Volume 61 (3). Published January, 01, 2019
www.jnsciences.org
E-ISSN 2286-5314

Genomic inbreeding and population structure in rams of Tunisian D'man sheep

S. BEN JEMAA[1*], S. KDIDI[2], M. BOUSSAHA[3], E. REBOURS[3], M. H. YAHYAOUI[2]

[1] Laboratory of animal and forage production, National Institute of Agronomic Research of Tunisia, University of carthage, Rue Hédi Karray, Ariana, 2049, Tunisia.
[2] Livestock and Wildlife Laboratory, Arid Lands Institute, Route Djorf Km 22, 4119 Medenine, Tunisia
[3] GABI, INRA, AgroParisTech, University of Paris Saclay, 78350 Jouy-en-Josas, France.

*Corresponding author: benjemaaslim@gmail.com

Abstract – D'man sheep is a minority breed in Tunisia that was established nearly two decades ago from a small Moroccan flock. Tunisian D'man is an isolated population which has experienced no gene flow since its introduction. Currently, this population is suffering from an increased incidence of stillbirths and recurrent abortions caused by inbreeding. The availability of dense SNP markers has facilitated the quantification of genomic inbreeding in farm animals. The aim of the present study was to estimate the kinship between the six D'man rams of a Tunisian breeding station and to assess their genomic structure using the Illumina OvineSNP50 BeadChip and comparisons with the French Lacaune and the Tunisian west thin tail breeds. We found moderate to high levels of inbreeding between the old rams ranging from 0.0069 to 0.1202 and low level of genetic diversity (expected heterozygosity ~ 0.26). Substantial level of west thin tail introgression was detected in the young D'man rams although these were supposed to be purebred. Analysis of runs of homozygosity (ROHs) showed that all individuals had at least one ROH > 10 Mb. ROH islands identified within the D'man rams harbored genes whose mutation leads to stillbirth, dystocia, embryo quality reduction and fleece phenotype variation. Our study demonstrates the usefulness of molecular markers in the management of inbreeding in small isolated populations and would be very helpful in the implementation of planned mating scheme based on sire genotypes.

Keywords: D'man sheep, Inbreeding, Single nucleotide polymorphism, Runs of homozygosity.

1. Introduction

The D'man sheep breed originated in the southern part of the Maghreb region (southwestern Algeria and southeastern Morocco). This breed was introduced for the first time in Tunisia in 1994 from a Moroccan founder flock composed of 200 ewes and 25 rams (Rekik *et al.* 2005). D'man population expanded widely in the country (particularly in the Tunisian oases) since its earliest introduction as it was appreciated by the breeders for its high prolificacy. A recent study showed that the high prolificacy of D'man is mainly due to the segregation of a major prolificacy locus, named FecL within the Tunisian D'man population (Ben Jemaa *et al*. 2019). This locus was originally shown to have a major effect on the prolificacy of the French Lacaune sheep and to be in complete linkage disequilibrium with an single nucleotide polymorphism (SNP) located in the intron 7 of B4GALNT2 gene (OAR11:36938224T>A) (Drouilhet *et al.* 2013). The estimated mutated allele effect of the $FecL^L$ locus on litter size was +0.4 to +0.5 lambs per lambing (Martin *et al.* 2014). Currently, the Tunisian D'man breed can be considered as an island population as it is isolated and it has a small size. Indeed, for sanitary reasons, Tunisian D'man has experienced no gene flow from neighboring countries since its establishment nearly two decades ago thus leading to an increase in the level of inbreeding within this population. Consequently, a frequent onset of lethal genetic disorders in new born lambs was observed within many farms. The inbreeding problem is made worse by the absence of a National pedigree and performance recording system allowing the use of optimal mating strategies. However, the problem is less marked within the state flock of the breeding station of Chenchou which belongs to the Livestock and Pasture Office of Tunisia (OEP) which possesses a pedigree recording system allowing a planned mating scheme based on sire rotation across the seven family lines forming the flock. Chenchou breeding station is the leading supplier of rams for most of the small farmers of Southern Tunisia.

The development of high-throughput SNP genotyping platforms coupled with the gradual reduction of genotyping costs makes it possible to accurately estimate kinship between individuals, thus helping control inbreeding in isolated populations. Furthermore, the availability of dense SNP genotypes allows the identification of potential continuous stretches of homozygous SNP genotypes known as runs of homozygosity (ROH). Characterization of ROHs, particularly their size and frequency, provides information about relatedness within a population (Mastrangelo et al., 2018).

In the present study we aimed to estimate the genomic kinship matrix in the six D'man rams of the Chenchou station and to assess genome-wide inbreeding ROH distribution within these genitors using the Illumina OvineSNP50 BeadChip. Furthermore, we provided, for the first time, a detailed assessment of the genetic structure of the D'man breed using comparisons with the Tunisian West thin tail (WTT) sheep and the French Lacaune (LAC) breed. The choice of these two breeds is prompted by the geographic proximity (for D'man and WTT) and the common presence of the prolific allele *FecLL* (for D'man and LAC).

2. Materials et Methods

2.1. Blood sampling, DNA extraction and genotyping

Jugular vein blood samples (5 ml per animal) were obtained from all the six rams of Chenchou breeding station under the Tunisian Veterinary Authorities' rules. Blood samples were collected in EDTA Vacutainer tubes. D'man DNA was extracted using salting out method as described by Miller *et al.* (1989). DNA samples belonging to the six D'man rams and to 17 WTT individuals (available from Kdidi *et al.* 2015) were then genotyped on the OvineSNP50 Genotyping BeadChip (Illumina, San Diego, CA, USA) using standard procedures (http://www.illumina.com) resulting in 54242 genotyped SNPs. We also included genotyping data of 25 Lacaune individuals available from the International Sheep Genomics Consortium (Kijas *et al*. 2012).

