

Differentiation of the domestic pig and wild boar using genotyping-by-sequencing

Anna KOSENIUK^{id}, Grzegorz SMOLUCHA^{id}, Artur GURGUL^{id}, Tomasz SZMATOŁA^{id}, Maria OCZKOWICZ^{id}
and Anna RADKO^{id}

Accepted January 19, 2023

Published online February 23, 2023

Issue online March 30, 2023

Original article

KOSENIAK A., SMOLUCHA G., GURGUL A., SZMATOŁA T., OCZKOWICZ M., RADKO A. 2023. Differentiation of the domestic pig and wild boar using genotyping-by-sequencing. *Folia Biologica (Kraków)* 71: 01-11.

Domestic pigs and wild boars have undergone frequent interspecies crossbreeding; therefore, the presence of hybrids makes it challenging to find genetic markers that distinguish both subspecies. The aim of this research is to identify the DNA regions that underwent strong selection during the domestication of the pig and to give an insight into the genetic diversity of the Polish wild boar and domestic pigs by implementing the genotyping-by-sequencing (GBS) technique. We studied two groups of animals: one consisted of domestic pigs (Landrace, Large White, Duroc, Puławska and Pietrain), while the second group included wild boars from Poland. The filtered single nucleotide polymorphisms (SNP) panel used in this study included 7,298 markers that were spread across 18 porcine autosomes and unmapped contigs. The maximum-likelihood phylogenetic trees and multidimensional scaling (MDS) clearly separated the populations of pigs from the wild boars. We also detected genome regions that demonstrated the most significant genetic differences between the domestic pigs and wild boars. These regions were distributed on eight different autosomes and overlapped with 48 different pig RefSeq genes. The KEGG pathway, Reactome and GO terms were further used to assign a functional significance to the identified genes that were associated with inter alia muscle development (MYOG, MEOX2), pre-weaning mortality stress (MYO7A) and sensory perception (TAS1R3).

Key words: GBS, population study, species identification, SNP.

Anna KOSENIUK[✉], Grzegorz SMOLUCHA, Tomasz SZMATOŁA, Maria OCZKOWICZ, Anna RADKO, National Research Institute of Animal Production, Department of Animal Molecular Biology.

E-mail: anna.koseniuk@iz.edu.pl

Artur GURGUL, Tomasz SZMATOŁA, University of Agriculture in Krakow, Center for Experimental and Innovative Medicine.

For many reasons, pigs are one of the most important livestock species for humans. First, pork is a major source of protein in global animal production. In addition, the pig is an excellent animal model to study the molecular components and processes of some human diseases including cancer (Groenen *et al.* 2012).

Food adulteration practices commonly include the substitution of wild boar meat products with widely accessible and cheaper, but poorer quality, pig meat (Koutsogiannouli *et al.* 2010). Therefore, the differentiation of both the pig (*Sus scrofa f. domestica*) and the wild boar (*Sus scrofa scrofa*) is of huge importance for the identification and

prevention of food fraud cases. A reliable method for species identification is needed by consumers, food and animal feed producers, as well as by the prosecuting authorities (Lorenzini *et al.* 2020). Molecular tests are commercially available that identify animal components in human food, as well as in animal feed. Based on PCR techniques (e.g. RFLP, Real-Time) it is possible to identify the DNA of cattle, horses and poultry in food and in feed for animals (Natonek-Wiśniewska *et al.* 2013; Safdar *et al.* 2014). Nonetheless, although a porcine component can be identified in a simple manner, the differentiation between a wild boar and a domestic pig in meat products remains challenging (Koutsogiannouli *et al.* 2010; Lorenzini *et al.* 2020; Kaltebrunner *et al.* 2019, 2020).

Fajardo *et al.* 2008 conducted studies for the first time that were aimed at the differentiation of wild boars and commercial pig breeds based on combining both nucleus (melanocortin 1 receptor, *MC1R*) and mitochondrial DNA (D-loop). The analysis of the *MC1R* gene proved to be more effective than the most common method used for species identification based on the polymorphism of mitochondrial DNA (Fajardo *et al.* 2008). The *NR6A1* gene has also been studied, which is a strong candidate gene for a vertebrae QTL. A proline to leucine substitution at *NR6A1* codon 192 (p.Pro192Leu) resulted in an increased number of vertebrae in commercial pig breeds (Mikawa *et al.* 2007). A method of differentiating pigs and wild boars was also tested using short tandem repeats (STRs), a marker type commonly used in parentage verification and population studies (Radko *et al.* 2021). The STRs alone (Rębala *et al.* 2016), combined with *MC1R* polymorphism (Nikolov *et al.* 2017) and *MC1R*, *NR6A1* or mitochondrial D-loop polymorphism, revealed the hybridisation that exists between pigs and wild boars (Lorenzini *et al.* 2020; Scandura *et al.* 2008). Szemethy *et al.* 2021 tested another approach based on identifying the insertions or deletions (InDels) specific to wild boars. In their case, the bioinformatics pipeline indicated three InDels to marker development during the performed routine genetic testing.

Introducing single nucleotide polymorphism (SNP) based molecular techniques to (sub-)species discrimination has many benefits. The most important of them is the significant number of markers (up to tens of thousands) that can be analysed for one individual. SNP-based molecular techniques have been implemented to seek the SNP markers which, under certain conditions (sufficient genotyping density and relevant reference populations), can

differentiate between a species and subspecies (Mary *et al.* 2022).

Rubin *et al.* (2012) used a whole-genome resequencing approach. They identified 33 selective sweeps localised on 18 *loci* covering the *NR6A1*, *PLAG1*, *LCORL*, and *OSTN* genes. All these genes have a known effect on the phenotype and are potential marker candidates. Another technology, based on a Porcine 60K SNP chip, was used to study the diversity of European pig breeds and to differentiate between domestic pigs and wild boars (Wilkinson *et al.* 2013). The genomes mostly differed in the regions covering the genes for bone formation (on *SSC1*), growth (on *SSC7*) and fat deposition (on *SSC7* and *SSCX*). However, the ability to genetically distinguish between phylogenetically close taxa like pigs and wild boar with a proven interspecies gene flow needs further elucidation.

Although microarrays are the most commonly used technology in genomic studies for small projects, their use is financially unjustified. A microarray study also requires details on the SNP type and the surrounding DNA sequence. For studying a poorly annotated species, the discovery of genome-wide SNPs and population-wide genotyping, genotyping-by-sequencing (GBS) could be an alternative to microarrays (Gurgul *et al.* 2019).

