Bolner, M., Schiavo, G., Bertolini, F., Bovo, S., Gallo, M., Fontanesi, L. (2026) Evaluating the impact of thirty years of directional genetic selection on shaping the population genomic structures of Italian pig breeds, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 23749: https://doi.org/10.31274/wcgalp.23749

Evaluating the impact of thirty years of directional genetic selection on shaping the population genomic structures of Italian pig breeds

M. Bolner[1], G. Schiavo[1], F. Bertolini[1*], S. Bovo[1], M. Gallo[2], and L. Fontanesi[1]

[1]*University of Bologna, Bologna, Italy;* [2]*ANAS, Roma, Italy;* [*]*Corresponding author: francesca.bertolini3@unibo.it*

Introduction

The breeding programs for the Italian heavy pig breeds (Italian Duroc, Italian Landrace and Italian Large White) were initiated in the early 1990s with the aim of meeting the specific requirements of the Italian dry-cured ham Protected Designation of Origin (PDO) value chain (Bosi and Russo, 2004). To this end, selection indices were developed to prioritize a combination of production, carcass, meat quality and reproductive traits, including growth, efficiency (average daily gain and feed conversion rate), carcass composition (backfat thickness and lean meat content), meat quality traits relevant for ham processing (curing loss at first salting and visible intermuscular fat), and reproductive performance (number of piglets born alive and teat number). The relative importance assigned to these traits differed among breeds and has been periodically updated over time, reflecting breed-specific selection objectives and evolving production needs within the PDO system (ANAS, 2022). While these breeding goals partially overlap with those of cosmopolitan pig populations, which are widely used in other intensive production systems, the Italian programs have been consistently oriented toward heavy pig production and ham quality rather than maximal lean growth efficiency.

Genomic differentiation among Italian heavy pig breeds has already been investigated, revealing clear breed-specific signatures of selection (Bertolini et al., 2024). Although such long-term directional selection is expected to leave detectable signatures on the genome, the extent to which the Italian heavy pig breeds have genetically differentiated from their cosmopolitan counterparts, and how this process has progressed over time, has not been systematically evaluated. Therefore, the aim of this study was to investigate the impact of approximately three decades of selection on the genomic structure of Italian Duroc, Italian Landrace and Italian Large White populations by tracking allele and haplotype frequency changes over time and comparing their genomic profiles with those of cosmopolitan Duroc, Landrace and Large White breeds.

Materials & Methods

All procedures involving animals complied with national and European regulations on animal welfare and were conducted in accordance with the ethical and legal requirements of Italy. A total of 12,157 Italian heavy pigs were sampled between 1994 and 2021. The animals belonged to the Italian Duroc (n = 1,802), Italian Landrace (n = 3,432) and Italian Large White (n = 6,923) breeds. All pigs were selected and raised within the breeding programs of the Italian pig breeders' association (Associazione Nazionale Allevatori Suini, ANAS) and were part of its routine performance recording system. Each pig was genotyped using either the PorcineSNP60 v2 BeadChip array (Illumina) or the GGP Porcine HD Array (GeneSeek), analyzing 64,232 and 51,544 single nucleotide polymorphisms (SNPs), respectively. Cosmopolitan pig breed data were obtained from two sources: (i) 60K SNP chip data derived from the materials made available in the study by Yang *et al.* (2017), which included 285 samples corresponding to 79

Duroc, 130 Landrace and 76 Large White pigs born between 1998 and 2010 (called GSE); and (ii) whole-genome sequencing (WGS) samples available from the European Nucleotide Archive (ENA), comprising a total of 145 samples corresponding to 43 Duroc, 32 Landrace and 70 Large White pigs, with an overall mean depth of coverage above 10×. These samples were aligned to the *Sus scrofa* reference genome (Sscrofa 11.1) and subsequent variant calling was performed following Bolner *et al*. (2023).

