

Scotland's Rural College

Novel genomic markers and genes related to reproduction in prolific Chios dairy sheep

Tsartsianidou, V.; Pavlidis, A.; Tosiou, E.; Arsenos, G.; Banos, G.; Triantafyllidis, A.

Published in:
Animal

DOI:
[10.1016/j.animal.2023.100723](https://doi.org/10.1016/j.animal.2023.100723)

Print publication: 01/03/2023

Document Version
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Tsartsianidou, V., Pavlidis, A., Tosiou, E., Arsenos, G., Banos, G., & Triantafyllidis, A. (2023). Novel genomic markers and genes related to reproduction in prolific Chios dairy sheep: a genome-wide association study. *Animal*, 17(3), Article 100723. <https://doi.org/10.1016/j.animal.2023.100723>

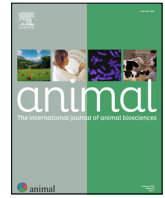
General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Novel genomic markers and genes related to reproduction in prolific Chios dairy sheep: a genome-wide association study



V. Tsartsianidou^{a,b,*}, A. Pavlidis^a, E. Tosiou^a, G. Arsenos^c, G. Banos^d, A. Triantafyllidis^{a,b}

^a Department of Genetics, Development & Molecular Biology, School of Biology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

^b Genomics and Epigenomics Translational Research (GENeTres), Center for Interdisciplinary Research and Innovation (CIRI-AUTH), Balkan Center, 57001 Thessaloniki, Greece

^c Laboratory of Animal Husbandry, School of Veterinary Medicine, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

^d Scotland's Rural College (SRUC), Easter Bush, Midlothian EH25 9RG, UK

ARTICLE INFO

Article history:

Received 1 November 2022

Revised 19 January 2023

Accepted 20 January 2023

Available online 27 January 2023

Keywords:

Association analysis

Candidate genes

Reproductive traits

Single-nucleotide polymorphisms

Small ruminants

ABSTRACT

Genetic architecture of sheep reproduction is increasingly gaining scientific interest due to the major impact on sheep production systems. In the present study, we conducted pedigree-based analyses and genome-wide association studies using the Illumina Ovine SNP50K BeadChip to explore the genetic mechanisms underlying the reproduction of the highly prolific Chios dairy sheep. First lambing age, total prolificacy and maternal lamb survival were selected as representative reproductive traits and estimated as significantly heritable ($h^2 = 0.07-0.21$) with no evident genetic antagonism among traits. We identified novel genome-wide and suggestive significant single-nucleotide polymorphisms (SNPs) on chromosomes 2 and 12 associated with age at first lambing. The new variants detected on chromosome 2 span a region of 357.79 kb with high pairwise linkage disequilibrium estimates ($r^2 = 0.8-0.9$). Functional annotation analysis revealed candidate genes, such as the collagen-type genes and the Myostatin gene, that participate in osteogenesis, myogenesis, skeletal and muscle mass development resembling the functionality of major genes affecting the ovulation rate and prolificacy. Additional functional enrichment analysis associated the collagen-type genes with multiple uterine-related disfunctions, such as cervical insufficiency, uterine prolapse and abnormalities of the uterine cervix. Several genes (e.g., *KAZN*, *PRDM2*, *PDPN*, *LRRC28*) localised close to the SNP marker on chromosome 12 were grouped in annotation enrichment clusters majorly involved in developmental and biosynthetic pathways, apoptosis, and nucleic acid-templated transcription. Our findings may further contribute to unravel the genomic regions that are important for sheep reproduction and could be incorporated into future selective breeding programmes.

© 2023 The Author(s). Published by Elsevier B.V. on behalf of The Animal Consortium. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Implications

Molecular mechanisms underlying sheep reproduction have been poorly clarified resulting in economic losses for sheep production systems. Reproductive traits are usually low heritable emphasising the need to detect relevant genomic regions and causal variants controlling the respective biological processes. This study focuses on genetic and genome-wide characterisation of total prolificacy, age at first lambing and maternal lamb survival of Chios sheep. All traits estimated significantly heritable and amenable to improvement. New single-nucleotide polymorphism markers and candidate genes significantly associated with age at

first lambing were detected on chromosomes 2 and 12 that could enhance the marker-assisted selection in sheep reproduction.

Introduction

Reproductive ability constitutes an important trait for female sheep production affecting the profitability of breeding systems (Mohammadi et al., 2011). Several relevant traits have been genetically studied with various measurement methods and trait definitions, including prolificacy (Xu et al., 2018; Smolucha et al., 2021), age at first lambing (Lôbo et al., 2009; Abdoli et al., 2019a), and maternal lamb survival (Ghavi Hossein-Zadeh et al., 2018). Reproduction is a complex process affected by multiple environmental and genetic factors (Abdoli et al., 2016). Following the disentanglement of the environment from the genetic component, the incorporation of quantitative trait loci (QTL) information into selection

* Corresponding author at: Department of Genetics, Development & Molecular Biology, School of Biology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece.

E-mail address: tsarvale@bio.auth.gr (V. Tsartsianidou).

purposes may be utilised for rapid genetic progress of fertility (Liu et al., 2017).

Previous genetic studies reported low to moderate heritability for quantitative polygenic reproductive traits (Safari et al., 2005; Cloete et al., 2009; Ghavi Hossein-Zadeh et al., 2018). Furthermore, genome-wide association studies (GWAS) have been conducted, utilising medium or high-density SNP BeadChip arrays, where new genetic variants and candidate genes were identified and associated with sheep prolificacy, representing potential targets for within-breed genomic selection (Xu et al., 2018; Abdoli et al., 2019a; 2019b); Smořucha et al., 2021). Indicatively, 19 candidate genes were identified in five prolific breeds (Wadi, Hu, Icelandic, Finnsheep, and Romanov) and the non-prolific Texel sheep associated to litter size, emphasising the diverse biological processes possibly involved in prolificacy control (Xu et al., 2018). GWAS on reproductive traits of Iranian fat-tailed sheep also revealed new genes (*LHCGR*, *TEX12*, *BCO2*, *WDR70*) involved in pathways relevant to fertility, growth, and litter size (Abdoli et al., 2019b) in addition to repeatedly detected major genes (e.g., *BMPR1B*, *BMP15*, *GDF9*) in other prolific sheep breeds (Gootwine, 2020). Few genome-wide scans have been reported for the first lambing age with limited genetic markers and relevant genes detected (Abdoli et al., 2019a; Pasandideh et al., 2020).

