



Genetic testing predicts appearance but not behavior in dogs

Kathryn A. Lord^{a,b,1}, Vista Sohrab^{a,b,c}, Kasia Bryc^{b,d}, Michelle E. White^b, Brittney Kenney^{a,b}, Kathleen Morrill Pirovich^{a,b,c}, Frances L. Chen^{a,b}, and Elinor K. Karlsson^{a,b,d,e,1}

Affiliations are included on p. 8.

Edited by Elaine Ostrander, National Human Genome Research Institute, Bethesda, MD; received December 12, 2024; accepted June 6, 2025

Genetic tests for behavioral and personality traits in dogs are now being marketed to pet owners, but their predictive accuracy has not been validated. To evaluate the reliability of such tests, we analyzed data from Darwin's Ark, a community science initiative that includes over 3,000 dogs with both genetic data and individual-level behavioral phenotypes. None of the candidate variants had significant associations or predictive power for behavioral traits as previously reported. However, we found strong associations with aesthetic traits that differentiate breeds, such as height, leg length, and ear shape. Our results suggest that earlier studies using breed-average phenotypes, rather than individually measured phenotypes, were confounded by population structure. Behavior in dogs is polygenic and complex, and thus cannot be accurately predicted using tests that consider only a few genetic variants. Furthermore, behavior in dogs is only moderately heritable, and environmental influences inherently limit the potential accuracy of genomic predictions. Developing meaningful, accurate genetic predictions for complex traits that can improve dog health and welfare will require very large cohorts of individually phenotyped dogs.

genomic prediction | genetic testing | behavior | dog | artificial selection

Genome-wide association studies (GWAS) are a powerful tool for identifying genetic variants, typically single-nucleotide polymorphisms (SNPs), associated with traits or diseases. GWAS are effective for both simple traits caused by mutations in a single gene and complex traits influenced by many genetic and environmental factors. In humans, GWAS have revealed novel pathways underlying complex diseases and opened new avenues for therapeutic development (1). The power of GWAS is largely determined by sample size, and the largest GWAS include millions of individuals (2).

To estimate an individual's risk of exhibiting a complex trait, researchers use polygenic risk scores (PRS), which aggregate the effects of many genetic variants across the genome discovered through GWAS. The potential for accurate prediction depends on both the heritability of a trait, meaning the proportion of variation explained by genetic factors, and its polygenicity, or the number of genetic variants that influence it. Traits with lower heritability require much larger sample sizes, and traits with very low heritability cannot be predicted accurately from genetic information alone (3, 4).

GWAS has been widely used in dogs to identify genomic loci associated with aesthetic traits, like size and color, behavioral traits, and disease risk. The history of dog breeding makes GWAS in dogs unusually powerful. About 200 years ago, humans began purposeful breeding of lineages of dogs with unusual physical characteristics (5), creating hundreds of modern breeds (6–11). The limited genetic diversity within dog breeds enhances the power to detect associations in GWAS (12, 13). However, this history also exacerbates issues of population structure known to confound genetic studies by causing both false positives (14) and missed true associations (15).

To date, genetic prediction in dogs has been limited to simple traits and highly heritable complex traits. For example, genotyping a dog at a single locus can determine whether they have an inherited form of progressive blindness (2, 3). Aesthetic traits under selection in breeds are often exceptionally heritable and explained by a small number of large-effect loci, making them particularly amenable to genetic prediction. We developed a highly accurate PRS for height, which is more than 99% heritable, by combining information from 2,733 SNPs identified in a GWAS with 1,730 dogs (16, 17).

Currently, the largest GWAS in dogs have fewer than 10,000 individuals (16, 18), making developing PRS for less heritable traits infeasible (19). This includes behavioral and personality traits in dogs, which are only weakly or moderately heritable in studies that use paired phenotype and genetic data from individual dogs (11, 16, 20–22). Many factors other than genetics influence behavior, including training, socialization, housing,

Significance

Using data from Darwin's Ark, a community science project with over 3,000 dogs with genetic and behavioral data, we critically evaluate reported behavior-associated genetic variants. We find no evidence that these variants predict behavior. Some variants are correlated with aesthetic traits that define breeds, suggesting that the earlier breed studies that linked these variants to behavior were confounded by the complex dog population structure. Genetic tests focusing on a few variants are unlikely to provide accurate predictions for polygenic behavioral traits or complex diseases in dogs. Instead, large, carefully designed studies that consider the whole genome are essential for developing predictions that can improve dog health and welfare.

Author contributions: K.A.L., K.B., F.L.C., and E.K.K. designed research; B.K. performed research; V.S., K.B., and K.M.P. analyzed data; K.A.L. edited manuscript; K.B. edited paper; M.E.W. and F.L.C. performed data acquisition; B.K. performed data acquisition and edited manuscript; K.M.P. edited Manuscript; E.K.K. conceived of project and edited manuscript; and K.A.L., V.S., K.B., M.E.W., F.L.C., and E.K.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2025 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](#).

K.A.L. is an organizer of this Special Feature.

¹To whom correspondence may be addressed. Email: kathryn.lord@umassmed.edu or elinor.karlsson@umassmed.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2421752122/-DCSupplemental>.

Published November 24, 2025.

and early-life experiences. Developing accurate PRS in dogs for such traits will likely require studies with tens or hundreds of thousands of dogs (23–26).

To circumvent the sample size problem, some studies combine behavioral scores from one set of dogs with genetic data from a different set of dogs. To test for association, dogs with genetic data are assigned a phenotype based on the average score of dogs in their breed (27–34). This approach has been effective when applied to aesthetic traits that have very little variability within breeds (35). However, behavioral traits in dogs vary nearly as much within breeds as between breeds (16, 36–39). Furthermore, the true sample size of these studies is much smaller than conventional genetic association studies. While they include large numbers of dogs, their statistical power is ultimately constrained by the number of independent units of analysis (the number of breeds), and the analysis vulnerable to pseudoreplication if nonindependent observations are treated as if they were independent (40). In this manuscript, we use the term GWAS exclusively to describe studies that pair genetic data with individual phenotype data, consistent with common terminology, and “breed-average studies” to describe those using population-level phenotypes.

To date, breed-average genetic studies have reported hundreds of SNPs associated with behavioral traits (27–34) conflicting with results from GWAS, which have identified only a few significant associations (16, 21, 22, 40, 41). Despite this, genetic tests based on these studies are being marketed to consumers (42, 43). For example, if a dog is labeled as genetically predisposed to aggression, an owner might limit essential social interactions, or a shelter might decide against adoption.

Here, we use Darwin’s Ark, a large community science project collecting individual phenotypes and genetic data (44), to investigate whether genetic variants from breed-average studies can accurately predict behavior in pet dogs from diverse breeds, mixed-breed dogs, and nonbreed dogs.

Results

Identifying 151 Candidate Behavior SNPs from Previous Studies. We reviewed recent studies that relied on breed-average phenotypes to identify candidate SNPs associated with behavior in dogs (27–30, 32–34, 45). These studies have reported hundreds of significantly associated SNPs, and several have proposed that behavior and personality traits of both breeds (27) and individual dogs (28) can be predicted from these variants. However, inferring individual-level associations from group-averaged data is statistically problematic (46). In genetic studies, individual-level data are essential for controlling for relatedness, population stratification, and other factors that lead to biased and spurious results (47, 48). These issues are especially pronounced in dogs due to high stratification between breeds and close relatedness within breeds (49).

Four of the studies scored behavioral traits that are also measured in Darwin’s Ark (*SI Appendix, Table S1*) (27–29, 34, 50). In each of these studies, the average score on each trait for each breed was calculated using survey data from dogs without accompanying genetic data (51), and those average scores were used to assign phenotypes to unphenotyped dogs genotyped or sequenced for other studies. This included genome-wide studies of (1) 14 traits in 101 breeds (34); (2) canine fear and aggression traits in 29 breeds across two published datasets (27); (3) 17 behavioral traits using genetic data from three studies with 29 to 50 breeds (29). The fourth study used SNPs from the previous three studies to predict problem behaviors in 397 dogs (28). In total, we identified 151 candidate behavior SNPs (*SI Appendix, Table S2*).

Many of the SNPs proposed to predict behavior in dogs occur near genes known to influence aesthetic traits that differentiate dog breeds and are artificially selected for by breeders (*SI Appendix, Table S3*) (30, 52–83), suggesting that population structure rather than associations with behavioral traits may explain these associations (15, 84, 85). For example, 18 of the 20 SNPs proposed to predict behavior in individual dogs in a recent publication (28) overlap loci associated with breed-defining traits: nine loci associated with size, including near the major size-associated genes IGF1 (2 SNPs), IGF2BP2 (1 SNP), HMGA2 (2 SNPs), and MSRB3 (2 SNPs) (16, 30, 74, 75, 77); five loci associated with skeletal morphology (73, 74, 76); and four associated with fur texture and coat color, including at RSPO2 (2 SNPs), FGF5 (1 SNP), and MITF (1 SNP) (53, 59, 73, 74).

