scientific reports

Check for updates

OPEN

An analysis of the genomic architecture of the worldwide Irish Wolfhound dogs

Francesca Bernini[1], Stefano P. Marelli[1], Marcello Poli[2,3] & Maria G. Strillacci[1]✉

The Irish Wolfhound (IW) is a dog breed characterised by a complex demographic history and reduced population size. In this study, we combined multiple population genomic approaches to characterise the genetic structure of 96 dogs collected from 23 countries worldwide, genotyped using the Illumina CanineHD BeadChip. Analyses of effective population size (Ne), linkage disequilibrium (LD) decay, heterozygosity and principal component analysis (PCA) consistently revealed limited genetic diversity. Complementary analyses of runs of homozygosity (ROH) and integrated haplotype score (iHS) identified extended homozygous segments and signatures of selection across the genome. ROH were predominantly short in length, and 40 samples showed ROH longer than 8 Mb. No ROH exceeding 16 Mb were detected, suggesting that these patterns reflect long-term demographic processes and historical selection rather than exclusively recent inbreeding. Particularly, 26 ROH islands were shared by at least 85% of the analysed individuals, with 3 ROHs shared by 100% of the population. Several ROH islands overlapped with regions previously reported in other hunting dog breeds and harboured genes associated with morphology, behaviour and diseases of major relevance to IWs, including osteosarcoma. Genomic regions identified by iHS also include genes involved in cancer and immune response. Compared with a previous IW population with publicly available genotypes, the dogs analysed here represent a more homogeneous subgroup. Overall, all approaches converged on a coherent genomic scenario, which highlights the combined effects of demographic history and selection in shaping the current genetic architecture of the IWs.

Keywords Irish Wolfhound dogs, Genetic variability, Inbreeding, Selection signature

Irish Wolfhounds (IWs) are a globally recognized breed, rich in history and distinctive traits (Fig. 1A and B). The simple idiom 'Lambs at home, lions in the chase', frequently used by breeders, aptly summarizes the behavioural and temperamental traits of the IWs. Today IWs are involved in coursing and racing events, sports that highlight their speed, agility, and endurance (Fig. 1C). In lure coursing, they chase a mechanically operated lure across open fields, simulating the pursuit of prey, while in racing, they compete over short distances, testing their sprinting ability. Although these activities are less intense than the hunting work for which the breed was originally developed, they provide opportunities for IWs to exercise their natural athleticism in a controlled and safe environment.

The ancestors of this breed were selected by Irish Celts for hunting wolves and deer. The first historical written description of the breed dates back to 391 AD, when the Roman consul Quintus Aurelius Symmachus described this precious gift to the Roman emperor[1]. From the Middle Ages to the 17th century, IWs were considered highly prized gifts and effective wolf hunters. A renewed interest in the breed emerged in the last decades of the 19th century, supported by Irish nationalism, when the breed became a living symbol of Irish culture and Celtic heritage. To save the breed from extinction, Captain G.A. Graham initiated planned outcrosses in 1862, using Scottish Deerhounds (descendants of the original IWs), Borzoi, and Great Danes[2]. The traceability of the breed's historical origins is consistent with the genomic analyses conducted by Li et al.[3], which demonstrated the clustering of IWs with other European breeds, with the Scottish Deerhound being their closest relative among sighthounds.

The first official breed standard was published in 1885 (*FCI Standard No. 160, 02.04.2001*). Since the breed's official recognition, pedigree data for IWs have been meticulously recorded. These records provide valuable

[1]Department of Veterinary Medicine and Animal Sciences, Università degli Studi di Milano, Lodi 26900, Italy. [2]Sighthound Club of Italy, Trento 38121, Italy. [3]Federation of Irish Wolfhound Clubs (FIWC), Harreveld, Netherlands. ✉email: maria.strillacci@unimi.it

Fig. 1. Irish Wolfhound dogs: (**A**) male; (**B**) female; (**C**) male during athletic activity. Photos provided by M. Poli (one of the authors).

tools for genealogical analyses aimed at preserving genomic diversity, as well as improving health and welfare[4]. However, selective matings within closed populations, together with historical genetic bottlenecks, have led to a strong reduction in genetic variability and increased inbreeding coefficients[5]. In this context, pedigree information alone may be insufficient to fully capture the nowadays breed's genetic structure and to monitor inbreeding accurately. In recent decades, the advent of genomic technologies has provided powerful tools to complement pedigree data and to study domestic animal populations with unprecedented resolution. High-density SNP genotyping arrays enable genome-wide analyses of genetic variation in domestic animals, complementing pedigree data and providing unprecedented resolution. They allow the identification of runs of homozygosity (ROHs), which are contiguous stretches of homozygous genotypes reflecting inbreeding across different timescales[6]. The length and frequency of ROHs can inform about demographic history, levels of autozygosity, and regions under selection[7]. SNP-based data also support accurate estimation of genomic inbreeding coefficients and effective population size (Ne), key parameters for monitoring genetic diversity and guiding sustainable breeding strategies[8]. When combined with pedigree information, these genomic analyses improve the precision of population assessments and help identify deleterious or adaptive variants[9]. In addition to ROH-based measures, genome-wide SNPs enable the detection of recent or ongoing selection using statistics such as the integrated haplotype score (iHS)[10]. Unlike ROHs, iHS is particularly sensitive to sweeps not yet fixed, allowing the identification of adaptive loci at intermediate frequencies. Together with ROH patterns, nucleotide diversity, and Ne estimates, iHS provides complementary insights into the interplay of drift, selection, and breeding management, aiding the reconstruction of population history and the identification of regions linked to fitness and breed-specific traits[11].

This study provides a comprehensive characterization of the genomic architecture of the Irish Wolfhound by (i) investigating population structure and pairwise genetic relationships, (ii) quantifying genetic diversity and genomic inbreeding levels, and (iii) identifying genomic regions showing signatures of selection. These analyses offer a genomic framework for understanding the breed's genetic background and for supporting its sustainable breeding and conservation.

Results

Genotyping IWs using a high-density canine SNP array (230 K SNPs) revealed a very high proportion of observed homozygosity, ranging from 80.5% to 85.5%. These values are consistent with the well-documented demographic history of the IWs and reflect the reduced genomic diversity characteristic of this breed[12]. The PCA (Fig. 2, Panel_1A) shows the genetic variation among IWs sampled across 23 countries. The first principal component

(PC_1) explains 14.7% of the total genetic variance, while the second principal component (PC_2) accounts for 11.5%. Individuals cluster tightly together with no evident country-specific substructure, indicating that IWs represent a largely homogeneous genetic population with minimal geographic differentiation. Population-level observed (Ho = 0.177) and expected (He = 0.181) heterozygosity indicate a low level of genetic diversity within the IW breed, consistent with the high levels of homozygosity observed. The historical trend of the Ne in the IW breed is shown in Fig. 2 (Panel_1B). The results indicate a marked decline in effective population size toward recent generations, followed by a progressive increase when moving backward in time. The most recent Ne estimate obtained with SNeP is 56, corresponding to approximately 13 generations ago, whereas values exceeding 200 are observed at more ancient time points.

The complementary LD-bin analysis targeting shorter genomic distances revealed higher Ne estimates at the most recent inferred time intervals (i.e., intervals derived from LD at short intermarker distances). The most recent estimate corresponds to the shortest distance class and reflects very recent ancestry rather than a precise single generation; specifically, the Ne estimate for the most recent interval was approximately 293, while estimates for the following interval ranged between 193 and 330, depending on the distance bins considered. These values show greater variability, reflecting the stochastic nature of LD at very short distances and the limited resolution of SNP-based estimates at extremely recent time scales.

Nucleotide diversity (π) and Linkage Disequilibrium (LD) decay analysis

As showed in Figure S1, the genome-wide plots of nucleotide diversity (π) provide a detailed visualization of how genetic variation is distributed across all the autosomes in the IWs. As expected for a breed that has undergone strong historical bottlenecks, π values are generally low throughout the genome, with most windows showing extremely reduced diversity ($\pi \approx 0.00002–0.00008$). This pattern is consistent with the high level of homozygosity observed through ROH analyses and reflects the long-term consequences of restricted effective population size.

Linkage disequilibrium (LD) decay analysis revealed a slow decline of pairwise r^2 values with increasing physical distance between SNPs (Figure S2). The initial LD was high ($r^2 \approx 0.8$ at < 10 kb), decreasing to ~0.3 at approximately 200–300 kb. Notably, r^2 values remained above 0.2 up to nearly 1 Mb and approached background levels ($r^2 \approx 0.15–0.18$) only beyond 1–2 Mb. This pattern indicates extensive long-range haplotype structure and is consistent with the reduced nucleotide diversity, and high ROH burden.

