ANIMAL PRODUCTION SCIENCE

A genome-wide scan of fat-tail sheep identifies signals of selection for fat deposition and adaptation

Journal:	*Animal Production Science*
Manuscript ID	AN17753.R1
Manuscript Type:	Research paper
Date Submitted by the Author:	02-Feb-2018
Complete List of Authors:	Mastrangelo, Salvatore; University Of Palermo, Demetra Moioli, Bianca; Consiglio per laRicerca e la sperimentazione in Agricoltura , Ahbara, Abdulgasim Latairish, Suliman Portolano, Baldassare; University Of Palermo, DEMETRA Pilla, Fabio Ciani, Elena
Keyword:	Adaptation, Genomics, Sheep

A genome-wide scan of fat-tail sheep identifies signals of selection for fat deposition and adaptation

S. Mastrangelo[1], B. Moioli[2], A. Ahbara[3,4], S. Latairish[4], B. Portolano[1], F. Pilla[5], E. Ciani[6]

[1] Dipartimento Scienze Agrarie Alimentari e Forestali, Università degli Studi di Palermo, 90128 Palermo, Italy

[2] Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (CREA), via Salaria 31, 00015 Monterotondo, Italy

[3] School of Life Science, University of Nottingham, NG7 2RD United Kingdom

[4] University of Misurata, Faculty of Sciences, Department of Zoology, Misurata, Libya

[5] Dipartimento Agricoltura, Ambiente e Alimenti, Università degli Studi del Molise, Campobasso, Italy

[6] Dipartimento di Bioscienze, Biotecnologie e Biofarmaceutica, Università degli Studi di Bari "Aldo Moro", Bari, Italy

Corresponding author: Bianca Moioli. E-mail: bianca.moioli@crea.gov.it

Short title: Genome-wide scan for fat-tail signatures in sheep

Abstract

Fat tail in sheep represent a valuable energy reserve for facing future climate changes. The identification of genes with a role in the fat-tail phenotype may contribute to understanding the physiology of fat deposition and the mechanisms of adaptation. Genotypic data obtained with the OvineSNP50K array in 13 thin-tail sheep breeds from Italy were used to identify selection signatures of fat tail through pairwise thin vs. fat-tail sheep breed comparisons, with the following fat-tail breeds of the Mediterranean area: two unique Italian fat-tail breeds (Barbaresca and

Laticauda), a Barbary sheep breed from Libya, Ossimi breed from Egypt, Cyprus Fat-Tail and Chios from the Greek islands Cyprus and Chios, respectively. F_{st} and χ^2 values obtained for over 40 thousand polymorphic markers allowed confirmation of twelve fat-tail associations which were previously reported in Chinese and Iranian breeds. Two of these signals – on OAR 7 and OAR 13 - are in proximity of two genes - *VRTN* and *BMP2* - with a role in the variation of vertebral number and in fat-tail formation, respectively. Two identified signals on OAR 6 and OAR 15 encompass two genes, *PDGFRA* and *PDGFD*, involved in the differentiation of preadipocytes. Further signals detected herein were reported in Chinese sheep as signatures of adaptation to desert areas. For a number of the detected associations the known role in either fat deposition or adaptation, thus contributing to unveiling the molecular basis underlying mechanisms of energy storage and climate adaptation.

Key words: fat-tail, adaptation, genomics, sheep.

Introduction

The fat tail characteristic of sheep has a role in the survival mechanism in harsh environments (Xu *et al.* 2017); it represents a valuable energy reserve during periods of food shortage and for facing future climate changes, but can also contribute to the identification of the genes with a role in lipid metabolism (Bakhtiarizadeh *et al*. 2013). Fat-tail sheep represent about 25% of the world's sheep population, and the genes with a role in this phenotype are likely not the same for every breed, since the fat tail was selected by humans in longstanding husbandry practices in different geographical regions (Moradi *et al*. 2012).

With the aim to identify putative candidate genes for the fat tail phenotype, studies were performed on various sheep populations, using anonymous markers distributed throughout the genome, and comparing allele frequencies of fat-tail sheep with various thin-tail breeds (Moradi *et al*. 2012; Moioli *et al*. 2015; Wei *et al*. 2015; Yuan *et al*. 2016). Such studies presented contrasting results, likely because of the used statistical methodology, and because of the complexity of the fat-tail

phenotype, well described in Chinese breeds by Wei *et al*. (2015), in Ethiopian breeds by Gizaw (2008) and in Sudanese breeds by Tibin (2007). These authors agree in categorizing the fat-tail types in their countries as short fat-tail, long fat-tail and fat-rumped, and such complexity suggests that many different genes have a role in the tail phenotype, considering that the three tail types were also associated to different productive purposes (Wei *et al*. 2015).

The wild ancestor of sheep was thin-tail, and the fat-tail phenotype was developed between 3000 and 1500 BCE in the Fertile Crescent (Moradi *et al*. 2010); at present these sheep are mainly distributed in the Middle East, North Africa and Central Asia because fat tails represents the energy reserve necessary to face drought seasons and food shortage (Xu *et al.* 2017). Thin-tail breeds on the contrary are predominant in areas far from the Fertile Crescent (Moradi *et al*. 2010) where fat tails are not required as an energy reserve (Nejati-Javaremi *et al.* 2007).

In Italy, the Barbaresca and the Laticauda are the only two fat-tail breeds; they have been introduced from North-Africa in different circumstances. The Barbaresca is an ancient Sicilian fat-tail sheep (Bigi and Zanon, 2008) with a long and pendulous tail and is a dual-purpose breed which originated from crosses between Tunisian Barbary sheep from North Africa and the Pinzirita breed during the Arab settling in Sicily (9th century), and is at present highly endangered (Mastrangelo *et al.* 2017). The Laticauda is an autochthonous breed of Southern Italy (Campania) derived from crossbreeding of local sheep from the Apennines with fat-tail North African sheep, likely imported under the Bourbons dynasty in the XVIII century. Reared under semi-extensive systems this breed shows high prolificacy and is a good meat producer (Ciani *et al*. 2013). Signatures of fat-tail in the Laticauda and Cyprus Fat-Tail were identified in a previous work (Moioli et *al.* 2015) where the contrasting groups were composed of the two fat-tail breeds on one side, and 13 thin-tail breeds on the other side. In this work few regions under selection were identified, likely because the number of analyzed animals were not sufficient to detect selective sweeps, which are much easier to detect if the analysis is performed on a large number of unrelated animals. Moreover, because the potential anonymous marker in linkage disequilibrium (LD) with the causal mutation is not the same in all

