Genetic and genomic analyses underpin the feasibility of concomitant genetic improvement of milk yield and mastitis resistance in dairy sheep

Georgios Banos[1,2,3], **Emily L. Clark**[2], **Stephen J. Bush**[2,4], **Prasun Dutta**[2], **Georgios Bramis**[3], **Georgios Arsenos**[3], **David A. Hume**[2,5], **Androniki Psifidi**[2,6]*

1 Scotland's Rural College, Edinburgh, Easter Bush, Midlothian, Scotland, United Kingdom, 2 The Roslin Institute, University of Edinburgh, Easter Bush, Midlothian, Scotland, United Kingdom, 3 School of Veterinary Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece, 4 Nuffield Department of Clinical Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, England, United Kingdom, 5 Mater Research Institute-University of Queensland, Translational Research Institute, Woolloongabba, Australia, 6 Royal Veterinary College, University of London, Hatfield, England, United Kingdom

* apsifidi@rvc.ac.uk, androniki.psifidi@roslin.ed.ac.uk

Abstract

Milk yield is the most important dairy sheep trait and constitutes the key genetic improvement goal via selective breeding. Mastitis is one of the most prevalent diseases, significantly impacting on animal welfare, milk yield and quality, while incurring substantial costs. Our objectives were to determine the feasibility of a concomitant genetic improvement programme for enhanced milk production and resistance to mastitis. Individual records for milk yield, and four mastitis-related traits (milk somatic cell count, California Mastitis Test score, total viable bacterial count in milk and clinical mastitis presence) were collected monthly throughout lactation for 609 ewes of the Chios breed. All ewes were genotyped with a mastitis specific custom-made 960 single nucleotide polymorphism (SNP) array. We performed targeted genomic association studies, (co)variance component estimation and pathway enrichment analysis, and characterised gene expression levels and the extent of allelic expression imbalance. Presence of heritable variation for milk yield was confirmed. There was no significant genetic correlation between milk yield and mastitis traits. Environmental factors appeared to favour both milk production and udder health. There were no overlapping of SNPs associated with mastitis resistance and milk yield in Chios sheep. Furthermore, four distinct Quantitative Trait Loci (QTLs) affecting milk yield were detected on chromosomes 2, 12, 16 and 19, in locations other than those previously identified to affect mastitis resistance. Five genes (*DNAJA1*, *GHR*, *LYPLA1*, *NUP35* and *OXCT1*) located within the QTL regions were highly expressed in both the mammary gland and milk transcriptome, suggesting involvement in milk synthesis and production. Furthermore, the expression of two of these genes (*NUP35* and *OXCT1*) was enriched in immune tissues implying a potentially pleiotropic effect or likely role in milk production during udder infection, which needs to be further elucidated in future studies. In conclusion, the absence of genetic

antagonism between milk yield and mastitis resistance suggests that simultaneous genetic improvement of both traits be achievable.

Introduction

The world's commercial dairy sheep industry is primarily concentrated in Mediterranean countries and linked to local breeds; milk is mostly used to produce high quality cheeses and other dairy products. Milk yield represents more than two thirds of the total income of the dairy sheep sector [1] and, therefore, increasing milk yield is the most important and some-times only objective of selective breeding. Milk production traits in dairy sheep are lowly to moderately heritable, with reported heritability estimates ranging from 0.13 to 0.51 [2, 3] and amenable to improvement with traditional selective breeding programmes based on pedigree and phenotypic data. Indeed, such programmes have been established in many sheep populations over recent decades [2, 4]. Incorporation of genomic information in some breeding programmes (e.g. French Lacaune, Spanish Churra, Italian Sarda) has led to an acceleration of the genetic improvement outcomes.

The Greek Chios breed is considered to be among the most productive and prolific dairy sheep breeds worldwide [5]. A traditional breeding programme for the enhancement of milk yield has been in place since year 2000 for this breed, leading to substantial improvement in this trait. However, further increases in milk yield may be achieved with the use of relevant genomic information.

Beyond simply increasing milk production, the dairy sheep industry faces challenges such as the need to offer healthy products to consumers, addressing animal welfare, and ensuring the long-term competitiveness and sustainability of the sector. Mastitis is the most prevalent and costly disease in the dairy industry due to reduced and discarded milk, early involuntary culling of animals, and veterinary services and labour costs [6, 7]. The disease also poses a potential threat of zoonosis and antimicrobial resistance if antibiotic treatment is not applied carefully [6–8]. Moreover, mastitis is a welfare concern because of associated pain, anxiety and restlessness, and upsets the normal feeding behaviour of the animals [9]. Host resistance to mastitis is generally a lowly heritable trait, with heritability estimates previously reported ranging from 0.10 to 0.20 [7]. Recently, an ovine custom made mastitis specific 960-SNP DNA array was built to facilitate genetic selection and improvement of animal resistance to mastitis in dairy sheep [10] [11] [12] [13]. We previously used this array in a targeted genomic association study and detected five quantitative trait loci (QTLs) for mastitis resistance in Chios sheep [10].

In the present study, we examined the genetic relationship between milk yield and mastitis resistance in the Chios sheep, using pedigree and genomic information. Mastitis resistance was manifested with four relevant measured traits, namely milk somatic cell count, California Mastitis Test score, total viable bacterial count in milk and clinical manifestation of the disease. The relationship between milk yield and these -mastitis traits is crucial if enhancing mastitis resistance is to be included in the selective breeding goal together with increasing milk production. We estimated genetic parameters and investigated whether SNP markers previously found to be associated with mastitis resistance in Chios sheep were also associated with milk yield, using the ovine custom-made mastitis-specific array. We also performed pathway analysis and examined gene expression and allelic expression imbalance to assess whether genes located under the QTL regions, were enriched in tissues relevant to milk yield and mastitis resistance.

Materials and methods

Ethical statement

The study was approved by the Ethics and Research Committee of the Faculty of Veterinary Medicine, Aristotle University of Thessaloniki, Greece. Permits for access and use of the commercial farms were granted by the farm owners, who were members of the Chios Sheep Breeders' Cooperative "Macedonia". During sampling, animals were handled by qualified veterinarians. Permission to qualified veterinarians to perform milk and blood sampling was granted by the National (Greek) Legislature for the Veterinary Profession, No. 344/29-12-2000.

Animals, sampling and phenotyping

Animals used in the present study included 609 purebred Chios dairy ewes raised in four commercial farms in Greece. Complete pedigree data were available comprising a total of 38,459 animals, 1,892 sires and 20,634 dams. Ewes were in their first or second lactation. Daily milk yield was recorded on each animal on the day of monthly visits to the farms during the first five months of lactation. The first milk yield record was obtained at least three days after lamb weaning (ca. 42 days post lambing), which signals the onset of lactation. The total number of individual animal records was 2,436. Animal records for clinical mastitis occurrence (CM) and three mastitis indicator traits (milk somatic cell count (SCC), California Mastitis Test (CMT) score and total viable bacterial count (TVC) in milk) were also collected at the time of these visits by a qualified veterinarian. On the day of visit, the presence or absence (0/1) of CM was recorded and two 50 ml milk samples were collected in the milking parlour under aseptic conditions for the measurement of CMT, SCC and TVC. CMT was scored on a scale from 0 to 4, with high values indicating the presence of elevated SCC and, potentially, pathogens in milk; this test was performed with a commercial kit according to manufacturer's instructions (Bovivet, Kruuse, Germany). SCC was measured with Fossomatic 360 (Foss Electric, Hillerød, Denmark) and expressed as the number of cells/ml of milk. TVC was measured with Bactoscan FC 50 (Foss Electric, Hillerød, Denmark) and expressed as the number of viable bacteria/ml of milk. The three mastitis indicator traits, CMT, SCC and TVC, may capture subclinical mastitis incidences and reflect the general health status of the udder. Peripheral blood samples were taken from each ewe in 9 ml K_2EDTA Vacutainer blood collection tubes (BD diagnostics) by jugular venepuncture for genomic DNA extraction.

