



## OPEN ACCESS

## EDITED BY

Klaus Wimmers,  
Leibniz Institute for Farm Animal Biology  
(FBN), Germany

## REVIEWED BY

Este van Marle-Köster,  
University of Pretoria, South Africa  
Ottmar Distl,  
University of Veterinary Medicine  
Hannover, Germany

## \*CORRESPONDENCE

Sowah Addo,  
sowah.addo@uni-kassel.de

## SPECIALTY SECTION

This article was submitted to Livestock  
Genomics, a section of the journal  
Frontiers in Genetics

RECEIVED 31 March 2022

ACCEPTED 30 September 2022

PUBLISHED 18 October 2022

## CITATION

Addo S and Jung L (2022), An insight  
into the runs of homozygosity  
distribution and breed differentiation in  
Mangalitsa pigs.  
*Front. Genet.* 13:909986.  
doi: 10.3389/fgene.2022.909986

## COPYRIGHT

© 2022 Addo and Jung. This is an open-  
access article distributed under the  
terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which does  
not comply with these terms.

# An insight into the runs of homozygosity distribution and breed differentiation in Mangalitsa pigs

Sowah Addo\* and Lisa Jung

Animal Breeding Section, University of Kassel, Witzenhausen, Germany

Mangalitsa pigs exhibit three distinct coat color patterns based on which they are described as Red, Blond, and Swallow-bellied. The current study investigated genome-wide diversity and selection signatures in the three breeds using fixation index, runs of homozygosity and population structure analyses. The analyses were originally based on quality-controlled data on 77 Mangalitsa animals from Germany, including 23 Blond, 30 Swallow-bellied and 24 Red Mangalitsa genotyped with a customized version of the ProcineSNP60 v2 Genotyping Bead Chip. Also, 20 Hungarian Mangalitsa genotypes were included as outgroup data for comparison. Estimates of observed heterozygosity were 0.27, 0.28, and 0.29, and inbreeding coefficients estimated based on runs of homozygosity were 24.11%, 20.82%, and 16.34% for Blond, Swallow-bellied and Red Mangalitsa, respectively. ROH islands were detected in all breeds, however, none of these were shared amongst them. The KIF16B gene previously reported to play a role in synaptic signaling was found in a ROH island (SSC17: 16–26) in Swallow-bellied Mangalitsa. The same gene was found to harbor a significantly differentiated SNP (MARC0032380) while contrasting either Blond or Red to Swallow-bellied Mangalitsa. In the Red Mangalitsa, some ROH islands were associated with genes that play a role in meat quality traits, i.e., ABCA12, VIL1, PLSCR5, and USP37. Our population structure analysis highlighted a separation of the three breeds, but also showed the closest relatedness between Red and Blond Mangalitsa pigs. Findings of this study improve our understanding of the diversity in the three breeds of Mangalitsa pigs.

## KEYWORDS

Mangalitsa pig, ROH islands, genetic diversity, selection signatures, runs of homozygosity

## Introduction

Domestication and selection events can lead to both favorable and unfavorable allele reconstitution in animal species. Following its creation by domestication of wild pigs (*Sus scrofa ferus*) in the 19th century, Sumadija pigs of Serbia were reared under favorable conditions that made them lose their original form to become one of the progenitors of

Mangalitsa pigs (Jukes 2017). Although originally from Serbia, Mangalitsa was systematically developed in Hungary at a time when market demand for good quality fat, bacon and less fibrous meat necessitated the crossing of small Hungarian sows such as Alföldi, Bakony and Szalonta with Serbia's improved Sumadija (Egerszegi et al., 2003; Jukes 2017). The product of such crosses was the Blond Mangalitsa, which was subsequently crossed either with Black Sarmian to develop Swallow-bellied Mangalitsa or with Szalonta to develop Red Mangalitsa (Zsolnai et al., 2013; Jukes 2017). Thus, three types of Mangalitsa pigs exist and these exhibit varying phenotypes, particularly coat color variation. In spite of the differences, Mangalitsa pigs are broadly described as fat-type, curly-haired pigs with strong motherliness and adaptability, but low reproductive performance (Egerszegi et al., 2003). They have geographical predominance in Hungary but are also distributed across Serbia, Romania, Austria, Croatia, Germany and Switzerland.

The three breeds of Mangalitsa are usually managed with restricted gene flow amongst them. Zsolnai et al. (2006) investigated the genetic relationships between the breeds using 10 microsatellite markers, and proposed a rejection of the hypothesis that Mangalitsa individual form just one unpartitioned population. They also found Blond and Swallow-bellied Mangalitsa to be genetically closer to each other than to Red Mangalitsa as did Marincs et al. (2013) who based their analysis on mitochondrial DNA D-loop sequences. Meanwhile, in a separate study, a conclusion has been drawn that mitochondrial DNA D-loop polymorphism could not distinguish between the three breeds (Molnár et al., 2013). Marincs et al. (2013) also unraveled the presence of both common European and Mangalitsa-specific mitochondrial DNA D-loop haplotypes in a Hungarian population and concluded that these pigs may have originated either by introgression of common European bloodline into the Mangalitsa breed or by isolation of some Mangalitsa ancestor species. Furthermore, Frank et al. (2017), for the first time found the mitogenomes of some Mangalitsa animals to be highly related to the Croatian Turopolje breed, which they attributed to either common origin of maternal lineages or admixture events. A comparative study of whole genome sequences of the three breeds of Mangalitsa and a Duroc pig highlighted on one hand, 52 Mangalitsa-specific genes involved in lipid metabolic processes, and on the other hand, several exonic polymorphisms unique to each of the breeds (Molnár et al., 2014). More recently, an investigation into the genetic basis of different colorations in Mangalitsa pigs revealed a display of signature of divergent directional selection in the solute carrier family 45-member 2 (SLC45A2) gene for the comparison of Red and Blond Mangalitsa breeds (Bâlțeanu et al., 2021).

In spite of their genetic differentiation, there was no mentioning of breed type in a number of studies involving Mangalitsa pigs, and it was not immediately clear if animals of different coat colors were used (García et al., 2006; Wilkinson et al., 2013; Herrero-Medrano et al., 2014; Manunza et al., 2016;

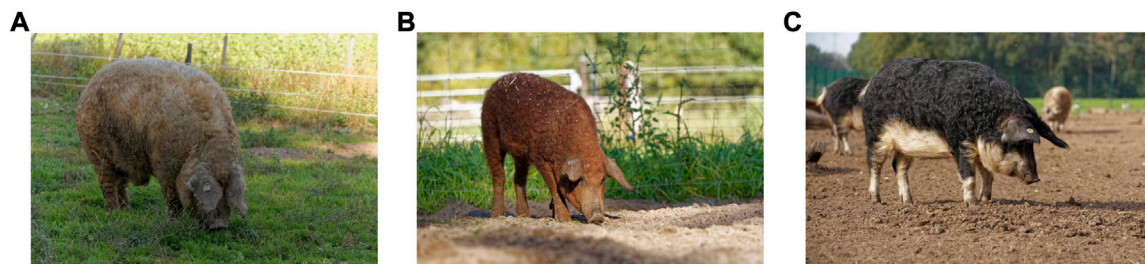
Schachler et al., 2020). In one of such studies, the mean genomic inbreeding coefficient estimated for 20 Hungarian Mangalitsa animals based on runs of homozygosity was high (0.22) for which reason the authors advocated for special conservation interventions to be put in place (Yang et al., 2017). Gorssen et al. (2021) recently presented a repository of ROH island for several breeds of eight animal species among which the Hungarian Mangalitsa pigs used in Yang et al. (2017) were featured. The repository shows high incidence of ROH occurring on chromosomes 11, 13, 14, and 17 in the Mangalitsa pigs. Bâlțeanu et al. (2019) found runs of homozygosity based inbreeding coefficient ranging from 0.09 to 0.14 in different populations of Mangalitsa pigs and argued that the pattern of homozygosity in these local breeds is comparable to those of the majority of cosmopolitan breeds. Albeit, they found a clear indication of strong and recent inbreeding in Romanian and Hungarian Red Mangalitsa pigs which was attributed to mating of related individuals and a reduction in population size. In Germany, Mangalitsa pigs are referred to as "Wollschweine", and an attempt to preserve their genetics is evidenced by the naming of Mangalitsa as breed of the year 1999 (Flegler 1999). Nowadays, there is growing interest of German breeders in breeding all three types of Mangalitsa, while the Society for the Conservation of Old and Endangered Livestock Breeds (GEH e.V.) is rapidly promoting the establishment of a herd book for the Mangalitsa pigs in Germany.

The availability of different Mangalitsa genotypes that exhibit phenotypic variability offers new possibilities of studying signatures of selection that may have played a role in the development of the three breeds. Besides, the ROH landscape of these breeds can be conveniently compared to those of other breeds published in a novel ROH repository to improve our understanding of selective sweeps in pigs. Therefore, the present study aimed at 1) investigating genome-wide relatedness among Blond, Red and Swallow-bellied Mangalitsa pigs, 2) identifying genomic regions with high level of inbreeding termed ROH islands, and 3) finding candidate genes that may be associated with significantly differentiated genomic regions in the breeds.

## Materials and methods

### Animal description

In this study, 109 animals belonging to the three main breeds of Mangalitsa pigs were initially considered. These include Blond, Red and Swallow-bellied Mangalitsa, hereafter referred to as BM, RM, and SM, respectively. The animals were from individual breeders and animal parks, and each had a registration number provided by the GEH e.V. in Germany. Animals were distinguishable, predominantly by their coat-color variation (Figure 1). BM individuals have a general grey to yellow to yellowish red coat-color with the head and leg regions often



**FIGURE 1**  
The three different breeds of Mangalitsa pigs including Blond (A), Red (B) and Swallow-bellied (C). Pictures were provided by Rudi Gosmann.

almost black (Figure 1A). RM tend to have darker or lighter shades of reddish-brown (Figure 1B) while SM have a blackish-brown coloration at the back and flanks and yellow, white or silvery grey at the underside, belly and cheek areas (Figure 1C). No extensive pedigree data were available at the time of sample collection, and a number of animals were said to have been produced by crossing two of the three breeds. Due to issues of misidentification and uncertainty, these animals, totaling 17, were labeled as mixed breed (MM) and removed from the analyses. The remaining animals included 34 SM, 29 RM, and 29 BM pigs. Additionally, we included data on 20 Hungarian Mangalitsa pigs (HUMA) of unspecified breed type from previous studies (Yang et al., 2017) for comparison.

