

Assessing Genomic Inbreeding and Homozygosity Patterns in Hungarian Merino Sheep and its Relatives

George Wanjala^{1,2,3}, Putri Kusuma Astuti^{1,2}, Nelly Kichamu^{1,2,4}, Zoltán Bagi¹, Szilvia Kusza^{1*}

¹Centre for Agricultural Genomics and Biotechnology, Faculty of Agricultural and Food Sciences and Environmental Management, University of Debrecen, 4032 Debrecen Egyetem tér 1., Hungary.

²Doctoral School of Animal Science, University of Debrecen, 4032 Debrecen Böszörményi út 138., Hungary.

³Directorate of livestock production, Bungoma County, Box 437-50200, Bungoma Kenya.

⁴Ministry of agriculture livestock, fisheries, and cooperatives, State department of livestock development, Naivasha sheep and goats breeding station, Box 20117, Naivasha Kenya.

Abstract

The Merino breed of sheep is one of the oldest breeds still in existence. All around the world, purebred and mixed populations of the Merino and Merino-derived sheep breeds can be found. This is a result of their wide distribution and increasing popularity. They constitute a huge genetic resource that has been applied historically to create numerous new breeds. The current Hungarian Merino sheep breed has evolved over many years as a result of crossbreeding with different breeds. Consequently, the objective of this study is to compare the genomic inbreeding coefficient (FROH) and patterns of runs of homozygosity (ROH) between the Hungarian Merino and other closely related Merino breeds. Following Rambouillet, Hungarian Merino had the second highest number of ROH class of 1-6 mbs. The Spanish Merinos Cordoba had the most ROH class of > 48Mbs, while the Hungarian Merino had none. The FROH of Hungarian Merinos was comparable to that of other studied breeds, while Spanish Merinos Cordoba had the highest. This study confirms that Hungarian Merino has several intermediate ancestors, and its breeding management is comparable to other Merino breeds and may also indicate greater genetic diversity within the breed. The results can therefore be used to improve genomic management of Hungarian Merino sheep breed.

Keywords: Genetic diversity, Hungary, Inbreeding coefficient, Merino, Sheep, SNP Markers.

1. Introduction

Merino sheep breeds are among the world's oldest established breeds. Although it is widely believed that the Merino breed originated in Spain (Iberia region [1,2]), the origin of the name "Merino" has been disputed. Some speculated that the name Merino came from the word "Marina," as the breed may have traveled across the sea to Spain. Others thought the name "Merino" came from the Spanish word for "wandering," while others thought it came from the Latin word "Majorinus"

for a head-man [2]. However, the widely accepted origin of the name "Merino" is that it came from the Moroccan people "Beni Merin," who invaded Spain in the 12th century and moved with large flocks of sheep, implying that the Merino breed has been in existence since the 12th century [2]. Since then, the breed has spread all over the world and is said to be approximately 20% of the total global sheep population.

In Hungary, the Merino sheep breed has a long history. Most of it is documented in [1,3,4]; in summary, the Merino sheep breed is one of the most important in the Hungarian sheep industry, accounting for approximately 80% of the national sheep population. Its breeding dates back to 1774, when the first Merino sheep were imported from

* Corresponding author: Szilvia Kusza, kusza@agr.unideb.hu

Spain at Queen Maria Terezia's request [1]. The breed "Hungarian Merino" was later developed through several stages across the years, depending on the breeding objective at the time. For instance, first Hungarian Merino breed was developed by crossing the imported Merinos with the German Mutton Merino and Rambouillet Merino [1], possibly with the intention of improving mutton production. Several crossings of Hungarian Merino and other Merino breeds, such as Russian Merino breeds, French Merino Precoce rams, and German Mutton Merino rams, were made to improve mutton production too. Hungarian-breed development then shifted to improving wool traits, with Hungarian Merinos being crossed with Australian Merinos, Kent, and Corriedale breeds. Furthermore, by crossing the Hungarian Merino with the Booroola Merino breed, prolificacy was improved [1]. With this brief history, it is clear that Hungarian Merino has a wide range of close "relatives" spread across Europe and Australia.

It is with this view that, this study comparatively examines the patterns of runs of homozygosity (ROH) between Hungarian Merino and the Merino breeds which are believed to have had a strong genetic contribution to its development.

