Negative selection on a *SOD1* mutation limits canine degenerative myelopathy while avoiding inbreeding

Hisashi Ukawa[a], Noriyoshi Akiyama[a,b], Fumiko Yamamoto[a], Ken Ohashi[a], Genki Ishihara[b], Yuki Matsumoto[*a,b,c]

[a]Genetic Testing Section, Anicom Pafe Inc., Kanagawa, Japan

[b]Research and Development Section, Anicom Specialty Medical Institute Inc., Kanagawa, Japan

[c]Data Science Center, Azabu University, Kanagawa, Japan

*Corresponding author: Yuki Matsumoto

Email: y-matsumoto@azabu-u.ac.jp

Running title: *SOD1* testing curbs canine degenerative myelopathy

Abstract

Several hundred mutations are known to cause disease in domestic dogs. Breeding management is therefore required to minimize their spread. Recently, genetic methods such as direct-to-consumer testing have gained popularity, but their effects on dog populations are unclear. Here, we aimed to evaluate the influence of genetic testing on the frequency of mutations causing canine degenerative myelopathy and changes to the genetic structure in the Pembroke Welsh corgi population from Japan. Genetic testing of 5,512 dogs for the causative mutation in *SOD1* (c.118G>A (p.E40K)) uncovered a recent decrease in frequency from 14.5% (95/657) in 2019 to 2.9% in 2022 (24/820). Genome-wide single nucleotide polymorphism (SNP) analyses of 117 selected dogs revealed no obvious changes in genetic diversity and inbreeding levels between the 2019 and 2022 populations. Further Fst-based analysis to detect selective pressure identified the signature of 143 SNPs, including an SNP located in the intron of *SOD1* adjacent to the c.118G>A mutation. Our study highlights that genetic testing can be supported with improved mating choices in breeding programs to avoid inbreeding. This combined strategy could decrease the genetic risk of canine degenerative myelopathy, a fatal dog disease, and its results would be evident within a few years.

Introduction

Genetic testing for disease-causing mutations in companion animals is increasingly performed among veterinarians for diagnosis, among breeders to reduce the incidence of inherited disease, and even among consumers to determine the genetic background of their pets (Moses et al. 2018). The latter group of genetic tests is termed direct-to-consumer (DTC) and can be broadly classified into two categories: 1) detection of specific mutations using sequencing or probes and 2) high-throughput genotyping using genome-wide marker sets designed to detect multiple mutations simultaneously. Beyond their use for consumers, however, data from these DTC tests have substantial implications for genetic research. A recent study using DTC test samples revealed breed-specific genetic mutations associated with hypertrophic cardiomyopathy in several cat breeds (Akiyama et al. 2023). Another study used over 10,000 DTC samples to determine the allele frequencies of 12 genes associated with canine coat color and the physical characteristics of different dog lineages (Dreger et al. 2019). The results indicated that random mating between certain dog breeds can produce unexpected phenotypes, including embryonic lethality (Dreger et al. 2019). Moreover, although at least 775 disease-associated mutations in dogs are already known (Rokhsar et al. 2021), a recent genome-wide association study using DTC genetic testing data revealed a novel and unexpected in-frame deletion that causes deafness in Rhodesian ridgebacks (Kawakami et al. 2022). Thus, the widespread adoption of animal genetic testing is beneficial for genetic studies.

In 2017, genetic testing for adult dogs was introduced in Japan, and genetic testing for puppies began in 2019. The results of such widespread testing could affect dog populations by preventing carriers from breeding, thus limiting the number of genetically affected animals. However, the actual effect of widespread testing on mutation frequency remains unclear. In

addition, while breeding to avoid mutant alleles could affect the genetic structure and inbreeding levels in dog populations, only a few studies have investigated this possibility.

Since the initial report of the draft genome of domestic dogs in 2005 (6), significant advancements have been made in canine genetic research. Inbreeding and genetic structures have been evaluated at the population level based on microsatellites and genome-wide single nucleotide polymorphisms (SNPs) (Boyko et al. 2009; Mellanby et al. 2013; Dreger et al. 2016; Chu et al. 2019). Accordingly, we can apply genome-wide SNP analyses to evaluate the influence of expanded DTC genetic testing on genetic structure in dog populations.

Canine degenerative myelopathy (DM) is a fatal neurodegenerative disease prevalent in several dog breeds, including the Pembroke Welsh corgi (PWC), German shepherd, and boxer (Neeves and Granger 2015). The c.118G>A mutation in the superoxide dismutase 1 (*SOD1*) gene (p.E40K on chromosome 31; 26,540,342 bp, based on CanFam 3.1) is reportedly a causative factor of DM in PWCs (Awano et al. 2009). SOD1 is one of the two antioxidant isozymes responsible for specifically eliminating free superoxide radicals in mammals. The homozygous A allele mutation is strongly associated with DM onset (Awano et al. 2009; Chang et al. 2013; Zeng et al. 2014), indicating that it is an autosomal recessive variant. The mutant homozygous A allele is fairly common in PWCs and is not geographically restricted (e.g., 48.4% or 59/122 in Japan (Chang et al. 2013) and 83% or 14/21 in Mexico (Ayala-Valdovinos and Gomez-Fernandez 2018)).