Chios sheep is considered a highly productive and prolific dairy breed mostly distributed in Northern Greece and usually reared in semi-intensive or intensive farming conditions (Gelasakis et al., 2018). Ewes usually lamb from September to May and their lactation period normally lasts five to six months (Basdagianni et al., 2019). Reproductive performance of Chios sheep has only been studied for common mutations (*FecB*, *FecX¹* and *FecX^B*) of the major genes *BMPR1B* and *BMP15* that increase fertility (Michailidis et al., 2008), but none of these mutations were identified, implying that Chios high prolificacy may be related to other causal variants.

Although genetic improvement has been applied for milk production by the Chios Sheep Breeder's Cooperative since 1997, the implementation of selective breeding for reproductive traits has not been examined to our knowledge. On the other hand, there is an increasing international scientific interest to understand the genomic architecture of reproductive process and genomic prediction has been applied to reproduction traits in sheep to increase the genetic gain (Bolormaa et al., 2017).

The present study focuses on the genomic characterisation of reproductive traits in Chios dairy sheep. Our specific objectives were to: (i) estimate genetic parameters for the number of lambs born (total prolificacy), age at first lambing and maternal lamb survival, (ii) perform independent GWAS for these animal traits and (iii) identify candidate genes involved in molecular pathways underlying Chios sheep reproduction.

Material and methods

Animal phenotypes and pedigree

Individual animal records of purebred Chios dairy sheep were obtained from the Chios Sheep Breeders' Cooperative Macedonia pertaining the 2005–2019 period. Total prolificacy, maternal lamb survival and age at first lambing were selected as representative reproductive traits. Total prolificacy refers to the number of total lambs born to a ewe during her lifetime, and maternal lamb survival is defined as the ratio of alive offspring (2 months following lambing) to the number of total offspring of a ewe.

The data were edited for age at first lambing (12–48 months) and years of productive life following the first lambing date (at least 1 year). Records obtained before 42 days of lambing were removed considering the duration of the suckling period. Finally,

animals with less than three milk records during their productive life were also removed. The edits resulted in a final dataset of 24 082 animals, raised in 126 farms (Supplementary Table S1).

These records were matched to lifetime milk yield and performance resilience to climate (Tavg10, Tavg25) from a previous study (Tsartsianidou et al., 2021). Individual total lifetime milk yield was estimated from test-day milk records following the rules of the International Committee for Animal Recording (ICAR, 2020). Individual animal performance resilience phenotypes were based on the slopes of individual reaction norms, reflecting milk yield change across certain temperature levels that represent cold (Tavg10) and heat (Tavg25) stress temperatures respectively.

Animal pedigree was also obtained from the Cooperative including 101 493 animals, 2 324 sires and 26 008 dams.

Genetic parameters estimation

Data were analysed with the following mixed-effect model:

$$y = Xa + Zu + \varepsilon \quad (1)$$

where y is the vector of animal phenotypes, a is the vector of associated fixed effects, X the design matrix, u represents the vector of the animal random additive genetic effect normally distributed ($0, V_a A$, with V_a being the additive genetic variance of the trait and A the pedigree-based numerator relatedness matrix), Z the design matrix that associates observations with the appropriate combination of random effects and ε is the vector of residual errors. Fixed effects included in the model for all traits were the farm and first lambing year and month. The total number of lactations, total days milked and age at first lambing were further fitted for prolificacy, birth-year and month for age at first lambing, and birth-year, total prolificacy, and lifetime milk production for maternal lamb survival.

Trait heritability and correlation between traits were estimated with univariate and bivariate analyses based on Model (1). All analyses were conducted with the ASREML software (Gilmour et al., 2021). Phenotypic and genetic correlations were also estimated between reproductive traits and the other animal phenotypes.

The statistical significance ($P < 0.05$) of all parameter estimates was assessed using the two-tailed Student's t -distribution. Additionally, the significance of all fitted fixed effects ($P < 0.001$) was assessed with the Wald F statistic.

Animal genotyping and quality control

A total of 538 Chios female sheep raised on three farms in Northern Greece were genotyped using the Ovine SNP50 Genotyping BeadChip containing 54 241 single-nucleotide polymorphisms (SNPs) and designed in collaboration with the International Sheep Genomics Consortium (International Sheep Genomics Consortium, 2015), which had included 23 Chios individuals. Quality control was performed by implementing Plink v1.07 software (Purcell et al., 2007). The process involved the removal of the following data: (i) SNPs on sex and zero chromosomes, (ii) SNPs with a call rate < 0.1 , (iii) SNPs with a minor allele frequency (MAF) < 0.02 , (iv) individuals with a call rate < 0.1 , and (v) duplicate samples. A total of 528 individuals and 45 936 SNPs comprised the final dataset across the 26 ovine chromosomes.

Genotype principal component analysis

A principal component analysis (PCA) was conducted to explore the population structure that, if present, should be considered for subsequent genomic analyses to avoid inflation. A genomic relatedness matrix was constructed and decomposed into eigenvalues

and eigenvectors. The analysis was performed with GEMMA software (Zhou and Stephens, 2012).

Genome-wide association study

GWAS was performed between Chios genotypes and individual phenotypes to detect SNPs related to total prolificacy, maternal lamb survival and age at first lambing using a linear mixed model via GEMMA software (Zhou and Stephens, 2012). The model's equation is described below:

$$y = Wa + x\beta + Zu + \varepsilon \quad (2)$$

where y is the vector of animal phenotypes, α is the vector of the same fixed effects as in model (1), x is the vector of SNP genotypes and β their associated fixed effects, u is the vector of random polygenic (additive genetic) effects normally distributed ($0, V_g G$, with G being the genomic relatedness matrix and V_g the genomic variance of the trait), W and Z are the corresponding design matrices, and ε is the vector of residual effect. A Bonferroni correction was applied for multiple testing to determine the adjusted genome-wide and suggestive (one false positive per genome scan) significance thresholds, resulting in $P < 1.09E-06$ (0.05/45 936 SNPs) and $P < 2.18E-05$ (1/45 936 SNPs), and corresponding to $-\log_{10}(P)$ of 5.96 and 4.66, respectively. Association results were plotted using RStudio software (RStudio Team, 2020).

