

Scotland's Rural College

Variants in the 3'UTR of the ovine Acetyl-Coenzyme A Acyltransferase 2 gene are associated with dairy traits and exhibit differential allelic expression

Miltiadou, D; Hager-Theodorides, AL; Symeou, S; Constantinou, C; Psifidi, A; Banos, G; Tzamaloukas, O

Published in:
Journal of Dairy Science

DOI:
[10.3168/jds.2016-12326](https://doi.org/10.3168/jds.2016-12326)

First published: 16/06/2017

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

[Link to publication](#)

Citation for published version (APA):

Miltiadou, D., Hager-Theodorides, AL., Symeou, S., Constantinou, C., Psifidi, A., Banos, G., & Tzamaloukas, O. (2017). Variants in the 3'UTR of the ovine Acetyl-Coenzyme A Acyltransferase 2 gene are associated with dairy traits and exhibit differential allelic expression. *Journal of Dairy Science*, *100*(8), 6285 - 6297. <https://doi.org/10.3168/jds.2016-12326>

General rights

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

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

Take down policy

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

1 **Interpretive summary**

2 **Variants in the 3'UTR of the ovine Acetyl-Coenzyme A Acyltransferase 2 gene**

3 **are associated with dairy traits and exhibit differential allelic expression.** *By*

4 *Miltiadou et al.* The present study shows that the ovine *ACAA2* gene is associated

5 with milk production, milk protein percentage and milk fat yield. Additionally, we

6 demonstrated that the mRNA expression in homozygous animals for each allele of the

7 *ACAA2* gene differs by several orders of magnitude in hepatic and mammary gland

8 tissues and the *ACAA2* variants are expressed at different levels in the udder of

9 heterozygous animals, suggesting the existence of regulatory variation within the

10 ovine *ACAA2* gene. These findings support the hypothesis that the *ACAA2* gene is a

11 functional candidate affecting dairy traits in sheep.

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28 **Variants in the 3'UTR of the ovine Acetyl-Coenzyme A Acyltransferase 2 gene**
29 **are associated with dairy traits and exhibit differential allelic expression.**

30

31 D. Miltiadou^{*1,2}, A. L. Hager-Theodorides^{†1}, S. Symeou^{*}, C. Constantinou^{*}, A.
32 Psifidi[‡], G. Banos^{‡,§} and O. Tzamaloukas^{*}

33

34 ^{*}Department of Agricultural Sciences, Biotechnology and Food Science, Cyprus
35 University of Technology, Lemesos, PO Box 50329, Cyprus

36 [†]Department of Animal Science, Agricultural University of Athens, Greece

37 [‡]Roslin Institute, University of Edinburgh, EH25 9RG Midlothian, UK

38 [§]Scotland's Rural College, EH25 9RG Midlothian, UK

39

40 ¹ Both authors contributed equally to this work.

41 ² Corresponding author: Department of Agricultural Sciences, Biotechnology and
42 Food Science, Cyprus University of Technology, PO Box 50329, Limassol, Cyprus,
43 tel: +357 25002590, fax: +357 25002767, Despoina.Miltiadou@cut.ac.cy

44

45

46

47

48

49

50

51

52

ABSTRACT

53 The Acetyl-CoA acyltransferase 2 (*ACAA2*) gene encodes an enzyme of the thiolase
54 family that is involved in mitochondrial fatty acid elongation and degradation by
55 catalyzing the last step of the respective β -oxidation pathway. The increased energy
56 needs for gluconeogenesis and triglyceride synthesis during lactation are met
57 primarily by increased fatty acid oxidation. Therefore, the *ACAA2* enzyme plays an
58 important role in the supply of energy and carbon substrates for lactation and may
59 thus affect milk production traits. This study investigates the association of the
60 *ACAA2* gene with important sheep traits and the putative functional involvement of
61 this gene in dairy traits. A single nucleotide substitution (SNP), a T to C transition
62 located in the 3' untranslated region (UTR) of the *ACAA2* gene, was used in mixed
63 model association analysis with milk yield, milk protein yield and percentage, milk fat
64 yield and percentage, and litter size at birth. The SNP was significantly associated
65 with total lactation production and milk protein percentage, with respective additive
66 effects of 6.81 ± 2.95 kg and -0.05 ± 0.02 %. Additionally, a significant dominance
67 effect of 0.46 ± 0.21 was detected for milk fat yield. Homozygous TT and
68 heterozygous CT animals exhibited higher milk yield compared to homozygous CC
69 animals, while the latter exhibited increased milk protein percentage. Expression
70 analysis from age, lactation and parity matched female sheep showed that mRNA
71 expression of the *ACAA2* gene from TT animals was 2.8 and 11.8 fold higher in liver
72 and mammary gland, respectively. In addition, by developing an allelic expression
73 imbalance (AEI) assay, it was estimated that the T allele was expressed at an average
74 of 18% more compared to the C allele in the udder of randomly selected ewes. We
75 demonstrated for the first time that the variants in the 3'UTR of the ovine *ACAA2*
76 gene are differentially expressed in homozygous ewes of each allele and exhibit AEI

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

77 within heterozygotes in a tissue specific manner, supporting the existence of cis-
78 regulatory DNA variation in the ovine *ACAA2* gene. This is the first study reporting
79 differential allelic imbalance expression of a candidate gene associated with milk
80 production traits in dairy sheep.

81

82 Key words (up to 4): *ACAA2* association, 3' *UTR* Cis-acting SNP, gene expression,

83 dairy sheep

THE OVINE ACAA2 GENE IS AFFECTED BY CIS-REGULATORY DNA VARIATION

84

INTRODUCTION

85 Milk yield represents more than two thirds of the total income of the dairy sheep
86 industry (Carta et al., 2009) and therefore the improvement of milk production is the
87 most important breeding objective. In Mediterranean countries, however, most of the
88 ovine milk produced is used for the production of cheese commercialized as products
89 of protected designation of origin (PDO) and other quality labels (Arranz and
90 Gutierrez-Gil, 2012). Thus, farmers' income is additionally determined by total solids
91 that affect cheese yield (De Rancourt et al., 2006) and, therefore, increased milk fat
92 and protein content is also highly desirable from an economic perspective (Ramon *et*
93 *al.*, 2010). Although traditional breeding programs have achieved appreciable genetic
94 gains mainly for milk yield, application of selection schemes assisted by molecular
95 information could expedite improvement (Carta et al., 2009). Moreover, marker
96 assisted selection could be of special interest for dairy sheep, due to the great regional
97 diversity of breeds, funding limitations, organizational difficulties (Arranz and
98 Gutierrez-Gil, 2012) and small population sizes (Garcia-Gamez et al., 2012).

99 To date, there are few reports of genome-wide association studies (GWAS) and
100 genome scans based on linkage mapping that detect quantitative trait loci (QTL) and
101 quantitative trait mutations (QTM) for ovine dairy traits (reviewed by Arranz and
102 Gutierrez-Gil, 2012; Garcia-Gamez et al., 2012; 2013; Gutierrez-Gil et al., 2014). In a
103 whole genome QTL study performed in Churra ewes, Gutierrez-Gil et al. (2009)
104 detected suggestive QTL for milk, fat and protein yields, mapped in a region of the
105 ovine chromosome 23 (OAR23) harbouring the Acetyl-CoA acyltransferase 2
106 (ACAA2) gene. The ACAA2 gene encodes an enzyme of the thiolase family, also
107 known as 3-oxoacyl-CoA thiolase or mitochondrial 3-ketoacyl-CoA thiolase. The
108 ACAA2 enzyme catalyzes the last step in mitochondrial fatty acid β -oxidation, thus

109 playing a central role in the supply of energy for the animal (Bartlett and Eaton,
110 2004). Therefore, due to the chromosomal location of the ovine *ACAA2* gene in
111 relation to the QTL described by Gutierrez-Gil et al., (2009) and its functional role in
112 lipid metabolism, it was regarded as a putative functional and positional candidate
113 gene that may affect milk yield and composition.

114 Genes encoding enzymes of the thiolase family have been correlated with
115 production traits in other livestock species. Single nucleotide polymorphisms (SNPs)
116 detected in the porcine *ACAA2* gene are reported to be associated with daily weight
117 gain and loin muscle area (Li HD, 2008). An important paralog of the *ACAA2* gene,
118 the acetyl-CoA acetyltransferase 2 (*ACAT2*) gene has been associated with production
119 and fertility traits (milk protein content, productive life and conception and pregnancy
120 rates) in Holstein cattle (Cochran et al., 2013), while SNPs within the swine *ACAT2*
121 gene were suggested to influence the metabolic functions of the corresponding
122 enzyme and thus may affect growth performance (Sodhi et al., 2014).

123 Our previous study showed that the entire mRNA (coding and untranslated
124 regions) of the ovine *ACAA2* gene is monomorphic in Chios sheep, one of the most
125 productive and extensively used breeds in Greece and Cyprus (Chatziplis et al., 2012),
126 with the exception of a SNP (HM537015:g.2982T>C) located in the 3' untranslated
127 region (*UTR*) of the gene (Orford et al, 2012). The SNP was significantly associated
128 with milk yield, at first lactation and across first to third lactations in Chios sheep.
129 Animals from a closed nucleus research flock at the Agricultural Research Institute of
130 Cyprus, carrying the g.2982TT or g.2982CT genotype had significantly higher milk
131 yield than those with the g.2982CC genotype and the g.2982T>C SNP explained 10%
132 of the additive genetic variance for milk yield when data up to third lactation from a
133 single flock were analysed (Orford et al., 2012).

134 It is well established that *UTRs* contain motifs involved in posttranscriptional
135 regulation of gene expression (Xie et al., 2005) that may lead to differential
136 expression of alleles associated with phenotypic diversity of production traits (Clop et
137 al., 2006; Khatib et al., 2007; Sugimoto et al., 2015). Studies in humans (Yan et al.,
138 2002, Bray et al., 2003), mice (Cowles et al., 2002), cattle (Khatib et al., 2007,
139 Olbromski et al., 2013) and pig (Muráni et al., 2009) have shown that alleles of non-
140 imprinted genes are not expressed equally at the mRNA level in heterozygous
141 animals, a phenomenon called allelic expression imbalance (AEI). AEI is the outcome
142 of the presence of at least one cis regulatory element in the regulatory sequences of a
143 gene (Campbell et al., 2008) and is therefore one of the possible mechanisms
144 underlying the effect of causative genetic variations that are not located on the
145 translated region of a gene.