In this study, we evaluated the allelic frequency of *SOD1* c.118G>A (p.E40K) in over 5,500 PWCs from Japan, analyzing DTC genetic testing data across 3 years (2019–2022). We also assessed inbreeding and population structures based on genome-wide SNPs to determine the effects of genetic testing on dog breeding. Finally, we performed a genome-wide scan to detect

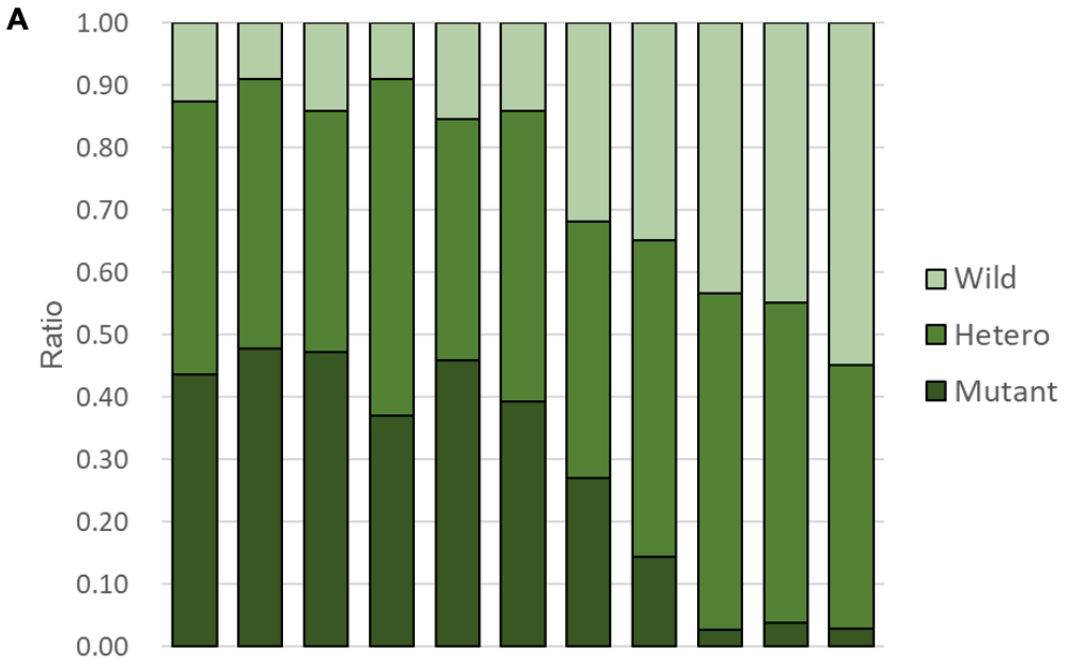
selection signatures across multiple years. The findings of this study could provide valuable insights into how widespread genetic testing controls the spread of genetic disorders among dogs.

Results

Allele frequencies of the SOD1 c.118G>A variant

We examined the distribution of *SOD1*:c.118G<A by birth year in 5,512 PWCs (Fig. 1A). Most PWCs born during 2012–2017 were homozygous for the mutant allele (Mutant) or heterozygous (Hetero), while homozygous wild type (Wild) was the lowest in frequency (9.0% or 12/134 in 2013 to 15.3% or 44/287 in 2016). With genetic testing for adult dogs having started in

2017, the frequency of mutations decreased from 39.4% (163/414) in 2017 to 2.9% (24/820) in 2022.



Birth Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Mutant	56	64	71	74	132	163	90	95	27	43	24
Hetero	56	58	58	108	111	193	137	333	538	588	346
Wild	16	12	21	18	44	58	106	229	431	514	450
Total	128	134	150	200	287	414	333	657	996	1145	820

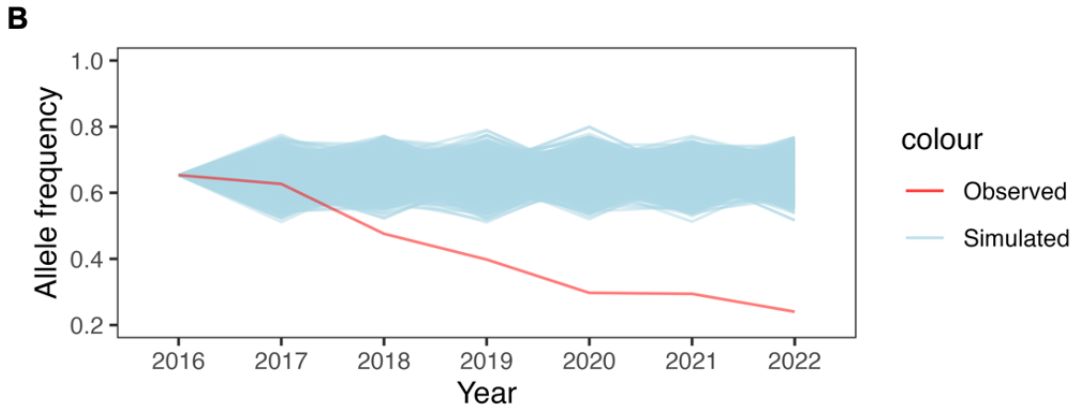


Figure 1. Trends in diploid genotype and allele frequencies of *SOD1* c.118G>A mutation.

A: Diploid genotype frequencies of the *SOD1* c.118G>A mutation from 2012 to 2022. The ratio of allele frequency is based on the birth year of the tested dogs. The lower chart shows the number of tested dogs divided by each diploid genotype. B: Real and simulated allele frequencies of *SOD1* c.118G>A mutation for 6 years before genetic testing in 2016. The red line indicates the observed allele frequency and light-blue lines indicate 10,000 simulated allele frequencies starting from 2016.

Since 2019, when genetic testing began to include puppies, the ratio of each genotype has changed significantly, considering the ratios in 2022 (Fisher's exact test, $P < 0.01$). In other words, the proportion of the mutants decreased from 14.5% (95/657) in 2019 to 2.9% (24/820) in 2022.

To further investigate the allele frequency of *SOD1*:c.118G<A, we performed a simulation based on random genetic drift (10,000 replicates). Deviation from genetic drift would indicate selection. The observed allele frequency was significantly lower than the simulated allele frequency after 2018 ($P < 0.01$; Fig. 1B, Table S1).

Inbreeding levels and genetic structure

We performed genome-wide SNP-based analyses on 117 PWC pups after determining their genetic backgrounds and separating them into four groups according to *SOD1*:c.118G<A genotype (Table S2). Observed heterozygosity (Ho) per group was calculated using 142,510 SNPs; the Ho was 0.305 for the Wild 2019 group and 0.306 for the Wild 2022 group. For mutant PWCs, the Ho was 0.296 and 0.314 in 2019 and 2022, respectively. We then compared the inbreeding coefficients between groups based on runs of homozygosity (F_{ROH}; Fig. 2). The mean F_{ROH} was 0.29 ± 0.067

(SD) for Wild 2019, 0.30 ± 0.082 for Mutant 2019, 0.28 ± 0.061 for Wild 2022, and 0.27 ± 0.038 for Mutant 2022. The four groups did not differ significantly (Table S3).