2.2. SNP quality control and marker selection

Among the 54242 genotyped SNPs, we used the 49034 ones that were in common between the LAC genotypes on one hand and the D'man and WTT genotypes on the other hand. We then used PLINK ver.1.09 (Purcell *et al*. 2007) for genotyping data quality control. Samples genotyped for less than 90 % of markers were excluded from the analysis. SNPs genotyped for less than 90 % of the animals and those with MAF less than 0.01 were also discarded. Using these criteria, we excluded one WTT individual due to a low genotyping rate (<90 %). Similarly, 253 SNPs were excluded because of low genotyping rates, and 251 SNPs were also excluded as they were monomorphic (MAF <0.01). An exact test for Hardy-Weinberg Equilibrium (HWE) was then carried within each population separately on the remaining SNPs, which led to 47038 SNPs retained for further analysis.

2.3. Kinship estimation and characterization of runs of homozygosity

Genomic kinship coefficient for each pair of D'man individuals was estimated using Loiselle estimator (LS) implemented in the R package RClone (Bailleul *et al*. 2016). Loiselle estimator was selected because it uses a correction for small sample sizes, thus making him more accurate than most of the other estimators.

Because kinship coefficients are better estimated if the markers used are not in linkage disequilibrium (LD) with each other, a subset of SNPs in approximate linkage equilibrium was selected by LD-based SNP pruning in PLINK.

ROHs were identified in sliding windows of 20 SNPs using PLINK. No more than five missing calls and one heterozygous SNP were allowed in each window. The minimum length of a ROH segment was set to 1 Mb. A ROH was declared if it contained at least 20 SNPs. The minimum required SNP density was one SNP per 200 kb and the maximum gap allowed between any two consecutive SNPs was 1000 kb. ROHs were classified into six categories (1 to 2, 2 to 3, 3 to 4, 4 to 8, 8 to 10, 10 to 20 Mb).

The --homozyg-group function implemented in PLINK was used to assess ROH islands shared among D'man individuals. In the present study, ROH islands were defined as the homozygous segments shared by at least 50% of individuals. Furthermore, gene content within the identified ROH islands was extracted using BioMart tool on Ensembl. Functional enrichment analysis was then performed on the list of genes overlapping ROH islands using Database for Annotation, Visualization and Integrated Discovery (DAVID) software by comparison to the sheep genome background supplied by DAVID.

2.4. Population structure and genetic diversity

We provided a fine-scale assessment of the genetic structure of the Dman, West thin tail and Lacaune individuals using two methods. First, we used Discriminant Analysis of Principal Components (DAPC) implemented in the R package adegenet (Jombart, 2008) without providing any a priori assignment of the individuals to their population of origin. In this analysis, the optimal number of genetic clusters that best describes the data was identified by running a clustering algorithm called k-means and comparing the different clustering solutions using the Bayesian Information Criterion (BIC). Individuals were then assigned to these clusters.

Second, we used ADMIXTURE 1.23 software (Alexander *et al.* 2009) to estimate proportions of ancestry from each contributing genetic cluster in the three populations. DISTRUCT software (Rosenberg, 2004) was then used to graphically display ancestry within each individual.

Finally, we used expected (Hs) and observed heterozygosities (Ho) to assess the genetic diversity of each of the three populations.

3. Results

3.1. Genomic relatedness and characterization of runs of homozygosity

Genomic pairwise kinship matrix between the six D'man rams was estimated using 3530 SNPs that are almost in linkage equilibrium with each other (Table 1). High positive kinship coefficients (> 0.11) were observed between ram DMAN1 on one hand and rams DMAN4 and DMAN6 on the other hand indicating that these individuals share more DNA identical by state than the rest of rams. Moderate kinship estimate was found between rams DMAN2, DMAN3 and DMAN5 (ranging from 0.011 to 0.0431). Importantly, negative kinship values were obtained between rams DMAN1, DMAN4 and DMAN6 on one hand and rams DMAN2, DMAN3 and DMAN5 on the other hand which indicates that these two groups are unrelated.

Table 1. Estimated genomic pairwise kinship matrix for the six Dman rams.

	DMAN1	DMAN2	DMAN3	DMAN4	DMAN5	DMAN6
DMAN1						
DMAN2	-0.0289					
DMAN3	-0.0412	0.0114				
DMAN4	0.1122	-0.0098	-0.0107			
DMAN5	-0.0366	0.0175	0.0432	-0.0061		
DMAN6	0.1202	-0.0189	-0.0046	0.0069	-0.0170	

Using our ROH definition (see Material and Methods), a total of 278 ROHs were identified in the six D'man rams. The number of identified ROHs per individual ranged between 36 (DMAN4) to 59 (DMAN 3). Total sum of ROHs per individual ranged between 71.28 Mb (DMAN4) to 224.22 Mb (DMAN3) (Table 2). A high positive correlation (0.98) was observed between the total sum of ROHs and the number of long ROHs (> 8 Mb). In all individuals, the highest number of ROHs was found within the [1-2 Mb] category. All individuals presented at least one ROH > 10 Mb. The number of ROHs per chromosome ranged between one and six (Figure 1). A high correlation was found between the number of ROHs and chromosome length (correlation = 0.90). Approximately 5.5%, on average, of the entire genome of the D'man individuals was covered by ROH segments. The highest coverage was shown by OAR24 (24.51%) and OAR11 (10.32%) (Figure 1).

Ten ROH islands were identified within the D'man rams located on OAR1 (2), OAR3 (2), OAR4 (2), OAR9 (1), OAR17 (1), OAR19 (1) and OAR25 (1). These ROH islands included 375 genes.

Two significant (enrichment score (ES) > 1.3) functional clusters were detected in the annotation clustering analysis (Table S1). The one with the highest enrichment score (ES=2.9) was related to Calcium-activated chloride channel protein and included chloride channel accessory (CLCA) genes. Proteins of the CLCA family may contribute to parasite expulsion through mucus hydration across the gut epithelium and smooth muscle contraction (Rowe *et al.* 2009).