Runs of Homozygosity detection

Overall, the two detection methods produced highly comparable ROH profiles. Given the substantial overlap between the segments identified with these two stringent approaches, only the results from the consecutive method are presented here (the sliding window results are reported in the Supplementary file, Figure S3). A total of 31,337 ROHs were detected, ranging from 220 to 433 per individual (mean: 326.42 ± 26.10). ROH lengths varied between 1.0 Mb and approximately 10.5 Mb. The total autosomal length encompassed by ROH per individual ranged from 328 Mb (14.5% of the canine autosomal genome) to 987 Mb (44.3% of the autosomal genome) (Fig. 2, Panel_2A).

The frequency distribution of ROH across length classes was nearly identical for both methods: 61% of ROH were shorter than 2 Mb, 30.0% were between 2 and 4 Mb, 0.8% between 4 and 8 Mb, and only 0.02% between 8 and 16 Mb. No ROH longer than 16 Mb were detected with either approach, indicating the absence of very recent inbreeding events in the analysed population[7,13] (Fig. 2, Panel_2B). Figure 2, Panel_2C illustrates the relationship between the number of ROHs and the total ROH length across samples, providing insight into the distribution of homozygous segments within individuals.

Overall, the similar distribution patterns demonstrate that both approaches provide consistent estimates of the extent and genomic structure of homozygosity, with the consecutive method showing a slightly higher sensitivity to short or borderline segments, whereas the sliding-window method appears more conservative in filtering out spurious ROHs.

A total of 26 ROH islands (defined by 3376 homozygous SNPs) were identified across the genome of the IWs population (Table 1; Fig. 2, Panel_2D). Among these, two regions were detected exclusively with the consecutive method, while the remaining 24 were shared by both detection approaches. Across all ROH islands, a total of 197 annotated genes were identified. Notably, three genomic regions were present in 100% of the individuals, representing the most conserved homozygous segments in the breed and suggesting strong historical selection or genetic drift leading to fixation.

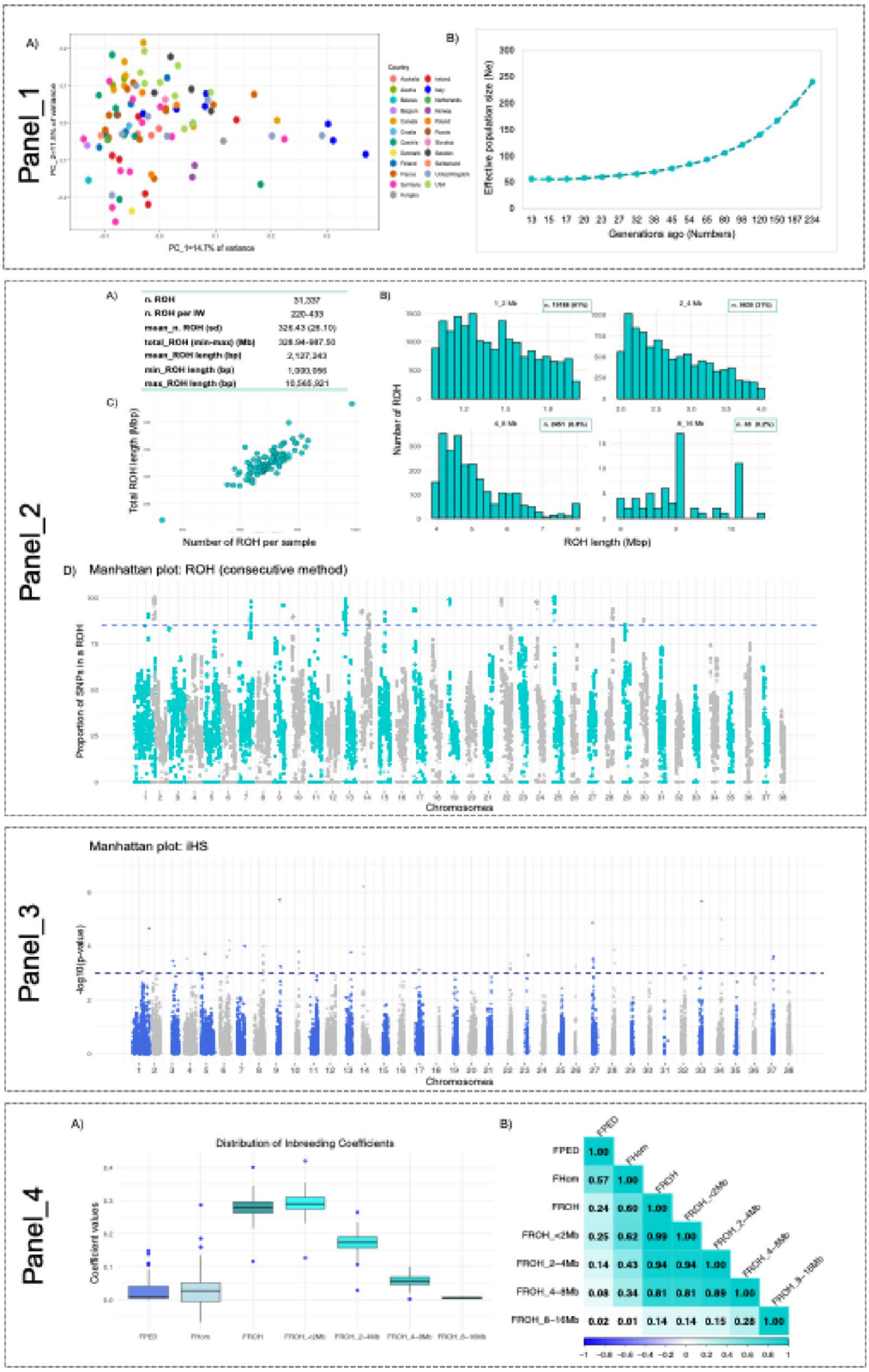
Table 1 provides detailed information on the ROH islands, including genomic coordinates, SNP density, and gene content.

iHS analysis

A total of 86,111 phased autosomal SNPs distributed across 38 chromosomes were evaluated using the iHS statistic. While several individual SNPs exceeded the conservative significance threshold ($-\log10(p) \geq 3$) across the genome (Fig. 2; Panel_3), candidate regions were defined using a clustering approach requiring at least five extreme SNPs within a 1 Mb window.

Under these criteria, three high-confidence candidate regions of recent positive selection were detected on chromosomes 8, 10, and 22 (Table 2). These regions spanned on average 1.64 Mb (median: 1.64 Mb; range: 1.28–1.99 Mb) and contained an average of 1,356 SNPs. The strongest iHS signal ($-\log10\ p = 3.81$) was observed on chromosome 8 at 59.4 Mb, indicating a localized and pronounced deviation from neutrality.

Functional enrichment analyses of the genes annotated in ROH islands (n. 197 genes) and genomic regions identified by iHS (n. 14 genes) are reported in Table S2.



Inbreeding coefficients: descriptive statistics and correlations

The descriptive statistics of the inbreeding coefficients are summarized in Table 3 and graphically represented in Fig. 2, Panel_4A, together with the correlations calculated between each pair of coefficients (Panel_4B).

The combined pedigree dataset includes 96 IWs, each with five complete generations of ancestors. When merged, the dataset spans 14 generations and 2,343 individuals, reflecting overlapping ancestral lineages across the sampled population. The average pedigree-based inbreeding coefficient was 0.025, reflecting the level of relatedness captured within the generations of recorded ancestry.

◂**Fig. 2**. Panel_1): (A) Principal component analysis of 96 IW samples collected form 23 countries, (B) Effective population size of the IW breed; Panel_2) ROH analysis results: (A) Descriptive statistics of identified ROH, (B) ROH distribution by classes of length, (C) Relation between number of ROH and total ROH length per sample, (D) Manhattan plot showing the proportion of SNPs in ROH across the genome, which corresponds to the % of individuals in which each SNP is homozygous within a ROH. The blue line indicates the 85% threshold, above which SNPs are considered part of an ROH island; Panel_3) Manhattan plot of iHS with blue line set at 3 (-log10p-value); Panel_4) (A) Relationships and distributions of inbreeding coefficients estimated from pedigree and genomic data: Boxplots showing the distribution of the inbreeding coefficients across individuals; (B) Correlation matrix among inbreeding coefficients.

