



Investigating whole genome selection signatures in local Moroccan sheep

Ouhrouch A., Boitard S., Boyer F., Da Silva A., Servin B., Pompanon F., Haddioui A., Benjelloun B.

in

López-Francos A. (ed.), Jouven M. (ed.), Porqueddu C. (ed.), Ben Salem H. (ed.), Keli A. (ed.), Araba A. (ed.), Chentouf M. (ed.). Efficiency and resilience of forage resources and small ruminant production to cope with global challenges in Mediterranean areas

Zaragoza : CIHEAM

Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 125

2021 pages 625-629

Article available on line / Article disponible en ligne à l'adresse :

http://om.ciheam.org/article.php?IDPDF=00008076

To cite this article / Pour citer cet article

Ouhrouch A., Boitard S., Boyer F., Da Silva A., Servin B., Pompanon F., Haddioui A., Benjelloun B. **Investigating whole genome selection signatures in local Moroccan sheep.** In : López-Francos A. (ed.), Jouven M. (ed.), Porqueddu C. (ed.), Ben Salem H. (ed.), Keli A. (ed.), Araba A. (ed.), Chentouf M. (ed.). *Efficiency and resilience of forage resources and small ruminant production to cope with global challenges in Mediterranean areas.* Zaragoza : CIHEAM, 2021. p. 625-629 (Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 125)



http://www.ciheam.org/ http://om.ciheam.org/



Investigating whole genome selection signatures in local Moroccan sheep

A. Ouhrouch^{1,2}, S. Boitard³, F. Boyer⁴, A. Da Silva⁵, B. Servin³, F. Pompanon⁴, A. Haddioui² and B. Benjelloun¹

¹Institut National de Recherche Agronomique (INRA Maroc). CRRA, Beni Mellal (Morocco)
²Université Sultan Moulay Slimane (FST-BeniMellal), LBVRP, Mghila, 523, Beni Mellal (Morocco)
³Laboratoire de Génétique Cellulaire. INRA, Toulouse (France)
⁴Laboratoire d'Écologie Alpine, Université Grenoble-Alpes. Avenue centrale, Grenoble (France)
⁵Univ. Limoges, INRA. EA7500, USC1061 GAMAA, 87000, Limoges (France)

Abstract. Sheep farming plays a key role in Moroccan agriculture and is one of the main sources of meat in the country. The main purpose of this study is to characterize diversity and demographic history as well as intra and inter-population selection signatures in the main sheep breeds raised in Morocco using their entire genomes. These local sheep breeds are characterized by a high genetic diversity and thus by a high adaptive potential. Therefore, they represent a valuable genetic resource for the conservation of the sheep species on a worldwide scale. The complete genome data from 87 individuals representing five predominant local sheep breeds in Morocco were used to infer demographic history, which has made it possible to estimate the evolution of the effective population size over time. Two methods were used to investigate selection signatures: one to detect putative regions under selection within each of these breeds and the second to detect selection signatures that differentiate the breeds one from the other. We identified several hundreds of regions/genes under selection from the studied breeds. We highlighted several biological processes involved in local adaptation as well as those linked to zootechnical performances characterizing each breed. Findings of this study increased our understanding on how genetic diversity is distributed in local breeds.

Keywords. Sheep – Whole genome sequences – Selection signatures – Local breeds.

Résumé. L'élevage ovin joue un rôle clé dans l'agriculture marocaine et constitue l'une des principales sources de viande dans le pays. Le but principal de cette étude est de caractériser la diversité et l'histoire démographique ainsi que les signatures de sélection intra et inter population dans les principales races ovines élevées au Maroc en utilisant leurs données de génomes entiers. Ces races ovines locales se caractérisent par une grande diversité génétique et donc par un potentiel d'adaptation élevé. Par conséquent, ils représentent une ressource génétique précieuse pour la conservation des espèces de moutons à l'échelle mondiale. Les données complètes du génome de 87 individus représentant cinq races ovines locales prédominantes au Maroc ont été utilisées pour déduire l'historique démographique, ce qui a permis d'estimer l'évolution de la taille efficace de la population au fil du temps. Deux méthodes ont été utilisées pour étudier les signatures de sélection : une pour détecter les régions présumées sous sélection dans chacune de ces races et la seconde pour détecter des signatures de sélection différenciant les races les unes des autres. Nous avons identifié plusieurs centaines de régions / gènes sous sélection parmi les races étudiées. Nous avons mis en évidence plusieurs processus biologiques impliqués dans l'adaptation locale ainsi que ceux liés aux performances zootechniques caractérisant chaque race. Les résultats de cette étude ont permis de mieux comprendre la répartition de la diversité génétique dans les races locales.

Mots-clés. Moutons – Séquences de génomes complets – Signatures de sélection – Races locales.