146 The objective of the current study was to provide novel insights into the
147 association of the ovine *ACAA2* gene with important sheep traits. Firstly, we
148 performed an association analysis of the previously identified g.2982T>C SNP with
149 milk yield, fat and protein contents, fat and protein yields and litter size, in a
150 population of Chios sheep from all available farms keeping production records in
151 Cyprus. Upon confirmation of the association of the SNP with total milk production
152 and additional detection of association with milk protein percentage, we tested the
153 hypothesis that the *ACAA2* is a putative functional gene affecting dairy traits, by
154 comparing the expression of the gene in different genotypes and by developing an
155 assay to test the presence of AEI.

156

157

MATERIALS AND METHODS

158 *Animals and phenotypic data*

159 Data were collected from 742 purebred Chios ewes from five farms. Those were
160 one governmental farm (Orites, Cyprus) and the only 4 commercial farms in Cyprus
161 keeping phenotypic records of purebred Chios sheep according to the regulations of
162 the International Committee for Animal Recording (ICAR).

163 For all animals, individual records included month of lambing, year of lambing,
164 lactation number and age of lambing. Phenotypic data were obtained for lambing
165 years between 2009 and 2016 and included 1,514 individual records of 742 milking
166 ewes; this data included total lactation milk yield, 1,242 observations for litter size at
167 birth, 1,203 measurements of milk fat content and yield and 615 measurements of
168 milk protein content and yield, respectively. Total lactation yield was calculated for
169 each animal with the Fleischmann method with monthly tests on actual yields (sum of
170 *am* and *pm* records, ICAR 2014). Milk samples were obtained for fat and protein
171 analysis by combined thermo-optical procedures (LactoStar 3510, Funke Gerber,
172 Berlin, Germany), previously calibrated for protein with the Lowry protein assay and
173 for fat with the method 989.05 (AOAC International, 2005).

174

175 ***DNA extraction and SNP genotyping***

176 Whole blood samples were obtained from all 742 ewes for DNA extraction and
177 genotyping of the g.2982T>C SNP of the ACAA2 gene. Genomic DNA was isolated
178 from all samples using the Genomic DNA Blood kit (Macherey-Nagel), according to
179 the manufacturer's instructions, and DNA quality and quantity were estimated by UV
180 absorption at 260 and 280nm. Genotyping of the Chios ewes was performed by a cost
181 effective direct DNA sequencing protocol (Miltiadou et al., 2017), with primers
182 amplifying the 10th exon of the gene, using conditions previously described (Orford et
183 al., 2012).

184

185 *Mixed model association analysis*

186 Each trait was analyzed separately with the following mixed linear model:

$$187 \quad Y_{ijklmn} = \mu + F_i + YS_j + L_k + b_1age + b_2dur + G_l + A_m + e_{ijklmn} \quad (1)$$

188 where:

189 Y = lactation milk yield, fat yield/percentage, protein yield/percentage or litter size

190 record n of animal m

191 μ = overall population mean for the trait

192 F = fixed effect of flock i (1-5)

193 YS = fixed effect of year (2009-2016) by season (1-2) of lambing interaction j

194 L = fixed effect of number of lactation k (1-4)

195 b_1 = linear regression on age at lambing (age)

196 b_2 = linear regression on lactation duration (dur); yield traits only

197 G = fixed effect of genotype l in the g.2982T>C locus (1-3; CC, CT, TT)

198 A = random effect of animal m

199 e = random residual effect

200

201 All lactation records were analyzed for all traits. In a separate analysis, only first

202 lactation records were considered for milk yield. In the latter case, the fixed effect of

203 number of lactation and the random animal effect were removed from the model.

204 In all cases, predicted trait values for each SNP genotype and respective standard

205 errors were derived; these values were reflective of the marginal genotypic effect on

206 each trait adjusted for all other effects fitted in the model. The predicted trait values

207 were used to estimate additive and dominance SNP effects on traits, and the

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

208 proportion of phenotypic variance for each trait accounted for by the SNP locus. The
209 equations used were:

210 Additive effect, $a = (TT - CC) / 2$;

211 Dominance effect, $d = CT - [(TT + CC) / 2]$,

212 Percentage of phenotypic variance (V_P) due to SNP = $100 * [2pq(a + d(q-p))^2] / V_P$,

213 where TT, CC, CT are the predicted trait values for each genotype class and p, q

214 are the allele frequencies at the SNP locus. Variance components were estimated with

215 model (1) after excluding the genotype effect. All statistical analyses were conducted

216 with the ASReml3 software (Gilmour et al., 2009).

217 Since multiple traits were analyzed, we assessed the number of independent tests in

218 order to adjust for multiple testing. For this reason, we conducted multivariate

219 analyses based on model (1) after excluding the genotype effect to assess the

220 correlation among traits. Very high correlations were found among the three

221 production traits (milk, fat, protein yield; 0.89-0.99), and between the two milk

222 concentration traits (fat and protein percentage; 0.75). Correlations between the two

223 groups of traits were near zero. Correlations of these two groups with litter size

224 ranged from 0.05 to 0.19. Consequently, we regarded three separate trait groups

225 (production, concentration and litter size) that corresponded to three distinct,

226 independent hypothesis tests. This was confirmed with a Principal Component

227 Analysis of the studied traits (Supplementary Figure 1). Subsequently, a Bonferroni

228 adjustment for multiple tests was implemented based on the Holm-Bonferroni method

229 (Holm 1979). This method works sequentially testing first the lowest nominal P-value

230 against the threshold value of $0.05/n$, where n is the number of independent

231 hypotheses tests (three in the present study). If this hypothesis is rejected, the next

232 lowest nominal P-value is compared to $0.05/(n-1)$ and so on, until a hypothesis is not

233 rejected (Holm, 1979). Supplementary Table 1 includes the corresponding P-value
234 thresholds in the case of three independent tests.

235

236 ***Animals and tissue sampling for expression analysis***

237 Blood samples were collected from 161 first parity ewes from the biggest
238 commercial farm in Cyprus, allowing the selection of age, parity and lactation stage
239 matched animals. Based on the methodology described in our previous work (Orford
240 et al., 2012), we identified animals with the genotypes g.2982TT, g.2982CT and
241 g.2982CC, and chose 3 ewes from each genotype for subsequent biopsies and RNA
242 extraction. All 9 selected ewes were first parity, 15-month old and had given birth to
243 two lambs each.

244 Mammary and liver biopsies were obtained under anaesthesia, by a professional
245 veterinarian, 42 ± 2 days after lambing, a week after weaning. Liver parenchyma was
246 sampled via puncture biopsy using ultrasonography for the selection of the
247 appropriate biopsy site and for avoiding large vessels and other organs. Udder
248 biopsies were taken from either the left or the right rear gland. The biopsy site was
249 carefully selected to avoid large subcutaneous blood vessels. Preparation of the site
250 involved shaving and washing with dilute betadine solution followed by sanitizing
251 with ethanol (70%). Ewes were given intravenous xylazine before anesthetizing the
252 biopsy site by subcutaneous injection of lidocaine hydrochloride. An incision was
253 made (~0.5-1.0 cm) on the outside of the quarter using a scalpel blade (size 22). A
254 Bard[®] Magnum[®] core biopsy instrument (Bard Peripheral Vascular, Inc., Tempe, AZ)
255 with a Bard[®] Magnum[®] core tissue biopsy needle (MN1210, 12G × 10 cm) was used.

256 In addition, liver and udder *ex vivo* tissue samples were obtained from 16 randomly
257 selected Chios ewes from an abattoir straight after slaughtering, under sterile

258 conditions. Although tissue samples obtained from the abattoir were at different and
259 unknown parities and stages of lactation, they were taken to increase the number of
260 heterozygous samples for allelic expression imbalance analysis, without increasing
261 the veterinary costs required for biopsies.

262 All samples were snap frozen in liquid N₂ immediately after the tissue was
263 obtained under sterile conditions and stored at -80°C until RNA extraction.

264 Animal sampling and handling and all procedures in this study were carried out in
265 strict accordance with the national legislation for animal welfare 46I/94 and no
266 animals were sacrificed for the purposes of this study.

267

268 ***DNA, total RNA extraction and reverse transcription***

269 Frozen tissue biopsies and ex vivo samples were homogenized using a mortar and
270 pestle constantly covered under liquid N₂ and aliquoted in tubes containing ~20 mg
271 each. DNA was extracted using the Genomic DNA Nucleospin Tissue kit (Macherey-
272 Nagel, Germany), while ~20 mg of frozen homogenized tissue was subjected to RNA
273 extraction using the RNA isolation Nucleospin RNA kit (Macherey-Nagel, Germany),
274 according to manufacturer's instructions. Before reverse transcription, extracted RNA
275 was incubated with 2U/μl rDNAase (Macherey-Nagel, Germany) for 10 min at 37°C
276 to eliminate putative contaminating genomic DNA and then purified using the RNA
277 clean up Nucleospin RNA Clean up XS kit (Macherey-Nagel, Germany). The
278 concentration and quality of DNA and RNA for all samples was measured using a
279 Nanodrop 1000 UV/VIS spectrophotometer (Thermoscientific Nanodrop
280 Technologies LLC, USA). The RNA integrity was assessed by electrophoretic
281 analysis of the 28S and 18S rRNA subunits.

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

282 Complementary DNA (cDNA) was synthesized from 0.5 µg RNA, using the High
283 Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA), using random
284 hexamer primers following the manufacturer's recommendations in a final volume of
285 20 µl. The lack of genomic DNA contamination was verified by a PCR amplification
286 using primers amplifying intron 2 of the prolactin gene (*PRL*), as described by Orford
287 et al. (2010) and by including a control without reverse transcriptase during cDNA
288 synthesis of each sample. These controls were used for conventional PCR
289 amplification of the *ACAA2* exon 10, as we have previously described (Orford et al.,
290 2012) and for subsequent real time PCR (RT-PCR) and verified lack of genomic
291 contamination.