Genotyping-by-sequencing (GBS), which was originally developed for use in plants (crops), is a relatively simple approach that is applied in population genomic studies. GBS is both versatile and can be used for any species. It is based on the high-throughput, next generation sequencing (NGS) of genomic *loci* that are targeted by restriction enzymes (Elshire *et al.* 2011). The technique involves a digestion of the genomic DNA with one of the available frequently-cutting endonucleases and, following the adapters ligation, the sequencing of the obtained short DNA fragments using NGS. This approach reduces the target genome regions for NGS, thus allowing for an analysis of multiple samples (multiplexing abilities) in a single sequencing run. The GBS technique can be used to identify many markers with a relatively low error rate, and the detected polymorphisms are distributed throughout the whole genome. Another advantage of the technique is that it allows for the simultaneous discovery and genotyping of SNPs within a large number of samples. These features facilitate different types of genome-wide analyses such as relatedness studies, genome-wide association studies and genomic selection (Kim *et al.* 2016), as well as population studies, including among other

things the identification of the selection signatures (Baazaoui *et al.* 2020; Wang *et al.* 2018). However, the analysis of GBS-derived SNP datasets is complicated by frequent cases of missing genotypes. This is especially relevant in highly heterozygous organisms, due to a dependence on the read depth to correctly identify both alleles (Dodds *et al.* 2015).

For the first time, we used a GBS approach to study the genetic differentiation of wild boars and domestic pigs in Poland. The aim of the research is to identify the DNA regions that underwent strong selection during the domestication of the pig and to give an insight into the genetic diversity of the Polish wild boar and domestic pigs. The study also focuses on the identification of admixed individuals in both subspecies.

Materials and Methods

Material

The material for the study was genomic DNA extracted with the use of the Sherlock AX Kit for DNA purification (A&A Biotechnology, Poland). The DNA was extracted from the ear tissues of 20 wild boars (WB) from the South region of Poland, as well as from the blood samples of unrelated and randomly selected pigs of the following breeds: Polish Landrace (PBZ, n=2), Large White Polish (WBP, n=4), Duroc (DUR, n=1), Puławska (PUL, n=3) and Pietrain (n=2). The wild boars were legally hunted in accordance with the national regulations during the 2009-2011 hunting seasons. The sample collection was performed by licensed hunters as a part of routine wildlife management measures. Blood samples from the pigs were taken by the representatives of the Polish Pig Breeders Association, "POLSUS", and were collected in accordance with the routine parentage controls conducted at the National Research Institute of Animal Production in Balice, Poland; therefore, no permission from the local ethics commission was needed.

Methods

Library preparation and sequencing

We prepared the GBS libraries according to the previously described protocol (Elshire *et al.* 2011) with minor modifications. In short, the PstI endonuclease was used for the DNA digestion and ligation of 48 indexed adapters designed with the

use of GBS Barcode Generator software. Before the library preparation process, the ligation-ready double-stranded adapters were prepared, i.e. the oligonucleotides (forward and reverse complements) were synthesised with unique barcodes (4-9 nt long) and annealed with a common adapter. The laboratory workflow covered the following steps: overnight incubation of the genomic DNA (200 ng in total) with a tenfold PstI-HF enzyme excess (100,000 U/ml, New England Biolabs, USA) at 37°C; digestion and ligation of the products at 22°C for 60 min with 4.8 ng of each adapter, using the T4 DNA ligase (400 U/μl; New England Biolabs, USA); clean-up step; amplification of the libraries by PCR in 18 cycles, using universal primers (12.5 pmol/μl each) and 2x Taq Master Mix (New England Biolabs, USA); purification of the amplified libraries; quality assessment using Agilent's TapeStation 2200 system (Agilent Technologies, USA) and quantification using a DNA binding dye (Qubit DS DNA assay; Thermo Fisher Scientific, USA). The oligonucleotide sequences used here were the same as the ones used in our previous study (Gurgul *et al.* 2019).

The library pools were ultimately sequenced in a single-end 50-bp run on the HiScanSQ system (Illumina, USA) using the TruSeq v3 HiSeq flowcell and v3 sequencing reagents (Illumina, USA). The raw reads were deposited in the NCBI Sequence Read Archive (SRA) database under the BioProject Number ID PRJNA658252.

SNPs calling and filtering

The quality of the generated sequencing reads was assessed with the use of FastQC software. The subsequent analysis of the reads was done using the TASSEL 5 GBS v2 Pipeline (Glaubitz *et al.* 2014) with the default parameters, except maximum kmer Length that was set to 40 and the minimum count of reads for a tag to be output that was set to 5. The pig Sscrofa11.1 genome assembly was used as a reference for the TASSEL pipeline including the read tags mapping with a Bowtie2 aligner (Langmead & Salzberg 2012). Polymorphisms resulting from the default TASSEL GBS v2 production pipeline were initially filtered as following: indels and multiallelic markers were removed; individual genotypes were removed if the obtained coverage was low (< 5 reads); and finally, SNPs located on the X and Y chromosomes were removed due to a lack of knowledge of the wild boars' sexes.

We also removed from the analysis markers with a minor allele frequency (MAF) lower than 0.01, missing genotypes in more than 40% of the

individuals and those deviating from the Hardy-Weinberg equilibrium with a Chi-Square test p -value < 0.000001 . Moreover, the individuals with a more than 40% missing genotypes (three) were also removed. All these filtering steps were done using PLINK (Purcell *et al.* 2007) or TASSEL (Bradbury *et al.* 2007) software.

Phylogenetic and admixture analysis

The analysis of the genetic distance between the individuals was done in a similar way as in our previous work (Gurgul *et al.* 2019). MEGAX software (Kumar *et al.* 2018) was used to construct the phylogenetic tree constructed using the maximum likelihood (ML) method, along with subsequent bootstrapping. The additional visualisation of the population differentiation was done by a multidimensional scaling (MDS) applied to the IBS (identity by state) distance matrix calculated using TASSEL software. To further analyse the population structure and the patterns of admixture among the pig and wild boar populations, a Structure software (Falush *et al.* 2003) was used and was run 10 times per each K, initially from K1 up to K6. An admixture model was used, where the parameters were as follows: 100,000 iterations and a 100,000 burn-in period. The best K for this analyse was inferred using Structure Harvester (Earl & Vonholdt 2012) software based on the Delta K parameter. Finally, CLUMPAK software was used to visualise the results (Kopelman *et al.* 2015).

Genetic differentiation between the populations

Apart from the global genetic differentiation, we also attempted to identify genome regions showing the largest genetic differences between the analysed domestic pigs and wild boars. To this end, based on the filtered SNPs, we calculated the SNP-by-SNP pairwise F_{ST} generic distances (Weir & Cockerham 1984). Next, a 5-SNP sliding window was applied to obtain the moving average for the pairwise F_{ST} ,

which allowed us to account for a random locus-by-locus variation in the alleles frequency. Subsequently, the F_{ST} -windows falling into the top 1% of the highest F_{ST} observations were analysed in detail for their gene content, in order to identify the genes that were strongly affected by the wild boar domestication and the artificial selection of pigs. The genes within regions with the highest F_{ST} were identified using the UCSC Genome Browser (Haeussler *et al.* 2019) and were analysed in terms of their functions using Kobas3.0 (Xie *et al.* 2011). The overrepresentation tests in the separate GO, KEGG and Reactome pathways categories were conducted with respect to all known pig genes using Fisher's exact test and the FDR (Benjamini & Hochberg 1995) correction for multiple testing. Some annotation analyses were completed using the PANTHER Classification System (Mi *et al.* 2013).