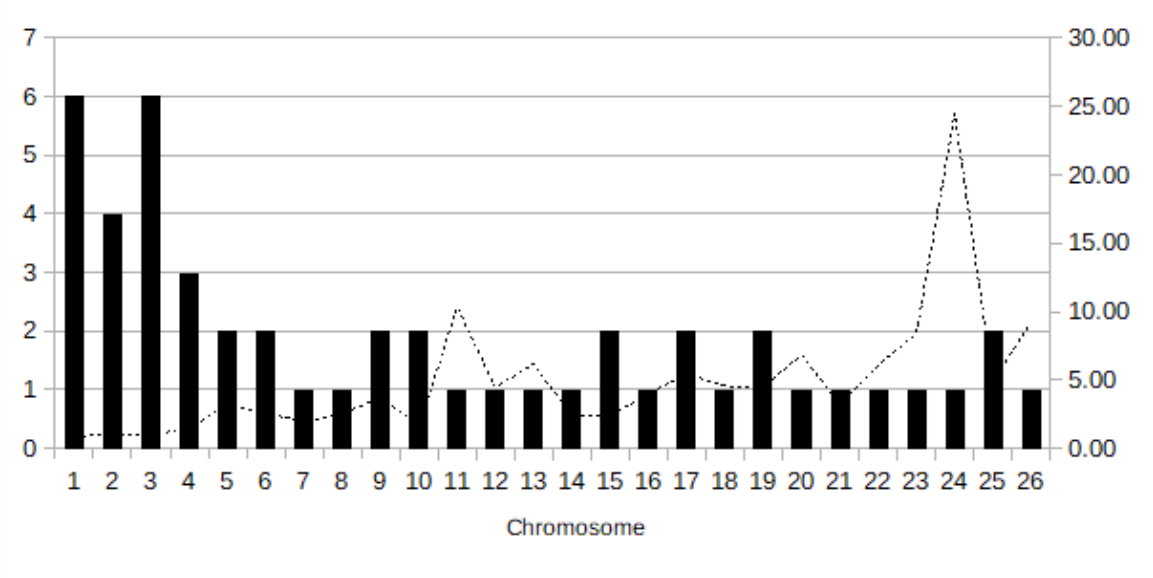

Figure 1. The number of runs of homozygosity (ROHs) per chromosome (bars) identified in the 6 D'man rams and the per-animal mean percentage of the chromosome covered by ROHs (line plot).

The second functional cluster was related to arylesterase activity (ES=2.5) and included Paraoxonase 1 (PON1), Paraoxonase 2 (PON2) and Paraoxonase 3 (PON3) genes. PON1 enzymes are present in ovarian follicular fluid. Mutations in PON1 affect embryo quality among women undergoing in vitro fertilization (Mackness and Durrington 1995).

Table 2. Total number and total length of detected ROHs and number of ROHs per individual for each of the six ROH categories.

Individual	Total number	Total length (Mb)	[1,2 Mb]	(2,3 Mb]	(3,4 Mb]	(4,8 Mb]	(8,10 Mb]	(10,20 Mb]
DMAN1	42	103.16	31	3	1	4	1	2
DMAN2	50	143.01	36	3	4	3	1	3
DMAN3	59	224.22	33	6	3	8	3	6
DMAN4	36	71.28	27	6	1	1	0	1
DMAN5	43	135.59	25	6	1	7	2	2
DMAN6	48	159.46	30	5	2	5	1	5

3.2. Genetic diversity and population structure

Mean observed heterozygosity was similar between D'man, Lacaune and WTT breeds (0.367, 0.368 and 0.375, respectively). In contrast, mean expected heterozygosity was clearly lower for D'man compared to Lacaune and WTT (0.399 Vs 0.492 and 0.494, respectively).

ADMIXTURE results for ancestral populations K=2 to K=4, for D'man, WTT and LAC, are shown in Figure 2. Assuming two ancestral populations (K = 2), D'man and WTT shared substantial level of ancestry (on average, 95 % of the D'man genome was shared with 75 % of WTT genome) while only 10 % genome was common between D'man and LAC. For K= 3, Lacaune was subdvided into two distinct genetic clusters while for K =4, half of D'man rams showed substantial signals of WTT genome introgression. These trends were confirmed by the DAPC (Figure 3). Indeed, Lacaune individuals were assigned to two different genetic clusters (1 and 3) while the three D'man individuals that showed signals of WTT introgression were assigned to the same genetic cluster as WTT (cluster 4). The three other individuals were assigned to cluster 2.

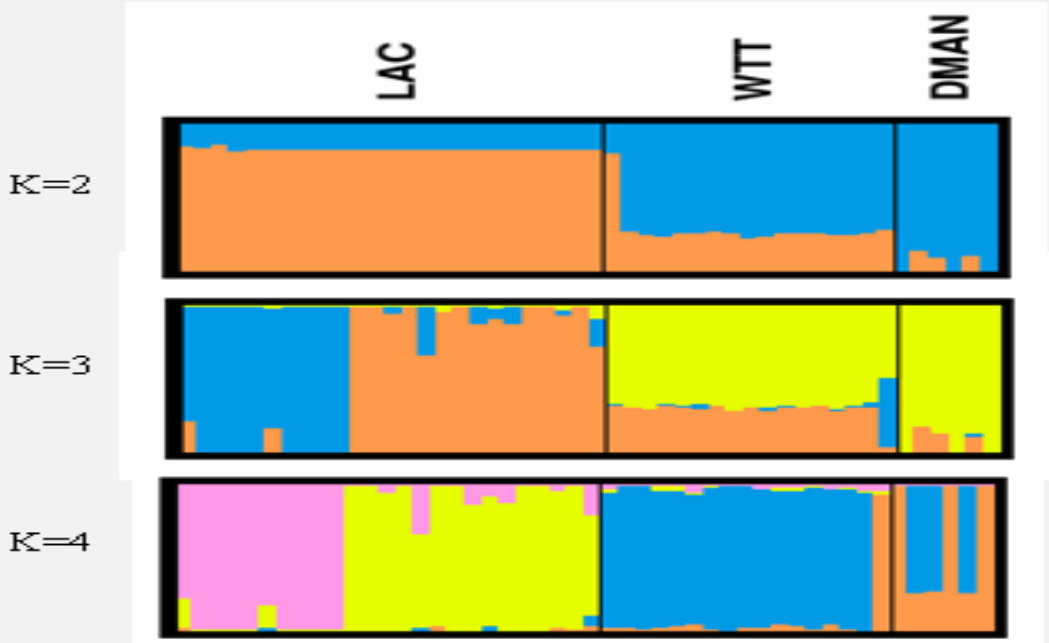

Figure 2. Unsupervised hierarchical clustering of the 48 individuals from the Lacaune (LAC), West thin tail (WTT) and D'man (DMAN) breeds. Results for k (number of clusters) = 2, k=3, k=4 are shown.