Genetic parameter estimation

Genetic parameters for milk yield were estimated using the following basic mixed model:

$$Y_{ijkmno} = \mu + F_i + YS_j + a_1 \cdot age + L_k + \sum_{n=1}^{2} b_n P_n W_m + g_o + pe_o + e_{ijkmno} \qquad (1)$$

Where: Y = record of ewe o in week of lactation m

μ = overall mean

F = fixed effect of flock (farm) i

YS = fixed effect of year-season of lambing j

α_1 = linear regression on age at lambing (*age*)

L = fixed effect of lactation number k

W = fixed effect of week of lactation (i.e. week post-lambing) m

b_n = fixed regression coefficient on week of lactation m (order n = 2)

P_n = orthogonal polynomial of week m (order n = 2)

g = random additive genetic effect of ewe o, including pedigree genetic relationships among animals

pe = random permanent environment effect of ewe o

e = random residual effect

Heritability and repeatability estimates were derived from the variance components calculated for the random effects in model (1). In a separate analysis, the additive genetic and permanent environment effects in model (1) were replaced by interactions of the latter with second-order polynomial functions of week of lactation. The choice of polynomial order was decided after testing sequentially increasing orders with the log-likelihood test. This analysis resulted in distinct variance component and genetic parameter estimates by week of lactation, which were then combined to derive average heritability and repeatability estimates for early (weeks of lactation 1–7), mid (weeks 8–17) and late (weeks 18–24) lactation. In addition, genetic correlations between milk yields measured at different lactation stages were calculated based on corresponding genetic covariance estimates. A smoothed lactation curve adjusted for all fixed effects in the model was also derived.

Finally, bivariate analyses of milk yield and each one of the four mastitis related traits were conducted using model (1). The four mastitis traits were analysed as described in [10]. Briefly, SCC and TVC data, which were originally significantly skewed, were log-transformed to ensure a normal distribution. CM was recorded as a 0/1 trait and, therefore, a logit function was fitted to account for its binary nature. Outcomes from the bivariate analyses were used to estimate phenotypic and genetic correlations between traits.

All statistical analyses in the present study were conducted with ASReml v4.0 [14].

Targeted genomic association analysis

DNA was extracted from blood buffy coat as described previously [15].

All animals were genotyped with a customised mastitis specific 960 SNP DNA array containing SNPs located on chromosomes 2, 3, 5, 12, 16 and 19. Briefly, this array was built based on QTLs for mastitis resistance found to segregate in multiple different dairy sheep breeds. For the design of this custom-made array, SNPs were selected from both 50K and 800K SNP ovine DNA arrays, as well as from re-sequencing data. The average density of the array was 1 SNP every 23 Kb (for more details see [10]). This genomic tool was built within an FP7 European research project (http://cordis.europa.eu/result/rcn/163471_en.html). Genotypes at each SNP locus were subjected to quality control measures using PLINK v1.9[16]. After excluding SNPs with minor allele frequency < 0.05 and/or call rate < 0.95%, 731 SNP markers remained for further analysis.

Possible population stratification was investigated with the use of the genomic relationship matrix among individual animals. This matrix was converted to a distance matrix that was then used to conduct multidimensional scaling analysis using the R package GenABEL v1.8[17].

Individual ewe phenotypes were residuals resulted after fitting a model that included all fixed effects of model (1); thus, phenotypic records were adjusted for all these environmental effects. Separate phenotypes were derived for the entire lactation (overall) and for each lactation stage (early, mid, late) as described above. In all cases, GEMMA v0.94.1 [18] was used to conduct genomic association analyses based on a mixed model that included the genomic relationship matrix among individual ewes as a polygenic effect. After Bonferroni correction for multiple testing, the significance threshold for nominal $P<0.05$ was set at $P<6.83 \times 10^{-5}$ and a suggestive threshold (accounting for one false positive per genome scan) was set at $P<1.36 \times 10^{-3}$.

Statistically significant SNPs from the genomic association analyses were further examined with a mixed model that included the fixed effects of model (1), the fixed effect of the SNP

genotype and the random effect of the animal including the pedigree relationship matrix. Additive (a) and dominance (d) effects were calculated as follows:

a = (AA-BB)/2

d = AB-((AA+BB)/2)

where AA, BB and AB were the marginal means of the respective genotype. All analyses were conducted with ASReml v4.0 [14].

Linkage disequilibrium (LD) among significant SNPs was calculated based on the r^2 value using PLINK v1.9 [16]. Blocks of LD in regions harbouring significant SNPs were visualised using Haploview v4.2 [19].

All significant (post-Bonferroni correction) and suggestive SNPs identified in the genomic analysis for milk yield were mapped to the reference genome and annotated using the Ensembl variant effect predictor (http://www.ensembl.org/Tools/VEP) tool and the Oar v3.1 assembly. Moreover, annotations for genes located both up- and down-stream (0.2 Mb) of the significant markers in the candidate regions for milk yield were obtained from Ensembl BioMart data mining tool (http://www.ensembl.org/biomart/martview/) and the Oar v3.1 assembly.

Pathway analysis

The list of annotated genes located within the QTL regions for milk yield identified in the present study were analysed with the Ingenuity Pathway Analysis (IPA) programme (www.ingenuity.com) in order to identify canonical pathways and gene networks constructed by the products of these genes. All genes located in the genomic regions targeted by the custom-made DNA array used in our study constituted the background of this analysis. IPA constructs multiple possible upstream regulators, pathways and networks which may be associated with the biological mechanism underlying the studied trait. The analysis is based on data from large-scale causal networks derived from the Ingenuity Knowledge Base. IPA then infers the most suitable pathways and networks based on their statistical significance, after correcting for a baseline threshold [20]. The IPA score in the constructed networks can be used to rank these networks based on the P-values obtained using Fisher's exact test (IPA score or P-score = $-\log_{10}$(P value)).

Gene expression analysis

We performed gene expression analyses to assess whether genes located within the candidate regions for milk yield were enriched in tissues relevant to milk yield and/or mastitis, assuming enrichment indicated functional relevance. Genes contributing to milk production are likely to be expressed in milk somatic cells, mammary gland, and other organs such as the liver and kidney that provide nutrients and regulate the electrolytes needed for lactosynthesis and the production of milk. We also reasoned that the expression of genes with pleiotropic effects would be associated with both milk yield and resistance to mastitis, and/or expressed in both mammary gland and immune related tissues. To assess the expression profiles of genes located in the candidate regions for milk yield, we obtained publicly available data from an RNA-seq characterisation of the milk transcriptome of two Spanish dairy sheep breeds, Churra and Assaf, where milk somatic cells of eight individual sheep (four from each breed) had been sampled throughout lactation at 10, 50, 120 and 150 days after lambing [21, 22]. To supplement this data, we used publicly available RNA-Seq data from a high-resolution atlas of gene expression across tissues and cell types from all major organ systems in sheep [23, 24]. The sheep gene expression atlas, which includes 437 RNA-Seq libraries was produced using six Texel x Scottish Blackface sheep [23]. An additional 83 RNA-Seq libraries from a Texel trio (ewe, lamb

and ram) were included in the sheep gene expression atlas [24]. We extracted data pertaining to the mammary gland, liver and kidney. Since we were interested in detecting genes related to both milk yield and mastitis, we also extracted the expression level of the genes under consideration in immune-related tissues, specifically hemolymph nodes, mesenteric, popliteal, prescapular and submandibular lymph nodes, peripheral blood mononuclear cells, blood leukocytes, monocyte-derived macrophages, bone marrow derived macrophages, alveolar macrophages, and tonsils.

Expression levels for all samples, were estimated using Kallisto v0.42.4 [25]. To reduce batch effects when combining data from different sources, particularly those employing different RNA selection methods [26], expression was quantified using a set of transcripts constituting a standardised transcriptomic space, as described in [27] and [28]. Expression was reported for each protein-coding transcript as the number of transcripts per million, and then summarised to the gene-level (as in [29]). Heatmaps were drawn using the heatmap.2 function of the R package gplots v3.0.1, in order to demonstrate expression enrichment in the different tissues and lactation stages.