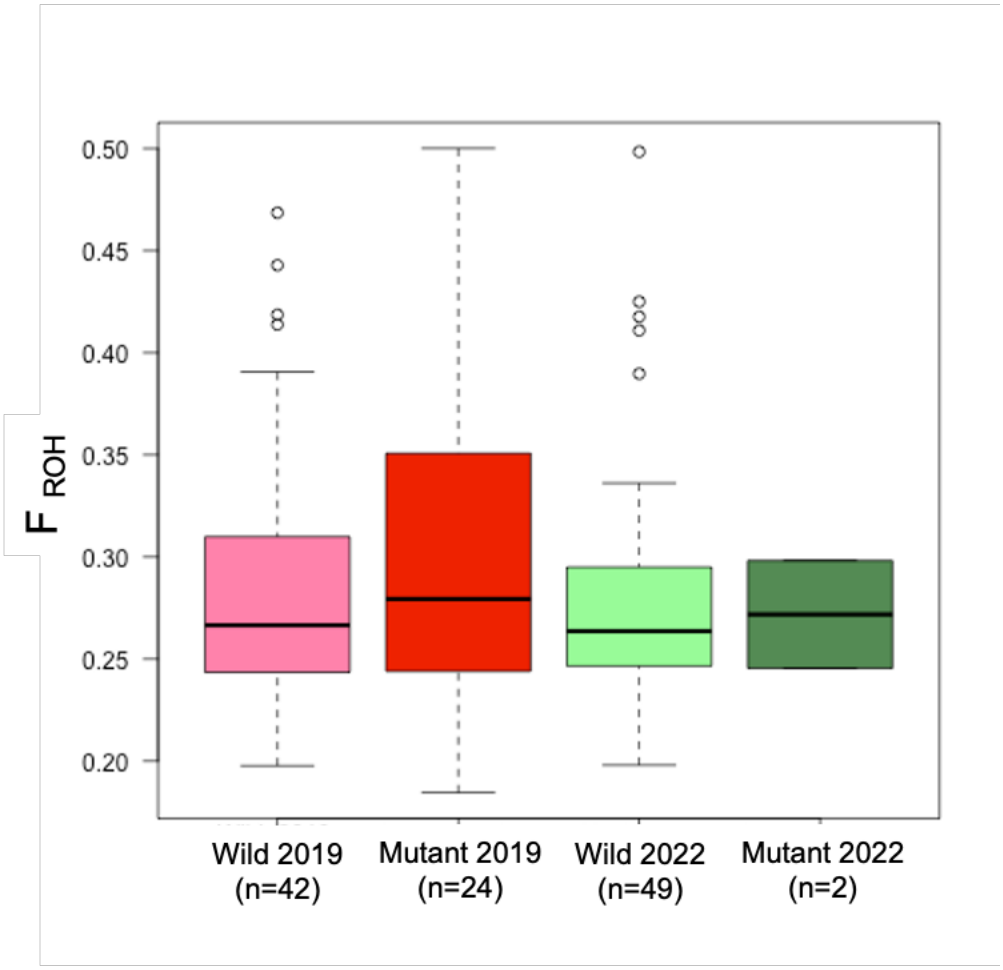


Figure 2. Inbreeding estimates based on runs of homozygosity (F_{ROH}) and genome-wide SNP data.

No difference was observed between populations (Welch's t-test, $P > 0.05$).

We analyzed the genetic structure through clustering using ADMIXTURE (Alexander et al. 2009) and principal component analysis (PCA). Clustering results showed that the cross-validation error was the lowest when K = 3 (Figure S1), with no obvious structure in the model with K = 3 (Fig.

3A). Likewise, PCA did not identify any obvious components that explained variations in the dogs (Fig. 3B).

