Genomic diversity and selection signatures underlying tail fatness inferred from genome-wide SNPs markers in local sheep in semi-arid area

I. Baazaoui[1], S. Bedhiaf-Romdhani[2*], S., Mastrangelo[3] and E. Ciani[4]

[1]Faculty of Sciences of Bizerte, University of Carthage, 7021 Jarzouna Bizerte,Tunisia;

[2]INRAT, Laboratory of Animal and Fodder Production, University of Carthage, 2049, Ariana, Tunisia.

[3]Department of Agricultural, Food and Forest Sciences, University of Palermo, 90128 Palermo, Italy.

[4] Department of Biosciences, Biotechnologies and Biopharmaceutics, University of Bari, 70121 Bari, Italy.

*Corresponding author: Sonia Bedhiaf-Romdhani, email: bedhiaf.sonia@gmail.com; romdhani.sonia@iresa.agrinet.tn

Genomic characterization of local Tunisian sheep

Abstract

In climate change perspective, the genetic make-up of local breeds showing adaptive traits of particular importance, should be explored and preserved as a priority. In this context, this study highlights the potential of genomic analysis to i) provide an overview of fine scale genomic variability within and among local breeds and ii) explore potential breed-specific selection footprints of local sheep under semi-arid conditions. Thus, we sampled 65 unrelated sheep breeds from Tunisia including 26 fat tailed Barbarine samples and 39 thin tailed sheep from the Noire de Thibar, the Queue fine de l'Ouest Thibar and the D'man breeds which are genotyped using Illumina OvineSNP50K array

data. The study of intra-breed genetic diversity based on estimation of genome-wide runs of homozygosity (ROH) distribution and inbreeding estimation (F_{ROH}) showed that Barbarine samples are less affected by inbreeding than other breeds. Moreover, the genomic relationships among 60 individuals based on network-based approach, using 46 232 genotypes confirmed the recent admixture between Barbarine and Queue Fine de l'Ouest animals resulting from the uncontrolled crossbreeding between two breeds, which suggest paying more attention to preserve the genetic integrity of those local breeds. As expected, all samples from Noire de Thibar breed were clearly differentiated from the remaining individuals explained by past introgression of European gene flow into the breed. The signals of selection identified across and within breeds were detected by selecting the top 10% of SNPs most commonly observed in a ROH region. For genome-wide scan of selective sweeps detection, we focused on the highest signal observed across and within analysis. Therefore, results across samples identified three significant ROH islands on chromosome 1,10 and 13 with highest signal on OAR 13. We uncover a strong signal of positive selection of an overlapping 3.03 Mb homozygous region spanning 46.58 – 49.61 Mb on chromosome 13 detected within Barbarine samples analysis with a peak of mean frequency P=0.31 harboring 20 candidate genes. These candidate putative genes preferentially selected in fat tailed Barbarine animals were *CDS2*, *PROKR1* and *BMP2* which are involved in fat deposition and lipid storage, a relevant sheep adaptive trait in semi-arid area. Our findings will primarily assist policy makers and livestock keepers for carrying out molecular breeding and conservation in sheep.

Keywords: local sheep, genomic relationships, selection footprints, fat deposition, *BMP2* gene.

Implications

In climate change context, fine-scale characterization of local sheep stock showing specific adaptive traits are urgently needed. Using ovine 50K SNP array data, this research highlights a potential threat of genetic homogenization between most important sheep breeds in Tunisia and uncover breed-specific footprints involved in fat tail formation in Barbarine sheep. This case study revealed how genomic data could be used to ascertain population structure, to exploit uniqueness of local sheep genetic stock to harsh conditions and give relevance for the conservation of native livestock breeds.