Contiguous regions of the genome which occur when an individual inherits a haplotype from both parents are called Runs of Homozygosity. They may arise from various population genetics events including and not limited to population bottlenecks, artificial and/or natural selection sweeps and consanguineous mating [5]. ROH studies have been conducted in several domestic livestock species, including sheep [6,7], cattle [8,9], pigs [10], just to mention a few. Since 1990, when ROH was first identified in humans using microsatellite markers [11]. The authors of that study came to the conclusion that long homozygous segments in human genomes may have a significant impact on gene mapping. In livestock species, ROH has been used to determine the inbreeding levels, population history, population bottle necks, selection sweeps [12] among others etc.

Therefore, this study will evaluate how runs of homozygosity pattern in Hungarian Merino is compared to closely related Merino breeds namely Spanish Merino Cordoba, Australian Merino, Australian Industry Merino, Chinese Merino, and Rambouillet.

2. Materials and methods

Twenty-eight Hungarian Merino samples from unrelated individuals were collected from several farms using vacutainer tubes. DNA isolation and genotyping by sequencing was outsourced to Neogen company inc [13]. Genotyping was done using the Illumina OvineSNP50 BeadChip on the Infinium assay platform spanning the whole genome. Variants were called based on OAR v4 reference genome. We received a set of files including final report and SNP map which were used to generate the working data hence forth referred to as "Hungarian dataset" set. This study also included selected Merino and Merino derived breeds which were studied by Ciani et al. [14], and further re-analyzed by Zsolnai et al. [3] to establish the genomic position of Hungarian Merino among other Merinos and their derived breeds. We therefore selected breeds that were shown to be phylogenetically close to Hungarian Merino. And thus, the breeds selected breeds included Spanish Merino Cordoba (SMC, n = 7), Australian Merino (AM, n = 24), Australian Industry Merino (AIM, n = 24), Chinese Merino (CM, n = 23), and Rambouillet (RM, n = 24)-hence forth referred to as "Ciani" dataset.

It is important to note here that, the Hungarian data set used in this study is not the same dataset used by Zsolnai et al. [3]. The Hungarian dataset was merged with the Ciani dataset using plink v1.9 [15], with the former dataset being treated as base genotypes. The final data set comprised of seven breeds of 155 animals before quality control (QC). Post merging genotyping quality was 0.72 with around 62481 SNPs.

The following quality control criteria were used, --maf 0.05, --autosome 0.3, --geno 0.1, and --hwe 10e-6. With one individual removed due to a missing genotype (--mind), 18695 variants removed due to a missing genotype (--geno), and 12 variants deleted due to the Hardy-Weinberg exact test (--hwe), the post-merging dataset had a genotyping quality of 0.74. Due to the minor allele threshold (-maf), 1569 variations were ultimately deleted. The final data set, which contained 154 samples and 42205 variations, was utilized to evaluate the ROH patterns.

ROHs were calculated using the detectRUN package [16] implemented in R software [17], applying the sliding window methodology. Since there are no universally defined criteria on how to

compute the ROHs, we used the criteria described by Dzomba et al. [18]. In summary the variants were not pruned based on linkage disequilibrium (LD), but the minimum length required to be considered ROH was increased to 1 megabases (mb) in order to exclude any short ROH that could have been caused by LD. The sliding window computation was performed using the following parameters: Window size of 15 mb with a threshold of 0.05 for an SNP to be called in the run. Further a minimum number of heterozygous genotypes was set to 1 with an allowance of a possible missing genotype of 1. The ROH window needed to have at least 30 SNPs in order to be considered complete. In addition, the minimum SNP density per ROH was set to one SNP every 100 kilo bases (kb) while the maximum gap between adjacent SNPs was set to 1mb. The ROHs that were found were categorized into the following four categories according to their length: 1–6 mb, 6–12 mb, 12–24mb, 24–48mb and above 48 mb. The total length of each animal's ROH was added up and then averaged across all breeds to determine the breed-specific mean sum of ROH length. In addition to that, this value was determined for each category. For each breed, the total number of ROH, the mean number of ROH,

and the average length of ROH were all determined. Also, the peak runs; the runs where SNPs occur in more than a specific percentage of individuals in a population were identified. The thresholds that were used to define the peak runs are SNPs that occurred in runs in at least 70%, 60%, and 50% of individuals in the population.

3. Results and discussion

A crucial genetic technique for identifying the population development history is the length of the ROH. ROHs, which are defined by their length, are simply haplotypes from the past common ancestors. Those ROHs between 1-6 mbs show haplotypes from very distant ancestors, whereas ROHs over 48 mbs show haplotypes from the most recent ancestors [19]. In Table 1, a summary of the ROH also referred to as “runs” and FROH per sheep breed are presented. The Hungarian sheep breed had the second highest number of runs of between 1-6 mbs, trailing only the Rambouillet breed. In terms of long runs of more than 48 mbs, the Spanish Merino Cordoba had the most, followed by Rambouillet suggesting that the later two breeds are highly inbred than others. Notably, Hungarian Merino alongside

Table 1. Summary of runs of homozygosity and inbreeding levels per sheep breed.

