Italian Journal of Animal Science

ISSN: 1828-051X (Online) Journal homepage: www.tandfonline.com/journals/tjas20

Genetic characterisation of Germanata Veneta and Mignon local duck breeds (*Anas platyrhynchos*) using whole-genome sequencing

Francesco Perini, Enrico Bortoletto, Filippo Cendron, Umberto Rosani, Emiliano Lasagna, Maristella Baruchello, Martino Cassandro & Mauro Penasa

To cite this article: Francesco Perini, Enrico Bortoletto, Filippo Cendron, Umberto Rosani, Emiliano Lasagna, Maristella Baruchello, Martino Cassandro & Mauro Penasa (2025) Genetic characterisation of Germanata Veneta and Mignon local duck breeds (*Anas platyrhynchos*) using whole-genome sequencing, Italian Journal of Animal Science, 24:1, 1621-1631, DOI: 10.1080/1828051X.2025.2541079

To link to this article: https://doi.org/10.1080/1828051X.2025.2541079

ITALIAN JOURNAL OF ANIMAL SCIENCE
2025, VOL. 24, NO. 1, 1621–1631
https://doi.org/10.1080/1828051X.2025.2541079

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Genetic characterisation of Germanata Veneta and Mignon local duck breeds (*Anas platyrhynchos*) using whole-genome sequencing

Francesco Perini[a], Enrico Bortoletto[b], Filippo Cendron[c], Umberto Rosani[b], Emiliano Lasagna[d], Maristella Baruchello[e], Martino Cassandro[c,f] and Mauro Penasa[c]

[a]Dipartimento di Chimica, Biologia e Biotecnologie, Università degli Studi di Perugia, Perugia, Italy; [b]Dipartimento di Biologia, Università degli Studi di Padova, Padova, Italy; [c]Dipartimento di Agronomia Animali Alimenti Risorse naturali e Ambiente, Università degli Studi di Padova, Legnaro, Italy; [d]Dipartimento di Scienze Agrarie, Alimentari e Ambientali, Università degli Studi di Perugia, Perugia, Italy; [e]Veneto Agricoltura, Legnaro, Italy; [f]Federazione delle Associazioni Nazionali di Razza e Specie, Roma, Italy

ABSTRACT

This study investigated the genetic diversity and conservation status of two endangered Italian duck breeds: Germanata Veneta (GER, $n = 50$) and Mignon (MIG, $n = 50$). The DNA was extracted from blood and analysed through whole-genome sequencing. The single nucleotide polymorphism (SNP) calling allowed the identification of 186,287 shared SNPs which were used to estimate expected and observed heterozygosity, and inbreeding coefficients based on homozygosity (F_{HOM}) and on runs of homozygosity. Both breeds showed low genetic variability and high inbreeding (GER: $F_{HOM} = 0.305 \pm 0.184$; MIG: $F_{HOM} = 0.369 \pm 0.163$). Multidimensional scaling revealed substructure within MIG, supported by phylogenetic and admixture analyses. Runs of homozygosity islands in GER were located on chromosomes 1, 2, 10, 11 and 21, with 524 SNPs enriched in immune response genes. In MIG, runs of homozygosity islands mapped to several chromosomes and included 614 SNPs related to muscle and cytoskeletal functions. Genes associated with small body size (e.g. *SHOX*, *IGF2R* and *PRKG2*) were found in MIG. A genome-wide association study comparing GER and MIG identified major genetic differences on chromosomes 2, 3, 5 and 13, including genes related to fatty acid metabolism (*HADHA* and *HADHB*) and dwarfism (*POC1A*). The results revealed low variability and high inbreeding, providing data to inform conservation programs aimed at mitigating the risk of genetic erosion.

HIGHLIGHTS

- Low genetic diversity and high inbreeding were observed in both Germanata Veneta (GER) and Mignon (MIG) duck breeds.
- ROH islands in GER were enriched in immune-related genes, while those in MIG were associated with muscle function and body size.
- GWAS identified genetic differences linked to fatty acid metabolism and dwarfism, particularly on chromosomes 3 and 13.

ARTICLE HISTORY
Received 16 May 2025
Revised 21 July 2025
Accepted 24 July 2025

KEYWORDS
Duck; local breed; WGS; genetic resource

Introduction

Ducks (*Anas platyrhynchos*) are the world's most widely distributed and agriculturally significant waterfowl species (Pingel 2011), serving as a crucial source of meat, eggs, and feathers. Local duck breeds often exhibit unique traits shaped by both natural selection and human-driven breeding practices, making them an essential resource for maintaining biodiversity and addressing future challenges in poultry production (Nielsen et al. 2023). Despite their agricultural and ecological importance, several local duck breeds remain uncharacterised at the genomic level. Understanding their genetic background is crucial for their effective management and conservation.

Genomic studies, with the identification of unique genetic variants and adaptive traits, can uncover the basis of economically important characteristics, such as growth rate (Wang et al. 2013), egg production (Perini et al. 2024), and resilience to environmental stressors (Hu et al. 2021), thereby supporting targeted breeding and conservation

CONTACT Filippo Cendron filippo.cendron@unipd.it

Supplemental data for this article can be accessed online at https://doi.org/10.1080/1828051X.2025.2541079.

strategies while maintaining genetic diversity (Besbes et al. 2007; Sponenberg et al. 2019). With the advent of high-throughput sequencing, it has become feasible to examine genetic variations at the molecular level. This advancement offers insights into the diversity and structure of populations and aids in the formulation of conservation strategies (Ghildiyal et al. 2023).

The present study exploits a comprehensive genomic approach to characterise the genetic architecture of two endangered Italian local duck breeds: Germanata Veneta (GER) and Mignon (MIG). Both breeds have been historically reared in the Veneto region, where they are valued for their meat production and, in the case of GER, for egg production, yielding approximately 120 eggs per year. By analysing genome-wide single nucleotide polymorphism (SNP) data, the present study aims to assess genetic diversity, population structure, and breed-specific variations. Furthermore, the research investigates possible introgression events and their impact on the genetic makeup of these breeds.

Materials and methods

Sample collection and sequencing

A total of 100 blood samples were collected from 50 randomly selected GER ducks (13 males and 37 females) and 50 randomly selected MIG ducks (9 males and 41 females) at the conservation centre 'Sasse Rami' of the Veneto Agricoltura Regional Agency (Ceregnano, Rovigo, Italy) and at the Institute of Agricultural Higher Education 'A. Della Lucia' (Feltre, Belluno, Italy) of Veneto region. Blood samples (about 2 mL) were collected from ulnar veins and stored at −20 °C in Vacutainers® tubes containing EDTA as anticoagulant. The main morphometric characteristics of the sampled adult animals, recorded according to FAO (2012) are summarised in Table 1.