292

293 *ACAA2 mRNA expression*

294 *ACAA2* relative expression in liver and udder was assessed by quantitative Reverse
295 Transcriptase PCR (qRT-PCR). *ACAA2* qRT-PCR was performed with the primers
296 used for genotyping as described above and expression levels were normalized by the
297 expression of reference genes, ovine β-actin (*ACTB*) for liver samples and ovine β2
298 microglobulin (*ovB2M*) for udder samples, utilizing primers: ovACTB Forward
299 (GCAAAGACCTCTACGCCAAC) and ovACTB Reverse
300 (TGATCTTGATCTTCATCGTGCT) (Sari et al., 2009) and for ovB2M Forward
301 (CTGTGCTGTCTGGACTGG) and ovB2M Reverse
302 TTTCCATCTTCTGGCGGGTG (designed using the NCBI/Primer-BLAST tool). For
303 each gene and cDNA sample qRT-PCR reactions were performed in triplicates.
304 Reactions of no template negative control and 5 point 1:5 serial dilutions of reference
305 cDNA, to obtain the reaction standard curves and calculate amplification efficiencies,
306 were also performed in triplicates. Reactions for *ACAA2* expression were performed

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

307 in a final volume of 20 μ l containing 1 \times KAPA SYBR fast qPCR Master Mix (Kapa
308 Biosystems), 100 nM of each of *ACAA2*_SNP Forward and Reverse primers, 1 \times ROX
309 low and cDNA. Reactions for *ACTB* and *B2M* were performed in a final volume of
310 10 μ l containing 1 \times KAPA SYBR fast qPCR Master Mix (Kapa Biosystems), 200nM
311 of each of the ov*ACTB* or ov*B2M* Forward and Reverse primers, 1 \times ROX low and
312 cDNA. Cycling conditions included a 3 min initial denaturation step at 95 $^{\circ}$ C followed
313 by 40 cycles of: (i) denaturation for 3 sec at 95 $^{\circ}$ C and primer annealing and template
314 extension for 30 sec at 60 $^{\circ}$ C, for *ACAA2* (ii) denaturation for 3 sec at 95 $^{\circ}$ C, primer
315 annealing at 62 $^{\circ}$ C for 20 sec and template extension at 72 $^{\circ}$ C for 30 sec, for *ACTB* and
316 (iii) denaturation for 3 sec at 95 $^{\circ}$ C, primer annealing at 64 $^{\circ}$ C for 15 sec and template
317 extension at 72 $^{\circ}$ C for 30 sec, for *B2M*. For all three genes, a melting curve analysis
318 was performed following the amplification cycles consisting of 15 s at 95 $^{\circ}$ C, 30 s at
319 the respective annealing temperature for each gene and continuous heating and data
320 collection up to 95 $^{\circ}$ C at a rate of 1% temperature increase per 30 seconds, to evaluate
321 specificity of the amplification. Reactions were optimized for amplification efficiency
322 to be between 85-95% and linear standard curve fit (r^2) to be greater than 0.990 for all
323 genes. Raw data were analyzed with the 7500 software v2.3 (Applied Biosystems,
324 USA), and mean Ct values and PCR reaction efficiencies were exported. The
325 normalized expression (nQ) of *ACAA2* was calculated according to the delta-Ct
326 method (Livak and Schmittgen, 2001), i.e. $nQ = E_{ACAA2}^{(\min_{ACAA2_Ct} - \text{sample}_{ACAA2_Ct})} / E_{ref}^{(\min_{ref_Ct} - \text{sample}_{ref_Ct})}$, where E is the PCR efficiency for a given
327 gene (values are greater than 1 and 2 corresponds to ideal doubling of templates per
328 cycle, 100% = 2) estimated based on standard cDNA dilution series reactions,
329 \min_{gene_Ct} is the minimum Ct value among all samples for a given gene and
330 sample_{gene_Ct} is the Ct value for a given gene and sample. To express *ACAA2* nQ as

332 fold change relative to nQ in CC genotype (nrQ) nQ for each sample was divided by
 333 mean nQ in CC samples. Data were then analyzed using the mixed models procedure
 334 of SAS software (SAS, 2005) with genotype as the fixed effect. Mean comparisons
 335 were performed with Tukey's adjustment with significance level set at 0.05.

336

337 ***Allelic Expression Imbalance (AEI) analysis***

338 Sequencing of the *ACAA2* gene from genomic DNA and cDNA of each sample
 339 was initially used as a semi-quantitative approach to detect differential expression of
 340 the *ACAA2* alleles. RT-PCR products amplified from heterozygous individuals were
 341 sequenced and data were analyzed using Applied Biosystems' Sequencing Analysis
 342 (Applied Biosystems, USA). The SNP was identified by visually inspecting each base
 343 in sequencing traces. Allelic variation was estimated by comparing the proportions of
 344 the peak heights of the two alternative alleles of the SNP.

345 Subsequently, the detection of AEI was based on quantitative analysis of mRNA
 346 transcripts using a TaqMan probe qRT-PCR assay (Livak, 1999, Chen et al., 2008) in
 347 order to detect deviations from the null hypothesis expecting equimolar ratio between
 348 the two alleles in heterozygous samples. TaqMan qPCR reactions were performed in
 349 10 µl reaction volumes containing 1× Type-it® Fast SNP Probe PCR Master Mix
 350 (Qiagen, USA), 400 nM of each PCR primer and 200 nM of each TaqMan probes
 351 (primer and probe sequences as described in the *DNA extraction and SNP genotyping*
 352 section), and liver or udder cDNA or gDNA from *ACAA2* heterozygous ewes. A 7500
 353 Real Time PCR system (Applied Biosystems, USA) was used with a cycling profile
 354 that included a 5 min denaturation step at 95°C and 40 cycles of 15 sec at 95°C and
 355 30 sec at 62°C.

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

356 Pooled genomic (g) DNA from three *ACAA2* heterozygous animals (genotype
357 confirmed by several replications of genotyping reactions) was used as standard
358 gDNA and 5 points of 1:4 serial dilutions were used to construct the standard curve
359 for relative *ACAA2* T and C allele quantification (Fig. 1A-C). The T:C allele copy
360 ratio is expected to equal 1 in all dilutions.

361 The quantitation relative to standard curve method was employed, using the
362 homonymous option of the ABI 7500v2.3 software to extract the value for
363 “normalized quantity” of T allele at each qPCR reaction, having set the C allele as the
364 “endogenous control”. The “normalized quantity” of the T allele was computed as the
365 quantity of T allele divided by the quantity of C allele and thus corresponds to the
366 ratio of T:C *ACAA2* allele quantity. T and C allele quantities in each qPCR reaction
367 were computed from the respective Ct values and the linear prediction function for
368 each of the *ACAA2* alleles, computed by the software based on standard gDNA
369 dilution series. The quantity of T and C alleles in the standard gDNA was set at the
370 same arbitrary value and thus T:C ratio equalled 1 at all points of the dilution series of
371 the standard gDNA. The T:C ratio for each sample, i.e. the mean normalized quantity
372 of the three replicates performed with each cDNA or gDNA, was divided by the
373 average normalized quantity of the T allele in all the gDNA samples to obtain the
374 corrected T:C ratio for each cDNA sample. Average corrected T:C ratio in the gDNA
375 samples equalled 1 and ranged between 0.90-1.09. Standard gDNA dilution series
376 reactions were performed in each qPCR run. In all runs, efficiency values for both T
377 and C allele amplification were very similar (102.01% for T and 101.23% for C in run
378 1 and 102.66% for T and 102.74% for C in run 2). Linear standard curve r^2 values
379 were 0.997 for T, 0.996 for C in run 1 and 0.998 for T and 0.998 for C in run 2.

380 Statistical analysis of the T:C ratios in cDNA or gDNA was performed using the
381 mixed models procedure of SAS software (SAS, 2005) with tissue (liver or udder) as
382 the fixed effect. Means are presented as ordinary means and se. Least square mean
383 pairwise comparisons were performed with Tukey's adjustment with significance
384 level set at 0.05.

385

386

RESULTS

387 *Association of the ACAA2 gene with sheep traits*

388 Allelic frequencies in the g.2982C/T SNP locus were 0.54 for the T allele and 0.46
389 for the C allele; genotypic frequencies were 0.27, 0.54 and 0.19 for TT, CT and CC,
390 respectively. Genotypic frequencies were found to deviate from the Hardy-Weinberg
391 equilibrium ($p = 0.019$).

392 Marginal predicted means for the three genotype classes for all production traits
393 are presented in Table 1. After applying Holm-Bonferroni correction, the SNP was
394 significantly associated with milk yield at first lactation ($P < 0.01$) and all lactations
395 ($P < 0.025$), with respective additive effects of 10.61 ± 3.56 kg and 6.81 ± 2.95 and
396 respective positive dominance effects of 13.02 ± 4.26 kg and 8.67 ± 3.53 (Table 2).
397 Significant differences were found between the CC and CT genotype pairs ($P < 0.01$)
398 and between the CC and TT classes ($P < 0.025$) at the g.2982T>C SNP locus for milk
399 yield (Table 1). These results overall suggest a complete dominance effect at the
400 locus, since heterozygous CT animals exhibit similar predicted mean values compared
401 to homozygous TT animals. Based on the estimated allelic effects and the allele
402 frequencies observed in the sample, it was estimated that the g.2982T>C SNP
403 explained 2.25% and 0.62% of the total phenotypic variance for 1st lactation and for
404 all lactations, respectively (Table 2). Overall, these results suggest a stronger

405 association in first lactation milk yield, whereas the effect was reduced in subsequent
406 lactations.

407 After applying Holm-Bonferroni correction, the SNP was also associated with
408 protein percentage ($P < 0.01$), with a significant additive effect of -0.05 ± 0.02 (Table
409 1). Pairwise contrasts between the predicted protein percentage values showed
410 significant differences between the genotype classes CC and CT ($P < 0.001$) and
411 between CC and TT ($P < 0.025$), with CC animals exhibiting higher protein
412 percentage, whereas differences between CT and TT were not significant (Table 1). It
413 was estimated that the SNP explained 1.68% of the phenotypic variance for protein
414 percentage (Table 1). Additionally, there was a dominance effect of 0.46 ± 0.21 ($p =$
415 0.027) and 0.41 ± 0.21 ($p = 0.048$) of the SNP on fat and protein yields, respectively.
416 Homozygous TT and heterozygous CT animals exhibited significantly higher milk fat
417 yield compared to homozygous CC ewes ($P < 0.025$ and $P < 0.001$, respectively). No
418 significant associations of the SNP genotype with fat percentage, protein yield or litter
419 size at birth were found.