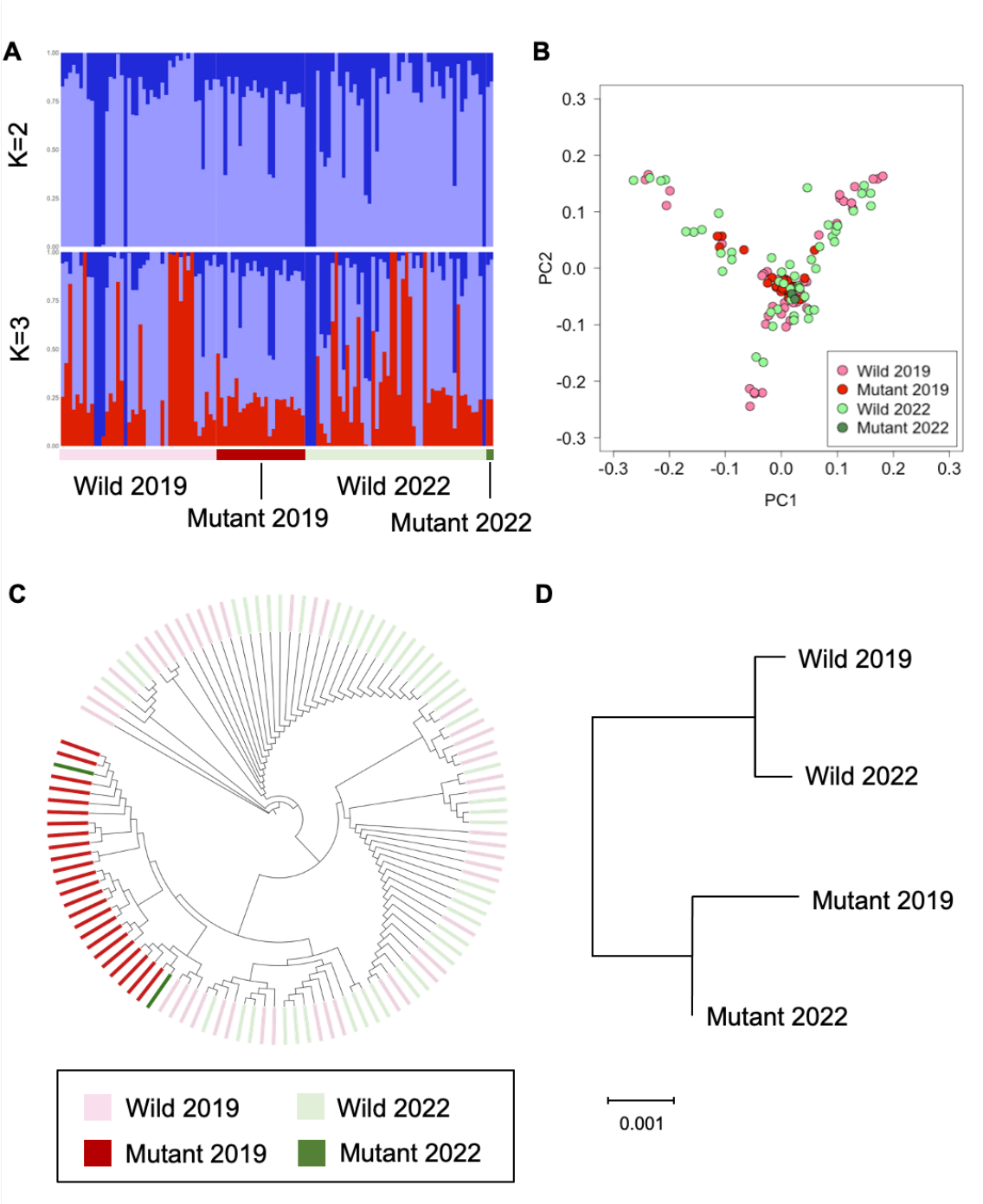


Figure 3. Genetic structure and relatedness of Pembroke Welsh corgis (PWCs).

A: ADMIXTURE. B: Principal component analysis. C: Neighbor-joining phylogenetic tree. D: Dendrogram based on Nei's genetic distance.

Next, we applied the neighbor-joining method for each PWC and Nei's genetic distance to infer genetic relationships between groups. The phylogenetic tree revealed one clade that included all Mutant dogs and another clade including all Wild dogs (Fig. 3C). The population-based Nei's genetic distance indicated genetic similarities within the Mutant groups as well as within the Wild groups (Fig. 3D). Both analyses revealed that PWCs homozygous for the *SOD1* mutation were more common in certain lineages.

To estimate the potential number of dogs bred, we used the linkage disequilibrium method to determine the contemporary effective population size (*N*e) (Do et al. 2014). The contemporary *N*e of Wild 2019 (n = 42) was 48.9, lower than the *N*e of Wild 2022 (73.1, n = 49).

Selection signature

We observed biased mating in all of our genetic analyses. We, therefore, compared the 2019 and 2022 groups to determine the selective signature in the PWC genome. Using the top 0.1% F*st* values as the threshold, we identified 143 SNPs as candidates for strong selection (Table S4). Of the 142,486 tested SNPs, "BICF2G630738971" ranked the highest at 26,550,871. This SNP is located in the intron of *SOD1* on canine chromosome 31 (Fig. 4A, B), 10,529 bp downstream of *SOD1* c.118G>A. We then calculated the extended haplotype homozygosity (EHH) of each population (2019 vs. 2022) for BICF2G630738971 (Fig. 4C). The 2019 group had longer EEH

haplotypes than the 2022 group, although the position of *SOD1* c.118G>A was highly linked to the top SNP in the 2022 group (0.47 < EHH < 1).