I – Introduction

Sheep farming plays a key role in Moroccan agriculture and is one of the main sources of meat in the country. It is mainly based on the breeding of indigenous breeds (e.g. Timahdite, Sardi, Dman, BeniGuil, etc.) well adapted to the specific and contrasting environmental conditions that prevail in their cradles as reported in the sheep plan (MARA, 1980). In the context of environmental change,

the improvement and conservation of this species are major challenges for the sustainable grant to the growing need for meat products both at national and international level. This conservation requires a perfect knowledge of their genetic heritage. The technological advances of the last decade have allowed access to the majority of the genome diversity of individuals and previous studies have shown via these technologies the importance of genome-environment relationships in Moroccan sheep (Benjelloun, 2015). However, the genomic specificities and genomic bases of the specific traits of local breeds are not known. In this context, the present work aims to determine the intra and inter-population selection signatures in five sheep breeds widely raised in Morocco using their complete genomic data. It identified the genomic regions/genes involved in local adaptation and those related to zootechnical performances characterizing each breed.

II – Material et methods

This study focused on the analysis of data from four officially identified sheep breeds (Timahdite, Sardi, Dman and BeniGuil) and a breed that is widely raised in the Oriental region and which is considered a priori as foreign: Ouled Jellal. The complete genomes of 87 sheep including 16 Timahdite, 27 Sardi, 30 Dman, 6 BeniGuil and 8 Ouled Jellal were analysed. The genotyped animals were sampled to represent the ecological and geographical diversity of each breeding area of the studied breed. Genomic data were produced at 12x coverage and then mapped on the reference genome OARv3.1 (Jiang *et al.*, 2014) and then processed using the pipeline described by Benjelloun (2015).

1. Inferring demographic history

The number of exclusive variants was calculated using Vcftools program (Danececk *et al.*, 2011). The effective population size (*Ne*) of the studied breeds was inferred using popsizeABC algorithm (Boitard *et al.*, 2016). Similarly, the demographic history was inferred by the same program, which made it possible to estimate the evolution of the effective population size over time from genomic data. For a basis of comparison, we have inferred the demographic history of a group of 13 wild sheep (*O. orientalis*) previously sampled in Iran and a group of 20 sheep representing 20 cosmopolitan breeds sampled in Eurasia, Africa and Australia (Benjelloun, 2015; Alberto *et al.*, 2018).

2. Investigating intra-population selection signatures using freqHMM

Following the results obtained on the demographic history, which made it possible to understand the demography of these breeds, a genomic scan was carried out using freqHMM software (Boitard *et al.*, 2012) to identify regions potentially under selection within each breed. This is a program that can detect variants under intra-population selection. The input files included all non-rare SNP variants identified in the five breeds (31,442,046 SNPs having MAFQ>=0.05), considering as ancestral alleles the homozygous alleles of the reference goat for the corresponding loci. This allowed to calculate the number of derived and ancestral alleles for each SNP. Then, the freqHMM analysis was done using the calibrated and adapted parameters for each breed.

3. Investigating inter-population selection signatures using hapFLK

The second part of the analyses consists of looking for breeding signatures differentiating breeds by FLK methods (Bonhomme *et al.*, 2010) and hapFLK (Fariello *et al.*, 2013) while considering the group of 13 Asiatic mouflons as outgroup. Both tests aim to identify typical differentiation regions between populations while considering their structures. The FLK test infers the Kinship matrix and Reynolds distances. The hapFLK algorithm has been applied to genomic data while considering sliding windows of 10Mb chromosomal segments with 1Mb overlaps.

III – Results and discussion

1. Demographic history results

The intra-breed diversity state shows very high numbers of exclusive variants within each breed and we note that the two Dman (1.8 million) and Sardi (1.6 million) breeds have the most exclusive variants (Table 1). These numbers of exclusive variants are not correlated with the effective population size of each breed (Table 1). The highest effective size is recorded by the Timahdite breed (Ne ~ 47,000), while other breeds have close effective sizes. The values obtained for these local breeds are substantially higher than those of wild sheep populations at the centre of domestication (Ne = 2,016). Moreover, as expected, cosmopolitan races (even when grouped) are marked by a very low effective size (Ne = 285), despite being present in very large numbers worldwide. This shows the threat of large-scale distribution of these breeds and their replacement of local breeds around the world. In addition, the effective sizes of the local breeds studied are very high compared to what is reported in the literature for sheep breeds (e.g. Maiwashe and Blackburn, 2004, Tapio *et al.*, 2005). This illustrates the exceptional adaptive potential of these local breeds and the opportunity to further intensify the reasoned programs of their breeding while maintaining this richness.

