Title: 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] and Androniki Psifidi[2,6]

[1]Scotland's Rural College, Edinburgh, Easter Bush, Midlothian, EH25 9RG, UK

[2]The Roslin Institute, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, UK

[3]School of Veterinary Medicine, Aristotle University of Thessaloniki, Thessaloniki, 54124, Greece

[4]Nuffield Department of Clinical Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, UK

[5]Mater Research Institute-University of Queensland, Translational Research Institute, Woolloongabba, Qld 4102, Australia

[6]Royal Veterinary College, University of London, Hatfield, AL9 7TA, UK

*Correspondence: Dr Androniki Psifidi

e-mail: androniki.psifidi@roslin.ed.ac.uk, apsifidi@rvc.ac.uk

Short title: Concomitant genetic improvement of milk yield and mastitis resistance in dairy sheep

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 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 array. We performed 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. Environmental factors appeared to favour both milk production and udder health. Four Quantitative Trait Loci (QTLs) affecting milk yield were detected on chromosomes 2, 12, 16 and 19, in locations distinct from those previously identified to affect mastitis resistance. Pathways, networks and functional gene clusters for milk yield were identified. Seven genes (*DNAJA1, DNAJC10, FGF10, GHR, HMGCS1, LYPLA1, 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 four genes (*DNAJC10, FGF10, OXCT1, EMB*) was enriched in immune tissues implying a favourable

43 pleiotropic effect or likely role in milk production during udder infection. In conclusion, the
44 absence of genetic antagonism between milk yield and mastitis resistance suggests that
45 simultaneous genetic improvement of both traits be achievable. The detection of milk yield
46 QTLs with the mastitis array underpins the latter's utility as a breeding tool for the genetic
47 enhancement of both traits.

48

49 Introduction

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

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

65

66 Beyond simply increasing milk production, the dairy sheep industry faces challenges such as
67 the need to offer healthy products to consumers, addressing animal welfare, and ensuring the
68 long-term competitiveness and sustainability of the sector. Mastitis is the most prevalent and
69 costly disease in the dairy industry due to reduced and discarded milk, early involuntary culling
70 of animals, and veterinary services and labour costs [6, 7]. The disease also poses a potential
71 threat of zoonosis and antimicrobial resistance if antibiotic treatment is not applied carefully
72 [6-8]. Moreover, mastitis is a welfare concern because of associated pain, anxiety and
73 restlessness, and upsets the normal feeding behaviour of the animals [9]. Host resistance to
74 mastitis is a moderately heritable trait [7]. Recently, an ovine custom made mastitis specific
75 960-SNP DNA array was built to facilitate genetic selection and improvement of animal
76 resistance to mastitis in dairy sheep [10] [11] [12] [13]. We previously used this array in a
77 genomic association study and detected five quantitative trait loci (QTLs) for mastitis
78 resistance in Chios sheep [10].

79

80 In the present study, we examined the genetic and genomic relationship between milk yield
81 and mastitis resistance in the Chios sheep, using pedigree and genomic information. The
82 relationship between the two traits is crucial if mastitis resistance is to be included in selective
83 breeding goals together with increased milk yield. We estimated genetic parameters and
84 investigated whether relevant QTLs for milk yield exist in previously identified mastitis-
85 specific genomic regions. We also performed pathway analysis and examined gene expression
86 and allelic expression imbalance to characterise the genes located under the QTL regions in
87 relation with milk yield and mastitis resistance.

88

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 for these animals. 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). Total number of individual animal records collected amounted to 2,436. Animal records for clinical mastitis occurrence and three mastitis indicator traits (milk somatic cell count, California Mastitis Test score and total viable bacterial count in milk) were also collected at the time of these visits by a qualified veterinarian, as described previously [10]. The three mastitis indicator traits 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

112 collection tubes (BD diagnostics) by jugular venepuncture for genomic DNA extraction.

113

114 Genetic parameter estimation

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

116 $$Y_{ijkmno} = \mu + F_i + YS_j + a_1 \cdot age + L_k + \sum_{n=1}^{2} b_n P_n W_m + g_o + pe_o + e_{ijkmno} \qquad (1)$$

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

118 μ = overall mean

119 *F* = fixed effect of flock (farm) *i*

120 *YS* = fixed effect of year-season of lambing *j*

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

122 *L* = fixed effect of lactation number *k*

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

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

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

126 *g* = random additive genetic effect of ewe *o*, including pedigree genetic relationships

127 among animals

128 *pe* = random permanent environment effect of ewe *o*

129 *e* = random residual effect

130

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); outcomes from these 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].

Genomic association studies

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

153 design of this custom-made array, SNPs were selected from 50K and 800K SNP ovine DNA
154 arrays, as well as available re-sequencing data. The average density of the array was 1 SNP
155 every 23 Kb (for more details see [10]). This genomic tool was built within an FP7 European
156 research project (http://cordis.europa.eu/result/rcn/163471_en.html). Genotypes at each SNP
157 locus were subjected to quality control measures using the following thresholds: minor allele
158 frequency >0.05, call rate >95% and Hardy-Weinberg equilibrium $P>10^{-4}$. After quality control,
159 710 SNP markers remained for further analysis.

160

161 Possible population stratification was investigated with the use of the genomic relationship
162 matrix among individual animals. This matrix was converted to a distance matrix that was then
163 used to conduct multidimensional scaling analysis using the GenABEL package of R [16].

164

165 Individual ewe phenotypes were residuals resulted after fitting a model that included all fixed
166 effects of model (1); thus, phenotypic records were adjusted for all these environmental effects.
167 Separate phenotypes were derived for the entire lactation (overall) and for each lactation stage
168 (early, mid, late) as described above. In all cases, GEMMA v0.94.1 [17] was used to conduct
169 genomic association analyses based on a mixed model that included the genomic relationship
170 matrix among individual ewes as a polygenic effect. After Bonferroni correction for multiple
171 testing, the significance threshold for nominal P=0.05 was set at $P=7.04\times10^{-5}$ and a suggestive
172 threshold (accounting for one false positive per genome scan) was set at $P=1.41\times10^{-3}$.

173

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

176 genotype and the random effect of the animal including the pedigree relationship matrix.
177 Additive (a) and dominance (d) effects, and the proportion of additive genetic variance due to
178 each SNP locus (pVA) were calculated as follows:

179 $$a = (AA-BB)/2$$

180 $$d = AB-((AA+BB)/2)$$

181 $$pVA = (2pq\ (a+d\ (q-p))^2)/VA$$

182 where AA, BB and AB were the marginal means of the respective genotype, p and q the
183 corresponding frequencies of alleles A and B at the SNP locus, and VA the estimated additive
184 genetic variance, derived with model (1). All analyses were conducted with ASReml v4.0 [14].

185

186 Linkage disequilibrium (LD) among significant SNPs was calculated based on the r^2 value
187 using PLINK v1.9 [18]. Blocks of LD in regions harbouring significant SNPs were visualised
188 using Haploview v4.2 [19].

189

190 All significant (post-Bonferroni correction) and suggestive SNPs identified in the genomic
191 analysis for milk yield were mapped to the reference genome and annotated using the variant
192 effect predictor (http://www.ensembl.org/Tools/VEP) tool within the Ensembl database and
193 the Oar v3.1 assembly. Moreover, genes located around (0.5 Mb windows upstream and
194 downstream) the significant markers -in the candidate regions for milk yield- were annotated
195 using the BioMart data mining tool (http://www.ensembl.org/biomart/martview/) and the Oar
196 v3.1 assembly.

197

198 ## Pathway and functional enrichment analysis

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

209

210 The list of candidate genes was also analysed against an *Ovis aries* background using the
211 Database for Annotation, Visualization and Integrated Discovery (DAVID) [21] to examine
212 gene set enrichment. We determined the corresponding gene ontology terms and performed
213 functional annotation clustering analysis to detect gene enrichment. The enrichment score
214 calculated by the DAVID software package is a modified Fisher's exact test P-value; an
215 enrichment score greater than unity reflects over-representation of the respective functional
216 category.