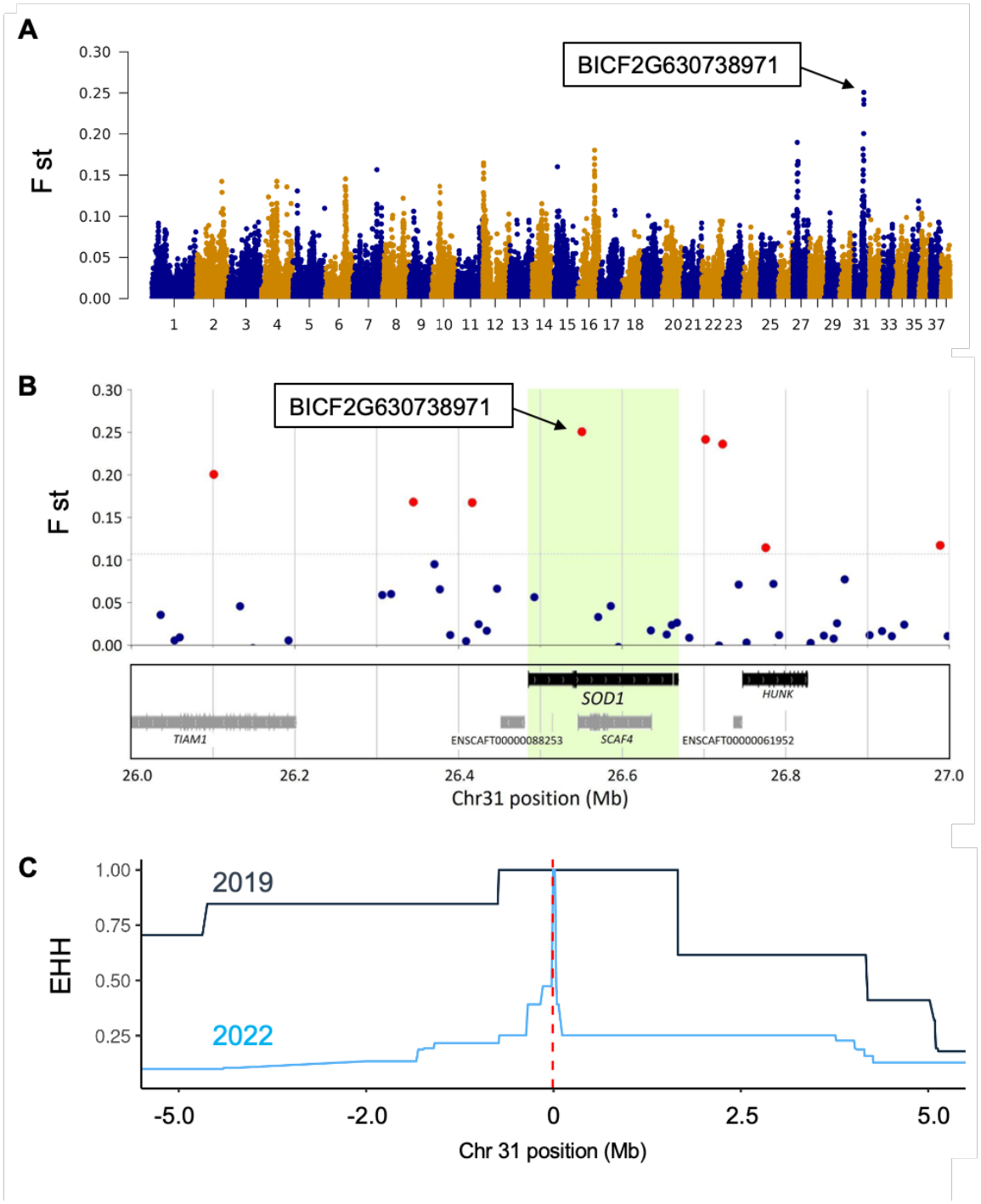


Figure 4. Selection signature observed in *SOD1.*

A: Manhattan plot based on population differentiation (*Fst*). B: Relationships of *Fst* per SNP and their locations around *SOD1* on chromosome 31. SNP BICF2G630738971 has the highest *Fst* value. Red dots indicate SNPs that exceeded the threshold of 0.1% Fst. C: Extended haplotype homozygosity (EHH) of each group from the top *Fst* SNP (BICF2G630738971) for the derived allele. Dark-blue and light-blue lines indicate the 2019 and 2022 groups, respectively. The red dotted line indicates the position of *SOD1* c.118G>A.

To identify genes under selection, we surveyed the 186 genes that included or were near strongly selected candidate SNPs (Table S4). Four protein-coding genes (*ABCA4*, *TNXB*, *COL11A2*, and *SOD1*) were known to affect phenotypes (Fig. S2, Table S5). Gene Ontology (GO) analysis identified three significantly enriched terms (Table 1 and Fig. S3), one related to peptide antigen binding and the others to the major histocompatibility complex (MHC).

Table 1. Gene Ontology (GO) analysis of protein-coding genes near SNPs with high Fst values.

Molecular function	GO	Number of genes	Fold enrichment	FDR
MHC class II protein complex binding	0023026	6	48	7.08E-05
MHC protein complex binding	0023023	6	39	9.01E-05
Peptide antigen binding	0042605	4	32	2.47E-02

Discussion

Our analysis revealed that the availability of DTC genetic testing coincided with a decrease in the frequency of homozygous *SOD1*:c.118G<A among PWCs, reflecting negative selection against the mutation. Our study provides valuable empirical evidence that genetic testing, coupled with selective breeding, can lower mutation frequency in the span of a few years. With the widespread, global availability of commercial genetic testing for *SOD1* mutations (Neeves and Granger 2015), breeding programs can apply the test results and inform systematic mate selection to decrease the frequency of this deleterious variant.

Our research also demonstrated that the mutation can be selected against without lowering the effective population size (i.e., generating inbred animals). Inbreeding avoidance is essential for effective animal breeding (Sams and Boyko 2019). Our genetic analysis revealed no differences in Ho and inbreeding levels between the 2019 and 2022 groups. In addition, we estimated a larger effective population size for the 2022 group (73.1) than for the 2019 group (48.9). The phylogenetic analysis (Fig. 3C) further implied that the genetic origin of some Wild dogs in 2022 was from other lineages. Taken together, our findings suggest that PWC breeders used genetic testing results to limit inbreeding while mating dogs from different families or lineages.

Studies investigating selection signatures during dog domestication (Akey et al. 2010; Wang et al. 2013; Plassais et al. 2019) have identified an influence on phenotypes such as body size, coat color, and behavior. Selection scans mainly focused on a given region over 10,000 years or similarly long periods (Akey et al. 2010). In contrast, selection scans over short periods (i.e., a few years) are rare in mammals. Such studies also focused on phenotype-driven selection rather than genotype-based selection, as we did in this study. Here, we identified four genes (*ABCA4*, *TNXB*, *COL11A2*, and *SOD1*) reported to be associated with disease phenotypes. First, our results indicated that selective breeding was based on the *SOD1* mutation, the cause of canine DM. Next, *ABCA4* is associated with an autosomal recessive retinal degenerative disease in Labrador retrievers (Mäkeläinen et al. 2019). Whole-genome sequencing of mixed-breed dogs identified two missense variants in *TNXB* that caused an Ehlers-Danlos syndrome-like signature (Bauer et al. 2019). Finally, the missense variant of *COL11A2* is associated with skeletal dysplasia 2 and is

inherited as a monogenic autosomal recessive trait with incomplete penetrance, primarily in Labrador retrievers (Frischknecht et al. 2013). No known association exists between these genes and *SOD1* or DM in dogs. Thus, why genomic regions other than *SOD1* experienced selection remains unknown. One likely explanation is that unidentified functional genes may be present in the vicinity of selected regions. These unknown genes may be associated with *SOD1* or DM while also being linked to valuable traits beneficial for individual survival and/or breeding performance.

The MHC plays an important role in pathogen resistance (Debenham et al. 2005). Canine MHC, also referred to as dog leukocyte antigen, exhibits genetic variations (Kennedy et al. 2007; Niskanen et al. 2013) that are associated with autoimmune disease (Jokinen et al. 2011; Gershony et al. 2019). In this study, we found two GO terms related to the MHC. Considering these terms were identified from the test results of pups, our findings suggest that selective pressure on the MHC region was related to survival in pups but not in adults, likely because the kennel environment necessitated the rapid development of a strong pathogen-resistant autoimmune system. Evidence of post-copulatory selection at *MHC* loci in puppies supports this hypothesis (Niskanen et al. 2016). However, our study did not sequence MHC haplotypes or obtain relevant phenotypes for each dog (e.g., survival rate). Thus, further research is required to uncover the mechanisms (e.g., targeting haplotypes) underlying MHC selection.