Genotype data preprocessing. Italian pig genotype data were pre-processed by first removing markers not shared between the two SNP chip platforms and subsequently excluding those with an overall genotyping rate below 90% using PLINK v1.9 (Purcell *et al.,* 2007). The GSE and ENA datasets were preprocessed by removing markers not shared with the Italian pig dataset, followed by the exclusion of markers with an overall genotyping rate below 90%.

Population structure analysis. Multi-dimensional scaling (MDS) was computed for each breed on the three merged SNP datasets using PLINK and used to evaluate the population structure of the analyzed samples. ADMIXTURE software v1.3 (Alexander et al., 2009) was applied to the combined dataset including all breeds and data sources to estimate individual ancestry proportions. The optimal number of ancestral populations (K) was determined by minimizing the cross-validation (CV) error.

Haplotype analysis. For the Italian pig dataset, haplotypes were built for each breed separately: SNeP v1.1 (Barbato *et al.,* 2015) was used to estimate the population size; then, genotypes were phased using shapeit v2.17 (Delaneau *et al.,* 2013), and haplotypes built with gHAP v3 (Utsunomiya *et al.,* 2016).

Detection of temporal genomic changes under selection. To evaluate genomic changes over time, Italian pigs were grouped by breed and year of birth. For each breed–year group, allele counts were extracted for each marker and used to compute allele frequencies; the same procedure was applied to the inferred haplotypes. Two complementary approaches were used to assess the significance of allele frequency changes across years: (i) the Mann–Kendall test to detect monotonic temporal trends, and (ii) logistic regression using a binomial generalized linear model (GLM), with yearly genotypes as input. For each breed, the top 1% of results based on p-values was selected independently for both tests and for both SNPs and haplotypes. The results were then intersected to retain only SNPs belonging to haplotypes ranked in the top percentile by both tests. To quantify the magnitude of allele frequency changes, median allele frequencies were calculated for the first and last five-year intervals. Moreover, pairwise F_{ST} values between the first and last five-year intervals were calculated within each breed using PLINK v1.9 (Purcell *et al.,* 2007). All retained haplotypes were annotated using Pig QTLdb information (Hu *et al.*, 2022).

Results & Discussion

The total number of markers retained after preprocessing was 41,128, 40,818 and 33,266 for the Italian, GSE and ENA pig datasets, respectively, including SNPs on the X chromosome and unplaced scaffolds; of these, between 25,091 and 27,696 were retained and used to construct haplotypes in the Italian pig dataset. The number of haplotypes identified in the Italian pig breeds was 34,958, 41,414 and 42,062 for Italian Duroc, Italian Landrace and Italian Large White, respectively. Figure 1 shows the MDS plot for all pigs included in the study. An overlap between cosmopolitan pig samples and older Italian pig breed samples is observed, consistent with a progressive genetic differentiation over time.

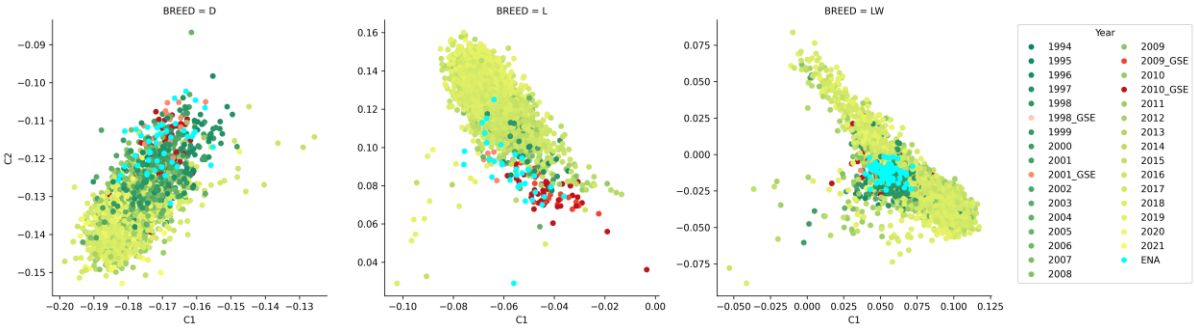

Figure 1. MDS plot for the Italian pig datasets including genomic data derived from the GSE and ENA datasets. Each figure represents one of the three groups of breeds (D, Duroc; L, Landrace; LW, Large White), and each point represents a pig sample. Samples are colored by year of birth.