## Genotyping and quality control

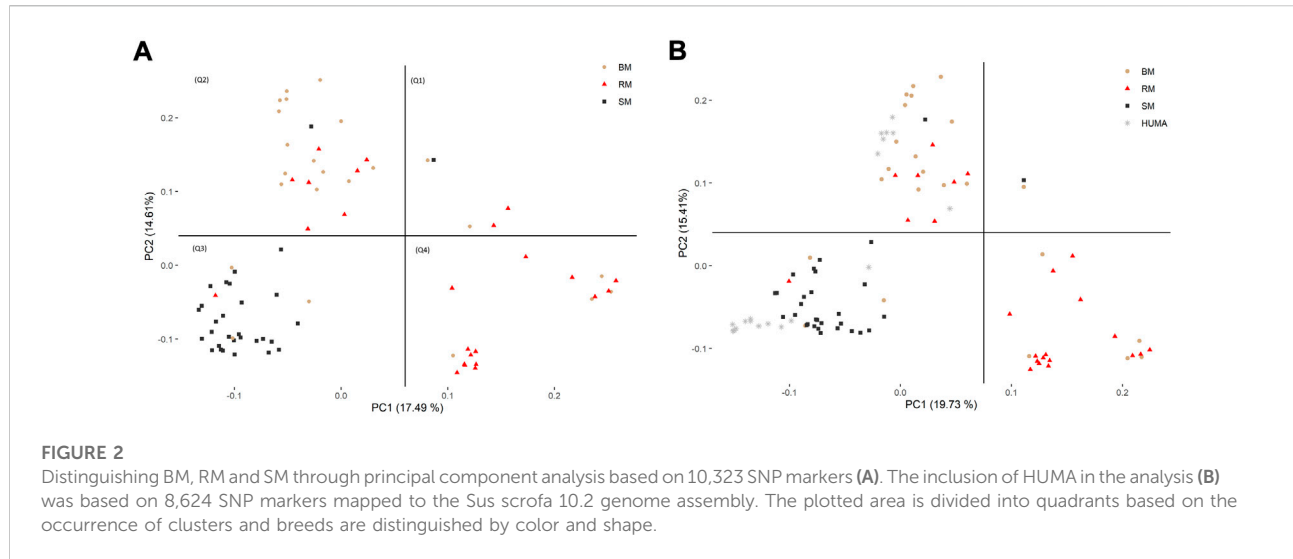
The collection of hair samples, DNA extraction and genotyping followed standard procedures and were performed in two batches, in 2018 and 2020. Animals were genotyped with customized versions of the ProcineSNP60 v2 Genotyping Bead Chip. SNP markers common to both batches were extracted and mapped to the *Sus scrofa* 11.1 genome assembly. A number of quality control procedures were conducted in PLINK v 1.9 (Chang et al., 2015) depending on the desired type of analysis. Unmapped and non-autosomal SNPs were broadly removed as were 2 (SM) and 1 (RM) animals whose genomic relationship coefficient with other pairs exceeded 0.95. Individual and marker genotyping rate thresholds were both set to 90%, and SNPs with minor allele frequency (MAF) lower than 0.05 or that deviated from Hardy Weinberg Equilibrium (HWE:  $10^{-6}$ ) were removed. Specifically, for ROH analyses, SNP call rate was set to 95% for easy comparison (Gorssen et al., 2021) and there was neither MAF nor HWE pruning as recommended in previous studies (Meyermans et al., 2020). MAF pruning was also not conducted on data destined for Fixation index ( $F_{st}$ ) analysis (Weir and Cockerham 1984) and the quality

control steps were performed for each breed separately. Furthermore, a minor linkage disequilibrium (LD) pruning was applied to the dataset destined for population structure analysis by using the PLINK (Chang et al., 2015) command “--indep-pairwise 50 25 0.5”. After all filtering steps, 23 BM, 24 RM, and 30 SM genotypes remained for further analyses.

## Diversity and population structure

Within-breed genetic diversity was investigated based on observed heterozygosity estimates calculated as a difference between the number of homozygous and non-missing genotypes, expressed as a proportion of non-missing genotypes. Also, the relationship between BM, RM, and SM was investigated using principal component analyses. LD-pruned SNPs totaling 10,323 were used to compute Plink-based (Chang et al., 2015) eigenvectors for the first 20 components for each individual. Subsequently, eigenvectors of the first (PC1) and second (PC2) principal components were visualized in R (R Core Team 2020) using breed as color code. This investigation was further expanded to include the 20 HUMA outgroup data and the analysis was based on 8,624 quality-controlled SNPs common to all animals. The population structure was also studied using ADMIXTURE (Alexander et al., 2009). For this, a 5-fold cross-validation procedure was performed for a range of  $k$  between 1 and 20, and the  $k$  with the lowest cross-validation error was considered as the optimal number of clusters for the data. Cluster assignments ranging from  $k = 2$  to  $k = 9$  were visualized using Pophelper (Francis 2017).

To detect genetic differentiation over time, the  $F_{st}$  (Weir and Cockerham 1984) between breeds was calculated for all loci using PLINK (Chang et al., 2015). Subsequently, an empirical  $p$ -value for each SNP was estimated following a  $z$ -score calculated from the distribution of  $F_{st}$  values. The  $F_{st}$  values were visualized using Manhattan plots implemented in the qqman R package (Turner 2018), and the top 0.1% were suggested as signatures of selection.



## ROH calling and analysis

The number of SNPs available for ROH analyses after quality control was 36,617 (BM), 38,085 (RM), and 36,964 (SM). ROH calling was performed using an R-script developed for standardized breed-by-breed quality control and analysis (Gorssen et al., 2021), which is available at Open Science Framework (<https://doi.org/10.17605/OSF.IO/XJTKV>). Default parameter settings as described in Gorssen et al. (2021) were followed in defining ROH, ROH incidence and ROH islands. Briefly, we considered a sliding window with minimal number of SNPs determined by an L-parameter (Purfield et al., 2012; Meyermans et al., 2020). Within this window, 1 SNP could be missing but no heterozygous SNP was allowed, and there was at least 1 SNP every 150 kb. Additionally, the maximal gap between two consecutive homozygous SNPs was set to 1,000 kb.

Genomic inbreeding coefficient ( $F_{ROH}$ ) was calculated for each animal as the total length of all ROH in the genome of an individual expressed as a proportion of the length of autosomal genome coverage expressed by SNPs in the analysis (McQuillan et al., 2008).  $F_{ROH}$  was calculated considering all ROH ( $F_{ROH\_all}$ ) in an individual but also for different ROH length categories including 1–2 Mb ( $F_{ROH1-2}$ ), 2–4 Mb ( $F_{ROH2-4}$ ), 4–8 Mb ( $F_{ROH4-8}$ ), 8–16 Mb ( $F_{ROH8-16}$ ) and >16 Mb ( $F_{ROH16}$ ).

For a given breed, the percentage of individuals with a specific SNP in ROH was defined as ROH incidence. From the distribution of ROH incidences, a threshold ( $p$ -value) was calculated based on standard normal  $z$ -scores. ROH islands were defined as the top 0.1% of SNPs with a  $p$ -value higher than 0.999 using a  $z$ -score table for ROH incidence (Purfield et al., 2017; Gorssen et al., 2020). Finally, a ROH must be present in at least 30% of the population to be part of a ROH island. ROH incidences and thresholds were visualized for

each breed *via* Manhattan plots using the qqman package (Turner 2018).

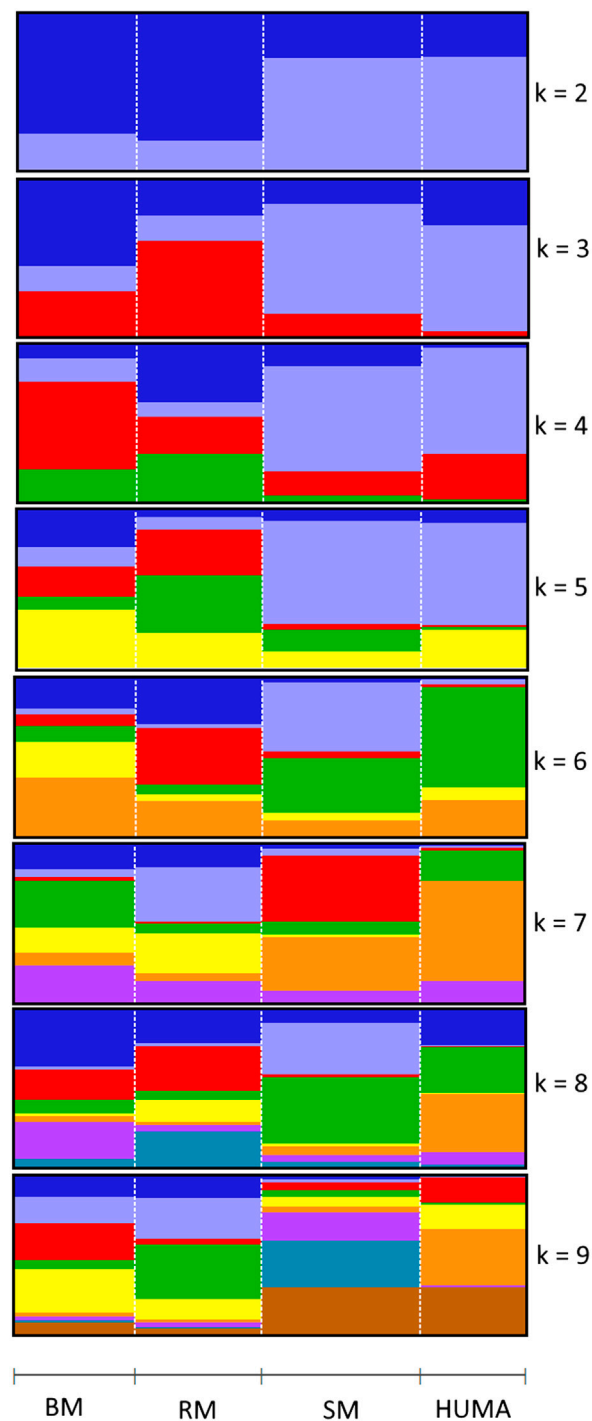
## Identification of candidate gene

Significant genome variants from both  $F_{st}$  and ROH analyses were annotated to the *Sus scrofa* genome version 11.1 reference assembly, and the “*susrofa\_gene\_ensembl*” dataset was explored using biomaRt v. 2.50.3 (Durinck et al., 2009). Furthermore, genes within 100 kb distance on either side of these variants were identified as candidate genes that may have played a role in the development of the breeds.