217

218 ## Gene expression analysis

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 [22, 23]. Individual milk yield and milk somatic cell count records were also available for the sheep used in the latter study [23]. 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 [24, 25]. The sheep gene expression atlas, which includes 437 RNA-Seq libraries was produced using six Texel x Scottish Blackface sheep [24]. An additional 83 RNA-Seq libraries from a Texel trio (ewe, lamb and ram) were included in the sheep gene expression atlas [25]. We extracted data pertaining to the mammary glands, liver and kidneys. 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 [26]. Expression was reported for each protein-coding transcript as the number of transcripts per million, and then

summarised to the gene-level (as in [27]). 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.

The relationship of the expression level of each gene in the milk transcriptome with milk yield and milk somatic cell count was assessed in the Spanish sheep data using the following linear model:

$$Y_{ij} = \mu + B_i + g_j + e_{ij} \qquad (2)$$

where Y = record of ewe (milk yield or milk somatic cell count), μ = overall mean, B = fixed effect of breed i, g = fixed effect of the mean expression of gene j, e = random residual effect.

The nominal significance threshold in this analysis was set at P=0.05. Since genes were located within four QTL regions, an FDR adjustment for multiple testing was applied, setting the significance threshold at P=0.0167. These analyses were conducted with ASReml v4.0 [14].

To identify significant expression differences amongst genes located in the milk yield candidate regions in sheep with low, medium and high milk somatic cell count we performed Tukey's Test using the statistical package R v3.0.1.

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 (Nicolae et al., 2010). 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 [28-31]. 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 [24]. The aligner HISAT2 (v2.0.4) [32], was used to produce the BAM files as previously described [24]. Variants were called using BCFtools [33] 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 [34] 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 [35] 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) [36] 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 [37] adjusted P-value ≤ 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 [38].

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. However, the genetic correlation between early and late lactation was moderate (0.60) and significantly less than one. In practical terms, lactation onset, peak lactation and lactation persistence may have partly separate genetic control. Genetic correlations between milk and mastitis traits 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).

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)

322

323

Table 2. Estimates of phenotypic and genetic correlations between milk yield and four mastitis traits 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)		

324

325 # Genomic association studies

326 Separate genomic association analyses were conducted for milk yield in early, mid, late and
327 overall lactation. Multidimensional scaling analysis of the studied population revealed no
328 substructure. In general, similar genomic associations were detected for milk yield in middle,
329 late and overall lactation but distinct associations were observed in early lactation. We
330 identified a genome-wide significant association after Bonferroni correction for multiple
331 testing on chromosome 19 (P= 1.28 x 10^{-5}) and three suggestive associations on chromosomes
332 2 (P= 4.30 x 10^{-4}), 12 (P= 3.65 x 10^{-4}) and 16 (P= 6.07 x 10^{-4}). Details of SNPs associated
333 with milk yield are shown in Table 3. Manhattan plots and corresponding Q-Q plots displaying
334 genomic association results are shown in Fig. 1 and Fig. 2, respectively.

335

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)	pVA	p	q
Early 1-7 weeks	OAR12_23075585	12(20050780)	3.35E-04	0.07(0.09)	0.07(0.10)	0.01	0.62	0.38
	oar3_OAR12_19689222	12(19689222)	3.65E-04	0.05(0.36)	0.12(0.03)	0.01	0.73	0.27
	oar3_OAR12_19269103	12(19269103)	4.66E-04	-0.03(0.33)	0.05(0.02)	0.01	0.73	0.27
	oar3_OAR12_19500329	12(19500329)	6.98E-04	0.08(0.02)	0.02(0.64)	0.02	0.6	0.4
	oar3_OAR12_19624437	12(19624437)	6.77E-04	0.05(0.30)	0.09(0.10)	0.01	0.72	0.28
	oar3_OAR12_19840123	12(19840123)	9.80E-04	0.04(0.30)	0.08(0.10)	0.01	0.68	0.32
	oar3_OAR16_33078067	16(33078067)	6.07E-04	-0.09(0.00)	0.20(0.05)	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)	0.16	0.48	0.52
	oar3_OAR19_24119431	19(24119431)	9.03E-04	-0.14(0.00)	-0.02(0.53)	0.15	0.48	0.52
	OAR19_25513179	19(24010793)	1.70E-03	-0.15(0.00)	-0.01(0.77)	0.16	0.58	0.42
	OAR16_34906481	16(32156238)	1.35E-03	-0.08(0.02)	0.07(0.05)	0.05	0.90	0.10
	OAR2_133418483	2(125230366)	1.48E-03	0.08(0.51)	-0.10(0.44)	0.06	0.92	0.08
	OAR2_133088440	2(124907852)	2.28E-04	0.23(0.00)	0.19(0.03)	0.05	0.82	0.18
	oar3_OAR2_124936445	2(124936445)	1.11E-03	0.12(0.06)	0.07(0.30)	0.03	0.78	0.22
Late 18-24 weeks	OAR19_25259444	19(23804520)	1.41E-04	-0.15(0.00)	-0.02(0.57)	0.14	0.48	0.52
	oar3_OAR19_24745933	19(24745933)	2.24E-04	0.10(0.00)	-0.10(0.00)	0.07	0.54	0.46
	OAR19_25830151	19(24342061)	1.35E-03	0.07(0.07)	-0.00(0.87)	0.02	0.72	0.28
	oar3_OAR19_24707843	19(24707843)	7.91E-04	-0.13(0.00)	-0.09(0.03)	0.06	0.64	0.36
	oar3_OAR19_23656789	19(23656789)	6.29E-04	-0.11(0.00)	-0.06(0.11)	0.07	0.52	0.48
	oar3_OAR2_124936445	2(125000000)	5.76E-04	0.11(0.02)	-0.00(0.93)	0.05	0.78	0.22
	OAR2_133088440	2(124907852)	4.30E-04	0.19(0.00)	0.09(0.18)	0.06	0.82	0.18
Overall	OAR19_25259444	19(23804520)	1.28E-05	-0.14(0.00)	-0.00(0.84)	0.12	0.48	0.52
	oar3_OAR19_24032312	19(24032312)	2.70E-04	-0.11(0.00)	-0.08(0.03)	0.05	0.6	0.4
	oar3_OAR19_24707843	19(24707843)	2.90E-04	-0.12(0.00)	-0.07(0.04)	0.06	0.64	0.36
	OAR19_25513179	19(24010793)	4.16E-04	-0.11(0.00)	-0.06(0.07)	0.06	0.58	0.42
	oar3_OAR19_24119431	19(24119431)	4.90E-04	-0.11(0.00)	-0.09(0.01)	0.08	0.48	0.52
	oar3_OAR19_23929524	19(23929524)	5.62E-04	0.11(0.00)	-0.03(0.29)	0.07	0.48	0.52
	oar3_OAR19_24745933	19(24745933)	9.80E-04	0.10(0.00)	-0.05(0.15)	0.06	0.54	0.46
	oar3_OAR19_23891277	19(23891277)	1.17E-03	-0.10(0.00)	-0.05(0.13)	0.05	0.54	0.46
	oar3_OAR19_23656789	19(23656789)	1.46E-03	-0.10(0.00)	-0.05(0.15)	0.06	0.52	0.48
	OAR2_133088440	2(124907852)	9.55E-04	0.20(0.00)	0.12(0.06)	0.05	0.82	0.18
	OAR2_133418483	2(125230366)	1.47E-03	0.10(0.30)	-0.08(0.43)	0.05	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; pVA: proportion of the genetic variance explained by the SNP; p and q allelic frequencies

336

337 **Fig 1. Manhattan plots displaying the genomic association results for milk yield in Chios**
338 **sheep.** Manhattan plots for milk yield in early (A), mid (B), late (C), and overall (D) lactation.
339 Genomic location is plotted against $-\log_{10}(P)$. Red and blue lines, respectively, are thresholds
340 for significance post-Bonferroni correction ($P<0.05$) and for suggestive significance
341 (accounting for one false positive per genome scan).

342 **Fig 2. Q-Q plots displaying the genomic association results for milk yield in Chios sheep.**
343 Q-Q plots in early (A), mid (B), late (C) and overall (D) lactation; observed *P*-values are plotted
344 against the expected *P*-values.