Chr	nSNP	From	To	Genes
1	85	78,000,733	79,320,730	
2*	231	12,971,540	16,091,790	KIF5B, ARHGAP12, ZEB1, ZNF438, SVIL, *JCAD*, *MTPAP*, *MAP3K8*, BAMBI, WAC, MPP7
2	132	19,972,339	21,746,601	TRDMT1, CUBN, RSU1, PTER, C1QL3, MINDY3, ITGA8, FAM171A1
7	82	65,205,683	66,137,236	RBBP8, GATA6
7	122	68,172,186	69,637,408	METTL4, NDC80, SMCHD1, EMILIN2
9	*115*	*51,470,031*	*52,482,812*	*RXRA*, *COL5A1*, *FCN2*, *OLFM1*, *PPP1R26*, *C9H9orf116*, *MRPS2*, *GBGT1*, *RALGDS*, *CEL*, *GTF3C5*, *GFI1B*, *TSC1*, *SPACA9*, *AK8*
10	77	7,425,726	8,355,539	RASSF3, GNS, TBC1D30, WIF1, LEMD3, MSRB3
13	296	184,812	3,883,373	LAPTM4B, MATN2, RPL30, ERICH5, RIDA, POP1, NIPAL2, KCNS2, STK3, OSR2, VPS13B, MIR599, MIR875, RGS22, FBXO43, POLR2K, SPAG1, RNF19A, ANKRD46, SNX31, PABPC1, YWHAZ, ZNF706, GRHL2, NCALD
13	173	5,475,801	7,480,276	RIMS2, DCSTAMP, DPYS, LRP12, ZFPM2, **OXR1**
13	50	7,530,118	8,141,752	OXR1, ABRA
13	329	8,224,958	12,079,619	ANGPT1, RSPO2, EIF3E, EMC2, TMEM74, TRHR, MIR8836, NUDCD1, ENY2, PKHD1L1, EBAG9, SYBU, KCNV1
13	85	12,114,766	13,230,859	CSMD3
14	126	2,719,987	4,251,815	AKR1B1, SLC35B4, LRGUK, EXOC4, CHCHD3
14	123	33,655,202	35,301,479	POLR1F, TMEM196, MACC1, ITGB8, ABCB5, SP8, SP4
14	120	42,664,169	43,836,777	SCRN1, FKBP14, PLEKHA8, ZNRF2, NOD1, GARS1, CRHR2, MINDY4, AQP1, GHRHR, **ADCYAP1R1**
15	98	35,433,529	36,679,720	NDUFA12, NR2C1, FGD6, VEZT, MIR331, METAP2, USP44, NTN4, SNRPF, CCDC38, AMDHD1, HAL, LTA4H, ELK3
17	122	4,292,294	5,749,246	CMPK2, RSAD2, RNF144A
17	88	5,820,182	6,962,923	ID2, KIDINS220, MBOAT2, ASAP2, ADAM17, YWHAQ
19	119	3,672,616	5,185,402	**NOCT**, SLC7A11, PCDH18
22	97	2,012,670	3,318,126	KPNA3, EBPL, RCBTB1, SETDB2, CAB39L, CDADC1, MLNR, FNDC3A, CYSLTR2, RCBTB2, RB1, LPAR6
24	85	3,894,614	5,037,218	RIN2, SLC24A3, DTD1
*25**	*98*	*984,065*	*2,453,286*	*COG6*, *LHFPL6*, *NHLRC3*, *PROSER1*, *STOML3*, *FREM2*
25*	261	2,491,987	5,908,676	UFM1, TRPC4, POSTN, SUPT20H, SMAD9, SERTM1, *CCNA1*, *SPART*, *SOHLH2*, *DCLK1*, *NBEA*, *MAB21L1*
28	131	36,607,901	37,848,840	DOCK1, NPS, CLRN3, PTPRE, MKI67
29	*24*	*11,666,854*	*11,911,242*	*CLVS1*, *ASPH*
30	106	18,173,791	19,598,373	MYO5A, ARPP19, ATOSA, ONECUT1, WDR72

Table 1. ROH Islands details and annotated genes: Genes annotated within each region are highlighted as follows: normal font indicates genes common to both ROH detection methods; **bold font** denotes genes specific to ROH islands identified by the Consecutive method; *italics* refer to genes located in ROH islands shared among all (100%) Irish Wolfhounds (IWs). *ROH Islands shared by all IWs (100% occurrence) ◊ chr 2:14678549–15,499,133; chr 25:984065–2,429,437; chr 25:4243224–5,193,533. *ROH Islands (and annotated genes) identified only with the consecutive method.*

iHS direction	Chr	nSNP	From	To	Genes
Positive	8	89	58,190,000	60,180,000	GALC, GPR65, KCNK10, SPATA7, PTPN21, ZC3H14, EML5
Negative	10	53	58,780,000	60,420,000	VRK2, FANCL
Negative	22	67	47,360,000	48,640,000	HS6ST3, OXGR1, MBNL2, RAP2A, IPO5

Table 2. Candidate regions and annotated genes under selection according to a threshold of 3 obtained with the iHS analysis.

F_type	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
FPED	0.025	0.032	0.000	0.004	0.009	0.040	0.148
FHom	0.028	0.054	−0.070	−0.006	0.024	0.057	0.259
FROH	0.278	0.033	0.116	0.262	0.279	0.296	0.401
FROH_<2 Mb	0.291	0.033	0.127	0.274	0.290	0.309	0.419
FROH_2-4 Mb	0.172	0.031	0.028	0.157	0.173	0.190	0.264
FROH_4-8 Mb	0.058	0.017	0.002	0.047	0.057	0.069	0.100
FROH_8-16 Mb	0.006	0.003	0.004	0.004	0.004	0.008	0.013

Table 3. Descriptive statistics of all estimated inbreeding coefficient.

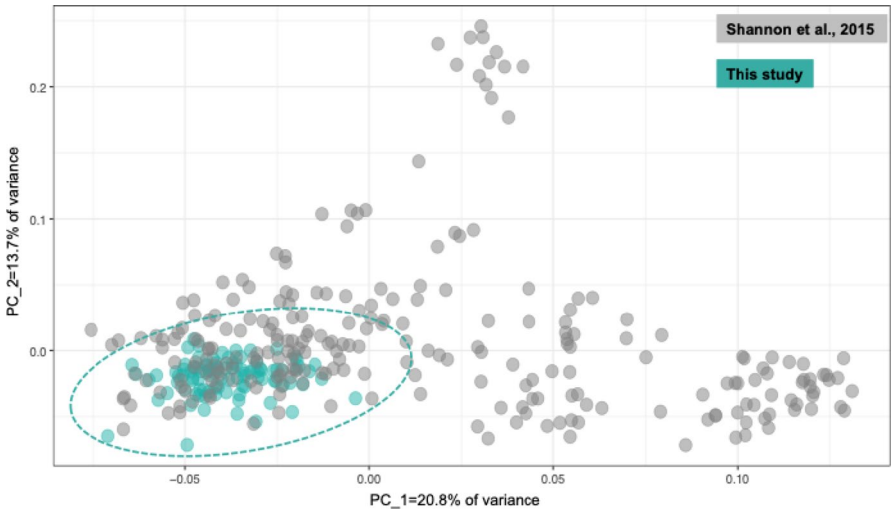


Fig. 3. Principal component analysis (PCA) comparing the IWs genotyped in this study (teal) with those reported by Shannon et al.[14](grey).

Genomic inbreeding coefficients (FROH and FHom) showed higher average values and greater variability than the pedigree-based coefficient (FPED). Among them, FROH exhibited the widest range (0.278 ± 0.033), followed by FHom (0.028 ± 0.054), while FPED displayed the lowest values 0.025 ± 0.032), and the most homogeneous distribution. This pattern is consistent with pedigree depth that constrains its ability to capture distant common ancestry. FROH had the highest mean values, confirming its higher sensitivity in detecting genome-wide autozygosity. In contrast, FPED values were clustered around zero, reflecting the depth of available pedigrees and the inability of pedigree-based metrics to capture ancient or breed-wide inbreeding. Overall correlations were moderate, with the strongest association observed between FHom and FROH ($r = 0.60$). FPED showed weaker correlations with genomic-based coefficients ($r = 0.60$ with FHom and $r = 0.24$ with FROH), again indicating that pedigree-based estimates capture only the most recent component of inbreeding. When ROH length classes were considered separately, FPED correlated weakly with FROH_8–16 Mb, likely because only 40 individuals exhibited ROH segments in this length range. Despite these differences, all inbreeding coefficients showed similar overall trends across individuals, with genomic-based measures providing finer resolution and more accurate quantification of both recent and historical inbreeding.