## Results

### Population structure and fixation index

In the principal component analysis, PC1 and PC2 together explained 32.1% of variation in the three breeds (Figure 2A). There was a clustering along the lines of breeds, but notably, clusters were not well defined. In the third quadrant (Q3) of the plotted area, SM was predominantly separated from RM and BM by PC1 and PC2, respectively. In the second quadrant, what seemed to be a BM cluster harbored several genotypes of RM. The RM breed formed a cluster in the fourth quadrant with the highest level of dispersion. This latter cluster also harbored genotypes of BM. By including HUMA genotypes in the analysis (Figure 2B), 35.14% of the variation in the data was explained. Majority of the HUMA clustered close to SM, and a few were in the

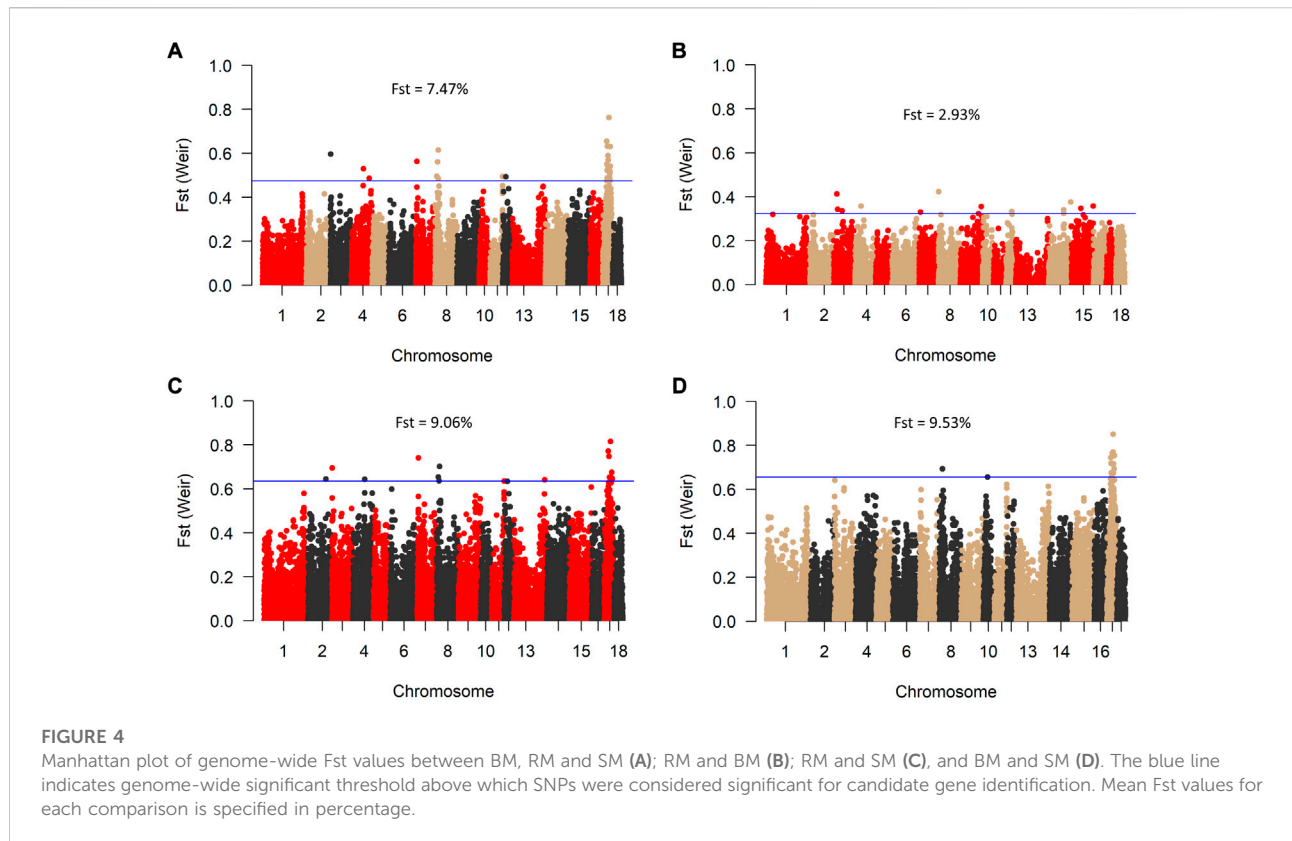


**FIGURE 3**  
Admixture analysis of four Mangalitsa pigs populations (BM, RM, SM and HUMA) with graphs representing cluster levels 2 through 9. The optimal cluster level was found at  $k = 6$ .

cluster of BM. The ADMIXTURE analysis for low levels of  $k$ , especially,  $k = 2$  showed a high degree of similarity of genetic background between BM and RM on one hand, and on the other hand, similarity between SM and HUMA (Figure 3).

However, there was no complete separation of all breeds as demonstrated by traces of admixture at all cluster levels including  $k = 6$ , which produced the lowest cross-validation error estimate (Supplementary Figure S1).





Additionally, genetic background was highly heterogeneous in all breeds.

Consistent with the principal component analysis, the lowest mean  $F_{st}$  (0.029) was found between RM and BM, while these two breeds were relatively distant from SM (Figure 4). By comparing all three breeds in a single  $F_{st}$  analysis, 30 genome-wide significant variants were detected across SSC 3, 4, 7, 8, 11, 12, and 17. Most of the significant variants (67%) were located on SSC17, which also had the topmost SNP (DRGA0016741) at position 38908507. The restriction of the analysis to a pairwise comparison revealed 14 significant variants between RM and BM with none occurring on chromosome 17. A number of significant variants were common to the comparison between RM and SM and between BM and SM. Five of these variants are depicted in Table 1 with superscript letters a–e. Additionally, Table 1 provides candidate genes located within 100 kb on either side of the  $F_{st}$ -based significant variants.

## Within-breed diversity

Average  $H_o$  estimates were 0.27, 0.28 and 0.29 for BM, SM and RM, respectively. The heterozygosity estimates correlated significantly and negatively with total genomic inbreeding ( $r =$

$-0.88$ ;  $p$ -value  $< 2.2e-16$ ) as shown in Figure 5. Across breeds,  $H_o$  ranged from 0.12 at the highest level of  $F_{ROH}$  (67.25%) to 0.42 at the lowest inbreeding level (0.39%).

The computed values of  $F_{ROH}$  for various ROH length considerations are presented with respect to breed in Figure 6. Considering all ROH, average inbreeding was highest in BM (24.11%) and lowest in RM (16.34%). This is also true for other ROH length considerations. Larger ROH segments played a major role in autozygosity in all breeds, and ROH segments below 2 Mb in size were completely absent in all animals. In general, fewer number of ROH segments were found in RM than in BM and SM (Table 2). Although BM and SM had about the same average number of ROH segments, the former had the highest sum of ROH (543626.6 kb) per individual. Across all breeds, 2715 ROH segments were found in this study.

The incidence of ROH and ROH islands varied across breeds (Figure 7). BM and SM had higher baseline ROH levels with ROH islands detected on SSC16 in BM, and on SSC11 and SSC17 in SM. The most significant variant in ROH island was found in 24 out of 30 SM animals on SSC17. In RM, ROH incidence levels were generally low, however, ROH islands were found on SSC7, SSC13 and SSC15. ROH incidence plots per chromosome show seemingly similar patterns across breeds (Supplementary Figures S2–S4). Furthermore, candidate genes proximal to ROH islands are presented in Table 3. Significant

**TABLE 1** Summary information of significant SNPs and candidate genes from  $F_{ST}$  comparison of three Mangalitsa pig breeds (BM, SM and RM). Superscripts a-e denote common SNPs across pairwise comparisons.

Breeds	CHR	SNP	Position	$F_{ST}$	Candidate genes	
BMSM	17	DRGA0016741 <sup>a</sup>	38908507	0.85	CEP250, ERGIC3, SPAG4, CPNE1, NFS1	
	17	MARC0026961	39771108	0.77	DLGAP4, TGIF2, MYL9	
	17	MARC0017379	33921649	0.76	SIRPB2, NSFL1C, SDCBP2, SNPH	
	17	ALGA0095426 <sup>b</sup>	46764111	0.75	OSER1, GDAP1L1, FITM2, R3HDML, HNF4A	
	17	MARC0032380 <sup>c</sup>	25017561	0.74	KIF16B	
	17	H3GA0048218	26643149	0.72	ZNF133, DZANK1, POLR3F, RBBP9	
	17	ALGA0094584 <sup>d</sup>	33341575	0.72	STK35	
	17	H3GA0049036	41739085	0.72	SLC32A1, ACTR5, PPP1R16B	
	8	ALGA0046856 <sup>c</sup>	20576000	0.69	STIM2	
	17	ASGA0077154	45575325	0.69	PTPRT	
	17	ALGA0094114	27748357	0.68	SLC24A3, RIN2	
	17	ALGA0095121	41180101	0.67	RPRD1B, TGM2, KIAA1755	
	17	ASGA0076251	30236910	0.66	THBD, CD93	
	10	H3GA0029711	26091506	0.66	ERCC6L2	
	RMSM	17	DRGA0016741 <sup>a</sup>	38908507	0.82	CEP250, ERGIC3, SPAG4, CPNE1, NFS1
		17	MARC0032380 <sup>c</sup>	25017561	0.77	KIF16B
		7	ALGA0038333	7188826	0.74	TFAP2A
8		ALGA0046856 <sup>c</sup>	20576000	0.70	STIM2	
3		H3GA0008443	2734718	0.70	SDK1	
17		ALGA0095426 <sup>b</sup>	46764111	0.68	OSER1, GDAP1L1, FITM2, R3HDML, HNF4A	
17		ALGA0094584 <sup>d</sup>	33341575	0.65	STK35	
17		ASGA0077190	46189837	0.65	SRSF6, SGK2, IFT52, MYBL2	
17		M1GA0022187	49729457	0.65	NCOA3, SULF2	
2		H3GA0007369	113554653	0.64	FBXL17	
4		ALGA0026041	75779884	0.64	PLAG1, MOS, RPS20, LYN	
13		M1GA0017756	193064698	0.64	GRIK1	
8		ALGA0046640	16400498	0.64	KCNIP4	
RMBM	8	M1GA0011680	2332002	0.42	ADRA2C	
	3	ALGA0017963	20715185	0.41	HS3ST4	
	4	MARC0024274	39449105	0.36	TSPYL5, CPQ	
	15	MARC0036536	137538236	0.36	RBM44, RAMP1, SCLY	
	9	ALGA0055521	130209873	0.35	RPS6KC1	
	15	ALGA0107554	60110316	0.35	FMNL2	
	3	ALGA0018104	25852288	0.34	GPR139, GPRC5B, IQCK	
	3	ALGA0019015	53054275	0.34	CREG2, RNF149, CNOT11, TBC1D8	
	12	MARC0039239	35751427	0.33	DHX40, CLTC	
7	ASGA0031021	8122178	0.33	NEDD9, TMEM170B		