420

421 *mRNA expression analysis*

422 The expression of *ACAA2* was found to be significantly increased in TT compared
423 to CC udder and liver biopsies from age matched, at the same lactation stage and
424 parity ewes that gave birth to two lambs each. In particular, homozygous TT animals
425 showed a 2.8 fold increase in mRNA expression levels compared to homozygous CC
426 animals in liver ($P < 0.05$), while TT ewes exhibited an 11.8 fold increase in mRNA
427 of the *ACAA2* gene in the udder compared to CC ewes ($P < 0.05$) (Fig. 2A). *ACAA2*
428 mRNA expression of heterozygous animals showed a tendency for reduction
429 compared to homozygous TT animals only in liver ($p = 0.07$), while it did not

430 significantly differ from homozygous CC animals. Heterozygous mRNA expression
431 in the udder exhibited a very high standard deviation (Fig. 2A) and did not differ
432 significantly from homozygote expression.

433 The expression of *ACAA2* was also assessed in liver and udder *ex vivo* samples
434 from 1 CC, 10 CT and 5 TT ewes (Fig. 2B). These ewes were selected at random and
435 were not age matched nor at the same milking season or lactation stage. Expression of
436 *ACAA2* was found to be significantly increased in the liver compared to the udder (P
437 < 0.0001), but no significant differences were observed between CT and TT ewes in
438 either the liver or the udder ($P > 0.05$). Since only one ewe from the slaughterhouse
439 was genotyped as CC, it was not possible to make comparisons between
440 homozygotes.

441

442 *Allelic expression imbalance analysis (AEI)*

443 To test whether the differences observed between homozygous TT and CC
444 animals, was due to AEI between the T and C alleles, the transcription of T and C
445 alleles was quantified in liver and udder of heterozygous individuals. Sequencing of
446 genomic DNA and cDNA from 12 heterozygous samples (3 from biopsies and 9 from
447 slaughterhouse) was initially performed to evaluate semi-quantitatively whether there
448 was allelic expression imbalance. The allele specific expression levels were evaluated
449 by comparing the peak heights of C and T in cDNA samples from heterozygous
450 individuals compared to the peak heights of the respective genomic DNA. Although
451 in all genomic DNA samples from heterozygotes, peak heights of both nucleotides at
452 the SNP position were tangential, suggesting no preferential amplification, peak
453 heights from heterozygous cDNA samples were different, with the T allele exhibiting
454 a higher peak compared to the C allele (Fig. 3).

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

455 To further confirm the initial results from the semi-quantitative method consistent
456 with AEI in *ACAA2* gene, we developed an AEI assay using TaqMan probes to
457 estimate the T:C transcription ratio in cDNA samples corrected by the mean T:C ratio
458 in gDNA samples and ultimately test for deviations from the expected ratio of 1 in the
459 absence of AEI. Results of the assay for the control gDNA reactions showed that
460 observed T:C ratios were, as expected, very close to 1 (average non-corrected ratio
461 was 1.052). *ACAA2* AEI was observed in individual ewes in both the liver and the
462 udder (Fig. 4). The average corrected T:C ratio in the liver cDNA was 1.026, not
463 significantly different from the average corrected gDNA ratio (mean = 1, $p = 0.67$).
464 On the contrary, corrected T:C ratio of transcription in the udder of heterozygous
465 animals was 1.18, significantly increased compared to both gDNA and liver cDNA
466 mean corrected T:C ratios ($p = 0.0054$ and 0.0134 respectively). Therefore, we
467 observed a moderate and tissue specific imbalance in the allelic expression of the two
468 *ACAA2* alleles in favour of allele T in the udder, but not in the liver of heterozygous
469 animals.

470

471

DISCUSSION

472 The study confirmed the previously observed association of the
473 HM537015:g.2982T>C *ACAA2* SNP with milk yield in an extended population of
474 Chios sheep from multiple flocks and showed that the g.2982T>C SNP of the *ACAA2*
475 gene is also associated with milk protein percentage and milk fat yield. Investigation
476 of the expression of the *ACAA2* gene from the three different genotypes and allelic
477 expression imbalance analysis (AEI) of heterozygous samples supported the
478 hypothesis that *ACAA2* is a functional gene affecting dairy traits. To the best of our

479 knowledge, this is the first study showing differential allelic imbalance expression of
480 a candidate gene associated with milk production traits in dairy sheep.

481 Consistent with our previous study (Orford et al., 2012), the g.2982T>C SNP, was
482 significantly associated with milk yield, with the T allele exhibiting positive additive
483 and dominance effects, mainly attributed to first lactation production data.
484 Confirmation of a previously detected association using a bigger and different data set
485 strengthens the evidence for the observed association (Sasaki et al., 2013). The overall
486 effects estimated in the present study, however, were lower compared to Orford et al.,
487 (2012), possibly due to the increased variation introduced from the use of multiple
488 flocks managed in different ways, as a substantial fraction of the environmental
489 variance for production traits is attributed to farm (Carta et al., 2009; Sasaki et al.,
490 2013). In the current analysis, we adjusted for the systematic effect of the flock and,
491 therefore, the estimates are more representative and indicative of the true effect of the
492 gene in the population. The reasons why the effects are more important in the first
493 lactation than in the rest of lactations are not clear at this stage and therefore further
494 research is needed.

495 In the present study, additional evidence about the correlation of the *ACAA2* gene
496 with important sheep traits is provided, as the g.2982T>C SNP was found to be
497 significantly associated with milk protein percentage. Homozygous CC animals
498 exhibited superior values for protein percentage compared to both homozygous TT
499 and heterozygous CT animals, in contrast to their inferior values estimated for total
500 milk yield (Table 1). This is consistent with the negative genetic correlation between
501 those two traits (Bencini and Pulina, 1997; Fuertes et al., 1998). Since, protein yield
502 was not significantly associated with this SNP, the decrease on protein content could
503 be attributed to a dilution effect, due to the increase of milk yield (Emery, 1988).

504 Similarly, milk yield is known to be negatively correlated to fat content (Fuertes et
505 al., 1998). However, although the marginal predicted mean for the fat content from
506 CC ewes was higher compared to that of the other genotypes in the present study, the
507 differences were not significant and the investigated SNP was not significantly
508 associated with milk fat percentage (Tables 1 and 2), in consistence with the results of
509 Orford et al. (2012). In agreement with other studies where fat is the most variable
510 component of ovine milk (Othmane et al., 2002; Pulina et al., 2005), in the current
511 study, standard errors for fat content were double compared to those for protein
512 content. Therefore, low precision of data for fat content may be a reason for not
513 detecting the association of the studied SNP with fat percentage. However, there is a
514 significant association of the SNP with fat yield (Table 2). In addition, homozygous
515 CC animals exhibit significantly lower fat yields, compared to heterozygous CT and
516 homozygous TT animals (Table 1), that could be attributed to decreased milk yield
517 with similar fat content.

518 Since fat and protein content are crucial for cheese-making, the potential use of the
519 studied SNP in a selection scheme requires applied research to identify whether each
520 additional unit of milk production compensates for the lower cheese yield due to
521 lower protein percentage, based on farmer prices and cheese yield for a certain type of
522 cheese. Using current prices in Cyprus for sheep milk and the values of each genotype
523 in Table 1, the producer is going to get 1.012 euros extra per additional liter, whereas
524 the price of the milk is reduced only by 0.012 (from 1.024 to 1.012) per liter due to
525 the lower protein and fat content. Therefore, selecting for the T allele is expected to
526 be beneficial. However, prior to incorporation of the SNP in any breeding program,
527 negative effects of selection on other important traits, for instance fertility, need to be
528 excluded. To facilitate further research, the SNP has been incorporated into the sheep

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

529 HD SNP chip developed by the International Sheep Genomics Consortium (ISGC) for
530 functional studies (James Kijas, Csiro, Australia, personal communication).

531 Allele frequencies of the g.2982T>C polymorphism genotyped were similar to the
532 previously observed (T:0.56; C:0.44; Orford et al., 2012). However, genotypic
533 frequencies were found to deviate from the Hardy-Weinberg (H-W) equilibrium in the
534 studied population, in contrast to the previously observed frequencies from a single
535 experimental flock (Orford et al., 2012), due to higher frequencies of TT and CT ewes
536 carrying the favourable for milk yield T allele. This could be possibly explained by
537 the fact that farmers select animals mainly based on milk yield, due to low meat prices
538 during the last decade, whereas the selection indices used at the experimental flock
539 (Orford et al., 2012) combine the individual capacity of young stock for growth and
540 milk production (Mavrogenis and Constantinou, 1991).

541 The candidate gene approach followed in the current study is an alternative to QTL
542 and GWA studies that can be very powerful in the identification of loci even with
543 small effects on the trait, if the candidate gene represents a true causative gene
544 (Andersson, 2001). Otherwise, an association detected could occur due to linkage
545 disequilibrium to linked or non-linked causative genes (Andersson, 2001). Therefore,
546 the polymorphism in the 3' UTR mRNA of the ACAA2 gene that was found
547 significantly associated with dairy traits, can be either the functional polymorphism to
548 which the genetic variance can be attributed or can be in linkage disequilibrium with
549 the functional polymorphism or both. Since the ACAA2 g.2982T>C SNP is in the
550 untranslated region of the gene and thus does not affect the protein produced, it may be
551 a functional polymorphism playing a regulatory role, by causing alterations in the
552 expression levels of the gene. Identification of regulatory variants requires studying
553 two alleles of a gene under identical circumstances and comparing the expression

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

554 associated with each allele (Cowles et al., 2002). For that reason, the expression of the
555 *ACAA2* variants was investigated in first parity, age, lactation stage and litter size
556 matched ewes that were managed, fed and sampled uniformly to eliminate
557 confounding sampling and environmental effects. The practical difficulties of
558 obtaining those uniform conditions resulted in the selection of 3 animals per
559 genotype. Three biological replicates are accepted as a minimum number of animals
560 for mRNA expression analysis, as shown by recent published studies (Shi H. et. al,
561 2015; Yao D. W. et al., 2016 a, b; Weller et al., 2016). The expression of the gene in
562 both the udder and the liver was significantly influenced by the animals' genotype,
563 with ewes homozygous for the T allele showing significantly increased expression, by
564 several orders of magnitude, compared to ewes homozygous for the C allele. The
565 reasons why heterozygous mRNA expression in the udder exhibited a very high
566 standard deviation needs to be elucidated by further research.