No Evidence of Association with Behavior in Pet Dogs. With the advent of large-scale community science in pets, we now have the capacity to do GWAS in dogs as a more powerful and less problematic alternative to breed-average studies. GWAS require individual-level phenotype and genetic data but are an exceptionally powerful tool for discovering variants associated with complex traits (1). Through our online community science platform Darwin’s Ark (www.darwinsark.org) we have collected both phenotypes and genetic data for thousands of dogs (16). Darwin’s Ark is a nonprofit, open data initiative that engages dog and cat owners in research by inviting them to complete short surveys and offering an opportunity to sequence their pet’s genome. In previous work, we showed that owner survey responses align with those from dog experts and are not strongly influenced by cultural breed stereotypes (*SI Appendix, Supporting Text*) (16).

In contrast to studies that rely on breed-average phenotypes, Darwin’s Ark includes both single-breed and mixed-breed dogs in proportions similar to the US pet dog population (16), making it ideal for assessing predictive tests marketed to dog owners. We generated high-density, highly accurate genotype data for 3,287 dogs enrolled in Darwin’s Ark through low-coverage sequencing (mean coverage = 0.82x) and imputation using reference genomes from Dog10K (86) (*SI Appendix, Supporting Text*). In total, we call genotypes for 29,089,701 autosomal SNPs. We estimated each dog’s breed ancestry using global admixture analysis, comparing their genotype to a reference panel of 1,212 dogs from 101 different breeds, following previously published methods (16). Only a minority of dogs (25%) have ancestry from just one breed (single-breed or “purebred” dogs) (*SI Appendix, Table S4*), and most (53%) have discernible ancestry from at least four breeds (Fig. 1A). The dogs span a wide range of ages, from infant to 22 years of age, with a median age of five years (IQR=6.7) (Fig. 1B). The most common breed ancestry is American pit bull terrier (11.7%) followed by Labrador retriever (6.8%) and Chihuahua (6.4%) (Fig. 1C).

The Darwin’s Ark survey data include 280 questions, from which we previously derived eight behavioral factors (16). Of these, 34 questions and four behavioral factors correspond to descriptions of seven of the fourteen C-BARQ traits used in the four breed-average studies (*SI Appendix, Tables S1 and S5*). Another ten questions measure aesthetic characteristics that differentiate breeds.

We ran GWAS in the current Darwin’s Ark dataset for all questions that overlapped with the breed-average studies, all eight behavioral factors and all aesthetic traits. This increased our sample size by about 1.4× (1,794 to 3,074 dogs per trait) compared to our previous dataset, including six additional behavioral traits and four additional aesthetic traits (*SI Appendix, Table S5*). We controlled for population structure and cryptic relatedness using a

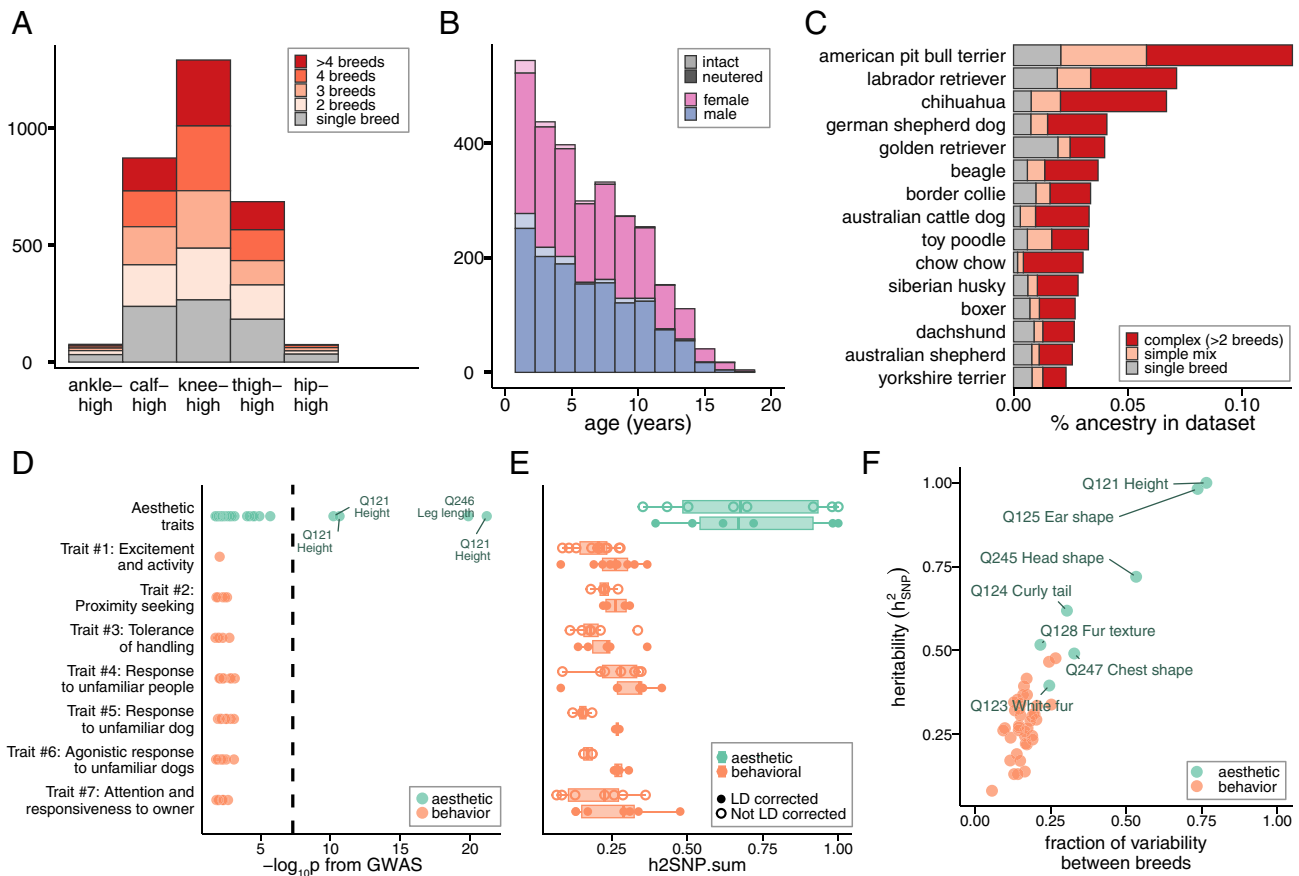


Fig. 1. No evidence for association with candidate behavior SNPs in a population of 3,295 pet dogs. (A) Dogs vary widely in size, and most have ancestry from more than one breed. About 1 in 4 (24.4%) are single-breed (“purebred”) dogs. (B) Ages range from 0 to 22 y (mean 5.7 ± 4.0 SD), with an even sex distribution (49.9% female; 93.8% spayed or neutered). (C) The 15 most common breed ancestries, drawn from at least 85 breeds contributing at least 5% ancestry to at least one dog. (D) In GWAS, some of the candidate behavior SNPs are significantly associated with aesthetic traits (green), but none are associated with any of the seven behavior traits (orange). The vertical dashed line marks the genome-wide significance threshold ($P < 5 \times 10^{-8}$); significant associations are labeled by trait. (E) Aesthetic traits are more heritable than behavior traits. SNP-based heritability estimates (h^2) are higher for aesthetic traits (green) than for behavioral traits (orange). Points show results from LD-corrected (open circles, *Top*) and uncorrected GREML (solid circles, *Bottom*). Boxplot shows the median (horizontal line), interquartile range (box), and whiskers extending to the smallest and largest values within $1.5 \times$ IQR of the lower and upper quartiles. (F) All behavioral traits (orange points) are more variable within breed than between breeds, but some aesthetic traits (green points with labels) are less variable within breeds than between breeds. The fraction of variability between breeds is strongly correlated with trait heritability (Pearson test; $r = 0.92$, $P = 5 \times 10^{-19}$).

mixed linear model approach incorporating a genetic kinship matrix. We included age, sex, and spay/neuter status as covariates in all GWAS, and size as a covariate in all but the size GWAS. In total, we identified 6596 SNPs significantly associated with 20 aesthetic traits (6570 SNPs) and 7 behavioral traits (26 SNPs).

The inclusion of mixed-breed dogs in our GWAS increases our statistical power. We compared GWAS in a cohort of dogs with exclusively mixed-breed ancestry (1,764 dogs with no more than 45% ancestry from any single breed) to GWAS in the same number of dogs randomly sampled from the full cohort, in which about 40% of the dogs are either single-breed or have at least half their ancestry from one breed. A majority of SNPs (61%; 690/1,132) that were genome-wide significant in the full cohort showed stronger associations in the mixed-breed subset than in the random subset (Wilcoxon $P = 1.6 \times 10^{-19}$). This is consistent with lower pairwise relatedness in the mixed-breed cohort increasing effective sample size, and thereby improving GWAS power.