Breed	Runs of homozygosity classes					FROH
	1-6	6-12	12-24	24-48	>48	
1	670	86	28	9	-	0.053 ± 0.022
2	807	149	62	21	1	0.085 ± 0.052
3	591	106	33	11	-	0.055 ± 0.027
4	663	95	36	16	1	0.063 ± 0.046
5	1041	116	20	7	-	0.060 ± 0.024
6	1206	178	61	10	2	0.101 ± 0.052
7	318	187	115	41	7	0.347 ± 0.051

1 = Australian Industry Merino, 2 = Australian Merino, 3 = Australian Poll Merino, 4 = Chinese Merino,

5 = Hungarian Merino, 6 = Rambouillet, 7 = Spanish Merinos Cordoba.

FROH = Genomic inbreeding coefficient

Australian Industry Merino, and Australian Poll Merino did not have long runs of above 48 mbs implying that the genetic management of these breeds could have successfully minimized levels of inbreeding as much as possible. In addition, the Spanish Merinos Cordoba had the highest intermediate classes of runs (6-12, 12-24, and 24-48 mb), whereas other breeds exhibited substantial variations in all these classes of ROH's. These findings demonstrate that, despite sharing a common ancestor, there are genetic differences amongst the populations under study.

Moreover, as length/ROH classes increased across genome, the number of ROHs decreased. In other words, the observed number of ROH for 1-6 mb decreased gradually as ROH classes increased to >48 mb in all populations. This suggests that the ancient populations were more inbred than the current populations. [20]. Figure 1 shows the number of SNPs found in each breed and ROH classes. Except for the > 48mb class, where Hungarian Merino, Australian Industry, and Australian Poll Merinos did not have any SNP, all breeds had a similar number of SNPs in

almost all the ROH classes. On average, Spanish Merino Cordoba had the highest mean number of SNPs in the ROH in every chromosome than all other breeds whereas Hungarian Merino's was relatively comparable to other studied breeds.

Since all breeds had a comparable number of ROH classes, it could be suggested that all the breeds had a series of intermediate ancestors indicating that they were developed through several stages just like the Hungarian Merino.

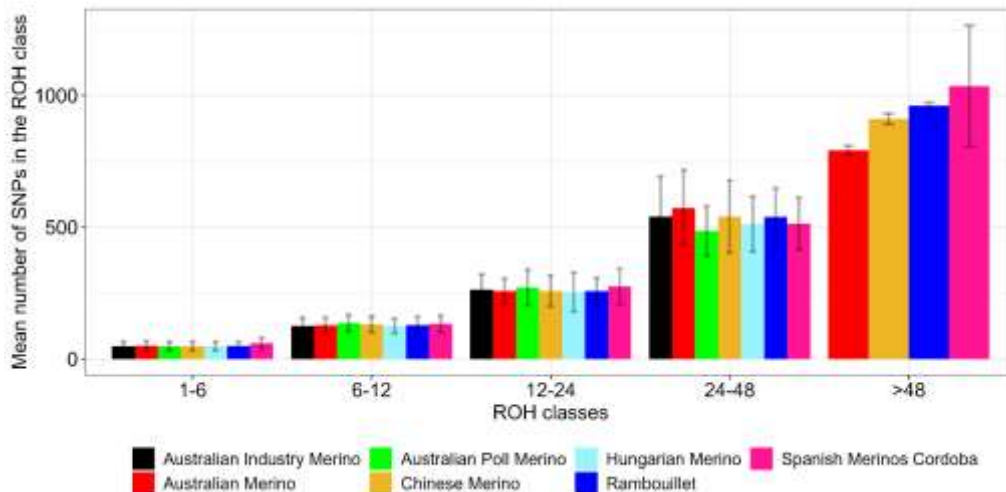


Figure 1. The average number of SNPs per ROH class in the Merino sheep breeds.

Except for Rambouillet ($FROH = 0.101 \pm 0.052$) and Spanish Merinos Cordoba ($FROH = 0.347 \pm 0.051$), all other studied Merino breeds had ROH inbreeding levels that were comparable to those of the Hungarian Merino ($FROH = 0.060 \pm 0.024$).