Introduction

In climate change perspective, adaptation and resilience of domestic animals to challenging environments is becoming increasingly important. In the last decades, many practices lead to loss of sheep genetic diversity, which serves as a raw material for livestock populations to adapt to changing environments (FAO, 2015). Therefore, an effective management program for their monitoring and preservation of purebred genetic stock are fundamental to meet future breeding needs, especially in the context of global climate change. Recently, the genome-wide ovine SNP data have become the marker of choice for underlying genetic diversity, inferring population structure and mapping genomic regions subject to selection. Indeed, the mapping of certain genomic regions often shows reduced genetic diversity and stretches of autozygosity at the individual and population levels, called runs of homozygosity (ROH) due to identical by descent chromosomal segments arising from a common ancestor, its identification and characterization can elucidate about recent inbreeding (Peripolli *et al.,* 2018) and selection pressure (Mastrangelo *et al.,* 2017) events. Sheep population in Maghreb represent a perfect case study of local adaptive sheep breeds raised under traditional management

system, threatened by an extensive anarchic crossbreeding among breeds, hence presenting a genetic homogenization and at risk to lose their adaptive traits to face future unpredictable environment. Most of Tunisian sheep are managed under transhumance in semi-arid area. Three adaptive local breeds make up the essential of meat sheep; there are the Barbarine with its black and red headed ecotypes, the Queue fine de l'Ouest and the Noire de Thibar in addition to small population of imported Moroccan D'man breed (supplementary Table S1). Previous fine scale genetic diversity of Tunisian sheep breeds has been to date investigated through RAPD-PCR (Khaldi *et al.,* 2010; Hentati *et al.,* 2012) and microsatellites markers (Sassi-Zaidy *et al.,* 2014b; Kdidi *et al.,* 2015). Thus, the present study is the first genome-wide study using 50K SNP ovine array in order to i) provide an assessment of genetic variability within and among four local Tunisian sheep population ii) identify genomic regions under selection within genomic ROH hotspots. This investigation is an important prerequisite to maintain local breed's integrity and ensure appropriate conservation

Materials and Methods

Sampling

A total of 65 sheep blood samples were randomly sampled from four meat breeds. The Barbarine (BAR), the only Tunisian fat tailed sheep breed was represented by 26 samples and 39 three thin-tailed sheep breeds including the Queue fine de l'ouest (QFO), the Noire de Thibar (NDT) and the D'man (DMN) breed. To avoid the probability of relatedness among individuals, samples were collected from different types of farms: public farms that have been tasked by the ministry of Agriculture with preserving the breed and private farms located mainly in the northern and central part of the country (Figure1). Details

about phenotypic descriptors, origin and specific traits of studied breeds were summarized in supplementary Table S1.

Genotyping and data quality control

The genomic DNA was isolated using the standard phenol-chloroform protocol (Sambrook *et al.,* 1989)**.** Samples were genotyped using the Illumina *Ovine* SNP50 BeadChip (San Diego, USA) and genotypic data was managed using PLINK 1.7 (Purcell *et al.,* 2007)**.** Only SNPs located on autosomes were considered in further analyses. A SNP was removed from data if the following criteria were not met: (i) a SNP call rate > 90%, (ii) minor allele frequency (MAF) > 0.01. Individuals with more than 10% of missing SNPs were removed from further analysis.

Identification of runs of homozygosity (ROH)

ROH estimation. Runs of homozygosity (ROHs) were computed across autosomes for each individual using PLINK 1.07. No pruning was performed based on linkage disequilibrium (LD), but the minimum length that constituted the ROH was set to 1 Mb, to exclude short ROH deriving from LD. The following criteria were used to define the ROH: (1) one missing SNP was allowed in the ROH and up to one possible heterozygous genotype (2) the minimum number of SNPs that constituted the ROH was set to 30 (3) the minimum SNP density per ROH was set to 1 SNP every 100 kb (100 000pb) (4) maximum gap between consecutive homozygous SNPs of 250 Kb.

ROH distribution and inbreeding coefficient (F_{ROH}). Samples from D'man breed were not considered in this analysis because of the reduced sample size (n=2). The distribution of genome-wide ROH is estimated by the mean number of ROH per individual (MN_{ROH}) and per breed (NT_{ROH}), the average length of ROH (AL_{ROH}) and the total number of ROH per

animal. The genomic inbreeding coefficient (F_{ROH}) was calculated for each breed using the following method (McQuillan *et al.,* 2008):

$$F_{ROH} = \frac{L_{ROH}}{L_{AUT}}$$

where L_{ROH} is the total length of all ROH per animal while L_{AUT} refers to the length of total autosomal SNP coverage.