DNA extraction, DNA quality check, library preparation and whole-genome sequencing (WGS) were performed by Neogen (Ayr, Scotland, UK) using standardised protocols and the InfiniSEEK™ platform, which applies low-coverage WGS (4x) combined with genotype imputation to generate high-resolution genomic data suitable for accurate genetic evaluation. The sequencing reads were quality-checked using fastQC (Brown et al. 2017) and adapter trimming was conducted with trimmomatic using the parameters LEADING:3 TRAILING:3 MINLEN:50 to remove low-quality bases and adapter sequences (Bolger et al. 2014). The cleaned reads were then mapped to the *Anas platyrhynchos* reference genome (ZJU1.0, NCBI ID: GCA_015476345.1) using the BWA-MEM software

Table 1. Means and standard deviation (SD) of the morphometric traits recorded on Germanata Veneta and Mignon duck breeds.

			Body weight (g)		Body length (cm)		Shank length (cm)		Shank circumference (cm)		Chest circumference (cm)		Wingspan (cm)	
Breed	Sex	*n*	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Germanata Veneta	F	37	2031.89	215.79	54.96	3.63	7.43	0.49	4.94	0.51	40.26	3.41	54.39	3.68
	M	13	1924.62	152.35	59.85	3.11	7.85	0.47	5.06	0.44	42.58	3.58	56.81	2.25
Mignon	F	41	1035.85	71.52	38.72	1.78	5.13	0.36	3.78	0.33	32.13	2.09	42.41	1.65
	M	9	1004.44	59.81	40.22	2.00	5.50	0.35	3.86	0.37	32.39	2.06	42.67	1.87

v0.7.17 (Pham et al. 2023), followed by duplicate marking using picard (https://github.com/broadinstitute/picard) and base quality score recalibration using GATK (Genome Analysis Toolkit, v4.2).

Variant calling was performed using HaplotypeCaller from GATK and BaseRecalibrator was used for recalibrating base quality scores. Insertion-deletion variants were filtered out to ensure higher genotype accuracy. Initially, 6,324,810 SNPs were identified for each animal, but only those shared among all the animals were retained for further analyses. The raw dataset was checked using the PLINK software (Chang et al. 2015) to remove SNPs with a call rate <95%, SNPs with minor allele frequency (MAF) <5%, and animals with more than 10% of missing genotypes. After applying these filters, 186,287 shared SNPs between GER and MIG remained for downstream analyses. PLINK software was used to calculate observed (Ho) and expected heterozygosity (He), inbreeding coefficient based on homozygosity (F_{HOM}), and inbreeding coefficient based on runs of homozygosity (ROH) values (F_{ROH}) (Chang et al. 2015).

Admixture and genetic relationships

Population structure was inferred using a model-based clustering algorithm implemented in the Admixture software, with K values from 1 to 25 (Alexander et al. 2009). The most likely number of ancestral populations was determined using the cross-validation procedure, minimising the cross-validation error, and finally plotted using the 'BITE2' package of the R software v4.5.0 (Milanesi et al. 2017). Genetic relationships within and between breeds were assessed through genome-wide identity-by-state genetic distances, computed using the 'cluster' command in PLINK (Chang et al. 2015), and subsequently visualised in a multidimensional scaling (MDS) plot, where the first two principal components (PC1 and PC2) were represented using the 'mds-plot' command in PLINK. Phylogenetic relationships between breeds were further explored using Reynolds' genetic distances, calculated with the 'ape' package of the R software (Paradis and Schliep 2019). Based on these genetic distances, neighbour-networks were constructed using FIGTREE (Huson and Bryant 2006). All graphical outputs and visualisations were generated in the R software (Hornik and R Core Team, 2024) with 'tidyverse', 'data.table' and 'ggplot2' packages.

Runs of homozygosity

The ROH in GER and MIG breeds were identified using the PLINK software (Chang et al. 2015) based on the following parameters: (1) minimum ROH length of 1 Mb (–homozyg-kb), (2) allowance for up to two missing SNPs and one heterozygous SNP per ROH (–homozyg-window-missing 2 and –homozyg-window-het 1), (3) minimum 100 SNPs required to define a ROH (–homozyg-snp 100), (4) minimum SNP density of one SNP per 100 kb (–homozyg-density 100), and (5) maximum gap of 1000 kb between consecutive homozygous SNPs (–homozyg-gap 1000) (Cendron et al. 2020; Moscarelli et al. 2021). To refine the identification of shorter ROH and those occurring in genomic regions with variable SNP density, a sliding window approach was applied. Additionally, a consecutive runs method was adopted using the 'detectRUNS' package of the R software (Biscarini et al. 2018). This method scans the genome using sliding windows of 50–100 kb and applies a lower homozygosity threshold of 80% in regions with sparse SNP coverage. By allowing for the detection of shorter or fragmented ROH, this approach ensures the identification of homozygous segments that might otherwise be overlooked due to low SNP density or uneven marker distribution. The parameters for ROH detection with this method were: (1) minimum 20 consecutive SNPs required to define a ROH, (2) allowance for 0–4 missing calls within a ROH segment, (3) maximum 2 heterozygous SNPs per ROH segment, (4) minimum ROH length of 250 kb, and (5) maximum gap of 1 Mb between consecutive ROH segments (Smaragdov 2023; Perini et al. 2025). This approach was particularly effective in detecting long, high-confidence ROH in regions with high SNP density. To ensure robust and reliable ROH calls, the approach focused on identifying continuous homozygous stretches with a minimum length of 500 kb to 1 Mb and a homozygosity threshold of 90%. By prioritising regions with high SNP density, the method enabled the accurate detection of uninterrupted homozygous segments, providing a more comprehensive analysis of ROH patterns in the studied breeds. To enhance robustness and minimise false positives, only ROH segments identified by both methods (the intersection of overlapping calls) were retained for downstream analyses. This consensus strategy aligns with recent workflows in livestock genomics, where intersecting ROH sets improves confidence in detected autozygous regions (Dixit et al. 2020; Smaragdov 2023).