Results

Sequencing read statistics

The sequencing of the libraries with the HiScanSQ System generated a total of 84.3 million (M) raw reads for all the animals. Of the reads, 36,285,567 (1.3 M on an average, per sample) had a proper barcode and PstI cut site. These reads were collapsed into 463,756 different sequence tags matching the reference genome. Finally, an average of about 1.1 M reads were mapped to the identified sequence tags and were used for SNP calling in the separate samples (Table 1).

SNPs panel characteristics

The initial TASSEL GBS pipeline allowed for the identification of 29,647 variants common for the domestic pigs and wild boars. Filtering of the markers for the average read depth and variants character (keeping only the biallelic SNPs) resulted in 7,785 SNPs, which were additionally evaluated

Table 1
Sequencing read and tags statistics

	Number of raw reads	Good bar-coded reads (sum)	Good barcoded reads per sample (mean)	Number of tags	Good reads mapped to tags per sample (mean)
Wild boar	–	16,178,714	1,348,226	–	1,183,829
Pig	–	20,106,853	1,340,457	–	1,123,428
Total	84,307,306	36,285,567	1,343,910	463,756	1,150,273

for their quality and variation parameters (MAF, percentage of genotypes called and Hardy-Weinberg equilibrium (HWE)). The final SNP panel included 7,298 markers (ranging from 20 on SSC18 to 587 on SSC6) showing satisfying polymorphism parameters and distributed across the Sscrofa11.1 genome assembly with an average inter-marker distance of 327kb (ranging from 205 to 484 kb for separate chromosomes) (Table 2). The mean MAF was 0.266 for all the markers and for both populations. A total of 5243 SNPs were polymorphic in both populations, and 2055 SNPs were monomorphic in at least one of them. The average observed heterozygosity of the SNPs was high and reached 0.346, on average. The polymorphism parameters were similar in both the analysed populations (Table 2). The average FIS for all individuals was 0.037, with a slightly higher value observed in the wild boars (0.044) than in the pigs (0.026).

Genetic differentiation of the pig and the wild boar captured using GBS

The final filtered panel of SNPs was used to analyse and visualise the genetic differentiation

between the studied populations. The MDS allowed for the identification of clearly separated populations of domestic pigs and wild boars, with a visibly lower genetic variation present within the wild boar cluster. Among the domestic pigs, a visible sub-clustering was observed that largely corresponded to the included breeds (Figure 1). Similar observations were made when the maximum-likelihood method was applied to create a phylogenetic tree. This analysis clearly separated the two analysed populations (Figure 2). To further validate this result, the Structure software was used to fully reveal the pattern of individual admixtures. For both the groups studied, we verified our hypothesis that more than two parental populations existed (at least one for the domestic representatives and a putative two parental lines, i.e. Asian and European, for the wild boar group). Therefore, we tested different K values ranging from one to six. However, the most appropriate K=2 was established using the Delta K parameter (Suppl. Mat. 1 – Fig. 1). This analysis showed a clear subdivision of the analysed animals into two subpopulations corresponding to the domestic pig and the wild boar populations (Figure 3).

Table 2
SNP panel statistics

CHR	SNPs #	# of shared SNPs	Distance between SNPs (bp)	All				Pig				Wild boar				Pig/Wild boar difference*		
				MAF	Ho	He	Fixed SNPs	MAF	Ho	He	Fixed SNPs	MAF	Ho	He	Fixed SNPs	MAF	Ho	He
1	566	428	483,995	0.279	0.366	0.370	138	0.264	0.356	0.345	34	0.246	0.352	0.313	104	0.018	0.004	0.032
2	383	273	396,869	0.276	0.363	0.366	110	0.267	0.349	0.353	26	0.228	0.349	0.294	84	0.038	0.000	0.058
3	487	333	272,844	0.262	0.331	0.354	154	0.262	0.338	0.345	32	0.215	0.302	0.281	122	0.047	0.037	0.064
4	390	284	333,230	0.265	0.334	0.355	106	0.261	0.336	0.346	28	0.217	0.309	0.285	78	0.043	0.027	0.060
5	348	261	299,874	0.274	0.356	0.365	87	0.258	0.348	0.344	24	0.242	0.337	0.313	63	0.016	0.011	0.030
6	587	369	290,681	0.256	0.326	0.346	218	0.280	0.360	0.363	25	0.190	0.274	0.246	193	0.090	0.086	0.117
7	379	282	320,387	0.271	0.332	0.363	97	0.254	0.313	0.336	30	0.230	0.327	0.303	67	0.024	-0.014	0.034
8	377	266	362,862	0.278	0.353	0.370	111	0.277	0.372	0.363	15	0.211	0.315	0.276	96	0.066	0.058	0.086
9	442	319	315,460	0.265	0.339	0.358	123	0.269	0.340	0.349	29	0.214	0.312	0.283	94	0.055	0.028	0.066
10	334	259	204,321	0.273	0.362	0.362	75	0.272	0.376	0.356	15	0.227	0.334	0.294	60	0.045	0.042	0.061
11	258	205	305,994	0.282	0.360	0.367	53	0.266	0.332	0.346	17	0.242	0.356	0.315	36	0.024	-0.024	0.030
12	287	186	213,221	0.253	0.304	0.346	101	0.276	0.330	0.364	13	0.179	0.257	0.239	88	0.097	0.073	0.125
13	441	314	473,241	0.279	0.358	0.367	127	0.262	0.354	0.340	43	0.239	0.338	0.305	84	0.023	0.016	0.035
14	510	358	278,147	0.267	0.337	0.360	152	0.267	0.329	0.352	29	0.226	0.318	0.288	123	0.041	0.011	0.063
15	386	283	364,286	0.261	0.359	0.349	103	0.253	0.351	0.334	28	0.220	0.344	0.285	75	0.034	0.007	0.049
16	226	156	353,030	0.250	0.316	0.342	70	0.247	0.327	0.332	13	0.196	0.284	0.260	57	0.050	0.043	0.072
17	307	218	205,745	0.267	0.337	0.357	89	0.266	0.344	0.353	14	0.214	0.307	0.277	75	0.052	0.037	0.075
18	203	160	272,869	0.257	0.344	0.352	43	0.229	0.311	0.310	20	0.244	0.355	0.323	23	-0.016	-0.044	-0.013
Cotigs	387	289	NA	0.223	0.383	0.303	98	0.228	0.383	0.299	34	0.206	0.370	0.271	64	0.021	0.013	0.028
All	7298	5243	326,918	0.266	0.346	0.356	2055	0.262	0.346	0.345	469	0.220	0.322	0.286	1586	0.042	0.024	0.059

CHR – chromosome; MAF – minor allele frequency; Ho – observed heterozygosity; He – expected heterozygosity; * – negative values point to the parameters which are lower in pigs than in wild boar; SNPs – polymorphic in both pig and wild boar populations; Fixed SNPs – monomorphic in at least one population.