Population genetic analysis

To better visualize the complex individual genetic relationship among sheep breeds, a network-based approach was performed implemented in NetView v.1.1 software (Neuditschko *et al.,* 2012; Steinig *et al.,* 2016) using identity by state (IBS) distance matrix to construct population networks using increasing K (i.e. parameter determining the number of mutual nearest neighbours) values up reaching an arbitrarily sufficient level of connectedness among samples from different breeds. The TreeMix (Pickrell and Pritchard, 2012) software uses allele frequencies at genome-wide polymorphisms and a Gaussian approximation of the genetic drift among populations to construct a Maximum Likelihood (ML) phylogeny connecting sampled populations by simple bifurcations. The graph-based method developed in TreeMix was applied to determine the directionality and quantify the extent of gene flow among species. To infer the history of population splits and admixtures, formal tests three-population (f3) and the four-population (f4) statistics, introduced in (Reich *et al.,* 2009).

Detection of common ROHs and putative genes under selection

Genomic regions subjected to selection represents reduced nucleotide diversity tending to generate ROH islands or hotspots, which have high levels of homozygosity around a selected locus compared with the rest of the genome (Purfield *et al.,* 2017). To identify

the genomic regions that were most commonly associated with ROH, the percentage (P) of the occurrences of a SNP in ROH was calculated by counting the number of times the SNP was detected in those ROH across individuals. The percentage values were plotted against the SNP position along the chromosomes. The SNPs showing a percentage higher than 10% were selected as an indication of a possible hotspot of ROH in the genome. The series of adjacent significant SNPs (P>10%) formed long genomic regions, called ROH islands. In order to annotate genes within the identified regions, we used the *Ovis aries* reference genome version 4.1 from the NCBI database. An extensive accurate literature search was conducted to investigate the biological function of annotated genes.

Results

ROH statistics and coefficient inbreeding

After filtering, the final number of animals and SNPs retained for analyses were 60 and 46,232 respectively (Supplementary Table S2). Analysis of ROH distribution for total breeds harbor 607 ROHs with a mean of 5.58 per individual where 89% among them were less than 5Mb (data not shown). The table 1 presented the descriptive statistics for runs of homozygosity in BAR, NDT and QFO breeds. The average length of ROH (AL_{ROH}) values showed low variation among sheep breeds, indicating that this value is not a good descriptor of ROH in Tunisian sheep population. The mean number of ROH per breed ranged from 7.20 (BAR) to 12.60 (NDT). Distribution of runs of homozygosity inbreeding coefficients F_{ROH} > 1 Mb for each sheep breed are reported in Supplementary Figure S1 and the mean values in Table 1. Indeed, the BAR breed showed the lowest value of F_{ROH}>1 Mb (0.017) in agreement with the MN_{ROH} results. However, the high standard deviation values within QFO (0.024±0.036) revealed high variability in autozygosity levels where F_{ROH} values ranged from zero to 0.13. The average genomic inbreeding coefficients

based on the difference between the observed vs. expected number of homozygous genotypes (F_{HOM}) were also estimated (Table 1). A high correlation was found between F_{HOM} and F_{ROH} within BAR and QFO breeds, r=0.974 with the most significant was mentioned for NDT breed (Pvalue<0.001).

Genomic relationship among breeds

A total of 60 sheep (BAR=25, NDT= 20, QFO=13 and DMN=2) were analyzed through a panel of 46,232 SNPs. In order to explore the genetic relationships among breeds, we used NetView software to visualize the genetic structure of the dataset along a gradient of k values (from 2 to 31), allowing for the investigation from fine to large scale population structure, revealed respectively, by small and large value of k (Figure 2). At fine scale, the topology of the network showed that most of samples fall in breed-specific clusters and at K=10, a first link among BAR and QFO appears, mediated by the probably admixed sample QF73. Thus, NetView from k= 2 to 30 (Figure 2) highlighted a clear differentiation between NDT samples and the remaining breeds. Moreover, notable exception was noticed in NetView at K=5, 10 and 20 , including four QFO samples (QF1, QF8, QF16, QF58) which are originally reared in a public farm institutionally engaged in QFO genetic conservation. The Treemix results at M = 1 (Supplementary Figure S2) no matter which breed out of the considered three (BAR, NDT, QFO) was selected for rooting the tree, a migration event was consistently detected between Barbarine and D'man, roughly corresponding to an edge's weight of 0.09. No evidence of gene flow was provided by both the f3 (Supplementary Table S3) and f4 (Supplementary Table S4) tests implemented in the TREEMIX computer package.