345

346

347 The significance of the above SNP markers was confirmed in mixed model analyses based on
348 the pedigree genetic relationship matrix. The additive and dominance genetic effects, and the
349 proportion of the total genetic variance explained by each of these SNPs in the corresponding
350 lactation stage are summarised in Table 3. Most SNPs had a significant additive effect and a
351 few a significant dominance effect on milk yield. The significant SNPs in the QTL region on
352 chromosome 19 accounted for 16% of the additive genetic variance, while collectively all the
353 SNPs in the four candidate regions accounted for 30% of the additive genetic variance of milk
354 yield. When located on the same chromosomes, the significant markers identified for milk yield
355 were in linkage disequilibrium (LD=0.27-0.97), implying that they correspond to the same
356 causative mutation (S1 Table). The significant SNPs identified in the present study were not in
357 LD with the SNPs previously associated with the mastitis related traits in Chios sheep [10] (S1
358 Table). Only small LD blocks were visualised with Haploview, probably due to a high number
359 of recombination events having taken place in the outbred population of study. All significant
360 SNP markers were located in intergenic or intronic regions. The candidate QTL regions for
361 milk yield contained a relatively small number of protein-coding genes (n=31) and microRNAs
362 (n=6) (S2 Table).

363

Pathway and functional clustering analysis

The genes located in the candidate regions for milk yield were enriched for pathways involved in electrolyte (Na^+, K^+, and H^+) transport and homeostasis, lipid metabolism (ketolysis, ketogenesis) and oxidative stress, as well as innate immune responses (Fig. 3). Moreover, two networks of molecular interactions were constructed, one of which was related with immunological disease and cell signalling and interaction, and another with the development, function and organ morphology of the endocrine and reproductive systems (Fig. 4).

Fig 3. Pathway analysis using the IPA software. The most highly represented canonical pathways derived from genes located within the studied candidate regions for milk yield in Chios sheep. The solid yellow line represents the significance threshold. The line joining squares represents the ratio of the genes within each pathway to the total number of genes in the pathway.

Fig 4. Network analysis using the IPA software. Two gene networks illustrating the molecular interactions between candidate gene products: (I) network related to immunological disease and cell signalling and interaction; (II) network related to the endocrine system development and function, reproductive system development and function, and organ morphology. 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.

The functional annotation clustering analysis showed that genes were organised into two clusters associated with the regulation of cellular processes (enrichment score =1.60) and metabolic processes (enrichment score =1.04). Both clusters contained the same genes

387 (*CNTN4, DNAJA1, ESRRG, FGF10, FRZB, GHR, HMGCS1, OXCT1, TGFB2*).

388

389 ## Gene expression analysis

390 Fourteen of the genes located in the candidate regions for milk yield (*CCL28, DNAJC10,*
391 *DUSP19, EMB, FGF10, GHR, LYPLAL1, NNT, HMGCS1, NCKAP1, NUP35, OXCT1, PAIP1*
392 and *ZNF131)* were expressed in either the milk transcriptome or the mammary gland (S1-3
393 Figs). The growth hormone receptor (*GHR*) and 3-oxoacid CoA transferase 1 (*OXCT1*) genes
394 were highly expressed in liver and kidney cortex tissue, respectively (S2 Fig). Five of the
395 candidate genes (*DNAJC10, EMB, HMGCS1, OXCT1, PAIP1*) detected in tissues related to
396 milk production (mammary gland, liver and kidney cortex) were also up-regulated in immune
397 related tissues, relative to the other tissues analysed (S3 Fig). The *EMB* gene in particular
398 exhibited a strong immune-specific profile with a high level of expression in macrophages
399 relative to the other tissues (S2 Fig).

400

401 Expression of genes *LYPLAL, PARP8, RRP15* and *TGFB2* had a suggestive (nominal $P<0.05$)
402 association with milk yield in the Churra and Assaf sheep data that did not remain significant
403 after the FDR correction.

404

405 A subset of five genes (*EMB, FGF10, DNAJC10, OXCT1, PARP8*) had a suggestive (nominal
406 $P<0.05$) association with milk somatic cell count in the Churra and Assaf sheep data. The
407 expression level of two of these genes, *EMB* and *FGF10,* were also significantly different

between sheep with high and low somatic cell count in milk (Tukey's Test; *EMB*: *P*=0.001648; *FGF10*: *P*=0.002085).

Allelic expression imbalance analysis

Exonic single nucleotide variation (SNP and indels) was observed in 24 of the protein coding genes located in the candidate regions for milk yield. Missense variants were identified in several genes including, *CNTN4, DNAJC1, DUSP19, GHR, HMGS1, MRPS30, NNT, NUP35* and *RRP15* genes. One-sampled MBASED analysis identified two genes, *CCL28* (*P= 1.8e-05)* and *RRP15* (*P= 3e-03*) with significant allelic expression imbalance. Specifically, seven SNPs in the 3' UTR region of *CCL28* (major allele frequency 0.72) and two synonymous 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 in both genes was evident in only one of the three individual sheep.

Candidate genes

Based on all above results, a total of seven genes (*DNAJA1, DNAJC10, FGF10, GHR, HMGCS1, LYPLA1, OXCT1*) were selected as candidate genes for milk yield located in mastitis genomic regions (Table 4). Genes were selected using a combination of their known biological function, involvement in relevant pathways and networks, enrichment in tissues relevant to milk production, and any previously known association with milk production in either dairy sheep or other species.

Table 4. Selected candidate genes for milk yield.

Gene Symbol	**Gene name**	**Genomic location**	**Function**
DNAJC10	DnaJ Heat Shock Protein (Hsp40) Member C10	2: 125867867-125913602	Required for efficient folding of proteins in the endoplasmic reticulum by catalysing the removal of non-native disulfide bonds formed during the folding of proteins, such as LDLR [39, 40].
LYPLAL1	Lysophospholipase-Like 1	12: 20785849-20814579	Hydrolyses fatty acids from S-acylated cysteine residues in proteins such as trimeric G alpha proteins or HRAS. Has depalmitoylating activity toward KCNMA1 and low lysophospholipase activity [41].
FGF10	Fibroblast Growth Factor 10	16: 30468674 - 30562188	This gene encodes a protein which is part of a family of proteins called fibroblast growth factors that are involved in important processes such as cell division, regulation of cell growth and maturation, formation of blood vessels, wound healing, and development before birth. By attaching to another protein known as a receptor, the FGF10 protein triggers a cascade of chemical reactions inside the cell that signals the cell to undergo certain changes, such as maturing to take on specialised functions [42].
GHR	Growth Hormone Receptor	16: 31832933 - 32000445	This gene encodes a member of the type I cytokine receptor family, which is a transmembrane receptor for growth hormone.

			Binding of growth hormone to the receptor leads to receptor dimerization and the activation of an intra- and intercellular signal transduction pathway leading to growth. Variation in this gene can affects glucose, lipid and insulin-like growth factor-I metabolism [43].
HMGCS1	3-Hydroxy-3-Methylglutaryl-CoA Synthase 1	16: 31409643 - 31430541	This enzyme condenses acetyl-CoA with acetoacetyl-CoA to form HMG-CoA, which is the substrate for HMG-CoA reductase [44].
OXCT1	3-Oxoacid CoA Transferase 1	16: 32617427-32779330	Key enzyme for ketone body catabolism. The encoded protein is a homodimeric mitochondrial matrix enzyme that plays a central role in extrahepatic ketone body catabolism by catalysing the reversible transfer of coenzyme A from succinyl-CoA to acetoacetate [45].
DNAJA1	DnaJ Heat Shock Protein Family (Hsp40) Member A1	19: 23692565-23720861	This gene encodes a member of the DnaJ protein family, which act as heat shock protein 70 cochaperones and facilitates protein folding, trafficking, prevention of aggregation, and proteolytic degradation. Stimulates ATP hydrolysis and plays a role in protein transport into mitochondria [46].

Discussion

The existence of a mastitis-specific ovine DNA array built on previously detected markers

433 associated with mastitis resistance in dairy sheep opens up opportunities for targeted genomic
434 and marker-assisted selection aiming to enhance animal resistance to the disease. The aim of
435 the present study was to investigate the association of this array with milk yield of dairy sheep
436 and assess the feasibility of a concomitant genetic improvement programme for the two traits.
437 Chios sheep were used as a study model.

438

439 According to our results, milk yield and mastitis traits in the Chios sheep are not genetically
440 correlated to each other. Genetic correlation estimates between milk somatic cell count and
441 milk yield are reportedly antagonistic in dairy cattle [47] but inconsistent amongst previous
442 sheep studies ranging from antagonistic [48] to favourable [3]. Our findings for the Chios sheep
443 indicate that selection for enhanced mastitis resistance could be incorporated into the current
444 genetic improvement programme without incurring adverse effects on milk yield.