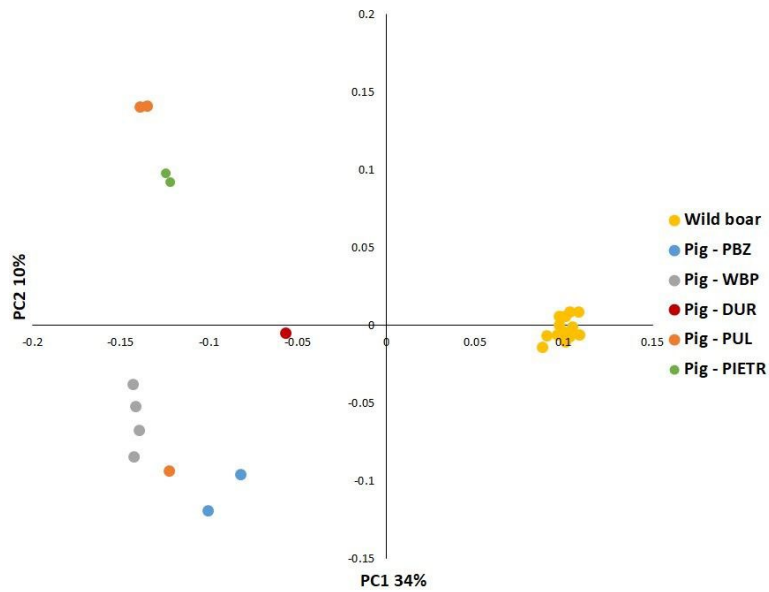


Fig. 1. Multidimensional scaling for the analysed individuals based on the IBS distance matrix. PBZ – Polish Landrace; WBP – Large White Polish; DUR – Duroc; PUL – Puławska; PIETR – Pietrain.

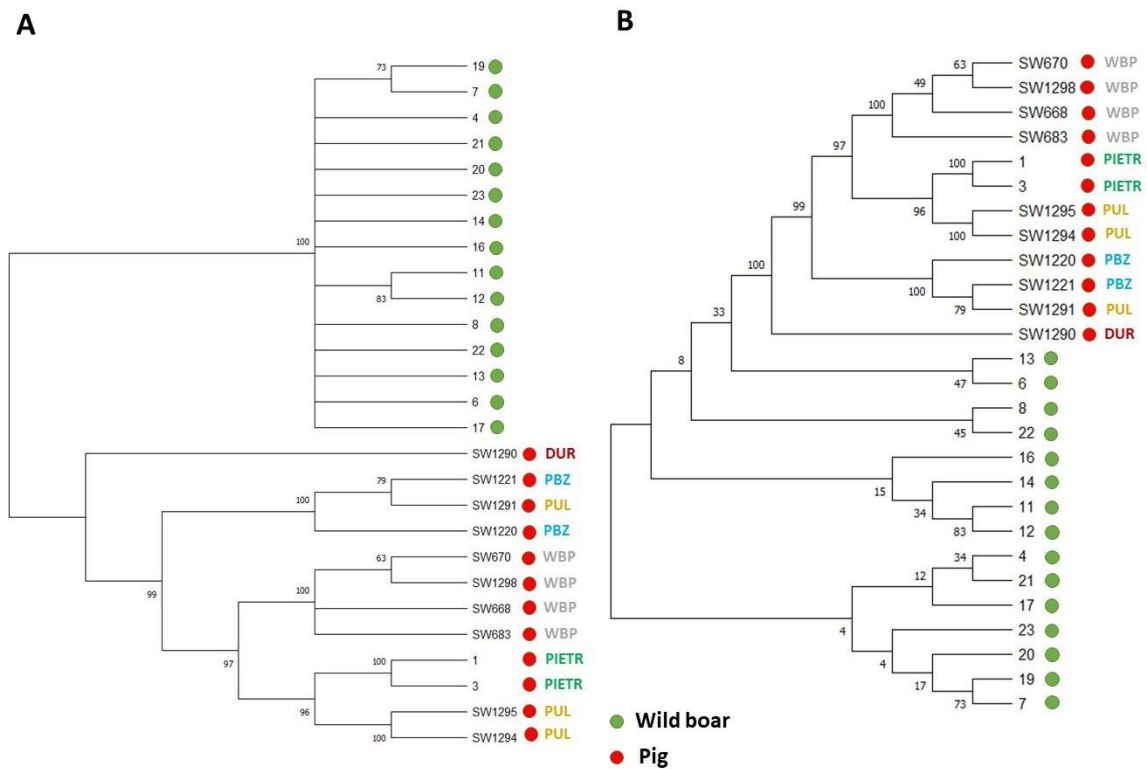


Fig. 2 Maximum-likelihood phylogenetic trees constructed with MEGA-X software. A – The evolutionary history was inferred by using the Maximum Likelihood method and the Jukes-Cantor model (Jukes & Cantor 1969). The tree with the highest log likelihood (-65198,14) is shown. The percentage of trees in which the associated taxa are clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with the superior log likelihood value. B – The bootstrap consensus tree inferred from 1000 replicates is shown to represent the evolutionary history of the taxa that were analysed (Felsenstein 1985). The branches corresponding to partitions reproduced in less than 50% of the bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa were clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). Initial tree(s) for the heuristic search were obtained automatically by applying the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with the superior log likelihood value. PBZ – Polish Landrace; WBP – Large White Polish; DUR – Duroc; PUL – Puławska; PIETR – Pietrain.

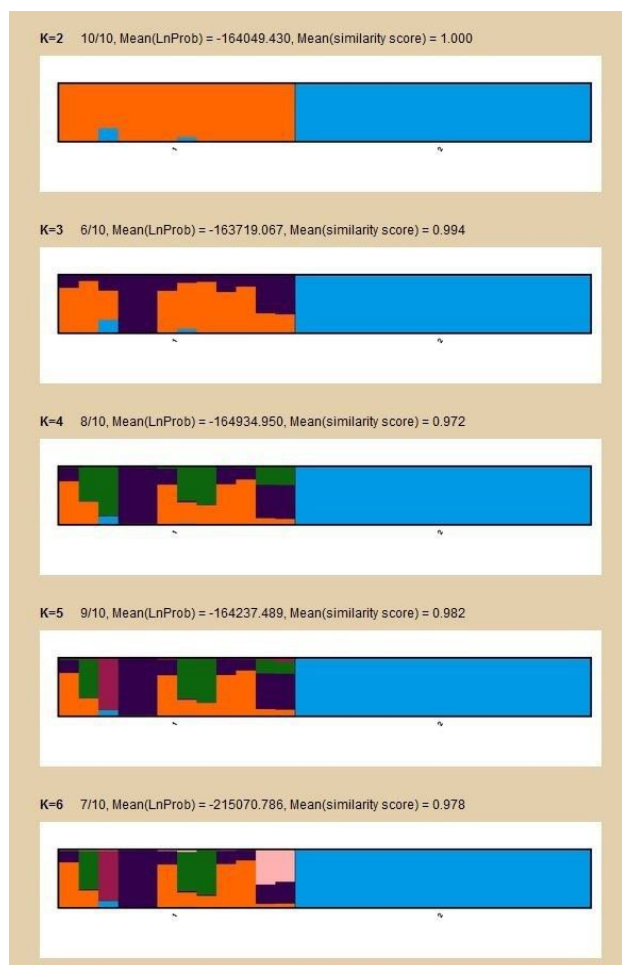


Fig. 3. Patterns of admixture among the studied populations revealed using Structure software. 1 – pig population; 2 – wild boar population; K – number of expected populations; K calculated from 2 to 6. The analysis was of locus-specific genetic differentiation.