81 breeds (Qanbari *et al.* 2010), contrasting groups composed of more than one breed, as in Moioli *et*
82 *al*. (2015) might have hindered the detection of putative selection sweeps.
83 The Barbaresca and the Laticauda were named by Mason (1967) Sicilian Barbary and Campanian
84 Barbary respectively, as to indicate their origin; however, despite of this common origin, in the
85 multi dimensional scaling plot of the first two components of the matrix of the pairwise identity by
86 state distances among Italian sheep (Mastrangelo *et al*. 2018) the Barbaresca represented a clearly
87 distinct cluster, while the Laticauda was within the same big cluster of the thin-tail breeds, but
88 closer to the breeds of its geographical area. Because crossbreeding of the Laticauda and Barbaresca
89 with North African Barbary sheep occurred at different times and the two breeds might have
90 differentiated from the use of different Barbary sheep stocks, the Libyan-Barbary was also included
91 in this study. Indeed, this breed belongs to the fat-tail, coarse-wool Barbary sheep, characterized by
92 multi-colored, large framed structure, with pendulous fat-tail (Akraim *et al*. 2008).
93 Prerequisite for dissecting the selection signals related to the fat tail from the ones associated to
94 other Barbary related traits was to conduct a similar analysis of other fat-tail, but non-Barbary
95 sheep. The Ossimi and the Cyprus Fat-Tail breeds were therefore included in the study. The Ossimi
96 is the most popular sheep breed in the Nile Valley and Delta. The Ossimi is reared for the
97 production of lambs and is expanding (over 1,000,000 sheep) at the expense of other less producing
98 breeds (Elshennawy, 1995). Origin of this breed is the Ossim village, near Cairo. The Cyprus Fat-
99 Tail breed is at present endangered, because of the low milk production (60-80 litre milk per
100 lactation) despite of the very high fat content (6.5-7%). Around 1000 purebred animals are left in
101 the island of Cyprus, where the breed originated (http://dad.fao.org). Genotypic data of the
102 Laticauda and Cyprus Fat-Tail breeds, previously analyzed (Moioli et al. 2015), were re-analyzed
103 here using the one-breed-to-one breed comparison.
104 To confirm the identification of genomic regions contributing to shaping the fat-tail phenotype, the
105 Chios sheep was also included in the study. This is a semi-fat-tailed sheep, with high milk
106 production and prolificacy (Theodoritis *et al.* 2012). It is the most productive among indigenous

Greek breeds, and is suitable for intensive farming, while all the other fat-tail breeds of the present study are low producing sheep and consequently, in most cases, they are highly endangered.

The aim of this study was then to identify loci influencing fat deposition in sheep; therefore genotypic data obtained with the OvineSNP50 BeadChip (Illumina, Inc.) were used in a genome-wide comparison of the six fat-tail breeds with 13 Italian thin-tail breeds. To strengthen the identification of common associations in the genome between the fat and thin-tail sheep, this study examined, through the relevant literature, selective signals shared by other fat-tail breeds worldwide.

Materials and methods

Animals and genotyping

Genotypic data of the Laticauda and the Italian thin-tail breeds were provided by a previous genome-wide analysis of genetic diversity performed in 21 Italian sheep breeds (Ciani *et al*. 2014). Genotypic data of the Barbaresca were available from a recent study on this breed (Mastrangelo *et al.* 2017). Genotypic data of the Chios and the Cyprus Fat-Tail were available from the HapMap project (Kijas *et al.* 2012). Libyan-Barbary and Ossimi samples were genotyped for this study using the Illumina OvineSNP50K BeadChip.

Ciani *et al*. (2014) described the pattern of genetic diversity in Italian breeds and provided the basis for excluding, from the comparison with the fat-tail breeds, the breeds most similar to the Laticauda, so to prevent that sheep carrying alleles involved in the fat tail phenotype be unintentionally included in the thin-tail group. The thin-tail breeds used in the comparison were then the following: Alpagota, Altamurana, Appenninica, Bergamasca, Biellese, Delle Langhe, Gentile di Puglia, Fabrianese, Istriana, Massese, Sambucana, Sarda, Sopravissana. Each sampled breed consisted of 24 animals, except the Barbaresca and the Cyprus Fat-Tail (30 animals each breed), the Chios (23 animals) and the Ossimi (9 animals). Animals were sampled from different flocks to avoid, when possible, sampling of related individuals.

Chromosomal coordinates for each SNP were obtained from the latest release of the ovine genome sequence assembly Oar_v4.0. Markers were filtered to exclude loci assigned to unmapped contigs. Only SNPs located on autosomes were considered for further analyses

The following filtering parameters were adopted to exclude certain loci and animals and to generate input files: (i) SNPs with call rate ≤99%, (ii) SNPs with minor allele frequency (MAF) ≤1%, (iii) animals displaying ≥10% of genotypes missing. File editing was carried out using PLINK (Purcell *et al*. 2007).

Genomic scan for selective sweeps

The F_{st} statistics identifies genetic differentiation between populations to multiple loci and compares the estimates with the expected under neutrality (Bonhomme *et al.* 2010). To detect genomic regions that may have been under positive selection, F_{st} values of differentiation for each marker were calculated in pairwise comparisons of each of the six fat-tail breeds with each of the 13 Italian thin tail breeds. Pairwise comparisons - one-breed vs. one-breed - were chosen in order to prevent that difference in LD between the causal mutation and the anonymous marker, in the various breeds, hindering the detection of putative selection sweep, because the marker associated to the gene might not be the same in all breeds.

Only the SNPs for which the pairwise locus-specific F_{st} had a rank percentile value of 0.01 or less were considered, so to strengthen the identification of fat-tail signals. Furthermore, in order to confirm by a statistical test which markers differed pairwise, since F_{st} and χ^2 values are highly correlated (Moioli *et al*., 2015), the χ^2 test of differences of allele frequency was performed for each marker, and only SNPs satisfying the threshold level of Bonferroni-adjusted χ^2 P-values ≤ 0.05 were then taken into consideration.

The following constraints were introduced for defining fat-tail sweeps: 1) they should include at least two consecutive significant markers at ≤ 500 Kb from each other; this value was deemed from the genome-wide association studies (GWAS) which use this distance to determine the potential candidate genes associated to the markers (Wu *et al.* 2013; Zare *et al.* 2014); 2) 10 or more pairwise

comparisons (i.e. in more than ¾ of the breeds) should indicate at least one marker within the region; 3) putative candidate genes in LD with the significant signals, performed on the Oar:v4.0 genome assembly, should fall either within or ± 200 Kb upstream or downstream from the region.

Results

The data set passing quality controls included 43,072 markers common to all the breeds. No animal was excluded by filtering for genotypes missing.

Under positive selection, the selected locus region shows decreased diversity levels within the population and increased levels between populations, leading to higher F_{st} than the expected under neutrality (Beaumont, 2005). According to F_{st} analysis, the genomic regions showing differentiation signals between one fat-tail breed and at least 10 of the Italian thin-tail breeds were reported in Table 1. F_{st} , χ^2 values and χ^2 P-values of the significant markers for each pairwise comparison were reported in the Supplementary Tables S1, S2, S3, S4 , S5 and S6 respectively for the six fat-tail breeds. Because the majority of the cited studies report the name of the marker as defined in the OAR v1.0 genome assembly, both this name as well as the "rs" number were reported in the tables; chromosome and position refer to the OAR v4.0 assembly. Because of the constraints used to define the genomic regions that are differentiated, the size of the signals was highly variable, both between regions and between breeds in the same region, ranging from few tens of base pairs to 1 Mb in the majority of cases (Table 1) It is worth noting that the Chios breed showed a peak of eight consecutive significant markers on OAR 13, encompassing over 6 Mb, from position 48,552,093 to 55,289,750 (Table 1).