Genome-wide association study

According to the morphometric phenotypes reported in Table 1, no sexual dimorphism was observed within each breed, whereas the MIG regularly showed lower values than the GER for all the traits. A genome-wide

association study (GWAS) was performed using MIG as a case study population and GER as a control population to search for genomic regions associated with body size. Additionally, a sex-based quality check was performed applying the same filtering parameters reported in Perini et al. (2024). The analysis was carried out through the GEMMA software v0.94.128 with linear mixed models, and corrected for the effect of gender and relatedness with the centred relatedness matrix (Zhou and Stephens 2014). The linear mixed model was applied to each chromosome (including sex Z chromosome) by incorporating random genetic effects with genomic relationships to correct for genetic structure in the chosen population. The Bonferroni approach was used to define the genome-wide significance threshold, dividing the standard *p*-value (0.05) by the number of tests as follows: cut-off = $-\log_{10}$(0.05/number of variants). The genomic inflation factor (λ) was calculated as the ratio between the median of the resulting chi-squared test statistics and the expected median of the chi-squared distribution. For this purpose, the 'qqman' package of the R software was used (Purcell et al. 2007). QQplots and λ values are reported in Supplementary Figure 1. For further analysis, only SNPs with significant *p*-value were considered. The 'GALLO' package of the R software was used to annotate the significant SNPs and to inspect genes around the identified SNPs within a range of 0.1 Mb (Fonseca et al. 2020). The annotation was carried out on *Anas platyrhynchos* (mallard) ZJU1.0 reference genomes' version (GCF_015476345.1). Finally, Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) analyses of predicted target genes were estimated using the Database for Annotation, Visualisation and Integrated Discovery (DAVID; https://david.ncifcrf.gov/) (Huang et al. 2007), and only pathways showing raw *p*-value < 0.05 were used for plotting.

Results

Population genetics and inbreeding

Initially, the admixture analysis identified $K = 2$ as the optimal number of populations, revealing a clear distinction between MIG and GER (Figure 1). At $K = 4$, an introgression in GER and a solid distinction of two subgroups in MIG, which is still present in all the subsequent *K,* were observed. A subpopulation structure appreciable from $K = 6$ was also detected in GER (Figure 1). The MDS plot shingly differentiated the two breeds, and all animals retained their breed identity, clustering within their respective group (Figure 2a). According to the neighbor-joining tree (Figure 2b), the MIG clearly showed two subpopulations, whereas the GER is characterised by different families all included in a single subpopulation (Figure 2b), as supported by admixture analysis (Figure 1).

The MAF was similar between the two breeds, with a slightly higher mean value for GER (0.193 ± 0.127) than MIG (0.190 ± 0.132). The He was higher in GER (0.269 ± 0.148) than in MIG (0.238 ± 0.147), indicating greater genetic diversity within the former breed (Table 2). The Ho followed the same trend, with GER having a higher value (0.272 ± 0.260) than MIG (0.245 ± 0.172). Inbreeding coefficients (F_{HOM} and F_{ROH}) differed between the two breeds. The F_{HOM} was higher in MIG (0.369 ± 0.163) than in GER (0.305 ± 0.184), indicating a greater deviation from Hardy-Weinberg equilibrium in MIG. Conversely, the F_{ROH} which reflects the proportion of the genome covered by ROH, was higher in GER (0.372 ± 0.222) than in MIG (0.313 ± 0.126), indicating a greater historical inbreeding burden in GER. The large standard deviations associated with genetic diversity indices suggest wide variation among animals within breed.

ROH islands analysis

A total of 16 ROH islands were detected in the two duck breeds: 7 in the GER and 9 in the MIG (Supplementary material 1). The length of the ROH islands in GER varied from 59,782 bp and 33 SNPs on chromosome 11 to 10,678,426 bp and 115 SNPs on chromosome 2, and the length of the ROH islands in MIG ranged from 585,385 bp and 20 SNPs on chromosome 2 to 8,475,172 bp and 146 SNPs on chromosome 1. Few GER animals shared the same ROH islands, with the maximum value (15% of the total individuals) observed on chromosome 10 (Figure 3). As regards MIG, more animals compared to GER shared the same ROH islands, with the maximum value (30% of the total individuals) observed on chromosome 19 (Figure 3). For each breed, all annotated genes within the ROH underwent to GO and KEGG analysis. In GER breed the GO analysis highlighted pathways belonging to immunity (e.g. GO:0006935 chemotaxis; GO:0004950 chemokine receptor activity) and the KEGG analysis showed pathways related to cytokine-cytokine receptor interaction (Figure 4 and Supplementary Material 1). Similar enrichment was observed in MIG for chemokine receptor activity and GABA activity, although the most valuable GO belonged to muscle movement (GO:0016459 myosin complex; GO:0003774 cytoskeletal motor activity), which was confirmed by the KEGG results

