



Genomic bases of morphometric variations in small ruminants of Morocco

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 645-649

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

http://om.ciheam.org/article.php?IDPDF=00008080

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. **Genomic bases of morphometric variations in small ruminants of Morocco.** 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. 645-649 (Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 125)



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



Genomic bases of morphometric variations in small ruminants of Morocco

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. Small ruminants are one of the main sources of meat and milk around the world. The improvement and conservation of these species constitute major challenges for their sustainable contribution to the needs of human populations. Some morphological variations of individuals reflect the diversity of their adaptive potential and their zootechnical performances. The aim of this work is to analyze whole genome data of 160 unrelated local sheep, at a depth of 12X, sampled to represent the ecological and geographical diversity of Morocco. It aims at studying the genomic bases of their phenotypic variations by associating the genomic variations noted with the best possible resolution of the genomes (sequences of the complete genomes) with a certain number of morphometric indicators which affect the adaptive potential (ie. ears length) or the productive potential (ie. shoulder height, barrel tour). This association, carried out via genome wide association approaches (GWAS) will make it possible to highlight the possible genomic mutations responsible for these variations, among which we can identify those involved in adaptation to different climatic environments and those linked to the best zootechnical performances. These mutations should make it possible to incorporate these variations, after their validation, into selection schemes to improve the productivity of these species and their adaptability to certain climatic conditions.

Keywords. Sheep – Adaptation – SNP – GWAS.

Résumé. Les petits ruminants sont l'une des principales sources de viande et de lait dans le monde. L'amélioration et la conservation de ces espèces constituent des défis majeurs pour leur soutien durable aux besoins des populations humaines dans ces produits. Certaines variations morphologiques d'individus reflètent la diversité de leur potentiel adaptatif et de leurs performances zootechniques. Le but de ce travail est d'analyser les données de génome entier de 160 moutons locaux non apparentés, à une profondeur de 12X, échantillonnés pour représenter la diversité écologique et géographique du Maroc. Il s'agit d'étudier les bases génomiques de leurs variations phénotypiques en associant les variations génomiques relevées à la meilleure résolution possible des génomes (séquences de génomes complets) avec un certain nombre d'indicateurs morphométriques qui affectent le potentiel adaptatif (longueur des oreilles) ou le potentiel de production (e.g. hauteur au garrot, tour des canons). Cette association, réalisée via des approches d'association pangénomique (GWAS), permettra de mettre en évidence les mutations génomiques possibles responsables de ces variations, parmi lesquelles nous pouvons identifier les acteurs de l'adaptation à différents environnements climatiques et ceux liés aux meilleures performances zootechniques. Ces mutations devraient permettre d'intégrer ces variations, après leur validation, dans des schémas de sélection visant à améliorer la productivité de ces espèces et leur adaptabilité à certaines conditions climatiques.

Mots-clés. Moutons – Adaptation – SNP – GWAS.

I – Introduction

Small ruminants, especially sheep, are one of the main sources of meat and milk around the world. The improvement and conservation of these species are major challenges in meeting the needs of human populations in these commodities. Some morphological variations reflect the diversity

of their adaptive potential (e.g. length of ears ...) and productivity (e.g. shoulder height, cannon tour). These variations can serve as proxies to study local adaptation to constraining environments and to identify genes associated with the best zootechnical performance (e.g. meat production).

This research is part of this perspective to study the genomic bases of certain morphometric variations in sheep via Genome Wide Association approaches (GWAS) using their whole genomes. This is a work following the UE FP7 NextGen European project (Benjelloun, 2015) analysing 160 sheep representing the ecological, climatic and phenotypic diversity of Morocco. Here we present the results of a descriptive analysis of their genomes with one morphometric variable, as well as a description of the approach taken for the association studies (GWAS).

II – Material and methods

The sampled sheep are unrelated and represent the eco-climatic diversity of Morocco. They were selected so as to have the maximum distance between individuals over a very wide region of about 400,000 km². For each individual, tissue samples were taken from the distal part of the ear and then placed in alcohol for one day, after which they are transferred to a silica gel tube pending the extraction of DNA. A total of 412 herds were covered. The most important criterion was to optimize the selection of individuals to be sequenced to represent a wide range of environmental conditions and to ensure spatial representativeness of all regions, as described by Benjelloun (2015).

The complete genomes of the 160 selected sheep, representative of the entire Moroccan climate variation gradient, were sequenced (Whole Genome Sequencing) with a coverage rate of 12X. (Benjelloun, 2015).

Illumina readings for sheep were aligned to the sheep reference genome [OAR v3.1, GenBank assembly GCA_000317765.1 (Jiang *et al.*, 2014)] and those for goats were aligned to the reference genome of the sheep.Variant discovery was performed using three different algorithms: Samtools Mpileup (Li *et al.*, 2009), GATK UnifiedGenotyper (McKenna *et al.*, 2010) and Freebayes (Garrison and Marth, 2012).