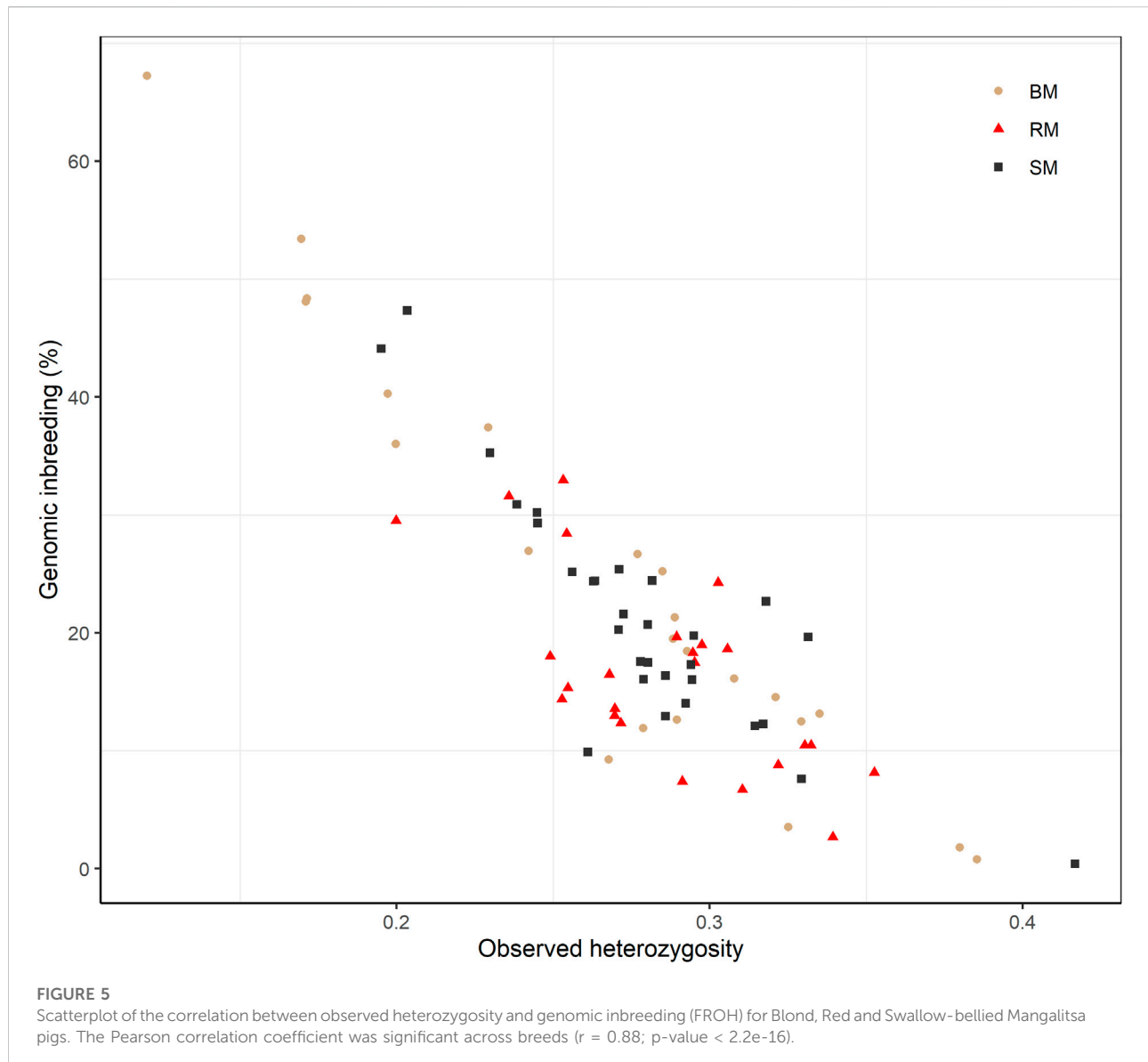
variants with no genes found within 100 kb distance on either side were not included.

In the separate analysis, where SNP positions were based on *Sus scrofa* genome version 10.2, ROH islands were detected on SSC7 (113–117 Mb) and SSC15 (125–134 Mb) in RM, and on SSC11 (80–86 Mb) and SS17 (17–29 Mb) in SM (Supplementary Table S1). For BM, no SNP reached a  $p$ -value > 0.999 for ROH due to the breed's high level of inbreeding. Therefore, we set a maximal threshold of 55%

( $p$ -value  $\geq$  0.998) for ROH island detection. Hence, ROH islands were found on SSC7 (104–106 Mb), SSC8 (29–31 Mb) and SSC16 (21–24 Mb; 60–62 Mb) in BM.

## Discussion

The Mangalitsa pig with its dense curly hair and forward falling ears is considered precious, offering advantages such as



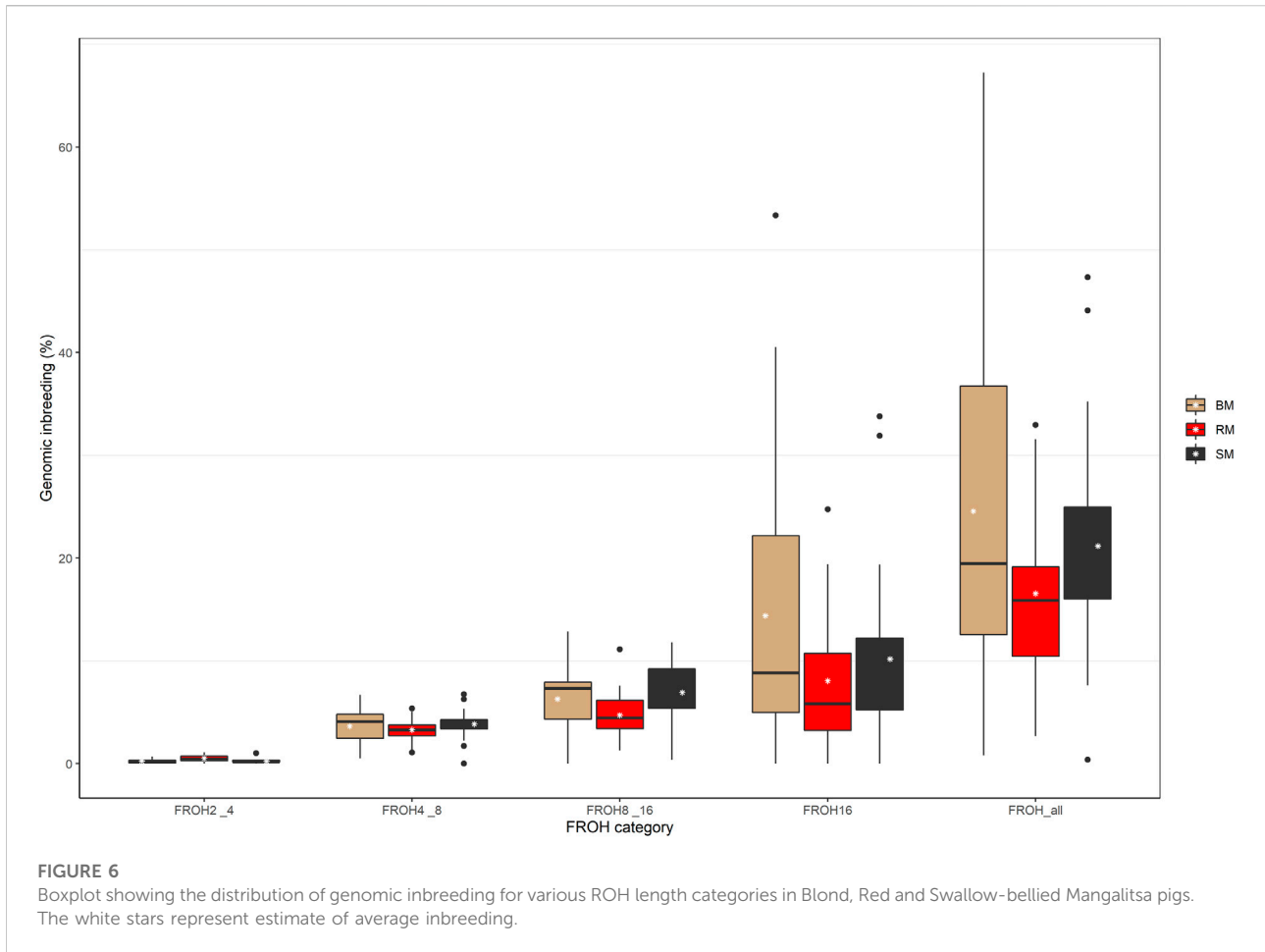
resistance to challenging weather conditions and diseases (Egerszegi et al., 2003). Also, it is not very demanding on housing conditions and feeding. Therefore, breeders in Balkan countries, Austria, Switzerland, and Germany aim to conserve the original Mangalitsa without deleterious effects of inbreeding (Egerszegi et al., 2003). There is a growing interest in the exploration of the genome resource of Mangalitsa pigs, relying on an array of genetic markers including microsatellites (Zsolnai et al., 2006; Druml et al., 2012; Kharzinova and Zinovieva 2020), mitochondrial DNA D-loop sequences (Marincs et al., 2013; Molnár et al., 2013; Frank et al., 2017), SNPs (Bâlțeanu et al., 2019; Bâlțeanu et al., 2021; Zorc et al., 2022) and whole genome sequences (Molnár et al., 2014). Previous studies have attempted to discriminate between Mangalitsa pigs of different colors, highlighting a rejection of the hypothesis that Mangalitsa

individuals belong to one indistinguishable group (Zsolnai et al., 2006). The current study did not only investigate genetic differentiation between the breeds, but also probed the occurrence of candidate genes in the vicinity of significantly differentiated marker loci. Additionally, we listed genes that are proximal to genomic regions displaying high levels of autozygosity. Whether or not these genes are differentially expressed in the studied breeds remains a subject for discussion.

## Population structure and diversity

The high degree of clustering of individual Mangalitsa pigs by breed in this study (Figure 2) is a notable observation echoing the findings of previous studies. Based on 10 microsatellite markers,