Identity-by-descent relatedness measure (IBD)

Pairwise genome-wide IBD-relatedness measures were estimated across all dogs using both SVS and PLINK, with mean values per individual of 0.018 (SVS) and 0.019 (PLINK) when averaged across all pairwise comparisons. These low averages indicate that most individuals are only weakly related at the genomic level (Figure S4, showing only SVS-IBD results). At the level of single pairwise comparisons, only one pair of individuals showed a PI_HAT consistent with first-degree relatedness (~ 0.5), while 11 individuals exhibited moderate relatedness (PI_HAT: 0.25–0.30), as expected in a closed breeding population. No evidence of duplicate samples (PI_HAT > 0.9) was detected. Overall, these results suggest that, although the breed's genetic diversity is constrained, the sampled population does not show high levels of recent relatedness at the pairwise level, and therefore does not indicate pervasive recent inbreeding.

Comparison with publicly available Irish Wolfhound dataset

The PCA shown in Fig. 3, which combines the 96 IWs from this study (green) and those reported by Shannon et al. (grey)[14], revealed a consistent genetic structure across the two datasets. In fact, the IWs from this study

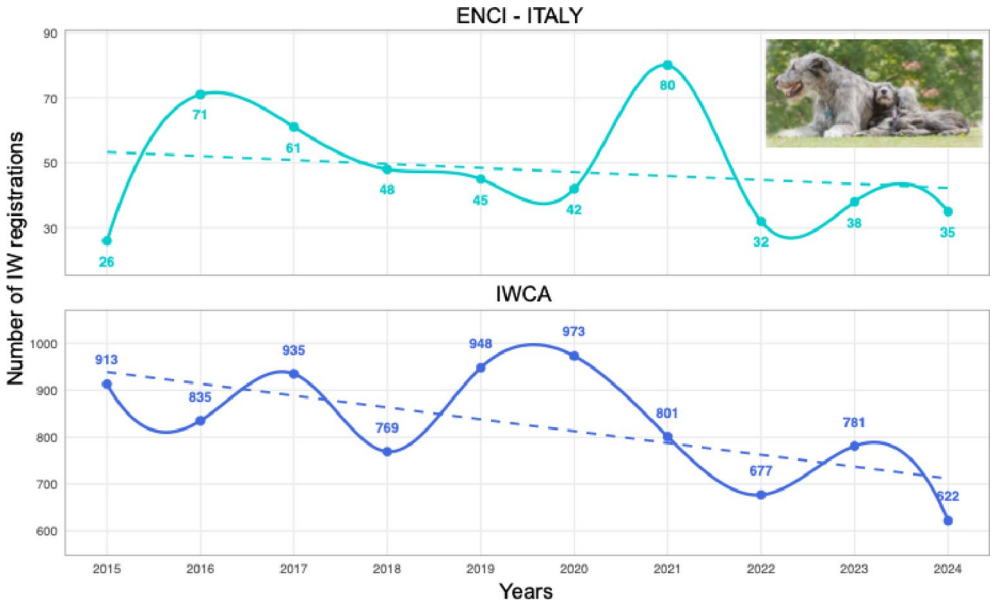


Fig. 4. Annual number of Irish Wolfhound registrations in the IWCA (United States) and ENCI (Italy) databases from 2015 to 2024. IWCA values represent minimum estimates derived from breeder-identified litters and database records, reflecting the absence of mandatory dog registration in the United States (data provided by the president of IWCA and available at the time of the analysis; data for 2024 may be subject to updates), whereas ENCI data correspond to officially registered dogs (https://www.enci.it/libro-genealogico/razze/irish-wolfhound). Trends are shown using smoothed curves to highlight temporal patterns. Photo provided by M. Poli (one of the authors).

clustered tightly within the main variation cloud formed by the public dataset, with no clear separation along either PC_1 (20.8% of variance explained) or PC_2 (13.7%).

This extensive overlap demonstrates that the individuals analysed here are representative of the broader Irish Wolfhound genetic background and exhibit no marked divergence from the previously characterized population. The compact distribution of the new samples also indicates limited population substructure and supports the absence of recent admixture or introgression events within the breed. Overall, the combined PCA confirms that both datasets capture a largely homogeneous genetic pool consistent with the demographic history of IWs. It should be noted that, as Shannon et al. did not report information on the geographic origins of the IW samples, we were not able to formally assess whether the observed clustering reflects country-specific structure. Therefore, no conclusions regarding geographic differentiation within the reference population can be drawn from this analysis.

Mitochondrial DNA Variation

Among the 335 mitochondrial DNA (mtDNA) markers available on the chip, only one SNP, chrM_2758, located in the MT-ND1 gene, was found to be polymorphic. At this locus, the frequency of the A allele in our population was 84.4%, which closely matched the value reported by Shannon et al. in their publicly available dataset (82.2%). The complementary G allele accounted for the remaining proportion. Thus, both datasets display a striking consistency in mitochondrial allele distribution at this locus across independent sampling. The nucleotide change (A→G) is synonymous, resulting in no amino acid substitution (Isoleucine → Isoleucine). Overall, the near-identical genotype frequencies and the presence of only a single polymorphic mitochondrial SNP across all individuals highlight the extremely low mtDNA diversity within the breed.

Discussion

The IW is one of the most historically significant yet numerically vulnerable giant dog breeds. Originally shaped by repeated population contractions and later reconstructed from a limited number of founders, the breed has long been characterized by reduced genetic diversity and a persistent risk of genetic erosion. Despite its worldwide distribution today, the IW remains a numerically rare breed when compared to more popular and widespread dog breeds. Annual registrations recorded by different kennel clubs reflect both the global distribution of the IW and the demographic scale of the countries involved. As expected, the IWCA registry, representing the North American population (data provided by the president of Irish Wolfhound Club of America[15]), reports substantially higher annual registrations compared with ENCI[16], whose figures are constrained by the smaller size of the national breeding population (Fig. 4). Nevertheless, even within the largest registry, annual registration numbers remain modest when compared with those of more widespread dog breeds, rarely exceeding one thousand individuals per year and showing a declining trend in recent years (e.g., German Shepherd at ENCI

website[17]). Both registries consistently indicate low annual registration numbers and a post-2020 decline in breeding activity, highlighting the overall demographic vulnerability of the breed.

Consistent with the demographic patterns observed across registries, the PCA analysis indicates that IWs from 23 different countries cluster closely together, without forming distinct country-based genetic groups. This overall genetic similarity reflects shared ancestry and extensive gene flow across international breeding lines. Importantly, the low pairwise IBD estimates and the predominance of short ROH segments indicate that, despite this global relatedness, recent inbreeding is limited. A few outliers may reflect local breeding histories or rare lineages maintained within specific kennels.

The slightly lower observed heterozygosity (Ho = 0.177) compared to the expected value (He = 0.181) suggests some degree of non-random mating or preferential use of particular breeding lines, but not enough to indicate pervasive recent inbreeding. Consistently, the corresponding levels of observed homozygosity in our population are high (0.85). Similar higher values, exceeding 0.80, have been reported in Bernese Mountain Dogs and in Norwegian Lundehund[18,19]. Overall, these findings are consistent with the PCA results, highlighting a largely homogeneous genetic structure at the international level. Although the sampled dogs were born between 2013 and 2024, the observed patterns likely reflect both historical founder effects and ongoing breeding practices.

The observed decline in effective population size (Ne) reflects the demographic history of the breed, which has undergone strong population contraction. In the present study, the most recent Ne estimates obtained using SNeP correspond to approximately 13 generations ago and indicate a value of 56, reflecting the Ne of the breed in the recent past rather than the current census population. This time frame is consistent with the resolution limits of LD-based methods, which rely on linkage disequilibrium between markers at short genomic distances and therefore cannot capture very recent generational changes. The additional LD-bin analysis was performed as an exploratory approach to assess Ne at more recent generational depths. The higher Ne estimates observed for the most recent time intervals likely reflect a combination of methodological differences and stochastic variation in LD at short distances. In particular, while SNeP relies on a model-based relationship between LD and recombination distance, the LD-bin approach directly averages r^2 within physical distance classes. As a result, LD at very short distances is more sensitive to stochastic variation and sampling effects, which can lead to inflated and more variable Ne estimates in the most recent time intervals. While these estimates should be interpreted with caution, they may also be consistent with reduced levels of recent inbreeding, as suggested by the genomic data.