This study plans to use the Genome-wide efficient mixed-model analysis for association studies; the GEMMA program (Zhou and Stephens, 2012), which is one of the most accurate methods designed to identify genetic associations with observable traits, and which incorporates a model. Linear Mixed Model that takes into account population structure as well as other genetic background factors that may affect association testing.

The diagram illustrated in Fig. 1, summarizes the different steps that were performed for this test association between variants discovered and a phenotypic variable measured in the sampled sheep. In these steps, other softwares were used for processing the data (as explained in Fig. 1): VCFTOOLS (Danecek *et al.*, 2011), PLINK (Purcell *et al.*, 2007) and BIMBAM (Guan and Stephens, 2010).

III – Results and discussion

1. Polymorphism in sheep genomes

Around 39 million variants were successfully identified in sheep, 6.8% of which were small insertions/deletions (indels) and 2.1% showed more than two alleles. These results show a very high polymorphism in this species. Admixture analysis suggested a weak population structure, whose data were better explained by considering the existence of a single group of sheep (Benjelloun, 2015). However, a soft pattern of weak geographical structure appears when considering the existence of three homogeneous groups.



Fig. 1. Steps taken for association tests using GEMMA software.

2. GWAS results

The very strong polymorphism of the genomes of this species. The weak structuring of diversity shows that the sheep of Morocco represent a favourable biological model to the genotype/phenotype association tests at the genomic scale.

GWAS analyses identified several important significant signals at the genomic level in sheep that would be associated to the variation of certain morphometric characters measured in the sampled sheep. We present here the results relating to the character: 'Shoulder height' (Fig. 2), that shows a significant correlation between the studied phenotype and some SNPs on chromosomes 11, 16 and 20.

The method adopted thus makes it possible to identify certain SNPs in potential association with the expression of the studied phenotype. The analyses are being carried out to elucidate the potential role of these SNPs in the variation of the shoulder height in sheep as well as for other morphometric characters.

IV – Conclusion

The results of this study revealed several sites putatively related to the phenotype studied, despite the relatively limited number of individuals used in this study. This would be linked to the sampling strategy, which allowed a wide coverage of the phenotypic and environmental diversity present in Morocco, and the use of whole genome data. This work shows that the approach adopted ensures the feasibility of these genotype-phenotype association studies with the biological model used. Thus, the further identification of genomic regions involved in a wide range of morphometric variations of sheep.

This project will allow, later, to highlight the possible genomic mutations responsible for phenotypic variations among which we can identify those related to adaptation to different climatic environments. This should make it possible to integrate these variations, after their validation, into selection schemes in order to improve the adaptability of this species to different climatic conditions.



Fig. 2. Manhattan plot of the association of genotypes with the trait Shoulder height in sheep. Each point represents a SNP along 26 chromosomes of the genome. The horizontal line in blue represents the Bonferroni 5% threshold for which the association is considered significant.

Acknowledgments

These data were generated as part of the UE FP7 NextGen project. We would like to thank the consortium of this project for the data availability. We would also like to thank Florian J. Alberto for the precious help in carrying out this analysis.

References

- Benjelloun B., 2015. Diversité des génomes et adaptation locale des petits ruminants d'un pays méditerranéen: le Maroc. PhD Thesis. Biodiversité, Ecologie, Environnement. Université Grenoble Alpes, 213 pp.
- Danecek Petr, Adam Auton, Goncalo Abecasis, Cornelis A. Albers, Eric Banks, Mark A. DePristo, Robert Handsaker, Gerton Lunter, Gabor Marth, Stephen T. Sherry, Gilean McVean, Richard Durbin and 1000 Genomes Project Analysis Group, 2011. 'The Variant Call Format and VCFtools', Bioinformatics 27: 2156-2158.
- Garrison E, Marth G., 2012. Haplotype-based variant detection from short-read sequencing. Vol 1207.3907v2. ArXiv.
- Guan and Stephens. BIMBAM Bayesian imputation-based association mapping. Copyright (C) 2008-2015. DOI: 10.1214/11-AOAS455.
- Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, Wu C, Muzny DM, Li Y, Zhang W et al., 2014. The sheep genome illuminates biology of the rumen and lipid metabolism. Science 344(6188): 1168-1173.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data P., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16): 2078-2079.

- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M et al., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing nextgeneration DNA sequencing data. Genome Research 20(9): 1297-1303.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ and Sham PC, 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics, 81:559-575.
- Zhou and Stephens, 2012. Software tool and univariate linear mixed models: Genome-wide efficient mixedmodel analysis for association studies. Nature Genetics. 44: 821-824.