Table 3

Top enriched GO teams for genes found in the genome regions the most divergent between wild boar and pig

GO term	GO ID	FDR	Genes
axonal transport of mitochondrion	GO:0019896	0.016194556	NEFL, UCHL1
UTP biosynthetic process	GO:0006228	0.016194556	NME1, NME2
GTP biosynthetic process	GO:0006183	0.016194556	NME1, NME2
CTP biosynthetic process	GO:0006241	0.016194556	NME1, NME2
endomembrane system	GO:0012505	0.016194556	RAB22A, PGRMC2, STX16
DNA-binding transcription activator activity, RNA polymerase II-specific	GO:0001228	0.016194556	WT1, PAX6, MYOG, NOBOX, MEOX2
nucleoside diphosphate phosphorylation	GO:0006165	0.016194556	NME1, NME2
positive regulation of transcription by RNA polymerase II	GO:0045944	0.01881921	WT1, NOBOX, NME2, AHR, PAX6, MAVS, MYOG
nucleoside diphosphate kinase activity	GO:0004550	0.01881921	NME1, NME2
positive regulation of protein kinase activity	GO:0045860	0.023408697	CDC25B, MMD
negative regulation of microtubule depolymerization	GO:0007026	0.023579161	SPEF1, APC
negative regulation of viral genome replication	GO:0045071	0.032865802	MAVS, ISG15
axon cytoplasm	GO:1904115	0.040083018	NEFL, UCHL1
cytoplasm	GO:0005737	0.042170286	SPEF1, CDH11, NME2, WT1, CDC25B, NPE-PL1, ISG15, APC, UCHL1, MEOX2, BZW2

The overall global genetic differentiation of the analysed population captured by the GBS SNPs was relatively low, with mean and weighted F_{ST} values of 0.099 and 0.118, respectively, meaning that only about 10% of the global genetic variation can be attributed to the differentiation between the populations.

The overlapping top 5-SNP windows were merged and the covered genome regions were analysed in detail. This approach allowed for the detection of the genome regions with fixed genetic differences between the pig and wild boar (Suppl. Mat. 1 – Fig. 2). The regions were distributed on 17 different autosomes, with the highest number located on SSC1, SSC6, SSC9 and SSC17 (3 regions each). The size of the regions ranged from 210 kb to 4.6 Mb, with a mean of 1.7 Mb. In the overlapped regions with 50 different RefSeq genes, 48 genes were annotated (Suppl. Mat. 2). Together, the genes enriched GO terms (FDR<0.05) associated inter alia with DNA-binding transcription activator activity, RNA polymerase II-specific and the regulation of oogenesis; KEGG ps was associated with the Wnt signalling pathway, arginine and proline metabolism and pyrimidine metabolism; and the Reactome pathways connected with, e.g., interconversion of nucleotide di- and triphosphates, metabolism of nucleotides and amino acids, VEGFR2 mediated cell proliferation and many other pathways (Table 3; Suppl. Mat. 2). In addition, the gene functions were supported by an extensive literature search.

Discussion

We used the GBS approach for the first time, with an aim to differentiate between closely related subspecies. GBS is a technique based on the random digestion of genomic DNA with one frequently-cutting restriction enzyme (REs) and ligating the cut site-specific adapters, including unique indexes following a common primer binding site. The selection of appropriate REs is a pivotal step in the GBS lab protocol, but it also generates the highest financial cost. The PstI enzyme was selected based on our previous experiences of farm animals as producing sufficient fragments for a genome-wide analysis (Gurgul *et al.* 2019). GBS takes advantage of a simplified hands-on lab protocol, which was achieved mostly by reducing the sequencing library preparation steps and simplifying the procedure for the generation of restriction fragments with barcoded adapters (Elshire *et al.* 2011). Finally, the time and cost were diminished by reducing the DNA purification steps and omitting DNA size selection (Gurgul *et al.* 2019).

While considering the GBS data analysis, the identification of the polymorphism and sample genotyping does not require a reference genome. This gives the GBS technique a broader purpose than more commonly used microarrays.

Genetic diversity and identification of hybrids and interspecies admixtures

The European and Asian wild boar populations diverged around 1 million years ago (mya), resulting in very different allele frequencies at millions of genomic *loci* and over a million *loci* representing alternative alleles (Groenen *et al.* 2012). While wild boars are the ancestors of modern domestic pigs, it is believed that the domestication process occurred separately for the species in Asia and Europe. Studies of the genetic diversity of wild boars and pigs can give insights into the phylogenetic background of the subspecies.

In our study, we found a low level of genetic diversity in the Polish wild boar population compared to the domestic pig population. Domestic pigs appear more divergent, as they originated from a highly diverse range of wild ancestors and were bred for several traits resulting in a large number of phenotypically diverse breeds (Yang *et al.* 2017). All the methods (MDS, Structure) we implemented to study the genetic structure of both subspecies proved that the analysed wild boars are not admixed with the gene pool of the domestic pigs. In the domestic pig group, two samples (one from the Duroc breed

and one Landrace) had a negligible proportion of the wild boar genome. This observation is also clearly demonstrated in the phylogenetic tree. Our studies did not coincide with that research on the European population that reported the frequent hybridisation of these two subspecies (Fontanesi *et al.* 2014; Iacolina *et al.* 2018; Scandura *et al.* 2008). There are farms (e.g. in Bulgaria and Sardinia) where pigs are kept in semi-wild conditions, and it is highly likely that in these cases the domestic pigs sporadically crossbred with wild boars (Scandura *et al.* 2008). The frequent hybridisation of pigs and wild boar was found in studies on the European population (Scandura *et al.* 2008; Fontanesi *et al.* 2014; Iacolina *et al.* 2018), while studies of the *MC1R* gene in the Polish population of wild boars also proved the existence of admixed genotypes (Dzialuk *et al.* 2018; Babicz *et al.* 2013). However, these results are probably due to the limited number of samples representing wild boars, which were restricted to one region of Poland. The breeding of Polish pigs also prevents crossing between domestic pigs and wild boars.

Genetic differentiation of the pig and the wild boar

In the present study, we identified genome regions that were differentially selected between the pig and the wild boar. Within these regions, 48 different genes were identified. Three databases – the KEGG pathway, Reactome and GO terms – were further used to assign a functional significance to the identified genes determined to have the greatest differences between the two groups. A detailed analysis of the enriched GO terms allowed for the identification of genes that are associated with the axonal transport of mitochondria (*NEFL*, *UCHL1*), DNA-binding transcription activator activity, RNA polymerase II-specific (*WT1*, *PAX6*, *MYOG*, *NOBOX*, *MEOX2*) and the positive regulation of transcription by RNA polymerase II (*WT1*, *NOBOX*, *NME2*, *AHR*, *PAX6*, *MAVS*, *MYOG*). The KEEG pathway and Reactome grouped the selected genes into biological processes: interconversion of nucleotide di- and triphosphates, myogenesis, RIG-I-like receptor signalling pathway, extra-nuclear oestrogen signalling and muscle contraction. These are processes that are presumably under selection during domestication and animal breeding.