Our GWAS replicate 62% of 97 previously published associations for aesthetic traits in dogs and find 30 new associations (*SI Appendix, Table S3* and *Dataset S3*), confirming that we have sufficient power to find associations with highly heritable traits at our current sample sizes. It also illustrates the power of owner surveys for finding loci associated with aesthetic traits using simple survey questions (Fig. 2). We find an association with black coat

color in the gene EPB41L1 (chr24:25568410; 3.91×10^{-10} ; $\beta = 0.110$), which is also associated with hair color in humans (87), and an association with red, brown, or tan coat color at the gene MFSD12 (chr20:56248427; $P = 9.39 \times 10^{-22}$; $\beta = -0.165$), which was previously linked to pigment intensity in both dogs (88) and humans (89). We found eight associations for behavioral traits not previously reported in dogs (*SI Appendix, Table S6*). The strongest association, to how affectionate owners say their dog is, is in the gene SH3BP4, which may be associated with Alzheimer’s disease in humans (90) (Fig. 3A).

We did not replicate any of the associations reported in the breed-average studies. Out of the 135 candidate behavior SNPs in the GWAS, none had significant associations with any behavioral trait (*SI Appendix, Table S7* and Fig. 1D). We found four significant associations of candidate behavior SNPs with aesthetic traits. None of the candidate behavior SNPs overlapped regions associated with behavioral traits, but 17 overlapped regions associated with an aesthetic trait, including nine associated with coat color, eight associated with ear shape, six associated with fur length, and six associated with height and leg length (*SI Appendix, Table S6*).

Behavioral traits are less heritable than aesthetic traits. We confirmed that behavioral traits have lower heritability than aesthetic traits, suggesting that even with much larger GWAS,

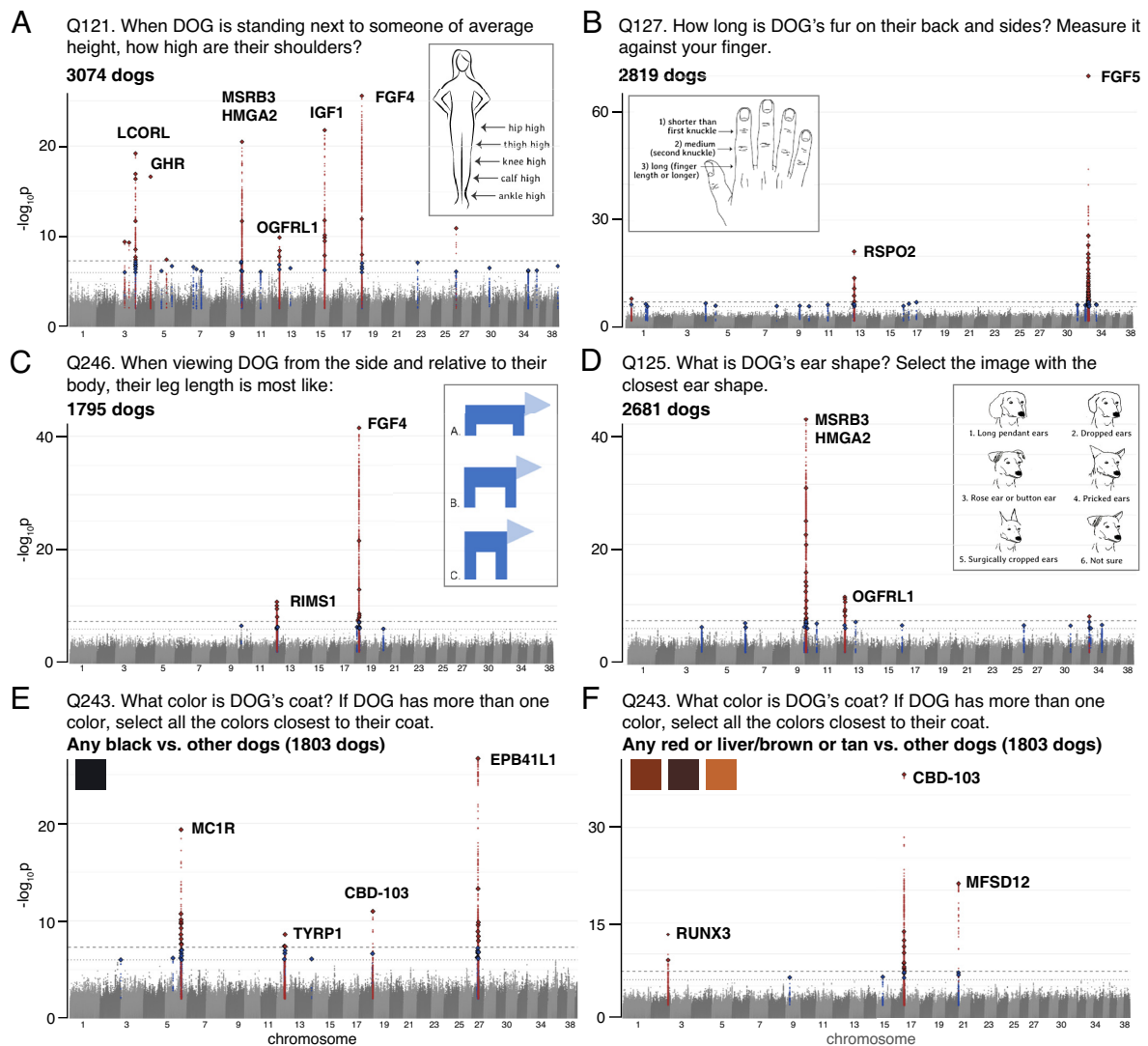


Fig. 2. GWAS in dogs using survey phenotypes finds known and novel associations with aesthetic traits. On the Darwin's Ark survey platform, the word "DOG" in the question is replaced with the dog's name to personalize the experience. Rather than requesting exact measurements, we collect estimates for (A) height relative to human height and (B) fur length relative to human finger length. We use visual aids to ask about (C) leg length and (D) ear shape. A question about coat color finds associations for (E) black and (F) red, liver, brown, or tan.

attempts to predict the behavioral characteristics of individual dogs using genetic information alone will have limited accuracy for many traits (3). Heritability for eight behavioral factors and 34 behavioral questions ranged from $h^2_{\text{SNP}} = 0.08$ to 0.48 (mean = 0.28 ± 0.089 SD) (Fig. 1E) (SI Appendix, Table S8). At current sample sizes, our most heritable behavior question (Q107 "Dog is easy to train"; $h^2_{\text{SNP}} = 0.48 \pm 0.07$ SD) had no significant associated

peaks (Fig. 3B). In contrast, aesthetic traits are highly heritable, ranging from $h^2_{\text{SNP}} = 0.48$ to 1.00 (mean = 0.68 ± 0.24 SD).

Candidate Behavior SNPs Are Correlated With Breed Ancestry. The candidate behavior SNPs from the breed-average studies are unusually differentiated between breeds, which may explain why they appeared associated with traits in breed-average

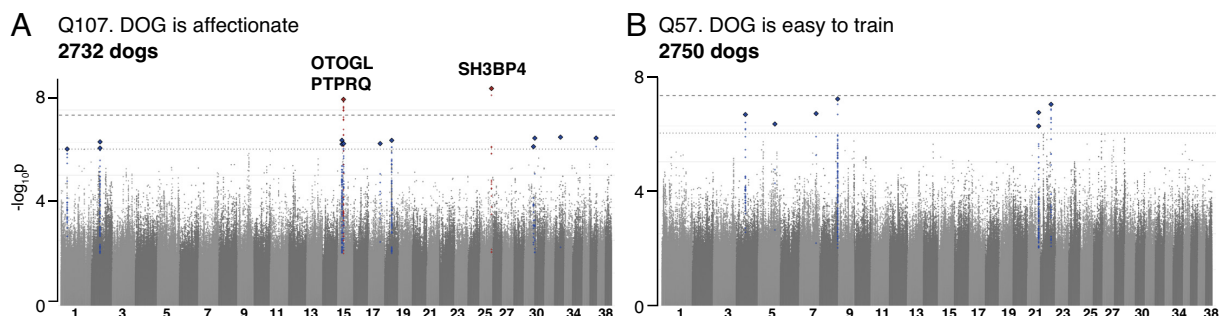


Fig. 3. Few significant associations with behavior traits. (A) The strongest associations are to how affectionate a dog is, which is a moderately heritable trait ($h^2_{\text{SNP}} = 0.31 \pm 0.082$ SE). (B) Some phenotypes like Q57 "DOG is easy to train" are highly heritable ($h^2_{\text{SNP}} = 0.48 \pm 0.07$ SE) but have no significant associations at current sample sizes, suggesting a polygenic architecture.

comparisons. Even in studies using individual-level data, population stratification, which is when allele frequencies and traits covary due to shared ancestry rather than causality, can produce spurious associations at highly differentiated markers (91). Many of the candidate behavior SNPs have unusually high differentiation, indicating that a large proportion of their genetic variation is structured by breed (Fig. 4A). Among breeds represented by more than 20 single-breed dogs in Darwin's Ark, Wright's F_{st} for the 135 candidate behavior SNPs in the GWAS averages 0.19 ± 0.13 SD and 23 SNPs have $F_{st} > 0.25$ (SI Appendix, Table S2) (92).