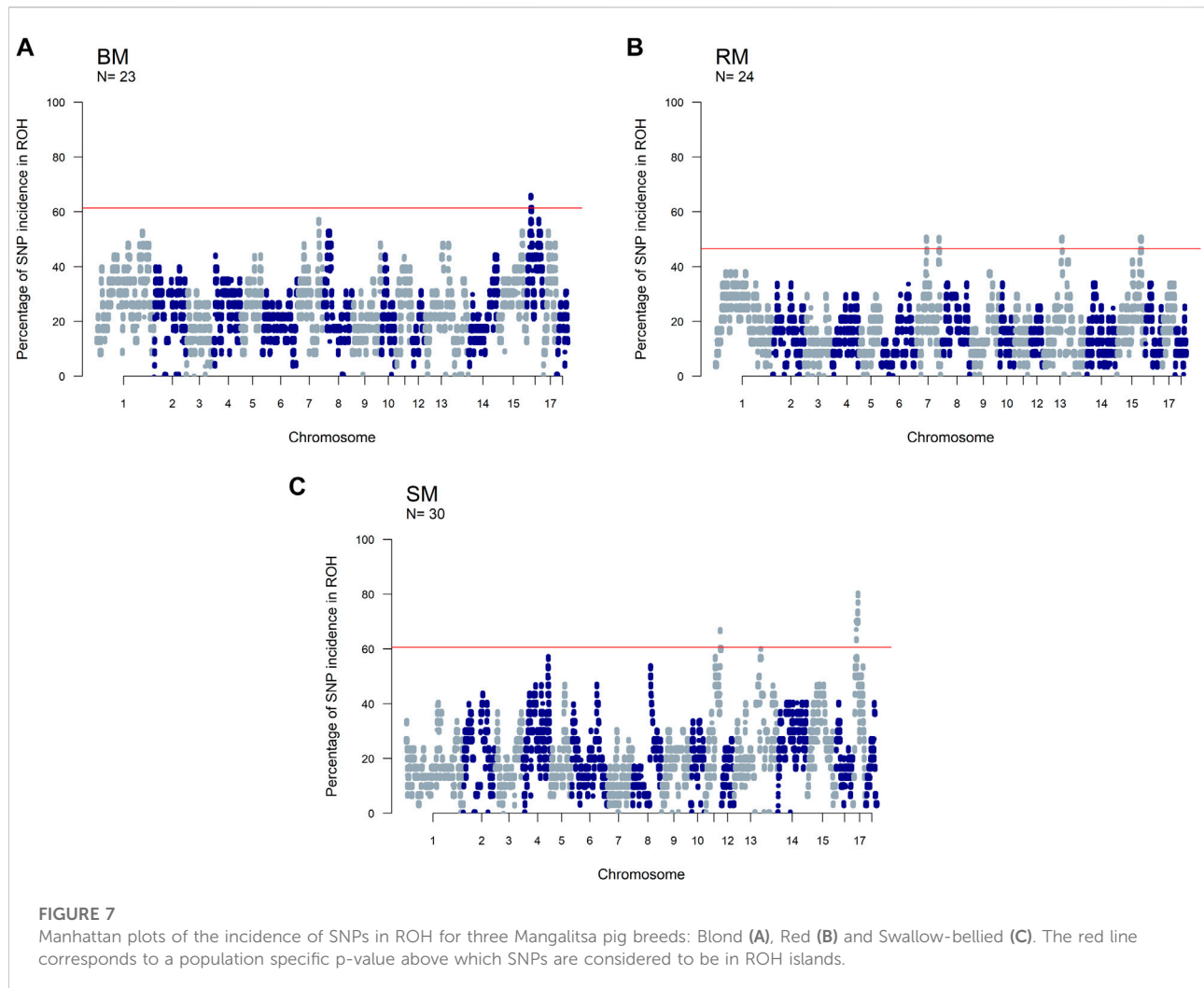


**TABLE 2** The average and total number of ROHs, and average sum of ROH length in BM, RM, SM and across all three breeds (ALL). Minimum and maximum values are provided in brackets.

Breed	Mean Nr. ROH (min + max)	Total Nr. ROH	Mean $\sum$ ROH in kb (min + max)
BM	37.00 (3–64)	851	543626.6 (17568–1487860)
RM	30.83 (9–50)	740	368431.3 (59347–733062)
SM	37.47 (1–56)	1,124	469469.6 (8,582–1049150)
ALL	35.26 (1–64)	2,715	460128.0 (8,582–1487860)

Zsolnai et al. (2006) separated the three breeds *via* genetic distance and Structure analyses. Bălteanu et al. (2019) also reported some degree of differentiation between the three breeds and other pig types through Admixture and principal component analysis using the Porcine SNP60 Beadchip. The persistency of traces of common genetic background across all cluster levels in our study reflects the developmental history of the breeds—RM and SM derived from BM. This is also consistent with previous studies that partly suggest that a present-day Mangalitsa population evolved by introgression of other

European breeds and wild boars (Marincs et al., 2013). Explaining their observation of a close relationship between Mangalitsa and Turopolje pigs, Frank et al. (2017) mentioned a common wild boar, the Siska pig, assumed to be amongst the ancestors of both breeds. Another study suggests a possible geneflow from Slavonian Black and Pietrain pigs into both Mangalitsa and Turopolje pig populations, and reported lower levels of differentiation (from  $F_{st} = 0.05$  to  $F_{st} = 0.09$ ) between Mangalitsa and Slavonian Black (Druml et al., 2012). The closeness between Mangalitsa (SB) and Slavonian Black was



confirmed by Zorc et al. (2022) who also showed potential gene flow between SB and Moravka pigs using both microsatellite and SNP data. What remains inconsistent across studies is the degree of differentiation among the breeds. We found in the current study, close connectedness between BM and RM than each of them to SM, which was corroborated by our  $F_{st}$  analysis, i.e., 0.029 (BMRM) vs. 0.091 (RMSM) or 0.095 (BMSM). On the contrary, Zsolnai et al. (2006) found BM and SM to be closer to each other ( $F_{st} = 0.064$ ) than each of them to RM ( $F_{st}$  equals 0.094 and 0.099, respectively). Similarly, Marincs et al. (2013) found lower estimates of both  $F_{st}$  and Nei's distance for BMSM than for either BMRM or RMSM. Their findings are consistent with the observation of a high degree of similarity of genetic background between BM and SM at low cluster levels in Bălțeanu et al. (2019). Differences between our findings and those of previous studies could stem from differences in breeding focus and management across populations since the genotypes in the

previous study were of Hungarian and Romanian origin. It is therefore, not surprising that the 20 HUMA genotypes, which we later included in our analysis, formed clusters predominantly with SM, and with BM to a lesser extent. Thus, we suspect the HUMA outgroup genotypes to belong to the SM breed, however, this cannot be immediately confirmed. Albeit, we consider the high resemblance between BM and RM (our finding) symbolic of the crossing of BM with Szalonta to develop RM (Zsolnai et al., 2013; Jukes 2017). Worth mentioning is that Szalonta is already one of the progenitors of BM (Egerszegi et al., 2003; Jukes 2017).

The dispersion of RM genotypes in our analysis reveals either a comparatively high level of genetic diversity or population substructures in this breed. Consistently, RM had the highest heterozygosity estimate (0.29) in our study. Even so, for the same breed types, observed heterozygosity estimates in previous studies are slightly higher than in ours, i.e., 0.32–0.38 vs. 0.29 for RM, 0.31 vs. 0.27 for BM and 0.29 vs. 0.28 for SM

TABLE 3 Summary information of significant SNPs and candidate genes from ROH islands in three Mangalitsa pig breeds (BM, SM and RM).

Breed	CHR	SNP	Position	ROH (%)	Candidate genes	
SM	17	MARC0086847	25118032	80.08	KIF16B	
	17	H3GA0048106	23564839	73.89	MACROD2	
	17	MARC0030901	24049827	73.54	MACROD2	
	17	ALGA0093805	22510190	70.63	ESF1, NDUFAF5, SEL1L2	
	17	MARC0109531	20937630	70.63	BTBD3	
	17	H3GA0048183	26170282	70.40	RRBP1, BANF2	
	17	H3GA0047989	17680444	63.92	PLCB1	
	11	H3GA0032842	77609850	60.66	ARHGEF7	
	11	DRGA0011540	74100734	60.65	EFNB2	
	BM	16	MARC0086770	20525463	65.77	RAI14, BRX1, RAD1, AGXT2, PRLR
		16	H3GA0046210	21372379	61.39	IL7R, CAPSL, LMBRD2, SKP2
16		ALGA0115908	22998327	60.99	WDR70, GDNF	
RM	15	DRGA0015508	120520780	50.81	CATIP, SLC11A1, CTDSP1, VIL1, USP37, CNOT9	
	15	CASI0006663	117450632	50.81	ABCA12	
	15	MARC0003725	119092405	50.81	TNP1	
	15	ALGA0086843	114044775	50.80	ERBB4	
	7	DIAS0001981	46539039	50.80	TMEM14A, GSTA1, GSTA4	
	13	ASGA0058507	87695642	50.73	ZIC4, ZIC1	
	13	H3GA0037098	89899499	50.72	WWTR1, COMMD2, ANKUB1	
	7	ALGA0041512	47324213	50.71	CHRNA3, CHRNA5	
	13	ALGA0071380	86992936	50.70	PLSCR5	
	7	ASGA0033712	48095614	50.57	MORF4L1, CTSH, RASGRF1	
	15	ASGA0070454	116910010	50.38	VWC2L	
	15	DRGA0015443	113969520	49.67	ERBB4	
	13	CASI0008610	90639691	46.65	TSC22D2	
	7	DIAS0001022	110412608	46.65	SPATA7, PTPN21, ZC3H14, EML5,	
	15	MARC0014079	115472905	46.50	IKZF2	
	7	ALGA0044724	111026023	46.45	FOXN3	
7	ALGA0041861	49748787	46.44	CEMIP, MESD, TLNRD1		

(Bălțeanu et al., 2019). Higher heterozygosity values in Bălțeanu et al. (2019) were associated with relatively low levels of inbreeding. Thus, across the three breeds, average  $F_{ROH}$  ranged from 0.09 to 0.14 in Bălțeanu et al. (2019) compared to our range of values being 0.16 to 0.24. A lower SNP-based observed heterozygosity value of 0.26 with a corresponding  $F_{ROH}$  estimate of 0.29 was previously estimated for SM by Zorc et al. (2022). They found SM to be second (after Turopolje:  $F_{ROH} = 0.51$ ), to have low level of genetic diversity amongst other local Balkan pig breeds. Contrary to previous finding (Bălțeanu et al., 2019) of RM exhibiting the strongest and most recent inbreeding, BM in our study had the highest level of autozygosity regardless of the age of inbreeding (Figure 6). Nevertheless, consanguinity increased in all three breeds in more recent generations, probably, about three generations ago as demonstrated by predominance of ROH larger than 16 Mb (Meszaros, 2015). More generally, Mangalitsa pigs are known to have undergone a serious demographic decline in the past (Egerszegi et al., 2003;

Posta et al., 2015; Bălțeanu et al., 2019). For the outgroup HUMA breed, observed heterozygosity estimated by Gorssen et al. (2021) was 0.26 with a corresponding high level of  $F_{ROH}$  estimate of 0.41. Since Gorssen et al. (2021) mentioned Yang et al. (2017) as source of the HUMA dataset used in their ROH analysis, we compared the estimate of  $F_{ROH}$  obtained for HUMA in the two studies. Average  $F_{ROH}$  for the same individuals was almost halved (0.22) in Yang et al. (2017). Variability in ROH statistics across studies are a direct consequence of differences in ROH calling criteria (Addo 2020; Meyermans et al., 2020) In this study, we used an R-script identical to Gorssen et al. (2021) for ROH calling, making it possible to compare our results to those available at the Open Science Framework (<https://doi.org/10.17605/OSF.IO/XJTKV>). For suitability of ROH island comparison, we further performed a separate analysis of our data using a previous genome assembly (*Sus scrofa* 10.2 genome build) as was used in the previous study. Autozygosity across breeds in our population of Mangalitsa pigs ( $n = 77$ ) was by far lower (~0.21) than was found

for the Hungarian population ( $n = 20$ ). The HUMA genotypes were sampled between 2008 and 2010 while we sampled our animals about a decade later, between 2018 and 2020. It could be that breed conservation strategies are being effectively implemented to save these endangered animals in recent years. Moreover, Bálteanu et al. (2019) argued that ROH statistics recorded in Mangalitsa pigs are comparable to those measured in most cosmopolitan breeds, adding the difficulty in predicting significant statistical differences due to high dispersion of  $F_{ROH}$  data in Mangalitsa pigs. Nevertheless, comparing  $F_{ROH}$  value of 0.14 (Bálteanu et al., 2019) to that of 0.29 in SB, Zorc et al. (2022) emphasized recent increase in inbreeding due to severe reductions in census numbers.