ADMIXTURE analysis with K=9 showed a clear differentiation between the three breeds, with comparable ancestry distribution across different datasets. In the Italian pig dataset, a trend for ancestry change over years was visible in all breeds, indicating a gradual breed genetic divergence. This was supported by Pearson correlation computed between K and year values; for Italian Duroc, Pearson *r* was 0.76 with K5 and -0.75 with K2; for Italian Landrace, *r* was 0.37 with K4 and -0.45 with K9; for Italian Large White, *r* was 0.46 with K3 and -0.70 with K9.

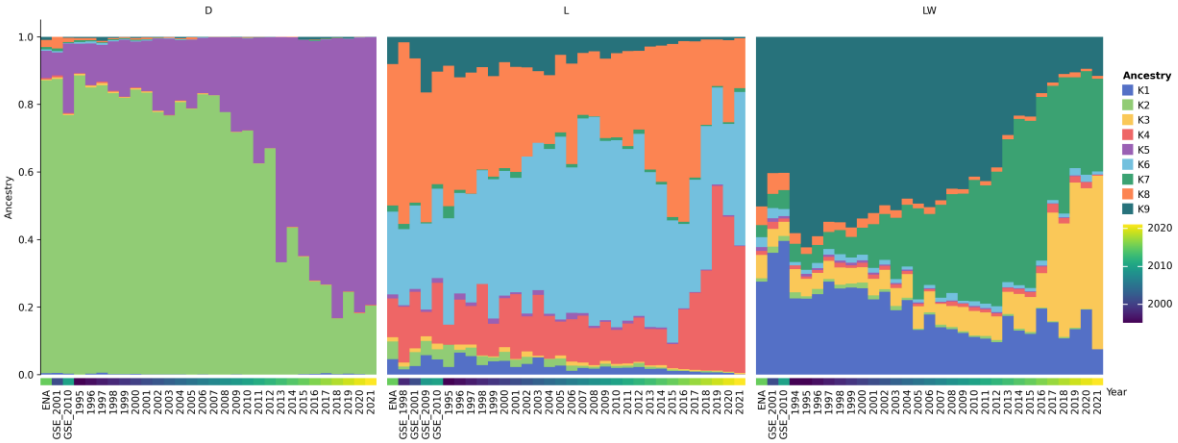

Figure 2. ADMIXTURE plots for the pig breed datasets. Each subplot represents the mean ancestry value for each different year for one of the three groups of breeds (D, Duroc; L, Landrace; LW, Large White), including ENA, GSE and Italian pigs. Different colors represent ancestry proportion. On the bottom, sampling year is shown.

The allele frequency and haplotype analyses in the three Italian breeds over time resulted in 250, 273 and 273 SNPs and 350, 414 and 420 haplotypes in the top 1% of results for Italian Duroc, Italian Landrace and Italian Large White, respectively. After intersecting the results of the two statistical approaches and retaining only SNPs located within significant haplotypes, a final high-quality subset of 32, 6 and 26 SNPs belonging to 40, 5 and 24 haplotypes was obtained for Italian Duroc, Italian Landrace and Italian Large White breeds, respectively. Across all SNPs, the absolute difference in median allele frequency between the first and last five-year intervals averaged 0.09 ± 0.07, 0.11 ± 0.08 and 0.10 ± 0.07 for Italian Duroc, Italian Landrace and Italian Large White, respectively, increasing to 0.32 ± 0.05, 0.38 ± 0.07 and 0.32 ± 0.05 when considering the top-ranked SNPs. To further quantify genomic differentiation

over time, pairwise F_{ST} values were calculated within each breed between the first and last five-year intervals. Genome-wide F_{ST} values were 0.043, 0.046 and 0.039 for Italian Large White, Italian Landrace and Italian Duroc, respectively, indicating a moderate level of temporal differentiation. When considering only the top-ranked SNPs identified in this study, F_{ST} values increased substantially to 0.30, 0.22 and 0.29 for Italian Large White, Italian Landrace and Italian Duroc, respectively, confirming the strong divergence at loci under selection.