567 To assess whether the observed genotype effects in the expression of the *ACAA2*
568 gene could be attributed to allelic expression imbalance (AEI), an AEI assay using
569 TaqMan probes was designed. The advantage of measuring AEI compared to total
570 transcript levels is the reduction of the confounding effect of trans-acting factors,
571 because the alleles are compared within and not across individuals (Bray et al., 2003;
572 Cowles et al., 2002; Forton et al., 2007). The results obtained are in support of
573 differential expression of the two alleles of the *ACAA2* gene due to a cis-acting
574 mechanism. Differential allelic expression has been found in a large proportion of
575 human (Bray et al., 2003; Yan et al., 2002), mouse (Campbell et al., 2008) and cattle
576 (Karim et al., 2011, Olbromski et al., 2013, Sasaki et al., 2013) genes and there is
577 growing evidence that polymorphisms in regulatory DNA and the resulting variability
578 in gene expression, can explain a significant proportion of disease susceptibility and

579 quantitative trait phenotypic variance (Bray et al., 2003, Khatib, 2007, Karim et al.,
580 2011, Sasaki et al., 2013).

581 Consistent with other studies (Yan et al., 2002; Bray et al., 2003), the differential
582 allelic expression of the *ACAA2* gene varied among individuals in post-mortem
583 random samples, with an overall observed increase of T relative to C allele expression
584 in the udder of heterozygous ewes close to 20%. Previous studies in cattle indicate
585 that even small gene expression imbalances may result in large phenotypic variance of
586 complex traits (Karim et al., 2011, Sasaki et al., 2013). Interestingly, mean AEI of the
587 *ACAA2* gene was only significant in the udder but not in the liver. Such tissue specific
588 AEI has also been reported for cattle genes (Olbromski et al., 2013, Chamberlain et
589 al., 2015). This finding, together with the observed difference between mammary
590 gland and liver in the average mRNA gene expression in TT relative to CC ewes (11.8
591 and 2.8 fold increase, respectively), suggest an organ specific differential expression.
592 Preferential, or higher expression in organs related to a quantitative trait is an
593 important criterion for the selection of functional candidate genes (Ron and Weller,
594 2007; Stickens et al., 2010).

595 Since no other polymorphisms have been detected in the mRNA sequence of the
596 *ACAA2* gene, the differential mRNA expression of homozygotes and the observed
597 AEI in the mammary gland could be attributed to the g.2982T>C SNP acting in cis to
598 modulate gene transcription or mRNA survival (Pesole et al., 2001). Analysis of the
599 region harbouring the SNP, using the MicroInspector program found differences in
600 potential miRNA binding sites between the two alleles (Orford et al., 2012). However,
601 since this in-silico analysis requires functional experimental support, the possibility of
602 another cis acting regulatory polymorphism in linkage disequilibrium with the
603 g.2982C>T SNP, e.g. in the promoter region of the gene, cannot be excluded.

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

604 Moreover, it is also likely that other trans-acting factors or mutations in other genes in
605 linkage disequilibrium with *ACAA2*, found here to be associated with milk yield and
606 protein content, may explain part of the QTL for milk, fat and protein yields observed
607 in the region harbouring the *ACAA2* gene (Gutierrez- Gil et al., 2009).

608 *ACAA2* is involved in mitochondrial fatty acid elongation and degradation (KEGG
609 database: <http://www.kegg.jp/>), by catalyzing the last step of the respective β -
610 oxidation pathway. In agreement with our data, although *ACAA2* is predominantly
611 expressed in liver, it has also been reported to be expressed in the mammary gland
612 (<http://www.genecards.org/cgi-bin/carddisp.pl?gene=ACAA2>; Paten et al., 2015). In
613 mammals, excess energy is stored primarily as triglycerides, which are mobilized
614 when energy demands arise. During periods of underfeeding or in early lactation,
615 ruminants cover their increased energy demands by mobilizing fat from the adipose
616 tissue (Drackley, 1999). Fatty acids coming from triglycerides are taken up by the
617 liver where they are either used as energy source or converted to ketone bodies that
618 may be released into the blood and used as energy or substrates for de novo fatty acid
619 synthesis in the mammary gland. In addition, it is estimated that gluconeogenesis is
620 usually increased two- to three-fold during early lactation to meet the demands of the
621 mammary gland for lactose and triglyceride synthesis (Drackley, 2000; Vernon,
622 2005). The increased energy needs for gluconeogenesis and triglyceride synthesis are
623 met primarily by increased fatty acid oxidation. Additionally, a recent transcriptomics
624 analysis of the ovine mammary gland has shown increased expression of genes of the
625 β -oxidation pathway during late pregnancy (Paten et al., 2015). This contributes to the
626 development of the udder that is crucial for subsequent lactation, as the number of
627 active secretory cells primarily determines milk yield (Pollott GE, 2004). Therefore,
628 although the mechanism by which the *ACAA2* gene could be linked to increased milk

629 yield and decreased protein percentage needs to be elucidated, increased *ACAA2*
630 expression at the mRNA level is likely to result in higher levels of the enzyme and
631 thus elevated amount of energy and carbon substrates for mammary development and
632 lactation.

633

634

CONCLUSION

635 In the current study, we demonstrated that the variants in the 3'UTR of the *ACAA2*
636 gene, which are associated with milk yield, protein percentage and fat yield, are
637 differentially expressed in homozygous ewes of each allele and exhibit AEI within
638 heterozygotes in a tissue specific manner, suggesting the existence of a cis-acting
639 regulatory DNA mechanism. These findings support the hypothesis that the *ACAA2*
640 gene is a functional candidate affecting dairy traits.

641

642

ACKNOWLEDGMENTS

643 This work was funded by the Cyprus Research Promotion Foundation, the
644 European Structural Funds and the Cyprus University of Technology. G. Banos and
645 A. Psifidi were supported by the BBSRC Institute Strategic Programme Grant
646 (BB/J004235/1; UK) and G. Banos also by the Rural & Environment Science &
647 Analytical Services Division of the Scottish Government. We thank the professional
648 veterinarian John Pierroua for obtaining biopsies. We also thank the head of the sheep
649 and goat farming sector of the Ministry of Agriculture, Mr. Paniko Christoforou, the
650 head (Mr. Giorgo Aspromalli) and technical staff of the governmental farm (Orites)
651 and the owners of the commercial farms for giving us access to their farms and help in
652 sampling and collection of data.

653

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

654 **Table 1.** Predicted genotype means and standard errors of the genotypic classes CC,
 655 CT and TT at the g.2982T>C ACAA2 locus and significance of genotype contrasts for
 656 each trait.

<i>Trait</i>	all lactations milk yield ¹	1 st lactation milk yield ¹	milk protein percentage	milk fat percentage	milk protein yield ¹	milk fat yield ¹
Genotyped means ± SE ²						
CC	245.23 ± 4.44 ^a	170.21 ± 5.28 ^a	5.29 ± 0.02 ^a	5.19 ± 0.05	8.78 ± 0.25	12.87 ± 0.26 ^a
CT	260.71 ± 2.76 ^c	193.85 ± 3.33 ^d	5.23 ± 0.01 ^c	5.20 ± 0.03	9.42 ± 0.16	13.62 ± 0.17 ^d
TT	258.85 ± 3.88 ^b	191.44 ± 4.77 ^c	5.20 ± 0.02 ^c	5.14 ± 0.04	9.23 ± 0.25	13.44 ± 0.23 ^b

657 ¹Yields are in kg.

658 ²Marginal genotype means (± standard error) predicted from the mixed model analyses

659 ^{a-e}Means within a column with two different superscripts differ as: ^{a,b}P < 0.025; ^{a,c}P < 0.010; ^{a,d}P <
 660 0.001 after Holm-Bonferroni adjustment.
 661

662 **Table 2.** SNP allelic effects and percentage of phenotypic variance explained by the
 663 g.2982T>C SNP of the ACAA2 gene.

<i>Trait</i>	all lactations milk yield	1 st lactation milk yield	milk protein percentage	milk fat percentage	milk protein yield	milk fat yield
a ¹ ± SE	6.81 ± 2.95	10.61 ± 3.56	-0.05 ± 0.02	-0.03 ± 0.03	0.23 ± 0.18	0.28 ± 0.17
p ²	0.021*	0.003*	0.003*	0.370	0.200	0.100
d ³ ± SE	8.67 ± 3.53	13.02 ± 4.26	0.02 ± 0.02	0.04 ± 0.04	0.41 ± 0.21	0.46 ± 0.21
p ⁴	0.014*	0.002*	0.41	0.322	0.048*	0.027*
% V _p due to SNP ⁵	0.62%	2.25%	1.68%	0.33%	0.44%	0.41%

664 ¹a = additive effect; positive additive genetic effect (a > 0) indicates T allele increased the trait.

665 ²P value for assessing the additive effect on the trait; *significant post Holm-Bonferroni adjustment for
 666 three independent tests.

667 ³d = dominance effect.

668 ⁴P value for assessing the dominance effect on the trait *significant post Holm-Bonferroni adjustment
 669 for three independent tests.