445 An overall moderate but significant heritability for milk yield was estimated in Chios sheep,
446 consistent with the dairy sheep literature (ranging from 0.16 to 0.30) as reviewed in [49] and
447 previous studies in Chios sheep ranging from 0.21 to 0.29 [50].

448

449 Genomic analyses conducted here revealed several SNPs on the mastitis array with a significant
450 effect on milk yield. These milk-associated SNPs were not in LD with genomic regions found
451 previously to affect mastitis resistance in the same population [10]. For example, the QTL for
452 milk yield on chromosome 2 was 75 Mb distant from the one previously identified for mastitis
453 resistance on the same chromosome [10]. The association of this QTL region with milk yield
454 is supported by results of a previous genomic selection mapping study that compared dairy
455 with meat sheep breeds to identify genomic regions for milk traits under selection [51]. 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 [51]. 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 [52]. No QTLs for mastitis resistance have been identified on this chromosome in Chios sheep [10]. The QTLs on chromosome 19 and 16 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, and milk protein and fat content have been also identified on chromosome 16 in Churra sheep [53], in close proximity with the QTL identified here in Chios sheep. To the best of our knowledge, the QTL on chromosome 19 is reported here for the first time. These results are also consistent with a previous study of Chios sheep [54], suggesting that a relatively major locus might be involved in ovine milk production. The QTL identified on chromosome 19 in the present study explained 16% of the genetic variance. Furthermore, the significant SNP markers identified for milk yield in our study collectively explained over 30% of the genetic variance of the trait, suggesting that the mastitis-specific targeted array can also be used for genomic selection to enhance milk yield.

In the QTL regions identified for milk yield on chromosomes 2 and 19 there are two candidate genes, *DNAJA1* and *DNAJC10*, both belonging to the same gene family. In the previous milk transcriptome study of the Churra and Assaf breeds, two other related genes, *DNAJA4* and *DNAJB2,* were reported as functional candidates for milk yield [55]. 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 [56], while *DNAJC10* promotes apoptosis

in response to endoplasmic reticulum stress. Therefore, these two genes 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 [57].

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 [58] and mastitis resistance [59], and has been suggested to regulate mammary gland metabolism and milk synthesis during mastitis infection in dairy cattle [60]. Using the gene expression atlas for sheep and the milk transcriptome dataset, *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 [60-64]. Selective sweeps were also identified in the *GHR* region when dairy with beef cattle were compared [65]. 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). Furthermore, fibroblast growth factor 10 (*FGF10)* and 3-Hydroxy-3-Methylglutaryl-CoA Synthase 1 (*HMGSC1)* genes have been previously associated with milk production in dairy cattle [61, 66]. The pleiotropic growth factor *FGF10* is reportedly required for mammary gland development in mice [67]. In sheep, this gene was highly expressed in the mammary gland and female reproductive tissues including the uterus and placenta (http://biogps.org/sheepatlas). In the present study, *FGF10* was found to be differentially expressed in sheep with different milk somatic cell counts, suggesting a possible role in mastitis resistance. Our pathway analysis showed that the products of genes *GHR, FGF10, HMGSC1* are part of a network related with the development, function and organ

morphology of the endocrine and reproductive systems. However, further studies are needed to confirm the relevance of this network with milk production and identify the causative genes and mutations.

Allelic expression imbalance was detected in the epithelia-associated chemokine (C-C motif) ligand 28 (*CCL28)* and ribosomal RNA processing 15 (*RRP15)* gene*s* in mammary gland tissue. The *CCL28* gene encodes a protein that has been previously reported to demonstrate direct antimicrobial activity against mastitis infection in dairy cattle [68] and in humans [69-71]. This gene is upregulated during pregnancy and lactation, and considered vital for the ability of IgA-producing B cells to migrate to the mammary tissue during lactation [71]. In the present study, *CCL28* was expressed highly in the milk somatic cells but not in immune tissues implying a protective role mainly in the mammary gland, as has been shown in humans [69, 70]. The *RRP15* gene plays a role in cell cycle, cell proliferation and apoptosis [72]. Both these genes are linked to protective immunity and, as such, are likely to be under strong selection pressure. The two variants detected in *PPR15* were synonymous SNPs and, therefore, less likely to be relevant to gene function. The seven variants in 3' UTR of *CCL28* could be more relevant to the genetic/transcriptional control of *CCL28* expression, since the turnover of mRNA is mostly regulated by *cis*-acting elements located in the 3'UTR regions [73]. Therefore, these SNPs might affect the corresponding phenotypic trait in sheep, possibly by disrupting miRNA binding as in the myostatin example from Texel sheep described in [74]. Further analysis using the miRWalk database [75] predicted that one of the SNPs on chromosome 16 (31362143 bp) that exhibited allelic expression imbalance is located within a microRNA binding site, in the 3' UTR of *CCL28*. Interestingly, the same microRNA (bta-mir-29e) in dairy cattle has been already reported to be differentially expressed in bovine mastitis caused by gram positive bacteria [76]. However, further studies are needed to confirm this,

since allelic expression imbalance in the two genes was identified in only one of the three individual sheep studied, implying this might be simply due to individual specific variation in expression. A wider analysis across multiple tissues would also help to determine if the allelic expression imbalance observed in the present study is indeed specific to genes associated with the mammary gland. Further studies could also investigate the relevance of the SNPs exhibiting allelic imbalance to gene function and quantify allelic expression imbalance in a wider subset of animals, preferably including animals of the Chios breed.

In conclusion, 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. Moreover, several candidate genes were highly expressed in immune tissues and in milk implying a favourable pleiotropic effect or likely role in milk production during udder infection. 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 or as novel therapeutics.

Competing interests

554 The authors declare that they have no competing interests regarding the publication of this
555 paper.

556

557 ## Funding

558 This research was partly supported by the Seven Framework Program of the European
559 Commission (project: '3SR: Sustainable Solutions for Small Ruminants'), a Biotechnology and
560 Biological Sciences Research Council (BBSRC) Grant 'Functional Annotation of the Sheep
561 Genome' ref.: BB/L001209/1 and BBSRC Institute Strategic Programme Grants 'Analysis and
562 Prediction in Complex Animal Systems' ref.: BB/J004235/1 and Farm animal genomics ref:
563 BBS/E/D/20211550, 'Blue Prints for Healthy Animals' ref.: BB/P013732/1 and 'Improving
564 Animal Production & Welfare' ref.: BB/P013759/1 as well as the Laboratory of Animal
565 Production, School of Veterinary Medicine, Aristotle University of Thessaloniki, Greece. The
566 contribution of GBa was also supported by the Rural & Environment Science & Analytical
567 Services Division of the Scottish Government.

568

569 # Availability of data and materials

570 All the raw gene expression data comprising the Texel and the Texel x Scottish Blackface gene
571 expression atlas is available in the European Nucleotide Archive (ENA) under accession
572 number PRJEB6169 and PRJEB19199, respectively. To supplement the data from these two
573 transcriptomic atlas projects, we also obtained expression levels from a milk transcriptomic
574 study of the milk somatic cells of two Spanish dairy sheep breeds, Churra and Assaf (Gene

575 Expression Omnibus (GEO) database accession number GSE74825 and NCBI BioProject ID
576 PRJNA301615).

577

578 Authors' contributions

579 GBa, AP and GA conceived and designed the genetic study of Chios sheep and secured
580 substantial funding. AP, GBr and GBa performed data collection, phenotyping, DNA
581 extractions and genetic parameter analysis. AP and GBa collated and edited the genotyping
582 data and performed the genomic analysis. DAH and ELC conceived and designed the sheep
583 gene expression atlas and DAH secured the substantial funding. ELC and SJB created the sheep
584 gene expression atlas and analysed the gene expression data for the atlas dataset and the milk
585 transcriptome. AP performed the pathway and the TFBS analyses. AP and SJB extracted and
586 annotated the re-sequencing data of the HapMap sheep. PD performed the allelic expression
587 imbalance analysis with input from AP, ELC, SJB and DAH. GBa, DH, GA, ELC and AP
588 interpreted these results. GBa and AP wrote the manuscript. All other co-authors provided
589 manuscript editing and feedback. All authors read and approved the final manuscript.