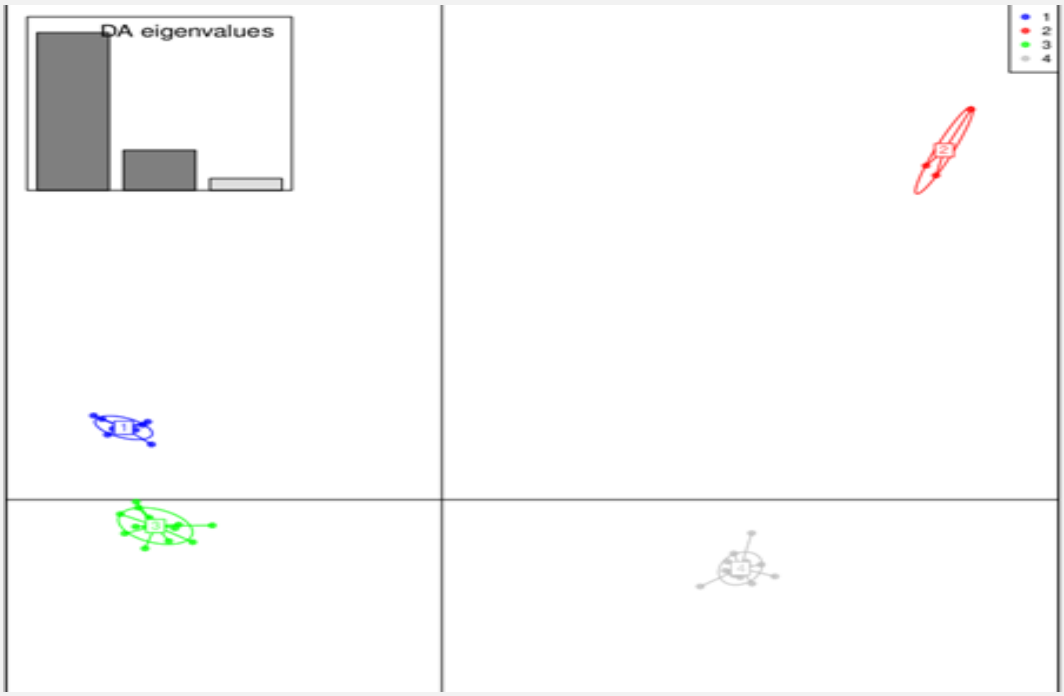

Figure 3. Scatter plot of the DAPC results where individuals were assigned to genetic clusters without providing any a priori population assignment. Cluster 1 & 3 : Lacaune individuals, Cluster 2 : the three old D'man rams, Cluster 4 : the West thin Tail individuals & the three young D'man rams.

4. Discussion

The main aim of the present study was to assess the genomic inbreeding and the population structure of the D'man rams in the breeding state farm of Chenchou which is the leading supplier of rams for most of the farmers in Southern Tunisia. The importance of this study stems from the inability to introduce new individuals or use ram semen from Morocco to reduce inbreeding in the Tunisian flocks (owing to sanitary reasons). Therefore, internal genetic management would be the only way to tackle the undesirable consequences of the increasing inbreeding levels in Tunisian D'man. Despite the possibility to estimate kinship based on pedigree information, relatedness between individuals is more accurately

estimated using molecular information. Indeed, variation of kinship due to Mendelian sampling is better captured using genomic information than the pedigree-based method. In the present study, we purposely used Loiselle estimator because it is less biased and more accurate than most of the other estimators when the sample is small (Wang 2017). The kinship matrix estimated from unlinked SNP markers showed that ram DMAN1 is closely related to rams DMAN4 and DMAN6 which implies that the former should not be mated with ewes from families 4 and 6. By contrast, rams DMAN2, DMAN3 and DMAN5 (which were younger than the three other rams) showed low kinship values (<0.043). Surprisingly, these three rams were shown to be unrelated with the first ones (i.e rams DMAN1, DMAN4 and DMAN6) although all the six rams are supposed to come from the same founder flock. This unexpected result was explained by the DAPC and admixture analysis showing significant levels of WTT introgression into the genome of these three rams (up to 75% according to ADMIXTURE). It is more likely that in order to tackle the growing rate of genetic problems arising from consanguinous matings, the managers of Chenchou station frequently practiced crosses with other local breeds. When we considered only non introgressed rams, expected heterozygosity dropped from 0.39 to only 0.26 thus becoming significantly lower than observed heterozygosity. This difference is mainly due to the high inbreeding rate between the three rams.

