conflict of interest.

Online First Article

REFERENCES

Akçurin, S., Türkkahraman, D., Kim, W.Y., Durmaz, E., Shin, J.G., and Lee, S.J., 2016. A novel null mutation in P450 Aromatase Gene (CYP19A1) associated with development of hypoplastic ovaries in humans. *J. Clin. Res. Pediatr. Endocrinol*., **8:** 205-210. https://doi.org/10.4274/jcrpe.2761

Arck, P., Hansen, P.J., Mulac, J.B., Piccinni, M.P., and Szekeres-Bartho, J., 2007. Progesterone during pregnancy: endocrine-immune cross talk in mammalian species and the role of stress. *Am. J. Reprod. Immunol*., **58:** 268-279. https://doi.org/10.1111/j.1600-0897.2007.00512.x

Balthazart, J. and Ball, G.F., 1998. New insights into the regulation and function of brain estrogen synthase (aromatase). *Trends Neurosci*., **21:** 243-249. https://doi.org/10.1016/S0166-2236(97)01221-6

Bao, B., and Garverick, H.A., 1998. Expression of steroidogenic enzyme and gonadotropin receptor genes in bovine follicles during ovarian follicular waves: A review. *J. Anim. Sci*., **76:** 1903-1921. https://doi.org/10.2527/1998.7671903x

Bershteĭn, L.M., 1997. Molecular genetic aspects of estrogen production. Aromatase gene. *Mol. Biol. Mosk*., **31:** 773-777. https://doi.org/10.1016/S0166-6851(97)00082-0

Bi, X.D., Chu, M.X., Jin, H.G., Fang, L., and Ye, S.C., 2005. Estrogen receptor as a candidate gene for prolificacy of small tail Han sheep. *Hereditas*, **32:** 1060-1065. 2005.

Bouchoucha, N., Samara-Boustani, D., Pandey, A.V., Bony-Trifunovic, H., Hofer, G., Aigrain, Y., Polak, M., and Flück, C.E., 2014. Characterization of a novel CYP19A1 (aromatase) R192H mutation causing virilization of a 46, X newborn, undervirilization of the 46, Y brother, but no virilization of the mother during pregnancies. *Mol. Cell Endocrinol*., **390:** 8-17. https://doi.org/10.1016/j.mce.2014.03.008

Chen, X., Fu, J., and Wang, A., 2016. Expression of genes involved in progesterone receptor paracrine signaling and their effect on litter size in pigs. *J. Anim. Sci. Biotechnol*., **7:** 31. https://doi.org/10.1186/s40104-016-0090-z

Cheong, R.Y., Porteous, R., Chambon, P., Abrahám, I., and Herbison, A.E., 2014. Effects of neuron-specific estrogen receptor (ER) α and ERβ deletion on the acute estrogen negative feedback mechanism in adult female mice. *Endocrinology*, **155:** 1418-1427. https://doi.org/10.1210/en.2013-1943

El-Bayomi, K.M., Saleh, A.A., Awad, A., El-Tarabany, M.S., El-Qaliouby, H.S., Afifi, M., El-Komy, S., Essawi, W.M., Almadaly, E.A., and El-Magd, M.A., 2018. Association of CYP19A1 gene polymorphisms with anoestrus in water buffaloes. *Reprod. Fertil. Dev*., **30:** 487-497. https://doi.org/10.1071/RD16528

Findlay, J.K., Britt, K., Kerr, J.B., O'Donnell, L., Jones, M.E., Drummond, A.E., and Simpson, E.R., 2001. The road to ovulation: The role of oestrogens. *Reprod. Fertil. Dev*., **13:** 543-547. https://doi.org/10.1071/RD01071

Ghosh, D., Griswold, J., Erman, M., and Pangborn, W., 2009. Structural basis for androgen specificity and oestrogen synthesis in human aromatase. *Nature*, **457:** 219-223. https://doi.org/10.1038/nature07614

Grumbach, M.M., and Auchus, R.J., 1999. Estrogen: consequences and implications of human mutations in synthesis and action. *J. clin. Endocrinol. Metab*., **84:** 4677-4694. https://doi.org/10.1210/jcem.84.12.6290

Guclu-Geyik, F., Coban, N., Can, G., and Erginel-Unaltuna, N., 2020. The rs2175898 polymorphism in the ESR1 gene has a significant Sex-Specific effect on obesity. *Biochem. Genet*., **58:** 935-952. https://doi.org/10.1007/s10528-020-09987-6

Gholami, M., Erbe, M., Gärke, C., Preisinger, R., Weigend, A., Weigend, S., and Simianer, H., 2014. Population genomic analyses based on 1 million SNPs in commercial egg layers. *PLoS One*, **4:** e94509. https://doi.org/10.1371/journal.pone.0094509

Hess, M.F., and Roser, J.F., 2004. Immunocytochemical localization of cytochrome P450 aromatase in the testis of prepubertal, pubertal, and postpubertal horses. *Theriogenology*, **61:** 293-299. https://doi.org/10.1016/S0093-691X(03)00237-1