670 ⁵Estimated using allele frequencies observed in sample (p = 0.46 for allele C and q = 0.54 for T).
 671

672

673

674

REFERENCES

- 675 Andersson, L. 2001. Genetic dissection of phenotypic diversity in farm animals.
676 Nature Reviews Genetics. 2:130-138.
- 677 AOAC International. 2005. official methods of analysis. 18th ed. 1st rev. AOAC
678 international gaithersburg, MD.
- 679 Arranz, J. J. and B. Gutiérrez-Gil. 2012. Detection of QTL Underlying Milk Traits in
680 Sheep: An Update. Milk Production - Advanced Genetic Traits, Cellular
681 Mechanism, Animal Management and Health, Prof. Narongsak Chaiyabutr (Ed.),
682 InTech, DOI: 10.5772/50762. Available
683 from: [Http://www.intechopen.com/books/milk-Production-Advanced-Genetic-
684 Traits-Cellular-Mechanism-Animal-Management-and-health/detection-of-Qtl-
685 Underlying-Milk-Traits-in-Sheep-an-Update](http://www.intechopen.com/books/milk-Production-Advanced-Genetic-Traits-Cellular-Mechanism-Animal-Management-and-health/detection-of-Qtl-Underlying-Milk-Traits-in-Sheep-an-Update). INTECH Open Access Publisher
- 686 Bartlett, K.; Eaton, S. Mitochondrial β -oxidation. Eur. J. Biochem. 2004, 271, 462-
687 469.
- 688 Bencini, R. and G. Pulina. 1997. The quality of sheep milk: A review. Animal
689 Production Science. 37:485-504.
- 690 Bray, N. J., P. R. Buckland, M. J. Owen and M. C. O'Donovan. 2003. Cis-acting
691 variation in the expression of a high proportion of genes in human brain. Hum.
692 Genet. 113:149-153.
- 693 Chatziplis D. G.; Tzamaloukas, O.; Miltiadou, D.; Ligda, C.; Koumas, A.;
694 Mavrogenis, A.P.; Georgoudis, A.; Papachristoforou, C. 2012. Evidence of major
695 gene(s) affecting milk traits in the Chios sheep breed. Small Rum. Res. 105, 61–
696 68.
- 697 Campbell, C. D., A. Kirby, J. Nemesh, M. J. Daly and J. N. Hirschhorn. 2008. A
698 survey of allelic imbalance in F1 mice. Genome Res. 18:555-563.
- 699 Carta, A., S. Casu and S. Salaris. 2009. Invited review: Current state of genetic
700 improvement in dairy sheep. J. Dairy Sci. 92:5814-5833.
- 701 Chamberlain, A. J., C. J. Vander Jagt, B. J. Hayes, M. Khansefid, L. C. Marett, C. A.
702 Millen, T. T. Nguyen and M. E. Goddard. 2015. Extensive variation between
703 tissues in allele specific expression in an outbred mammal. BMC Genomics.
704 16:1.
- 705 Chen, X., J. Weaver, B. A. Bove, L. A. Vanderveer, S. C. Weil, A. Miron, M. B. Daly
706 and A. K. Godwin. 2008. Allelic imbalance in BRCA1 and BRCA2 gene
707 expression is associated with an increased breast cancer risk. Hum. Mol. Genet.
708 17:1336-1348.
- 709 Clop, A., F. Marcq, H. Takeda, D. Pirottin, X. Tordoir, B. Bibé, J. Bouix, F. Caiment,
710 J. Elsen and F. Eychenne. 2006. A mutation creating a potential illegitimate
711 microRNA target site in the myostatin gene affects muscularity in sheep. Nat.
712 Genet. 38:813-818.

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

- 713 Cochran, S. D., J. B. Cole, D. J. Null and P. J. Hansen. 2013. Discovery of single
714 nucleotide polymorphisms in candidate genes associated with fertility and
715 production traits in Holstein cattle. *BMC Genetics*. 14:1.
- 716 Cowles, C. R., J. N. Hirschhorn, D. Altshuler and E. S. Lander. 2002. Detection of
717 regulatory variation in mouse genes. *Nat. Genet.* 32:432-437.
- 718 De Rancourt, M., N. Fois, M. Lavín, E. Tchakérian and F. Vallerand. 2006.
719 Mediterranean sheep and goats production: An uncertain future. *Small Ruminant*
720 *Research*. 62:167-179.
- 721 Drackley, J. K. 1999. ADSA foundation scholar award: Biology of dairy cows during
722 the transition period: The final frontier? *J. Dairy Sci.* 82:2259-2273.
- 723 Drackley, J. 2000. Lipid metabolism. *Farm Animal Metabolism and Nutrition*. 97-
724 119.
- 725 Emery, R.S. 1988. Milk fat depression and the influence of diet on milk composition.
726 *The Veterinary clinics of North America. Food animal practice* 4 (2), pp. 289-
727 305
- 728 Forton, J. T., I. A. Udalova, S. Campino, K. A. Rockett, J. Hull and D. P.
729 Kwiatkowski. 2007. Localization of a long-range cis-regulatory element of IL13
730 by allelic transcript ratio mapping. *Genome Res.* 17:82-87.
- 731 Fuertes, J.; Gonzalo, C.; Carriedo, J.; San Primitivo, F. 1998. Parameters of test day
732 milk yield and milk components for dairy ewes. *J. Dairy Sci.* 81, 1300-1307.
- 733 Garcia-Gámez, E.; Gutiérrez-Gil, B.; Sahana, A. G.; Sanchez, J.P.; Bayon, Y.; Arranz,
734 J. J. 2012. GWA analysis for milk production traits in dairy sheep and genetic
735 support for a QTN influencing milk protein percentage in the LALBA gene. *PloS*
736 *One*, 7(10), e47782.
- 737 Garcia-Gámez, E., B. Gutiérrez-Gil, A. Suarez-Vega, L. F. de la Fuente and J. J.
738 Arranz. 2013. Identification of quantitative trait loci underlying milk traits in
739 spanish dairy sheep using linkage plus combined linkage disequilibrium and
740 linkage analysis approaches. *J. Dairy Sci.* 96:6059-6069.
- 741 Gilmour, A. R., B. Gogel, B. Cullis, R. Thompson and D. Butler. 2009. ASReml user
742 guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.
- 743 Gutiérrez-Gil, B., J. J. Arranz, R. Pong-Wong, E. García-Gámez, J. Kijas and P.
744 Wiener. 2014. Application of selection mapping to identify genomic regions
745 associated with dairy production in sheep. *Plos One*. 9:.
- 746 Gutiérrez-Gil, B., M. F. El-Zarei, L. Alvarez, Y. Bayón, L. F. De La Fuente, F. San
747 Primitivo and J. -. Arranz. 2009. Quantitative trait loci underlying milk
748 production traits in sheep. *Anim. Genet.* 40:423-434.
- 749 Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian*
750 *Journal of Statistics* 6:65-70.

751

752 ICAR. 2014. ICAR (international committee for animal recording), 2014. section 2
753 ICAR rules, standards and guidelines for dairy production recording. approved
754 by the general assembly held in berlin, germany, on may 2014.

755 Karim, L., H. Takeda, L. Lin, T. Druet, J. A. Arias, D. Baurain, N. Cambisano, S. R.
756 Davis, F. Farnir and B. Grisart. 2011. Variants modulating the expression of a
757 chromosome domain encompassing PLAG1 influence bovine stature. *Nat. Genet.*
758 43:405-413.

759 Khatib, H., V. Schutzkus, Y. Chang and G. Rosa. 2007. Pattern of expression of the
760 uterine milk protein gene and its association with productive life in dairy cattle. *J.*
761 *Dairy Sci.* 90:2427-2433.

762 Khatib, H. 2007. Is it genomic imprinting or preferential expression? *Bioessays.*
763 29:1022-1028.

764 Li HD (2008) I: Segregation, Cloning, SNP Detection and Its Association with
765 Economic Traits of 5 Bovine Genes II: Quantitative Trait Loci Analysis of Meat
766 Quality Traits in Pigs. Ph.D. thesis. Dissertation ID: 1647914.
767 Available:<http://www.dissertationtopic.net/doc/1647914>.

768 Livak, K. J. 1999. Allelic discrimination using fluorogenic probes and the 5' nuclease
769 assay. *Genet. Anal. : Biomol. Eng.* 14:143-149.

770 Livak, K. J. and T. D. Schmittgen. 2001. Analysis of relative gene expression data
771 using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods.* 25:402-
772 408.

773 Mavrogenis, A. and A. Constantinou. 1991. Selection Index and Expected Genetic
774 Progress in Chios Sheep. Agricultural Research Institute, Ministry of Agriculture
775 and Natural Resources.

776 Miltiadou, D., Orford M., Symeou S. and G. Banos. 2017. Identification of variation
777 in the ovine prolactin gene with a cost effective sequence based typing (SBT)
778 assay. *J. Dairy Sci.* 100 (2), pp. 1290-1294.

779 Muráni, E., S. Ponsuksili, T. Srikanchai, S. Maak and K. Wimmers. 2009. Expression
780 of the porcine adrenergic receptor beta 2 gene in longissimus dorsi muscle is
781 affected by cis-regulatory DNA variation. *Anim. Genet.* 40:80-89.

782 Olbromski, R., E. Siadkowska, B. Zelazowska and L. Zwierzchowski. 2013. Allelic
783 gene expression imbalance of bovine IGF2, LEP and CCL2 genes in liver, kidney
784 and pituitary. *Mol. Biol. Rep.* 40:1189-1200.

785 Orford, M., O. Tzamaloukas, C. Papachristoforou and D. Miltiadou. 2010. Technical
786 note: A simplified PCR-based assay for the characterization of two prolactin
787 variants that affect milk traits in sheep breeds. *J. Dairy Sci.* 93:5996-5999.