Furthermore, we found several genes that were potentially under selection in early domestication, as well as in the later artificial selection of pigs. These are associated with inter alia muscle development (*MYOG*, *MEOX2*), pre-weaning mortality stress (*MYO7A*), immune response (*CSTF3*, *APC*) and sensory perception (*TASIR3*).

The breeding selection of pigs was aimed at increasing the meat production efficiency through both changes in the bone and muscle development processes and changes in the metabolism that affects the growth rate (Amaral *et al.* 2011). We found the *MYOG* and *MEOX2* genes in divergent gene regions, which influence muscle development (Otto *et al.* 2010; Stupka *et al.* 2012; Verner *et al.* 2007). *MYOG* belongs to the *MYOD* family of genes, which determines the lean meat content capacity in pork (Verner *et al.* 2007). The *MYOG* gene has been thoroughly studied in pigs, owing to its potential effect on different production and meat quality traits (Roman *et al.* 2012; Verner *et al.* 2007). The meat quality is positively correlated with the lean meat content, which relies on the muscle fibres. The key moment for developing muscle fibres is the period of embryogenesis (Verner *et al.* 2007). The *MEOX2* gene belongs to the mesoderm/mesenchyme homeobox family of genes with a pivotal role in the development and patterning of somites (Mankoo *et al.* 2003). In mice, it was found that a genetic deletion of *MEOX2* led to a serious reduction of the adult muscle size, as well as a modification of the myofibre composition, muscle architecture and physiological features (Otto *et al.* 2010).

Parameters related to reproduction affect the efficiency of the slaughter production. Derks *et al.* (2021) studied a stop-gain variant (g.11280403C>T) in the *MYO7A* gene. They proved a significant 23% decline in the pre-weaning survival while analysing 31 past carrier-by-carrier litters. In humans, a variant of the *MYO7A* gene is associated with Usher syndrome 1 (deaf-blindness) (Liu *et al.* 1997). *MYO7A* knocked-out mice suffered from various neurological diseases and reduced male fertility (Williams 2008).

To date, comparative studies the genomes of both pigs and wild boars have identified different genes related to the immune system, host defence and sensory perception (Amaral *et al.* 2011; Groenen *et al.* 2012). The genes associated with the immune system and olfactory receptors belong to the most rapidly evolving regions of the genome (Amaral *et al.* 2011). The immune response mechanism is always faced with strong challenges caused by changing environments and pathogens (Iacolina *et al.* 2016). Among the differentiating pigs and wild boar genes, we found several genes related to the immune response. These genes code the receptors for interleukins (e.g. *CSTF3*) or play a significant role in the immune response to viruses (e.g. *APC*). (Groenen *et al.* 2012) pointed out that among the porcine evolutionary breakpoint regions (EBPs) is a group related to the sensory perception of taste

in pigs (mostly related to salty, umami and sweet tastes). Among the genes coding taste receptors, we found the one related to the sweet taste (*TAS1R3*) in the divergent gene regions.

Despite a limited sample number, our results proved that the GBS as a whole-genome screening technique is a useful and valuable tool for population genetic studies and for the differentiation of two closely-related species and their possible hybrids. A limited number of SNPs (5K SNPs) was sufficient to grasp the genetic distance between pigs and wild boars. However, the number of SNPs might be improved in further experiments regarding this subject. We revealed novel genes with different patterns of variation in domestic pigs and wild boars. The genes differentiating both subspecies were related to both breeding selection (fertility and meat efficiency) and to biological processes associated with adaptation to the altered environment (host defence and sensory perception.). The polymorphism of these genes has the potential to differentiate wild boars from domestic pigs, but further studies with more samples representing both subspecies and their hybrids are needed.

Author Contributions

Research concept and design: A.K.; Collection and/or assembly of data: G.S., A.R.; Data analysis and interpretation: A.K., G.S., A.G., T.S., M.O.; Writing the article: A.K., M.O.; Critical revision of the article: G.S., A.G., T.S., M.O., A.R.; Final approval of article: A.K.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary Material to this article can be found online at:

<http://www.isez.pan.krakow.pl/en/fovia-biologica.html>

Supplementary files:

Suppl. Mat. 1 – Fig. 1. Delta K values for different numbers of populations assumed in the Structure analysis.

Suppl. Mat. 1 – Fig. 2. Sliding-window averaged FST values between pigs and wild boars.

Suppl. Mat. 2. Data for genetic differentiation analysis.