Genomic regions within runs of homozygosity

The incidence of each SNP in a run of homozygosity (ROH) among sheep breeds (Figure 3A, Table 2) revealed three ROH signals positioned on chromosome OAR1, OAR10 and OAR13 harbored by 84 significant SNPs (P>10%). The ROH length vary across positions in the genome ranging from 0.47 Mb in OAR 10 to 3.03 Mb in OAR 13. Indeed, the chromosome 13 had the largest extent of overlapping ROH positions among all sheep breeds spanning positions from 46.58 to 49.61 Mb encompassed the highest peak consisted of seven adjacent SNPs with occurrence in ROH of (P= 16.67 %). In order, to identify the characteristics of several genomic regions of potential autozygosity islands in each breed, the top 10% of SNPs was also plotted within each breed (Figure 3). The genomic distribution of ROH islands was clearly non-uniform among breeds and chromosomes. The chromosome position, number of SNPs, start and end position of ROH and number of annotated genes within the genomic regions of extended homozygosity were reported in Supplementary Table S5. The longest ROH island was observed in QFO breed on OAR 16 (42.16 Mb), while the shortest one was observed in NDT on OAR 13 (0.15Mb) where most of SNPs occurred in poor gene content regions. Interestingly, the within breed results show the highest ROH signal within Barbarine samples (mean frequency = 0.31) overlapping to the previously founded significant ROH peak across breeds (Figure 3A, Figure 3B). Indeed, this homozygous region starts at position 46.270 456 and ends at 49 619 573 pb with a total size of 3.3 Mb including 57 loci, out of them, 28 consecutives homozygous SNPs showing the highest SNP frequency (P=36%). Moreover, common ROHs hotspots between breeds were detected between QFO and BAR breeds on OAR 1 (51 027 822 – 54 156 664 pb) and between QFO and NDT on OAR 10 (30 908 858 – 34 551 330pb) separated into three ROH batches.

Discussion

Genetic diversity and genomic relationship among breeds

The study of intra-breed genetic diversity based on ROH abundance, length and inbreeding estimation highlighted that BAR appeared as the less consanguineous breed whereas the NDT breed suggest that recent inbreeding typically produces the observed ROH patterns. Moreover, the high correlation between F_{ROH} and F_{HOM} pointed out that F_{ROH} is a good alternative to estimate the IBD (identity by descent) autozygosity level in the absence of pedigree data, as may be the case for most local breeds in developing countries. Indeed, (Iniguez, 2005) reported that an increase of inbreeding in NDT breed is evidenced by cases of lamb genetic defect, due to the lack of scheme providing farmers with unrelated males. Moreover, the distribution of F_{ROH} for each breed reveal that QFO individuals show high variability in autozygosity level due probably to mating scheme adopted in their related farms. These results reflect in general the breed management practices that allow for uncontrolled mating of related individuals especially in smallholder farms where the exchange of rams among flocks is quite unusual.

The results of population structure among breeds overshadow two clear genetic patterns of local sheep breeds in Tunisia i) clear differentiation of NDT breed from the remaining breeds ii) high genetic closeness between BAR and QFO breeds. Indeed, the genetic distinction of NDT breed from local Tunisian breeds resulted from high European gene flow when the breed is created through crossbreeding between the French Merino d'Arles and QFO breed originated from Ouled Djellal sheep in Algeria (Porter *et al.,* 2016). In addition, the investigation of genomic variation between NDT breed and its parental breed

(QFO), using genotyping-by-sequencing data, was assumed as a consequence of selection implicated in tolerance to photosensitization by local toxic weed (Baazaoui *et al.,* 2020). In fact, the genetic closeness between BAR and QFO would not be surprising, if we consider the size reduction observed in Barbarine population in favor to QFO in recent decades (Sassi-Zaidy et al., 2014a) caused by massive cross-breeding between two breeds. In fact, the shift toward thin tail sheep breeds at the expense of the fat tail Barbarin was mainly due to the butchers' interests. Because of the difficulty in selling the fat of the carcass tail, butchers were reluctant to buy fat tail animals and farmers admit that butchers' preferences are influencing income because they are paying favorably thin tailed animals. This shift to thin tail breeds, if not controlled, will have a negative impact on the Barbarin breed, which is perfectly adaptated in a variety of production systems (Bedhiaf-Romdhani et al., 2008). Indeed, crossbreeding among autochthonous Northern African sheep breeds is a current practice (Othman *et al.,* 2016; Harkat *et al.,* 2017) in order to improve production performances and cope with economic pressure. Moreover, this results are consistent with those observed in Maghrebin breeds which evidenced a clear genetic homogenization among local sheep populations (Belabdi *et al.,* 2019), attesting for similar anthropological and demographic events of Northern African sheep ressources.