These findings are consistent with Figure 2, which shows that the Spanish Merinos Cordoba had most FROHs across the genome, while Hungarian Merino had least spread FROHs. It is worthy to emphasize the impact of inbreeding on decreased fitness and productivity. Consanguineous mating may cause the inheritance of haplotypes that are identical by descent (IBD), which may result in long homozygous stretches in the progeny (long ROH > 48mbs in this case). These regions potentially increase the risks of recessive alleles to be co-

expressed reducing the fitness traits of the population.

Also, the occurrence of ROH could be an indicator of selection since it has been found that the genomic regions under selection overlap at an "ROH island" [21]. ROH islands are defined by Bosse [22] as genomic regions with highest distribution of ROHs.

The Hungarian Merino had only one ROH island with thirty-four (34) SNPs on chromosome 6, stretching between 40.5 mb and 43.0 mb, On the same chromosome Rambouillet had also one ROH island of 38 SNPs spanning between 40.2 mb and 43.01 mb while Spanish Merino Cordoba had multiple islands on different chromosomes. In particular, chromosome 6 island in Spanish Merino Cordoba had 139 SNPs spanning 10.0 mb and 10.8 mb.

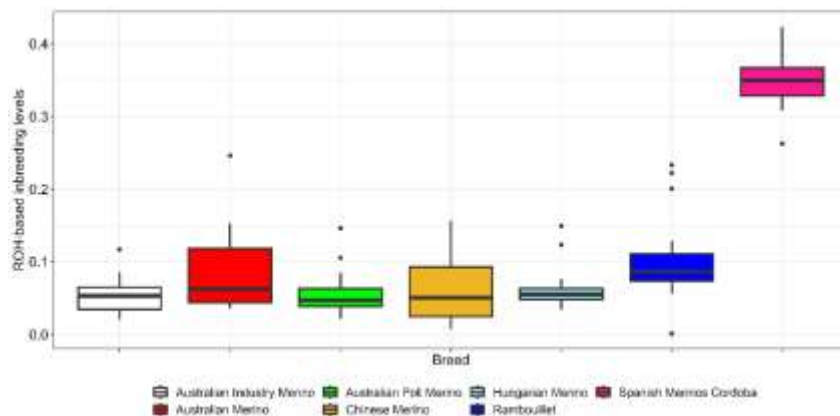


Figure 2. Average ROH-based inbreeding levels per sheep breed.

All the above islands were defined as peak regions across the genome which has SNPs in the ROH in more than 50% individuals in the population. Following this criterion, only Hungarian Merino, Rambouillet and Spanish Merino Cordoba were represented as described above. Although we did not dive into gene hunting, we hypothesize that these regions could be underpinning ancestral selection signatures. The present signatures identified differ from the signatures identified by Megdiche [23] in a multi-cohort approach of analysis of selection of signatures in a combined group of Merinos involving Spanish Merino, Australian Merino, Chinese Merino, and Sopravissana. The authors identified several signatures that underly wool traits.

4. Conclusions

Based on the present results, the Hungarian Merino breed has exhibited genetic uniqueness when compared with the studied breeds. The breed has its unique numbers of ROH classes suggesting a totally different genetic resource from other Merino breeds. Besides, the low FROH exhibited by Hungarian Merino shows the success of Merino genetic resource management.

Having singled out chromosome 6 which was a top hit ROH island in Hungarian Merino, both Hungarian Merino and Rambouillet had almost similar number of SNPs and genomic regions making in ROH island where Spanish Merino Cordoba had more than 100 SNPs on the same chromosome with but different coordinates. This

suggests that Hungarian Merino and Rambouillet share some ancestral selection signatures.

Acknowledgements

G.W., P.K.A, and N.K were supported by Tempus Public Foundation within the Stipendium Hungaricum Programme.

The work was partially funded by the project European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 772787 (SMARTER).

References

1. Loukovitis, D., Szabó, M., Chatziplis, D., Monori, I. and Kusza, S., 2022. Genetic diversity and substructuring of the Hungarian Merino sheep breed using microsatellite markers. *Animal Biotechnology*, pp.1-9.
2. Mallett, J.A., 1960. Famous sheep breeds. 4. The Border Leicester. *Journal of the Department of Agriculture, Western Australia, Series 4, 1(4)*, pp.305-307.
3. Zsolnai, A., Egerszegi, I., Rózsa, L., Mezőszentgyörgyi, D. and Anton, I., 2022. Position of Hungarian Merino among other Merinos, within-breed genetic similarity network and markers associated with daily weight gain. *Animal Bioscience*, 36(1), pp.10-18.
4. Nagy, Z., Németh, A., Mihályfi, S., Toldi, G., Gergátz, E. and Holló, I., 2011. The short history of Hungarian sheep breeding and Hungarian Merino breed. *Acta Agraria Kaposváriensis*, 15(1), pp.19-26.
5. Ceballos, F.C., Hazelhurst, S. and Ramsay, M., 2018. Assessing runs of Homozygosity: a comparison of SNP Array and whole genome sequence low coverage data. *BMC genomics*, 19, pp.1-12.