The IWs breed-level genetic assessments (including UC Davis breed report[12]), indicate a limited genetic diversity and a small pool of influential lineages. Breeding-club reports, such as the IW Breeding Strategy (2019), acknowledge major reductions in diversity over the past 150 years and pervasive relatedness within the breed[20]. Although no study to our knowledge has provided a detailed temporal estimate of Ne in the IWs or its close sighthound relatives, available analyses in other sighthounds (e.g., Greyhound, Italian Greyhound) and broader purebred dog surveys suggest that low Ne and loss of genetic diversity are common in sighthound lineages and closed-breed populations[3]. Instead, similar temporal trends in Ne have been observed in other dog breeds, with extremely low contemporary Ne in the Norwegian Lundehund (about 10–13), reduced Ne in the Dalmatian (69 from genomic data; 116 from pedigrees), and comparable declines in the Finnish Spitz (57) and Nordic Spitz (49) following breed establishment[21–23].

Regarding π, the very low values found along all autosomes is consistent with the high level of homozygosity observed through ROH analyses and reflects the long-term consequences of restricted effective population size. As showed in Figure S1, peaks of comparatively higher π are rare and typically narrow, indicating localized regions where genetic variation has been retained, possibly due to recombination hotspots or the introduction of diversity through historical outcrosses. In contrast, extended regions with π approaching zero correspond to ROH islands shared by most dogs in the breed, highlighting ancient selection or persistent founder effects. Taken together, the π landscape supports the hypothesis of a globally homogeneous and genetically constrained population, further emphasized by the PCA results showing minimal geographic structure. Literature reporting π estimates are limited in dogs, however, Brouillette et al. reported π values on the order of 10^{-3} across both coding and non-coding regions, reflecting the reduced genome-wide variation resulting from domestication, founder effects, and the formation of modern breeds[24].

The slow LD decay observed in IW, with r^2 values remaining elevated over megabase scales, reflects long-term reduced Ne and extensive autozygosity. This genomic signature aligns with the breed's documented demographic history, including severe bottlenecks and a limited founder base, and helps explain the markedly low SNP diversity and increased levels of ROH observed across the genome.

Runs of Homozygosity detection

A particularly relevant finding of this study is the identification of 26 ROH islands shared by 85% of the analysed IWs (corresponding to at least of 81 dogs), indicating the presence of genomic regions that are consistently homozygous across the population. These ROH islands are not excessively extended, ranging in size from approximately 277 kb to 3.7 Mb, with a mean length of about 1.6 Mb. This size distribution suggests that the observed ROH islands are unlikely to be solely the result of very recent inbreeding, but rather reflect older, stable homozygous segments shaped by the breed's demographic history and long-term selection. Five of the identified ROH islands overlap, either partially (chr 13 at 8.2 Mbp, chr 14, chr 22, and chr 25) or completely (chr 10 and chr 13 at 12 Mbp), with ROH regions previously reported in other hunting dog breeds, suggesting the presence of shared genomic segments potentially shaped by common functional selection related to hunting ancestry[25–29].

Notability, three IWs ROH islands are shared by 100% of the individuals, highlighting genomic regions that appear to be fixed within the population. These fully shared regions contain a total of 15 annotated genes, some of which are involved in human diseases that also impact the lifespan and health of IWs, including cancer (*DCLK1*, *MAP3K8*, also known as *TPL2*, and *CCNA1*), cardiac disease (*JCAD*), and neurological disorders,

including epilepsy or ataxia (*MTPAP*, *SPART*, *COG6*, and *NBEA*). Based on current knowledge, ROH islands shared by 100% of individuals have so far only been reported in the Bernese Mountain Dog; however, none of these regions overlap with the ROH islands identified in our IW population, indicating breed-specific patterns of homozygosity[18].

ROH islands shared by at least 85% of dogs, contain genes associated with morphological characteristics (e.g., drop ear: *WIF1*and *MSRB3*)[29] and behavioural traits (11 traits: n. 30 genes, Table S1)[30]. Four of these ROH islands overlap with regions previously reported in other hunting dog breeds: on chr 10, chr 13 (at 7.5 Mpb), and chr 25[26,28,31]. In contrast, the ROH islands on chr14 (at 2.7 Mbp) and chr 22 have been observed in the different Shepherd dog breeds[31,32].

In addition, these regions harbour genes which, based on their functional characterization in humans, may be involved in one of the major causes of mortality in IWs, the osteosarcoma: (i) mutations in the *RB1* gene are strongly associated with osteosarcoma, as *RB1* acts as a tumour suppressor that normally prevents uncontrolled cell proliferation[33–35]; (ii) *POSTN* has been reported to regulate osteosarcoma cell migration and invasion, suggesting a potential role as a biomarker and therapeutic target; (iii) *DOCK1* has been implicated in osteosarcoma progression through regulatory mechanisms affecting cell signalling pathways[36].

iHS selection signatures

The strongest positive iHS region on chromosome 8 (58.19–60.18 Mb) encompasses several genes with roles in neuroinflammation, immune regulation, and cellular homeostasis, suggesting recent selection favouring the derived haplotype. *GALC*, the causal gene for globoid cell leukodystrophy, directly affects lysosomal sphingolipid metabolism and neuroimmune pathways[37]. *GPR65*, a key regulator of Th17 pathogenicity and lysosomal pH, links this region to inflammatory response modulation in minks[38]. Additional genes include *KCNK10*, previously implicated as a microRNA target in metastatic canine mast cell tumours; *SPATA7*, associated with early-onset retinal degeneration[39]; and cancer-relevant loci such as *ZC3H14*, whose phosphorylation is affected by *CDK13* mutations in multiple tumour types[40], and *EML5*, a microtubule-associated protein recurrently mutated in cancer models[41].

Two negative iHS regions, where selection favours the ancestral haplotype, were detected on chromosomes 10 and 22. The region on chromosome 10 (58.78–60.42 Mb) contains *FANCL*, a core component of the Fanconi anaemia DNA repair pathway essential for maintaining chromosome stability[42]. The region on chromosome 22 (47.36–48.64 Mb) includes genes linked to inflammatory and structural pathways, such as *HS6ST3*, which modulates heparan sulfate–dependent signalling favouring myoblast proliferation[43]; *OXGR1*, a GPCR involved in various inflammatory disorders[44]. These patterns suggest that ancestral alleles in these regions may confer advantageous regulation of immune–metabolic balance and cellular stress responses.

Inbreeding coefficients

The comparison between pedigree- and genome-based inbreeding coefficients (Fig. 2, Panel_4, a) revealed moderate concordance among estimators, reflecting their distinct biological and methodological foundations. In line with the descriptive statistics (Table 3), the pedigree-based coefficient (FPED) tended to underestimate inbreeding compared with genomic estimators[5]. The distribution patterns shown in Panel_4 further support this interpretation. FROH and FHom exhibited wider variability and higher mean values, indicating that genomic estimators are more sensitive to genome-wide autozygosity. In particular, FROH quantifies the proportion of the genome contained in homozygous-by-descent segments, providing a direct measure of inbreeding. This contrasts with pedigree-based estimates, which reflect the expected inbreeding probabilities derived from pedigree relationships[6,45,46]. Finally, the correlation analysis (Panel_4, b) confirmed that FHom and FROH were the most strongly associated ($r = 0.60$), while FPED correlated weakly with both genomic coefficients. When FROH was partitioned by ROH length, correlations with FPED further decreased for longer segments (e.g., FROH_8–16 Mb), likely because only a limited number of individuals showed ROHs in these length classes. These long ROHs typically reflect recent common ancestry; therefore, their scarcity suggests limited recent inbreeding within the population[47,48]. In contrast, in other dog breeds such as the Bernese Mountain Dog, where overall homozygosity exceeds 80%, long ROH segments can account for more than 4% of the genome, with corresponding FROH values exceeding 0.40 (mean value), indicating substantial recent inbreeding[18].

Overall, the three coefficients capture different but complementary aspects of inbreeding:1) FPED reflects expected inbreeding based on known ancestry and depends on pedigree completeness and accuracy; ii) FHom provides a genome-wide measure of overall homozygosity; iii) FROH, particularly when partitioned by ROH length, offers temporal resolution of inbreeding history, distinguishing recent from ancient events. Together, these coefficients provide a comprehensive view of inbreeding patterns, integrating genealogical expectations with realized genomic information and enabling a more nuanced interpretation of population structure and management strategies[46,48]. This interpretation is supported by the IBD results, which show generally low pairwise relatedness among individuals and only a few cases of close kinship, confirming that recent inbreeding is limited in the sampled population.