Many of the candidate SNPs are also strongly correlated with breed ancestry, which is further evidence of population stratification. We measured the correlation between genotype and breed ancestry estimates for the 70 candidate behavior SNPs overlapping with Darwin's Ark (Dataset S4). For each of 101 breeds, we calculated the correlation between SNP genotype and the proportion of ancestry from that breed in each dog, including both mixed-breed and single-breed dogs. Two-thirds (47 SNPs) are correlated with ancestry from at least one breed ($|r_{\text{spearman}}| > 0.1$), and 30% (21 SNPs) are correlated with ancestry from at least five breeds (Fig. 4B). At a subset of SNPs, the correlation of allele frequency with breed varies with breed standard height in an almost textbook illustration of population stratification (Fig. 4C). This is particularly evident at SNPs near genes previously associated with height

or leg length such as IGF1, IGF2BP2, HMGA2, and the Fgf4 retrogene on chromosome 18 (30, 72, 74, 75, 77).

Breed-Average Studies Were Confounded by Population Structure. The most likely explanation for the strong associations observed in breed-average studies is confounding due to population structure. None of the candidate SNPs were significantly associated with behavior phenotypes in GWAS in a genetically diverse population of pet dogs. Furthermore, we detected very few significant associations for any other behavioral traits, consistent with our sample size of ~3,000 dogs being underpowered for complex trait GWAS. Consequently, developing meaningful polygenic risk scores, which would require substantially larger cohorts and independent validation datasets, remains unfeasible.

To assess whether these candidate SNPs may have initially been flagged in breed-average studies because of population structure, we took a straightforward modeling approach. For each of the 38 behavioral traits and four aesthetic traits, we tested the predictive value of each candidate SNP using a likelihood ratio test comparing a full regression model (including the SNP) to a reduced model (excluding the SNP). To evaluate whether associations were mediated by other factors, we included covariates such as sex (including spay/neuter status), age, body size (height and leg length), and breed ancestry. We excluded two SNPs significantly associated with aesthetic traits not included in this model from the behavior

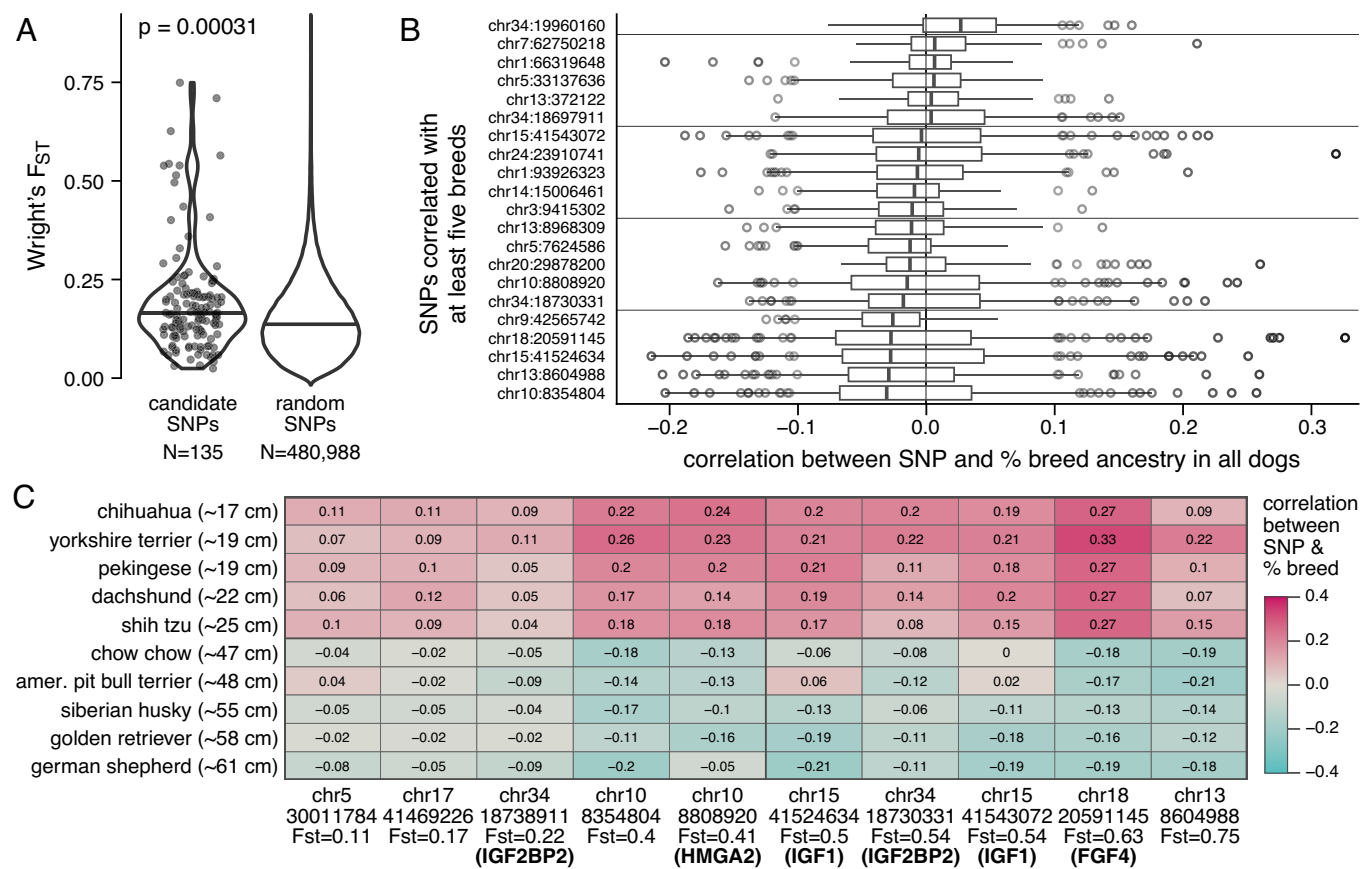


Fig. 4. Evidence of population stratification at candidate behavior SNPs. (A) Candidate SNPs have significantly higher differentiation between breeds compared to a set of random SNPs. Violin plots show the density of F_{ST} values, with horizontal lines indicating the median. Jittered points are overlaid for candidate SNPs. Significance measured with a one-sided Wilcoxon rank-sum test. (B) SNP-level correlations between genotype and breed ancestry for candidate SNPs correlated with ancestry in at least 5 breeds ($|r| > 0$). Each boxplot summarizes correlations across breeds for one SNP (gray boxes) and shows the median (horizontal line), interquartile range (box), and whiskers extending to the smallest and largest values within $1.5 \times$ IQR of the lower and upper quartiles. Breeds beyond the whiskers are plotted as outliers. (C) Heatmap of the 10 SNPs most strongly associated with breed height ($|r|$ with height > 0.3 , $P < 0.05$) and their correlations with the ancestry proportion of 10 dog breeds spanning the height spectrum. Fill color and overlaid text indicate the Pearson correlation between SNP genotype and breed ancestry. SNPs are ordered by correlation with breed height; breeds are ordered by height. Labels include breed name and approximate median height, as well as gene names for SNPs near known size-associated loci (e.g., IGF1, IGF2BP2).

prediction analysis. For the analysis of body size traits (height and leg length), the corresponding trait was excluded from the covariate set. All dogs were included in the analysis ($N = 3,277$), and P -values were adjusted for multiple testing using the Benjamini–Hochberg procedure to control the false discovery rate (FDR).

We validated our modeling approach by testing four SNPs previously implicated in aesthetic traits, each of which was also significantly associated with the same trait in our GWAS. All four variants showed strong predictive effects on their corresponding traits even after adjusting for covariates ($p_{\text{FDR}} = 2.4 \times 10^{-55}$, 3.1×10^{-27} , 2.8×10^{-17} , and 2.7×10^{-11}) (Fig. 5A). When we extended this analysis to all SNP/trait pairs, SNPs previously associated with specific aesthetic traits were significantly more predictive of those same traits ($N = 19$; median $p_{\text{FDR}} = 0.031$; range = 2.4×10^{-55} – 0.93) than of other aesthetic traits ($N = 261$; median $p_{\text{FDR}} = 0.75$; range = 4.8×10^{-7} – 1.0) (Fig. 5B).