590

591 Acknowledgements

592 The authors thank the cooperation of the commercial farmers who allowed access to their
593 respective flocks.

594

595 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. doi: https://doi.org/10.3168/jds.2016-12326.
2. Barillet F. Genetic improvement for dairy production in sheep and goats. Small Ruminant Research. 2007;70(1):60-75. doi: 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. doi: https://doi.org/10.3168/jds.S0022-0302(05)72899-X.
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. doi: 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. doi: 10.1017/s1751731108003522. PubMed 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. doi: 10.3389/fmicb.2016.00319. PubMed PMID: PMC4789554.
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. doi: 10.1186/s12864-017-3982-1. PubMed PMID: 28814268; PubMed Central PMCID: PMCPMC5559839.
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. doi: 10.1371/journal.pgen.1005629. PubMed PMID: 26658352; PubMed Central PMCID: PMCPmc4676722.
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.
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, Sechi T, Usai MG, Ligios C, Foucras G, Carta A. 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. doi: 10.1371/journal.pone.0115960.
16. 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. doi: 10.1093/bioinformatics/btm108. PubMed PMID: 17384015.

17. 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. doi: 10.1038/nmeth.2848. PubMed PMID: 24531419; PubMed Central PMCID: PMCPMC4211878.
18. 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. doi: 10.1086/519795. PubMed PMID: 17701901; PubMed Central PMCID: PMCPMC1950838.
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. doi: 10.1093/bioinformatics/bth457. PubMed 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. doi: 10.1093/bioinformatics/btt703. PubMed PMID: PMC3928520.
21. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003;4(5):P3. Epub 2003/05/08. PubMed PMID: 12734009.
22. 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. doi: 10.1038/sdata.2016.51. PubMed PMID: 27377755; PubMed Central PMCID: PMCPmc4932878.
23. 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. doi: 10.1038/srep18399. PubMed PMID: 26677795; PubMed Central PMCID: PMCPmc4683406.
24. 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. doi: 10.1371/journal.pgen.1006997.
25. 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. doi: 10.1126/science.1252806. PubMed PMID: 24904168; PubMed Central PMCID: PMCPmc4157056.
26. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotech. 2016;34(5):525-7. doi: 10.1038/nbt.3519

http://www.nature.com/nbt/journal/v34/n5/abs/nbt.3519.html#supplementary-information.
27. 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. doi: 10.12688/f1000research.7563.1. PubMed PMID: 26925227; PubMed Central PMCID: PMCPMC4712774.
28. 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. doi: 10.1038/nature02797. PubMed PMID: 15269782; PubMed Central PMCID: PMCPMC2966974.
29. 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. doi: 10.1371/journal.pgen.0010078. PubMed PMID: 16362079; PubMed Central PMCID: PMCPMC1315281.
30. 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. doi: 10.1007/s00439-003-0956-y. PubMed PMID: 12728311.
31. 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. doi: 10.1186/s12864-015-2174-0. PubMed PMID: 26596891; PubMed Central PMCID: PMCPMC4657355.

32. 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. doi: 10.1038/nmeth.3317. PubMed PMID: 25751142; PubMed Central PMCID: PMCPMC4655817.

33. 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. doi: 10.1093/bioinformatics/btr509. PubMed PMID: 21903627; PubMed Central PMCID: PMCPMC3198575.

34. 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. doi: 10.1093/bioinformatics/btq033. PubMed PMID: 20110278; PubMed Central PMCID: PMCPmc2832824.

35. 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. doi: 10.4161/fly.19695. PubMed PMID: 22728672; PubMed Central PMCID: PMCPMC3679285.

36. 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. doi: 10.1186/s13059-014-0405-3. PubMed PMID: 25315065; PubMed Central PMCID: PMCPMC4165366.

37. 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.

38. 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. doi: 10.3168/jds.2008-1719. PubMed PMID: 19528590.

39. Dong M, Bridges JP, Apsley K, Xu Y, Weaver TE. ERdj4 and ERdj5 are required for endoplasmic reticulum-associated protein degradation of misfolded surfactant protein C. Molecular biology of the cell. 2008;19(6):2620-30. Epub 2008/04/11. doi: 10.1091/mbc.E07-07-0674. PubMed PMID: 18400946; PubMed Central PMCID: PMCPMC2397301.

40. Oka OB, Pringle MA, Schopp IM, Braakman I, Bulleid NJ. ERdj5 is the ER reductase that catalyzes the removal of non-native disulfides and correct folding of the LDL receptor. Molecular cell. 2013;50(6):793-804. Epub 2013/06/19. doi: 10.1016/j.molcel.2013.05.014. PubMed PMID: 23769672; PubMed Central PMCID: PMCPMC3906653.

41. Burger M, Zimmermann TJ, Kondoh Y, Stege P, Watanabe N, Osada H, et al. Crystal structure of the predicted phospholipase LYPLAL1 reveals unexpected functional plasticity despite close relationship to acyl protein thioesterases. Journal of lipid research. 2012;53(1):43-50. Epub 2011/11/05. doi: 10.1194/jlr.M019851. PubMed PMID: 22052940; PubMed Central PMCID: PMCPMC3243480.

42. Emoto H, Tagashira S, Mattei MG, Yamasaki M, Hashimoto G, Katsumata T, et al. Structure and expression of human fibroblast growth factor-10. J Biol Chem. 1997;272(37):23191-4. Epub 1997/09/12. PubMed PMID: 9287324.

43. Sorensen K, Aksglaede L, Munch-Andersen T, Aachmann-Andersen NJ, Leffers H, Helge JW, et al. Impact of the growth hormone receptor exon 3 deletion gene polymorphism on glucose metabolism, lipids, and insulin-like growth factor-I levels during puberty. The Journal of clinical endocrinology and metabolism. 2009;94(8):2966-9. Epub 2009/05/07. doi: 10.1210/jc.2009-0313. PubMed PMID: 19417039.

44. Shafqat N, Turnbull A, Zschocke J, Oppermann U, Yue WW. Crystal structures of human HMG-CoA synthase isoforms provide insights into inherited ketogenesis disorders and inhibitor