The six Manhattan plots depicting signals of differentiation between the fat-tail breeds and the 13 thin-tail were reported in Fig.1. The y axis shows the χ^2 values of the 43072 SNPs. The different scales of the Y-axis gives evidence of the global genomic differences between the fat-tail and the Italian thin-tail breeds, which is more accentuated in the Barbaresca, Cyprus Fat-Tail, Chios and Ossimi compared to the Laticauda and Libyan-Barbary. The plots give also moderate evidence on one peak on OAR 13 in the Laticauda, Cyprus Fat-Tail, Ossimi and Chios breeds; one peak on

187 OAR 6 in the Barbaresca, Laticauda and Libyan-Barbary, and one peak on OAR 10 in the three
188 Barbary breeds and the Ossimi.
189 Remarkably, no signal was shared by all six fat-tail breeds, but one signal (OAR 7: 82.0-82.9 Mb)
190 was shared by all breeds except the Libyan-Barbary. This peak on OAR 7 appears less evident
191 because of the higher significant values of the peak of OAR 13. Several signals on the other hand
192 were shared by either two or three fat-tail breeds. All the detected signals are reported no matter
193 whether shared by more breeds or not, because they are meant to provide an information reservoir
194 for future studies on selection signatures in the sheep genome.
195

196 **Discussion**

197 The complexity of the fat-tail phenotype (Tibin 2007; Gizaw 2008; Wei *et al*. 2015) partly justifies
198 the extremely high number of signals detected in the pairwise comparisons (Table 1); many signals
199 may therefore be ascribed to other phenotypes and functions, and not just fat-tail phenotype or
200 adipogenesis.
201 Under the hypothesis that the signals directly connected to fat deposition and adipogenesis were to
202 be found among the signals shared by the breeds of different origin (Cyprus Fat-Tail, Chios and
203 Ossimi on one side and the Barbary breeds on the other side), the signals shared by the Libyan-
204 Barbary with the Italian breeds may represent Barbary related traits not automatically involved in
205 fat deposition. The low number of sampled animals of the Ossimi breed would not allow
206 conclusions on genes influencing any trait of this specific breed; however, this breed was included
207 in the analysis because SNPs associated to the fat-tail, if detected in the Ossimi as well as in other
208 breeds will corroborate the hypothesis of association.
209 All signals therefore will be discussed taking with special consideration the literature focusing on
210 either sheep fat-tail or adaptation to arid areas, in agreement with the literature which reported the
211 connection of fat-tail phenotype with adaptation traits (Atti *et al.* 2004; Kashan *et al*. 2005).
212 To the best of our knowledge, five GWAS targeting the fat-tail phenotype have been performed to
213 date; they used different fat-tail breeds, including the Iranian Lori-Bakhtiari (Moradi *et al*. 2012),

the Cyprus Fat-Tail and Laticauda in Europe (Moioli *et al*. 2015), ten indigenous Chinese fat-tail sheep (Wei *et al*. 2015; Yuan *et al*. 2016) and two Chinese Han sheep with tails of different size (Xu *et al*. 2017). Moreover, Yang *et al*. (2016) performed a whole-genome sequencing of 21 native Chinese sheep breeds and gave insight into signatures in the genome differentiating breeds from extreme environments: plateau, high altitude, desert and arid areas, while Lv *et al*. (2014) identified the genomic regions where climate-mediated selective pressure had shaped phenotypic variation in sheep. The following discussion will first examine those signals which confirm the signals previously mentioned in the literature as fat-tail signatures. In a second stage, the signals for which no connection with the fat-tail was to date reported will be examined, as these new signals might encode genes involved in fat deposition or adaptation or simply be the signals that Italian Barbary sheep have inherited by the North-African sheep.

Putative fat-tail signals

Twelve of the fat-tail signals reported previously were identified also in the present study. The first (OAR 2: 52.0-53.3 Mb) was detected in the Libyan-Barbary, Cyprus Fat-Tail, Chios and Ossimi, but not in the two Italian breeds. It was reported as fat-tail signature by Moradi *et al.* (2012) in the fat-tail Lori-Bakhtiari, while Wei *et al.* (2015) who identified the same signal in the fat-tail Duolang breed associated it to growth traits.

The signal on OAR 3:154.0-155.6 Mb, shared by the two Italian Barbary breeds and the Chios, was reported by Yuan *et al*. (2017) in a study including seven indigenous Chinese sheep (three thin-tail of Tibetan origin, two short fat-tail and two long fat-tail), by contrasting fat-tail vs. thin-tail phenotypes. It is hypothesized that this is a signal of fat deposition because the fat-tail group analyzed by Yuan *et al*. (2017) included both long and short-tail breeds; similarly, in our study, the signature was identified in Barbaresca, Laticauda and Chios breeds, therefore both Barbary and non-Barbary breeds, the former having a definitely longer fat-tail than the Laticauda and the Chios. However, it must be noted that, in the same region, Fariello *et al*. (2014) reported a signature differentiating the Red Maasai from the Ethiopian Menz sheep. The Red Maasai sheep is the only

241 breed resistant to the Haemonchus contortus parasite; genetic differences in the impact of infection
242 and resistance to this parasite have been demonstrated in the comparison of this breed with the
243 Dorper sheep (Baker *et al*. 2004). The Red Maasai have been shown to survive under high
244 trypanosome challenge (Wanyangu *et al.* 1993). It would be worth substantiating through specific
245 experimental trials whether the two Italian and the Chios breeds, which share the signature on OAR
246 3:154. 0-155.6 Mb, also possess higher resistance to gastro-intestinal parasites.
247 On OAR 4:48.5-48.6 Mb the Libyan-Barbary showed a signal of differentiation from the Italian
248 thin-tail breeds not detected in the other fat-tail breeds considered here. This region was reported as
249 fat-tail signature in Chinese sheep by Yuan *et al.* (2017). However, in the study of adaptation to
250 extreme environments, Yang *et al*. (2016) detected the same signal in sheep from arid areas. It is
251 likely, therefore, that this signal encodes genes influencing both fat-tail and adaptation, being
252 evident only in the Libyan-Barbary, which tolerates arid environments much more extreme than
253 those in the Mediterranean and Southern Italy.
254 The signal on OAR 6:36.0-36.4 Mb was here detected only in the Barbaresca breed, but was
255 reported as fat-tail signal by Yuan *et al.* (2017) in Chinese sheep. However, according to Yang *et al*.
256 (2016), in Chinese sheep this was a signal of adaptation to arid areas. Because these two signals
257 were not shared by other breeds, the role of these regions should be further investigated in other
258 sheep populations to elucidate their role in shaping phenotypes.
259 The signal on OAR 7: 82.1-82.9 Mb was the most shared one in the present study, being identified
260 in five breeds out of the analyzed six: the two Italian Barbary and Cyprus Fat-Tail, the Chios and
261 the Ossimi. The signal was already reported as fat-tail signature in the Iranian Lori-Bakhtiari
262 (Moradi *et al*. 2012). This region encodes the *VRTN* gene, which is associated to variation in
263 vertebral number (Mikawa *et al*. 2011). The variability within and between breeds of thoracolumbar
264 vertebrae number in sheep (17 to 21) was recently associated to carcass traits (Zhang *et al*. 2017);
265 these authors reported that another gene (*NR6A1*) influences thoracic and lumbar vertebral number
266 independently. The tail of the five breeds of the present study where this signature was detected is
267 definitively longer than the tail of the Italian thin-tail breeds used in the pairwise comparisons, and

268 particularly the tails of the Barbaresca and the Cyprus Fat-Tail (21 vertebrae) are so long as to trail
269 on the ground.
270 The signal on OAR 10:26.6-27.6 Mb, shared by the two Italian breeds, was reported as fat-tail
271 signature in Chinese sheep by Yuan et al. (2017). This region encodes the RFC3 gene, recognized
272 to play a role in cattle environmental responses and adaptation by Wang et al. (2015).
273 The signal on OAR 10:29.1-30.7 Mb, shared by the three Barbary and the Ossimi breeds, was
274 already reported as fat-tail signature in Chinese sheep (Yuan *et al*. 2017). This signature was also
275 reported by Seroussi *et al*. (2017) in sheep obtained through crossbreeding of the highly prolific
276 Afec-Assaf, the fat-tail Awassi and the dairy East Friesian and Boroola Merino breeds. The authors
277 referred to this region as signal of climate adaptation. This conclusion being supported by the
278 presence of the *ALOX5AP* gene, which encodes a protein that is required for the synthesis of lipid
279 mediators involved in various types of inflammatory responses (Seroussi *et al.* 2017). However, the
280 crossbred sheep analyzed by Seroussi *et al*. (2017) received genes also from the fat-tail Awassi and
281 Afec Assaf. Yang *et al.* (2016), in Chinese sheep, previously reported this region as a signal of
282 adaptation to desert areas. Therefore it is likely that this region of the genome, as well as the
283 upstream region previously mentioned (OAR 10:29.1-30.7 Mb) encode genes influencing more than
284 one quantitative trait, and this confirming also that fat-tail and adaptation are strictly connected
285 traits (Atti *et al.* 2004; Kashan *et al.* 2005).
286 The signal on OAR 11: 18.1-18.4 Mb, detected only in the Ossimi breed (Table 1), was reported
287 previously as signature of fat tail by Wei *et al*. (2015), but also as signal of adaptation to arid areas
288 by Yang *et al.* (2016).
289 The signal on OAR 13:48-49 Mb, shared by Cyprus Fat-Tail, Chios, Ossimi and Laticauda, was
290 already reported as fat-tail signature in Chinese sheep (Wei *et al*. 2015; Yuan *et al*. 2017). This
291 region is particularly large and evident in the Chios breed (Table 1, Figure 1). The strong LD
292 between the SNPs in this OAR 13 region with a missense mutation in exon 1 of the *BMP2* gene
293 (OAR13: 48,552,093-48,897,111) was demonstrated by Moioli *et al*. (2015) in the Laticauda fat-