The inverse relationship between inbreeding coefficients and heterozygosity was generally studied across breeds in the current study. Our finding of a strongly negative correlation ( $-0.88$ ) between the two variables is not limited to this study. Slate et al. (2004) investigated the use of multi-locus heterozygosity as a robust surrogate for inbreeding coefficients for subsequent investigation of inbreeding depression, which is especially important for captive populations. Based on a limited number of microsatellite markers, not only Slate et al. (2004) but also other studies, e.g., Hedrick et al. (2001), Overall et al. (2005), Jensen et al. (2007) and Alho et al. (2009) generally reported negative weak correlation coefficient values between heterozygosity and pedigree-based inbreeding coefficients, although the estimate for one of the populations (Scandinavian wolves) reached  $-0.72$ . The authors identified the effect of marker density and pedigree errors among other factors that impact the estimate of correlation coefficients between the two variables. Higher estimate of correlation in our study could be legitimate owing to high marker density and a better precision of genomic inbreeding estimates. Our findings reemphasize the utility of heterozygosity as proxy for inbreeding in the study of inbreeding depression in light of the availability of current genetic markers and where phenotypes are available.

## Genes under selection

The detection of ROH island in all our breeds allows for the investigation of selection signatures, which are a consequence of selection, recombination and genetic drift (Ceballos et al., 2018; Gorssen et al., 2020). Surprisingly, identified ROH islands were not shared across the three breeds in our study, rather, SM and the outgroup HUMA shared the same ROH island on SSC11 (80–86 Mb). By contrast, in an investigation of breed substructures in Pietrain pigs, Gorssen et al. (2020) found several ROH islands common to different populations of the same pig type, and even reported a large ROH island on SSC8 (34–126 Mb), which appears to be fixed. For our findings, it is compelling that BM, RM, and SM have been managed as separate breeds whose ROH patterns have been predominantly shaped

through within-breed selection. Assuming that the HUMA genotypes are from Swallow-bellied Mangalitsa as suggested by our principal component analysis, the shared ROH island region of this breed with SM indicates a breed-specific signature of selection present in both the Hungarian and German populations. The shared ROH island between SM and HUMA was mapped to SSC11 (73–78 Mb) in our original analysis, where an updated reference assembly (*Sus scrofa* genome version 11.1) was used. One out of nine previously detected ROH islands in SM was mapped to SSC11, however, the exact position (33,516,188–36,493,548) differs from our finding (Zorc et al., 2022). The authors identified PCDH20 which is associated with brain activity and tameness in this region. They also found many other genes in regions some of which overlapped with those in Black Slavonian, Banija spotted and Moravka pig breeds. To investigate candidate genes in our study, we limited our scope to 100 kb distance flanking the significant SNPs defining ROH islands (Table 3). The proximal, genes ARHGEF7 (SSC11: 77,427,368–77,530,995) and EFNB2 (SSC11: 74,186,64–4,233,354) appear to be important in the SM and HUMA. ARHGEF7 is associated with the storage of materials, including nutrients, pigments and waste products (GO:0000322), and might influence the coat color pattern in SM. Among others, EFNB2 plays a role in the negative regulation of neuron projection development (GO:0010977), a positive regulation of neuron death (GO:1901216) and the movement of cells in response to specific chemical signals (GO:0050920). The gene has been suggested as a candidate for hearing and visual impairment, and pigmentary anomalies in human (Lévy et al., 2018). Distinguishing between Landrace and Yorkshire pigs *via* fixation index analysis, Wang et al. (2021) found EFNB2 to have been under intense selection pressure.

The second ROH island in SM (SSC17: 16–26), which was neither shared by HUMA nor BM and RM revealed a high level of selection for the kinesin family member 16B (KIF16B) gene. Interestingly, the same gene also popped up while comparing SM to both BM and RM (Table 1). Located on SSC17 (24,871,985–25,188,204), the 31,6219 BP length KIF16B gene is predominantly involved in microtubule-based movement (GO: 0007018). By this, it plays an important role in regulating early development and organogenesis such as in embryos, kidneys and in stem cells (Hirokawa and Tanaka 2015). KIF16B gene was found to significantly influence wool length and greasy yield in fine wool sheep (Wang et al., 2014; Zhao et al., 2021). In humans it has been associated with synaptic signaling, which confers intellectual abilities (Loo et al., 2012). Therefore, SM may differ from the other breeds in terms of cognitive ability. To our knowledge, this has not yet been investigated in our breeds. Nevertheless, several studies have shown general differences in cognitive abilities in pigs (Gieling et al., 2011). The ROH region in SM also harbors other important genes—BANF2 and BTBD3 are associated with male (Omolaoye et al., 2022) and female (Kim et al., 2012) fertility in pigs; ESF1 is associated with

meat quality in pigs (Puig-Oliveras et al., 2014); and MACROD2 is reported to be associated with disease resistance in cattle (González-Ruiz et al., 2019).

Autozygosity in RM is especially related to genes associated with meat quality including ABCA12 (Wang Xiao et al., 2019), VIL1 (Zhang et al., 2015; Verardo et al., 2017), PLSCR5 (Wang Zezhao et al., 2019) and USP37 (Verardo et al., 2017). Meat quality assessment in Mangalitsa pigs has been very general and shown that meat from these pigs have higher intramuscular fat content compared to a commercial pig breed (Stanišić et al., 2016). Nistor et al. (2012) found significant differences in the ratio of saturated and unsaturated fatty acids in RM and BM meat (35.88%:62.76% and 38.42%:59.94%, respectively). Although not properly documented, we know that among German Mangalitsa pig breeders, RM is mostly preferred owing to a perceived higher meat quality in the breed. Also, of importance in RM are genes previously reported to be associated with characteristics such as feed efficiency—CTSH (Russo et al., 2008) and VWC2L (Kejun et al., 2015); growth—RASGRF1 (Magee et al., 2010); and fertility—SPATA7 (Marques et al., 2018), TNP1 (Gòdia et al., 2020) and WWTR1 (Budna et al., 2017). In Egerszegi et al. (2003), RM had the highest body weight of 220 kg and 180 kg compared to lowest values of 165 kg and 170 kg in SM boars and sows, respectively.

Identified in the ROH island of BM, the PRLR gene is reported to play a role in prolactin signaling (GO:0038161), lactation (GO:0007595), mammary gland development (GO:0060644) and response to bacterium stimuli (GO:0009617). Studies on lactation in Mangalitsa pigs are lacking, however, there is one that only reported a mean lactation length of 52.57 days in BM (Petrović et al., 2013).

The occurrence of identical loci under divergent selection for the pairwise comparison of BMSM and RMSM signals similarity between BM and RM on one hand, and on the other hand, differences between each of the pair and SM. Inferring from the associated genes, characteristics such as behavior influenced by SPAG4 (Bélteky et al., 2016; Bélteky et al., 2017); body size by STIM2 (Osei-Amponsah et al., 2017); carcass and fat deposition by HNF4A (Fan et al., 2010; Alexandre et al., 2014; MoreiraMonterio et al., 2018; Criado-Mesas et al., 2020) and CEP250 (Damon et al., 2012); muscularity by CPNE1 (Dawei et al., 2010); and disease resistance by NFS1 (Dauben et al., 2021) are more likely to be similar between BM and RM than between each of these and SM. However, this may not be entirely true if several other differentially expressed gene control these traits. The RMBM comparison showed the lowest differentiation in this study, for which the SNP with the highest signal (M1GA0011680) was found proximal to an autoregulatory  $\alpha$ -adrenergic receptor 2C (ADRA2C) gene involved in the regulation of smooth muscle contraction among others (GO:0006940). Contrasting RM and BM, Bálteanu et al. (2021) previously found the region SSC16 (18–20 Mb) to have a potential effect on hair pigmentation. Their findings were, however, based on selection scans with HapFLK, BAYESACN

and GWAS, and furthermore, the analysis of gene content which revealed the solute carrier family 45-member 2 (SLC45A2) locus as a candidate gene.

Since their development, Mangalitsa pigs have evolved differentially as evidenced by both intra- and inter-population statistics in the current and previous studies. Adjudged by  $F_{st}$  estimates, the degree of differentiation, which peaked at about 9.5% (BMSM) in the current study is low compared to those in other breeds such as the Turopolje, being 21% on average (Druml et al., 2012). Besides, Mangalitsa pigs are collectively described as endangered and require conservation interventions. The maintenance of small populations under restricted geneflow between animals of different coat colors raises concerns about long-term implications for the conservation of genetic diversity. In our original data, 17 animals were said to be crossbreds from the three main breeds, meaning that some German breeders already practice crossbreeding amongst different Mangalitsa pig breeds. The lack of pedigree records makes it difficult to confirm this type of cross, and to use such animals in our analyses. The crossing of different Mangalitsa breeds by German breeders is not yet documented, however, possible reason for crossbreeding in these pigs can be the favorable effect of breed complementarity and the avoidance of inbreeding.

## Conclusion

In this study, we provided an insight into the differences in the three Mangalitsa pig breeds using a medium density SNP information. The Blond and Red Mangalitsa breeds are more similar to each other than each of them to the Swallow-bellied Mangalitsa. This finding sharply contrasts previous reports of the lowest differentiation between Blond and Swallow-bellied Mangalitsa pigs. Genetic diversity was highest in Red Mangalitsa; however, inbreeding was considerably high in all the breeds. Highly homozygous genomic regions were not shared across breeds and this, to a large extent, emphasizes the restriction of geneflow among the breeds. We found several breed specific signatures of selection including those that may underline growth and meat quality traits in Red Mangalitsa or that suggest intellectual ability in Swallow-bellied Mangalitsa. By providing a list of candidate genes for all genome-wide significant variants, we propose further investigations that would link these genes to actual phenotypes where available. Furthermore, we would add Manhattan plots of our ROH incidence to the repository of ROH islands for comparisons with future studies.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and



accession number(s) can be found in the article/[Supplementary Material](#).

## Ethics statement

Ethical review and approval was not required for the animal study because Ethical review and approval were waived for this study, since no experimental handling of animals was involved. The pigs were kept in line with national law and guidelines. Written informed consent was obtained from the owners for the participation of their animals in this study.

## Author contributions

SA and LJ conceived the study. SA wrote the codes. SA and LJ performed the data analysis. SA drafted the manuscript and both finalized the writing.