746 design. Journal of molecular biology. 2010;398(4):497-506. Epub 2010/03/30. doi:
747 10.1016/j.jmb.2010.03.034. PubMed PMID: 20346956.
748 45. Song XQ, Fukao T, Mitchell GA, Kassovska-Bratinova S, Ugarte M, Wanders RJ, et al. Succinyl-
749 CoA:3-ketoacid coenzyme A transferase (SCOT): development of an antibody to human SCOT and
750 diagnostic use in hereditary SCOT deficiency. Biochimica et biophysica acta. 1997;1360(2):151-6.
751 Epub 1997/04/12. PubMed PMID: 9128180.
752 46. Tzankov S, Wong MJ, Shi K, Nassif C, Young JC. Functional divergence between co-
753 chaperones of Hsc70. J Biol Chem. 2008;283(40):27100-9. Epub 2008/08/08. doi:
754 10.1074/jbc.M803923200. PubMed PMID: 18684711; PubMed Central PMCID: PMCPMC5026489.
755 47. Rupp R, Boichard D. Genetics of resistance to mastitis in dairy cattle. Veterinary research.
756 2003;34(5):671-88. Epub 2003/10/15. doi: 10.1051/vetres:2003020. PubMed PMID: 14556700.
757 48. Tolone M, Riggio V, Portolano B. Estimation of genetic and phenotypic parameters for
758 bacteriological status of the udder, somatic cell score, and milk yield in dairy sheep using a threshold
759 animal model. Livestock Science. 2013;151(2-3):134-9. doi: 10.1016/j.livsci.2012.11.014.
760 49. Juan José Arranz BGr-G. Detection of QTL Underlying Milk Traits in Sheep: An Update, Milk
761 Production. In: Chaiyabutr PN, editor. Advanced Genetic Traits, Cellular Mechanism, Animal
762 Management and Health: InTech; 2012.
763 50. Ligda C, Papadopoulos T, Mavrogenis A, Georgoudis A. Genetic parameters for test day milk
764 traits and somatic cell counts in Chios dairy sheep. In: Gabiña D, Sanna S, editors. Breeding
765 programmes for improving the quality and safety of products New traits, tools, rules and
766 organization? Options Méditerranéennes : Série A. Séminaires Méditerranéens. 55: Zaragoza :
767 CIHEAM; 2003. p. 55-9.
768 51. Gutierrez-Gil B, Arranz JJ, Pong-Wong R, Garcia-Gamez E, Kijas J, Wiener P. Application of
769 selection mapping to identify genomic regions associated with dairy production in sheep. PLoS One.
770 2014;9(5):e94623. Epub 2014/05/03. doi: 10.1371/journal.pone.0094623. PubMed PMID: 24788864;
771 PubMed Central PMCID: PMCPmc4006912.
772 52. Mateescu RG, Thonney ML. Genetic mapping of quantitative trait loci for milk production in
773 sheep. Anim Genet. 2010;41(5):460-6. Epub 2010/04/17. doi: 10.1111/j.1365-2052.2010.02045.x.
774 PubMed PMID: 20394603.
775 53. Garcia-Gamez E, Gutierrez-Gil B, Sahana G, Sanchez JP, Bayon Y, Arranz JJ. GWA analysis for
776 milk production traits in dairy sheep and genetic support for a QTN influencing milk protein
777 percentage in the LALBA gene. PLoS One. 2012;7(10):e47782. Epub 2012/10/25. doi:
778 10.1371/journal.pone.0047782. PubMed PMID: 23094085; PubMed Central PMCID:
779 PMCPMC3475704.
780 54. Chatziplis DG, Tzamaloukas O, Miltiadou D, Ligda C, Koumas A, Mavrogenis AP, et al.
781 Evidence of major gene(s) affecting milk traits in the Chios sheep breed. Small Ruminant Research.
782 2012;105(1):61-8. doi: https://doi.org/10.1016/j.smallrumres.2011.12.009.
783 55. Suárez-Vega A, Gutiérrez-Gil B, Klopp C, Tosser-Klopp G, Arranz JJ. Variant discovery in the
784 sheep milk transcriptome using RNA sequencing. BMC Genomics. 2017;18(1):170. doi:
785 10.1186/s12864-017-3581-1.
786 56. Gotoh T, Terada K, Oyadomari S, Mori M. hsp70-DnaJ chaperone pair prevents nitric oxide-
787 and CHOP-induced apoptosis by inhibiting translocation of Bax to mitochondria. Cell death and
788 differentiation. 2004;11(4):390-402. Epub 2004/01/31. doi: 10.1038/sj.cdd.4401369. PubMed PMID:
789 14752510.
790 57. Stefanon B, Colitti M, Gabai G, Knight CH, Wilde CJ. Mammary apoptosis and lactation
791 persistency in dairy animals. The Journal of dairy research. 2002;69(1):37-52. Epub 2002/06/06.
792 PubMed PMID: 12047109.
793 58. Bionaz M, Loor JJ. Gene networks driving bovine milk fat synthesis during the lactation cycle.
794 BMC Genomics. 2008;9:366. Epub 2008/08/02. doi: 10.1186/1471-2164-9-366. PubMed PMID:
795 18671863; PubMed Central PMCID: PMCPMC2547860.

59. 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. doi: https://doi.org/10.3168/jds.2013-7480.
60. 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. doi: 10.1371/journal.pone.0114919. PubMed PMID: PMC4319771.
61. 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. doi: 10.1186/1297-9686-46-29. PubMed PMID: 24779965; PubMed Central PMCID: PMCPMC4036308.
62. 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. doi: 10.1111/j.1365-2052.2009.01945.x.
63. 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. doi: 10.1007/s13353-011-0051-3. PubMed PMID: 21660490.
64. 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. doi: 10.1111/j.1365-2052.2009.01921.x. PubMed PMID: 19508288; PubMed Central PMCID: PMCPMC2779988.
65. 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). doi: 10.3389/fgene.2015.00167.
66. Yodklaew P, Koonawootrittriron S, Elzo MA, Suwanasopee T, Laodim T. Genome-wide association study for lactation characteristics, milk yield and age at first calving in a Thai multibreed dairy cattle population. Agriculture and Natural Resources. 2017;51(3):223-30. doi: https://doi.org/10.1016/j.anres.2017.04.002.
67. Cui Y, Li Q. Expression and functions of fibroblast growth factor 10 in the mouse mammary gland. International journal of molecular sciences. 2013;14(2):4094-105. Epub 2013/02/26. doi: 10.3390/ijms14024094. PubMed PMID: 23434672; PubMed Central PMCID: PMCPMC3588087.
68. Pallister KB, Mason S, Nygaard TK, Liu B, Griffith S, Jones J, et al. Bovine CCL28 Mediates Chemotaxis via CCR10 and Demonstrates Direct Antimicrobial Activity against Mastitis Causing Bacteria. PLoS One. 2015;10(9):e0138084. Epub 2015/09/12. doi: 10.1371/journal.pone.0138084. PubMed PMID: 26359669; PubMed Central PMCID: PMCPMC4567263.
69. Bourges D, Meurens F, Berri M, Chevaleyre C, Zanello G, Levast B, et al. New insights into the dual recruitment of IgA+ B cells in the developing mammary gland. Mol Immunol. 2008;45(12):3354-62. Epub 2008/06/06. doi: 10.1016/j.molimm.2008.04.017. PubMed PMID: 18533264.
70. Hieshima K, Ohtani H, Shibano M, Izawa D, Nakayama T, Kawasaki Y, et al. CCL28 has dual roles in mucosal immunity as a chemokine with broad-spectrum antimicrobial activity. Journal of immunology (Baltimore, Md : 1950). 2003;170(3):1452-61. Epub 2003/01/23. PubMed PMID: 12538707.
71. Low EN, Zagieboylo L, Martino B, Wilson E. IgA ASC accumulation to the lactating mammary gland is dependent on VCAM-1 and alpha4 integrins. Mol Immunol. 2010;47(7-8):1608-12. Epub 2010/02/23. doi: 10.1016/j.molimm.2010.01.015. PubMed PMID: 20171738; PubMed Central PMCID: PMCPMC2849929.
72. Wu T, Ren MX, Chen GP, Jin ZM, Wang G. Rrp15 affects cell cycle, proliferation, and apoptosis in NIH3T3 cells. FEBS open bio. 2016;6(11):1085-92. Epub 2016/11/12. doi: 10.1002/2211-5463.12128. PubMed PMID: 27833849; PubMed Central PMCID: PMCPMC5095146.

846 73. Mignone F, Gissi C, Liuni S, Pesole G. Untranslated regions of mRNAs. Genome biology.
847 2002;3(3):REVIEWS0004-REVIEWS. Epub 02/28. PubMed PMID: 11897027.
848 74. Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibe B, et al. A mutation creating a potential
849 illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. Nature
850 genetics. 2006;38(7):813-8. Epub 2006/06/06. doi: 10.1038/ng1810. PubMed PMID: 16751773.
851 75. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: An online resource for prediction of
852 microRNA binding sites. PLoS One. 2018;13(10):e0206239. Epub 2018/10/20. doi:
853 10.1371/journal.pone.0206239. PubMed PMID: 30335862.
854 76. Lawless N, Foroushani AB, McCabe MS, O'Farrelly C, Lynn DJ. Next generation sequencing
855 reveals the expression of a unique miRNA profile in response to a gram-positive bacterial infection.
856 PLoS One. 2013;8(3):e57543. Epub 2013/03/09. doi: 10.1371/journal.pone.0057543. PubMed PMID:
857 23472090; PubMed Central PMCID: PMCPMC3589390.

858

859

860

861

862 ## Supplementary files

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

866 **S2 Table.** Genes located in the candidate genomic regions identified for milk yield in Chios
867 sheep.

868 **S3 Table.** Allelic expression imbalance analysis using the one-sampled MBASED method.