In the present study, we used runs of homozygosity to assess autozygosity for the six D'man rams. The mean number of identified ROH and mean total sum of ROHs per individual were 46 ± 8 and 139 ± 52 Mb, respectively. Using a similar ROH definition, Michailidou *et al.* (2018) found significantly lower values in the greek Chios (n=11.88), Boutsko (n=11.59) and Karagouniko (n=3.1) breeds. Total sum of ROHs per individual found by these authors did not exceeded 60 Mb (Michailidou *et al.* 2018). Similarly, Mastrangelo *et al.* (2017) found that mean number of ROH per individual in the Italian Barbaresca breed was 15.9 which is also by far lower than the average of D'man individuals. Among the six rams, DMAN3 and DMAN6 showed the highest number of long (> 8 Mb) ROHs (9 and 6 ROHs, respectively). Long autozygous ROH segments have long been implicated as a cause of the higher prevalence of genetic disorders stemming from the accumulation of strongly deleterious mutations as well as mildly deleterious variants (Szpiech *et al.* 2013). Under the assumption that the expected length of an autozygous segment follows an exponential distribution with mean equal to 1/2 g Morgans, where g is the number of generations since the common ancestor (Howrigan *et al.* 2011), we can hypothesize that the autozygous segments of these two rams originated from a common ancestor 6-7 generations ago (assuming that 1 cM ~ 1 Mb). Assuming that the generation interval is 2.5 years, this common ancestor lived about 15-17.5 years ago. Our results confirm the trend that ROHs occur more frequently on large chromosomes. The longest ROH was found on OAR03 (19 Mb), however, OAR24 and OAR11 showed the highest proportion in coverage by ROH segments. On the other hand, only DMAN6 showed a ROH on OAR24 while all but ram DMAN5 showed a ROH on OAR11. Among these, rams DMAN1, DMAN2 and DMAN6 showed a ROH beginning at 41.38, 46.66 and 43.37 Mb, respectively. It is worth noting that the major prolificacy locus FecL which was shown to have a significant effect on D'man litter size (Ben Jemaa *et al.* 2019) is located on the OAR11 (36.929322 Mb -36.992982 Mb). Both observations lead us to hypothesize that these ROHs resulted from a positive selection surrounding the FecL locus. Indeed, Ben Jemaa *et al.* (2019) found that the prolific allele, named $FecL^{L}$ exhibited a high frequency in the D'man population (0.65) owing to a decades-old breeding strategy based on the selection of ewe lambs born from large litter size which are most likely born from either L+ or LL parents. Such positive selection could have generated long ROH surrounding the FecL locus.

Among the ten identified ROH islands in the D'man rams, the one located on OAR4 (8.363 Mb-13.125 Mb) is particularly interesting. This ROH island included PON1, PON2 and PON3 genes. It was hypothesized that PON genes are involved in follicle maturation as it was observed that an increase in the activity of these genes is positively correlated with follicle size in humans (Meijide *et al.* 2017). Interestingly, only non introgressed D'man rams (DMAN1, DMAN4 and DMAN6) showed patterns of autozygosity islands at this position. Therefore, we can hypothesize that this ROH island may have occurred due to selection for traits related to D'man prolificacy. Furthermore, by examining individually the list of genes located in the ten ROH islands, we identified two interesting genes located on OAR25 within the [4 Mb-8.609 Mb] interval. The first one is Interferon Regulatory Factor 2 Binding Protein gene (IRF2BP2). This gene is involved in fleece phenotype in sheep (Demars *et al.* 2017). The second interesting gene is the Beta-1,3-N-Acetylgalactosaminyltransferase 2 (B3GALNT2). Mutations in this gene were shown to be associated with stillbirth and dystocia in horse (Ducro *et al.* 2015).

Supplementary Table S1. ROH islands within the D'man individuals. Gene content within these ROHs is also reported.

BTA	ROH position (Pb)	Length (Mb)	% of animals	Gene content (Stable ID)	Gene content (Gene name)
1	60404477-63386591	2.982	50	ENSOARG00000013816, ENSOARG00000013816, ENSOARG00000013828, ENSOARG00000013841, ENSOARG00000013841, ENSOARG00000013854, ENSOARG00000013867, ENSOARG00000013925, ENSOARG00000013942, ENSOARG00000013959, ENSOARG00000013977, ENSOARG00000014058, ENSOARG00000014086, ENSOARG00000014086, ENSOARG00000014101, ENSOARG00000014238, ENSOARG00000014263, ENSOARG00000014278, ENSOARG00000014293, ENSOARG00000014319, ENSOARG00000014336, ENSOARG00000014350, ENSOARG00000014461, ENSOARG00000014549, ENSOARG00000014621, ENSOARG00000014659, ENSOARG00000014659, ENSOARG00000014794, ENSOARG00000014819, ENSOARG00000014843, ENSOARG00000014861, ENSOARG00000014861, ENSOARG00000014881, ENSOARG00000014893	PRKACB, PRKACB, DNASE2B, RPF1, SPATA1, CTBS, SSX2IP, LPAR3, MCOLN2, MCOLN2, MCOLN3, WDR63, SYDE2, C1orf52, BCL10, DDAH1, CYR61, ZNHIT6, COL24A1, ODF2L, CLCA2, CLCA1, CLCA1, CLCA4, SH3GLB1, SH3GLB1, SELENOF
1	130774107-132661094	1.886	83	ENSOART00000016806, ENSOART00000016820	
3	182242247-184334066	2.091	50	ENSOART00000021103, ENSOART00000021104, ENSOART00000021117, ENSOART00000021128, ENSOART00000021129, ENSOART00000021145, ENSOART00000022017, ENSOART00000021178, ENSOART00000021179, ENSOART00000021217, ENSOART00000021241, ENSOART00000021275, ENSOART00000021302	BICD1, KIAA1551, AMN1, AMN1, DENND5B, DENND5B, SINHCAF, CAPRIN2, IPO8
3	213086338-221471895	8.385	50	ENSOARG00000013505, ENSOARG00000013519, ENSOARG00000013607, ENSOARG00000013724, ENSOARG00000013739, ENSOARG00000013750, ENSOARG00000013761, ENSOARG00000013836, ENSOARG00000013931, ENSOARG00000013947, ENSOARG00000013956, ENSOARG00000013969, ENSOARG00000014071, ENSOARG00000020372, ENSOARG00000014198, ENSOARG00000014264, ENSOARG00000014355, ENSOARG00000014439, ENSOARG00000014561, ENSOARG00000014681, ENSOARG00000014692, ENSOARG00000014774, ENSOARG00000014774, ENSOARG00000014944, ENSOARG00000015015, ENSOARG00000015099, ENSOARG00000015125, ENSOARG00000015125, ENSOARG00000015236, ENSOARG00000015251, ENSOARG00000015331, ENSOARG00000015331, ENSOARG00000015413, ENSOARG00000015489, ENSOARG00000015590, ENSOARG00000015674, ENSOARG00000015734, ENSOARG00000015745, ENSOARG00000015813, ENSOARG00000015877, ENSOARG00000015976, ENSOARG00000016073, ENSOARG00000016145, ENSOARG00000016213, ENSOARG00000016290, ENSOARG00000016400, ENSOARG00000016412, ENSOARG00000016495, ENSOARG00000016513, ENSOARG00000016603, ENSOARG00000016677, ENSOARG00000020375, ENSOARG00000016693, ENSOARG00000016693, ENSOARG00000016775, ENSOARG00000016785, ENSOARG00000016884, ENSOARG00000017027, ENSOARG00000017108, ENSOARG00000017179, ENSOARG00000017238, ENSOARG00000017303, ENSOARG00000017466, ENSOARG00000017567, ENSOARG00000020382, ENSOARG00000020388, ENSOARG00000017634, ENSOARG00000017728, ENSOARG00000017792, ENSOARG00000017843, ENSOARG00000017943, ENSOARG00000018085, ENSOARG00000018130, ENSOARG00000018140, ENSOARG00000018228, ENSOARG00000018238, ENSOARG00000025175, ENSOARG00000018246, ENSOARG00000018307, ENSOARG00000018319, ENSOARG00000018326, ENSOARG00000018385, ENSOARG00000018392, ENSOARG00000018400, ENSOARG00000018451, ENSOARG00000018496, ENSOARG00000020395, ENSOARG00000020400,	BCL2L13, MICAL3, PEX26, TUBA8, CDC42EP1, LGALS2, GGA1, PDXP, LGALS1, TRIOBP, H1F0, GCAT, ANKRD54, EIF3L, MICALL1, C22orf23, POLR2F, SOX10, SOX10, PICK1, SLC16A8, BAIAP2L2, PLA2G6, PLA2G6, MAFF, TMEM184B, KDELR3, DDX17, DMC1, FAM227A, CBY1, TOMM22, JOSD1, GTPBP1, SUN2, DNAL4, APOBEC3Z1, APOBEC3F, CBX7, PDGFB, TAB1, MGAT3, MIEF1, MIEF1, ATF4, RPS19BP1, CACNA1I, ENTHD1, GRAP2, FAM83F, TNRC6B, ADSL, SGSM3, MRTFA, MCHR1, SLC25A17, ST13, XPNPEP3, RBX1, EP300, L3MBTL2, RANGAP1, ZC3H7B, TEF, TOB2, PHF5A,