294 tail as well as in the Altamurana thin-tail sheep. Yuan *et al*. (2017) emphasized that *BMP2* gene
295 may play important roles in fat tail formation.
296 The Laticauda and the Libyan-Barbary, together with the Ossimi, shared one signal (OAR 15:3.5-
297 3.9 Mb) which was reported as fat-tail signature in Chinese sheep by Wei *et al.* (2015) and Yuan *et*
298 *al.* (2017). Because the region encodes the *PDGFD* gene (*PDGF* family), Yuan et al. (2017)
299 suggested a role of *PDGF* gene in the fat-tail phenotype because it promotes cell proliferation,
300 inhibits differentiation of preadipocytes, and is expressed at a higher level in adipose tissue.
301 Yuan et al. (2017) reported also a fat-tail signature on OAR 22, corresponding to the signal detected
302 in the present study in the Libyan-Barbary and the Cyprus Fat-Tail (OAR 22:36.3-36.5 Mb).
303 Interestingly, one SNP (s19503.1) located only 280 KB downstream the OAR 22 region, was
304 reported by these authors as the top SNP of the region, with Fst value = 0.36 between fat and thin
305 tail Chinese breeds.
306 Finally, on OAR 25:7.0-7.3 Mb, only the Ossimi breed showed the signal previously reported by
307 Yuan *et al*. (2017) as fat-tail signature. However, one gene (*TARBP1*) encoded in this region, was
308 reported by Keane *et al*. (2006) to influence the mechanisms of genetic resistance to gastro
309 intestinal nematodes in sheep.
310
311 *Other signals*
312 The signal on OAR 1:68.9-69.5 Mb was identified only in the Cyprus Fat-Tail, and according to Lv
313 *et al*. (2014) it indicated genomic variation distinguishing the sheep breeds under investigation
314 based on the percent maximum sunshine in the geographical area where they were raised. These
315 authors had in fact clustered 32 native sheep breeds from all over the world according to the
316 environment they were adapted to inhabit.
317 The signal on OAR 2: 28.6-28.7 Mb, detected only in the Ossimi breed (Table 1), was reported
318 previously as signal of adaptation to arid areas by Yang *et al*. (2016).

On OAR 3:186.2-186.4 Mb, the Libyan-Barbary and the Cyprus Fat-Tail breeds shared a region previously identified by Lv *et al.* (2014), in their investigation on the 32 native sheep breeds, as signal of selective response to both the percent maximum sunshine and the relative humidity.

The signal on OAR 6:37-39 Mb was shared only by the Italian fat-tail breeds. According to the literature, it is likely that this region of the genome encode genes influencing more than one quantitative trait. In fact, a GWAS on Australian Merino sheep (Al-Mamun *et al.* 2015) described a region on OAR 6, from 36.0 to 38.0 Mb, which includes 13 SNPs significantly associated ($P < 0.001$) to body weight. Seven of the 13 SNPs identified by Al-Mamun *et al*. (2015): OAR6_40370293.1, OAR6_40409402.1, OAR6_40449774.1, OAR6_41558126.1, OAR6_41768532.1, OAR6_41936490.1 and OAR6_42247197.1, were highly significant in the Barbaresca breed (Supplementary Table S1) in the pairwise comparisons with all 13 breeds. In the other fat-tail breeds some of the SNPs were also significant in the comparisons with only a few breeds (Supplementary Tables S2, S3, S4 and S5), including the reported positive correlation between long tail and body weight of the lamb (Oltenacu *et al.* 1974). This region encodes some biologically relevant genes (*LAP3, MED28, FAM184B, DDB1, DCAF16, NCAPG* and *LCORL*; Table 1) expressed in ovine adipose tissue depots and skeletal muscle (Al-Mamun *et al.* 2015) so substantiating the involvement of this region also in fat deposition.

The genomic region on OAR 6: 69.6-69.9 Mb was shared by all the three breeds of Barbary origin, and was reported previously as signature of adaptation to arid areas by Yang *et al.* (2016). On the other hand, Fariello *et al*. (2014) ascribed to this region it a role in pigmentation, because it differentiated the Valais Blacknose sheep from three other Swiss sheep, and because it encodes the *KIT* gene, a candidate for development and migration of melanocytes. However, this region encodes also the *PDGFRA* gene (platelet-derived growth factor receptor A) (Table 1) a potent stimulator of proliferation, which requires the *PDGF* gene as ligand for its activation. The role *PDGF* gene in the fat-tail phenotype was already described for the signature on OAR 15:3.5-3.9 Mb which encoded the *PDGFD* gene, which is highly expressed in adipose tissue (Yuan *et al.,* 2017), further supporting the role of this region in influencing the fat-tail phenotype.

346 The signal on OAR 7: 31.6-34.5 Mb, shared by Barbaresca, Libyan-Barbary, Cyprus Fat Tail and
347 Chios (Table 1), was previously reported as signal of adaptation to extreme environments, because
348 it was detected by Yang *et al*. (2016) only in sheep from arid areas. However, since the analyzed
349 samples by Yang *et al*. (2016) included also fat-tail sheep, then it could be assumed that also fat-tail
350 phenotype might be influenced by the genes in this region, since the signal is shared by both
351 Barbary and non-Barbary breeds.
352 The signal on OAR 11: 51.0-51.5 Mb, detected only in the Cyprus Fat-Tail breed (Table 1), was
353 also reported as signal of adaptation to arid areas by Yang *et al*. (2016).
354 One signal, at OAR 23:62.1-62.2 Mb, was shared by the Cyprus Fat-Tail and the Ossimi, but by
355 none of the Barbary breeds. Although not reported previously in the literature, this signal is worth
356 examination because it falls in the genomic region encoding three genes of the Serpin family (Table
357 1). The major physiological function of these genes is the protection of the respiratory tract in
358 mammals, by preserving cellular protein integrity through the inhibition of proteolysis
359 (http://www.genecards.org) so supporting their role in adaptation to harsh environments. A
360 differential expression of Serpina1 between start and peak of lactation in dairy sheep was reported
361 by Signorelli et al. (2012).
362

363 **Conclusion**

364 Detecting genetic signatures of selection is of great interest for many research questions. A novel
365 putative fat-tail signature, shared by the three breeds of Barbary origin, was identified by the
366 present study on OAR 6:69.6-69.9 Mb. This signature encodes the *PDGFRA* gene, which is
367 activated by the *PDGF* gene; the role of this gene in the fat-tail phenotype was strengthened by the
368 detection, in the Libyan-Barbary and the Laticauda breeds, of the fat-tail signature on OAR 15
369 encoding a gene of the same family (*PDGFD*). The results reported here confirm that fat-tail and
370 adaptation are strictly connected traits, and in two cases also genes involved in nematode resistance
371 were detected within the signatures of adaptation. The genetic connection between fat deposition
372 and adaptation was previously observed in the Chinese sheep, but this is the first study which