Variant calling and allelic expression imbalance analysis

Much of the genetic variation in genes that control a quantitative trait is likely to affect their transcriptional regulation. In fact, many quantitative traits associated with altered gene expression, and trait-associated loci are enriched for eQTLs [30]. If an individual is heterozygous for a *cis*-acting mutation it is expected that the two alleles of the gene will be expressed unequally causing allelic expression imbalance. Measuring the relative expression levels of two alleles using RNA-Seq may lead to the identification of *cis*-acting SNPs or haplotypes [31–34]. To identify any *cis*-QTLs affecting the genes located in the candidate regions for milk yield we obtained the raw RNA-Seq data for mammary gland tissue from three adult female Texel x Scottish Blackface sheep from the sheep gene expression atlas [23]. The aligner HISAT2 (v2.0.4) [35], was used to produce the BAM files as previously described [23]. Variants were called using BCFtools [36] mpileup (v1.4) with parameters—max-depth 1000000—min-MQ 60, followed by BCFtools call (v1.4) with parameters -m (allow multiallelic variants) and -v (variant only). The minimum MAPQ (mapping quality) score was chosen to focus on uniquely mapped reads for variant calling. The resulting VCF file contained both SNPs and indels. The exonic variants of the protein coding genes located in the milk yield candidate regions were obtained from each VCF file using the program GTF_Extract (v0.9.1) (https://github.com/fls-bioinformatics-core/GFFUtils/blob/master/docs/GTF_extract.rst) and BEDtools [37] intersect (v2.25.0) based on gene annotations from Ovis_aries.Oar_v3.1. The putative functional impact of each variant on the encoded proteins was predicted using SnpEff v4.3 [38] with the parameter–onlyProtein (only annotate protein-coding variants). BCFtools norm (v1.4) with parameter–d was used to remove duplicated VCF records that arose due to duplicated exon coordinates in the GTF file (that is, exons present in more than one transcript). Finally, VCFs from each animal were filtered to obtain only biallelic heterozygous SNPs, using BCFtools 'view' (v1.4). For the allelic expression imbalance analysis we focused on biallelic heterozygous exonic SNPs, since the non-exonic variants may signify transcriptional noise in mRNA sequencing and contribute potential errors in the analysis.

Read counts for both the reference and alternate allele were obtained using allelecounter v0.6 (https://github.com/secastel/allelecounter) with parameters—min_cov 4,—min_baseq 20 and—min_mapq 60 and—max_depth 10000. Allelic expression imbalance, per gene, was estimated using MBASED (Meta-analysis Based Allele-Specific Expression Detection) [39] with parameters isPhased = FALSE, numSim = 10^6, BPPARAM = SerialParam(). MBASED allelic

expression imbalance estimates were derived by combining information across individual heterozygous SNP within a gene. Only variants with >10 reads in either reference or alternate allele were used. We retained only those genes with Benjamin-Hochberg [40] adjusted $P < 0.05$ and major allele frequency > 0.7.

Results

Descriptive statistics

An average daily milk yield of 1,912 grams (g) was produced in the studied sheep population with a standard deviation of 713 g, a maximum of 4,597 g and a minimum of 210 g. As expected, milk yield decreased as lactation progressed [41].

All fixed effects that were included in the model of statistical analysis accounted for a significant ($P < 0.05$) proportion in milk variation. This can be exemplified by the average daily milk yield ranging from 1,787 g in lactation 1 (368 ewes) to 2,134 g in lactation 2 (241 ewes). Including these significant sources of systematic variation in the model as fixed effects ensured the unbiasedness of the variance component estimates of the random effects and corresponding genetic parameters presented next.

Genetic parameters

Estimates of heritability and repeatability of milk yield (Table 1) were derived for the entire lactation as well as different stages of lactation defined as early, mid and late. Statistically significant ($P < 0.05$) moderate trait heritabilities (0.19–0.28) and repeatabilities (0.69–0.76) were estimated across all lactation stages. Moreover, the genetic correlations between milk yield in different lactation stages were significantly ($P < 0.05$) positive. Specifically, average genetic correlation between milk production in early and mid lactation was 0.89, early and late lactation 0.60, and mid and late lactation 0.86. The genetic correlation between early and late lactation was significantly ($P < 0.05$) lower than unity. In practical terms, lactation onset, peak lactation and lactation persistence may have partly separate genetic control.

Genetic correlations between milk and mastitis traits measured throughout the entire lactation were not significantly different from zero ($P > 0.05$). Negative phenotypic correlations were observed between these traits ($P < 0.05$), indicative of favourable environmental effects to both production and health (Table 2).

Targeted genomic association analysis

Separate targeted genomic association analyses were conducted for milk yield in early, mid, late and overall lactation. Multidimensional scaling analysis of the studied population revealed no substructure. In general, similar genomic associations were detected for milk yield in middle, late and overall lactation but distinct associations were observed in early lactation. We identified a statistically significant association after Bonferroni correction for multiple testing on chromosome 19 ($P = 1.28 \times 10^{-5}$) and three suggestive associations on chromosomes 2 ($P = 2.27 \times 10^{-4}$), 12 ($P = 3.35 \times 10^{-4}$) and 16 ($P = 6.03 \times 10{-4}$). Details of SNPs associated with

Table 1. Heritability (h^2) and repeatability (r) estimates of daily milk yield in Chios sheep by lactation stage and across the entire lactation; standard errors in parentheses.

Parameter	Early lactation (1–7 weeks)	Mid lactation (8–17 weeks)	Late lactation (18–24 weeks)	Overall lactation
h^2	0.28 (0.06)	0.19 (0.06)	0.23 (0.06)	0.23 (0.06)
r	0.76 (0.02)	0.69 (0.02)	0.71 (0.02)	0.71 (0.02)

Table 2. Estimates of phenotypic and genetic correlations between milk yield and four mastitis traits measured throughout the entire lactation in Chios sheep; standard errors in parentheses.

Mastitis trait	Phenotypic correlation	Genetic correlation
SCC	-0.18 (0.04)*	-0.12 (0.14)
CMT	-0.18 (0.04)*	-0.12 (0.13)
TVC	-0.10 (0.03)*	-0.11 (0.14)
CM	-0.07 (0.04)	-0.09 (0.19)

SCC: milk somatic cell count, CMT: California Mastitis Test score, TVC: total bacterial count in milk, CM: clinical mastitis occurrence

*Significantly different from zero ($P<0.05$)

https://doi.org/10.1371/journal.pone.0214346.t002

milk yield are shown in Table 3. Manhattan plots and corresponding Q-Q plots displaying genomic association results are shown in Fig 1 and Fig 2, respectively.

The significance of the above SNP markers was confirmed in mixed model analyses based on the pedigree genetic relationship matrix. The additive and dominance genetic effects of each SNP in the corresponding lactation stage, are summarised in Table 3. Most SNPs had a significant additive effect and a few a significant dominance effect on milk yield. When located on the same chromosomes, the significant markers identified for milk yield were in linkage disequilibrium (LD measured as $r^2 = 0.27–0.97$), implying that they correspond to the same causative mutation (S1 Table). Most importantly, the significant SNPs identified in the present study were not in LD with the SNPs previously associated with the mastitis related traits in Chios sheep [10] (S1 Table). Only small (less than 200kb length) LD blocks were visualised with Haploview, probably due to a high number of recombination events having taken place in the studied outbred population. All significant SNP markers were located in intergenic or intronic regions. The candidate QTL regions for milk yield contained a relatively small number of protein-coding genes (n = 13), microRNAs (n = 3) and non-coding RNAs (n = 3) (S2 Table).

Pathway analysis

One significant (IPA score = 34) network of molecular interactions related with organ development, organismal development and embryonic development was constructed from the genes located in the candidate regions for milk yield (Fig 3).

Gene expression analysis

Six of the genes located in the candidate regions for milk yield (*DNAJA1*, *GHR*, *LYPLAL1*, *NUP35*, *OXCT1* and *RRP15)* were expressed in either the milk transcriptome or the mammary gland (S1–S3 Figs). The growth hormone receptor (*GHR*) and 3-oxoacid CoA transferase 1 (*OXCT1*) genes were highly expressed in liver and kidney cortex tissue, respectively (S2 Fig). Moreover, *OXCT1* and *NUP35*, detected in tissues related to milk production, were also enriched in immune related tissues, relative to the other tissues analysed (S3 Fig).

Allelic expression imbalance analysis

Exonic SNP and indels were observed in all the protein coding genes located in the candidate regions for milk yield. Missense variants were identified in several genes including *CNTN4*, *DNAJC1*, *GHR*, *NUP35* and *RRP15*. One-sampled MBASED analysis identified only one gene *RRP15* (*P = 3e-03*) with significant allelic expression imbalance. Specifically, two synonymous

Table 3. List of Single Nucleotide Polymorphisms (SNPs) associated with milk yield in Chios sheep.