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

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

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

883

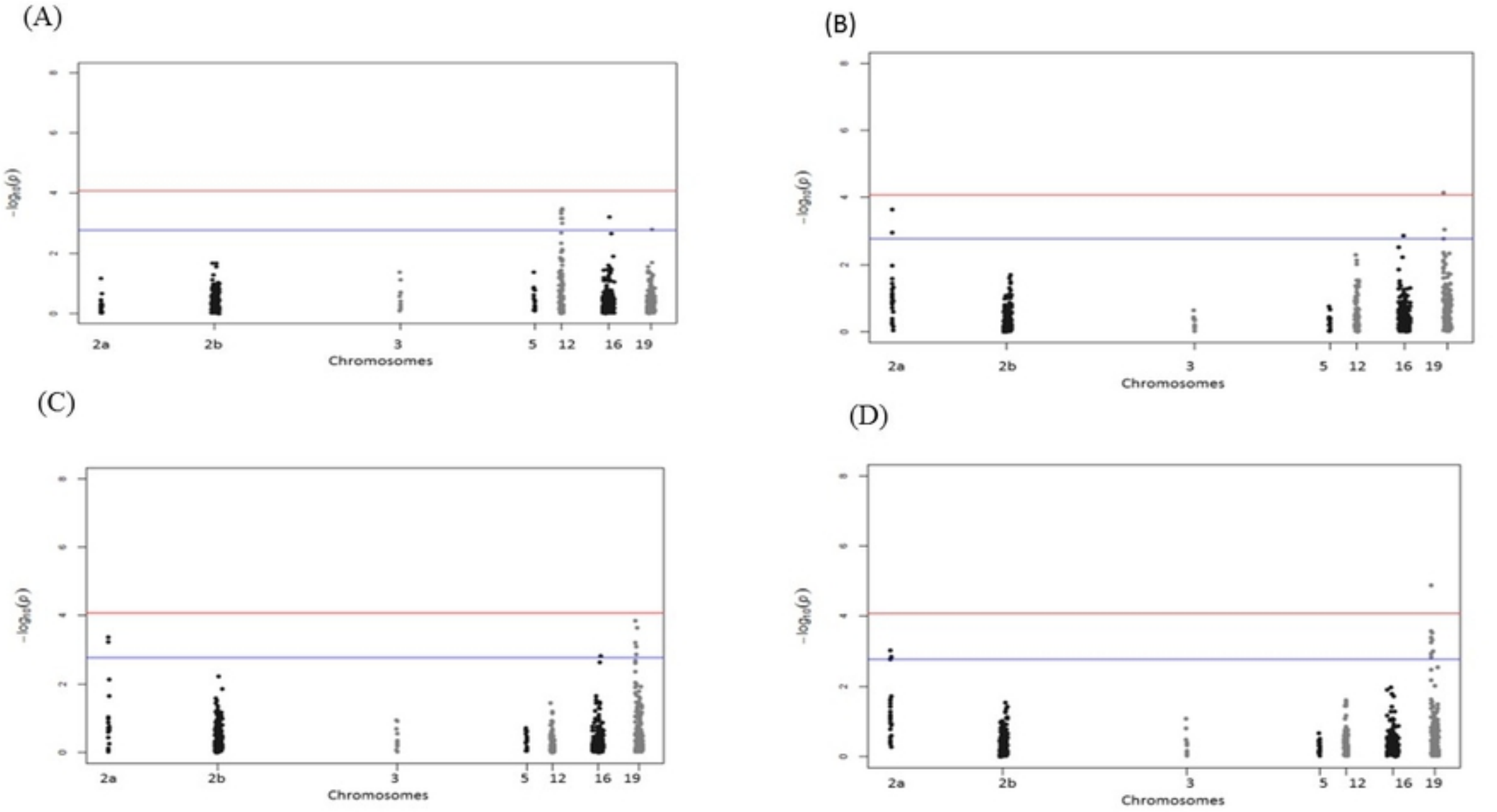

Figure 1

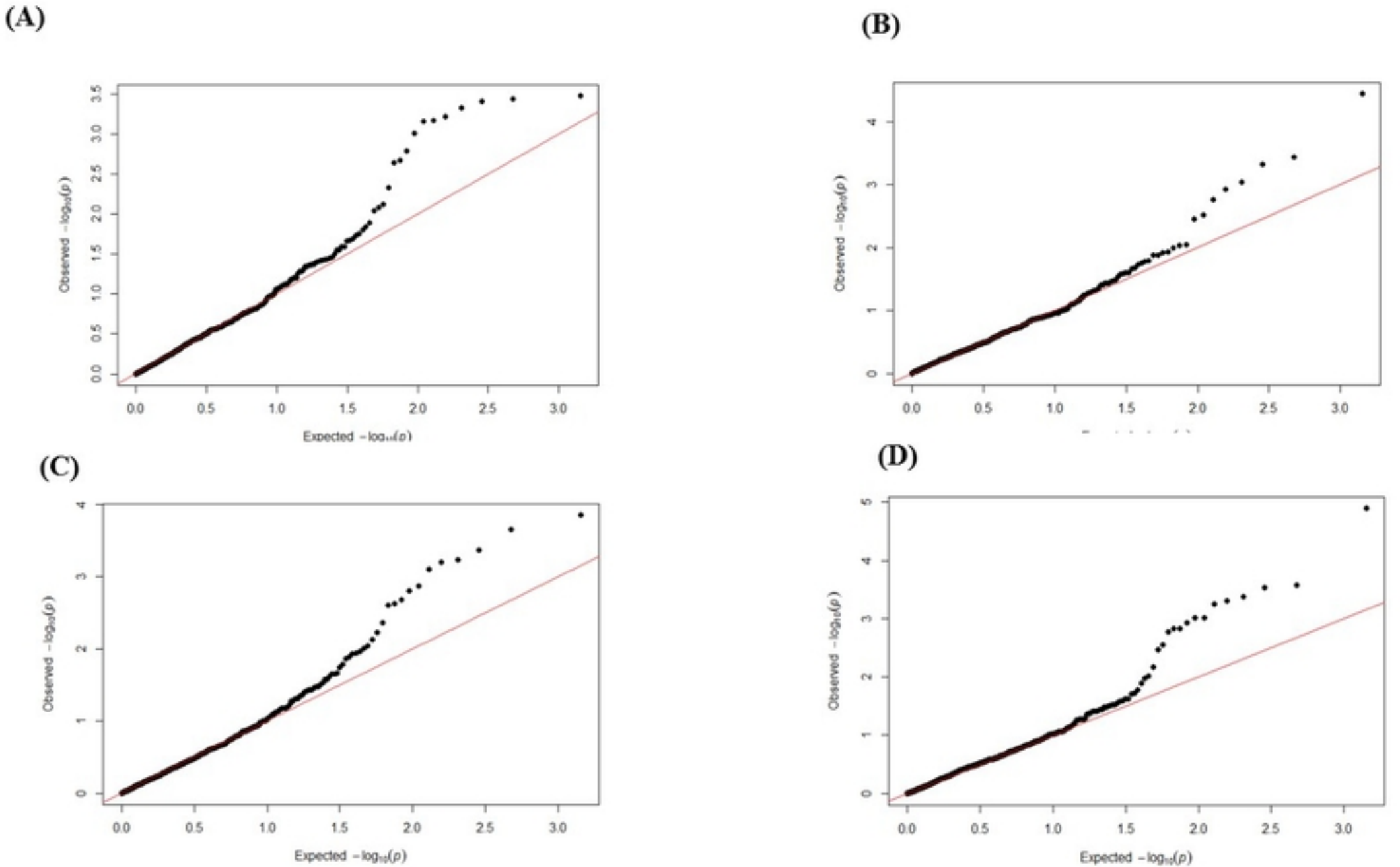

Figure 2

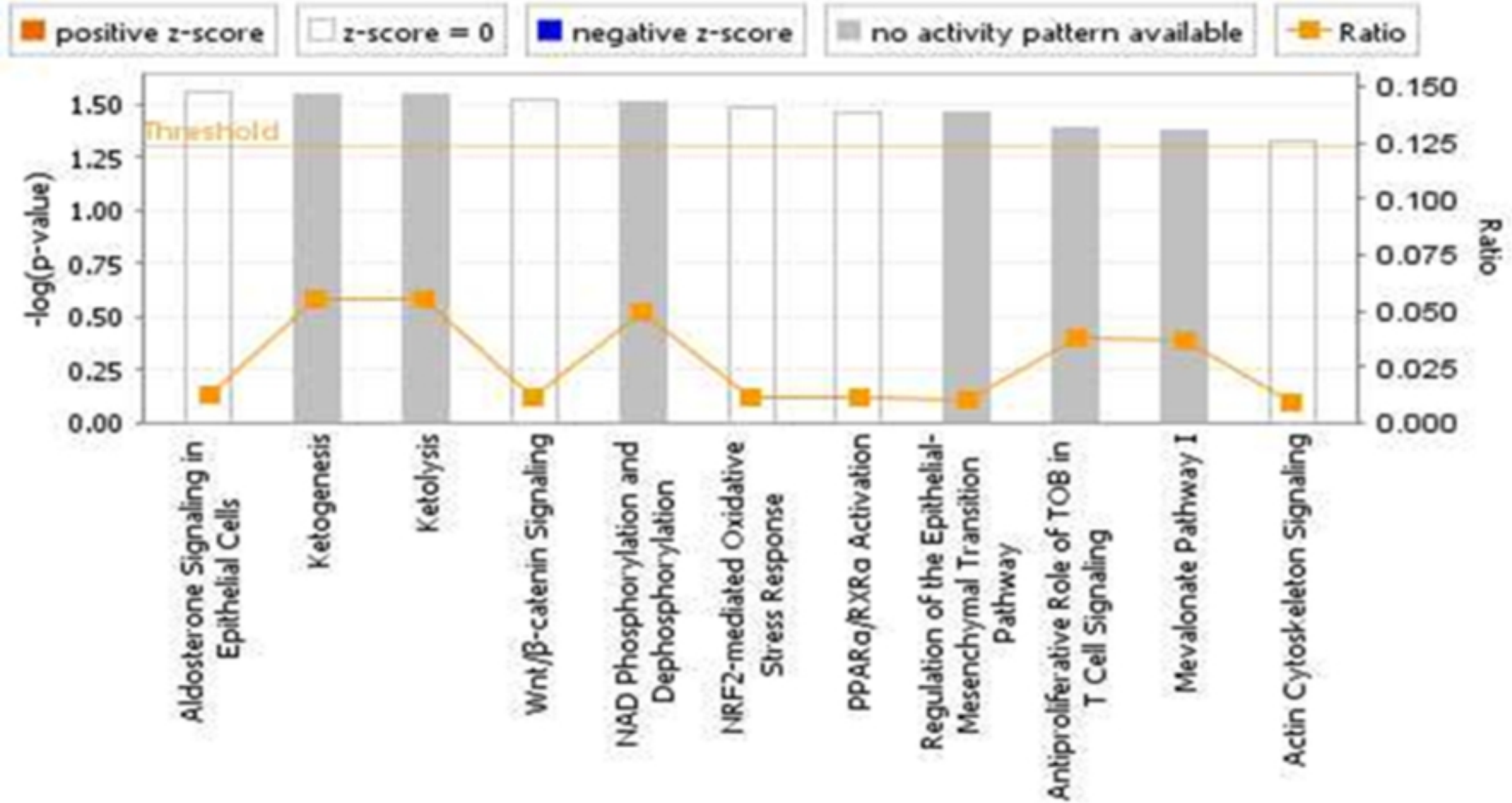
positive z-score
z-score = 0
negative z-score
no activity pattern available
Ratio
-log(p-value)
Ratio
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0.150
0.125
0.100
0.075
0.050
0.025
0.000
Threshold
Aldosterone Signaling in Epithelial Cells
Ketogenesis
Ketolysis
Wnt/β-catenin Signaling
NAD Phosphorylation and Dephosphorylation
NRF2-mediated Oxidative Stress Response