While our interpretations rely on widely used population genetic frameworks, it is important to acknowledge that such approaches are inherently model-dependent. Although recent studies have questioned some general assumptions underlying these approaches[49,50], in domestic dog breeds genomic patterns are predominantly shaped by demographic history and artificial selection. Therefore, our results should be interpreted within this specific context.

Conclusion

The combined results of all analyses provide a consistent and nuanced picture of the genomic evolutionary history of the IW. While reduced effective population size and elevated homozygosity are evident, the scarcity

of long ROH (> 8 Mb) and the absence of very long ROH (> 16 Mb) indicate that recent inbreeding has been limited, consistent with limited close mating events in the population. Instead, the observed patterns are best explained by ancient demographic bottlenecks and long-term population constraints. Importantly, haplotype-based selection scans identified only a small number of regions under recent positive selection, which were largely distinct from ROH islands, suggesting that widespread recent selection is not the primary driver of homozygosity. Taken together, these findings demonstrate that, despite historical genetic constraints, recent breeding practices have been effective in controlling inbreeding over the last generations, as supported by the absence of long ROH and the limited evidence of recent positive selection.

This highlights the role of informed and coordinated breeding strategies in maintaining genomic stability and supports the view that current homozygosity patterns in IW largely reflect historical rather than ongoing processes. From a conservation perspective, future efforts should continue to prioritise the maintenance of genome-wide diversity while avoiding further reductions in effective population size, with the integration of genomic information providing a valuable framework for sustainable breed management and mating practices. Overall, these results underscore the importance of long-term breeding strategies in preserving both the genetic integrity and health of the IW and highlight that genomic information represents a fundamental asset to pursue the maintenance of the IW breed. In addition, our results underline the importance of strong collaborative planning regarding the breed selection strategies within all countries where the IW is bred.

Materials and methods

A total of 96 purebred IWs (41 males and 55 females) were previously sampled from 23 countries worldwide, representing a broad and diverse subset of the global population (Table S3). Sampling was coordinated by the Italian Irish Wolfhound Club under the supervision of Marcello Poli, President of the Sighthound Club of Italy and Chairman of the Federation of Irish Wolfhound Clubs (FIWC). Biological samples were provided by breeders who voluntarily participated in the project. For remotely collected samples, owners received sampling kits (sent by M. Poli) and returned them, together with signed consent forms and individual identification data, to the research team. The Sighthound Club of Italy performed the genotyping and made the genotypes available for this study.

Ethics statement

For this study, approval from the Animal Care and Welfare Committee was not required, as only the genotypes made available by the President of the Sighthound Club of Italy were used.

Genotyping and Quality Control

The genotypes were based on a genotyping performed using the Illumina CanineHD BeadChip (230 K SNPs). Before all genetic analyses, SNP positions were updated from the CanFam3.1 assembly to the ROS_Cfam_1.0 reference genome. The CanFam3.1 assembly was downloaded from the Ensembl database, which provides reference information including the rs identifiers assigned to known variants. Each SNP for which genotypes were available was first matched to its corresponding rs identifier based on its position in the CanFam3.1 assembly. These rs identifiers were then used as input for the Ensembl Variant Effect Predictor (VEP) tool[51] to retrieve the updated genomic coordinates in the ROS_Cfam_1.0 assembly. In total, the positions of 194,458 SNPs were successfully remapped. On this dataset quality control filtering was performed in SNP & Variation Suite (SVS) v8.9 (Golden Helix Inc., Bozeman, MT, USA) retaining SNPs with: (i) individual call rate > 90%; (ii) minor allele frequency (MAF) ≥ 1%; (iii) and unique genomic positions (no duplicated loci). Also, only autosome SNPs were retained, resulting in a total of 188,681 SNPs used for downstream analyses.

Genetic variability and population structure

The genetic structure among individuals was assessed through Principal Component Analysis (PCA), performed with the SVS software. Graphical visualization of PCA results was generated in R using the ggplot2 package[52]. Observed and expected heterozygosity (Ho and He) were estimated using PLINK v1.9 to provide a measure of within-population genetic diversity[53]. Observed homozygosity was directly computed by SVS software as part of the estimation of the inbreeding coefficient FHOM (see below). The effective population size (Ne) of the Irish Wolfhound population was primarily inferred using the software SNeP[54], which estimates historical Ne trajectories based on linkage disequilibrium (LD) between SNP markers. This method allows reconstruction of Ne across past generations, with the temporal resolution determined by the genomic distance between markers. To further explore Ne at more recent generational timescales, beyond the lower resolution limit of SNeP, an additional LD-based analysis was performed following the approach described in previous genomic studies on dogs and other domestic animals[55–57]. In detail, pairwise LD (r^2) values were calculated using PLINK from the 188,681 autosomal SNPs. SNP pairs located on the same chromosome were grouped into physical distance bins of 0.1 Mb, considering distances between 1 kb and 33.3 Mb. For each bin, the mean r^2 value was computed, applying a correction for sample size[55]. Finally, nucleotide diversity (π) was estimated across the autosomal genome using VCFtools[58] with a sliding-window approach (50 kb non-overlapping windows). π estimates were calculated for each chromosome and visualized in R. Linkage disequilibrium (LD) decay was assessed using PopLDdecay v.3.42[59]. SNP data were filtered in PLINK v.1.9 to retain autosomal variants with MAF ≥ 0.01 and missingness ≤ 5%. LD statistics (pairwise r^2) were computed up to 2 Mb between SNPs (MaxDist = 2000). LD decay curves were generated using Plot_OnePop.pl with 10 kb and 100 kb bin sizes for short- and long-range LD, respectively[60]. Mean r^2 values were plotted against inter-marker distance to characterize the rate of LD decline across the genome.

Runs of Homozygosity detection

Runs of Homozygosity (ROH) were identified for all 96 dogs using both sliding-window and consecutive-based methods implemented in the DetectRUNS library of R software[61]. For the sliding-window-based method, a window size of 60 SNPs and a homozygosity threshold of 0.05 were applied. ROH were called when the window contained at least 60 SNPs, no opposite or missing genotypes, a maximum gap of 1 Mb between consecutive SNPs, a minimum length of 1,000 kb, and a minimum SNP density of 1 per 50 kb. The same conservative criteria were applied to limit false-positive short calls for the consecutive approach, which required a minimum of 60 SNPs per ROH, a minimum length of 1,000 kb, a maximum gap of 1 Mb between consecutive SNPs, and no heterozygous or missing SNPs permitted within an ROH.

ROH were classified into five length categories (< 2 Mb, 2–4 Mb, 4–8 Mb, 8–16 Mb, and > 16 Mb), and descriptive statistics were calculated per individual and across the population. ROH islands were defined as genomic regions where SNPs were located within ROH in more than 85% of individuals. This threshold was selected to focus on highly shared autozygous regions, given the high prevalence of short ROH segments across the genome in this population[62]. Genes located within ROH islands, identified through both methods, were identified using the Genome Data Viewer (NCBI) (https://www.ncbi.nlm.nih.gov/gdv). In addition, to explore potential gene functions the DAVID online annotation tool was used[63].

iHS analysis

Genotype data in PED/MAP format were converted into VCF after standard QC (removal of non-biallelic SNPs, loci with > 5% missingness, and variants with MAF < 0.01) that left a total of 86,111 variants. Autosomal SNPs were then phased chromosome-by-chromosome using BEAGLE v5.4, and the phased VCF was used for downstream analyses.

Haplotype homozygosity–based statistics were computed using the rehh R package (v.3.2.2)[64]. Phased SNP data were imported directly from the VCF file using data2haplohh() with a minimum per-marker call rate of 95% and allele polarization based on major/minor allele frequencies. Extended haplotype homozygosity (EHH) decay and integrated haplotype homozygosity (iHH) were estimated with scan_hh() using the parameters from Jemaa et al.[65] (limehh = limehhs = 0.05; limhaplo = limhomohaplo = 2). Preliminary exploration of genome-wide SNP spacing revealed substantial variability in inter-marker distance, including high-frequency gaps above 300 kb and maximums exceeding 3 Mb across several autosomes. To ensure robust EHH integration in the presence of these discontinuities, we adopted a scale gap of 200 kb and allowed the algorithm to integrate across gaps up to 600 kb. These values represent a conservative compromise between minimizing artefactual integration and avoiding systematic loss of iHH estimates in regions characterized by sparse SNP density. Standardized integrated haplotype scores (iHS) were obtained using ihh2ihs() with a minimum MAF of 5% and allele-frequency bins of width 0.05.