Gene and functional annotation analysis

Pairwise linkage disequilibrium (LD) was estimated for the chromosomes where genome-wide and suggestive significant SNPs were detected with GWAS using Plink v1.07 (Purcell et al., 2007). Gene annotation was conducted utilising the Variant Effect Predictor software (McLaren et al., 2016) with distances of 1 Mb according to pairwise LD results. Finally, functional annotation was performed using the gProfiler genomic tool (Raudvere et al., 2019) and the DAVID software (Huang et al., 2009) to identify candidate genes relevant to livestock fertility. Uniprot (The UniProt Consortium, 2021) and Animal QTL (Hu et al., 2022) databases were utilised to identify molecular functions, pathways and QTL.

Results

Data descriptive statistics

Descriptive statistics of Chios reproduction, lifetime milk production and resilience traits are summarised in Table 1. The highest CV among fertility traits was observed for total prolificacy (74.15%). The respective estimates for milk performance (93.61%) and climate resilience to cold (109.64%) and hot weather (815.07%) were considerably higher compared to fertility traits (29.47–74.15%).

Table 1
Descriptive statistics of Chios reproduction, lifetime milk production and resilience traits.

| Phenotype | Mean | SD | CV (%) | Minimum | Maximum |
|---|--------|--------|--------|---------|----------|
| Total prolificacy | 4.31 | 3.20 | 74.15 | 1 | 25 |
| Age at 1st lambing (days) | 595.88 | 232.05 | 38.94 | 345 | 1 454 |
| Maternal lamb survival | 0.88 | 0.26 | 29.47 | 0.00 | 1.00 |
| Lifetime milk production (kg) | 552.73 | 517.44 | 93.61 | 10.04 | 4 751.27 |
| Resilience to cold (10 °C) weather (kg/°C) ¹ | 0.012 | 0.01 | 109.64 | −0.050 | 0.086 |
| Resilience to hot (25 °C) weather (kg/°C) ¹ | 0.002 | 0.01 | 815.07 | −0.060 | 0.045 |

¹ Resilience to cold (10 °C) and hot (25 °C) weather refers to milk yield change in response to changing air temperature at the respective temperature levels, based on the slopes of individual reaction norms.

Genetic parameters

Trait heritability estimates for the studied traits are presented in Table 2. Heritability estimates were low but significantly greater than zero for age at first lambing ($h^2 = 0.07$, SE = 0.01) and total prolificacy ($h^2 = 0.08$, SE = 0.01), and moderate for maternal lamb survival ($h^2 = 0.21$, SE = 0.01). Lifetime milk production and resilience to climate were moderate ($h^2 = 0.14$ – 0.16 , SE = 0.01).

Phenotypic and genetic correlations between reproductive, productive and resilience traits are also shown in Table 2, attesting to significantly positive genetic association between (i) total prolificacy and lifetime milk production (0.16) and (ii) age at first lambing and resilience to cold weather changes (0.18), respectively. A negligible negative phenotypic correlation was estimated between total prolificacy and maternal lamb survival (−0.05). Additionally, associations between resilience to cold or hot weather changes (Tavg10, Tavg25) and lifetime milk production were significantly negative. All other correlations were not statistically different from 0.

Population structure

Results of the principal component analysis conducted on the genomic relatedness matrix are illustrated in Fig. 1. The first three principal components accounted for almost 11% of the total variance (Fig. 1A). Strong population stratification was observed by the first two principal components attributed to the farm of origin, manifested in different colours in Fig. 1B. These results may be associated with limited genetic connections and targeted breeding practices influencing the genetic makeup of the individuals raised at the respective farms. Therefore, the farm of origin was included as a fixed effect for the subsequent genomic analyses to account for population structure.

Genome-wide association study

Genome-wide association results of Chios total prolificacy, maternal lamb survival and age at first lambing are presented in Fig. 2. The respective Quantile–Quantile (Q–Q) plots are provided in Supplementary Fig. S1, indicating there is no significant inflation. Specific regions of interest were mainly detected on chromosomes 1 and 2 for maternal lamb survival and age at first lambing, respectively. Table 3 summarises the candidate single-nucleotide polymorphisms (SNPs) with genome-wide or suggestive significant association with the fertility-related traits under study. Particularly, one suggestive significant SNP was detected on chromosome 1 for lamb survival. Two genome-wide significant SNPs and one suggestive significant were detected on chromosome 2, and one suggestive significant on chromosome 12 for age at first lambing. No genome-wide or suggestive significant SNPs were identified for the total prolificacy of Chios sheep.

Table 2

Heritability (diagonal) and correlation (above: genetic below: phenotypic) estimates (SE) between the traits under study for Chios sheep with SE in parentheses.

| Phenotype | tprol | Afl | mls | tmilk | Tavg10 | Tavg25 |
|-----------|---------------|---------------|---------------|---------------|---------------|---------------|
| tprol | 0.11 (0.01)* | 0.08 (0.08) | 0.01 (0.06) | 0.16 (0.06)* | -0.08 (0.07) | -0.10 (0.07) |
| afl | 0.00 (0.01) | 0.07 (0.01)* | -0.04 (0.07) | 0.01 (0.07) | 0.18 (0.08)* | 0.12 (0.08) |
| mls | -0.05 (0.01)* | 0.02 (0.01) | 0.21 (0.01)* | 0.07 (0.05) | 0.00 (0.06) | 0.05 (0.06) |
| tmilk | 0.02 (0.01)* | 0.00 (0.01) | -0.01 (0.01) | 0.14 (0.01)* | -0.38 (0.06)* | -0.79 (0.03)* |
| Tavg10 | 0.00 (0.01) | -0.02 (0.01)* | -0.02 (0.01)* | -0.25 (0.01)* | 0.14 (0.01)* | 0.72 (0.03)* |
| Tavg25 | 0.00 (0.01) | 0.00 (0.01) | 0.00 (0.01) | -0.66 (0.00)* | 0.72 (0.00)* | 0.16 (0.01)* |

Abbreviations: tprol = total prolificacy; afl = age at first lambing (days); mls = maternal lamb survival; Tavg10 Tavg25 = animal resilience at 10 °C and 25 °C, respectively; tmilk = lifetime milk yield (kg).