## Acknowledgments

We are sincerely grateful to the German Mangalitsa Breeders Association for their valuable support.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Addo, S. (2020). Assessment of genetic diversity in local breeds. Dissertation thesis. Kassel, Germany: University of Kassel. <https://kobra.uni-kassel.de/handle/123456789/12613>
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19 (9), 1655–1664. doi:10.1101/gr.094052.109
- Alexandre, P. A., Gomes, R. C., Santana, M. H. A., Silva, S. L., Leme, P. R., Mudadu, M. A., et al. (2014). Bovine NR1I3 gene polymorphisms and its association with feed efficiency traits in Nellore cattle. *Meta Gene* 2014, 206–217. doi:10.1016/j.mgene.2014.01.003
- Alho, J. S., Lillandt, Bo-G., Jaari, S., and Merilä, J. (2009). Multilocus heterozygosity and inbreeding in the Siberian jay. *Conserv. Genet.* 10 (3), 605–609. doi:10.1007/s10592-008-9588-z
- Bálteanu, V. A., Cardoso, T. F., Amills, M., Egerszegi, I., Anton, I., Beja-Pereira, A., et al. (2019). The footprint of recent and strong demographic decline in the genomes of Mangalitsa pigs. *Animal* 13 (11), 2440–2446. doi:10.1017/s1751731119000582
- Bálteanu, V. A., Cardoso, T. F., Amills, M., Luigi-Sierra, M. G., Egerszegi, I., Anton, I., et al. (2021). Red and blond Mangalitsa pigs display a signature of divergent directional selection in the SLC45A2 gene. *Anim. Genet.* 52 (1), 66–77. doi:10.1111/age.13031

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.909986/full#supplementary-material>

### SUPPLEMENTARY FIGURE 1

Cross-validation for varying number of clusters ranging from  $k = 2$  to  $k = 20$  in Mangalitsa pigs.

### SUPPLEMENTARY FIGURE 2

ROH incidence plots for each chromosome in Blond Mangalitsa pigs.

### SUPPLEMENTARY FIGURE 3

ROH incidence plots for each chromosome in Red Mangalitsa pigs.

### SUPPLEMENTARY FIGURE 4

ROH incidence plots for each chromosome in Swallow-bellied Mangalitsa pigs.

### SUPPLEMENTARY TABLE 1

SNPs in ROH island in Blond (BM), Red (RM) and Swallow-bellied (SM) Mangalitsa pigs. SNP positions were mapped to the *Sus scrofa* genome version 10.2.

### SUPPLEMENTARY DATA 1

Plink formatted files consisting of Blond (BM), Red (RM) and Swallow-bellied Mangalitsa (SM) genotypes.

Bélteky, J., Agnvall, B., and Jensen, P. (2017). Gene expression of behaviorally relevant genes in the cerebral hemisphere changes after selection for tameness in Red Junglefowl. *PLOS ONE* 12 (5), e0177004. doi:10.1371/journal.pone.0177004

Bélteky, J., Agnvall, B., Johnsson, M., Wright, D., and Jensen, P. (2016). Domestication and tameness: Brain gene expression in red junglefowl selected for less fear of humans suggests effects on reproduction and immunology. *R. Soc. Open Sci.* 3 (8), 160033. doi:10.1098/rsos.160033

Budna, J., Bryja, A., Celichowski, P., Kahan, R., Kranc, W., Ciesiółka, S., et al. (2017). Genes of cellular components of morphogenesis in porcine oocytes before and after IVM. *Reprod. Camb. Engl.* 154 (4), 535–545. doi:10.1530/REP-17-0367

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. *Nat. Rev. Genet.* 19 (4), 220–234. doi:10.1038/nrg.2017.109

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4.

Criado-Mesas, L., Ballester, M., Crespo-Piazuelo, D., Castelló, A., Fernández, A. I., and Folch, J. M. (2020). Identification of eQTLs associated with lipid metabolism in Longissimus dorsi muscle of pigs with different genetic backgrounds. *Sci. Rep.* 10 (1), 9845. doi:10.1038/s41598-020-67015-4

Damon, M., Wyszynska-Koko, J., Vincent, A., Héroult, F., and Lebret, B. (2012). Comparison of muscle transcriptome between pigs with divergent meat quality



- phenotypes identifies genes related to muscle metabolism and structure. *PLOS ONE* 7 (3), e33763. doi:10.1371/journal.pone.0033763
- Dauben, C. M., Pröll, C., Maren, J., Heuß, E. M., Appel, A. K., Henne, H., et al. (2021). Genome-wide associations for immune traits in two maternal pig lines. *BMC Genomics* 22 (1), 717. doi:10.1186/s12864-021-07997-1
- Druml, T., Salajpal, K., Dikic, M., Urosevic, M., Grilz-Seger, G., and Baumung, R. (2012). Genetic diversity, population structure and subdivision of local Balkan pig breeds in Austria, Croatia, Serbia and Bosnia-Herzegovina and its practical value in conservation programs. *Genet. Sel. Evol.* 44 (1), 5–9. doi:10.1186/1297-9686-44-5
- Durinck, S., Soellmen, P. T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* 4, 1184–1191. doi:10.1038/nprot.2009.97
- Egerszegi, I., Rátky, J., Solti, L., and Brüssow, K.-P. (2003). Mangalica – An indigenous swine breed from Hungary (review). *Arch. Anim. Breed.* 46 (3), 245–256. doi:10.5194/aab-46-245-2003
- Fan, B., Du, Z.-Q., and Rothschild, M. F. (2010). The Hepatocyte nuclear factor-1 alpha (HNF1A) gene is associated with fatness and loin muscle area in the pig. *Animal* 4 (10), 1619–1627. doi:10.1017/S175173111000087X
- Flegler, J. (1999). *Das Wollschwein. Gefährdete Nutztierasse des Jahres 1999*. Linz, Austria: Informationsbroschüre GEH.
- Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* 17 (1), 27–32. doi:10.1111/1755-0998.12509
- Frank, K., Molnár, J., Barta, E., and Marincs, F. (2017). The full mitochondrial genomes of Mangalica pig breeds and their possible origin. *Mitochondrial DNA. B Resour.* 2 (2), 730–734. doi:10.1080/23802359.2017.1390415
- García, D., Martínez, A., Dunner, S., Vega-Pla José, L., Fernández, C., et al. (2006). Estimation of the genetic admixture composition of Iberian dry-cured ham samples using DNA multilocus genotypes. *Meat Sci.* 72 (3), 560–566. doi:10.1016/j.meatsci.2005.09.005
- Gieling, E. T., Nordquist, R. E., and Josef, F. (2011). Assessing learning and memory in pigs. *Anim. Cogn.* 14 (2), 151–173. doi:10.1007/s10071-010-0364-3
- Gòdia, M., Reverter, A., González-Prendes, R., Ramayo-Caldas, Y., Castelló, A., Rodríguez-Gil, J.-E., et al. (2020). A systems biology framework integrating GWAS and RNA-seq to shed light on the molecular basis of sperm quality in swine. *Genet. Sel. Evol.* 52 (1), 72. doi:10.1186/s12711-020-00592-0
- González-Ruiz, S., Strillacci, M. G., Durán-Aguilar, M., Cantó-Alarcón Germinal, J., Herrera-Rodríguez Sara, E., Bagnato, A., et al. (2019). Genome-wide association study in Mexican holstein cattle reveals novel quantitative trait loci regions and confirms mapped loci for resistance to bovine tuberculosis animals: An open access journal from MDPI 9 (9). doi:10.3390/ani9090636
- Gorssen, W., Meyermans, R., Janssens, S., and Buys, N. (2021). A publicly available repository of ROH islands reveals signatures of selection in different livestock and pet species. *Genet. Sel. Evol.* 53 (1), 2. doi:10.1186/s12711-020-00599-7
- Gorssen, W., Meyermans, R., Buys, N., and Janssens, S. (2020). SNP genotypes reveal breed substructure, selection signatures and highly inbred regions in Piétrain pigs. *Anim. Genet.* 51 (1), 32–42. doi:10.1111/age.12888
- Hedrick, P., Fredrickson, R., and Ellegren, H. (2001). Evaluation of d<sub>2</sub>, a microsatellite measure of inbreeding and outbreeding, in wolves with a known pedigree. *Evolution* 55 (6), 1256–1260. doi:10.1111/j.0014-3820.2001.tb00646.x
- Herrero-Medrano, J. M., Megens, H.-J., Martien, A. G., Bosse, M., Pérez-Enciso, M., and Crooijmans, R. P. M. A. (2014). Whole-genome sequence analysis reveals differences in population management and selection of European low-input pig breeds. *BMC Genomics* 15 (1), 601–612. doi:10.1186/1471-2164-15-601
- Hirokawa, N., and Tanaka, Y. (2015). Kinesin superfamily proteins (KIFs): Various functions and their relevance for important phenomena in life and diseases. *Exp. Cell Res.* 334 (1), 16–25. doi:10.1016/j.yexcr.2015.02.016
- Jensen, H., Bremset, E. M., Ringsby, T. H., and Saether, B.-E. (2007). Multilocus heterozygosity and inbreeding depression in an insular house sparrow metapopulation. *Mol. Ecol.* 16 (19), 4066–4078. doi:10.1111/j.1365-294X.2007.03452.x
- Jukes, E. (2017). Mason's world encyclopedia of livestock breeds and breeding. *RR* 31 (5), 28–30. doi:10.1108/RR-02-2017-0039
- Kejun, W., Dewu, L., Jules, H.-S., Jie, C., Chengkun, L., Zhenfang, W., et al. (2015). Genome wide association analysis reveals new production trait genes in a male Duroc population. *PLOS ONE* 10 (9), e0139207. doi:10.1371/journal.pone.0139207
- Kharzinova, V. R., and Zinovieva, N. A. (2020). The pattern of genetic diversity of different breeds of pigs based on microsatellite analysis. *Vavilovskii Zhurnal Genet. Sel.* 24 (7), 747–754. doi:10.18699/VJ20.669
- Kim, M., Seo, H., Choi, Y., Shim, J., Kim, H., Lee, C.-K., et al. (2012). Microarray analysis of gene expression in the uterine endometrium during the implantation period in pigs. *Asian-Australas. J. Anim. Sci.* 25 (8), 1102–1116. doi:10.5713/ajas.2012.12076
- Lévy, J., Haye, D., Marziliano, N., Casu, G., Guimiot, F., Dupont, C., et al. (2018). EFN2 haploinsufficiency causes a syndromic neurodevelopmental disorder. *Clin. Genet.* 93 (6), 1141–1147. doi:10.1111/cge.13234
- Loo, S. K., Shtir, C., Doyle, A. E., Mick, E., McGough, J. J., McCracken, J., et al. (2012). Genome-wide association study of intelligence: Additive effects of novel brain expressed genes. *J. Am. Acad. Child. Adolesc. Psychiatry* 51 (4), 432–440. e2. doi:10.1016/j.jaac.2012.01.006
- Magee, D. A., Sikora, K. M., Berkowicz, E. W., Berry, D. P., Howard, D. J., Mullen, M. P., et al. (2010). DNA sequence polymorphisms in a panel of eight candidate bovine imprinted genes and their association with performance traits in Irish Holstein-Friesian cattle. *BMC Genet.* 11, 93. doi:10.1186/1471-2156-11-93
- Manunza, A., Amills, M., Noce, A., Cabrera, B., Zidi, A., Eghbalsaid, S., et al. (2016). Romanian wild boars and Mangalita pigs have a European ancestry and harbour genetic signatures compatible with past population bottlenecks. *Sci. Rep.* 6 (1), 29913–29919. doi:10.1038/srep29913
- Marincs, F., Molnár, J., Tóth, G., Stéger, V., and Barta, E. (2013). Introgression and isolation contributed to the development of Hungarian Mangalica pigs from a particular European ancient bloodline. *Genet. Sel. Evol.* 45 (1), 22–26. doi:10.1186/1297-9686-45-22
- Marques, D. B. D., Bastiaansen, J. W. M., Broekhuijsen, M. L. W. J., Lopes, M. S., Knol, E. F., Harlizius, B., et al. (2018). Weighted single-step GWAS and gene network analysis reveal new candidate genes for semen traits in pigs. *Genet. Sel. Evol.* 50 (1), 40. doi:10.1186/s12711-018-0412-z
- McQuillan, R., Leutenegger, A.-L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., et al. (2008). Runs of homozygosity in European populations. *Am. J. Hum. Genet.* 83 (3), 359–372. doi:10.1016/j.ajhg.2008.08.007
- Meyermans, R., Gorssen, W., Buys, N., and Janssens, S. (2020). How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC genomics* 21 (1), 94. doi:10.1186/s12864-020-6463-x
- Molnár, J., Nagy, T., Stéger, V., Tóth, G., Marincs, F., and Barta, E. (2014). Genome sequencing and analysis of Mangalica, a fatty local pig of Hungary. *BMC genomics* 15, 761. doi:10.1186/1471-2164-15-761
- Molnár, J., Tóth, G., Stéger, V., Zsolnai, A., Jánosi, A., Mohr, A., et al. (2013). Mitochondrial D-loop analysis reveals low diversity in Mangalica pigs and their relationship to historical specimens. *J. animal Breed. Genet. = Zeitschrift für Tierzucht und Zuchtungsbiologie* 130 (4), 312–320. doi:10.1111/j.1439-0388.2012.01014.x
- Moreira, G. C. M., Boschiero, C., Cesar, A. S. M., Reecy, J. M., Godoy, T. F., Trevisoli, P. A., et al. (2018). A genome-wide association study reveals novel genomic regions and positional candidate genes for fat deposition in broiler chickens. *BMC Genomics* 19 (1), 374. doi:10.1186/s12864-018-4779-6
- Nistor, E., Bampidis, V., Pentea, M., Prundeanu, H., and Ciolac, V. (2012). Nutritional quality of pork produced by Mangalita breed. *Bio Anim. Husb.* 2012(45), 386–389. doi:10.2298/BAH1002021P
- Omolaoye, T. S., Hachim, M. Y., Du, P., and Stefan, S. (2022). Using publicly available transcriptomic data to identify mechanistic and diagnostic biomarkers in azoospermia and overall male infertility. *Sci. Rep.* 12 (1), 2584. doi:10.1038/s41598-022-06476-1
- Osei-Amponsah, R., Skinner, B. M., Adjei, D. O., Bauer, J., Larson, G., Affara, N. A., et al. (2017). Origin and phylogenetic status of the local Ashanti Dwarf pig (ADP) of Ghana based on genetic analysis. *BMC Genomics* 18 (1), 193. doi:10.1186/s12864-017-3536-6
- Overall, A. D. J., Byrne, K. A., Pilkington, J. G., and Pemberton, J. M. (2005). Heterozygosity, inbreeding and neonatal traits in Soay sheep on St Kilda. *Mol. Ecol.* 14 (11), 3383–3393. doi:10.1111/j.1365-294X.2005.02682.x
- Petrović, M., Savić, R., Parunović, N., Radokjović, D., and Radović, C. (2013). “Reproductive traits of pigs of Mangalita breed,” in 8th International Symposium on the Mediterranean Pig, Slovenia, Ljubljana, October 10th–12th, 2013.
- Posta, J., Szabó, P., and Komlósi, I. (2015). Pedigree analysis of mangalica pig breeds. *Ann. Animal Sci.* 16 (3), 701–709. doi:10.1515/aoas-2015-0075
- Puig-Oliveras, A., Ramayo-Caldas, Y., Corominas, J., Estellé, J., Pérez-Montarelo, D., Hudson, N. J., et al. (2014). Differences in muscle transcriptome among pigs phenotypically extreme for fatty acid composition. *PLOS ONE* 9 (6), e99720. doi:10.1371/journal.pone.0099720
- Purfield, D. C., Berry, D. P., McParland, S., and Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC Genet.* 13, 70. doi:10.1186/1471-2156-13-70