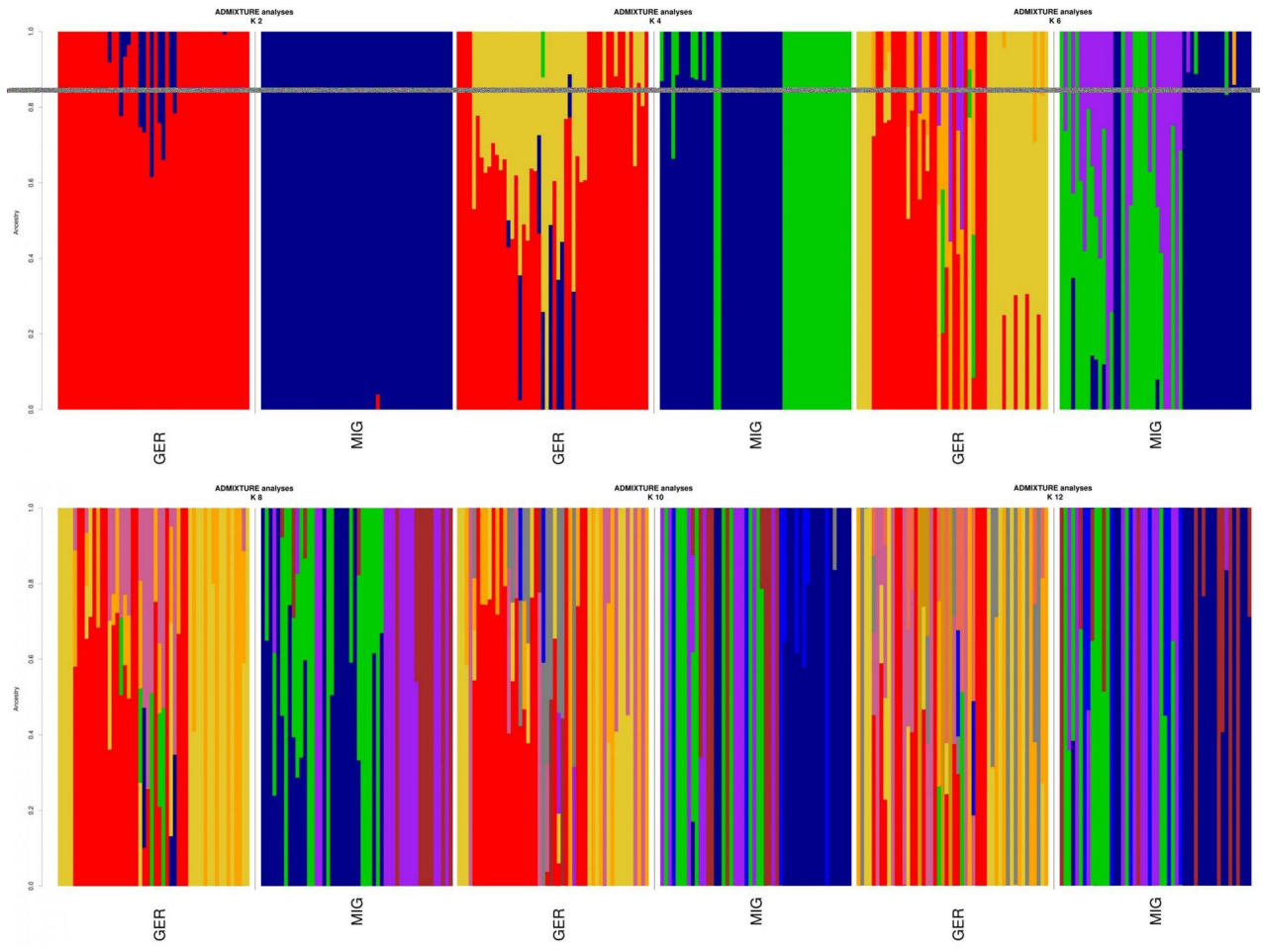

Figure 1. Admixture plot based on different number of assumed ancestors (K) for Germanata Veneta (GER) and Mignon (MIG) duck breeds.

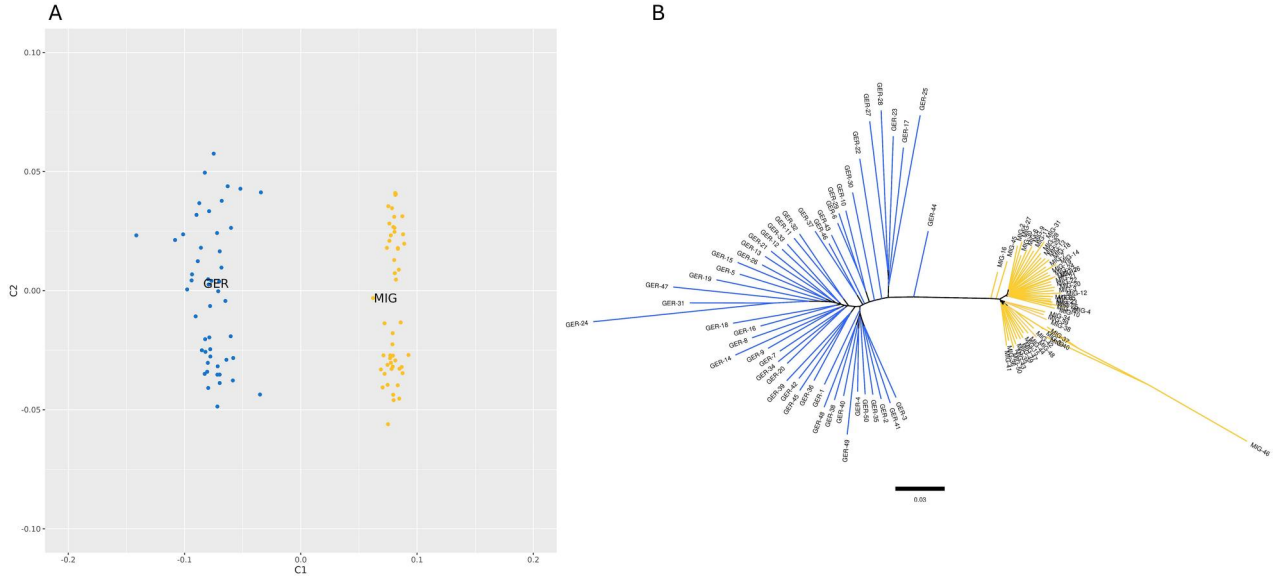

Figure 2. (A) multidimensional scaling plot and (B) neighbour-joining tree plot for Germanata Veneta (GER) and Mignon (MIG) duck breeds.

Table 2. Means and standard deviation (SD) of genetic diversity indices[1] for Germanata Veneta (GER, $n = 50$) and Mignon (MIG, $n = 50$) duck breeds.

	MAF		He		Ho		F_{HOM}		F_{ROH}	
Breed	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
GER	0.193	0.127	0.269	0.148	0.272	0.260	0.305	0.184	0.372	0.222
MIG	0.190	0.132	0.238	0.147	0.245	0.172	0.369	0.163	0.313	0.126

[1]MAF = minor allele frequency; He = expected heterozygosity; Ho = observed heterozygosity; F_{HOM} = inbreeding coefficient based on homozygosity; F_{ROH} = inbreeding coefficient based on runs of homozygosity values.

(cytoskeleton in muscle cells and motor proteins) (Figure 4). In MIG the genes related to the above-mentioned pathways were *SHOX Homeobox* (*SHOX*) on chromosome 1, *Fibronectin Type III Domain Containing 1* (*FNDC1*) and *Insulin Like Growth Factor 2 Receptor* (*IGF2R*) on chromosome 3, and *Protein Kinase CGMP-Dependent 2* (*PRKG2*) and *Secreted Phosphoprotein 1* (*SPP1*) on chromosome 4 (Supplementary material 1).