* Estimates significantly different from 0 ($P < 0.05$).

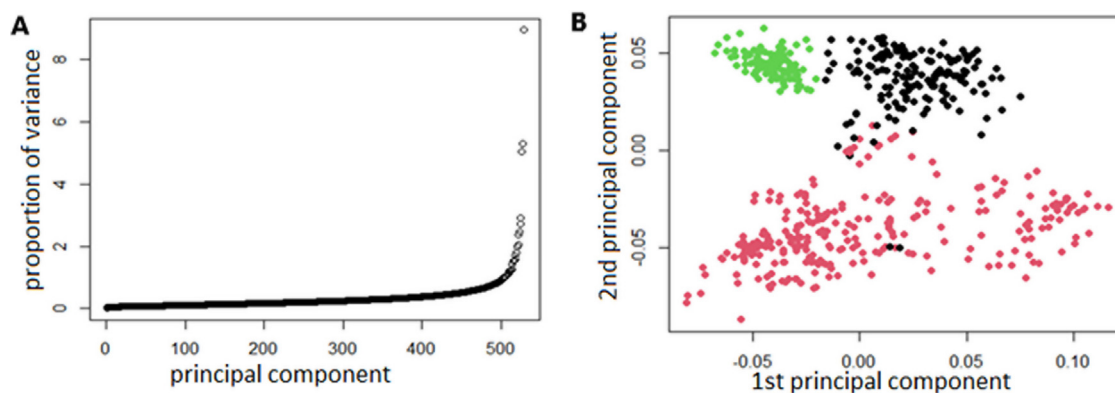


Fig. 1. Principal component analysis (PCA) of Chios sheep. (A) Proportion of variance of each principal component of the decomposed genomic relatedness matrix. (B) First and second principal components depicting population structure explained and coloured by the sheep farm of origin.

Functional annotation analyses

High pairwise LD estimates were obtained for chromosome 2, where significant SNPs were detected for age at first lambing ($r^2 = 0.79-0.93$), spanning a region of 352.79 kb (Supplementary Table S2). Based on these results, an SNP annotation analysis was followed in 1 Mb regions upstream and downstream of the SNPs. The same distance was applied also for chromosomes, where only one marker was detected. Candidate genes detected for each reproductive trait are presented in Supplementary Table S3, while most genes identified were related to age at first lambing. Indicatively, the collagen-type V alpha 2 chain (*COL5A2*) gene was located 171.94 kb upstream the OAR2_127110460.1 SNP marker and 16.18 kb upstream the OAR2_127469865.1 SNP respectively, whereas the OAR2_127499743.1 constitutes an intronic polymorphism of collagen-type III alpha 1 chain (*COL3A1*) gene. Functional annotation and enrichment analysis revealed that these collagen-type genes belong to the extracellular matrix structure (Table 4, Supplementary Fig. S2). Furthermore, the zona pellucida like domain containing 1 (*ZPLD1*) is located 473.36 kb upstream the OAR1_180150573.1 SNP marker detected for lamb survival. According to the molecular and functional characterisation, many of the annotated genes are involved in the skin, skeletal and muscle system development, and growth (Table 4, Supplementary Table S4).

203 QTLs previously reported in Animal QTL database are localised within a subset of the detected candidate genes. Specifically, these loci have been majorly associated with carcass, skeletal and muscle traits, BW and fat composition (Myostatin (*MSTN*), *PMS1*, *COL3A1*), emphasising on the functional role of these chromosomal regions while 3 QTLs were related to cattle milk traits (*MSTN*, *KAZN*) (Supplementary Table S5).

Discussion

The present study set out to generate new insights into the genetic architecture of reproductive traits in Mediterranean dairy sheep linked to total prolificacy, age at first lambing and maternal lamb survival. Focus was on the highly producing and prolific Greek Chios breed. For this purpose, we estimated trait heritability and correlations between reproductive, productive and resilience traits of interest, with a view to implement a multi-trait index in future selective breeding programmes. The genetic mechanisms underlying reproduction in Chios sheep were further explored with GWAS, where several novel variants and candidate genes were identified mostly related to age at first lambing (Table 3, Fig. 2).

Low heritability of the number of lambs born per ewe during lifetime was consistent with previously reported estimates in the Mediterranean dairy breeds Churra (0.05) (El-Saied et al., 2005), Latxa and Manchega (0.08) (Legarra et al., 2007). Similar results have been published for litter size (number of lambs born per ewe lambing) in Friesian (0.04) (Hamann et al., 2004) and cross-bred dairy sheep (0.07) (Murphy et al., 2017). It is worth mentioning that prolificacy has been mostly studied in meat and wool breeds (average $h^2 = 0.13$) (Safari et al., 2005). Additionally, the moderately positive genetic correlation estimated between total prolificacy and lifetime milk production (0.16) suggests that the traits are partially under common genetic control, which was also reported for Spanish Churra ewes but with a higher estimate (0.67) (El-Saied et al., 2005).