By contrast, candidate behavior-associated SNPs showed little such specificity. After adjusting for covariates, they were similarly predictive of the previously associated behavioral trait ($N = 481$; median $p_{\text{FDR}} = 0.82$; range = 0.064 to 1.0) and of other behavioral

traits ($N = 2,383$; median $p_{\text{FDR}} = 0.87$; range = 0.031 to 1.0) (Fig. 5B). Only one SNP (chr14:15006461) showed a marginally significant effect on a behavioral trait (“Easy to train”; $p_{\text{FDR}} = 0.03$), and this was not the same trait reported in earlier studies.

The strong predictive effects of candidate SNPs on behavioral traits reported in previous studies are likely due to confounding from population stratification and artificial selection on aesthetic traits. When body size and breed are not included in our model, many candidate SNPs appear to replicate previous associations: 18 SNPs showed a significant predictive effect on at least one previously reported behavioral trait (Fig. 5C), and 12 SNPs on a reported aesthetic trait (Fig. 5D). However, after adjusting for both body size and breed ancestry, none of the SNPs remained significantly predictive of behavior. In contrast, 75% aesthetic trait associations persisted, with 9 of the 12 SNP–trait pairs still significant (Fig. 5D).

Discussion

We find no evidence that genetic variants identified through breed-average studies have utility for predicting behavior. Instead,

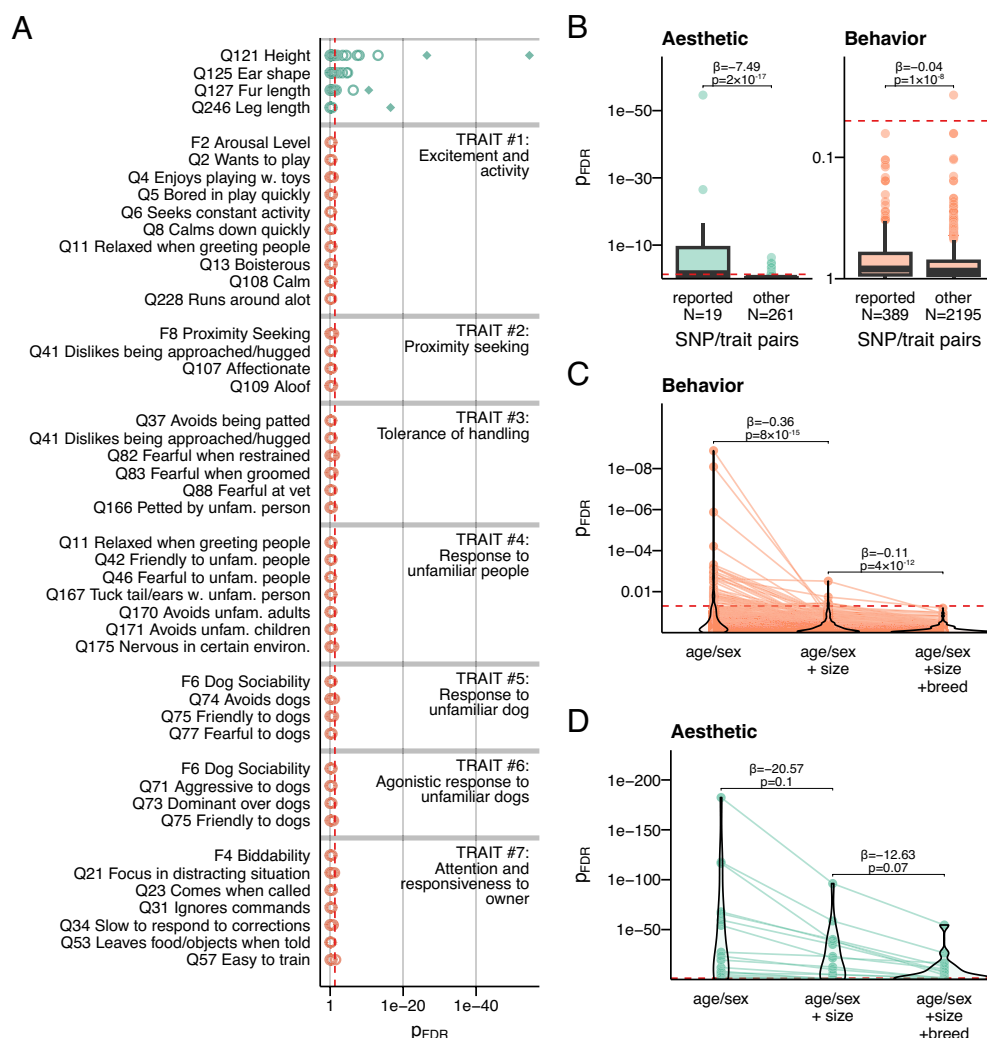


Fig. 5. Predictive power of candidate behavior SNPs diminishes after accounting for size and breed ancestry. (A) Adjusted P -values ($-\log_{10} p_{\text{FDR}}$) from predictive models for both aesthetic (Top) and behavioral (Bottom) traits. Each point represents one SNP–trait pair. Some of the SNPs improved predictions of aesthetic traits (green) but not of behavioral traits (orange). Filled diamonds are the four validation SNPs, and circles are all other SNPs (B) Boxplots comparing the predictive power of SNPs previously associated with specific traits versus all other SNP/trait pairs, shown separately for aesthetic (Left) and behavioral (Right) traits. (C) Violin plots showing the distribution of $-\log_{10}$ -transformed P -values across increasingly stringent covariate models for behavioral traits and (D) aesthetic traits. Lines connect values for the same SNP/trait pair across models. In all panels, the dashed red line indicates the significance threshold. In panels B–D, differences between groups were assessed using a linear mixed-effects model of $-\log_{10}$ -transformed p_{FDR} values, with SNP included as a random intercept to account for nonindependence. Reported values represent the estimated fixed effect (β) and associated P -value.

we find that the apparent associations were likely driven by population structure and artificial selection on aesthetic traits such as body size, leg length, and ear shape. Thus, behavioral predictions based on these SNPs could reinforce misleading and potentially harmful breed stereotypes (93, 94). Our findings demonstrate a fundamental flaw in breed-average study design: Averaging phenotypes by breed and then testing for genetic predictors is inherently circular. Furthermore, breed-level comparisons lack the within-breed variance necessary to distinguish causal effects from spurious correlations.

Our results also underscore the broader challenges of developing genetic predictions for behavioral traits in dogs. While our sample sizes (~1,000 to ~3,000 dogs per trait) are smaller than typical human GWAS, they remain among the largest available for pet dogs. We show that behavioral traits consistently have lower heritability and fewer significant associations compared to aesthetic traits, highlighting behavior's polygenic, environmentally influenced nature. Thus, much larger sample sizes are needed to detect robust associations that explain a greater proportion of trait variation (95).

Behavioral traits in dogs cannot be meaningfully predicted using a handful of candidate SNPs. Instead, more sophisticated approaches developed for complex traits are needed. Polygenic risk scores (PRS), now widely used in human medicine (26), and genomic estimated breeding values (GEBVs), established in animal breeding (96), are promising options. Both methods account for polygenicity and population structure when estimating and incorporating SNP effects across the whole genome, resulting in greater predictive power (4).

Dogs present unique challenges. The population structure of pet dogs is unlike that of any other species commonly used in genetic studies. In the Darwin's Ark dataset alone, at least 85 breeds contribute detectable ancestry and approximately 40% of dogs carry ancestry from at least four breeds. Each breed reflects a history of differing origins, population bottlenecks, strong selection for various aesthetic traits, and ongoing inbreeding (97, 98). Some breeds, like the basenji, were derived from larger free-living dog populations (99), while others, like the great Pyrenees, trace their origins to landrace breeds shaped without human-directed breeding (100, 101). Many modern breeds were created by mixing older breeds, often leading to unexpected relatedness (11, 102). For example, the golden retriever was established in the early 1,900s by Lord Tweedmouth who mixed Labrador retrievers, Irish setters, bloodhounds, and now-extinct breeds (103). Free-breeding village dogs without evident breed ancestry make up ~80% of the global dog population (104) and are also represented in the US pet population (16).

This complex structure introduces a major confounding factor in genetic studies of polygenic traits (15, 84, 105). Prediction accuracy is sensitive to ancestry and often decreases substantially when applied to individuals from underrepresented genetic backgrounds (91, 92). For working dog programs, where a few breeds dominate and breeding is controlled, genomic prediction with GEBVs can help select for complex traits such as behavior and health, accelerating the production of successful assistance and service dogs (93, 94). In contrast, accurately predicting trait or disease risk in pet dogs is likely to prove more challenging.