373 demonstrated the connection in the Mediterranean breeds. Particularly, the signatures on OAR
374 6:36.0-36.5 Mb, OAR 10:29.0-30.8 Mb and OAR 13:48.5-49.1 Mb, notably reported as signature of
375 fat-tail as well as of adaptation to desert areas, have been confirmed as signatures differentiating
376 five Mediterranean fat-tail breeds from the Italian thin-tail breeds. Indeed, the only two fat-tail
377 Italian breeds, Barbaresca and Laticauda, are reared in Italian areas of hot and dry climate. There
378 are many fat-tail breeds of sheep in the world, and only few have been investigated at the genome-
379 wide level; therefore further studies on different fat-tail breeds will be profitable to clarify the
380 complexity of this phenotype which might represent an asset to face climate change.
381
382 **Conflicts of interest**
383 The authors declare no conflicts of interest.
384
385 **Acknowledgements**
386 The authors are grateful to Drs. H. Lenstra (Utrecht University, The Netherlands), P. Ajmone-
387 Marsan and Licia Colli (University of Piacenza, Italy) for providing the DNA samples of the Ossimi
388 breeds collected within the “Econogene” project.
389
390 **References**
391 Akraim F, Milad IS, Abdulkarim AA, Ganem M (2008) Wool characteristics of Libyan Barbary
392 sheep in north-eastern Libya: I. fiber diameter and staple length. Livestock Research for Rural
393 Development **20**: 2008. Retrieved on 24 May 2017 from http://www.lrrd.org/lrrd20/8/
394 akra20118.htm.
395 Al-Mamun HA, Kwan P, Clark SA, Ferdosi MH, Tellam R, Gondro C (2015) Genome-wide
396 association study of body weight in Australian Merino sheep reveals an orthologous region on
397 OAR6 to human and bovine genomic regions affecting height and weight. *Genetics Selection*
398 *Evolution* **47,** 66.

399 Atti N, Bocquier F, Khaldi G. (2004) Performance of the fat-tailed Barbarine sheep in its
400 environment: adaptive capacity to alternation of underfeeding and re-feeding periods. A review.
401 *Animal Research* **53**, 165–176.
402 Baker RL, Mugambi JM, Audho JO, Carles AB, Thorpe W (2004) Genotype by environment
403 interactions for productivity and resistance to gastro-intestinal nematode parasites in Red Maasai
404 and Dorper sheep. *Animal Science* **79,** 343–353.
405 Bakhtiarizadeh MR, Moradi-Shahrbabak M, Ebrahimie E (2013) Underlying functional genomics
406 of fat deposition in adipose tissue. *Gene* **521**, 122–8.
407 Beaumont MA (2005) Adaptation and speciation: what can FST tell us? *Trends in Ecology and*
408 *Evolution* **20**,8.
409 Bigi D, Zanon A (2008) Atlante delle razze autoctone: Bovini, equini, ovicaprini, suini allevati in
410 Italia (in Italian). Edagricole, Milan, Italy.
411 Bonhomme M, Chevalet C, Servin B, Boitard S, Abdallah J, Blott S, San Cristobal M (2010)
412 Detecting Selection in Population Trees: The Lewontin and Krakauer Test Extended. *Genetics* **186**,
413 241–262.
414 Ciani E, Ciampolini R, D'Andrea M, Castellana E, Cecchi F, Incoronato C, d'Angelo F, Albenzio
415 M, Pilla F, Donato M, Cianci D (2013) Analysis of genetic variability within and among Italian
416 sheep breeds reveals population stratification and suggests the presence of a phylogeographic
417 gradient. *Small Ruminant Research* **112,** 21-27.
418 Ciani E, Crepaldi P, Nicoloso L, Lasagna E, Sarti F M, Moioli B, Napolitano F, Carta A, Usai G,
419 D'Andrea M, Marletta D, Ciampolini R, Riggio V, Occidente M, Matassino D, Kompan D,
420 Modesto P, Macciotta NPP, Ajmone-Marsan P, Pilla F (2014) Genome-wide analysis of Italian
421 sheep diversity reveals a strong geographic pattern and cryptic relationships between breeds.
422 *Animal Genetics* **45**, 256-266.
423 Elshennawy M (1995) Sheep development program in Egypt. In : Gabiñ a D.(ed.). Strategies for
424 sheep and goat breeding. Zaragoza.: CIHEAM, p. 27 32 (Cahiers Options Méditerranéennes; n. 11).

425 Fariello M-I, Servin B, Tosser-Klopp G, Rupp R, Moreno C, San Cristobal M, Boitard S (2014)
426 Selection signatures in worldwide sheep populations. *PLoS ONE* **9**, e103813.
427 Gizaw G (2008) Sheep resources of Ethiopia: genetic diversity and breeding strategy. Thesis PhD,
428 Wageningen University, The Netherlands.
429 Keane OM,, Zadissa A, Wilson T, Hyndman DL, Greer GJ, Baird DB, McCulloch AF, Crawford
430 AM, McEwan JC (2006) Gene expression profiling of Naïve sheep genetically resistant and
431 susceptible to gastrointestinal nematodes. *BMC Genomics* **7**,42.
432 Lv FH, Agha S, Kantanen J, Colli L, Stucki S, Kijas JW, Joost S, Li M-H, Ajmone Marsan P (2014)
433 Adaptations to Climate-Mediated Selective Pressures in Sheep. *Molecular Biology and Evolution*,
434 **31**(12),3324–3343
435 Kashan NEJ, Azar GMH, Afzalzadeh A,Salehi A (2005) Growth performance and carcass quality
436 of fattening lambs from fat-tailed and tailed sheep breeds. *Small Ruminant Research* **60**,267-271.
437 Kijas JW, Lenstra JA, Hayes B, Boitard S, Porto Neto LR, San Cristobal M, Servin B, McCulloch
438 R, Whan V, Gietzen K, Paiva S, Barendse W, Ciani E, Raadsma H, McEwan J, Dalrymple B (2012)
439 Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong
440 recent selection. *PLoS Biology*,10, e1001258.
441 Mason IL (1967) Sheep breeds of the Mediterranean. Food and Agricultural Organization of the
442 United Nations, Rome, Italy.
443 Mastrangelo S, Portolano B, Di Gerlando R, Ciampolini R, Tolone M, Sardina MT, International
444 Sheep Genomics Consortium (2017) Genome-wide analysis in endangered populations: a case
445 study in Barbaresca sheep. *Animal* **12:** 1-10.
446 Mastrangelo S, Ciani E, Sardina MT, Sottile G, Pilla F, Portolano B, Bi.Ov.Ita Consortium (2018)
447 Runs of homozygosity reveal genome-wide autozygosity in Italian sheep breeds. *Animal Genetics,*
448 DOI: 10.1111/age.12634.
449 Mikawa S, Sato S, Nii M, Morozumi T, Yoshioka G, Imaeda N, Yamaguchi T, Hayashi T, Awata
450 T(2011) Identification of a second gene associated with variation in vertebral number in domestic
451 pigs. *BMC Genetics* **12:** 5.