788 Orford, M., G. Hadjipavlou, O. Tzamaloukas, D. Chatziplis, A. Koumas, A.
789 Mavrogenis, C. Papachristoforou and D. Miltiadou. 2012. A single nucleotide

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

- 790 polymorphism in the acetyl-coenzyme A acyltransferase 2 (ACAA2) gene is
791 associated with milk yield in chios sheep. *J. Dairy Sci.* 95:3419-3427.
- 792 Othmane M.H., Carriedo J.A., De la Fuente L.F., San Primitivo F. (2002): Factors
793 affecting test-day milk composition in dairy ewes, and relationships amongst
794 various milk components. *J. Dairy Res.*, 69, 53–62.
- 795 Paten, A.M., J.E. Duncan, S.J. Pain, S.W. Peterson, P. R. Kenyon, H.T. Blair and P.K.
796 Dearden. 2015. Functional development of the adult ovine mammary gland—
797 insights from gene expression profiling. *BMC Genom.* 16:748
- 798 Pesole, G., F. Mignone, C. Gissi, G. Grillo, F. Licciulli and S. Liuni. 2001. Structural
799 and functional features of eukaryotic mRNA untranslated regions. *Gene.* 276:73-
800 81.
- 801 Pollott G.E. 2004. Deconstructing milk yield and composition during lactation using
802 biologically based lactation models. *J Dairy Sci.* Aug;87(8):2375-87.
- 803 Pulina G., Macciotta N., Nudda A. (2005): Reproductive performance and milk
804 production of Assaf sheep in an intensive management system. *Ital. J. Anim. Sci.*,
805 4, 5–14.
- 806 Ramón, M., A. Legarra, E. Ugarte, J. J. Garde and M. D. Pérez-Guzmán. 2010.
807 Economic weights for major milk constituents of manchega dairy ewes. *J. Dairy*
808 *Sci.* 93:3303-3309.
- 809 Ron, M. and J. Weller. 2007. From QTL to QTN identification in livestock—winning
810 by points rather than knock-out: A review. *Anim. Genet.* 38:429-439.
- 811 Sari, I. P., A. Rao, J. T. Smith, A. J. Tilbrook and I. J. Clarke. 2009. Effect of RF-
812 amide-related peptide-3 on luteinizing hormone and follicle-stimulating hormone
813 synthesis and secretion in ovine pituitary gonadotropes. *Endocrinology.*
814 150:5549-5556.
- 815 SAS. 2005. Statistical Analysis Systems User's Guide. Version 9.1.3. SAS Institute,
816 Inc., Cary, NC.
- 817 Sasaki, S., T. Ibi, T. Watanabe, T. Matsushashi, S. Ikeda and Y. Sugimoto. 2013.
818 Variants in the 3' UTR of general transcription factor IIF, polypeptide 2 affect
819 female calving efficiency in japanese black cattle. *BMC Genet.* 14:.
- 820 Shi, H., J. Zhu, J. Luo, W. Cao, H. Shi, D. Yao, J. Li, Y. Sun, H. Xu, K. Yu and J. J.
821 Looor. 2015. Genes regulating lipid and protein metabolism are highly expressed
822 in mammary gland of lactating dairy goats. *Funct. Integr. Genomics.* 15:309-321.
- 823 Sodhi, S. S., M. Ghosh, K. D. Song, N. Sharma, J. H. Kim, N. E. Kim, S. J. Lee, C.
824 W. Kang, S. J. Oh and D. K. Jeong. 2014. An approach to identify SNPs in the
825 gene encoding acetyl-CoA acetyltransferase-2 (ACAT-2) and their proposed role
826 in metabolic processes in pig. *PloS One.* 9:e102432.
- 827 Stinckens, A., M. Schroyen, L. Peeters, S. Janssens and N. Buys. 2010. Quantitative
828 trait mutations in cattle, sheep and pigs: A review. *Animal Science Reviews.* 59.

CIS-REGULATORY DNA VARIATION IN THE OVINE ACAA2 GENE

- 829 Sugimoto, M., Y. Gotoh, T. Kawahara and Y. Sugimoto. 2015. Molecular effects of
830 polymorphism in the 3'UTR of unc-5 homolog C associated with conception rate
831 in holsteins. *PloS One*. 10:e0131283.
- 832 Vernon, R. G. 2005. Lipid metabolism during lactation: A review of adipose tissue-
833 liver interactions and the development of fatty liver. *J. Dairy Res.* 72:460-469.
- 834 Weller, M. M. D. C. A., R. L. Albino, M. I. Marcondes, W. Silva, K. M. Daniels, M.
835 M. Campos, M. S. Duarte, M. L. Mescouto, F. F. Silva and S. E. F. Guimarães.
836 2016. Effects of nutrient intake level on mammary parenchyma growth and gene
837 expression in crossbred (holstein × gyr) prepubertal heifers. *J. Dairy Sci.*
838 99:9962-9973.
- 839 Xie, X., J. Lu, E. Kulbokas, T. R. Golub, V. Mootha, K. Lindblad-Toh, E. S. Lander
840 and M. Kellis. 2005. Systematic discovery of regulatory motifs in human
841 promoters and 3' UTRs by comparison of several mammals. *Nature*. 434:338-
842 345.
- 843 Yan, H., W. Yuan, V. E. Velculescu, B. Vogelstein and K. W. Kinzler. 2002. Allelic
844 variation in human gene expression. *Science*. 297:1143.
- 845 Yao, D. W., J. Luo, Q. Y. He, J. Li, H. Wang, H. B. Shi, H. F. Xu, M. Wang and J. J.
846 Loo. 2016a. Characterization of the liver X receptor-dependent regulatory
847 mechanism of goat stearoyl-coenzyme A desaturase 1 gene by linoleic acid. *J.*
848 *Dairy Sci.* 99:3945-3957.
- 849 Yao, D. W., J. Luo, Q. Y. He, H. F. Xu, J. Li, H. B. Shi, H. Wang, Z. Chen and J. J.
850 Loo. 2016b. Liver X receptor a promotes the synthesis of monounsaturated fatty
851 acids in goat mammary epithelial cells via the control of stearoyl-coenzyme A
852 desaturase 1 in an SREBP-1-dependent manner. *J. Dairy Sci.* 99:6391-6402.
- 853

854

FIGURE CAPTIONS

855 **Figure 1. Allelic expression imbalance TaqMan Assay for the quantification of**
 856 **the ACAA2 allele T:C transcription ratio. A.** Amplification plots of TaqMan probe
 857 PCR from ACAA2 TT (top), TC (middle) and CC (bottom) individuals. **B.**
 858 Amplification plot of serial dilution of standard gDNA pooled from three ACAA2 TC
 859 ewes. **C.** Log₁₀Quantity and threshold cycle (Ct) relationship graph for T and C
 860 alleles. Red dots represent standard gDNA dilution series reactions. Linear standard
 861 curves for T and C alleles (targets), standard curve slope, Y-intersection, r^2 and
 862 efficiency values are shown.

863 **Figure 2. Expression of the ACAA2 gene in liver and udder. A.** ACAA2 expression
 864 in liver and udder samples (taken 42 ± 2 days after lambing) of CC, CT and TT first
 865 parity, 15 month old ewes that gave birth to two lambs each relative to the mean
 866 normalized expression in CC ewes. Expression was normalized by β -actin and β -
 867 microglobulin (B2M) expression in liver and udder, respectively. n=3 for all means
 868 except udder CT (n=2). **B.** Mean normalized expression of the ACAA2 gene in *ex vivo*
 869 liver and udder samples from CC, CT and TT ewes, normalized by corresponding *b*-
 870 actin transcription and relative to mean normalized transcription in the liver of TT
 871 individuals. CC: n=1, CT: n=10 (udder) n=9 (liver), TT: n=4 (udder) n=5 (liver). Bars
 872 represent mean relative normalized transcription and standard error of the mean is
 873 shown. Means with different letters are significantly different ($P < 0.05$).

874 **Figure 3.** Sequencing of heterozygous samples from gDNA and cDNA from liver (A)
 875 and udder (B). The T/C SNP is shown in light blue shade. The peak heights of the T
 876 and C alleles are the same in gDNA, whereas the peak height of the T allele is higher
 877 compared to the C allele in cDNA of the same samples.

878 **Figure 4.** Allelic expression imbalance of the *ACAA2* gene. Graph shows TaqMan
879 assay corrected T:C ratios in the liver and udder of heterozygous ewes. Horizontal
880 lines represent mean corrected T:C allele in the cDNA from liver (mean=1.03, n=12)
881 and udder (mean=1.18, n=11) and in gDNA (mean=1, n=11) of *ACAA2* heterozygous
882 ewes. Open circles represent corrected T:C ratios for individual ewes. * $P < 0.05$, ** P
883 < 0.01 .