Meaningful, accurate genetic predictions for complex traits could transform veterinary medicine by identifying individuals at high risk of developing diseases like cancer (106) and helping dog breeders improve population health (96). To build accurate and generalizable predictors, large and diverse cohorts are essential. Efforts such as Darwin's Ark (16), Dog10K (86), and the Dog Aging Project (18) are sequencing thousands of dogs. Expanding

these efforts is essential for developing effective genomic tools that can enhance the health and welfare of all dogs.

Materials and Methods

Phenotypes and Genetic Data for GWAS. *Darwin's Ark.* Data were collected through Darwin's Ark, an open-access resource for collecting owner-reported phenotypes and genetic data in dogs, as previously described (16). Briefly, dog owners were invited to participate in an online community science project (www.darwinsark.org), where they completed profile information and behavioral and physical trait surveys about their dogs.

Age. Multiple fields related to age were available across different phases of the project. For each dog, we computed a composed age value using the following logic: 1) When available, we used the auto-generated "age_norm" field, which calculates age from the dog's date of birth; 2) If "age_norm" was unavailable (in earlier-enrolled dogs), we estimated age using the profile creation date; 3) If neither was available, we used the legacy free-text field "age," remapping month- and week-based entries into numerical values. After applying this logic, 3,082 dogs had usable age values.

Body size. Body height was assessed using a categorical survey question (question #121; *SI Appendix, Fig. S2*), in which owners indicated whether their dog was ankle-high, calf-high, knee-high, thigh-high, or hip-high. A total of 285 dogs did not have responses to this question. For analyses requiring height information, we first calculated the average height for each breed using genotyped dogs with responses to question #121. We then imputed height for the dogs without height phenotypes based on the breed contributing the highest proportion to their inferred genetic ancestry. One dog lacked both ancestry information and a height response and was excluded from height-based analyses.

Survey data analysis. To prepare the data for input to GWAS, we scaled the survey questions; factor scores were already normalized. We calculated factor scores by applying the factor pattern loadings derived from our previous varimax rotation factor analysis (16), to the survey responses to generate a factor score for our updated set of dogs. The quantitative covariates used in the GWAS and heritability calculations are age and height, while the discrete covariates are sex and neuter status. For the height GWAS, the only quantitative covariate used was age.

Genetic Data. Dogs were sequenced using Illumina NovaSeq 6000 with low-coverage sequencing with a mean coverage of ~0.82X. A total of 3,285 FASTQ files (available on SRA under PRJNA675863) were aligned to the canFam4 (UU_Cfam_GSD_1.0 + ROSY) genome assembly using NVIDIA Clara Parabricks (version 4.0) fq2bam wrapper for BWA-MEM on the Sol supercomputer of Arizona State University (107). The parameter `--bwa-options="-K 100000000 -Y"` was included in the pbrun fq2bam command as recommended by the Dog10K consortium's genome alignment code (<https://github.com/jmkidd/dogmap>). For dogs with multiple sequencing runs, sequence data were merged to improve coverage, with the exception of dogs 9881 and 5954, where higher-coverage (4×) data were retained and lower-coverage (0.2×) data were excluded. We excluded 16 dogs with mean sequencing depth < 0.1x. After filtering, 3,277 samples remained.

We performed genotype imputation with GLIMPSE2 (108). Briefly, the Dog10K phased imputation reference panel of 1,929 dogs was used to create chunks and binary reference panels for each autosome (chromosomes 1 to 38). The 3,277 binary alignment map files (BAMs) were provided via a list to GLIMPSE2_phase for phasing and imputation. The resulting individual binary variant call format files (BCFs) per chunk were ligated together into a single BCF file containing the imputed genotype calls for all dogs.

Heritability and GWAS. Workflow. We developed a reproducible and scalable Nextflow workflow for heritability estimation and GWAS using PLINK (v1.90b6.21 and v2.00a5LM) (92) and Genome-wide Complex Trait Analysis (GCTA v1.94.1) (109). This framework, packaged in a Docker container image, enables consistent execution across computing environments and parallelization across multiple phenotypes. The workflow follows the methods we used previously (16) and includes quality control filtering, estimation of SNP-based heritability using restricted maximum likelihood (REML), and GWAS using a mixed linear model with a leave-one-chromosome-out (LOCO) approach. We note that previous work has proposed that GCTA-based heritability estimates are inflated by population stratification (109), but this is disputed by the software's authors (110).

Quality control. After filtering using parameters defined in the Nextflow workflow script, the dataset included genotypes for 10,113,058 autosomal biallelic SNPs in 3,277 dogs. SNPs were excluded if they had a minor allele frequency (MAF) below 0.01, were missing in more than 5% of individuals, or showed extreme deviation from Hardy-Weinberg equilibrium ($P < 1 \times 10^{-20}$ in the exact test with mid-p adjustment and excess heterozygosity; --hwe 1e-20 midp keep-fewhet).

Heritability and GWAS. For LD-adjusted heritability estimation, we calculated linkage disequilibrium (LD) scores in 250-kb windows (params.ld_window_kb = 250) using a block size of 10,000 kb with 5,000 kb overlap between blocks. These LD scores were then used to construct four LD-stratified GRMs for use in GCTA's GREML-LDMS (LD score-stratified genomic REML) heritability analysis (111). We performed GWAS on autosomal chromosomes using the MLMA-LOCO method in GCTA (112). To define the associated regions, we used linkage-disequilibrium-based clumping. We defined clumps as index SNPs (SNPs with P -values $< 1 \times 10^{-6}$) and any other SNPs in linkage disequilibrium ($r^2 > 0.2$) and within 250 kb of an index SNP.

Ancestry estimation and population differentiation. Breed ancestry estimates for each dog were generated using ADMIXTURE (113), as previously described (16). Our dataset of 2,454 mixed-breed dogs excludes 504 dogs whose owners self-reported that they were "purebred" and 304 dogs with over 85% ancestry from a single breed. Our dataset of single-breed dogs included all dogs excluded from the mixed-breed set. We validated these thresholds previously (16). For all single-breed dogs from the ten breeds with ≥ 20 dogs (297 dogs total), we computed F_{ST} using plink1.9 for the 135 candidate SNPs and 480,988 random sampled SNPs. (SI Appendix, Table S2). The 10 breeds were Labrador retriever, golden retriever, American pit bull terrier, border collie, miniature schnauzer, Australian shepherd, dachshund, Yorkshire Terrier, German shepherd, and boxer.

Genotype correlation with breed. For each candidate SNP, we computed the Spearman correlation coefficient with breed ancestry in python using the pandas library corr method (114) for all 101 breeds (Dataset S4). We used breed ancestry estimates for the five most highly correlated breed ancestries in our prediction analysis as covariates.

Effect on Prediction. For each of the candidate SNPs in the prediction analysis (SI Appendix, Table S2), we tested whether the genetic variant had an effect on each trait using a likelihood ratio test and the imputed genotype dosages for each SNP for each dog. We used ordinary least squares regression, implemented in the python package statsmodels in statsmodels.formula.api.ols, to fit a model predicting the phenotype from factors. We tested for significance within both the dataset of all dogs, as well as within just the mixed-breed dog dataset (Dataset S2).

First, we fit a "full" model, predicting the phenotype using the SNP of interest, in addition to covariates, including age, sex, spay/neuter status, height, leg length (Q246), and the top five breed ancestries for the SNP of interest. For testing the effect on phenotypes of height or leg length, we excluded the corresponding phenotype from the covariates used in prediction. We then fit a "reduced" model using the same covariates but without the SNP of interest. When using height as a covariate, we included the imputed height phenotypes to maximize our sample size. However, when testing for an effect on height predictions, we used only the reported height phenotypes. We excluded from the behavior prediction analysis two SNPs that were significantly associated in our GWAS with aesthetic traits not included as covariates in the full model. SNP chr13:8604988 was associated with "Q127 Fur length" ($P = 6.8 \times 10^{-11}$) and "Q128 Fur texture" ($P = 3.6 \times 10^{-8}$) and SNP chr24:23910741 was associated with "Q243 Coat color" ($P = 5.8 \times 10^{-25}$) (Dataset S3).

To determine significance of the SNP in predicting the phenotype, we used a likelihood ratio statistic of $-2 \times (\text{reduced_ll} - \text{full_ll})$, where reduced_ll and full_ll correspond to the log likelihood of each corresponding model. We determined significance of the likelihood ratio statistic by computing a P -value using the chi squared (scipy.statst.chi2.sf) with one degree of freedom.