In conclusion, this study highlights the value of genetic testing as a tool to lower the risk of canine DM while avoiding inbred animals. Only a few years are required to decrease the number of homozygous mutant dogs. Notably, our study focused on the genotype of the *SOD1* mutation and genome-wide SNPs rather than on the phenotype. In the current work, we carried out a genome-wide analysis of short-term selection in mammals. Our results highlight that genetic testing could reduce the prevalence of predictable genetic conditions, thus contributing to improved animal welfare. Future studies focused on elucidating gene function will improve our understanding of the short-term selection process described herein.

The median onset of DM in PWCs is 11 years (Coates et al. 2007). Considering when DTC genetic testing on PWCs began in Japan, the corresponding decreases in *SOD1* mutation

frequency (and the prevalence of associated disease) should be noticeable around the 2030s. Accordingly, further research using comprehensive phenotypic datasets, such as those available from pet insurance companies, is needed to monitor DM onset in PWCs.

Methods

Ethics statement

We obtained all swab samples from dogs with the consent of their owners. The ethics committee at the Anicom Specialty Medical Institute approved the study procedures (ID:2022-01).

Genotyping for the SOD1:c.118G<A variant and statistical testing

Buccal swabs from 5,512 PWCs were sampled by breeders/owners and pet stores from all over Japan between Jan 1, 2017, and Dec 31, 2022, with their consent. DNA was extracted from oral mucosal tissue using a commercial kit (chemagic™ DNA Buccal Swab Kit, PerkinElmer and DNAdvance Kit, Beckman Coulter). We performed real-time PCR to determine genotypes of DM-associated mutations (*SOD1*:c.118G<A), specifically wild-type homozygotes (G/G), heterozygous carriers (G/A), and variant homozygotes (A/A), as previously reported (Zeng et al. 2014).

We used Fisher's exact test in R to compare the decrease in diploid genotypes between 2016 and 2022 (RCore 2016). Next, we employed the learnPopGen package in R to simulate allele frequencies under genetic drift (Revell 2019). Because large-scale genetic testing began in 2017, we used data from the previous year (2016) in the simulation. In Japan, an estimated 10% of dogs sold annually are used for breeding (https://www.env.go.jp/nature/dobutsu/aigo/2_data/pamph/rep_h1503.html). Thus, 540 dogs were set as the effective population size in 2016, given that 5,395 PWCs were registered by the Japan Kennel Club (https://www.jkc.or.jp/archives/enrollment/4598). The breeding cycle was set as 0.67 (2/3) per year, indicating that four breeding generations had passed over 6 years (2016–2022). Using these parameters, we simulated genetic drift 10,000 times (simulated allele frequency) and

set the resultant distribution as the null for subsequent comparisons with observed allele frequency through 6 years. Next, we calculated the probability that the observed allele frequency was lower than simulated allele frequencies. Significance was set at $P < 0.05$ for both tests.

Genome-wide SNP genotyping

We used an SNP-genotyping array (Canine 230 K Consortium BeadChip Array, Illumina, San Diego, CA, USA) to investigate the effects of genetic testing on the whole genome. We compared genome-wide SNPs from four subpopulations of wild-type and variant homozygotes of the *SOD1* risk allele, all derived from dogs born in 2019 and 2022. Considering the effect of genetic background on array-based analyses, we selected all dogs from the two largest clients for array-based analyses (Table S1). We used an SNP-genotyping array (Canine 230 K consortium) and the Illumina iScan system to detect over 230 K SNP genotypes. Genotype coordinates corresponded to the CanFam3.1 genome assembly.

We conducted quality control using PLINK version 1.90 (Chang et al. 2015). Missing rates for each locus (geno option) and individual dogs (mind option) were both < 1%. In this study, we did not examine the Hardy-Weinberg equilibrium as we did not assume random mating for pedigreed PWC. In general, population genetic analyses should not include close relatives. Therefore, we conducted robust relationship-based pruning (Manichaikul et al. 2010) using the king-cutoff option in PLINK version 2.00a2LM, with a threshold of 0.176 to exclude pairings between monozygotic twins and first-degree relatives. For allele frequency-based quality control, we excluded SNPs with a minor allele frequency of < 0.01. We also removed sex chromosome variants and autosomal indels (insertions-deletions) because their effects on allele frequency differ from those of autosomal SNPs. The final analysis retained 143,013 SNPs. After quality control for individual dogs, including the removal of related animals, 117 dogs remained: homozygous wild-type (G/G) dogs tested in 2019 (Wild 2019, n = 42) and in 2022 (Wild 2022, n = 49) as well as homozygous mutant dogs tested in 2019 (Mutant 2019, n = 24) and in 2022 (Mutant 2022, n = 2;

Table S2). We employed this dataset for subsequent analyses. SNP data for the PWCs are available from the Dryad database (DOI: 10.5061/dryad.rbnzs7hhk)

Inbreeding levels

We inferred inbreeding levels using specimens from genome-wide SNP genotyping. The R package DetectRuns was used to obtain the proportion of times each SNP fell inside a run per population, corresponding to locus homozygosity or heterozygosity in the respective population. We used the following DetectRuns parameters: minSNP = 41, maxGap 10^6, minLengthBps = 50000, and minDensity = 1/5000.

Population genetics

To clarify the genetic structure and phylogenetic relationships in the PWC population, we performed PCA in PLINK version 1.9 with default settings and maximum-likelihood ancestry analysis in ADMIXTURE version 1.3 (Alexander et al. 2009), respectively. For ADMIXTURE, we set the number of populations (K) between 2 and 10. We used the option cv for cross-validation error calculation to select the optimal K value based on the lowest error.

For genetic relationships, we first converted the PLINK PED format into FASTA format. We then converted the FASTA file to NEXUS format in MEGA X (Kumar et al. 2018). We constructed a phylogenetic tree using the neighbor-joining algorithm with p-distance in MEGA X.