Lactation stage	SNP	Chr (position)	P-value	Add(P-value)	Dom(P-value)	p	q
Early **1–7 weeks**	OAR12_23075585	12(20050780)	3.35E-04	0.07(0.09)	0.07(0.10)	0.62	0.38
	oar3_OAR12_19689222	12(19689222)	3.65E-04	0.05(0.36)	0.12(0.03)	0.73	0.27
	oar3_OAR12_19269103	12(19269103)	4.66E-04	-0.03(0.33)	0.05(0.02)	0.73	0.27
	oar3_OAR12_19500329	12(19500329)	6.98E-04	0.08(0.02)	0.02(0.64)	0.6	0.4
	oar3_OAR12_19624437	12(19624437)	6.77E-04	0.05(0.30)	0.09(0.10)	0.72	0.28
	oar3_OAR12_19840123	12(19840123)	9.80E-04	0.04(0.30)	0.08(0.10)	0.68	0.32
	oar3_OAR16_33078067	16(33078067)	6.03E-04	-0.09(0.00)	0.20(0.05)	0.96	0.04
Middle **8–17 weeks**	OAR19_25259444	19(23804520)	7.03E-05	-0.15(0.00)	0.05(0.25)	0.48	0.52
	oar3_OAR19_24119431	19(24119431)	9.03E-04	-0.14(0.00)	-0.02(0.53)	0.48	0.52
	OAR19_25513179	19(24010793)	1.70E-03	-0.15(0.00)	-0.01(0.77)	0.58	0.42
	OAR16_34906481	16(32156238)	1.29E-03	-0.08(0.02)	0.07(0.05)	0.90	0.10
	OAR2_133418483	2(125230366)	1.45E-03	0.08(0.51)	-0.10(0.44)	0.92	0.08
	OAR2_133088440	2(124907852)	-2.27E-04	0.23(0.00)	0.19(0.03)	0.82	0.18
	oar3_OAR2_124936445	2(124936445)	1.11E-03	0.12(0.06)	0.07(0.30)	0.78	0.22
Late **18–24 weeks**	OAR19_25259444	19(23804520)	1.38E-04	-0.15(0.00)	-0.02(0.57)	0.48	0.52
	oar3_OAR19_24745933	19(24745933)	2.17E-04	0.10(0.00)	-0.10(0.00)	0.54	0.46
	OAR19_25830151	19(24342061)	1.35E-03	0.07(0.07)	-0.00(0.87)	0.72	0.28
	oar3_OAR19_24707843	19(24707843)	7.83E-04	-0.13(0.00)	-0.09(0.03)	0.64	0.36
	oar3_OAR19_23656789	19(23656789)	6.25E-04	-0.11(0.00)	-0.06(0.11)	0.52	0.48
	oar3_OAR2_124936445	2(124936445)	5.53E-04	0.11(0.02)	-0.00(0.93)	0.78	0.22
	OAR2_133088440	2(124907852)	4.15E-04	0.19(0.00)	0.09(0.18)	0.82	0.18
Overall	OAR19_25259444	19(23804520)	1.28E-05	-0.14(0.00)	-0.00(0.84)	0.48	0.52
	oar3_OAR19_24032312	19(24032312)	2.70E-04	-0.11(0.00)	-0.08(0.03)	0.6	0.4
	oar3_OAR19_24707843	19(24707843)	2.90E-04	-0.12(0.00)	-0.07(0.04)	0.64	0.36
	OAR19_25513179	19(24010793)	4.16E-04	-0.11(0.00)	-0.06(0.07)	0.58	0.42
	oar3_OAR19_24119431	19(24119431)	4.79E-04	-0.11(0.00)	-0.09(0.01)	0.48	0.52
	oar3_OAR19_23929524	19(23929524)	5.62E-04	0.11(0.00)	-0.03(0.29)	0.48	0.52
	oar3_OAR19_24745933	19(24745933)	9.35E-04	0.10(0.00)	-0.05(0.15)	0.54	0.46
	oar3_OAR19_23891277	19(23891277)	1.17E-03	-0.10(0.00)	-0.05(0.13)	0.54	0.46
	oar3_OAR19_23656789	19(23656789)	1.44E-03	-0.10(0.00)	-0.05(0.15)	0.52	0.48
	OAR2_133088440	2(124907852)	9.47E-04	0.20(0.00)	0.12(0.06)	0.82	0.18
	OAR2_133418483	2(125230366)	1.44E-03	0.10(0.30)	-0.08(0.43)	0.92	0.08

P–value: P-value from genomic association study; additive allele substitution effect (ADD) and corresponding P-value; dominance effect (DOM) and corresponding P-value; p and q allelic frequencies; SNP position is based on Oar v3.1 assembly.

SNPs in *RRP15* (major allele frequency 0.71) were detected exhibiting allelic expression imbalance (S3 Table). However, these results should be interpreted with caution since allelic expression imbalance was evident in only one of the three individual sheep.

Candidate genes

Based on all above results, a total of four genes (*DNAJA1*, *GHR*, *LYPLA1* and *OXCT1*) were selected as putative candidate genes for milk yield in Chios sheep (S4 Table). Genes were selected using a combination of their known biological function, involvement in relevant networks, enrichment in tissues relevant to milk production, and any previously known association with milk production in either dairy sheep or other species.
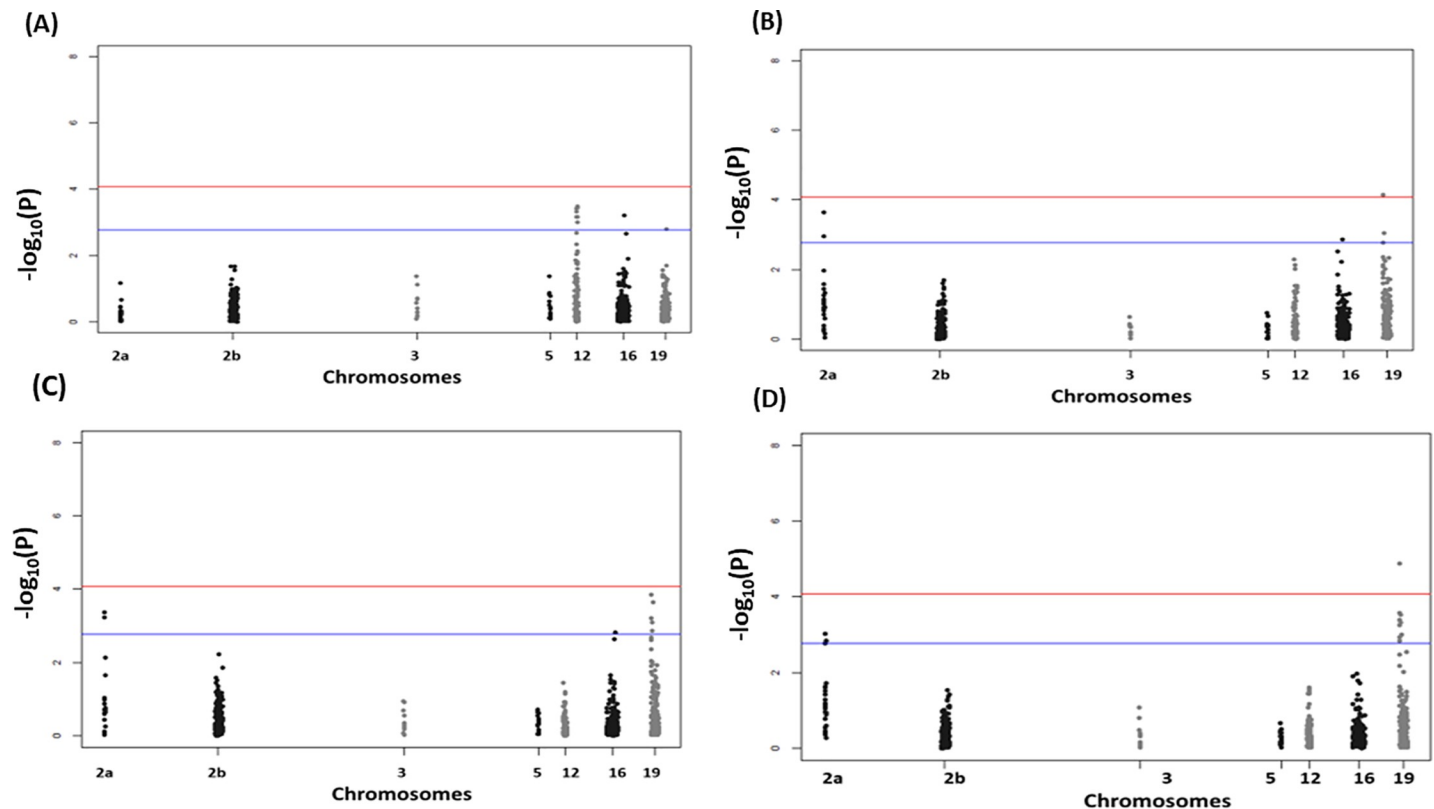
Fig 1. Manhattan plots displaying the genomic association results for milk yield in Chios sheep. Manhattan plots for milk yield in early (A), mid (B), late (C), and overall (D) lactation. Genomic location is plotted against $-\log_{10}(P)$. Red and blue lines, respectively, are thresholds for significance post-Bonferroni correction ($P<0.05$) and for suggestive significance (accounting for one false positive per genome scan).

https://doi.org/10.1371/journal.pone.0214346.g001

Discussion

The existence of a mastitis-specific ovine DNA array built on previously detected QTL regions associated with mastitis resistance in dairy sheep opens up opportunities for targeted genomic and marker-assisted selection aiming to enhance animal resistance to the disease. The aim of the present study was to investigate the association of this array with milk yield of dairy sheep and assess the feasibility of a concomitant genetic improvement programme for the two traits.