Hewitt, S.C., and Korach, K.S., 2003. Oestrogen receptor knockout mice: Roles for oestrogen receptors alpha and beta in reproductive tissues. *Reproduction*, **125:** 143-149. https://doi.org/10.1530/rep.0.1250143

Jiao, J., Wang, T., Zhou, J., Degen, A.A., Gou, N., Li, S., Bai, Y., Jing, X., Wang, W., and Shang, Z., 2020. Carcass parameters and meat quality of Tibetan sheep and Small-tailed Han sheep consuming diets of low-protein content and different energy yields. *J. Anim. Physiol. Anim. Nutr. (Berl.)*, **104:** 1010-1023. https://doi.org/10.1111/jpn.13298

Jin, K., Zuo, Q., Song, J., Zhang, Y., Chen, G., and Li, B., 2020. CYP19A1 (aromatase) dominates female gonadal differentiation in chicken (Gallus gallus) embryos sexual differentiation, *Biosci. Rep*., **40:** 10. https://doi.org/10.1042/BSR20201576

Johansen, P., Andersen, J.D., Børsting, C., and Morling, N., 2013. Evaluation of the iPLEX® Sample ID Plus Panel designed for the Sequenom MassARRAY® system. A SNP typing assay developed for human identification and sample tracking based on the SNPforID panel. *Forensic Sci. Int. Genet*., **7:** 482-487. https://doi.org/10.1016/j.fsigen.2013.04.009

Kwon, J.Y., McAndrew, B.J., and Penman, D.J., 2001. Cloning of brain aromatase gene and expression of brain and ovarian aromatase genes during sexual differentiation in genetic male and female Nile tilapia oreochromis niloticus. *Mol. Reprod. Dev*., **59:** 359-370. https://doi.org/10.1002/mrd.1042

Langmia, I.M., Apalasamy, Y.D., Omar, S.Z., and Mohamed, Z., 2015. Progesterone Receptor (PGR) gene polymorphism is associated with susceptibility to preterm birth. *BMC Med. Genet*., **16:** 63. https://doi.org/10.1186/s12881-015-0202-1

Li, X., Li, H., Jia, L., Li, X., and Rahman, N., 2015. Oestrogen action and male fertility: experimental and clinical findings. *Cell Mol. Life Sci*., **72:** 3915-3930. https://doi.org/10.1007/s00018-015-1981-4

Notter, D.R., 2012. Genetic improvement of reproductive efficiency of sheep and goats. *Anim. Reprod. Sci*., **130:** 147-151. https://doi.org/10.1016/j.anireprosci.2012.01.008

Osz, J., Brélivet, Y., Peluso-Iltis, C., Cura, V., Eiler, S., Ruff, M., Bourguet, W., Rochel, N., and Moras, D., 2012. Structural basis for a molecular allosteric control mechanism of cofactor binding to nuclear receptors. *Proc. natl. Acad. Sci. U.S.A*., **109:** E588-

Online First Article

594. https://doi.org/10.1073/pnas.1118192109

Oñate S.A., Tsai, S.Y., Tsai, M.J., and O'Malley, B.W., 1995. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science*, **270:** 1354-1357. https://doi.org/10.1126/science.270.5240.1354

Paterni, I., Granchi, C., Katzenellenbogen, J.A., and Minutolo, F., 2014. Estrogen receptors alpha (ERα) and beta (ERβ): subtype-selective ligands and clinical potential. *Steroids*, **90:** 13-29. https://doi.org/10.1016/j.steroids.2014.06.012

Pelch, K.E., Beeman, J.M., Niebruegge, B.A., Winkeler, S.R., and Nagel, S.C., 2011. Chapter 14- Endocrine-disrupting chemicals (EDCs) in mammals. In: *Hormones and reproduction of vertebrates*, (eds. D.O. Norris and K.H. Lopez). Academic Press, London, pp. 329-371. https://doi.org/10.1016/B978-0-12-374932-1.00061-5

Pramod, R.K., Sharma, S.K., Kumar, R., and Rajan, A.J.V.W., 2013. Genetics of ovulation rate in farm animals. *Vet. World*, **6:** 833-838. https://doi.org/10.14202/vetworld.2013.833-838

Rahman, M., Phookan, A., Zaman, G.U., Das, A., Akhtar, F., Tamuly, S., Choudhury, H., and Sarma, L.M., 2021. Allelic variability of estrogen receptor (ESR) gene and its effect on litter traits of Doom pigs. *Trop. Anim. Hlth. Prod*., **53:** 316. https://doi.org/10.1007/s11250-021-02756-6

Song, H.Z., 2017. *Polymorphism and association analysis of breeding related genes in Rex rabbits* (in Chinese). Shandong Agricultural University.