References

- Amaral A.J., Ferretti L., Megens H.J., Crooijmans R.P.M.A., Nie H., Ramos-Onsins S.E., Perez-Enciso M., Schook L.B., Groenen M.A.M. 2011. Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. *PLoS One* **6**. <https://doi.org/10.1371/journal.pone.0014782>
- Baazaoui I., McEwan J., Anderson R., Brauning R., McCulloch A., van Stijn T., Bedhiaf-Romdhani S. 2020. GBS data identify pigmentation-specific genes of potential role in skin-photosensitization in two tunisian sheep breeds. *Animals* **10**. <https://doi.org/10.3390/ani10010005>
- Babicz M., Pastwa M., Skrzypczak E., Buczyński J. T. 2013. Variability in the melanocortin 1 receptor (MC1R) gene in wild boars and local pig breeds in Poland, *Anim. Genet.* **44**: 357-358. <https://doi.org/10.1111/age.12023>
- Benjamini Y., Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**: 289-300.
- Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler E.S. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**: 2633-2635. <https://doi.org/10.1093/bioinformatics/btm308>
- Derks M.F.L., Megens H.J., Giacomini W.L., Groenen M.A.M., Lopes M.S. 2021. A natural knockout of the MYO7A gene leads to pre-weaning mortality in pigs. *Anim. Genet.* **52**: 514-517. <https://doi.org/10.1111/age.13068>
- Dodds K.G., Mcewan J.C., Brauning R., Anderson R.M., Stijn T.C., Kristjánsson T., Clarke S.M. 2015. Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genomics* **16**: 1-15. <https://doi.org/10.1186/s12864-015-2252-3>
- Działuk A., Zastempowska E., Skórzewski R., Twarużek M., Grajewski J. 2018. High domestic pig contribution to the local gene pool of free-living European wild boar: a case study in Poland. *Mamm Res.* **63**: 65-71. <https://doi.org/10.1007/s13364-017-0331-3>
- Earl D.A., Vonholdt B.M. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**: 359-361. <https://doi.org/10.1007/s12686-011-9548-7>
- Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., Mitchell S.E. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**: 1-10. <https://doi.org/10.1371/journal.pone.0019379>
- Fajardo V., González I., Martín I., Rojas M., Hernández P.E., García T., Martín R. 2008. Differentiation of European wild boar (*Sus scrofa scrofa*) and domestic swine (*Sus scrofa domestica*) meats by PCR analysis targeting the mitochondrial D-loop and the nuclear melanocortin receptor 1 (MC1R) genes. *Meat Sci.* **78**: 314-322. <https://doi.org/10.1016/j.meatsci.2007.06.018>
- Falush D., Stephens M., Pritchard J.K. 2003. Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics* **164**: 1567-1587.
- Felsenstein J. 1985. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution (N. Y.)* **39**: 783. <https://doi.org/10.2307/2408678>
- Fontanesi L., Ribani A., Scotti E., Utzeri V.J., Veličković N., Dal'olio S. 2014. Differentiation of meat from European wild boars and domestic pigs using polymorphisms in the MC1R and NR6A1 genes. *Meat Sci.* **98**: 781-784. <https://doi.org/10.1016/j.meatsci.2014.07.026>
- Glaubitz J.C., Casstevens T.M., Lu F., Harriman J., Elshire R.J., Sun Q., Buckler E.S. 2014. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS One* **9**. <https://doi.org/10.1371/journal.pone.0090346>
- Groenen M.A.M., Archibald A.L., Uenishi H., Tuggle C.K., Takeuchi Y., Rothschild M.F., Rogel-Gaillard C., Park C., Milan D., Megens H.J., Li S., Larkin D.M., Kim H., Frantz L.A.F., Caccamo M., Ahn H., Aken B.L., Anselmo A., Anthon C., Auvil L., Badaoui B., Beattie C.W., Bendixen C., Berman D., Blecha F., Blomberg J., Bolund L., Bosse M., Botti S., Bujie Z., Bystrom M., Capitanu B., Carvalho-Silva D., Chardon P., Chen C., Cheng R., Choi S.H., Chow W., Clark R.C., Clee C., Crooijmans R.P.M.A., Dawson H.D., Dehais P., De Sapio F., Dibbitts B., Drou N., Du Z.Q., Eversole K., Fadista J., Fairley S., Faraut T., Faulkner G.J., Fowler K.E., Fredholm M., Fritz E., Gilbert J.G.R., Giuffra E., Gorodkin J., Griffin D.K., Harrow J.L., Hayward A., Howe K., Hu Z.L., Humphray S.J., Hunt T., Hornshøj H., Jeon J.T., Jern P., Jones M., Jurka J., Kanamori H., Kapetanovic R., Kim J.H., Kim J.H., Kim K.W., Kim T.H., Larson G., Lee K.T.K., Lee K.T.K., Leggett R., Lewin H.A., Li Y., Liu W., Loveland J.E., Lu Y., Lunney J.K., Ma J., Madsen O., Mann K., Matthews L., McLaren S., Morozumi T., Murtaugh M.P., Narayan J., Nguyen D.T., Ni P., Oh S.J., Onteru S., Panitz F., Park E.W., Park H.S., Pascal G., Paudel Y., Perez-Enciso M., Ramirez-Gonzalez R., Reecy J.M., Rodriguez-Zas S., Rohrer G.A., Rund L., Sang Y., Schachtschneider K., Schraiber J.G., Schwartz J., Scobie L., Scott C., Searle S., Servin B., Southey B.R., Sperber G., Stadler P., Sweedler J.V., Tafer H., Thomsen B., Wali R., Wang J.J., Wang J.J., White S., Xu X., Yerle M., Zhang G., Zhang J.J., Zhang J.J., Zhao S., Rogers J., Churcher C., Schook L.B. 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **491**: 393-398. <https://doi.org/10.1038/nature11622>
- Gurgul A., Miksza-Cybulska A., Szmatola T., Jasielczuk I., Piestrzyńska-Kajtoch A., Fomal A., Semik-Gurgul E., Bugno-Poniewierska M. 2019. Genotyping-by-sequencing performance in selected livestock species. *Genomics* **111**: 186-195. <https://doi.org/10.1016/j.ygeno.2018.02.002>
- Haeussler M., Zweig A.S., Tyner C., Speir M.L., Rosenbloom K.R., Raney B.J., Lee C.M., Lee B.T., Hinrichs A.S., Gonzalez J.N., Gibson D., Diekhans M., Clawson H., Casper J., Barber G.P., Hausler D., Kuhn R.M., Kent W.J. 2019. The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res.* **47**: D853-D858. <https://doi.org/10.1093/nar/gky1095>
- Iacolina L., Pertoldi C., Amills M., Kusza S., Megens H.J., Bálteanu V.A., Bakan J., Cubric-Curic V., Oja R., Saarma U., Scandura M., Šprem N., Stronen A.V. 2018. Hotspots of recent hybridization between pigs and wild boars in Europe. *Sci. Rep.* **8**. <https://doi.org/10.1038/s41598-018-35865-8>
- Iacolina L., Scandura M., Goedbloed D.J., Alexandri P., Crooijmans R.P.M.A., Larson G., Archibald A., Apollonio M., Schook L.B., Groenen M.A.M., Megens H.J. 2016. Genomic diversity and differentiation of a managed island wild boar population. *Heredity (Edinb.)* **116**: 60-67. <https://doi.org/10.1038/hdy.2015.70>
- Jukes TH, Cantor CR. 1969. Evolution of protein molecules. (In: *Mammalian Protein Metabolism*. H.N. Munro ed. Academic Press, New York): 21-132. <https://doi.org/10.1016/B978-1-4832-3211-9.50009-7>
- Kaltenbrunner M., Mayer W., Kerkhoff K., Epp R., Rüggeberg H., Hochegger R., Cichna-Markl M. 2020. Applicability of a duplex and four singleplex real-time PCR assays for the qualitative and quantitative determination of wild boar and domestic pig meat in processed food products. *Sci Rep.* **10**: 17243. <https://doi.org/10.1038/s41598-020-72655-7>
- Kaltenbrunner M., Mayer W., Kerkhoff K., Epp R., Rüggeberg H., Hochegger R., Cichna-Markl M. 2019. Differentiation between wild boar and domestic pig in food by targeting two gene loci by real-time PCR. *Sci. Rep.* **9**: 9221. <https://doi.org/10.1038/s41598-019-45564-7>
- Kim C., Guo H., Kong W., Chandnani R., Shuang L.S., Paterson A.H. 2016. Application of genotyping by sequencing technology to a variety of crop breeding programs. *Plant Sci.* **242**: 14-22. <https://doi.org/10.1016/j.plantsci.2015.04.016>