According to our results, milk yield and mastitis traits in the Chios sheep are not genetically correlated to each other. Genetic correlation estimates between milk somatic cell count and milk yield are reportedly antagonistic in dairy cattle [42] but inconsistent amongst previous sheep studies ranging from antagonistic [43] to favourable [3]. Here we considered more than 600 carefully monitored and densely phenotyped individual animals, and more than 38,000 pedigrees. We believe the genetic correlation estimates derived, ranging from -0.09 to -0.12 (Table 2), are unbiased. Even if we had a larger dataset available, rendering the standard errors small enough to qualify these estimates as significant, the practical implications would not really change; estimates would still demonstrate a very weak connection between traits. Indeed, an absolute correlation of 0.09–0.12 suggests that a very small proportion of the variation in two traits is common. Therefore, our findings indicate that selection for enhanced mastitis resistance could be incorporated into the current genetic improvement programme of the Chios sheep without incurring adverse effects on milk yield.
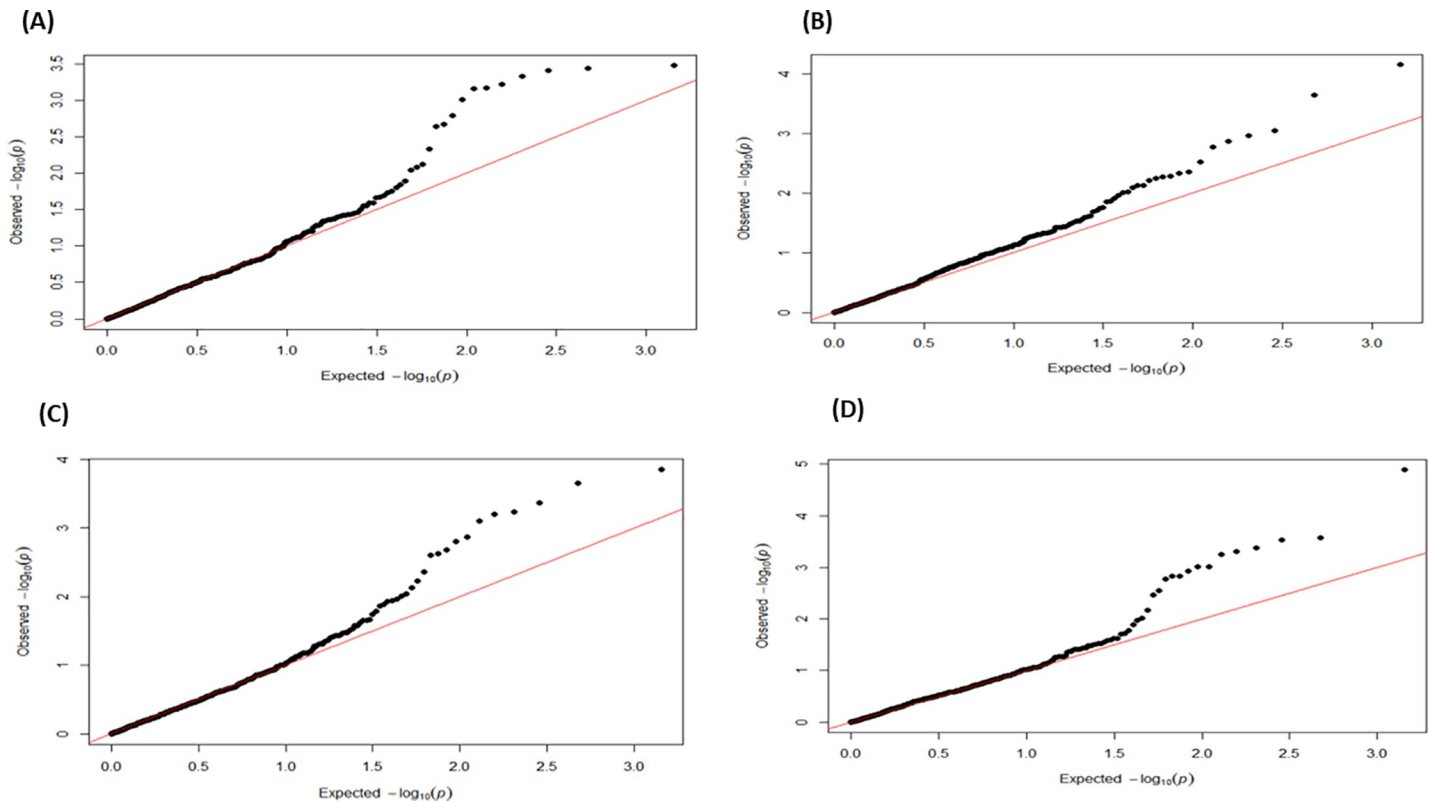
Fig 2. Q-Q plots displaying the genomic association results for milk yield in Chios sheep. Q-Q plots in early (A), mid (B), late (C) and overall (D) lactation; observed *P*-values are plotted against the expected *P*-values.

An overall moderate but significant heritability for milk yield was estimated in Chios sheep, consistent with the dairy sheep literature (ranging from 0.16 to 0.30) as reviewed in [44] and previous studies in Chios sheep ranging from 0.21 to 0.29 [45].

Targeted genomic analyses were conducted to further investigate the underlying correlation between milk yield and mastitis, in the context of utilising a mastitis-specific DNA array in genomic selection aiming to improve mastitis resistance. These analyses revealed several SNPs on the mastitis array with a significant effect on milk yield. However, these milk-associated SNPs were not overlapping or being in LD with genomic regions that had been previously found to affect mastitis resistance in the same population [10]. For example, the QTL for milk yield on chromosome 2 was 75 Mb distant from the one previously identified for mastitis resistance on the same chromosome [10]. The association of this QTL region with milk yield is supported by results of a previous genomic selection mapping study that compared dairy with meat sheep breeds to identify genomic regions for milk traits under selection [46]. In that study a highly homozygous region was detected in both Chios and Churra sheep in close proximity with our QTL region on chromosome 2 [46]. Furthermore, the QTL for milk yield on chromosome 12 identified in the present study was located within a previously identified QTL region for milk yield in East Friesian X Dorset cross sheep [47]. The QTLs on chromosomes 16 and 19 identified in the present study were also independent from those previously identified for mastitis resistance on the same chromosomes in the Chios sheep; the latter were located 2–4 Mb away and were in zero LD with the milk-associated region of the present study. QTLs for milk yield, milk protein and fat content have also been identified on chromosome 16 in Churra sheep [48], in close proximity with the
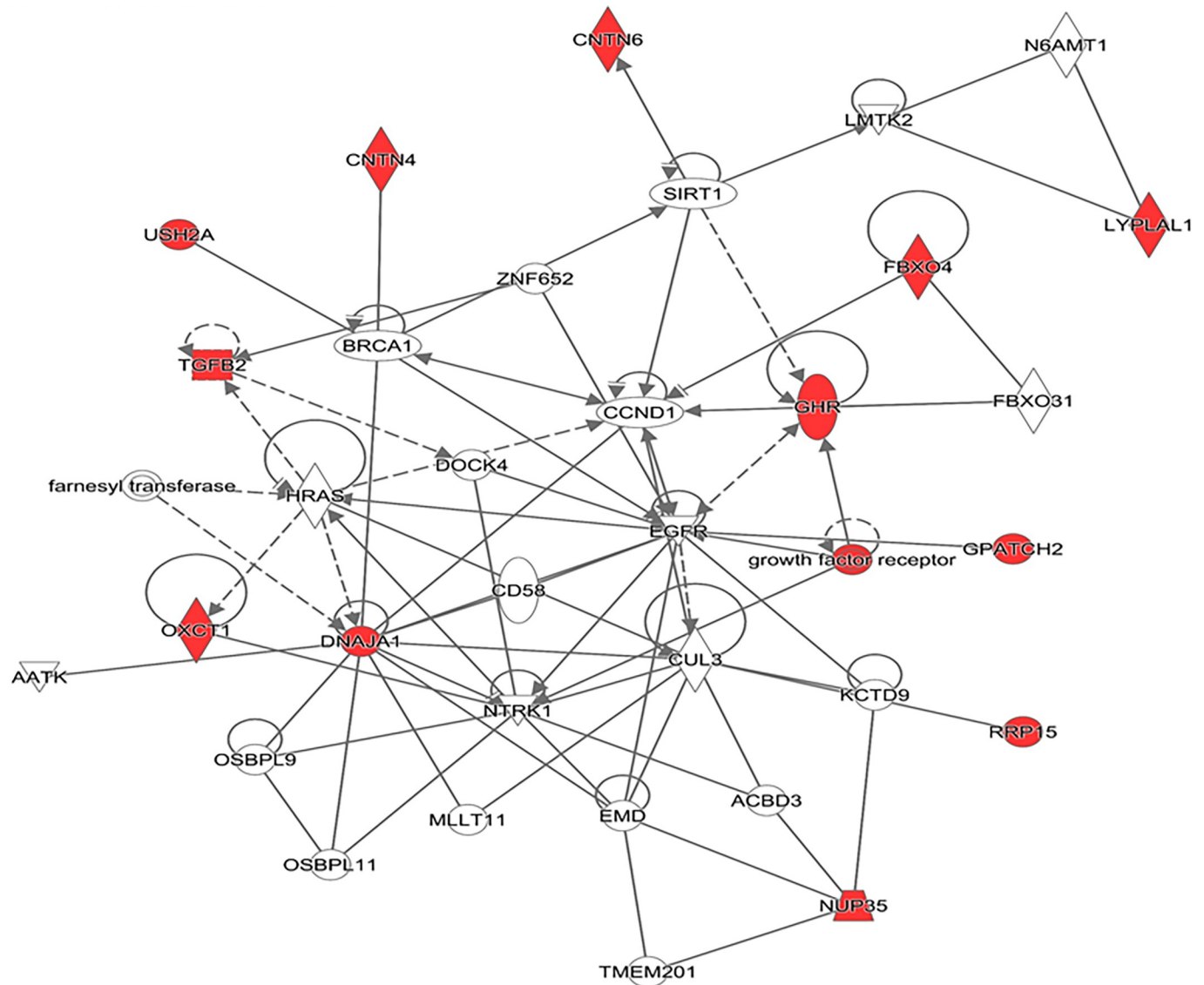
Fig 3. Network analysis using the IPA software. A gene network related to organ, organismal and embryonic development. The network illustrates the molecular interactions between candidate gene products for milk yield. Arrows with solid lines represent direct interactions and arrows with broken lines represent indirect interactions. Genes with white labels are those added to the IPA analysis because of their interaction with the target gene products (in red).