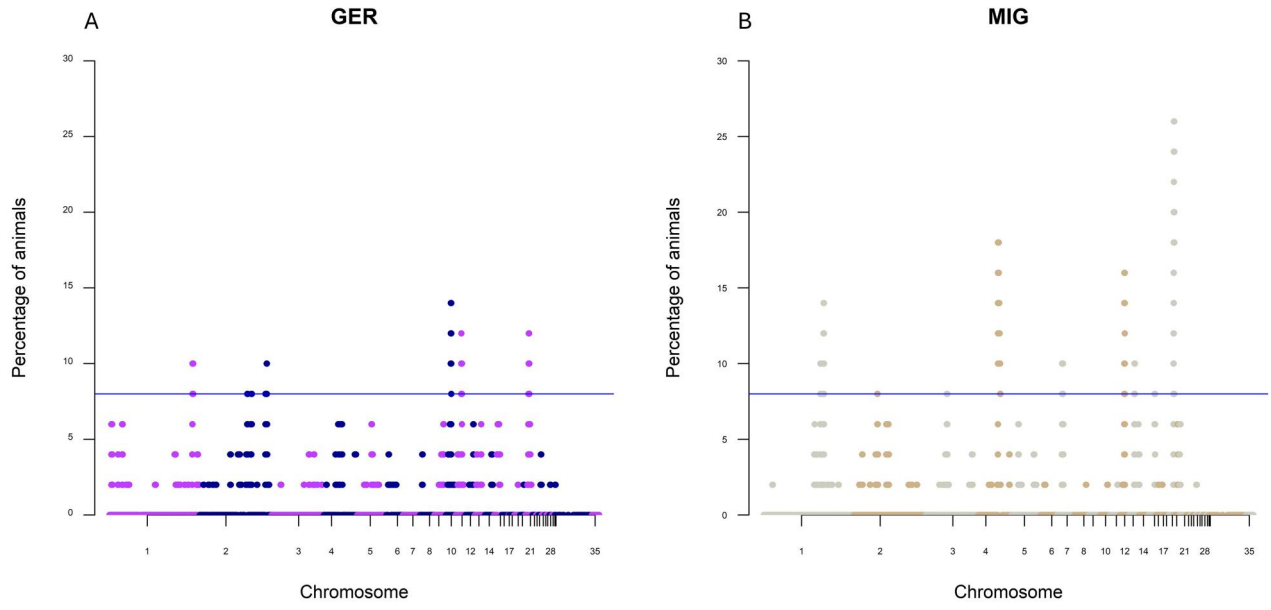

Figure 3. Frequency distribution plot (%) of SNPs in runs of homozygosity islands. (A) Germanata Veneta (GER) breed and (B) Mignon (MIG) breed.

Figure 4. Top enriched traits identified in the GO and KEGG analysis of annotated genes in ROH analysis for Germanata Veneta (GER) and Mignon (MIG) duck breeds. The bubble size represents the number of genes involved in the pathway and the colours indicate the $-\log_{10}$ of false discovery rate. The *x*-axis reports the $-\log_{10}(p\text{-value})$.

GWAS analysis

A total of 314 genes were annotated through the 'GALLO' package in the R software. The two most significant *p*-value signals on Manhattan plot were on chromosomes 3 and 13 (Figure 5). Regarding chromosome 3, the significant SNPs were annotated in genes involved in the KEGG fatty acid elongation pathway, such as *hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex subunit alpha and beta* (*HADHA* and *HADHB*) (Figure 6), and chromosome 13 which included two main genes with unknown activity and role, namely *PDZ Domain Containing Family Member 1* (*GIPC1*) and *A-Kinase Anchoring Protein 8* (*AKAP8*). Other genes were annotated around SNPs of interest in the same chromosome: *POC1 Centriolar Protein* A (*POC1A*), *Ribosomal RNA Processing 9, U3 Small Nucleolar RNA Binding Protein* (*RRP9*) and *Glutamate Metabotropic Receptor 2* (*GRM2*) (Supplementary Material 2).

Discussion

The genetic diversity indices calculated for GER and MIG duck breeds pointed out their high level of inbreeding. Carcò et al. (2018) conducted a study on GER and MIG using a panel of 23 microsatellite markers and reported greater heterozygosity values compared to our results. Although both breeds are under conservation, the risk of genetic erosion remains significant and the beneficial effect of these programs on inbreeding mitigation is limited. A noteworthy result is that Ho was slightly higher than He for both GER and MIG, which aligns with findings of Carcò et al. (2018) on the same duck breeds. However, this result is unusual, as small populations typically exhibit the opposite trend due to genetic drift and inbreeding (Restoux et al. 2022). This indicates that non-random mating or recent gene flow may have influenced heterozygosity levels. Similar patterns were observed in indigenous Asian duck breeds (Huang et al. 2025), in which Ho exceeded He, indicating the presence of segregation-destroying effect. This phenomenon suggests that the population may be a mixture of previously segregated subpopulations. Additionally, the populations that exhibited similar levels of He and Ho in Huang et al. (2025) shared high inbreeding coefficients, as in our study. Nevertheless, the higher F_{HOM} and F_{ROH} in GER and MIG suggested high levels of inbreeding. The GER breed had the highest values, potentially reflecting a history of matings between relatives or a reduction of population size (Evelyne et al. 2022). The admixture analysis revealed that the sub-population structure is present in both GER and MIG. While MIG sub-population