Detection of fat tail selection sweeps

The detection of ROH hotspots across and within sheep breeds consitently highlighted releavant genomic regions under positive selection shared between QFO and other breeds in OAR 1 and 10 and the most significant was exlusively detected in chromosome 13 highlighted in Barbarine breed, distinguished from other local breeds by its fat tail as

adaptness trait to harsh environment. Thus, in what follows, we will focus our discussion on annotated genes with putative selection sweeps involved in fat deposition in Tunisian sheep and specifically the tail type in Barbarine breed.

Gene annotation of genomic region within runs of homozygosity (ROHs) hotspots spanning 46.5–49.6 Mb on OAR 13 commonly shared with fat-tailed Barbarine samples (Table 2, Supplementary Table S5) has encompassed several known genes with specific relevance in lipid metabolic process elucidating the mechanism of tail fat deposition in sheep which provide insights into fat tail formation. It is evident that the fat tail phenotype is the result of the action of more than one gene. Among the candidate genes detected within this ROH islands were the *CDS2* gene that is involved in the phospholipid biosynthetic process is an important novel regulator of lipid storage and has a crucial role in mammalian lipid storage (Qi *et al.,* 2016). The prokineticin receptor1 (*PROKR1*) controls obesity through suppression of preadipocyte proliferation and differentiation in an animal model and humans (Yuan *et al.,* 2017). A recent study by (Pan *et al.,* 2019) suggested that the fixation of fat tail in domestic sheep is caused by a selective sweep near a retro-transposable hotspot, IBH region (47 993 040- 49 270 447 pb) on chromosome 13 affecting the expression of *BMP2* gene which is differentially expressed in fat-tailed in tissues. Moreover, the same study *LOC101117953* is a novel gene copy derived from a retro-transposable event, originated from protein phosphatase 1 catalytic subunit gamma (*PPP1CC*) gene, located also at ovine chromosome 13, which was evidenced as determinant of fat deposition and tail type differences in Chinese sheep (Wei *et al.,* 2015). Indeed, the *BMP2* gene played an important role in fat tail development in sheep breeds. It

has been detected in fat tailed Chinese sheep (Wei *et al.,* 2015; Yuan *et al.,* 2017), in Italian Laticauda breeds and in Mediterranean breeds including the Lybian Barbarine (Mastrangelo *et al.,* 2019). Therefore, the strong signal of selection is widespread in fat tailed sheep breeds from different geographic origins, suggesting that those breeds may share a common ancestor and the selection for tail fatness was occurred during the sheep domestication event.

Conclusion

The present SNP based molecular study shows that fat tailed Barbarine sheep is a perfectly adapted breed to harsh conditions in semi-arid area, is not affected by inbreeding practices. However, genomic structure results revealed that this purebred is threatened by a genetic erosion due to genetic homogenization with thin-tailed Queue fine de l'Ouest breed, resulting of crossbred animals with intermediate tail size. Interestingly, the genomic selection sweeps the potential genetic make-up of fat tail deposition which is considered as a valuable trait that should be conserved as a priority for the purpose to valorize this ancient irreversible heritage.

Acknowledgements

This work was supported by LABERPAR" Project (code 67), financed by the Apulia Region, Italy. We would like to thank the following members of AU-IBAR, Simplice Nouala, Mary Mbole-Kariuki and Edward Nengomasha for partial funding through the "Genetics" project. The authors are grateful to OEP members' Mohsen Ben Sassi, Mounir Hemdène and Zied Abdennebi who contributed samples used in this study. The mobility grant was

supported by GLOBAL-DOC Project (CUP H96J17000160002) funded by University of Bari,Italy.

Declaration of interest

The authors declare that they have no competing interests.

Ethics statement

Blood samples from animals were collected by veterinarians during routine blood sampling for medical care or follow up. All the samples in our study were obtained upon the breeder’s and breeding organizations' consent. Those animals were not linked to any experimental trials. The manuscript is not currently being considered for publication in another journal.