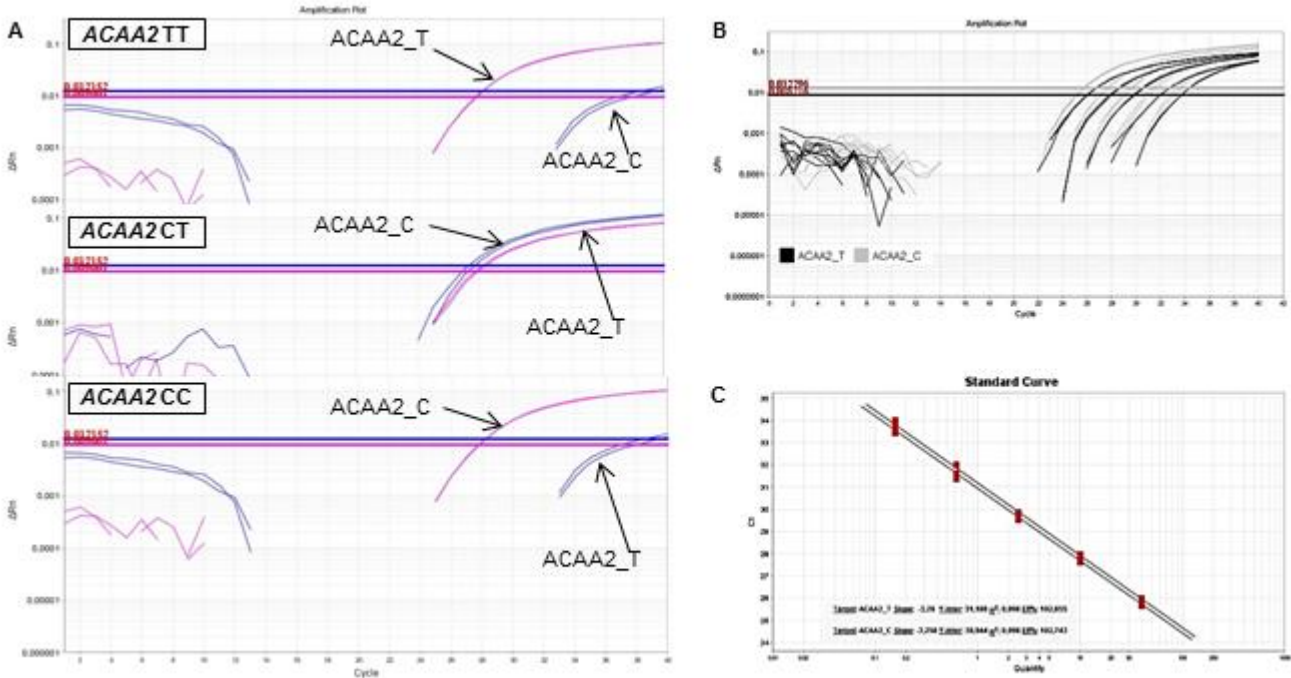
884

885

886

887

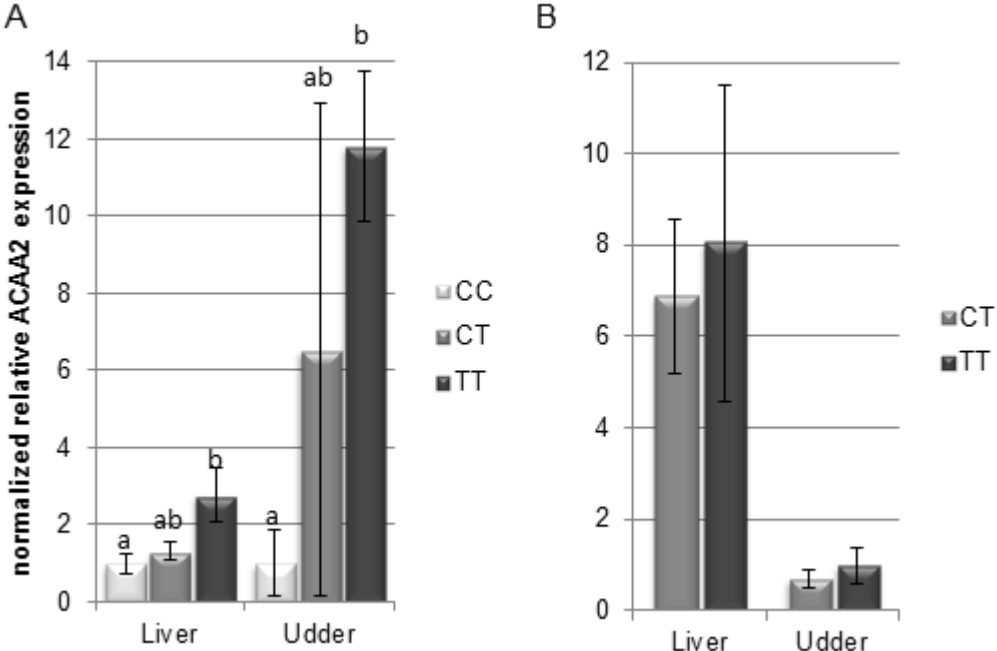
Miltiadou et al. **Figure 1**



888

889

Miltiadou et al. **Figure 2**

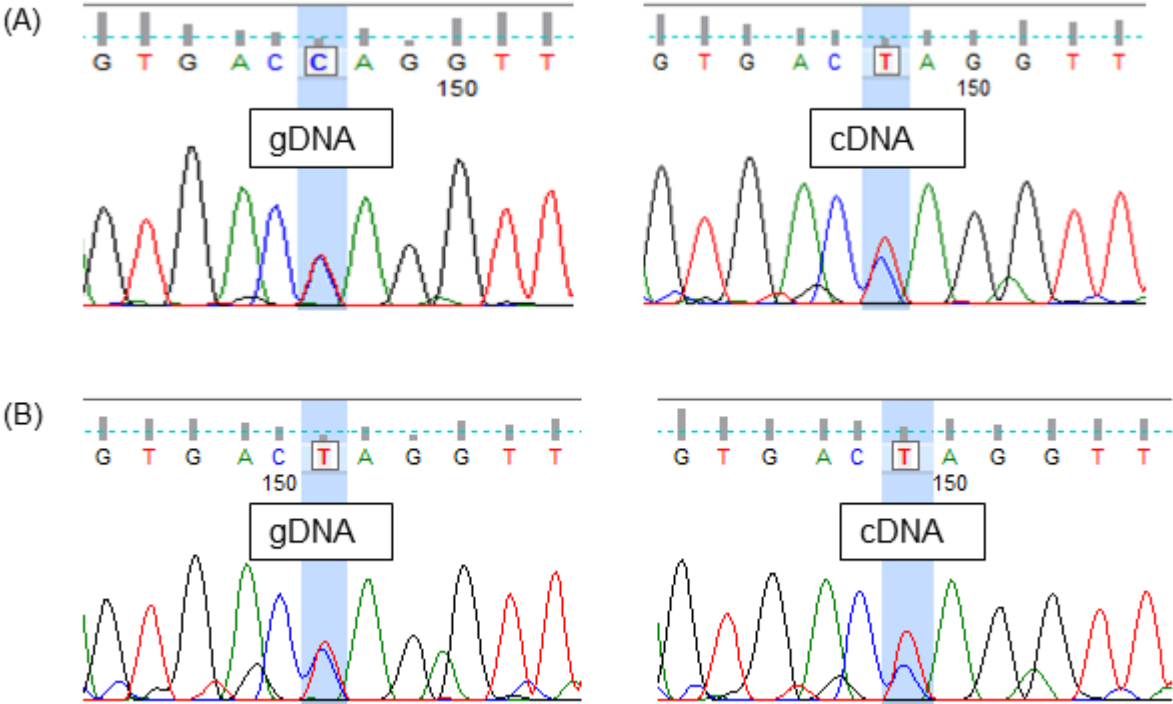


890

891

892

Miltiadou et al. **Figure 3**

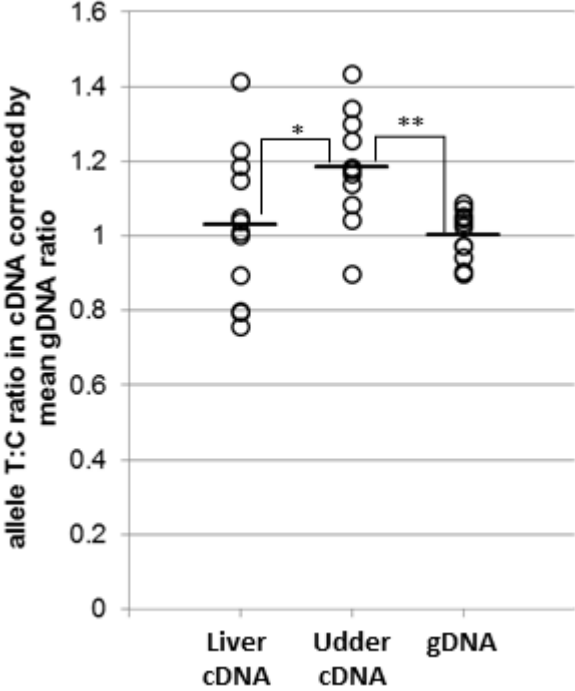


893

894

895

Miltiadou et al. **Figure 4**

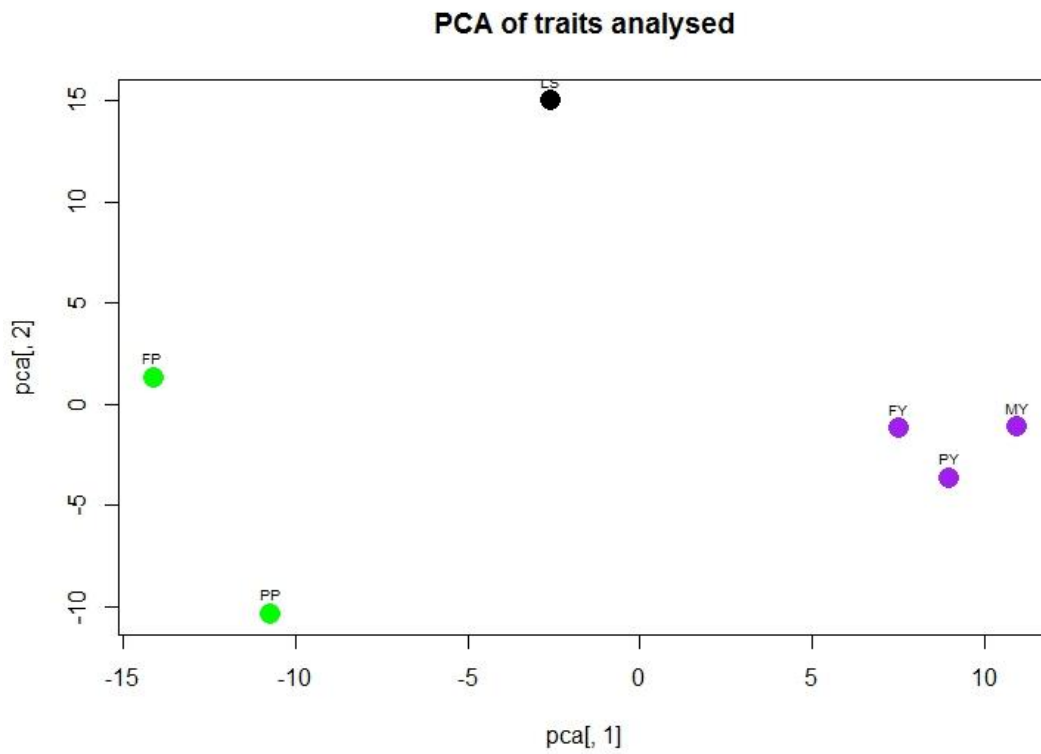


896

897

898 **Appendix**

899 **Supplementary Figure 1:** Principal Component Analysis of studied traits: milk yield
 900 (MY), protein yield (PY), fat yield (FY), protein percentage (PP), fat percentage (FP)
 901 and litter size (LS).



902

903 Supplementary Table 1

<i>P</i> values after Holm-Bonferroni adjustment		
Lowest <i>P</i> value	Second lowest <i>P</i> value	Third and later lowest <i>P</i> value
<i>P</i> = 0.0167	<i>P</i> = 0.025	<i>P</i> = 0.05

904