- Purfield, D. C., McParland, S., Wall, E., and Berry, D. P. (2017). The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLoS ONE* 12 (5), e0176780. doi:10.1371/journal.pone.0176780
- R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Core Team.
- Russo, V., Fontanesi, L., Scotti, E., Beretti, F., Davoli, R., Nanni Costa, L., et al. (2008). Single nucleotide polymorphisms in several porcine cathepsin genes are associated with growth, carcass, and production traits in Italian Large White pigs. *J. Anim. Sci.* 86 (12), 3300–3314. doi:10.2527/jas.2008-0920
- Schachler, K., Distl, O., and Metzger, J. (2020). Tracing selection signatures in the pig genome gives evidence for selective pressures on a unique curly hair phenotype in Mangalitza. *Sci. Rep.* 10 (1), 22142. doi:10.1038/s41598-020-79037-z
- Slate, J., David, P., Dodds, K. G., Veenvliet, B. A., Glass, B. C., Broad, T. E., et al. (2004). Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: Theoretical expectations and empirical data. *Heredity* 93 (3), 255–265. doi:10.1038/sj.hdy.6800485
- Stanišić, N., Parunović, N., Stajić, S., Petrović, M., Radović, Č., Živković, D., et al. (2016). Differences in meat colour between free-range Swallow Belly Mangalitza and commercially reared Swedish Landrace pigs during 6 days of vacuum storage. *Arch. Anim. Breed.* 59 (1), 159–166. doi:10.5194/aab-59-159-2016
- Turner, S. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* 3 (25), 731. doi:10.21105/joss.00731
- Verardo, L. L., Sevón-Aimonen, M.-L., Serenius, T., Hietakangas, V., and Uimari, P. (2017). Whole-genome association analysis of pork meat pH revealed three significant regions and several potential genes in Finnish Yorkshire pigs. *BMC Genet.* 18 (1), 13. doi:10.1186/s12863-017-0482-x
- Wang, Kai, Wu, Pingxian, Chen, Dejuan, Zhou, Jie, Yang, Xidi, Jiang, Anan, et al. (2021). Detecting the selection signatures in Chinese Duroc, Landrace, Yorkshire, liangshan, and qingyu pigs. *Funct. Integr. Genomics* 21 (5–6), 655–664. doi:10.1007/s10142-021-00809-5
- Wang, Xiao, Cao, Chunwei, Li, Yongshun, Hai, Tang, Jia, Qitao, Zhang, Ying, et al. (2019a). A harlequin ichthyosis pig model with a novel ABCA12 mutation can be rescued by acitretin treatment. *J. Mol. Cell Biol.* 11 (12), 1029–1041. doi:10.1093/jmcb/mjz021
- Wang, Zezhao, Zhu, Bo, Niu, Hong, Zhang, Wengang, Xu, Ling, Xu, Lei, et al. (2019b). Genome wide association study identifies SNPs associated with fatty acid composition in Chinese Wagyu cattle. *J. Anim. Sci. Biotechnol.* 10, 27. doi:10.1186/s40104-019-0322-0
- Wang, Zhipeng, Zhang, Hui, Yang, Hua, Wang, Shouzhi, Rong, Enguang, Pei, Wenyu, et al. (2014). Genome-wide association study for wool production traits in a Chinese Merino sheep population. *PLoS one* 9 (9), e107101. doi:10.1371/journal.pone.0107101
- Weir, B. S., and Cockerham, C. Clark (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38 (6), 1358–1370. doi:10.1111/j.1558-5646.1984.tb05657.x
- Wilkinson, S., Lu, Zen H., Megens, Hendrik-Jan, Archibald, Alan L., Haley, Chris, Jackson, Ian J., et al. (2013). Signatures of diversifying selection in European pig breeds. *PLoS Genet.* 9 (4), e1003453. doi:10.1371/journal.pgen.1003453
- Yang, Bin, Cui, Leilei, Perez-Enciso, Miguel, Traspov, Aleksei, Crooijmans, Richard P. M. A., Zinovieva, Natalia, et al. (2017). Genome-wide SNP data unveils the globalization of domesticated pigs. *Genet. Sel. Evol.* 49 (1), 71–15. doi:10.1186/s12711-017-0345-y
- Zhang, Chunyan, Wang, Zhiquan, Bruce, Heather, Kemp, Robert Alan, Charagu, Patrick, Miar, Younes, et al. (2015). Genome-wide association studies (GWAS) identify a QTL close to PRKAG3 affecting meat pH and colour in crossbred commercial pigs. *BMC Genet.* 16, 33. doi:10.1186/s12863-015-0192-1
- Zhao, Haiyu, Hu, Ruixue, Li, Fadi, and Yue, Xiangpeng (2021). Two strongly linked blocks within the KIF16B gene significantly influence wool length and greasy yield in fine wool sheep (*Ovis aries*). *Electron. J. Biotechnol.* 53, 23–32. doi:10.1016/j.ejbt.2021.05.003
- Zorc, M., Škorput, D., Gvozdanić, K., Margeta, P., Karolyi, D., Luković, Z., et al. (2022). Genetic diversity and population structure of six autochthonous pig breeds from Croatia, Serbia, and Slovenia. *Genet. Sel. Evol.* 54 (1), 30. doi:10.1186/s12711-022-00718-6
- Zsolnai, A., Radnóczy, L., Fésüs, L., and Anton, I. (2006). Do mangalica pigs of different colours really belong to different breeds? *Arch. Anim. Breed.* 49 (5), 477–483. doi:10.5194/aab-49-477-2006
- Zsolnai, A., Tóth, G., Molnár, J., Stéger, V., Marincs, F., Jánosi, A., et al. (2013). Looking for breed differentiating SNP loci and for a SNP set for parentage testing in Mangalica. *Arch. Anim. Breed.* 56 (1), 200–207. doi:10.7482/0003-9438-56-019