Although the first lambing age is an important reproductive trait affecting sheep longevity and lifetime milk production (Schoeman and Albertyn, 1991), the genetic parameters have not been extensively evaluated. Trait heritability in the Chios sheep was low (0.07) in concordance with few reported estimates in

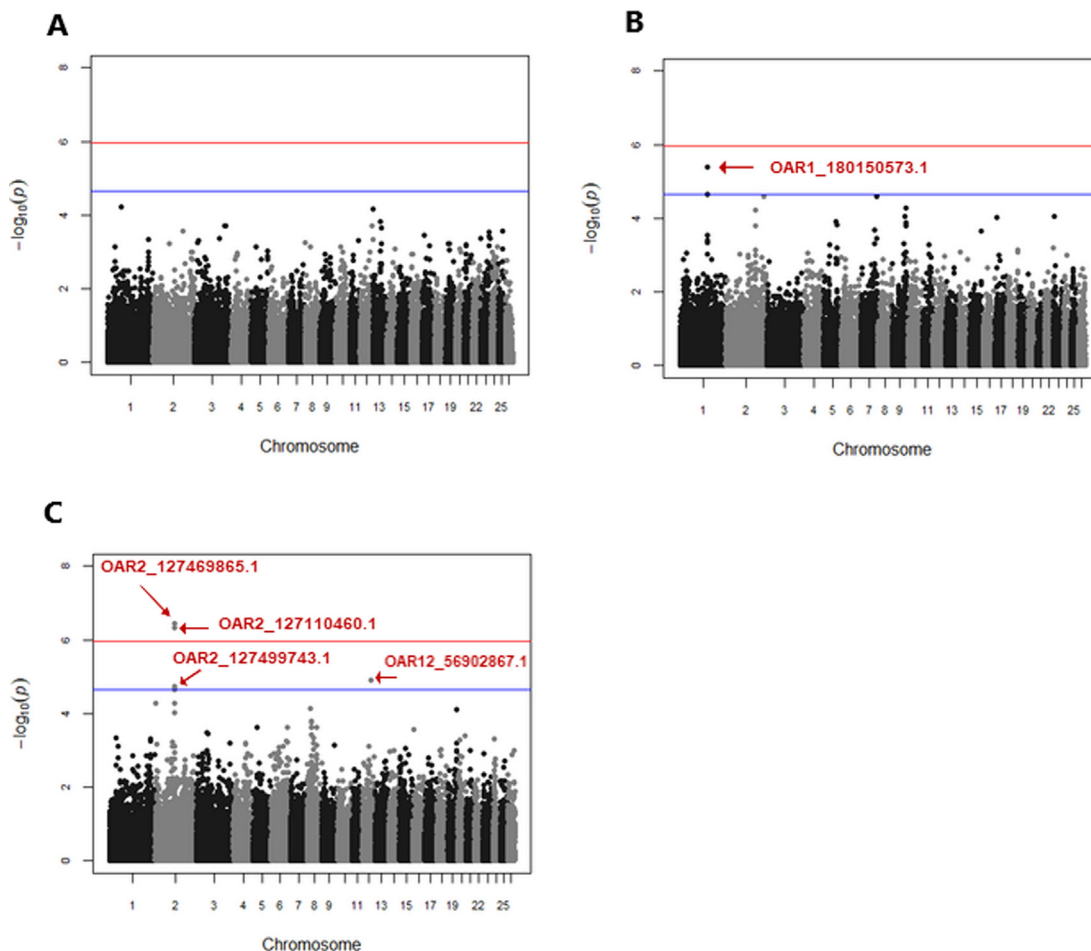


Fig. 2. Genome-wide association results for Chios sheep reproductive traits. (A) Total prolificacy, (B) maternal lamb survival, (C) age at first lambing. Red and blue lines indicate the genome-wide (Bonferroni correction) and suggestive (one false positive per genome scan) significance thresholds, respectively.

Table 3

Genome-wide and suggestive significant SNPs associated with Chios reproductive traits. The respective chromosomal locations refer to the Rambouillet v1.0 reference genome.

| Phenotype | Chr | SNP | Position (bp) | Location/Distance Gene | Alleles | MAF | Effect ¹ | P-value ² |
|------------------------|------------------|------------------|------------------------|--------------------------|---------|-------|---------------------|----------------------|
| Maternal lamb survival | 1 | OAR1_180150573.1 | 183 753 694 | D/473361 ZPLD1 | C/T | 0.081 | -0.101 | 3.99E-06* |
| Age at first lambing | 2 | OAR2_127110460.1 | 129 887 607 | D/821304 <i>MSTN</i> | G/A | 0.392 | 0.858 | 3.63E-07** |
| | | | | U/546124 <i>PMS1</i> | | | | |
| | | | | U/435100 <i>ASDURF</i> | | | | |
| | | | | D/237863 <i>SLC40A1</i> | | | | |
| 2 | OAR2_127469865.1 | 130 210 267 | D/16181 <i>COL5A2</i> | A/G | 0.391 | 0.848 | 4.57E-07** | |
| | | | D/2456 <i>COL3A1</i> | | | | | |
| 2 | OAR2_127499743.1 | 130 240 392 | I/- <i>COL3A1</i> | T/C | 0.355 | 0.862 | 1.26E-05* | |
| 12 | OAR12_56902867.1 | 57 299 327 | D/433007 <i>GULP1</i> | A/G | 0.055 | 1.799 | 1.84E-05* | |
| I/- <i>KAZN</i> | | | | | | | | |
| D/375528 <i>PRDM2</i> | | | | | | | | |
| D/585939 <i>PDPN</i> | | | | | | | | |
| | | | | U/687853 <i>LRRC28</i> | | | | |

Abbreviations: SNP = single-nucleotide polymorphism; Chr = Chromosome; bp = base pairs; MAF = minor allele frequency; D, U, and I = downstream, upstream, intron variant.

¹ Allelic substitution effect.

² Wald test-derived P-values exceeding (*) the Bonferroni genome-wide ($P < 1.09E-06$) and (**) the suggestive significance threshold ($P < 2.18E-05$), respectively.

meat (0.04) (Lôbo et al., 2009) and dairy sheep (0.02) (El-Saied et al., 2005). A first insight into the relationship between resilience and reproduction was manifested in a positive favourable genetic correlation between resilience to cold weather (Tavg10) and age at first lambing (Table 2), assuming that the direction of genetic selection would lead to intermediate optimal values for both animal traits.