Software and data repository resources

The data generated in the current study are available from the corresponding author on reasonable request.

References

Baazaoui I, McEwan J, Anderson R, Brauning R, McCulloch A, Van Stijn T, Bedhiaf-Romdhani S 2020. GBS Data Identify Pigmentation-Specific Genes of Potential Role in Skin-Photosensitization in Two Tunisian Sheep Breeds. Animals 10, 5.

Bedhiaf-Romdhani S, Djemali M, Zaklouta M, Iniguez L 2008. Monitoring crossbreeding trends in native Tunisian sheep breeds. Small ruminant research 74, 274-278.

Belabdi I, Ouhrouch A, Lafri M, Gaouar SBS, Ciani E, Benali AR, Ouelhadj HO, Haddioui A, Pompanon F, Blanquet V 2019. Genetic homogenization of indigenous sheep breeds in Northwest Africa. Scientific reports 9, 7920.

FAO 2015. The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture, edited by B.D. Scherf & D. Pilling. FAO Commission on Genetic Resources for Food and Agriculture Assessments. Rome.

Harkat S, Laoun A, Belabdi I, Benali R, Outayeb D, Payet-Duprat N, Blanquet V, Lafri M, Da Silva A 2017. Assessing patterns of genetic admixture between sheep breeds: Case study in Algeria. Ecology and evolution 7, 6404-6412.

Hentati H, Hamouda MB, Chriki A 2012. Genetic diversity of two Tunisian sheep breeds using random amplified polymorphic DNA (RAPD) analysis. African Journal of Biotechnology 11,

Iniguez L 2005 Characterization of small ruminant breeds in West Asia and North Africa, Vol 1. ICARDA Aleppo.

Kdidi S, Calvo J, González-Calvo L, Sassi MB, Khorchani T, Yahyaoui M 2015. Genetic relationship and admixture in four Tunisian sheep breeds revealed by microsatellite markers. Small Ruminant Research 131, 64-69.

Khaldi Z, Rekik B, Haddad B, Zourgui L, Souid S 2010. Genetic characterization of three ovine breeds in Tunisia using randomly amplified polymorphic DNA markers. Livest. Res. Rural Dev 22.

Mastrangelo S, Bahbahani H, Moioli B, Ahbara A, Al Abri M, Almathen F, Da Silva A, Belabdi I, Portolano B, Mwacharo JM 2019. Novel and known signals of selection for fat deposition in domestic sheep breeds from Africa and Eurasia. PloS one 14, e0209632.

Mastrangelo S, Tolone M, Sardina MT, Sottile G, Sutera AM, Di Gerlando R, Portolano B 2017. Genome-wide scan for runs of homozygosity identifies potential candidate genes associated with local adaptation in Valle del Belice sheep. Genetics Selection Evolution 49, 84.

McQuillan R, Leutenegger A-L, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A 2008. Runs of homozygosity in European populations. The American Journal of Human Genetics 83, 359-372.

Neuditschko M, Khatkar MS, Raadsma HW 2012. NetView: a high-definition network-visualization approach to detect fine-scale population structures from genome-wide patterns of variation. PloS one 7, e48375.

Othman OEM, Payet-Duprat N, Harkat S, Laoun A, Maftah A, Lafri M, Da Silva A 2016. Sheep diversity of five Egyptian breeds: Genetic proximity revealed between desert breeds: Local sheep breeds diversity in Egypt. Small Ruminant Research 144, 346-352.

Pan Z, Li S, Liu Q, Wang Z, Zhou Z, Di R, An X, Miao B, Wang X, Hu W 2019. Rapid evolution of a retro-transposable hotspot of ovine genome underlies the alteration of *BMP2* expression and development of fat tails. BMC genomics 20, 261.

Peripolli E, Stafuzza NB, Munari DP, Lima ALF, Irgang R, Machado MA, do Carmo Panetto JC, Ventura RV, Baldi F, da Silva MVGB 2018. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (Bos indicus) dairy cattle. BMC genomics 19, 34.

Pickrell JK, Pritchard JK 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS genetics 8, e1002967.

Porter V, Alderson L, Hall SJ, Sponenberg DP 2016 .Mason's World Encyclopedia of Livestock Breeds and Breeding, 2 Volume Pack. Cabi.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics 81, 559-575.