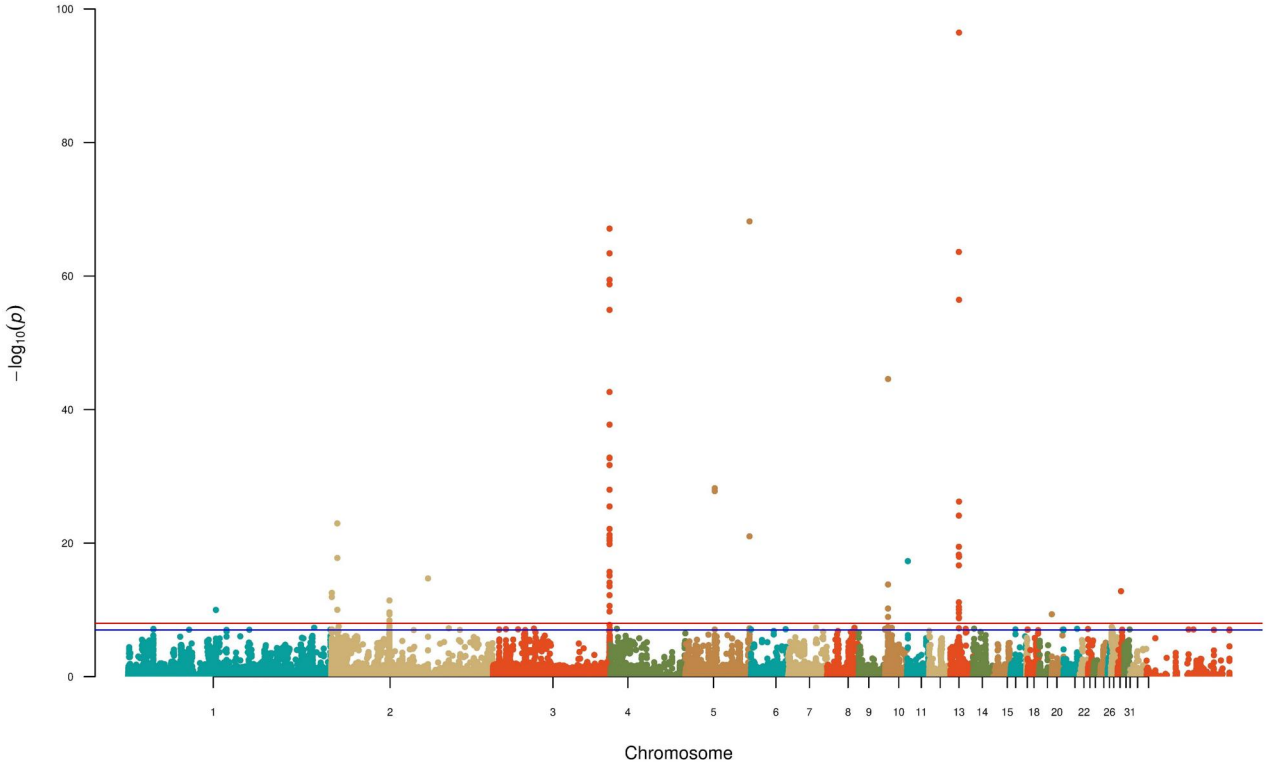

Figure 5. Manhattan plot of GWAS analysis. The red line represents the *p*-value $< -\log_{10}(1e-8)$ and the blue line represents suggestive significance, with a value of $-\log_{10}(1e-7)$.

Figure 6. Top enriched traits identified in the GO and KEGG analysis of annotated genes in GWAS analysis. The bubble size represents the number of genes involved in the pathway and the colours represent the $-\log_{10}$ of false discovery rate. The *x*-axis reports the $-\log_{10}$(*p*-value).

structure is corroborated by neighbour-joining and MDS analyses, GER had a distinct sub-structure at $K = 6$ in the admixture analysis, which was not evident in the other analyses. Both breeds were sampled in two flocks and this could explain the highlighted sub-population structure. In GER, despite different ancestries, controlled mating practices prevented the emergence of substructure in neighbour-joining and MDS analyses. In contrast, MIG experienced a clear separation within the breed. This type of substructure is likely due to the endangered status of the breed and the consequent difficulty in finding new breeders to integrate into the flocks. This also explains the level of inbreeding reported in this study.

The ROH islands were analysed for both breeds to identify genomic regions under potential selective pressure. The GO analysis for the GER duck breed highlighted an enrichment of immune-related pathway, indicating that selection has favoured immune system function. As regards the MIG duck breed, findings were more intriguing as ROH analysis identified several genes associated with body size regulation. Notably, the *SHOX* gene on chromosome 1 plays a crucial role during development, as it is involved in cell cycle regulation and functions as a transcriptional activator in osteogenic cells in humans (Rao et al. 2001). The *SHOX* gene deficiency and haploinsufficiency have been linked to skeletal abnormalities, including short stature, in humans (Wu et al. 2009). Furthermore, *SHOX* has been associated with *CRLF2* (*type 1 cytokine receptor*) in the context of skeletal atavism in Shetland ponies (Rafati et al. 2016). Notably, *CRLF2* is located within the ROH islands on chromosome 1 of the MIG breed.

Additional genes annotated in the ROH islands on chromosomes 3 and 4 of the MIG breed associated

with body size and skeletal/muscle development. On chromosome 3, the *FNDC1* gene is noteworthy, as its expression has been linked to femoral head separation, a condition impacting growth in broilers (Ibelli et al. 2022). Given that the plumage colour of the MIG breed is entirely white, the expression of the *FNDC1* gene appears to be correlated with white feathering, like in Muscovy ducks (Sun et al. 2024). Also located on chromosome 3 is the *IGF2R* gene, a negative regulator of the *IGF2* gene (Farmer et al. 2013) which plays a significant role in muscle mass development in pigs (Van Laere et al. 2003) and mice (Clark et al. 2015). Furthermore, Xu et al. (2019) observed differential expression of *IGF2R* between two duck populations artificially selected for lean and fat meat, suggesting that it is a candidate gene for the regulation of skeletal muscle mass in ducks. On chromosome 4, the *PRKG2* gene may contribute to the small body size observed in MIG. Notably, a mutation in the *PRKG2* gene is associated with dwarfism in Angus cattle (Koltes et al. 2009). Similarly, the *SPP1* gene on chromosome 4 was identified as significant in MIG and is documented in the literature as a regulator of muscle development (Liu et al. 2020). Specifically, a transcriptome study in ducks identified *SPP1* as a key gene involved in skeletal muscle development. The genes identified for MIG in the ROH analysis suggest genomic pressure on this breed, resulting in a reduction of body size, known as mignon. Additional evidence from the GWAS analysis on chromosome 13, which yielded highly significant values, underlines the involvement of the *POC1A* gene, which is implicated in primordial dwarfism in humans (Koparir et al. 2015) and skeletal dysplasia in mice (Geister et al. 2015). In the comparison between GER and MIG duck breeds through GWAS, morphological traits differed substantially. Specifically, MIG and GER animals had an average body weight of ∼1000 g and ∼2000 g, respectively. Consequently, *POC1A* may be considered a significant factor contributing to the small body size of the MIG breed.

Conclusions

The genetic analysis of the two duck breeds revealed high levels of inbreeding, with GER having higher historical inbreeding burden. Admixture and neighbour-joining analyses confirmed subpopulation structures in MIG, whereas GER showed family-based differentiation. The ROH analysis highlighted immune-related genes in GER, while MIG had genes associated with body size, suggesting selective pressures towards smaller stature. The GWAS further supported these results, with *POC1A* potentially playing a key role in the reduced body size of MIG. To further validate our findings, it is essential to undertake an in-depth tissue expression study. These findings point out the importance of conservation efforts to manage genetic diversity and mitigate inbreeding risks in both breeds.