				ENSOARG00000018558, ENSOARG00000018655, ENSOARG00000018717, ENSOARG00000018753, ENSOARG00000018792, ENSOARG00000020407, ENSOARG00000018800, ENSOARG00000018806, ENSOARG00000018845, ENSOARG00000018850, ENSOARG00000018854, ENSOARG00000018883, ENSOARG00000018936, ENSOARG00000018945, ENSOARG00000018951, ENSOARG00000018955, ENSOARG00000018955, ENSOARG00000018983, ENSOARG00000018986, ENSOARG00000018993, ENSOARG00000019024, ENSOARG00000019031, ENSOARG00000019057, ENSOARG00000019057, ENSOARG00000019085, ENSOARG00000019104, ENSOARG00000019123, ENSOARG00000019125, ENSOARG00000019140, ENSOARG00000019154, ENSOARG00000019187, ENSOARG00000019199, ENSOARG00000019214, ENSOARG00000019215, ENSOARG00000019227, ENSOARG00000019246, ENSOARG00000019249, ENSOARG00000019265, ENSOARG00000019286, ENSOARG00000019289, ENSOARG00000019291, ENSOARG00000019304, ENSOARG00000019312, ENSOARG00000019323, ENSOARG00000019337, ENSOARG00000019351, ENSOARG00000019370, ENSOARG00000019386, ENSOARG00000019401, ENSOARG00000019404, ENSOARG00000019420, ENSOARG00000019422, ENSOARG00000019427, ENSOARG00000019442, ENSOARG00000019447, ENSOARG00000019450, ENSOARG00000019455, ENSOARG00000019475, ENSOARG00000019502, ENSOARG00000019524, ENSOARG00000019539, ENSOARG00000019563, ENSOARG00000019581	ACO2, POLR3H, CSDC2, PMM1, DESI1, C22orf46, MEI1, CCDC134, SREBF2, TNFRSF13C, CENPM, SEPT3, WBP2NL, NAGA, SMDT1, NDUFA6, TCF20, NFAM1, NFAM1, SERHL2, RRP7A, POLDIP3, CYB5R3, A4GALT, ARFGAP3, ARFGAP3, PACSIN2, TTLL1, MCAT, TSPO, TTLL12, SCUBE1, MPPED1, EFCAB6, SULT4A1, PNPLA5, PNPLA3, SAMM50, PARVB, PARVG, SHISAL1, PRR5, PHF21B, NUP50, KIAA0930, UPK3A, FAM118A, SMC1B, RIBC2, FBLN1, ATXN10, WNT7B, PPARA, CDPF1, PKDREJ, TTC38, TRMU, CELSR1, GRAMD4, CERK
4	320333-1978315	1.657	50	ENSOART00000014374, ENSOART00000014428, ENSOART00000014478, ENSOART00000014486, ENSOART00000014518	VSTM2A
4	8363934-13125455	4.761	50	ENSOARG00000016384, ENSOARG00000016532, ENSOARG00000016629, ENSOARG00000016682, ENSOARG00000016682, ENSOARG00000017010, ENSOARG00000017477, ENSOARG00000017669, ENSOARG00000017767, ENSOARG00000017958, ENSOARG00000018089, ENSOARG00000018184, ENSOARG00000018309, ENSOARG00000018481, ENSOARG00000018522, ENSOARG00000018572, ENSOARG00000018615, ENSOARG00000018713, ENSOARG00000018713, ENSOARG00000000197, ENSOARG00000000197, ENSOARG00000000708, ENSOARG00000000918, ENSOARG00000001035, ENSOARG00000001101, ENSOARG00000001206, ENSOARG00000001508, ENSOARG00000002097, ENSOARG00000002378, ENSOARG00000002378, ENSOARG00000002475, ENSOARG00000002679, ENSOARG00000002913, ENSOARG00000003042, ENSOARG00000003259, ENSOARG00000003383, ENSOARG00000003501, ENSOARG00000003766	CDK14, FZD1, MTERF1, MTERF1, AKAP9, CYP51A1, LRRD1, ANKIB1, GATAD1, PEX1, RBM48, FAM133B, CDK6, SAMD9, HEPACAM2, HEPACAM2, VPS50, VPS50, CALCR, TFPI2, GNGT1, GNG11, COL1A2, CASD1, SGCE, SGCE, PEG10, PPP1R9A, PON1, PON3, PON2, ASB4, PDK4, DYNC1I1
9	76652429-78607295	1.954	50	ENSOARG00000001013, ENSOARG00000001261, ENSOARG00000001480, ENSOARG00000001577, ENSOARG00000001692, ENSOARG00000002300, ENSOARG00000002300, ENSOARG00000014618, ENSOARG00000014632, ENSOARG00000002902, ENSOARG00000003080, ENSOARG00000003080, ENSOARG00000014656, ENSOARG00000003274	RNF19A, SPAG1, POLR2K, FBXO43, RGS22, VPS13B, VPS13B, STK3, STK3, KCNS2, NIPAL2
17	60404477-63386591	2.982	50	ENSOARG00000007078, ENSOARG00000007263, ENSOARG00000007382, ENSOARG00000007553, ENSOARG00000007699, ENSOARG00000007858, ENSOARG00000008129, ENSOARG00000008363, ENSOARG00000008557, ENSOARG00000008557, ENSOARG00000008612, ENSOARG00000008734,	LHX5, SDSL, SDS, PLBD2, DTX1, RASAL1, CFAP73, DDX54, TPCN1, SLC8B1,