6. Nosrati, M., Nanaei, H.A., Javanmard, A. and Esmailzadeh, A., 2021. The pattern of runs of homozygosity and genomic inbreeding in world-wide sheep populations. *Genomics*, 113(3), pp.1407-1415.
7. Selli, A., Ventura, R.V., Fonseca, P.A., Buzanskas, M.E., Andrietta, L.T., Balieiro, J.C. and Brito, L.F., 2021. Detection and Visualization of Heterozygosity-Rich Regions and Runs of Homozygosity in Worldwide Sheep Populations. *Animals*, 11(9), p.2696.
8. Purfield, D.C., Berry, D.P., McParland, S. and Bradley, D.G., 2012. Runs of homozygosity and population history in cattle. *BMC genetics*, 13, pp.1-11.
9. Goszczynski, D., Molina, A., Terán, E., Morales-Durand, H., Ross, P., Cheng, H., Giovambattista, G. and Demyda-Peyrás, S., 2018. Runs of homozygosity in a selected cattle population with extremely inbred bulls: Descriptive and functional analyses revealed highly variable patterns. *PLoS One*, 13(7), p.e0200069.
10. Xu, Z., Mei, S., Zhou, J., Zhang, Y., Qiao, M., Sun, H., Li, Z., Li, L., Dong, B., Oyelami, F.O. and Wu, J., 2021. Genome-wide assessment of runs of homozygosity and estimates of genomic inbreeding in a chinese composite pig breed. *Frontiers in genetics*, 12, p.720081.
11. Broman, K.W. and Weber, J.L., 1999. Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. *The American Journal of Human Genetics*, 65(6), pp.1493-1500.
12. Zhang, Q., Guldbandsen, B., Bosse, M., Lund, M.S. and Sahana, G., 2015. Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC genomics*, 16, pp.1-16.
13. <https://www.neogen.com/en-gb/>, Neogen, (n.d.).
14. Ciani, E., Lasagna, E., D'andrea, M., Alloggio, I., Marroni, F., Ceccobelli, S., Delgado Bermejo, J.V., Sarti, F.M., Kijas, J., Lenstra, J.A. and Pilla, F., 2015. Merino and Merino-derived sheep breeds: a genome-wide intercontinental study. *Genetics Selection Evolution*, 47(1), pp.1-12.
15. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. and Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, 4(1), pp.s13742-015.
16. Filippo, B., Paolo, C., Gaspa, G. and Gabriele, M., 2019. Detect Runs of Homozygosity and Runs of Heterozygosity in Diploid Genomes.
17. Team, R.D.C., 2009. A language and environment for statistical computing. <http://www.R-project.org>.
18. Dzomba, E.F., Chimonyo, M., Pierneef, R. and Muchadeyi, F.C., 2021. Runs of homozygosity analysis of South African sheep breeds from various production systems investigated using OvineSNP50k data. *BMC genomics*, 22(1), pp.1-14.
19. Ceballos, F.C., Joshi, P.K., Clark, D.W., Ramsay, M. and Wilson, J.F., 2018. Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics*, 19(4), pp.220-234.
20. Granero, A., Anaya, G., Demyda-Peyrás, S., Alcalde, M.J., Arrebola, F. and Molina, A., 2022. Genomic Population Structure of the Main Historical Genetic Lines of Spanish Merino Sheep. *Animals*, 12(10), p.1327.
21. Zhang, Q., Guldbandsen, B., Bosse, M., Lund, M.S. and Sahana, G., 2015. Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC genomics*, 16, pp.1-16.
22. Bosse, M., Megens, H.J., Madsen, O., Paudel, Y., Frantz, L.A., Schook, L.B., Crooijmans, R.P. and Groenen, M.A., 2012. Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. *PLoS genetics*, 8(11), p.e1003100.
23. Megdiche, S., Mastrangelo, S., Ben Hamouda, M., Lenstra, J.A. and Ciani, E., 2019. A combined multi-cohort approach reveals novel and known genome-wide selection signatures for wool traits in Merino and Merino-derived sheep breeds. *Frontiers in Genetics*, 10, p.1025.