All top-ranked SNPs were breed-specific, and the corresponding haplotypes overlapped genes associated with a range of production-related traits, as retrieved from the PigQTLDB, including body weight, androstenone levels, muscle development and teat number. Clear breed-specific patterns were observed. In Italian Duroc, several genes were associated with QTLs related to growth and body development (e.g., *FGF14*, *CARMIL1*, *SCGN*), as well as reproductive and meat quality traits, such as *FRAS1* (teat number), *ETV5* (stillbirth rate) and *TAC3* (androstenone levels). In Italian Landrace, fewer signals were detected, with genes such as *MED27* associated with QTLs for meat quality traits (e.g., cooking loss). In Italian Large White, genes involved in QTLs for growth and skeletal development (*PRIM2*, *KIF6*) and fat deposition (*MOCS1*) were identified, along with morphological traits such as *KCNIP4*.

Conclusions

Long-term directional selection in Italian heavy pig breeding programs has led to measurable temporal changes in allele and haplotype frequencies and population structure in Italian Duroc, Italian Landrace and Italian Large White breeds. The integration of allele frequency trends, haplotype analysis and F_{ST} estimates highlights a consistent pattern of genomic divergence over time, with stronger signals concentrated in loci under selection. Comparison with cosmopolitan populations suggests that these changes reflect selection pressures distinct from those of intensive lean-type production systems. Overall, the results indicate that PDO-oriented breeding strategies have contributed to the progressive genomic differentiation of Italian heavy pig breeds and provide new insights into the genomic consequences of long-term selection within specialized production systems.

Acknowledgements

Funded by the Programma di Sviluppo Rurale Nazionale (PSRN) SUIS-2 (co-funded by the European Agricultural Fund for Rural Development of the European Union and by the MASAF) and by the Horizon Europe Re-Livestock project, grant agreement no. 101059609.

References

ANAS (2022) Circolare 202200001: Mission and Genetic Progress in Italian Heavy Pig Breeds. Available at: https://www.anas.it/files/circolari/202200001.PDF, visited on: 27-01-2026.

Alexander D.H., Novembre J., and Lange K. (2009) Genome Res. 19:1655-1664. doi.org/10.1101/gr.094052.109

Babraham Bioinformatics—FastQC A Quality Control tool for High Throughput Sequence Data. Available at: www.bioinformatics.babraham.ac.uk/projects/fastqc/, visited on: 13-06-2025.

Barbato M., *et al.* (2015) Front. Genet. 6:109. doi.org/10.3389/fgene.2015.00109

Bertolini F, et al. (2024) Animal. doi.org/10.1016/j.animal.2024.101335

Bolner M., et al. (2023) Animals 14: 5. doi.org/10.3390/ani14010005

Bosi P., and Russo V. (2004) Ital. J. Anim. Sci. 3.4: 309-321. doi.org/10.4081/ijas.2004.309

Delaneau O., Zagury J.-F., and Marchini J. (2013) Nat. Methods 10:5-6. doi.org/10.1038/nmeth.2307

Hu Z.-L., Park C.A., and Reecy J.M. (2022) Nucleic Acids Res. 50:D956-D961. doi.org/10.1093/nar/gkab1116

Purcell S., Neale B., *et al.* (2007) Am. J. Hum. Genet. 81:559-575. doi.org/10.1086/519795

Utsunomiya Y.T., Milanesi M., Utsunomiya A.T.H., Ajmone-Marsan P., and Garcia J.F. (2016) Bioinformatics 32:2861-2862. doi.org/10.1093/bioinformatics/btw356

Yang B., Cui L., *et al.* (2017) Genet. Sel. Evol. 49:71. doi.org/10.1186/s12711-017-0345-y