Genome-wide candidate regions under putative recent positive selection were not defined based solely on individual significant SNPs, but rather using a clustering approach implemented in the calc_candidate_regions function. Specifically, regions were required to contain at least five extreme SNPs (min_n_extr_mrk = 5) within a 1 Mb sliding window. This approach is commonly applied in iHS analyses to reduce false positives due to isolated outlier SNPs and to identify biologically meaningful selection signals supported by multiple linked markers[66]. Results are reported using a threshold ($-\log10\ p \geq 3$) in line with recent studies. Genes located within the identified candidate regions were annotated using the Genome Data Viewer (NCBI). Functional enrichment analysis was then performed using the same approach described for ROH-associated genes.

Inbreeding coefficients and IBD estimation

Pedigree-based inbreeding coefficients (FPED)

Genealogical information for the 96 Irish Wolfhounds was obtained from the Irish Wolfhound Database (https://iwdb.org), a privately maintained international pedigree resource. Data extraction was performed by the database curator in collaboration with Marcello Poli, who provided dedicated pedigree files for each sampled individual, including all known ancestors up to the fifth generation.

To manage and analyze the genealogical data, an in-house R script was developed to merge all individual pedigree files into a single integrated dataset, enabling population-level analyses of pedigree structure, ancestor overlap, and genealogical completeness for all the 96 dogs included in the study.

The pedigree-based inbreeding coefficients (F) were estimated using the pedInbreeding() function implemented in the optiSel R package[67].

Genomic Inbreeding coefficient (FHom and FROH)

Two genomic inbreeding coefficients were estimated:

- F_{HOM} calculated as:

$$FHom = \frac{Hom_{Ob} - Hom_{Ex}}{n.\,of\ non\ missing\ SNPs - \ HomEx}$$

where Hom_{Ob} and Hom_{Ex} are the observed and expected homozygous genotypes, respectively[68].

- F_{ROH}, representing the proportion of the genome covered by ROH, computed as:

$$FROH = \frac{L_{ROH}}{L_{aut}}$$

where L_{ROH} is the total ROH length per individual and L_{aut} is the total length of the autosomal genome covered by SNPs (~ 2.23 Gb; 2,228,735,030 bp)[61].

Correlations among inbreeding coefficients were assessed using Pearson's correlation coefficient, computed in R.

Identity-by-descent relatedness measure (IBD)

Pairwise IBD was estimated between all individuals to assess genetic relatedness as part of quality control. Two complementary approaches were applied: PLINK and SVS. PLINK was used to compute standard pairwise IBD metrics (PI_HAT), while SVS provided an alternative haplotype-based estimate. These analyses were primarily performed to identify potential duplicate or closely related individuals in the dataset.

Comparison with publicly available Irish Wolfhound dataset

To contextualize the genomic variability of the studied dogs, the dataset was integrated with 247 publicly available Irish Wolfhounds from Shannon et al.[14] in the SVS software exclusively for principal component analysis (PCA). This reference dataset (Table S4) was used to assess the genetic consistency of the current population with previously characterized individuals, whereas all subsequent analyses were restricted to the present cohort to provide a contemporary snapshot of the breed's genetic status. The combined dataset (95,423 SNPs resulting after applying a filter including MAF ≥ 0.01 and SNP call rate ≥ 0.9) enabled a comparative PCA and assessment of population structure between the newly genotyped individuals and the broader reference population, allowing exploration of potential clustering patterns associated with geographic origin or breeding history.

Mitochondrial DNA variability

The Illumina CanineHD BeadChip includes 335 mitochondrial SNPs, allowing partial assessment of mitochondrial diversity. Although this approach does not capture the full mitochondrial variability, particularly the D-loop region, which is the most variable portion of the mtDNA, the retained markers, available also for the Shannon et al.[14], may provide insight into conserved mitochondrial haplotypes within the breed. Functional annotation of variable sites was performed using Ensembl's Variant Effect Predictor (VEP) to determine gene context and possible consequences.

Data availability

Genotype data for the 96 IWs are provided in the supplementary file "Supplementary_IWF_genotypes.zip" in PLINK format.

Received: 11 January 2026; Accepted: 28 May 2026
Published online: 06 June 2026

References

1. Hogan, E. I. *The History of the Irish Wolfdog* (Privately reprinted for JA McAleenan, 1897).
2. Denise Flaim. Irish Wolfhound History: Behind the Ancient Breed of Ireland. (2023). https://www.akc.org/expert-advice/dog-breeds/irish-wolfhound-history-behind-the-breed/
3. Li, W.-L. et al. Multiple origins and genomic basis of complex traits in sighthounds. *Mol. Biol. Evol.* **40**, msad158 (2023).
4. Urfer, S. How the past affects the present: A genetic history of the Irish wolfhound. *Dogs: Biology, Behavior and Health Disorders; DeGiovine, VM, Ed* 69–91 (2011).
5. Leroy, G. Genetic diversity, inbreeding and breeding practices in dogs: results from pedigree analyses. *Vet. J.* **189**, 177–182 (2011).
6. Purfield, D. C., Berry, D. P., McParland, S. & Bradley, D. G. Runs of homozygosity and population history in cattle. *BMC Genet.* **13**, 70 (2012).
7. Curik, I., Ferenčaković, M. & Sölkner, J. Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livest. Sci.* **166**, 26–34 (2014).
8. Howrigan, D. P., Simonson, M. A. & Keller, M. C. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics* **12**, 460 (2011).
9. Boichard, D., Ducrocq, V., Croiseau, P. & Fritz, S. Genomic selection in domestic animals: Principles, applications and perspectives. *C. R. Biol.* **339**, 274–277 (2016).
10. Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72 (2006).
11. Gouveia, J., Silva, M., Paiva, S. R. & Oliveira, S. Identification of selection signatures in livestock species. *Genet. Mol. Biol.* **37**, 330–342 (2014).
12. UC Davis IW breed report. https://vgl.ucdavis.edu/sites/g/files/dgvnsk15116/files/media/documents/IrishWolfhound20190819c.pdf
13. Biscarini, F., Cesarani, A., Criscione, A. & Mastrangelo, S. Invited review: The application of runs of homozygosity and heterozygosity-rich regions to infer population history in livestock species. *Livest. Sci.* **307**, 105950. https://doi.org/10.1016/j.livsci.2026.105950 (2026).
14. Shannon, L. M. et al. Genetic structure in village dogs reveals a Central Asian domestication origin. *Proc. Natl. Acad. Sci. U. S. A.* https://doi.org/10.1073/pnas.1516215112 (2015).
15. The Irish. Wolfhound Club of America. https://iwclubofamerica.org/
16. ENCI, Irish Wolfhound. https://www.enci.it/libro-genealogico/razze/irish-wolfhound
17. ENCI, German Shepard. https://www.enci.it/libro-genealogico/razze/pastore-tedesco
18. Letko, A. et al. Genomic diversity and runs of homozygosity in Bernese Mountain Dogs. *Genes* **14**, 650 (2023).
19. Stronen, A. V. et al. Genetic rescue of an endangered domestic animal through outcrossing with closely related breeds: A case study of the Norwegian Lundehund. *PLoS One* **12**, e0177429 (2017).
20. Norsk Kennel Klub report. https://www.iukn.no/wp-content/uploads/2016/05/Breeding-strategy-IW-Norway.pdf
21. Vasiliadis, D., Metzger, J. & Distl, O. Demographic assessment of the Dalmatian dog–Effective population size, linkage disequilibrium and inbreeding coefficients. *Canine Med. Genet.* **7**, 3 (2020).
22. Kumpulainen, M. et al. Founder representation and effective population size in old versus young breeds—Genetic diversity of Finnish and Nordic Spitz. *J. Anim. Breed. Genet.* **134**, 422–433 (2017).