QTL identified here in Chios sheep. To the best of our knowledge, the QTL on chromosome 19 is reported here for the first time.

In the QTL region identified for milk yield on chromosome 19 *DNAJA1* was identified as a good candidate gene. In the previous milk transcriptome study of the Churra and Assaf breeds, two other genes belonging to the same gene family, *DNAJA4* and *DNAJB2*, were reported as functional candidates for milk yield [49]. The *DNAJ* family of proteins regulate ATP hydrolysis activity, and facilitate protein folding, trafficking, prevention of aggregation and proteolytic degradation; *DNAJA1* functions as a co-chaperone and protects cells against apoptosis in response to cellular stress [50]. Therefore, this gene might affect milk yield through both metabolism and mammary apoptosis; the latter has been associated negatively with lactation persistency (daily milk yield decline in late lactation stages) in dairy species [51].

Some of the candidate genes for milk yield identified in the present study have been previously reported in dairy cattle. For example, 3-oxoacid CoA transferase 1 *(OXCT1)* has been associated favourably with both milk production [52] and mastitis resistance [53], and has been suggested to regulate mammary gland metabolism and milk synthesis during mastitis infection [54]. In our study, *OXCT1* was found to be expressed in both mammary gland and immune tissues, and highly expressed in the kidney cortex indicating that it may play a similar role in sheep. Growth hormone receptor (*GHR*) has been previously associated with increased milk yield and reduced milk somatic cell count in several dairy cattle studies [54–58]. Selective sweeps were also identified in the *GHR* region after comparing dairy and beef cattle [59]. In the present study, *GHR* was expressed in the mammary gland and the milk transcriptome, and was highly expressed in liver, relative to the other tissues sampled for the sheep gene expression atlas (http://biogps.org/sheepatlas). However, further studies, preferably including animals of the Chios breed, are needed to confirm the relevance of these genes with the regulation of milk production.

The significant SNP markers identified for milk yield in our study are mostly located in QTLs that overlap with previously identified QTLs for milk yield in other dairy sheep populations [55–57]. The only QTL reported here for the first time is on chromosome 19, which attained the highest significance level in the present study. These results are also consistent with a previous study of Chios sheep [60], suggesting that a relatively major locus might be involved in ovine milk production. Nevertheless, these SNPs were not associated with any of the studied mastitis traits.

Conclusions

Results of the present study suggest that genetic selection for enhanced host resistance to mastitis will not antagonise milk yield in Chios sheep. Therefore, a genetic improvement programme for enhancing both mastitis resistance and milk production is feasible for this breed. In addition, there are QTLs within the mastitis specific DNA array that may be used to further increase milk production with genomic selection. Genes within genomic regions associated with ovine milk production exhibited tissue-specific expression patterns and pathways similar to those observed in cattle indicating that the underlying genetic mechanisms are likely to be, at least partially, conserved between the two species. These genes are suitable candidates for further investigation to determine if they can be exploited in breeding programmes for concomitant improvement of milk production and mastitis resistance. Admittedly, the detection of QTLs for milk yield was performed using a targeted SNP panel and not a genome-wide array; therefore our scan was very focussed and major loci associated with milk production in Chios sheep might not have been detected. Further studies using genome-wide DNA arrays are needed to identify novel QTLs for milk yield.

Supporting information

S1 Table. Linkage disequilibrium (LD) estimates (expressed as r ^{2}) for the significant SNP markers identified in the genomic association analyses of milk yield and mastitis resistance in Chios sheep.
(XLSX)

S2 Table. Genes located in the candidate genomic regions identified for milk yield in Chios sheep.
(XLSX)

S3 Table. Allelic expression imbalance analysis using the one-sampled MBASED method.
(XLSX)

S4 Table. Selected candidate genes for milk yield in Chios sheep.
(XLSX)

S1 Fig. Expression level of genes located in the milk yield candidate regions, as extracted from the Churra and Assaf sheep milk transcriptome analysis. Expression level is estimated as the mean number of transcripts per million of all (5) experimental replicates and is represented here as a z-score per individual animal.
(TIF)

S2 Fig. Expression level of genes located in the milk yield candidate regions, across all cell lines/tissues. Expression level is estimated as the mean number of transcripts per million (TPM) of all five (5) experimental replicates and is represented here as a z-score per cell line/tissue. Data is obtained from the sheep gene expression atlas which includes data from Texel X Scottish Blackface and Texel sheep.
(TIF)

S3 Fig. Expression level of genes, located in the milk yield candidate regions, across both mammary gland and immune cell lines/tissues. Expression level is estimated as the mean number of transcripts per million of all five (5) experimental replicates and is represented here as a z-score per cell line/tissue. Data is obtained from the sheep gene expression atlas which includes data from Texel X Scottish Blackface and Texel sheep.
(TIF)

Acknowledgments

Author Contributions

Conceptualization: Georgios Banos, Georgios Arsenos, David A. Hume, Androniki Psifidi.

Data curation: Georgios Banos, Emily L. Clark, Stephen J. Bush, Georgios Bramis, Androniki Psifidi.

Formal analysis: Emily L. Clark, Stephen J. Bush, Prasun Dutta, Georgios Bramis, Androniki Psifidi.

Funding acquisition: Georgios Banos, Georgios Arsenos, David A. Hume, Androniki Psifidi.

Supervision: Georgios Banos.

Writing – original draft: Georgios Banos, Androniki Psifidi.

Writing – review & editing: Emily L. Clark, Prasun Dutta, Georgios Bramis, Georgios Arsenos, David A. Hume.

References

1. Miltiadou D, Hager-Theodorides AL, Symeou S, Constantinou C, Psifidi A, Banos G, et al. Variants in the 3′ untranslated region of the ovine acetyl-coenzyme A acyltransferase 2 gene are associated with dairy traits and exhibit differential allelic expression. Journal of Dairy Science. 2017; 100(8):6285–97. https://doi.org/10.3168/jds.2016-12326 PMID: 28624287

2. Barillet F. Genetic improvement for dairy production in sheep and goats. Small Ruminant Research. 2007; 70(1):60–75. http://dx.doi.org/10.1016/j.smallrumres.2007.01.004.