452 Moioli B, Pilla F, Ciani E (2015) Signatures of selection identify loci associated with fat tail in
453 sheep. *Journal of Animal Science* **93**, 4660–4669.
454 Moradi MH, Nejati_Navaremi A, Moradi-Shahrbabak M, Dodds K, McEwan JC (2012) Genomic
455 scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions
456 association with fat deposition. *BMC Genetics* **13**, 10.
457 Nejati-Javaremi A, Izadi F, Rahmati G, Moradi M, Izadi F (2007) Selection in fat-tailed sheep
458 based on two traits of fat-tail and body weight versus single-trait total body weight. *International*
459 *Journal of Agriculture & Biology* **4**, 645–8.
460 Theodoridis A, Ragkos A, Roustemis D, Galanopoulos K, Abas Z, Sinapi E (2012) Assessing
461 technical efficiency of Chios sheep farms with data envelopment analysis. *Small Ruminant*
462 *Research* **107**, 85-91.
463 Tibin MAM (2007) Performance, carcass characteristics and meat quality attributes of desert sheep
464 under range conditions. Thesis BSc, University of Khartoum, Sudan.
465 Oltenacu EAB, Boylan WJ (1974) Inheritance of tail length in crossbred Finnsheep. J*ournal of*
466 *Heredity* **65**: 331-334.
467 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de
468 Bakker PI, Daly MJ and Sham PC 2007. PLINK: a tool set for whole-genome association and
469 population-based linkage analyses. *American Journal of Human Genetics* **81**, 559‑575.
470 Qanbari S, Pimentel EC, Tetens J, Thaller G, Lichtner P, Sharifi AR, Simianer H (2010) The pattern
471 of linkage disequilibrium in German Holstein cattle. *Animal Genetics* **41**(4),346-56.
472 Seroussi E, Rosov A, Shirak A, Lam A, Gootwine E (2017) Unveiling genomic regions that
473 underlie differences between Afec‑Assaf sheep and its parental Awassi breed. *Genetics Selection*
474 *Evolution* **49,**19.
475 Signorelli F, Cifuni GF, Miarelli M (2012) Differentially expressed mammary proteins during
476 lactation in dairy sheep. *Livestock Science* **149**, 224–231.
477 Wang MD, Dzama K, Hefer CA, Muchadeyi FC (2015) Genomic population structure and
478 prevalence of copy number variations in South African Nguni cattle. *BMC Genomics* **16,**894.

Wanyangu SW, Mugambi JM, Bain RK, Duncan JL, Murray M, Stear MJ (1997) Response to artificial and subsequent natural infection with Haemonchus contortus in Red Maasai and Dorper ewes. *Veterinary Parasitology* **69,** 275-282.

Wei C, Wang H, Liu G, Wu M, Cao J, Liu Z, Liu R, Zhao F, Zhang L, Lu J, Liu C, Du L (2015) Genome-wide analysis reveals population structure and selection in Chinese indigenous sheep breeds. *BMC Genomics* **16**,194.

Wu X, Fang M, Liu L, Wang S, Liu J, Ding X, Zhang S, Zhang Q, Zhang Y, Qiao L, Lund MS, Su G, Sun D (2013) Genome wide association studies for body conformation traits in the Chinese Holstein cattle population. *BMC Genomics* 14,897.

Xu SS, Ren X, Yang GL, Xie XL, Zhao YX, Zhang M, Shen ZQ, Ren YL, Gao L, Shen M, Kantanen J, Li MH (2017) Genome-wide association analysis identifies the genetic basis of fat deposition in the tails of sheep (Ovis aries) *Animal Genetics* 48,560-569.

Yang J, Li WR, Lv FH, He SG, Tian SL, Peng WF, Sun YW, Zhao YX, Tu XL, Zhang M, Xie XL, Wang YT, Li JQ, Liu YG, Shen ZQ, Wang F, Liu GJ, Lu HF, Kantanen J, Han JL, Li MH, Liu MJ (2016) Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Molecular Biology and Evolution* **33**, 2576-2592.

Yuan Z, Liu E, Kijas JW, Zhu C, Hu S, Ma X, Zhang L, Du L, Wang H, Wei C (2017). Selection signature analysis reveals genes associated with tail type in Chinese indigenous sheep. *Animal Genetics* **48**: 55-66.

Zhang Z, Sun Y, Du W, He S, Liu M, Tian C (2017) Effects of vertebral number variations on carcass traits and genotyping of Vertnin candidate gene in Kazakh sheep. *Asian-Australasian Journal of Animal Science* **30**(9),1234-1238.

Figure captions:

Figure 1. Manhattan plots depicting signals of differentiation between six fat-tail breeds and 13 thin-tail Italian breeds.

Supplementary materials:

Supplementary Table S1. Fst, χ^2 and χ^2 P-values of the markers shaping the genomic regions of Table 1 in the pairwise comparisons of the Barbaresca with 13 Italian thin-tail breeds.

Supplementary Table S2. Fst, χ^2 and χ^2 P-values of the markers shaping the genomic regions of Table 1 in the pairwise comparisons of the Laticauda with 13 Italian thin-tail breeds.

Supplementary Table S3. Fst, χ^2 and χ^2 P-values of the markers shaping the genomic regions of Table 1 in the pairwise comparisons of the Libyan-Barbary with 13 Italian thin-tail breeds.

Supplementary Table S4. Fst, χ^2 and χ^2 P-values of the markers shaping the genomic regions of Table 1 in the pairwise comparisons of the Cyprus Fat-Tail with 13 Italian thin-tail breeds.

Supplementary Table S5. Fst, χ^2 and χ^2 P-values of the markers shaping the genomic regions of Table 1 in the pairwise comparisons of the Ossimi with 13 Italian thin-tail breeds.

Supplementary Table S6. Fst, χ^2 and χ^2 P-values of the markers shaping the genomic regions of Table 1 in the pairwise comparisons of the Chios with 13 Italian thin-tail breeds.

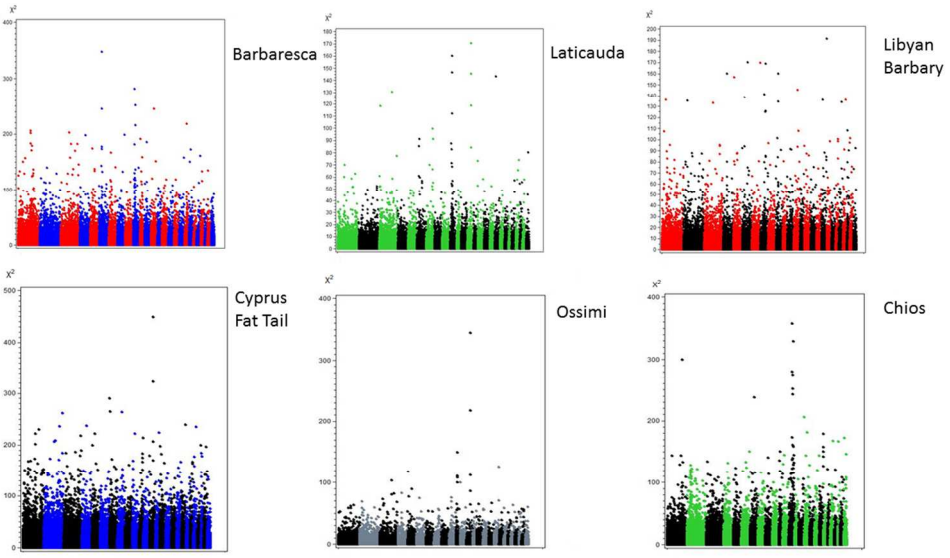

Figure 1. Manhattan plots depicting signals of differentiation between six fat-tail breeds and 13 thin-tail Italian breeds

338x190mm (96 x 96 DPI)

Table 1. Significant sweeps registered in fat-tail breeds. Start/end position are expressed based on the ovine genome sequence assembly Oar_v4.0.