Szaflik, T., Smolarz, B., Romanowicz, H., Bryś, M., Forma, E., and Szyłło, K., 2020. Polymorphisms in the 3'UTR region of ESR2 and CYP19A1 genes in women with endometriosis. *Eur. J. Obstet. Gynecol. Reprod. Biol*., **250:** 241-245. https://doi.org/10.1016/j.ejogrb.2020.05.023

Seitz, M., Maillard, L.T., Obrecht, D., and Robinson, J.A., 2008. Molecular characterization of the NCoA-1-STAT 6 interaction. *Chembiochem*, **9:** 1318-1322. https://doi.org/10.1002/cbic.200700773

Tan, G.C., Chu, C., Lee, Y.T., Tan, C.C.K., Ashburner, J., Wood, N.W., and Frackowiak, R.S., 2020. The influence of microsatellite polymorphisms in sex steroid receptor genes ESR1, ESR2 and AR on sex differences in brain structure. *Neuro Image*, **221**: 117087. https://doi.org/10.1016/j.neuroimage.2020.117087

Tian, Z.L., Liu, Q.Y., and Wang, X.Y., 2019. Expression of CYP11A1 CYP17A1 CYP19A1 gene in hypothalamic pituitary ovarian axis of small tail Han sheep (in Chinese). *J. China Agric. Univ*., **24:** 87-93.

Tian, Z.L., Liu, Q.Y., Wang, X.Y., Di, R., Hu, W.P., Wang, W.P., Zhang, X.S., Zhang, J.L., and Chu, M.X., 2018. Expression of *ESR1* and *PGR* genes in hypothalamus-pituitary-ovary axis in Small Tail Han Sheep with different fertility (in Chinese). *Chinese J. Anim. Sci*., **54:** 36-40.

Tsai, M.J., and O'Malley, B.W., 1994. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu. Rev. Biochem*., **63:** 451-486. https://doi.org/10.1146/annurev.bi.63.070194.002315

Turner, N.C., Swift, C., Kilburn, L., Fribbens, C., Beaney, M., Garcia-Murillas, I., Budzar, A.U., Robertson, J.F.R., Gradishar, W., Piccart, M., Schiavon, G., Bliss, J.M., Dowsett, M., Johnston, S.R.D., and Chia, S.K., 2020. ESR1 mutations and overall survival on fulvestrant versus exemestane in advanced hormone Receptor-Positive breast cancer: A combined analysis of the phase III SoFEA and EFECT trials. *Clin. Cancer Res*., **26:** 5172-5177. https://doi.org/10.1158/1078-0432.CCR-20-0224

Vega, W.H.O., Quirino, C.R., Bartholazzi-Junior, A., Rua, M.A.S., Serapião, R.V., and Oliveira, C.S., 2018. Variants in the CYP19A1 gene can affect in vitro embryo production traits in cattle. *J. Assist. Reprod. Genet*., **35:** 2233-2241. https://doi.org/10.1007/s10815-018-1320-4

Wang, C., Wang, H., Zhang, Y., Tang, Z., Li, K., and Liu, B., 2015. Genome-wide analysis reveals artificial selection on coat colour and reproductive traits in Chinese domestic pigs. *Mol. Ecol. Resour*., **15:** 414-424. https://doi.org/10.1111/1755-0998.12311

Wang, T., Sun, Y., Xiong, Z., Wu, J., Ding, X., Guo, X., and Shao, Y., 2021. Association of ST6GAL1 and CYP19A1 polymorphisms in the 3'-UTR with astrocytoma risk and prognosis in a Chinese Han population. *BMC Cancer*, **21:** 391. https://doi.org/10.1186/s12885-021-08110-1

Wu, Y., Pan, A.L., Pi, J.S., Pu, Y.J., Du, J.P., Liang, Z.H., Shen, J., 2015. SNP analysis reveals estrogen receptor 1 (ESR1) gene variants associated with laying traits in quails. *Arch. Anim. Breed*., **58:** 441-444. https://doi.org/10.5194/aab-58-441-2015

Yuan, Z., Zhang, J., Li, W., Wang, W., Li, F., and Yue, X., 2019. Association of polymorphisms in candidate genes with the litter size in two sheep breeds. *Animals*, **9:** 958. https://doi.org/10.3390/ani9110958

Xu, S.S., Gao, L., Xie, X.L., Ren, Y.L., Shen, Z.Q., Wang, F., Shen, M., Eyþórsdóttir, E., Hallsson,

Online First Article

J.H., Kiseleva, T., Kantanen, J., and Li, M.H., 2018. Genome wide association analyses highlight the potential for different genetic mechanisms for litter size among sheep breeds. *Front. Genet*., **9:** 118. https://doi.org/10.3389/fgene.2018.00118

Zhang, Y., and Xu, L., 2021. Preliminary study of Yulin mixture affecting the miR-320/SF-1/Cyp19a1 on mouse polycystic ovary syndrome model. *Gynecol. Endocrinol*., **37:** 546-553. https://doi.org/10.1080/09513590.2020.1843623

Zhang, Z., Liu, Q., Di, R., Hu, W., Wang, X., He, X., Ma, L., and Chu, M., 2019. Single nucleotide polymorphisms in BMP2 and BMP7 and the association with litter size in Small Tail Han sheep. *Anim. Reprod. Sci*., **204:** 183-192. https://doi.org/10.1016/j.anireprosci.2019.04.001

Online First Article