Ethical approval

Blood samples were collected during routine health controls by the public veterinary service and thus ethical approval was not required for the current study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by Ministero dell'Agricoltura, della Sovranità Alimentare e delle Foreste (MASAF) [PSRN 2014–2022, Sottomisura 10.2 Biodiversità/Comparto avicoli – project n. 2015.99.42203.5924 – 'TuBAvI-2'].

ORCID

Francesco Perini http://orcid.org/0000-0003-2235-3926
Enrico Bortoletto http://orcid.org/0000-0002-7226-4887
Filippo Cendron http://orcid.org/0000-0002-8221-7566
Umberto Rosani http://orcid.org/0000-0003-0685-1618
Emiliano Lasagna http://orcid.org/0000-0003-2725-2921
Martino Cassandro http://orcid.org/0000-0002-8709-2870
Mauro Penasa http://orcid.org/0000-0001-9984-8738

Data availability statement

The data that support the findings presented in this study are available from the corresponding author upon reasonable request.

References

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19(9):1655–1664. doi: 10.1101/gr.094052.109.

Besbes B, Tixier-Boichard M, Hoffmann I, Jain GL. 2007. Future trends for poultry genetic resources. In: Proceedings of the International Conference of Poultry in the 21st Century: Avian influenza and beyond, Bangkok, Thailand, 5–7 November.

Biscarini F, Cozzi P, Gaspa G, Marras G. 2018. detectRUNS: an R package to detect runs of homozygosity and heterozygosity in diploid genomes. https://orca.cardiff.ac.uk/108906/.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30(15): 2114–2120. doi: 10.1093/bioinformatics/btu170.

Brown J, Pirrung M, McCue LA. 2017. FQC Dashboard: integrates FastQC results into a web-based, interactive, and

extensible FASTQ quality control tool. Bioinformatics. 33(19):3137–3139. doi: 10.1093/bioinformatics/btx373.

Carcò G, Grajewski B, Cassandro M, Lisowski M, Szwaczkowski T. 2018. Genetic variability of some Italian and Polish duck breeds. Ital J Anim Sci. 17(4):899–906. doi: 10.1080/1828051X.2018.1436006.

Cendron F, Perini F, Mastrangelo S, Tolone M, Criscione A, Bordonaro S, Iaffaldano N, Castellini C, Marzoni M, Buccioni A, et al. 2020. Genome-wide SNP analysis reveals the population structure and the conservation status of 23 Italian chicken breeds. Animals. 10(8):1441. doi: 10.3390/ani10081441.

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 4(1):7. doi: 10.1186/s13742-015-0047-8.

Clark DL, Clark DI, Hogan EK, Kroscher KA, Dilger AC. 2015. Elevated insulin-like growth factor 2 expression may contribute to the hypermuscular phenotype of myostatin null mice. Growth Horm IGF Res. 25(5):207–218. doi: 10.1016/j.ghir.2015.06.007.

Dixit SP, Singh S, Ganguly I, Bhatia AK, Sharma A, Kumar NA, Dang AK, Jayakumar S. 2020. Genome-wide runs of homozygosity revealed selection signatures in *Bos indicus*. Front Genet. 11:92. doi: 10.3389/fgene.2020.00092.

Evelyne M, Nguyen Van D, Amelyne D, Nguyen Hoang T, Duc LD, Nassim M, Dinh TV, Frédéric F. 2022. High-resolution genomic analysis of four local Vietnamese chicken breeds. J Anim Breed Genet. 139(5):583–595. doi: 10.1111/jbg.12723.

FAO. 2012. Phenotypic characterization of animal genetic resources. FAO Animal Production and Health Guidelines No. 11. Rome. Available at www.fao.org/docrep/015/i2686e/i2686e00.pdf.

Farmer WT, Farin PW, Piedrahita JA, Bischoff SR, Farin CE. 2013. Expression of antisense of insulin-like growth factor-2 receptor RNA non-coding (AIRN) during early gestation in cattle. Anim Reprod Sci. 138(1–2):64–73. doi: 10.1016/j.anireprosci.2013.01.009.

Fonseca PAS, Suárez-Vega A, Marras G, Cánovas Á. 2020. GALLO: an R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. GigaScience. 9(12):1–9. doi: 10.1093/gigascience/giaa149.

Geister KA, Brinkmeier ML, Cheung LY, Wendt J, Oatley MJ, Burgess DL, Kozloff KM, Cavalcoli JD, Oatley JM, Camper SA. 2015. LINE-1 mediated insertion into *Poc1a* (Protein of Centriole 1 A) causes growth insufficiency and male infertility in mice. PLoS Genet. 11(10):e1005569. doi: 10.1371/journal.pgen.1005569.

Ghildiyal K, Nayak SS, Rajawat D, Sharma A, Chhotaray S, Bhushan B, Dutt T, Panigrahi M. 2023. Genomic insights into the conservation of wild and domestic animal diversity: a review. Gene. 886:147719. doi: 10.1016/j.gene.2023.147719.

Hornik. and R Core Team. 2024. The R FAQ. https://CRAN.R-project.org/doc/manuals/R-FAQ.html.

Hu Z, Liu J, Cao J, Zhang H, Liu X. 2021. Ovarian transcriptomic analysis of black Muscovy duck at the early, peak and late egg-laying stages. Gene. 777:145449. doi: 10.1016/j.gene.2021.145449.

Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC, et al. 2007. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 35:169–175. doi: 10.1093/nar/gkm415.

Huang M, Zhou J, Yihao Z, Luo K, Zheng S, Tang H, Wu Y, Xuan R, Huang Y, Li J, et al. 2025. Whole genome sequencing revealed genetic structure, domestication, and selection of Chinese indigenous ducks. Int J Biol Macromol. 306:141724. doi: 10.1016/j.ijbiomac.2025.141724.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 23(2):254–267. doi: 10.1093/molbev/msj030.

Ibelli AMG, Peixoto J, Zanella R, Gouveia J, Cantão ME, Coutinho LL, Marchesi JAP, Pizzol M, Marcelino DEP, Ledur MC. 2022. Downregulation of growth plate genes involved with the onset of femoral head separation in young broilers. Front Physiol. 13:941134. doi: 10.3389/fphys.2022.941134.