				ENSOARG00000008875, ENSOARG00000009097, ENSOARG00000009430, ENSOARG00000009565, ENSOARG00000009765, ENSOARG00000009871, ENSOARG00000010003, ENSOARG00000010093, ENSOARG00000010295, ENSOARG00000010504, ENSOARG00000010548, ENSOARG00000004713, ENSOARG00000004740, ENSOARG00000004766, ENSOARG00000010578, ENSOARG00000010650, ENSOARG00000010904, ENSOARG00000011008, ENSOARG00000011301, ENSOARG00000011805, ENSOARG00000011850, ENSOARG00000011982, ENSOARG00000011987, ENSOARG00000012036, ENSOARG00000012097, ENSOARG00000012248, ENSOARG00000012253, ENSOARG00000012290, ENSOARG00000012315, ENSOARG00000012338, ENSOARG00000012402, ENSOARG00000012402, ENSOARG00000012429, ENSOARG00000012543, ENSOARG00000012664, ENSOARG00000012764, ENSOARG00000012897, ENSOARG00000012941, ENSOARG00000013039, ENSOARG00000013118, ENSOARG00000013244, ENSOARG00000013244, ENSOARG00000013764, ENSOARG00000013872, ENSOARG00000013935, ENSOARG00000014003, ENSOARG00000014044, ENSOARG00000014141, ENSOARG00000014274, ENSOARG00000014399, ENSOARG00000014443, ENSOARG00000014606, ENSOARG00000014797, ENSOARG00000014824, ENSOARG00000014942, ENSOARG00000015640, ENSOARG00000015640, ENSOARG00000015859, ENSOARG00000015998	RPH3A, PTPN11, TRAFD1, NAA25, ERP29, TMEM116, MAPKAPK5, ALDH2, BICDL1, RAB35, GCN1, PXN, SIRT4, PLA2G1B, MSI1, COX6A1, SRSF9, SRSF9, COQ5, RNF10, POP5, MLEC, UNC119B, ACADS, SPPL3, SPPL3, HNF1A, C12orf43, OASL, ANKRD13A, GIT2, GLTP, TRPV4, MVK, MMAB, UBE3B, UBE3B, KCTD10, MYO1H
19	19089105-26629558	7.54	50	ENSOARG00000007668, ENSOARG00000007719, ENSOARG00000007723, ENSOARG00000014064, ENSOARG00000007760, ENSOARG00000007805, ENSOARG00000007840, ENSOARG00000008132, ENSOARG00000008479, ENSOARG00000008566, ENSOARG00000014076, ENSOARG00000014093, ENSOARG00000008602, ENSOARG00000008631, ENSOARG00000008664, ENSOARG00000008743, ENSOARG00000008794, ENSOARG00000008838, ENSOARG00000009067	GRM7, EDEM1, ARL8B, BHLHE40, ITPR1, SUMF1, LRRN1, CRBN, TRNT1, IL5RA, CNTN4, CNTN6, CHL1
25	4000324-8609367	4.609	67	ENSOARG00000003216, ENSOARG00000003225, ENSOARG00000003238, ENSOARG00000003258, ENSOARG00000003265, ENSOARG00000003271, ENSOARG00000003276, ENSOARG00000003291, ENSOARG00000003327, ENSOARG00000003327, ENSOARG00000003337, ENSOARG00000003366, ENSOARG00000003390, ENSOARG00000003401, ENSOARG00000003407, ENSOARG00000003415, ENSOARG00000003420, ENSOARG00000003434, ENSOARG00000017793, ENSOARG00000003459, ENSOARG00000003469, ENSOARG00000003481, ENSOARG00000003489, ENSOARG00000003496, ENSOARG00000003507, ENSOARG00000003514, ENSOARG00000003571, ENSOARG00000003571	TRIM67, C1orf131, GNPAT, EXOC8, SPRTN, EGLN1, TSNAX-DISC1, SIPA1L2, SIPA1L2, NTPCR, PCNX2, MAP3K21, KCNK1, SLC35F3, TARBP1, IRF2BP2, TOMM20, ARID4B, GGPS1, B3GALNT2, GNG4, LYST

5. Conclusion

To our knowledge, this is the first study to describe the population structure of the D'man breed based on medium-density SNP array data. We show that the youngest D'man rams of the breeding station of Chenchou display consistent introgression signals from the Tunisian WTT. On the other hand, the oldest rams have preserved the ancestral genomic structure of the breed but have high average level of inbreeding. The introduction of new rams from flocks located in the Northern part of the country can remedy this problem. However, a molecular verification of racial authenticity of the selected rams is essential. Finally, the ROH analysis suggests the possibility of presence of other genes influencing the D'man prolificacy.