We calculated Nei's standard genetic distance D (Nei 1987) between the four populations (Wild 2019, Wild 2022, Mutant 2019, and Mutant 2022). Nei's D was estimated in GenoDive version 3.06 with default settings (Meirmans and Van Amsterdam 2019).

Estimating effective population size (Ne)

The contemporary *Ne* of PWC was estimated from genome-wide SNP data using a linkage disequilibrium method in NeEstimator version 2.1 (Do et al. 2014). The lowest allele frequency was set at 0.01.

Detecting selection pressure on SNPs

We obtained all gene datasets from the Ensembl Genome Browser (release 104; CanFam 3.1) and retrieved Ensembl gene IDs using Bedtools version 2.27.1 (Quinlan and Hall 2010; Danecek et al. 2021).

Artificial selection alters the frequency of genes associated with phenotypes. To uncover selection signatures in the PWC genome, we estimated Weir and Cockerham population differentiation (*F*st) in the 2019 and 2022 groups. We estimated *F*st values using the –fst option implemented in PLINK version 1.9. The thresholds for selective pressure were SNPs with *F*st values in the upper 0.1%. Genes in the surrounding ± 100 kb of SNPs under selection were considered selected genes.

To investigate haplotype length, we calculated the EHH of wild-type populations. The EHH is a measure used to detect genome regions under recent selective pressure (Sabeti et al. 2002). First, target SNPs under selective pressure in the Wild 2019 and 2022 populations were phased using beagle 5.4 (version 22Jul22.46e) (Browning et al. 2018) with default settings. After phasing, EHH for the target SNP was estimated with a 1 Mbp window in Selscan version 2.0.0 (Szpiech and Hernandez). To compare other genetic regions with the target SNP, we calculated EHH using 10,000 randomly selected SNPs per group.

To determine the functional classes of genes associated with the analyzed traits, we performed GO analysis powered by PANTHER (Thomas et al. 2022). We selected target genes using linkage disequilibrium analysis and the PANTHER Overrepresentation Test (release 20221013). 'Molecular function' was used for domestic dog (*Canis lupus familiaris*) dataset annotation in PANTHER 17.0. We used Ensembl gene IDs for data annotation. Statistical differences were determined with Fisher's exact test. We determined significance at a false discovery rate (FDR) of $P < 0.05$.

Acknowledgments

We thank all participating breeders, veterinarians, and dog owners for collecting buccal swabs from their dogs. We also thank all the Anicom Specialty Medical Institute members for their feedback on this study. This study was supported by the Research Foundation of Anicom Specialty Medical Institute.

Author Contributions

YM and HU designed the study; HU, NA, FY, GI, KO, and YM collected dog samples, performed experiments, contributed to the analytical tools, analyzed the data, and wrote the paper.

Competing Interest Statement

HU, FY, and KO received salaries from Anicom Pafe, a genetic testing company in Japan. NA, GI, and YM received wages from Anicom Specialty Medical Institute. YM also received a stipend from Azabu University.

References

Akey JM, Ruhe AL, Akey DT, Wong AK, Connelly CF, Madeoy J, Nicholas TJ, Neff MW. 2010. Tracking footprints of artificial selection in the dog genome. *Proc Natl Acad Sci U S A* **107**: 1160–1165.

Akiyama N, Suzuki R, Saito T, Yuchi Y, Ukawa H, Matsumoto Y. 2023. Presence of known feline ALMS1 and MYBPC3 variants in a diverse cohort of cats with hypertrophic cardiomyopathy in Japan. *PLoS One* **18**: e0283433.

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**: 1655–1664.

Awano T, Johnson GS, Wade CM, Katz ML, Johnson GC, Taylor JF, Perloski M, Biagi T, Baranowska I, Long S, et al. 2009. Genome-wide association analysis reveals a SOD1 mutation in canine degenerative myelopathy that resembles amyotrophic lateral sclerosis. *Proc Natl Acad Sci U S A* **106**: 2794–2799.

Ayala-Valdovinos MA, Gomez-Fernandez A. 2018. Frequency of canine degenerative myelopathy SOD1: c. 118G> A mutation in 22 dog breeds in Guadalajara, Mexico. *de Ciencias Pecuarias*. http://www.scielo.org.co/scielo.php?pid=S0120-06902018000200150&script=sci_arttext&tlng=en.

Bauer A, de Lucia M, Leuthard F, Jagannathan V, Leeb T. 2019. Compound heterozygosity for TNXB genetic variants in a mixed-breed dog with Ehlers-Danlos syndrome. *Anim Genet* **50**: 546–549.

Boyko AR, Boyko RH, Boyko CM, Parker HG, Castelhano M, Corey L, Degenhardt JD, Auton A, Hedimbi M, Kityo R, et al. 2009. Complex population structure in African village dogs and its implications for inferring dog domestication history. *Proc Natl Acad Sci U S A* **106**: 13903–13908.

Browning BL, Zhou Y, Browning SR. 2018. A One-Penny Imputed Genome from Next-Generation Reference Panels. *Am J Hum Genet* **103**: 338–348.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**: 7.

Chang H-S, Kamishina H, Mizukami K, Momoi Y, Katayama M, Rahman MM, Uddin MM, Yabuki A, Kohyama M, Yamato O. 2013. Genotyping assays for the canine degenerative myelopathy-associated c.118G>A (p.E40K) mutation of the SOD1 gene using conventional and real-time PCR methods: a high prevalence in the Pembroke Welsh Corgi breed in Japan. *J Vet Med Sci* **75**: 795–798.

Chu ET, Simpson MJ, Diehl K, Page RL, Sams AJ, Boyko AR. 2019. Inbreeding depression causes reduced fecundity in Golden Retrievers. *Mamm Genome* **30**: 166–172.

Coates JR, March PA, Oglesbee M, Ruaux CG, Olby NJ, Berghaus RD, O'Brien DP, Keating JH, Johnson GS, Williams DA. 2007. Clinical characterization of a familial degenerative myelopathy in Pembroke Welsh Corgi dogs. *J Vet Intern Med* **21**: 1323–1331.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* **10**. http://dx.doi.org/10.1093/gigascience/giab008.