Purfield DC, McParland S, Wall E, Berry DP 2017. The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. PLoS One 12, e0176780.

Qi Y, Kapterian TS, Du X, Ma Q, Fei W, Zhang Y, Huang X, Dawes IW, Yang H 2016. CDP-diacylglycerol synthases regulate the growth of lipid droplets and adipocyte development. Journal of lipid research 57, 767-780.

Reich D, Thangaraj K, Patterson N, Price AL, Singh L 2009. Reconstructing Indian population history. Nature 461, 489.

Sambrook J, Fritsch EF, Maniatis T 1989 Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.

Sassi-Zaidy YB, Maretto F, Zanetti E, Hajji G, Charfi-Cheikrouha F, Cassandro M 2014a. Genetic structure and variability within and among populations of the fat-tailed Barbarine sheep breed using microsatellites markers. African Journal of Biotechnology 13

Sassi-Zaidy YB, Maretto F, Charfi-Cheikrouha F, Cassandro M 2014b. Genetic diversity, structure, and breed relationships in Tunisian sheep. Small Ruminant Research 119, 52-56.

Steinig EJ, Neuditschko M, Khatkar MS, Raadsma HW, Zenger KR 2016. netview p: a network visualization tool to unravel complex population structure using genome-wide SNP s. Molecular Ecology Resources 16, 216-227.

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

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

Table 1. *Descriptive statistics for runs of homozygosity (ROH) and inbreeding coefficient in the Tunisian sheep breeds.*

Breed	NT_{ROH}[1]	MN_{ROH}[2]	AL_{ROH}[3]	$F_{ROH>1Mb}$[4]	F_{HOM}[5]	$r(F_{ROH}- F_{HOM})$[6]
BAR	180	7.20	5.87	0.017±0.028	0.027±0.034	0.974
NDT	252	12.60	4.31	0.021±0.009	0.005±0.013	0.793***
QFO	128	9.85	6.58	0.024±0.036	0.033±0.042	0.974

[1] Total number of ROH per breed

[2] Mean number of ROH per individual

[3] Average length of ROH in Mb;

[4] Mean ROH-based inbreeding coefficient with standard deviation

[5] Inbreeding coefficient based on the difference between observed *vs.* expected number of homozygous genotypes

[6] Correlation between F_{ROH} and F_{HOM}

*** P-value<0.001

Table 2. *Details of Runs of homozygosity (ROH) hotspots detected across Tunisian sheep breeds*

OAR	Mean percent[1]	SNPs[2]	ROH length (Kb)	Positions (pb)	Genes	Genes in ROH hotspot
1	11.67	25	1185	49,985,611 - 51,170,678	8	*LOC101120030, ERICH3, TRNAW-CCA,CRYZ, TYW3, LHX8, SLC44A5, LOC105611835*
10	11.11	9	473	32,442,135 - 32,915,074	10	*LOC101113604,LOC101114283, LOC101114533, LNX2, MTIF3, GTF3A, RASL11A, RPL21, LOC105609570, USP21*
13*	14.77	50	3037	46,582,744 - 49,619,573	20	*CDS2, LOC101109379, PROKR2, GPCPD1, LOC101109635, C13H20orf196, TRNAF-GAA, CHGB, TRMT6, MCM8, LOC105609936, CRLS1, LRRN4, FERMT1, LOC106991507, LOC101117437, BMP2, LOC101117953, LOC101118207, LOC101110166*

[1] Mean percent of significant SNPs (P>10%) of occurrence in a ROH region,

[2] Number of loci identified within ROH hotspots.

*chromosome 13 include the highest SNPs frequency in ROH region

Figure captions

Figure 1 Sampling sites and sheep breeds

Figure 2 Mutual nearest-neighbour graphs obtained from NetView considering the following k values: $k = 2$, $k = 5$, $k = 10$, $k = 20$, k=30 and k=31 in Tunisian sheep breeds. Color shades code for different sheep breeds. BAR= Barbarine, QFO = Queue Fine de l'ouest; NDT= Noire de Thibar; DMN=D'man.

Figure 3 Genome-wide frequency of single nucleotide polymorphisms (SNPs) occurrence into runs of homozygosity (ROH) (A) across and within (B) Barbarine (C) Noire de Thibar) and (D) Queue fine breeds sheep breeds.