Acknowledgements

This research was supported by the International Foundation for Science (IFS grant B/5478-2). The authors wish to thank Tunisian Livestock and Pasture Office (OEP) for their precious help in blood sample collection.

6. References

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19 :1655–1664. doi:101101/gr094052109

Bailleul D, Stoeckel S, Arnaud-Haond S (2016) RClone: a package to identify MultiLocus Clonal Lineages and handle clonal data sets in r. Methods in Ecology and Evolution 7 : 966–970. doi:101111/2041-210X12550

Ben Jemaa S, Ruesche J, Sarry J, Woloszyn F, Lassoued N, Fabre S. (2019) The high prolificacy of D'man sheep is associated with the segregation of the *FecL* ^{L} mutation in the *B4GALNT2* gene. *Reprod Domest Anim*. doi:10.1111/rda.13391

Demars J, Cano M, Drouilhet L, Plisson-Petit F, Bardou P, Fabre S, Servin B, Sarry J, Woloszyn F, Mulsant P, Foulquier D, Carrière F, Aletru M, Rodde N, Cauet S, Bouchez O, Pirson M, Tosser-Klopp G, Allain D (2017) Genome-Wide Identification of the Mutation Underlying Fleece Variation and Discriminating Ancestral Hairy Species from Modern Woolly Sheep. Mol Biol Evol 34 : 1722–1729. doi:101093/molbev/msx114

Drouilhet L, Mansanet C, Sarry J, Tabet K, Bardou P, Woloszyn F, Lluch J, Harichaux G, Viguié C, Monniaux D, Bodin L, Mulsant P, Fabre S (2013) The Highly Prolific Phenotype of Lacaune Sheep Is Associated with an Ectopic Expression of the B4GALNT2 Gene within the Ovary. PLOS Genetics 9, e1003809. doi:101371/journalpgen1003809

Ducro BJ, Schurink A, Bastiaansen JWM, Boegheim IJM, van Steenbeek FG, Vos-Loohuis M, Nijman IJ, Monroe GR, Hellinga I, Dibbits BW, Back W, Leegwater PAJ (2015) A nonsense mutation in B3GALNT2 is concordant with hydrocephalus in Friesian horses. BMC Genomics 16 : 761. doi:101186/s12864-015-1936-z

Howrigan DP, Simonson MA, Keller M C (2011) Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. BMC Genomics 12 : 460. doi:101186/1471-2164-12-460

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers *Bioinformatics* 24 : 1403–1405. doi:101093/bioinformatics/btn129

Kdidi S, Calvo JH, González-Calvo L, Sassi M, Khorchani T, Yahyaoui M H (2015) Genetic relationship and admixture in four Tunisian sheep breeds revealed by microsatellite markers. *Small Ruminant Research* 131 :64–69. doi:101016/jsmallrumres201508012

Kijas JW, Lenstra JA, Hayes B, Boitard S, Porto Neto LR, San Cristobal M, Servin B, McCulloch R, Whan V, Gietzen K, Paiva S, Barendse W, Ciani E, Raadsma H, McEwan J, Dalrymple B, International Sheep Genomics Consortium Members (2012) Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol* 10, e1001258. doi:101371/journalpbio1001258

Mackness M I, and Durrington P N (1995) HDL, its enzymes and its potential to influence lipid peroxidation Atherosclerosis 115 : 243–253

Martin P, Raoul J, Bodin L (2014) Effects of the FecL major gene in the Lacaune meat sheep population *Genet Sel Evol* 46 : 48. doi:101186/1297-9686-46-48

Mastrangelo S, Ciani E, Sardina M T, Sottile G, Pilla F, Portolano B, and Bi.Ov. Ita Consortium (2018). Runs of homozygosity reveal genome-wide autozygosity in Italian sheep breeds. Anim Genet 49 : 71–81. doi:10.1111/age.12634

Mastrangelo S., Portolano B., Di Gerlando R., Ciampolini R., Tolone M., Sardina M. T., and International Sheep Genomics Consortium (2017). Genome-wide analysis in endangered populations: a case study in Barbaresca sheep. Animal 11 : 1107–1116. doi:10.1017/S1751731116002780

Meijide S, Pérez-Ruiz I, Hernández ML, Navarro R, Ferrando M, Larreategui Z, Ruiz-Sanz J-I, Ruiz-Larrea MB (2017) Paraoxonase activities in human follicular fluid: role in follicular maturation Reproductive BioMedicine Online 35 : 351–362. doi:101016/jrbmo201706008

Michailidou S, Tsangaris G, Fthenakis G C, Tzora A, Skoufos I, Karkabounas S C, Banos G, Argiriou A, and Arsenos G (2018) Genomic diversity and population structure of three autochthonous Greek sheep breeds assessed with genome-wide DNA arrays Mol Genet Genomics 293 : 753–768 doi:101007/s00438-018-1421-x

Miller S A, Dykes D D, and Polesky H F (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16 :1215

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M A R, Bender D, Maller J, Sklar P, de Bakker P I W, Daly M J, and Sham P C (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet 81 : 559–575

Rekik M, Ben Salem I, Ben Hamouda M, Diallo H, Ammar H, Aloulou R (2005) Numerical and weight productivities of crossbred D'man by local Queue Fine de l'Ouest ewes Revue Elev Méd vét Pays trop 58 : 81–88

Rosenberg N A (2004) distruct: a program for the graphical display of population structure. Molecular Ecology Notes 4 : 137–138 doi:101046/j1471-8286200300566x

Rowe A, Gondro C, Emery D, Sangster N (2009) Sequential microarray to identify timing of molecular responses to Haemonchus contortus infection in sheep. *Veterinary Parasitology* 161 : 76–87 doi:101016/jvetpar200812023

Szpiech ZA, Xu J, Pemberton TJ, Peng W, Zöllner S, Rosenberg NA, Li JZ (2013) Long Runs of Homozygosity Are Enriched for Deleterious Variation. Am J Hum Genet 93 : 90–102 doi:101016/jajhg201305003

Wang J (2017) Estimating pairwise relatedness in a small sample of individuals. Heredity (Edinb) 119 : 302–313. doi:101038/hdy201752