Debenham SL, Hart EA, Ashurst JL, Howe KL, Quail MA, Ollier WER, Binns MM. 2005. Genomic sequence of the class II region of the canine MHC: comparison with the MHC of other mammalian species. *Genomics* **85**: 48–59.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Mol Ecol Resour* **14**: 209–214.

Dreger DL, Hooser BN, Hughes AM, Ganesan B. 2019. True Colors: Commercially-acquired morphological genotypes reveal hidden allele variation among dog breeds, informing both trait ancestry and breed …. *PLoS*. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0223995&fbclid=IwAR23tHir7I-GlC1yr-pCuZPuYl46Rp3bIvZdcH4Rob97a6iTHTHYu4yEjKk.

Dreger DL, Rimbault M, Davis BW, Bhatnagar A, Parker HG, Ostrander EA. 2016. Whole-genome sequence, SNP chips and pedigree structure: building demographic profiles in domestic dog breeds to optimize genetic-trait mapping. *Dis Model Mech* **9**: 1445–1460.

Frischknecht M, Niehof-Oellers H, Jagannathan V, Owczarek-Lipska M, Drögemüller C, Dietschi E, Dolf G, Tellhelm B, Lang J, Tiira K, et al. 2013. A COL11A2 mutation in Labrador retrievers with mild disproportionate dwarfism. *PLoS One* **8**: e60149.

Gershony LC, Belanger JM, Short AD, Le M, Hytönen MK, Lohi H, Famula TR, Kennedy LJ, Oberbauer AM. 2019. DLA class II risk haplotypes for autoimmune diseases in the bearded collie offer insight to autoimmunity signatures across dog breeds. *Canine Genet Epidemiol* **6**: 2.

Jokinen P, Rusanen EM, Kennedy LJ, Lohi H. 2011. MHC class II risk haplotype associated with canine chronic superficial keratitis in German Shepherd dogs. *Vet Immunol Immunopathol* **140**: 37–41.

Kawakami T, Raghavan V, Ruhe AL, Jensen MK, Milano A, Nelson TC, Boyko AR. 2022. Early onset adult deafness in the Rhodesian Ridgeback dog is associated with an in-frame deletion in the EPS8L2 gene. *PLoS One* **17**: e0264365.

Kennedy LJ, Barnes A, Short A, Brown JJ, Lester S, Seddon J, Fleeman L, Francino O, Brkljacic M, Knyazev S, et al. 2007. Canine DLA diversity: 1. New alleles and haplotypes. *Tissue Antigens* **69 Suppl 1**: 272–288.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**: 1547–1549.
Mäkeläinen S, Gòdia M, Hellsand M, Viluma A, Hahn D, Makdoumi K, Zeiss CJ, Mellersh C, Ricketts SL, Narfström K, et al. 2019. An ABCA4 loss-of-function mutation causes a canine form of Stargardt disease. *PLoS Genet* **15**: e1007873.
Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**: 2867–2873.
Meirmans P, Van Amsterdam I. 2019. GenoDive. http://www.bentleydrummer.nl/software/GenoDiveManual.pdf.
Mellanby RJ, Ogden R, Clements DN, French AT, Gow AG, Powell R, Corcoran B, Schoeman JP, Summers KM. 2013. Population structure and genetic heterogeneity in popular dog breeds in the UK. *Vet J* **196**: 92–97.
Moses L, Niemi S, Karlsson E. 2018. Pet genomics medicine runs wild. *Nature Publishing Group UK*. http://dx.doi.org/10.1038/d41586-018-05771-0 (Accessed May 25, 2023).
Neeves J, Granger N. 2015. An update on degenerative myelopathy in dogs. *Companion Animal* **20**: 408–412.
Nei M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press.
Niskanen AK, Hagström E, Lohi H, Ruokonen M, Esparza-Salas R, Aspi J, Savolainen P. 2013. MHC variability supports dog domestication from a large number of wolves: high diversity in Asia. *Heredity* **110**: 80–85.
Niskanen AK, Kennedy LJ, Lohi H, Aspi J, Pyhäjärvi T. 2016. No evidence of prenatal diversifying selection at locus or supertype levels in the dog MHC class II loci. *Canine Genet Epidemiol* **3**: 9.
Plassais J, Kim J, Davis BW, Karyadi DM, Hogan AN, Harris AC, Decker B, Parker HG, Ostrander EA. 2019. Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nat Commun* **10**: 1489.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.
RCore T. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Revell LJ. 2019. learnPopGen: An R package for population genetic simulation and numerical analysis. *Ecol Evol* **9**: 7896–7902.
Rokhsar JL, Canino J, Raj K, Yuhnke S, Slutsky J, Giger U. 2021. Web resource on available DNA variant tests for hereditary diseases and genetic predispositions in dogs and cats: An Update. *Hum Genet* **140**: 1505–1515.
Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, et al. 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**: 832–837.
Sams AJ, Boyko AR. 2019. Fine-Scale Resolution of Runs of Homozygosity Reveal Patterns of Inbreeding and Substantial Overlap with Recessive Disease Genotypes in Domestic Dogs. *G3* **9**: 117–123.
Szpiech ZA, Hernandez RD. selscan: An Efficient Multithreaded Program to Perform EHH-Based Scans for Positive Selection. *Evolution* **31**: 2824–2827.
Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou L-P, Mi H. 2022. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci* **31**: 8–22.
Wang G-D, Zhai W, Yang H-C, Fan R-X, Cao X, Zhong L, Wang L, Liu F, Wu H, Cheng L-G, et al. 2013. The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nat Commun* **4**: 1860.
Zeng R, Coates JR, Johnson GC, Hansen L, Awano T, Kolicheski A, Ivansson E, Perloski M, Lindblad‐Toh K, O'Brien DP, et al. 2014. Breed Distribution of *SOD 1* Alleles Previously Associated with Canine Degenerative Myelopathy. *Journal of Veterinary Internal Medicine* **28**: 515–521. http://dx.doi.org/10.1111/jvim.12317.