	Barbaresca	Laticauda	Libyan-Barbary	Cyprus FatTail	Ossimi	Chios	Citation and target trait	Genes in the region
OAR	start/end	start/end	start/end	start/end				
1				17138949 17167984				*ZNF691, SLC2A1*
1			29000386 29049438		29049438 29078098			*DHCR24, BSND, TMEM61, USP24*
1				64022048 64465382				*LOC105609236, LOC101113557*
1				68912445 69482536			Lv et al. (2014) adaptation	*EVI5, RPL5, FAM69A, MTF2, TMED5*
1					84281785 84329444			*VAV3*
1			103040031 103455454		100532202 100584364			*AQP10, HAX1, ATP8B2, IL6R, UBE2Q1, TDRD10, SHE, LOC105608947, ADAR, CHRNB2*
1				154937981 154966674				*LOC105616630*
1	161588670 161912812							*ST3GAL6, DCBLD2, LOC101107320*
1	163538737 163965269			164645667 164882801	164645667 164882801			*ABI3BP, ADGRG7, TMTM45A, LOC101109067*
						178706990 178755797		*GAP43*
1			192523354 192876915	192523354 192580340				*MB21D2*
1					198342295 198427122			*ST6GAL1, ADPOQ, RFC4, EIF4A2*
						210198899 210478841		-
						217103660 217352839		*GOLIM4, PDCD10, SERPINI1, WDR49*
						218603664 218692027		-
1			239050505 239453954					*LOC105604038*
1				244308436 244308436				*PLS1, ATR, XRN1*
2					28642475		Yang et al.	*IPPK, ECM2, ASPN, CENPP, OMD, OGN*

					28685512		(2016) arid zone	
2			52081095 52463944	52508876 52770626	52106037 52463944	52106037 53305743	Moradi et al. (2012) tail; Wei et al. (2015) growth	*GNE CLTA CCIN GLI PR2 RECK TMEM8B FAM221B NPR2 SPAG8 HINT2*
2	56006296 56331862							*LOC101111526*
2			58659378 58973759			58659378 58805835		*GNAQ, GNA14*
2	87286590 87291060							*PLIN2, DENND4C, RPS6, ACER2*
2			104109704 104273138	104109704 104396663	104109704 104451808			*LOC105608545, LOC101118164, FAM167A, BLK*
2				145652259 145769645				*KCNH7*
2				175091617 175412735				*TMEM153, MGAT5*
2				193503875 193684368				*TMEFF2*
3			42296486 42427263					*LOC101117153*
3	122515163 122887650							*MGAT4C*
3	136341082 136433664							*FAIM2, BCDIN3D, NCKAP5L, LOC105612627, TMBIM6, PRPF40B, FMNL3, FAM186B, KCNH3, MCRS1, SPATS2, C1QL4, DNAJC22, PRPH*
3		129079963 129440404						*LOC105614718*
3	154033734 154318689	154033734 155439736				155554559 155842760	Yuan *et al*. (2017) tail; Fariello *et al*. (2015) nematode resistance	*MSRB3, LOC105609947, LEMD3, WIF1, TBC1D30*
3	169813339 169988856							*ANO4*
3				182602767 182650034				*AMN1, METTL20*
3			186226037 186477020	186226037 186447855			Lv et al. (2014) adaptation	*PTHLH, KLHL42, MANSC4, MRPS35, REP15*
3			209539484 209693423					*FGF6, FGF23, TIGAR, CCND2, LOC105608579, LOC106608577, PARP11*

3				212824931 212994019		212039601 212438523	*CECR2, ATP6V1E1, BCL2L13, BID, MICAL3*
4				24138603 24201492			*MEOX2*
4			47074699 47235537		47074699 47189034		*LOC105609617, CDHR3, SYPL1, NAMPT*
4			48510936 48649775			Yuan *et al*. (2017) tail; Yang *et al*. (2016) adaptation arid area	*COG5, GPR22, DUS4L, BCAP29, SLC26A4, SLC26A3, LOC105611148, CBLL1, DLD*
4				108772903 108930060			*LOC105615127*
5					6010087 6333143		*MED26, SMIM7, CHERP, SLC35E1, CALR3, EPS15L1, KLF2*
5					47543428 47785464		*CTNNA1, LRRTM2, SIL1*
5			51067599 51930923			51388369 51459909	*ARHGAP26, LOC105608313, NR3C1*
5	59126508 59475280						*CAMK2A, TCOF1, CD74, LOC105606717, RPS14, NDST1*
5						95208526 95481386	*RGMB, CHID1*
6	36034915 36390529					Yuan *et al* (2017) tail ; Yang *et al*. (2016) adaptation arid area; Lv et al. (2014) adaptation	*HERC3, PYURF, PIGY, HERC5, HERC6, PPM1K, ABCG2, PKD2, SPP1, MEPE*
6	37126564 39487124	36502071 39816933				Al-Mamun et al. (2015) weight	*IBSP, LAP3, MED28, LOC105615455, LOC105608051, MED28, FAM184B, NCAPG, LCORL, DCAF16, FAM184B, LOC105615456, LOC105608050, LOC105608049, LOC101122950, SLIT2*
6	69675370 69867326	69675370 69867326	69816517 69867326			Yang et al. (2016) arid zone Fariello *et al*. (2015) pigmentation	*CHIC2, LOC105613061, PDGFRA, GSX2, KIT*

6	77012913 77459287							*ADGRL3, LOC101114018*
6				103130311 103321439				*EVC2, STK32B*
6	115244531 115439490							*GRK4, NOP14, LOC106991221, LOC106991246, MFSD10, ADD1, SH3PB2, FAM193A*
7	33565208 33736820		34067457 34497866	31623228 34497866		33565208 33971063	Yang *et al*. (2016) adaptation arid area and desert	*RHOV, VPS18, DLL4, CHAC1, INO80, EXD1, CHP1, MGA, LOC10561677, MARKBP1, JMJD7, PLAG2G4B, PLAG2G4E, EDH4, SPTBN5*
7						51600923 51732837		*PRTG, PYGO1*
7						62056516 62140808 62169254		*GATM,SLC28A2, DUOX1*
7	82172198 82279625	82117886 82968403		82279625 82968403	82067322 82279625	82067322 82968403	Moradi *et al*. (2012) tail; Moioli *et al*. (2015) tail	*ELMSAN1, PNMA1, PTGR2, ZNF410, FAM16B, COQ6, ENTPD5, BBOF1, ALDH6A1, LIN52, VSX2, ABCD4, VRTN SYNDIG1L*
7				99115579 99230967				*TTC7B, RPS6KA5*
8			51733711 52318337	51856297 51875114				*LOC105611297, TBX18, CEP162*
8			62552594 62941437	62665338 62766989	62552594 62941437			*IFNGR1, IL22RA2, OLIG2, LOC105608892, TNFAIP3, PERP, LOC101113418*
8			90070184 90488251					*PHF10, TCTE3, ERMRD, LOC106991326, DLL1, FAM120B*
9	36057235 36326237							*PLAG1, CHCHD7, SDR16C5, SDR16C5, LOC101116323, PENK*
9						37040246 37495510		*FAM11B, UBXN2B, SDCBP, NSMAF*
9				53606563 53907509				*LOC101118031*
10				12090226 12454552				*VWA8*
10	17765501 17904934							*SUCLA2*
10	19626592 20969799			20538098 20571271				*DLEU7, FAM124A, LOC101119651*
10	22012281 23976329	23628017 23929416						*FOX01, MRPS3I, LOC105616150, LOC105608777, LOC105608775, LOC105608779, LOC105608780, LOC101122286, COG6, LHFP, NHLRC3, PROSER1,*