3. Legarra A, Ugarte E. Genetic Parameters of Udder Traits, Somatic Cell Score, and Milk Yield in Latxa Sheep. Journal of Dairy Science. 2005; 88(6):2238–45. https://doi.org/10.3168/jds.S0022-0302(05)72899-X PMID: 15905453

4. Baro J, San Primitivo F, Facultad De Veterinaria I, Spa L. Breeding programme for the Spanish Churra sheep breed1995.

5. Mavrogenis AP, Papachristoforou C. Genetic and phenotypic relationships between milk production and body weight in Chios sheep and Damascus goats. Livestock Production Science. 2000; 67(1/2):81–7. https://doi.org/10.1016/S0301-6226(00)00187-1

6. Davies G, Genini S, Bishop SC, Giuffra E. An assessment of opportunities to dissect host genetic variation in resistance to infectious diseases in livestock. Animal: an international journal of animal bioscience. 2009; 3(3):415–36. Epub 2009/03/01. https://doi.org/10.1017/s1751731108003522 PMID: 22444313.

7. Bishop SC, Axford RFE, Nicholas FW, Owen JB. Breeding for disease resistance in farm animals: CABI Publishing; 2010.

8. Merz A, Stephan R, Johler S. Staphylococcus aureus Isolates from Goat and Sheep Milk Seem to Be Closely Related and Differ from Isolates Detected from Bovine Milk. Frontiers in Microbiology. 2016; 7:319. https://doi.org/10.3389/fmicb.2016.00319 PMC4789554. PMID: 27014240

9. Authority EFS. Scientific opinion on the welfare risks related to the farming of sheep for wool, meat and milk production. EFSA J. 2014; 12:128

10. Banos G, Bramis G, Bush SJ, Clark EL, McCulloch MEB, Smith J, et al. The genomic architecture of mastitis resistance in dairy sheep. BMC Genomics. 2017; 18(1):624. Epub 2017/08/18. https://doi.org/10.1186/s12864-017-3982-1 PMID: 28814268; PubMed Central PMCID: PMC5559839.

11. Rupp R, Senin P, Sarry J, Allain C, Tasca C, Ligat L, et al. A Point Mutation in Suppressor of Cytokine Signalling 2 (Socs2) Increases the Susceptibility to Inflammation of the Mammary Gland while Associated with Higher Body Weight and Size and Higher Milk Production in a Sheep Model. PLoS genetics. 2015; 11(12):e1005629. Epub 2015/12/15. https://doi.org/10.1371/journal.pgen.1005629 PMID: 26658352; PubMed Central PMCID: PMC4676722.

12. Gutiérrez-Gil B G-GE, Suárez-Vega A., Arranz JJ Detection of QTL influencing somatic cell score in Churra sheep employing the OvineSNP50 BeadChip. EAAP, 64th Annual Meeting, Nantes 2013;http://old.eaap.org/Previous_Annual_Meetings/2013Nantes/Nantes_2013_Abstracts.pdf.

13. Sechi S CS, Casula M, Congiu GB, Miari S, Mulas G, Salaris S, et al. Genome -wide association analysis of resistance to paratuberculosis and mastitis in dairy sheep. EAAP, 64th Annual Meeting, Nantes 2013. 2013.

14. Gilmour AR, Cullis B.R. and Thompson R. ASREML User Guide, Release 3.0, NSW Department of Primary Industries, Australia. 2009.

15. Psifidi A, Dovas CI, Bramis G, Lazou T, Russel CL, Arsenos G, et al. Comparison of eleven methods for genomic DNA extraction suitable for large-scale whole-genome genotyping and long-term DNA banking using blood samples. PLoS One. 2015; 10(1):e0115960. https://doi.org/10.1371/journal.pone.0115960 PMID: 25635817

16. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3):559–75. Epub 2007/08/19. https://doi.org/10.1086/519795 PMID: 17701901; PubMed Central PMCID: PMC1950838.

17. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics (Oxford, England). 2007; 23(10):1294–6. Epub 2007/03/27. https://doi.org/10.1093/bioinformatics/btm108 PMID: 17384015.

18. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nat Methods. 2014; 11(4):407–9. Epub 2014/02/18. https://doi.org/10.1038/nmeth.2848 PMID: 24531419; PubMed Central PMCID: PMC4211878.

19. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics (Oxford, England). 2005; 21(2):263–5. Epub 2004/08/07. https://doi.org/10.1093/bioinformatics/bth457 PMID: 15297300.

20. Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics (Oxford, England). 2014; 30(4):523–30. https://doi.org/10.1093/bioinformatics/btt703 PMC3928520. PMID: 24336805

21. Suarez-Vega A, Gutierrez-Gil B, Klopp C, Tosser-Klopp G, Arranz JJ. Comprehensive RNA-Seq profiling to evaluate lactating sheep mammary gland transcriptome. Scientific data. 2016; 3:160051. Epub

2016/07/06. https://doi.org/10.1038/sdata.2016.51 PMID: 27377755; PubMed Central PMCID: PMC4932878.

22. Suarez-Vega A, Gutierrez-Gil B, Klopp C, Robert-Granie C, Tosser-Klopp G, Arranz JJ. Characterization and Comparative Analysis of the Milk Transcriptome in Two Dairy Sheep Breeds using RNA Sequencing. Scientific reports. 2015; 5:18399. Epub 2015/12/19. https://doi.org/10.1038/srep18399 PMID: 26677795; PubMed Central PMCID: PMC4683406.

23. Clark EL, Bush SJ, McCulloch MEB, Farquhar IL, Young R, Lefevre L, et al. A high resolution atlas of gene expression in the domestic sheep (Ovis aries). PLoS genetics. 2017; 13(9):e1006997. https://doi.org/10.1371/journal.pgen.1006997 PMID: 28915238

24. Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, et al. The sheep genome illuminates biology of the rumen and lipid metabolism. Science. 2014; 344(6188):1168–73. Epub 2014/06/07. https://doi.org/10.1126/science.1252806 PMID: 24904168; PubMed Central PMCID: PMC4157056.

25. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotech. 2016; 34(5):525–7. https://doi.org/10.1038/nbt.3519 http://www.nature.com/nbt/journal/v34/n5/abs/nbt.3519.html#supplementary-information. PMID: 27043002

26. Lahens NF, Kavakli IH, Zhang R, Hayer K, Black MB, Dueck H, et al. IVT-seq reveals extreme bias in RNA sequencing. Genome biology. 2014; 15(6):R86–R. https://doi.org/10.1186/gb-2014-15-6-r86 PMID: 24981968.

27. Bush SJ, Freem L, MacCallum AJ, O'Dell J, Wu C, Afrasiabi C, et al. Combination of novel and public RNA-seq datasets to generate an mRNA expression atlas for the domestic chicken. BMC genomics. 2018; 19(1):594–. https://doi.org/10.1186/s12864-018-4972-7 PMID: 30086717.

28. Bush SJ, McCulloch MEB, Summers KM, Hume DA, Clark EL. Integration of quantitated expression estimates from polyA-selected and rRNA-depleted RNA-seq libraries. BMC bioinformatics. 2017; 18 (1):301–. https://doi.org/10.1186/s12859-017-1714-9 PMID: 28610557.

29. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. 2015; 4:1521. Epub 2016/03/01. https://doi.org/10.12688/f1000research.7563.2 PMID: 26925227; PubMed Central PMCID: PMC4712774.

30. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS genetics. 2010; 6(4):e1000888. Epub 2010/04/07. https://doi.org/10.1371/journal.pgen.1000888 PMID: 20369019; PubMed Central PMCID: PMC2848547.

31. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, et al. Genetic analysis of genome-wide variation in human gene expression. Nature. 2004; 430(7001):743–7. Epub 2004/07/23. https://doi.org/10.1038/nature02797 PMID: 15269782; PubMed Central PMCID: PMC2966974.

32. Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R, et al. Genome-wide associations of gene expression variation in humans. PLoS genetics. 2005; 1(6):e78. Epub 2005/12/20. https://doi.org/10.1371/journal.pgen.0010078 PMID: 16362079; PubMed Central PMCID: PMC1315281.

33. Bray NJ, Buckland PR, Owen MJ, O'Donovan MC. Cis-acting variation in the expression of a high proportion of genes in human brain. Human genetics. 2003; 113(2):149–53. Epub 2003/05/03. https://doi.org/10.1007/s00439-003-0956-y PMID: 12728311.