- Kopelman N.M., Mayzel J., Jakobsson M., Rosenberg N.A., Mayrose I. 2015. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15**: 1179-1191. <https://doi.org/10.1111/1755-0998.12387>
- Koutsogiannouli E.A., Moutou K.A., Sarafidou T., Stamatis C., Mamuris Z. 2010. Detection of hybrids between wild boars (*Sus scrofa scrofa*) and domestic pigs (*Sus scrofa f. domestica*) in Greece, using the PCR-RFLP method on melanocortin-1 receptor (MC1R) mutations. *Mamm. Biol.* **75**: 69-73. <https://doi.org/10.1016/j.mambio.2008.08.001>
- Kumar S., Stecher G., Li M., Knyaz C., Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**: 1547-1549. <https://doi.org/10.1093/molbev/msy096>
- Langmead B., Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**: 357-359. <https://doi.org/10.1038/nmeth.1923>
- Lorenzini R., Fanelli R., Tancredi F., Siclari A., Garofalo L. 2020. Matching STR and SNP genotyping to discriminate between wild boar, domestic pigs and their recent hybrids for forensic purposes. *Sci. Rep.* **10**: 3188. <https://doi.org/10.1038/s41598-020-59644-6>
- Liu X., Newton V.E., Steel K.P., Brown S.D.M. 1997. Identification of a New Mutation of the Myosin VII Head Region in Usher Syndrome Type 1. *Human Mutation* **10**: 168-170.
- Mankoo B.S., Skuntz S., Harrigan I., Grigorieva, E., Candia A., Wright C.V.E., Arnheiter H., Pachnis V. 2003. The concerted action of Meox homeobox genes in required upstream of genetic pathways essential for the formation, patterning and differentiation of somites. *Development* **130**: 4655-4664. <https://doi.org/10.1242/dev.00687>
- Mary N., Iannuccelli N., Petit G., Bonnet N., Pinton A., Barasc H., Faure A., Calgaro A., Grosbois V., Servin B., Ducos A., Riquet J. 2022. Genome-wide analysis of hybridization in wild boar populations reveals adaptive introgression from domestic pig. *Evol. Appl.* **2022**, **15**: 1115-1128.
- Mi H., Muruganujan A., Thomas P.D. 2013. PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* **41**: 377-386. <https://doi.org/10.1093/nar/gks1118>
- Mikawa S., Morozumi T., Shimanuki S.I., Hayashi T., Uenishi H., Domukai M., Okumura N., Awata T. 2007. Fine mapping of a swine quantitative trait locus for number of vertebrae and analysis of an orphan nuclear receptor, germ cell nuclear factor (NR6A1). *Genome Res.* **17**: 586-593. <https://doi.org/10.1101/gr.6085507>
- Natonek-Wiśniewska M., Krzyściński P., Piestrzyńska-Kajtoch A. 2013. The species identification of bovine, porcine, ovine and chicken components in animal meals, feeds and their ingredients, based on COX I analysis and ribosomal DNA sequences. *Food Control* **34**: 69-78. <https://doi.org/10.1016/j.foodcont.2013.04.014>
- Nikolov I., Stoeckle B., Markov G., Kuehn R. 2017. Substantial hybridisation between wild boars (*Sus scrofa scrofa*) and East Balkan pigs (*Sus scrofa f. domestica*) in natural environment as a result of semi-wild rearing in Bulgaria. *Czech J. Anim. Sci.* **62**: 1-8.
- Otto A., Macharia R., Matsakas A., Valasek P., Mankoo B.S., Patel K. 2010. A hypoplastic model of skeletal muscle development displaying reduced foetal myoblast cell numbers, increased oxidative myofibres and improved specific tension capacity. *Dev. Biol.* **343**: 51-62. <https://doi.org/10.1016/j.ydbio.2010.04.014>
- Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., De Bakker P.I.W., Daly M.J., Sham P.C. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**: 559-575. <https://doi.org/10.1086/519795>
- Radko A., Smolucha G., Koseniuk A. 2021. Microsatellite DNA Analysis for Diversity Study, Individual Identification and Parentage Control in Pig Breeds in Poland. *Genes* **12**: 595.
- Rębała K., Rabtsava A.A., Kotova S.A., Kipen V.N., Zhurina N.V., Gandzha A.I., Iosif S. Tsybovsky. 2016. STR Profiling for Discrimination between Wild and Domestic Swine Specimens and between Main Breeds of Domestic Pigs Reared in Belarus. *PLoS ONE* **11**: e0166563. <https://doi.org/10.1371/journal.pone.0166563>
- Rubin C.-J., Megens H.-J., Barrio A.M., Maqbool K., Sayyab S., Schwochow D., Wang C., Carlborg Ö., Jern P., Jørgensen C.B. 2012. Strong signatures of selection in the domestic pig genome. *Proc. Natl. Acad. Sci. USA*, **109**: 19529-19536.
- Safdar M., Junejo Y., Arman K., Abasiyanik M.F. 2014. A highly sensitive and specific tetraplex PCR assay for soybean, poultry, horse and pork species identification in sausages: Development and validation. *Meat Sci.* **98**: 296-300. <https://doi.org/10.1016/j.meatsci.2014.06.006>
- Scandura M., Iacolina L., Crestanello B., Pecchioli E., Di Benedetto M.F., Russo V., Davoli R., Apollonio M., Bertorelle G., Bologna U. 2008. Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: Are the effects of the last glaciation still detectable? *Mol. Ecol.* **17**: 1745-1762. <https://doi.org/10.1111/j.1365-294X.2008.03703.x>
- Stupka R., Citek J., Sprysl M., Okrouhla M., Brzobohaty L. 2012. The impact of MYOG, MYF6 and MYOD1 genes on meat quality traits in crossbred pigs. *African J. Biotechnol.* **11**: 15405-15409. <https://doi.org/10.5897/ajb12.1820>
- Szemethy D., Mihalik B., Frank K., Nagy T., Újváry D., Kusza S., Szemethy L., Barta E., Stéger V. 2021. Development of Wild Boar Species-Specific DNA Markers for a Potential Quality Control and Traceability Method in Meat Products. *Analytical Methods* **14**: 18-27. <https://doi.org/10.1007/s12161-020-01840-1>
- Verner J., Humpolíček P., Knoll A. 2007. Impact of MYOD family genes on pork traits in Large White and Landrace pigs. *J. Anim. Breed. Genet.* **124**: 81-85. <https://doi.org/10.1111/j.1439-0388.2007.00639.x>
- Wang K., Wu P., Yang Q., Chen D., Zhou J., Jiang A., Ma J., Tang Q., Xiao W., Jiang Y., Zhu L., Li X., Tang G. 2018. Detection of selection signatures in Chinese Landrace and Yorkshire pigs based on genotyping-by-sequencing data. *Front. Genet.* **9**: 1-9. <https://doi.org/10.3389/fgene.2018.00119>
- Weir B.S., Cockerham C.C. 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution (N. Y.)* **38**: 1358-1370. <https://doi.org/10.2307/2408641>
- Wilkinson S., Lu Z.H., Megens H.-J., Archibald A.L., Haley C., Jackson I.J., Groenen M.A.M., Crooijmans R.P.M.A., Ogdén R., Wiener P. 2013. Signatures of Diversifying Selection in European Pig Breeds. *PLoS Genet.* **9**: e1003453.
- Williams D.S. 2008. Usher syndrome: Animal models, retinal function of Usher proteins, and prospects for gene therapy. *Vision Res.* **48**: 433-441. <https://doi.org/10.1016/j.visres.2007.08.015>
- Xie C., Mao X., Huang J., Ding Y., Wu J., Dong S., Kong L., Gao G., Li C.Y., Wei L. 2011. KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* **39**: 316-322. <https://doi.org/10.1093/nar/gkr483>
- Yang B., Cui L., Perez-Enciso M., Traspov A., Crooijmans R.P.M.A., Zinovieva N., Schook L.B., Archibald A., Gatphayak K., Knorr C., Triantafyllidis A., Alexandri P., Semiadi G., Hanotte O., Dias D., Dovč P., Uimari P., Iacolina L., Scandura M., Groenen M.A.M., Huang L., Megens H.J. 2017. Genome-wide SNP data unveils the globalization of domesticated pigs. *Genet. Sel. Evol.* **49**: 1-15. <https://doi.org/10.1186/s12711-017-0345-y>