Data, Materials, and Software Availability. Genetic analysis files; FASTQ files for all genetic information; code for curating the phenotypes and predicting SNP effects on phenotype; pipeline for GWAS and heritability; software packaged in a Docker container data have been deposited in Dryad: SRA; GitHub; Docker Hub [https://doi.org/10.5061/dryad.83bk3jb4r (44); PRJNA675863 (115); https://github.com/broadinstitute/dog_behavioral_gwas_paper (116); https://github.com/VistaSohrab/dog-gwas-heritability-nextflow/ (117); https://hub.docker.com/r/vistasohrab/terra-dogagingproject-gwas (118)]. All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank the Darwin's Ark community scientists and their dogs. We thank D. Genereux, K. Tengvall, J. Meadows, and K. Lindblad-Toh for analysis support and N. Snyder-Mackler and Research Computing at Arizona State University for computational support. This study was funded by NIH, including NCI 5R37CA218570, NCI R01CA255319, and NHGRI R56HG013639, Morris Animal Foundation DCA24-527, Hill's Pet Nutrition, and Mars Petcare.

Author affiliations: ^aGenomics and Computational Biology, UMass Chan Medical School, Worcester, MA 01655; ^bMedical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, MA 02142; ^cMorningside Graduate School of Biomedical Sciences, UMass Chan Medical School, Worcester, MA 01655; ^dDarwin's Ark Foundation, Seattle, WA 98026; and ^eProgram in Molecular Medicine, UMass Chan Medical School, Worcester, MA 01655

1. A. Abdellaoui, L. Yengo, K. J. H. Verweij, P. M. Visscher, 15 years of GWAS discovery: Realizing the promise. *Am. J. Hum. Genet.* **110**, 179–194 (2023).
2. L. Yengo *et al.*, A saturated map of common genetic variants associated with human height. *Nature* **610**, 704–712 (2022).
3. Y. Wang, K. Tsuo, M. Kanai, B. M. Neale, A. R. Martin, Challenges and opportunities for developing more generalizable polygenic risk scores. *Annu. Rev. Biomed. Data Sci.* **5**, 293–320 (2022).
4. N. R. Wray, K. E. Kemper, B. J. Hayes, M. E. Goddard, P. M. Visscher, Complex trait prediction from genome data: Contrasting EBV in livestock to PRS in humans: Genomic prediction. *Genetics* **211**, 1131–1141 (2019).
5. H. Ritvo, Pride and pedigree: The evolution of the Victorian dog fancy. *Vict. Stud.* **29**, 227–253 (1986).
6. H. G. Parker *et al.*, Genetic structure of the purebred domestic dog. *Science* **304**, 1160–1164 (2004).
7. A. R. Boyko *et al.*, 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 (2009).
8. H. G. Parker *et al.*, Breed relationships facilitate fine-mapping studies: A 7.8-kb deletion cosegregates with Collie eye anomaly across multiple dog breeds. *Genome Res.* **17**, 1562–1571 (2007).
9. B. M. Vonholdt *et al.*, Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **464**, 898–902 (2010).
10. A. H. Freedman *et al.*, Genome sequencing highlights the dynamic early history of dogs. *PLoS Genet.* **10**, e1004016 (2014).
11. E. V. Dutrow, J. A. Serpell, E. A. Ostrander, Domestic dog lineages reveal genetic drivers of behavioral diversification. *Cell* **185**, 4737–4755.e18 (2022).
12. E. K. Karlsson, K. Lindblad-Toh, Leader of the pack: Gene mapping in dogs and other model organisms. *Nat. Rev. Genet.* **9**, 713–725 (2008).
13. E. A. Ostrander, L. Kruglyak, Unleashing the canine genome. *Genome Res.* **10**, 1271–1274 (2000).
14. A. Helgason, B. Yngvadottir, B. Hrafnkelsson, J. Gulcher, K. Stefansson, An Icelandic example of the impact of population structure on association studies. *Nat. Genet.* **37**, 90–95 (2005).
15. J. H. Sul, L. S. Martin, E. Eskin, Population structure in genetic studies: Confounding factors and mixed models. *PLoS Genet.* **14**, e1007309 (2018).
16. K. Morrill *et al.*, Ancestry-inclusive dog genomics challenges popular breed stereotypes. *Science* **376**, eabk0639 (2022).
17. K. L. Moon *et al.*, Comparative genomics of Balto, a famous historic dog, captures lost diversity of 1920s sled dogs. *Science* **380**, eabn5887 (2023).
18. K. E. Creevy, J. M. Akey, M. Kaeberlein, D. E. L. Promislow, Dog aging project consortium, an open science study of ageing in companion dogs. *Nature* **602**, 51–57 (2022).
19. F. Dudbridge, Power and predictive accuracy of polygenic risk scores. *PLoS Genet.* **9**, e1003348 (2013).
20. N. Bhowmik *et al.*, Heritability and genome-wide association study of dog behavioral phenotypes in a commercial breeding cohort. *Genes (Basel)* **15**, 1611 (2024).
21. J. Ilkka *et al.*, Genetic characterization of dog personality traits. *Genetics* **206**, 1101–1111 (2017).
22. J. Friedrich *et al.*, Genetic dissection of complex behaviour traits in German Shepherd dogs. *Heredity (Edinb.)* **123**, 746–758 (2019).
23. A. R. Martin *et al.*, Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).
24. N. J. Lennon *et al.*, Selection, optimization and validation of ten chronic disease polygenic risk scores for clinical implementation in diverse US populations. *Nat. Med.* **30**, 480–487 (2024).
25. J. B. Nielsen *et al.*, Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat. Genet.* **50**, 1234–1239 (2018).
26. A. Torkamani, N. E. Wineinger, E. J. Topol, The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* **19**, 581–590 (2018).
27. I. Zapata, J. A. Serpell, C. E. Alvarez, Genetic mapping of canine fear and aggression. *BMC Genomics* **17**, 572 (2016).
28. I. Zapata, M. L. Lilly, M. E. Herron, J. A. Serpell, C. E. Alvarez, Genetic testing of dogs predicts problem behaviors in clinical and nonclinical samples. *BMC Genomics* **23**, 102 (2022).