PPARα/RXRα Activation
Regulation of the Epithelial-Mesenchymal Transition Pathway
Antiproliferative Role of TOB in T Cell Signaling
Mevalonate Pathway I
Actin Cytoskeleton Signaling

Figure 3

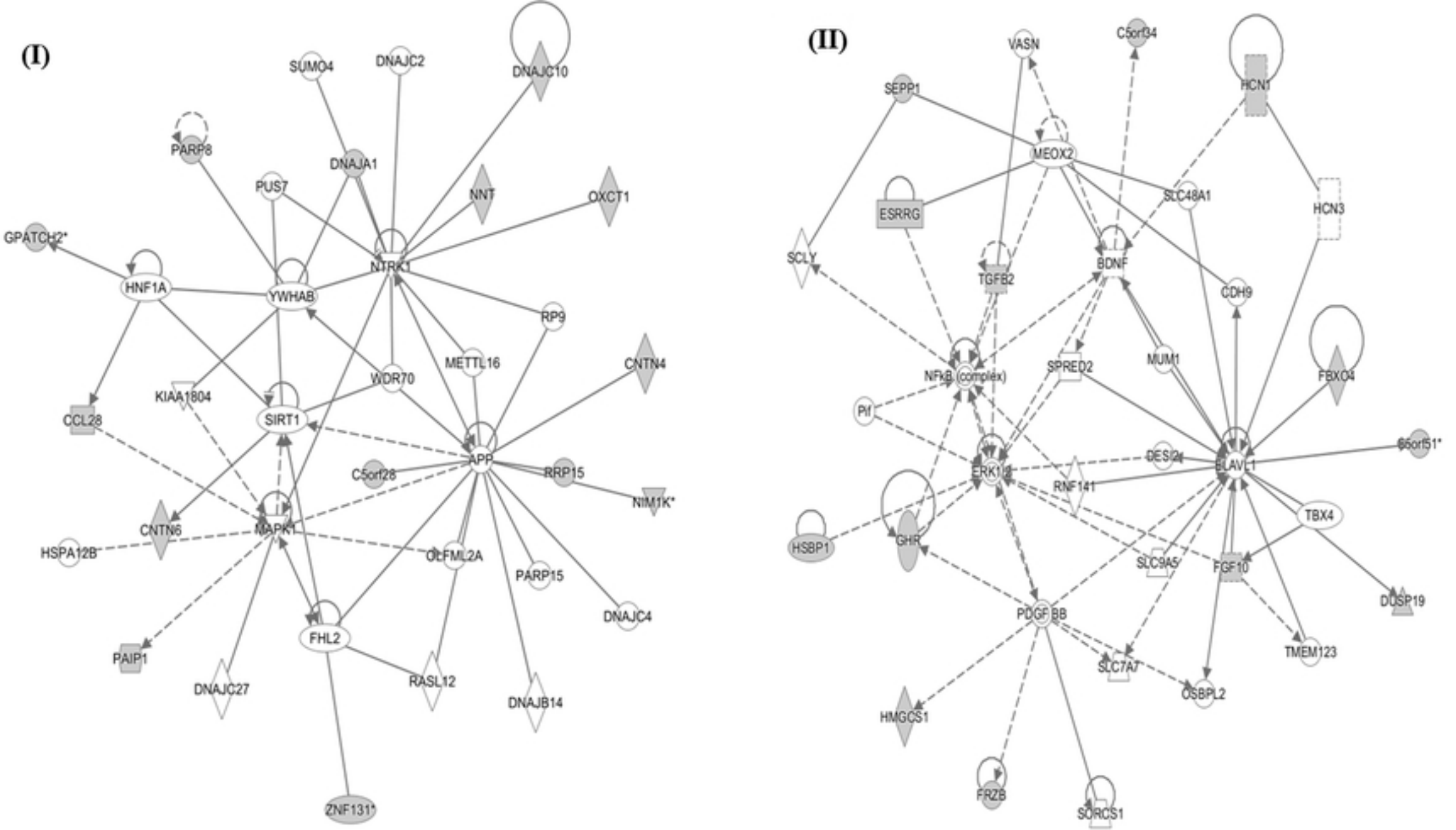
(I)
SUMO4
DNAJC2
DNAJC10
PARP8
DNAJA1
PUS7
NNT
OXCT1
GPATCH2*
NTRK1
HNF1A
YWHAB
RP9
METTL16
CNTN4
WDR70
KIAA1804
SIRT1
CCL28
APP
C5orf28
RRP15
NIM1K*
CNTN6
MAPK1
HSPA12B
OLFML2A
PARP15
DNAJC4
FHL2
PAIP1
DNAJC27
RASL12
DNAJB14
ZNF131*
(II)
VASN
C5orf34
SEPP1
HCN1
MEOX2
SLC48A1
ESRRG
HCN3
SCLY
TGFB2
BDNF
CDH9
NFkB (complex)
SPRED2
MUM1
FBXO4
Pif
DES(2)
ELAVL1
65orf51*
ERK1/2
RNF141
TBX4
HSBP1
GHR
SLC9A5
FGF10
DUSP19
PDGF BB
TMEM123
SLC7A7
OSBPL2
HMGCS1
FRZB
SORCS1

Figure 4