Chios lamb survival was found to be a moderately heritable trait (0.21) in agreement with a previous estimate (0.16) in the closely related Sakiz sheep (Ceyhan et al., 2009). However, several other studies have reported a lower inherent capacity of lamb survival (Safari et al., 2005). These discrepancies may be related to the inconsistent trait definition usually referring to the lamb rather than the maternal ewe trait or different time periods of lamb sur-

Table 4
Indicative list of the candidate causal genes associated with Chios reproductive traits and annotated using UniProt database.

| Phenotype | Chr | Gene | Description | Molecular Function, Biological process, Pathway |
|------------------------|-----|---------|--|--|
| Maternal lamb survival | 1 | ZPLD1 | zona pellucida like domain containing 1 | vestibular reflex |
| Age at first lambing | 2 | MSTN | Ovis aries myostatin | growth factor activity, muscle cell cellular homeostasis, myoblast migration involved in skeletal muscle regeneration |
| | 2 | PMS1 | PMS1 homolog 1, is match repair system component | ATP binding, ATP hydrolysis activity, mismatched DNA-binding, DNA-binding, mismatch repair, DNA damage |
| | 2 | ASDURF | ASNSD1 upstream reading frame | asparagine synthase (glutamine-hydrolysing) activity, amino acid biosynthesis, asparagine biosynthetic process |
| | 2 | SLC40A1 | solute carrier family 40 member 1 | iron ion transmembrane transporter activity, peptide hormone binding, ferroptosis, mineral absorption |
| | 2 | WDR75 | WD repeat domain 75 | positive regulation of rRNA processing, positive regulation of transcription by RNA polymerase, ribosome biogenesis in eukaryotes |
| | 2 | COL5A2 | collagen-type V alpha 2 chain | extracellular matrix structural constituent, cellular response to amino acid stimulus, collagen fibril organisation, eye morphogenesis, skeletal system development |
| | 2 | COL3A1 | collagen-type III alpha 1 chain | extracellular matrix structural constituent, aorta smooth muscle tissue morphogenesis, endochondral bone morphogenesis, fibroblast proliferation, heart and lung development |
| | 2 | GULP1 | GULP PTB domain containing engulfment adaptor 1 | phagocytosis, engulfment |
| | 12 | PRDM2 | PR/SET domain 2 | histone methyltransferase and transferase activity |

Abbreviations: Chr = Chromosome.

vival (Ghavi Hossein-Zadeh et al., 2018), and the implemented modelling (Sawalha et al., 2007).

All reproductive traits in the present study had heritability and genetic variance estimates that were significantly greater than zero and no clear genetic antagonism was evident among traits, indicating that reproduction may be amenable to improvement through genetic selection in a breeding programme.

Evident population structure was demonstrated by principal component analysis for Chios sheep raised in three farms of Northern Greece. This genetic stratification was attributed to the sheep farm of origin, as depicted by different colours (Fig. 1B), suggesting distinct within-farm selective breeding and mating practices.

Our GWAS revealed three SNPs on chromosome 2 and one on chromosome 12 (Table 3, Fig. 2) significantly associated with age at first lambing. These genetic variants are reported for the first time, to our knowledge, while very few genome association analyses have been conducted for this reproductive trait (Abdoli et al., 2019a; Pasandideh et al., 2020). Although no QTLs linked to reproduction have been identified on these chromosomal areas, functional annotation revealed a set of genes closely located to the detected SNP markers on chromosome 2, such as *MSTN*, *COL5A2* and *COL3A1* (Table 4, Supplementary Table S3), that participate in biological processes strongly affecting prolificacy. Indicatively, the OAR2_127469865.1 SNP is located 16.18 kb downstream the *COL5A2* gene while the OAR2_127499743.1 SNP marker is positioned within the fourth intron of *COL3A1* gene. High linkage disequilibrium estimates between all significant SNPs on chromosome 2 further suggest that they are possibly inherited as a haplotype block functionally related to the first lambing age.

Characteristically, the *MSTN* gene (Myostatin or growth and differentiation factor 8) is a member of the TGF β superfamily, similar to reported major genes (*BMP1B*, *BMP15*, *GDF9*) that affect the ovulation rate (Davis, 2005) and participate in bone formation and growth processes strongly associated with high prolificacy (Grochowska et al., 2020). Interestingly, the *MSTN* gene has been associated with sheep QTLs that influence relevant biological processes such as sheep growth attributed to skeletal and muscle mass development (Osman et al., 2021). The collagen-type genes *COL3A1* and *COL5A2* are only 40 kb apart, and they contribute to osteogenesis, normal bone formation, skeletal development, myogenesis and muscle growth (Shen et al., 2021) also resembling the functionality of the major genes. Indeed, functional enrichment

analysis confirmed the association of these collagen genes with multiple uterine-related disfunctions, such as cervical insufficiency, uterine prolapse and abnormalities of the uterine cervix (Supplementary Fig. S2). Within this genomic area, the *SLC40A1* gene was detected and annotated in the same family with *SLC5A1* (solute carrier family 40 member 1) that has been associated with major effect on litter size in Chinese sheep breeds (La et al., 2019).

The SNP marker OAR12_56902867.1 was located close to several candidate genes (*KAZN*, *PRDM2*, *PDPN* and *LRRRC28*) (Supplementary Table S3) grouped in enrichment clusters majorly involved in the regulation of several developmental processes such as cell differentiation, proliferation and death, developmental and biosynthetic pathways, apoptosis, and nucleic acid-templated transcription (Supplementary Fig. S2). Previous studies have reported candidate genes located on chromosome 12 (*NR5A2*, *GDF11*, *GORAB*) affecting animal reproduction through cell differentiation and development (Abdoli et al., 2016).

Significant SNPs and candidate genes detected in the present study were associated with development and body growth, indicating a potential association with first lambing age. Consistently, sufficient physical growth and BW have been proven necessary for the first lambing age without decreasing productivity according to previously reported results for sheep (de Souza Arco et al., 2021) and cattle (Lacerda et al., 2018) breeds. Further research on point mutations of the detected genes and the respective gene transcription may shed light into the cellular mechanisms affecting sheep reproduction related to age at first lambing.

No genome-wide or suggestive significant SNP markers were detected for total prolificacy of Chios ewe that may be related to low trait heritability or small variant effects that could not be detected in our population of study. Regional heritability mapping analysis may permit the identification of genomic regions of interest, capturing consecutive small variant effects associated with this trait.