23. Pfahler, S. & Distl, O. Effective population size, extended linkage disequilibrium and signatures of selection in the rare dog breed Lundehund. *PLoS One* **10**, e0122680 (2015).
24. Brouillette, J. A., Andrew, J. R. & Venta, P. J. Estimate of nucleotide diversity in dogs with a pool-and-sequence method. *Mamm. Genome* **11**, 1079–1086 (2000).
25. Perfilyeva, A. et al. Homozygosity mapping in the Kazakh national dog breed Tazy. *Sci. Rep.* **13**, 10735 (2023).
26. Djurkin Kušec, I. et al. Genomic Characterization of the Istrian Shorthaired Hound. *Animals* **10**, 2013 (2020).
27. Yu, Y. et al. Unique genetic signature and selection footprints in Dutch population of German Longhaired Pointer dogs. *Anim. Genet.* **53**, 829–840 (2022).
28. Boccardo, A. et al. The German shorthair pointer dog breed (Canis lupus familiaris): Genomic inbreeding and variability. *Animals* 10, (2020).
29. Vaysse, A. et al. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genet.* **7**, e1002316 (2011).
30. MacLean, E. L., Snyder-Mackler, N., VonHoldt, B. M. & Serpell, J. A. Highly heritable and functionally relevant breed differences in dog behaviour. *Proceedings of the Royal Society B* **286**, 20190716 (2019).
31. Bionda, A., Cortellari, M., Liotta, L. & Crepaldi, P. The shepherd and the hunter: a genomic comparison of Italian dog breeds. *Animals* **13**, 2438 (2023).
32. Friedrich, J. et al. Unravelling selection signatures in a single dog breed suggests recent selection for morphological and behavioral traits. *Adv. Genet.* **1**, e10024 (2020).
33. Hansen, M. F. Genetic and molecular aspects of osteosarcoma. *J. Musculoskelet. Neuronal Interact.* **2**, 554–560 (2002).
34. Imbert-Bouteille, M. et al. Osteosarcoma without prior retinoblastoma related to RB1 low-penetrance germline pathogenic variants: A novel type of RB1-related hereditary predisposition syndrome? *Mol. Genet. Genomic Med.* **7**, e913 (2019).
35. Xu, C. et al. Periostin promotes the proliferation and metastasis of osteosarcoma by increasing cell survival and activates the PI3K/Akt pathway. *Cancer Cell. Int.* **22**, 34 (2022).
36. Xu, G., Zhang, H., Shi, Y. & Yang, F. Circular RNA circDOCK1 contributes to osteosarcoma progression by acting as a ceRNA for miR-936 to regulate LEF1. *J. Bone Oncol.* **36**, 100453 (2022).
37. Hammack, S. et al. Novel genetic variant associated with globoid cell leukodystrophy in a family of mixed breed dogs. *J. Vet. Intern. Med.* **37**, 1710–1715 (2023).
38. Karimi, K., Farid, A. H., Myles, S. & Miar, Y. Detection of selection signatures for response to Aleutian mink disease virus infection in American mink. *Sci. Rep.* **11**, 2944 (2021).
39. Murgiano, L. et al. Two genes, one culprit-a functional candidate validation of a SPATA7 deletion in dogs with day blindness/retinal degeneration. *PLoS Genet.* **21**, e1011961 (2025).
40. Rambout, X. & Maquat, L. E. Nuclear mRNA decay: regulatory networks that control gene expression. *Nat. Rev. Genet.* **25**, 679–697 (2024).
41. Davis, L. E. et al. Integration of genomic, transcriptomic and functional profiles of aggressive osteosarcomas across multiple species. *Oncotarget* **8**, 76241 (2017).
42. Ferris, E., Abegglen, L. M., Schiffman, J. D. & Gregg, C. Accelerated evolution in distinctive species reveals candidate elements for clinically relevant traits, including mutation and cancer resistance. *Cell Rep.* **22**, 2742–2755 (2018).
43. Ghadiali, R. S., Guimond, S. E., Turnbull, J. E. & Pisconti, A. Dynamic changes in heparan sulfate during muscle differentiation and ageing regulate myoblast cell fate and FGF2 signalling. *Matrix Biol.* **59**, 54–68 (2017).
44. Hervás-Rivero, C. et al. Runs of Homozygosity Islands in Autochthonous Spanish Cattle Breeds. *Genes (Basel)*. **15**, 1477 (2024).
45. Keller, M. C., Visscher, P. M. & Goddard, M. E. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics* (1), https://doi.org/10.1534/genetics.111.130922 (2011).
46. Kardos, M., Luikart, G. & Allendorf, F. W. Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity (Edinb)*. **115**, 63 (2015).
47. Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M. & Wilson, J. F. Runs of homozygosity: windows into population history and trait architecture. *Nat. Rev. Genet.* **19**, 220 (2018).
48. Peripolli, E. et al. Runs of homozygosity: current knowledge and applications in livestock. *Anim. Genet.* **48**, 255–271 (2017).
49. Lynch, M., Wei, W., Ye, Z. & Pfrender, M. The genome-wide signature of short-term temporal selection. *Proceedings of the National Academy of Sciences* **121**, e2307107121 (2024).
50. Song, S., Chen, P., Shen, X. & Zhang, J. Adaptive tracking with antagonistic pleiotropy results in seemingly neutral molecular evolution. *Nat. Ecol. Evol.* **9**, 2358–2373 (2025).
51. The Ensembl Variant Effect Predictor (VEP). https://www.ensembl.org/info/docs/tools/vep/index.html
52. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*New York,. (2016).
53. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
54. Barbato, M., Orozco-terWengel, P., Tapio, M. & Bruford, M. W. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front. Genet.* **6**, 109 (2015).
55. Michels, P. W. & Distl, O. Genetic variability in polish lowland sheepdogs assessed by pedigree and genomic data. *Animals* **10**, 1520 (2020).
56. Salehi, R., Javanmard, A., Mokhber, M. & Alijani, S. Estimating Linkage Disequilibrium and Effective Population Size Across Generations in Holstein Cattle. *Vet. Med. Sci.* **11**, e70684 (2025).
57. Kijas, J. W. et al. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol.* **10**, e1001258 (2012).
58. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
59. Zhang, C., Dong, S. S., Xu, J. Y., He, W. M. & Yang, T. L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **35**, 1786–1788 (2019).
60. PopLDdecay script. https://github.com/BGI-shenzhen/PopLDdecay/blob/master/bin/Plot_OnePop.pl
61. Biscarini, F., Cozzi, P., Gaspa, G. & Marras, G. *DetectRUNS: An R Package to Detect Runs of Homozygosity and Heterozygosity in Diploid Genomes Genomes. R Package Version 0.9.5. Available at* (2019). https://cran.r-project.org/package=detectRUNS.
62. Ablondi, M., Viklund, Å., Lindgren, G., Eriksson, S. & Mikko, S. Signatures of selection in the genome of Swedish warmblood horses selected for sport performance. *BMC Genomics* **20**, 717 (2019).
63. Sherman, B. T. et al. DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* **50**, W216–W221 (2022).
64. Gautier, M., Klassmann, A. & Vitalis, R. rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Mol. Ecol. Resour.* **17**, 78–90 (2017).
65. Ben Jemaa, S. et al. A genome-wide comparison between selected and unselected Valle del Belice sheep reveals differences in population structure and footprints of recent selection. *J. Anim. Breed. Genet.* **140**, 558–567 (2023).
66. Criscione, A. et al. Detecting the footprint of selection on the genome of Girgentana goat, a popular ancient breed. *animal* **19**, 101466 (2025).
67. Wellmann, R. Optimum contribution selection for animal breeding and conservation: The R package optiSel. *BMC Bioinformatics* **20**, 25 (2019).
68. Wright, S. *Genetics of Populations*. *Encyclopaedia Britannica* vol. 10Benton Chicago, (1948).

Acknowledgements
The authors gratefully acknowledge the owners of the dogs for providing the biological samples on which this study is based. They also wish to thank Jocelyne Gagnè, Sarah Backer, Anthony Dunne, Outi Piisi, and Per Hamrén Lundström for coordinating sample collection. The authors also acknowledge the Federation of Irish Wolfhound Clubs (FIWC) for supporting this study, as well as the Irish Wolfhound Club of America (IWCA) and the Irish Wolfhound Database (IWDB) for sharing birth registration records and pedigree data.

Author contributions
M.G.S, S.P.M and F.B Writing—original draft, M.G.S. Methodology, M.G.S, F.B. Formal analysis, M.G.S Conceptualization. M.P. Funding acquisition. All authors discussed and approved the manuscript.

Funding
This study was supported by a voluntary donation from Marcello Poli to the Department of Veterinary Medicine and Animal Sciences (DIVAS), University of Milan, within the research project "Use of genomic information for the reproductive management of Irish Wolfhound dogs", led by Maria Giuseppina Strillacci as principal investigator.

Declarations

Competing interests
Marcello Poli is one of the authors of this manuscript and provided financial support for genotyping as well as the genotype data. He had no role in study design, data analysis, interpretation of the results, or decision to publish the manuscript. The other authors declare no conflicts of interest.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-56093-5.

Correspondence and requests for materials should be addressed to M.G.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.