Koltes JE, Mishra BP, Kumar D, Kataria RS, Totir LR, Fernando RL, Cobbold R, Steffen D, Coppieters W, Georges M, et al. 2009. A nonsense mutation in cGMP-dependent type II protein kinase (*PRKG2*) causes dwarfism in American Angus cattle. Proc Natl Acad Sci USA. 106(46):19250–19255. doi: 10.1073/pnas.0904513106.

Koparir A, Karatas OF, Yuceturk B, Yuksel B, Bayrak AO, Gerdan OF, Sagiroglu MS, Gezdirici A, Kirimtay K, Selcuk E, et al. 2015. Novel *POC1A* mutation in primordial dwarfism reveals new insights for centriole biogenesis. Hum Mol Genet. 24(19):5378–5387. doi: 10.1093/hmg/ddv261.

Liu H, Chen Y, Zhu X, Zhang M, Li W. 2020. Identification of key genes and pathways associated with chronic obstructive pulmonary disease based on bioinformatics analysis. In Proceedings of the Fourth International Conference on Biological Information and Biomedical Engineering. New York, NY: Association for Computing Machinery, p. 1–7. doi: 10.1145/3403782.3403803.

Milanesi M, Capomaccio S, Vajana E, Bomba L, Garcia JF, Ajmone-Marsan P, Colli L. 2017. BITE: an R package for biodiversity analyses. BioRxiv. 181610 doi: 10.1101/181610.

Moscarelli A, Sardina MT, Cassandro M, Ciani E, Pilla F, Senczuk G, Portolano B, Mastrangelo S. 2021. Genome-wide assessment of diversity and differentiation between original and modern Brown cattle populations. Anim Genet. 52(1):21–31. doi: 10.1111/age.13019.

Nielsen SS, Alvarez J, Bicout DJ, Calistri P, Canali E, Drewe JA, Garin-Bastuji B, Gonzales Rojas JL, Schmidt CG, Herskin M, et al. 2023. Welfare of ducks, geese and quail on farm. Efsa J. 21(5) doi: 10.2903/j.efsa.2023.7992.

Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 35(3):526–528. doi: 10.1093/bioinformatics/bty633.

Perini F, Cendron F, Castellini C, Iaffaldano N, Marzoni M, Buccioni A, Soglia D, Schiavone A, Cerolini S, Lasagna E, et al. 2025. Genome-wide analysis of Collo Nudo Italiana and Millefiori Piemontese local chicken breeds: genetic variability and structure analysis in the context of Italian chicken biodiversity. Ital J Anim Sci. 24(1):137–148. doi: 10.1080/1828051X.2024.2445172.

Perini F, Cendron F, Lasagna E, Cassandro M, Penasa M. 2024. Genomic insights into shank and eggshell color in Italian local chickens. Poult Sci. 103(6):103677. doi: 10.1016/j.psj.2024.103677.

Pham M, Tu Y, Lv X. 2023. Accelerating BWA-MEM read mapping on GPUs. In Proceedings of the 37th International Conference on Supercomputing, Orlando, USA, 21–23 June, pp. 155–166. doi: 10.1145/3577193.3593703.

Pingel H. 2011. Waterfowl production for food security. Lohmann Inf. 46(2):32–42.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 81(3):559–575. doi: 10.1086/519795.

Rafati N, Andersson LS, Mikko S, Feng C, Raudsepp T, Pettersson J, Janecka J, Wattle O, Ameur A, Thyreen G, et al. 2016. Large deletions at the SHOX locus in the pseudoautosomal region are associated with skeletal atavism in shetland ponies. G3 Genes, Genomes, Genet. 6(7):2213–2223. doi: 10.1534/g3.116.029645.

Rao E, Blaschke RJ, Marchini A, Niesler B, Burnett M, Rappold GA. 2001. The Leri-Weill and Turner syndrome homeobox gene SHOX encodes a cell-type specific transcriptional activator. Hum Mol Genet. 10(26):3083–3091. doi: 10.1093/hmg/10.26.3083.

Restoux G, Rognon X, Vieaud A, Guemene D, Petitjean F, Rouger R, Brard-Fudulea S, Lubac-Paye S, Chiron G, Tixier-Boichard M. 2022. Managing genetic diversity in breeding programs of small populations: the case of French local chicken breeds. Genet Sel Evol. 54(1):56. doi: 10.1186/s12711-022-00746-2.

Smaragdov MG. 2023. Identification of homozygosity-rich regions in the Holstein genome. Vestn VOGiS. 27(5):471–479. doi: 10.18699/VJGB-23-57.

Sponenberg DP, Martin A, Couch C, Beranger J. 2019. Conservation strategies for local breed biodiversity. Diversity. 11(10):177. doi: 10.3390/d11100177.

Sun GB, Lu YF, Duan XJ. 2024. Investigation into the association of *FNDC1* and *ADAMTS12* gene expression with plumage coloration in Muscovy ducks. Open Life Sci. 19(1):2022877. doi: 10.1515/biol-2022-0877.

Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, Archibald AL, Haley CS, Buys N, Tally M, et al. 2003. A regulatory mutation in *IGF2* causes a major QTL effect on muscle growth in the pig. Nature. 425(6960):832–836. doi: 10.1038/nature02064.

Wang C, Li SJ, Li C, Yu GH, Feng YP, Peng XL, Gong YZ. 2013. Molecular cloning, expression and association study with reproductive traits of the duck LRP8 gene. Br Poult Sci. 54(5):567–574. doi: 10.1080/00071668.2013.819488.

Wu XL, Gu MM, Huang L, Liu XS, Zhang HX, Ding XY, Xu JQ, Cui B, Wang L, Lu SY, et al. 2009. Multiple synostoses syndrome is due to a missense mutation in exon 2 of FGF9 gene. Am J Hum Genet. 85(1):53–63. doi: 10.1016/j.ajhg.2009.06.007.

Xu T, Gu L, Yu H, Jiang X, Zhang Y, Zhang X, Rong G, Zhou Z, Schachtschneider KM, Hou S. 2019. Analysis of *Anasplatyrhynchos* genome resequencing data reveals genetic signatures of artificial selection. PLoS ONE. 14(2):e0211908. doi: 10.1371/journal.pone.0211908.

Zhou X, Stephens M. 2014. Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nat Methods. 11(4):407–409. doi: 10.1038/nmeth.2848.