Conclusion

Our results provide novel insights into the genetic architecture of Chios sheep reproductive traits, focusing on the total prolificacy, the age at first lambing and the maternal lamb survival. Heritable genetic variance ($h^2 = 0.07\text{--}0.21$) was estimated that may be uti-

lised in selective breeding schemes to enhance reproductive efficiency together with other economically important animal traits. Our GWAS results revealed four new SNP markers (OAR2_127110460.1, OAR2_127469865.1, OAR2_127499743.1, OAR12_56902867.1) on chromosomes 2 and 12 and relevant candidate genes (*MSTN*, *PMS1*, *ASDURF*, *SLC40A1*, *WDR75*, *COL5A2*, *COL3A1*, *GULP1*, *PRDM2*) significantly associated with age at first lambing that may enhance the list of causal variants related to prolific sheep breeds. These genes are involved in bone and muscle formation and growth processes linked to high sheep prolificacy. Future functional studies may further disentangle the molecular characterisation of the candidate genes related to reproductive ability.

Supplementary material

Supplementary material to this article can be found online at <https://doi.org/10.1016/j.animal.2023.100723>.

Ethics approval

Not applicable.

Data and model availability statement

Data analysed during the current study are available from the corresponding author upon reasonable request. Phenotypic and genotypic datasets generated are deposited in Mendeley public repository available at <https://data.mendeley.com/datasets/nvjn8mvwtg/1>.

Author ORCIDs

Valentina Tsartsianidou: <https://orcid.org/0000-0003-3723-9835>.

Alexandros Pavlidis: <https://orcid.org/0000-0002-7062-851X>.

Eleftheria Tosiou: <https://orcid.org/0000-0002-0351-8054>.

Georgios Banos: <https://orcid.org/0000-0002-7674-858X>.

Georgios Arsenos: <https://orcid.org/0000-0003-3224-5128>.

Alexandros Triantafyllidis: <https://orcid.org/0000-0003-0469-011X>.

Author contributions

V. Tsartsianidou, A. Triantafyllidis, G. Arsenos and G. Banos conceptualised and designed the current study. **V. Tsartsianidou, A. Pavlidis** and **E. Tosiou** performed the formal analyses and visualisation. **V. Tsartsianidou, A. Pavlidis** and **E. Tosiou** drafted the first manuscript. All authors contributed to the manuscript revision, read and approved the submitted version. **A. Triantafyllidis** and **G. Arsenos** acquired funding.

Declaration of interest

The authors declare no conflict of interest.

Acknowledgements

The authors would like to thank Chios Sheep Breeder's Cooperative Macedonia for the animal records provided.

Financial support statement

This work was funded by SMARTER Horizon 2020 project: "Small Ruminants breeding for Efficiency and Resilience" (772787, <https://www.smarterproject.eu/>).

References

- Abdoli, R., Zamani, P., Mirhoseini, S.Z., Ghavi Hossein-Zadeh, N., Nadri, S., 2016. A review on prolificacy genes in sheep. *Reproduction in Domestic Animals* 51, 631–637.
- Abdoli, R., Mirhoseini, S.Z., Ghavi Hossein-Zadeh, N., Zamani, P., Moradi, M.H., Ferdosi, M.H., Gondro, C., 2019a. Genome-wide association study of first lambing age and lambing interval in sheep. *Small Ruminant Research* 178, 43–45.
- Abdoli, R., Mirhoseini, S.Z., Ghavi Hossein-Zadeh, N., Zamani, P., Ferdosi, M.H., Gondro, C., 2019b. Genome-wide association study of four composite reproductive traits in Iranian fat-tailed sheep. *Reproduction, Fertility and Development* 31, 1127–1133.
- Basdagianni, Z., Sinapis, E., Banos, G., 2019. Evaluation of reference lactation length in Chios dairy sheep. *Animal* 13, 1–7.
- Bolormaa, S., Brown, D.J., Swan, A.A., van der Werf, J.H.J., Hayes, B.J., Daetwyler, H.D., 2017. Genomic prediction of reproduction traits for Merino sheep. *Animal Genetics* 48, 338–348.
- Ceyhan, A., Sezenler, T., Erdoğan, I., 2009. The estimation of variance components for prolificacy and growth traits of Sakiz sheep. *Livestock Science* 122, 68–72.
- Cloete, S.W.P., Misztal, I., Olivier, J.J., 2009. Genetic parameters and trends for lamb survival and birth weight in a Merino flock divergently selected for multiple rearing ability. *Journal of Animal Science* 87, 2196–2208.
- Davis, G.H., 2005. Major genes affecting ovulation rate in sheep. *Genetics Selection Evolution* 37 (suppl. 1), S11–S23.
- de Souza Arco, T.F.F., Itavo, C.C.B.F., Itavo, L.C.V., Souza, M.I.L., Dias, A.M., dos Santos Difante, G., da Costa E Silva, E.V., de Almeida Borges, F., Borges, D.G.L., da Silva Monteiro Ferelli, K.L., Rodrigues, B.J., da Silva Miguel, A.A., de Godoy, C., de Melo Soares, E.S., 2021. Effects of supplementation on production and age anticipation at first mating of ewe lambs in tropical pastures. *Tropical Animal Health and Production* 53, 481.
- El-Saied, U.M., De La Fuente, L.F., Carriedo, J.A., San Primitivo, F., 2005. Genetic and phenotypic parameter estimates of total and partial lifetime traits for dairy ewes. *Journal of Dairy Science* 88, 3265–3272.
- Gelasakis, A.I., Valergakis, G.E., Fortomaris, P., Arsenos, G., 2018. Farm conditions and production methods in Chios sheep flocks. *Journal of the Hellenic Veterinary Medical Society* 61, 111–119.
- Ghavi Hossein-Zadeh, N., Noori, R., Shadparvar, A.A., 2018. Genetic analysis of longevity and lamb survival from birth to yearling in Moghani sheep. *Journal of Applied Animal Research* 46, 1363–1369.
- Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J., Thompson, R., 2021. ASReml User Guide Release 4.2 Functional Specification. Retrieved on 5 October 2022 from <https://asreml.kb.vsnr.co.uk/wp-content/uploads/sites/3/ASReml-4.2-Functional-Specification.pdf>.
- Gootwine, E., 2020. Invited review: Opportunities for genetic improvement toward higher prolificacy in sheep. *Small Ruminant Research* 186, 106090.
- Grochowska, E., Borys, B., Mroczkowski, S., 2020. Effects of intronic SNPs in the myostatin gene on growth and carcass traits in colored polish merino sheep. *Genes (Basel)* 11, 2.
- Hamann, H., Horstick, A., Wessels, A., Distl, O., 2004. Estimation of genetic parameters for test day milk production, somatic cell score and litter size at birth in East Friesian ewes. *Livestock Production Science* 87, 153–160.
- Hu, Z.L., Park, C.A., Reecy, J.M., 2022. Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. *Nucleic Acids Research* 50, D956–D961.
- Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research* 37, 1–13.
- International Committee for Animal Recording (ICAR), 2020. ICAR Guidelines: Dairy Sheep and Goats. Retrieved on 5 October 2022 from <http://www.icar.org>.
- International Sheep Genomics Consortium, 2015. Retrieved on 5 October 2022 from <https://www.sheepmap.org/>.
- La, Y., Liu, Q., Zhang, L., Chu, M., 2019. Single Nucleotide Polymorphisms in *SLC5A1*, *CCNA1*, and *ABCC1* and the Association with Litter Size in Small-Tail Han Sheep. *Animals* 9, 432.
- Lacerda, V.V., Campos, G.S., Roso, V.M., Souza, F.R.P., Brauner, C.C., Boligon, A.A., 2018. Effect of mature size and body condition of Nelore females on the reproductive performance. *Theriogenology* 118, 27–33.
- Legarra, A., Ramón, M., Ugarte, E., Pérez-Guzmán, M.D., 2007. Economic weights of fertility, prolificacy, milk yield and longevity in dairy sheep. *Animal* 1, 193–203.
- Liu, A., Wang, Y., Sahana, G., Zhang, Q., Liu, L., Lund, M.S., Su, G., 2017. Genome-wide Association Studies for Female Fertility Traits in Chinese and Nordic Holsteins. *Scientific Reports* 7, 8487.
- Lôbo, A.M.B.O., Lôbo, R.N.B., Paiva, S.R., de Oliveira, S.M.P., Facó, O., 2009. Genetic parameters for growth, reproductive and maternal traits in a multibreed meat sheep population. *Genetics and Molecular Biology* 32, 761–770.

- McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., Cunningham, F., 2016. The Ensembl Variant Effect Predictor. *Genome Biology* 17, 122.
- Michailidis, G., Avdi, M., Pappa, V., 2008. Reproductive performance and investigation of BMPR-IB and BMP-15 gene mutation in Greek Chios and Florina sheep breeds. *Archiva Zootechnica* 11, 24–31.
- Mohammadi, A.R., Abbasi, M.A., Moghaddam, A.A., Zare, S.A., 2011. Determination of some reproductive traits in Iranian Afshari sheep breed. *Australian Journal of Basic and Applied Sciences* 5, 2742–2751.
- Murphy, T.W., Berger, Y.M., Holman, P.W., Baldin, M., Burgett, R.L., Thomas, D.L., 2017. Estimates of genetic parameters, genetic trends, and inbreeding in a crossbred dairy sheep research flock in the United States. *Journal of Animal Science* 95, 4300–4309.
- Osman, N.M., Shafey, H.I., Abdelhafez, M.A., Sallam, A.M., Mahrous, K.F., 2021. Genetic variations in the Myostatin gene affecting growth traits in sheep. *Veterinary World* 14, 475–482.
- Pasandideh, M., Gholizadeh, M., Rahimi-Mianji, G., 2020. Identification of two novel SNPs affecting lambing traits in sheep by using a 50K SNP-Chip. *Small Ruminant Research* 191, 106193.
- 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., Sham, P.C., 2007. PLINK: A Ttool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American Journal of Human Genetics* 81, 559–575.
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., Vilo, J., 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research* 47, W191–W198.
- RStudio Team, 2020. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. Retrieved on 5 October 2022 from <http://www.rstudio.com/>.
- Safari, E., Fogarty, N.M., Gilmour, A.R., 2005. A review of genetic parameter estimates for wool, growth, meat and reproduction traits in sheep. *Livestock Production Science* 92, 271–289.
- Sawalha, R.M., Conington, J., Brotherstone, S., Villanueva, B., 2007. Analyses of lamb survival of Scottish blackface sheep. *Animal* 1, 151–157.
- Schoeman, S.J., Albertyn, J.R., 1991. Lifetime reproduction of Karukul ewes as influenced by season of birth, age at first lambing and lambing interval. *South African Journal of Animal Science* 21, 169–172.
- Shen, J., Hao, Z., Wang, J., Hu, J., Liu, X., Li, S., Ke, N., Song, Y., Lu, Y., Hu, L., Qiao, L., Wu, X., Luo, Y., 2021. Comparative Transcriptome Profile Analysis of Longissimus dorsi Muscle Tissues From Two Goat Breeds With Different Meat Production Performance Using RNA-Seq. *Frontiers in Genetics* 11, 619399.
- Smolucha, G., Gurgul, A., Jasielczuk, I., Kawęcka, A., Miksza-Cybulska, A., 2021. A genome-wide association study for prolificacy in three Polish sheep breeds. *Journal of Applied Genetics* 62, 323–326.
- The UniProt Consortium, 2021. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research* 49, D480–D489.
- Tsartsianidou, V., Sánchez-Molano, E., Kapsona, V.V., Basdagianni, Z., Chatziplis, D., Arsenos, G., Triantafyllidis, A., Banos, G., 2021. A comprehensive genome-wide scan detects genomic regions related to local adaptation and climate resilience in Mediterranean domestic sheep. *Genetics Selection Evolution* 53, 90.
- Xu, S.S., Gao, L., Xie, X.L., Ren, Y.L., Shen, Z.Q., Wang, F., Shen, M., Eypórsdóttir, E., Hallsson, J.H., Kiseleva, T., Kantanen, J., Li, M.H., 2018. Genome-Wide Association Analyses Highlight the Potential for Different Genetic Mechanisms for Litter Size Among Sheep Breeds. *Frontiers in Genetics* 9, 118.
- Zhou, X., Stephens, M., 2012. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* 44, 821–824.