34. Chamberlain AJ, Vander Jagt CJ, Hayes BJ, Khansefid M, Marett LC, Millen CA, et al. Extensive variation between tissues in allele specific expression in an outbred mammal. BMC Genomics. 2015; 16:993. Epub 2015/11/26. https://doi.org/10.1186/s12864-015-2174-0 PMID: 26596891; PubMed Central PMCID: PMC4657355.

35. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015; 12(4):357–60. Epub 2015/03/10. https://doi.org/10.1038/nmeth.3317 PMID: 25751142; PubMed Central PMCID: PMC4655817.

36. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics (Oxford, England). 2011; 27 (21):2987–93. Epub 2011/09/10. https://doi.org/10.1093/bioinformatics/btr509 PMID: 21903627; PubMed Central PMCID: PMC3198575.

37. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics (Oxford, England). 2010; 26(6):841–2. Epub 2010/01/30. https://doi.org/10.1093/bioinformatics/btq033 PMID: 20110278; PubMed Central PMCID: PMC2832824.

38. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012; 6(2):80–92. Epub 2012/06/26. https://doi.org/10.4161/fly.19695 PMID: 22728672; PubMed Central PMCID: PMC3679285.

39. Mayba O, Gilbert HN, Liu J, Haverty PM, Jhunjhunwala S, Jiang Z, et al. MBASED: allele-specific expression detection in cancer tissues and cell lines. Genome Biol. 2014; 15(8):405. Epub 2014/10/16. https://doi.org/10.1186/s13059-014-0405-3 PMID: 25315065; PubMed Central PMCID: PMC4165366.

40. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995; 57(1):289–300.

41. Hagnestam-Nielsen C, Emanuelson U, Berglund B, Strandberg E. Relationship between somatic cell count and milk yield in different stages of lactation. J Dairy Sci. 2009; 92(7):3124–33. Epub 2009/06/17. https://doi.org/10.3168/jds.2008-1719 PMID: 19528590.

42. Rupp R, Boichard D. Genetics of resistance to mastitis in dairy cattle. Veterinary research. 2003; 34 (5):671–88. Epub 2003/10/15. https://doi.org/10.1051/vetres:2003020 PMID: 14556700.

43. Tolone M, Riggio V, Portolano B. Estimation of genetic and phenotypic parameters for bacteriological status of the udder, somatic cell score, and milk yield in dairy sheep using a threshold animal model. Livestock Science. 2013; 151(2–3):134–9. https://doi.org/10.1016/j.livsci.2012.11.014

44. Juan José Arranz BGr-G. Detection of QTL Underlying Milk Traits in Sheep: An Update, Milk Production. In: Chaiyabutr PN, editor. Advanced Genetic Traits, Cellular Mechanism, Animal Management and Health: InTech; 2012.

45. Ligda C, Papadopoulos T, Mavrogenis A, Georgoudis A. Genetic parameters for test day milk traits and somatic cell counts in Chios dairy sheep. In: Gabiña D, Sanna S, editors. Breeding programmes for improving the quality and safety of products New traits, tools, rules and organization? Options Méditerranéennes: Série A. Séminaires Méditerranéens. 55: Zaragoza: CIHEAM; 2003. p. 55–9.

46. Gutierrez-Gil B, Arranz JJ, Pong-Wong R, Garcia-Gamez E, Kijas J, Wiener P. Application of selection mapping to identify genomic regions associated with dairy production in sheep. PLoS One. 2014; 9(5): e94623. Epub 2014/05/03. https://doi.org/10.1371/journal.pone.0094623 PMID: 24788864; PubMed Central PMCID: PMC4006912.

47. Mateescu RG, Thonney ML. Genetic mapping of quantitative trait loci for milk production in sheep. Anim Genet. 2010; 41(5):460–6. Epub 2010/04/17. https://doi.org/10.1111/j.1365-2052.2010.02045.x PMID: 20394603.

48. Garcia-Gamez E, Gutierrez-Gil B, Sahana G, Sanchez JP, Bayon Y, Arranz JJ. GWA analysis for milk production traits in dairy sheep and genetic support for a QTN influencing milk protein percentage in the LALBA gene. PLoS One. 2012; 7(10):e47782. Epub 2012/10/25. https://doi.org/10.1371/journal.pone.0047782 PMID: 23094085; PubMed Central PMCID: PMC3475704.

49. Suárez-Vega A, Gutiérrez-Gil B, Klopp C, Tosser-Klopp G, Arranz JJ. Variant discovery in the sheep milk transcriptome using RNA sequencing. BMC Genomics. 2017; 18(1):170. https://doi.org/10.1186/s12864-017-3581-1 PMID: 28202015

50. Gotoh T, Terada K, Oyadomari S, Mori M. hsp70-DnaJ chaperone pair prevents nitric oxide- and CHOP-induced apoptosis by inhibiting translocation of Bax to mitochondria. Cell death and differentiation. 2004; 11(4):390–402. Epub 2004/01/31. https://doi.org/10.1038/sj.cdd.4401369 PMID: 14752510.

51. Stefanon B, Colitti M, Gabai G, Knight CH, Wilde CJ. Mammary apoptosis and lactation persistency in dairy animals. The Journal of dairy research. 2002; 69(1):37–52. Epub 2002/06/06. https://doi.org/10.1017/s0022029901005246 PMID: 12047109.

52. Bionaz M, Loor JJ. Gene networks driving bovine milk fat synthesis during the lactation cycle. BMC Genomics. 2008; 9:366. Epub 2008/08/02. https://doi.org/10.1186/1471-2164-9-366 PMID: 18671863; PubMed Central PMCID: PMC2547860.

53. Zarrin M, Wellnitz O, van Dorland HA, Gross JJ, Bruckmaier RM. Hyperketonemia during lipopolysaccharide-induced mastitis affects systemic and local intramammary metabolism in dairy cows. Journal of Dairy Science. 2014; 97(6):3531–41. https://doi.org/10.3168/jds.2013-7480 PMID: 24679930

54. Tiezzi F, Parker-Gaddis KL, Cole JB, Clay JS, Maltecca C. A Genome-Wide Association Study for Clinical Mastitis in First Parity US Holstein Cows Using Single-Step Approach and Genomic Matrix Re-Weighting Procedure. PLoS ONE. 2015; 10(2):e0114919. https://doi.org/10.1371/journal.pone.0114919 PMC4319771. PMID: 25658712

55. Raven LA, Cocks BG, Goddard ME, Pryce JE, Hayes BJ. Genetic variants in mammary development, prolactin signalling and involution pathways explain considerable variation in bovine milk production and milk composition. Genet Sel Evol. 2014; 46:29. Epub 2014/05/02. https://doi.org/10.1186/1297-9686-46-29 PMID: 24779965; PubMed Central PMCID: PMC4036308.

56. Sun D, Jia J, Ma Y, Zhang Y, Wang Y, Yu Y, et al. Effects of DGAT1 and GHR on milk yield and milk composition in the Chinese dairy population. Animal Genetics. 2009; 40(6):997–1000. https://doi.org/10.1111/j.1365-2052.2009.01945.x PMID: 19781040

57. Rahmatalla SA, Muller U, Strucken EM, Reissmann M, Brockmann GA. The F279Y polymorphism of the GHR gene and its relation to milk production and somatic cell score in German Holstein dairy cattle.

Journal of applied genetics. 2011; 52(4):459–65. Epub 2011/06/11. https://doi.org/10.1007/s13353-011-0051-3 PMID: 21660490.

58. Ogorevc J, Kunej T, Razpet A, Dovc P. Database of cattle candidate genes and genetic markers for milk production and mastitis. Anim Genet. 2009; 40(6):832–51. Epub 2009/06/11. https://doi.org/10.1111/j.1365-2052.2009.01921.x PMID: 19508288; PubMed Central PMCID: PMC2779988.

59. Gutiérrez-Gil B, Arranz JJ, Wiener P. An interpretive review of selective sweep studies in Bos taurus cattle populations: identification of unique and shared selection signals across breeds. Frontiers in genetics. 2015; 6(167). https://doi.org/10.3389/fgene.2015.00167 PMID: 26029239

60. Chatziplis DG, Tzamaloukas O, Miltiadou D, Ligda C, Koumas A, Mavrogenis AP, et al. Evidence of major gene(s) affecting milk traits in the Chios sheep breed. Small Ruminant Research. 2012; 105(1–3):61–8. http://dx.doi.org/10.1016/j.smallrumres.2011.12.009.