Breed	Beni Guil	Dman	Ouled Jellal	Sardi	Timahdite	Asian Mouflons	20 cosmopolitan breeds
No. of exclusive variants	341,296	1,783,651	460,461	1,621,728	934,212	_	_
Effective population size (Ne)	30,921	29,789	25,811	28,523	47,097	2,016	285

Table 1	. Demographic	characteristics	of the	studied	populations
---------	---------------	-----------------	--------	---------	-------------

2. Intra population selection signatures in local sheep breeds

A whole analysis pipeline was set up to calibrate the research programs of the selection signatures according to the demographic histories that the studied breeds had experienced. Once these programs were calibrated, we were able to identify a large number of variants and genes selected in each one of the studied breeds. From a total of 31,442,046 SNPs analysed, 43,616 variants associated with 115 genes in the Sardi breed, 34,909 SNPs associated with 96 genes were identified under selection in Timahdite, 56,218 variants associated with 153 genes in Dman, 54,845 associated with 127 genes in the Beni Guil breed and 53,346 variants associated with 139 genes in Ouled Jellal, were identified under selection. Table 2 illustrates the number of SNPs variants and associated genes that are commonly identified under selection between each couple of breeds. We note that the three breeds Timahdite, Sardi and Dman; share more SNPs/genes under selection between each other than with the other two breeds. More investigation would elucidate the reasons and the biological processes associated.

······································							
Breed	Beni Guil	Dman	Ouled Jellal	Sardi	Timahdite		
Beni Guil	_	8.180	5.307	6.311	5.880		
Dman	36	_	5.795	14.568	9.436		
Ouled Jellal	30	32	_	8.238	6.846		
Sardi	30	70	36	_	9.708		
Timahdite	25	60	33	45	-		

Table 2. Number of selected SNPs (above the diagonal) and selected genes (below the diagonal) common to each two sheep breeds

Efficiency and resilience of forage resources and small ruminant production to cope with global challenges in Mediterranean areas

The number of variants identified under selection commonly in all studied breeds is 219. These variants are associated with 4 genes (Fig. 1): *HMGA2*, *RCOR1*, *SBF2* and *U6*. These genes are generally involved in the cellular functioning of sheep. The genes identified in common by considering different combinations of breeds and also the signatures associated exclusively in a given breed (Fig. 1) will shed light on biological processes underlying both adaptive and zootechnical traits selected in each breed.



Fig. 1. Venn diagram ilustrating the number of genes selected commonly in flve local sheep breeds and those exclusively selected in each breed.

3. Inter population selection signatures in local sheep breeds

The hapFLK scores were subjected to Bonferroni correction to determine the selection threshold. The analysis process identified 3,536 variants under selection that differentiate the five local breeds. The main selective sweeps identified are associated with the genes *RXFP2, FANCA-201, SPIRE2, TCF25 and MC1R-201*. (Fig. 2). The association between all the SNPs/genes identified and the breeds under consideration would highlight the selected biological processes within each breed.

IV – Conclusion

The results obtained at this stage identified several regions and genes selected in the studied breeds. This would allow understanding the mechanisms differentiating the main sheep breeds in Morocco and their genomic traceability. Considerable work is planned to identify the associated biological processes. Outlier genomic variants identified here will be ranked according to their importance in the context of climate changes and some of them would be very precious when designing selection programs based on adapted genotyping tools.



Fig. 2. Manhattan plot depicting hapFLK scores in five Moroccan sheeps along their genome. Each point represents a SNP. The horizontal red line represents the 5% Bonferroni threshold of significance.

Acknowledgments

These data were generated as part of the UE FP7 NextGen project. We would like to thank the consortium of this project for data availability.

References

- Alberto, F.J., Boyer, F., Orozco-terWengel, P., Streeter, I., Servin, B., de Villemereuil, P., Benjelloun, B., Pompanon, F., 2018. Convergent genomic signatures of domestication in sheep and goats. *Nature Communications*, 9(1), 813. doi: 10.1038/s41467-018-03206-y
- Benjelloun, B., 2015. Diversité des génomes et adaptation locale des petits ruminants d'un pays méditerranéen: le Maroc. Biodiversité, Ecologie, Environnement. Université Grenoble Alpes, 213 pp.
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M-A., Handsaker, R., Lunter, G., Marth, G., Sherry, S.T., McVean, G., Durbin, R. and 1000 Genomes Project Analysis Group, 2011. 'The Variant Call Format and VCFtools', Bioinformatics 27: 2156-2158.
- Boitard, S., Rodríguez, W., Jay, F., Mona, S., Austerlitz, F., 2016. Inferring Population Size History from Large Samples of Genome-Wide Molecular Data – An Approximate Bayesian Computation Approach. *PLoS Genet* 12(3): e1005877. doi: 10.1371/journal.pgen.1005877.
- Boitard, et al., 2012. Detecting Selective Sweeps from Pooled Next-Generation Sequencing Samples. Mol. Biol. Evol. doi:10.1093/molbev/mss090.
- Bonhomme, et al., 2010, Detecting selection in population trees: The Lewontin and Krakauer test extended. Genetics 186(1) 241-262.
- Fariello, et al., 2013, Detecting Signatures of Selection Through Haplotype Differentiation Among Hierarchically Structured Populations. Genetics 193(3): 929-941.
- Maiwashe, A.N., Blackburn, H.D., 2004. Genetic diversity and conservation strategy considerations for Navajo Churro sheep. *Journal of Animal Science*, 82, 2900-2905.
- Tapio, M., Tapio, I., Grislis, Z. et al., 2005. Native breeds demonstrate high contributions to the molecular variation in northern European sheep. *Molecular Ecology*, 14, 3951-3963.