								STOLM3, FREM2, LOC101108400, UFM1
10	27402323 27623690	26688054 26917895					Yuan *et al*. (2017) tail	*LOC101109717, LOC101109981, RFC3, STARD13*
10	29101583 30823905	30591945 30717766	29588481 30717766		30591945 30717766		Yuan *et al*. (2017) tail; Moioli *et al*. (2015) tail; Seroussi *et al*. (2017) climate adaptation; Yang *et al*. (2016) desert adaptation	*BRCA2, ZAR1L, FRY, LOC106991357, RXFP2, LOC101110773, LOC106991379, B3GLCT, HSPH1, LOC105616258, TEX26, MEDAG, LOC105610262, ALOX5AP, USPL1, LOC111112330, LOC101112071, KATNAL1, LOC106991380, UBL3*
10		35552637 35728821					Yang *et al*. (2016) adaptation arid area	*LOC105609559, FGF9, MICU2, LOC105611671, ZDHHC20, SKA3, SAP18, MRPL57*
10					55986703 55992595			*LOC105607734*
11					4523179 4727109			*TOM1L1*
11					18103177 18408552		Wei et al. (2015) tail; Yang et al., (2016) desert adaptation	*RAB11FIP4, EVI2B, OMG, NF1*
11						24322534 28478905		*NCBP3, CAMKK1, P2RX1, ZZF1, ANKFY1, UBE2G1, SPN53, SPN52, TEKT1, MYBBP1A, GGT6, SMTNL, XAF1, FBOX39, SLC13A5, TXDC17, MED31, KIAA0753, CLDN7, SLC2A4, SPEM1, FGF11, CHRNB1, ZBTB4, SLC35G6, POLR2A, TNFSF12, TNFSF12, SAT2, SENP3, EIF4R1, FXR2, SHBG, CD68, MPDU1, SOX15, ATP1B2, WRAP53, EFNB3, DNAH2, KSM6B, TMEM88, NAA38, PER1*
11				41708179 41732873				*GHDC, HCRT, STAT5B, STAT5A, STAT3,*
11				51097504 51437604			Yang et al. (2016) arid zone	*ENDOV, NPTX1, RNF213, SLC26A11, CARDS14, EIF4A3, GAA, CCDC40, TBC1D16*
12	29027811 29441975							*SMYD3, LOC105616512, LOC105616511*
12			39193469 39430517					*AADACL4, DHRS3, VPS13D, LOC105606462*

12			42119363 42234968					*LOC105606443, LOC106991486, LOC105606444, SPSB1, LOC105616528, H6PD, LOC105606442,GPR157, SLC2A5*
12					66648281 66888838			*KCNK2*
13						27612267 28478097		*FAM107B*
13		48512859 49197707		48696401 49197707	48512859 49137513	48552093 55289750	Moioli *et al*. (2015) tail; Yuan (2017) tail; Wei *et al*. (2015) tail; Yang *et al*. (2016) adaptation arid area	*BMP2, LOC101117953*
13						56508160 58466919		*PRELTD3B, TUBB1, CTSZ, NELFCD, NPEPL1, STX16, APCDD1L, VAPB, RAB22A, ANKRD60, PMEPA, CTCFL, ZPB1, PCK1, RBM38, RAE1, BMP7,SPO11*
13						60103477 63347314		*DEFB125, DEFB115, REM1, HM13, ID1, BCL2L1, PLAGL2, POUFUT1, ASXL1, NOL4L, MAPRE1, BPIFB4, BPIFB1, BPIAF3, BPIAF1, DCDK5RAP1, SNTA1, E2F1, ASIP*
13						72530927 73010600		*WISP3, KCNK15, RIMS4, YWHAB, TOMM34, PABPC1L, STK4, KCNS1, PI3, MATN4, SCD4, SLPI*
14	12515054 12841601							*MAP1, LC3B, ZCCHC14, JPH3, KLHDC4, SLC7A5, LOC106991585, CA5A, BANP*
15		3505599 3879770	3505599 3879770		3505599 3709662		Yuan *et al*. (2017) tail	*PDGFD, DDI1*
15				16614403 16737249				*ALKBH8,ELMOD1, SLN*
15					24040958 24311283			*ZBTB16, RMB7*
15	30012057 30425982		30012057 30425982					*TRIM29, OAF, POU2F3, TMEM136, ARHGEF12*
15			72605179 72987859					*EXT2, ALX4, CD82, TSPAN18*
15					77751134 77797673			*P2RX3,SLC43A3,RTN4RL2, TIMM10, SMTNL1, UBE2LG*
16			26288412 26407287					*ITGA2, ITGA1, LOC106991663, LOC105602547*
17					44631305		Yang et al.,	*DDX51, EP400, PUS1, ULK1, MMP17, SFSWAP*

				44954100		(2016) desert adaptation	
17			53320204 53480272				*KDM2B, RNF34, CAMKK2, P2RX4*
17	70788082 70947463	70788082 70947463					*SMARCB1, SLC2A11, MIF, DERL3, MP11, ZNF70, VPREB3, CHCHD10, CABIN1, SUSD2, GGT5, SNRPD3, GUCD1, ADORA2A, UPB1, SPECC1L, BCR, RSPH14, GNAZ*
18		19330647 19479424					*LOC105603292, LOC105603293, LOC101119588, LOC101119079, LOC101119853, LOC101120114, LOC101120360, LOC101120610, LOC101120869, LOC101121378, LOC101121632, LOC10112139, LOC10112633, LOC101122387, ACAN, HAPLN3, MGE8*
18		38397777 38481481					*LOC105603166, LOC106991724, LOC106991725, LOC106991738, FOXG1*
18		66257896 66371408		66172063 66371408			*RCOR1, TRF3, AMN, CDC42BPB, LOC101104348, LOC101104530, LOC106990142, LOC105603310, EIFS5*
19					7674951 8432176	Yang et al., (2016) desert adaptation	*UBP1, CLASP2, PDCD6IP*
19			16027834 16415561				*TOPAZ1, TCAIM, ZNF445, ZNF852, KIAA1143, KIF15*
19	31614145 31942689	31614145 31942689					*LOC105607729, LOC105603449, MITF*
19	33342669 33399965	33342669 33399965					*FAM19A1*
19					56017177 56845772		*PLXND1, RHO, IFT122, MBD4, EFCAB12, CAND2, TMEM40, RAF1, MKRN2, MKRN20S, TSEN2, PPARG, TIMP4, SYN2*
21					1655570 1802139		*FAT3*
21		42325238 42492873	42325238 42492873				*SF1, MEN1, MAP4K2, CDC42BPG, EHD1, MIR194, ATG2A, PPP2R5B, GPHA2, BATF2, ARL2, SNX15, NAALADL1, SAC3D1, CDCA5, VPS51, TM7SF2, ZNHIT2, SYVN1, TMEM262, MRPL49, FAU, ZFPL1, SPDYC, CAPN1, SLC22A20*
22			15286671 15488173				*NOC3L, HELLS*
22		36385977 36465970	36385977			Yuan *et al*. (2017) tail	*ENO4, SHTN1, VAX1, KCNK18, SLC18A2, PDZD8*

23	25864051 25996978		25917146 25996978			*TRAPPC8, LOC105604435, B4GALT6, TTR, LOC106990157, DSG2, DSG3, DSG4, DSG1, LOC105604437*
23		62194482 62251113	62194482 62251113			*SERPINB11, SERPINB7, SERPINB10*
25			7011212 7338667		Yuan et al. (2017) tail Kean et al. (2006) nematode resistance	*TARBP1, IRF2BP2*
26			34592661 34731523			*ZMAT4, SFRP1*
26				36053084 37486784		*THAP1, RFN170, HOOK3, FNTA, INTS10, HGSNAT, SH2D4A, PSD3*