29. I. Zapata, E. E. Hecht, J. A. Serpell, C. E. Alvarez, Genome scans of dog behavior implicate a gene network underlying psychopathology in mammals, including humans. *bioRxiv* [Preprint] (2020), <https://doi.org/10.1101/2020.07.19.211078> (Accessed 12 February 2024).
30. K. Chase, P. Jones, A. Martin, E. A. Ostrander, K. G. Lark, Genetic mapping of fixed phenotypes: Disease frequency as a breed characteristic. *J. Hered.* **100**, S37–S41 (2009).
31. P. Jones *et al.*, Single-nucleotide-polymorphism-based association mapping of dog stereotypes. *Genetics* **179**, 1033–1044 (2008).
32. D. A. Akkad, W. M. Gerding, R. B. Gasser, J. T. Epplen, Homozygosity mapping and sequencing identify two genes that might contribute to pointing behavior in hunting dogs. *Canine Genet. Epidemiol.* **2**, 5 (2015).
33. G. E. Gnanadesikan *et al.*, Breed differences in dog cognition associated with brain-expressed genes and neurological functions. *Integr. Comp. Biol.* **60**, 976–990 (2020).
34. E. L. MacLean, N. Snyder-Mackler, B. M. vonHoldt, J. A. Serpell, Highly heritable and functionally relevant breed differences in dog behaviour. *Proc. Biol. Sci.* **286**, 20190716 (2019).
35. A. R. Boyko *et al.*, A simple genetic architecture underlies morphological variation in dogs. *PLoS Biol.* **8**, e1000451 (2010).
36. L. R. Mehrkam, C. D. L. Wynne, Behavioral differences among breeds of domestic dogs (*Canis lupus familiaris*): Current status of the science. *Appl. Anim. Behav. Sci.* **155**, 12–27 (2014).
37. K. Svartberg, Breed-typical behaviour in dogs—Historical remnants or recent constructs? *Appl. Anim. Behav. Sci.* **96**, 293–313 (2006).
38. A. Hammond, T. Rowland, D. S. Mills, M. Pilot, Comparison of behavioural tendencies between “dangerous dogs” and other domestic dog breeds—Evolutionary context and practical implications. *Evol. Appl.* **15**, 1806–1819 (2022).
39. J. P. Scott, J. L. Fuller, *Genetics and the Social Behavior of the Dog* (University of Chicago Press, 1965).
40. Y. Matsumoto, A. Konno, G. Ishihara, M. Inoue-Murayama, Genetic dissection of behavioral traits related to successful training of drug detection dogs. *Sci. Rep.* **13**, 7326 (2023).
41. M. E. Persson, D. Wright, L. S. V. Roth, P. Batakis, P. Jensen, Genomic regions associated with interspecies communication in dogs contain genes related to human social disorders. *Sci. Rep.* **6**, 33439 (2016).
42. Ancestry® Launches Know Your Pet DNA. Ancestry.com (2023), <https://web.archive.org/web/20240603025415/https://www.ancestry.com/corporate/newsroom/press-releases/ancestry-launches-know-your-pet-dna> [Accessed 2 June 2024].
43. N. E. Bennett, S. E. Mirabal Torres, P. B. Gray, Exploratory content analysis of direct-to-consumer pet genomics: What is being marketed and what are consumers saying? *PLoS One* **17**, e0261694 (2022).
44. E. Karlsson, V. Sohrab, K. Lord, Data from “Genetic testing predicts looks but not behavior in dogs”. Dryad. <https://doi.org/10.5061/dryad.83bk3jb4r>. Deposited 12 May 2025.
45. S. Shan, F. Xu, B. Brenig, Genome-wide association studies reveal neurological genes for dog herding, predation, temperament, and trainability traits. *Front. Vet. Sci.* **8**, 693290 (2021).
46. W. S. Robinson, Ecological correlations and the behavior of individuals. *Am. Sociol. Rev.* **15**, 351 (1950).
47. C. Tian, P. K. Gregersen, M. F. Seldin, Accounting for ancestry: Population substructure and genome-wide association studies. *Hum. Mol. Genet.* **17**, R143–R150 (2008).
48. M. L. Chang *et al.*, Inbreed stratification related to divergent selection regimes in purebred dogs may affect the interpretation of genetic association studies. *J. Hered.* **100**, S28–S36 (2009).
49. K. Lindblad-Toh *et al.*, Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819 (2005).
50. J. A. Serpell, BSC, C-BARQ: Canine behavioral assessment and research questionnaire. (2024), <https://vetapps.vet.upenn.edu/cbarq/> [Accessed 4 December 2024].
51. J. A. Serpell, D. L. Duffy, “Dog Breeds and their Behavior” in *Domestic Dog Cognition and Behavior* (Springer, Berlin Heidelberg, 2014), pp. 31–57.
52. J. M. Akey *et al.*, Tracking footprints of artificial selection in the dog genome. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 1160–1165 (2010).
53. D. L. Bannasch *et al.*, Dog colour patterns explained by modular promoters of ancient canid origin. *Nat. Ecol. Evol.* **5**, 1415–1423 (2021).
54. A. Bauer, A. Kehl, V. Jagannathan, T. Leeb, A novel MLPH variant in dogs with coat colour dilution. *Anim. Genet.* **49**, 94–97 (2018).
55. C. Dierks, S. Mömke, U. Philipp, O. Distl, Allelic heterogeneity of *FGF5* mutations causes the long-hair phenotype in dogs. *Anim. Genet.* **44**, 425–431 (2013).
56. E. A. Brown *et al.*, FGF4 retrogene on CFA12 is responsible for chondrodysplasia and intervertebral disc disease in dogs. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11476–11481 (2017).
57. E. Cadieu *et al.*, Coat variation in the domestic dog is governed by variants in three genes. *Science* **326**, 150–153 (2009).
58. S. I. Candille *et al.*, A -defensin mutation causes black coat color in domestic dogs. *Science* **318**, 1418–1423 (2007).
59. D. L. Dreger, S. M. Schmutz, A SINE insertion causes the black-and-tan and saddle tan phenotypes in domestic dogs. *J. Hered.* **102**, S11–S18 (2011).
60. C. Drögemüller, U. Philipp, B. Haase, A.-R. Günzel-Apel, T. Leeb, A noncoding melanophilin gene (MLPH) SNP at the splice donor of exon 1 represents a candidate causal mutation for coat color dilution in dogs. *J. Hered.* **98**, 468–473 (2007).
61. C. Drögemüller *et al.*, A mutation in hairless dogs implicates *FOXI3* in ectodermal development. *Science* **321**, 1462 (2008).
62. N. Dürig *et al.*, Two MC1R loss-of-function alleles in cream-coloured Australian Cattle Dogs and white Huskies. *Anim. Genet.* **49**, 284–290 (2018).
63. J. J. Hayward *et al.*, Complex disease and phenotype mapping in the domestic dog. *Nat. Commun.* **7**, 10460 (2016).
64. L. Honkanen, R. Loechel, S. Davison, J. Donner, H. Anderson, Canine coat color E locus updates: Identification of a new MC1R variant causing “sable” coat color in English Cocker Spaniels and a proposed update to the E locus dominance hierarchy. *Anim. Genet.* **55**, 291–295 (2024).
65. B. C. Hoopes, M. Rimbault, D. Liebers, E. A. Ostrander, N. B. Sutter, The insulin-like growth factor 1 receptor (IGF1R) contributes to reduced size in dogs. *Mamm. Genome* **23**, 780–790 (2012).
66. D. J. E. Housley, P. J. Venta, The long and the short of it: Evidence that FGF5 is a major determinant of canine “hair”-itability. *Anim. Genet.* **37**, 309–315 (2006).
67. E. K. Karlsson *et al.*, Efficient mapping of Mendelian traits in dogs through genome-wide association. *Nat. Genet.* **39**, 1321–1328 (2007).
68. J. A. Kerns *et al.*, Characterization of the dog Agouti gene and a nonagouti mutation in German Shepherd Dogs. *Mamm. Genome* **15**, 798–808 (2004).
69. I. Baranowska Körberg *et al.*, A simple repeat polymorphism in the MITF-M promoter is a key regulator of white spotting in dogs. *PLoS One* **9**, e104363 (2014).
70. J. M. Newton *et al.*, Melanocortin 1 receptor variation in the domestic dog. *Mamm. Genome* **11**, 24–30 (2000).
71. M. Ollivier *et al.*, Evidence of coat color variation sheds new light on ancient canids. *PLoS One* **8**, e75110 (2013).
72. H. G. Parker *et al.*, An expressed fgf4 retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *Science* **325**, 995–998 (2009).
73. H. G. Parker, K. Chase, E. Cadieu, K. G. Lark, E. A. Ostrander, An insertion in the RSPO2 gene correlates with improper coat in the Portuguese water dog. *J. Hered.* **101**, 612–617 (2010).
74. J. Plassais *et al.*, Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nat. Commun.* **10**, 1489 (2019).
75. M. Rimbault *et al.*, Derived variants at six genes explain nearly half of size reduction in dog breeds. *Genome Res.* **23**, 1985–1995 (2013).
76. J. J. Schoenebeck *et al.*, Variation of BMP3 contributes to dog breed skull diversity. *PLoS Genet.* **8**, e1002849 (2012).
77. N. B. Sutter *et al.*, A single IGF1 allele is a major determinant of small size in dogs. *Science* **316**, 112–115 (2007).
78. E. H. Turnova *et al.*, A novel mutation in the TYRP1 gene associated with brown coat colour in the Australian Shepherd Dog Breed. *Animal Genetics* **48**, 626–627 (2017).
79. S. L. Van Buren *et al.*, A third MLPH variant causing coat color dilution in dogs. *Genes (Basel)* **11**, 639 (2020).
80. A. Vaysse *et al.*, Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genet.* **7**, e1002316 (2011).
81. H. Wright, E. Schofield, C. Mellersh, L. Burnmeister, A novel TYRP1 variant is associated with liver and tan coat colour in Lancashire Heelers. *Anim. Genet.* **50**, 783–783 (2019).
82. J. Plassais *et al.*, Analysis of large versus small dogs reveals three genes on the canine X chromosome associated with body weight, muscling and back fat thickness. *PLoS Genet.* **13**, e1006661 (2017).
83. S. M. Schmutz, T. G. Berryere, A. D. Goldfinch, TYRP1 and MC1R genotypes and their effects on coat color in dogs. *Mamm. Genome* **13**, 380–387 (2002).
84. J. J. Berg *et al.*, Reduced signal for polygenic adaptation of height in UK biobank. *Elife* **8**, e39725 (2019).
85. A. Korte, A. Farlow, The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* **9**, 29 (2013).
86. J. R. S. Meadows *et al.*, Genome sequencing of 2000 canids by the Dog10K consortium advances the understanding of demography, genome function and architecture. *Genome Biol.* **24**, 187 (2023).
87. G. Kichaev *et al.*, Leveraging polygenic functional enrichment to improve GWAS power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).
88. B. Hédan *et al.*, Identification of a missense variant in MFSD12 involved in dilution of pheomelanin leading to white or cream coat color in dogs. *Genes (Basel)* **10**, 386 (2019).
89. N. G. Crawford *et al.*, Loci associated with skin pigmentation identified in African populations. *Science* **358**, eaan8433 (2017).
90. M. Lundberg *et al.*, Novel Alzheimer’s disease genes and epistasis identified using machine learning GWAS platform. *Sci. Rep.* **13**, 17662 (2023).
91. A. L. Price, N. A. Zaitlen, D. Reich, N. Patterson, New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* **11**, 459–463 (2010).
92. C. C. Chang *et al.*, Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
93. L. M. Gunter, R. T. Barber, C. D. L. Wynne, What’s in a name? Effect of breed perceptions & labeling on attractiveness, adoptions & length of stay for pit-bull-type dogs. *PLoS One* **11**, e0146857 (2016).
94. B. Dickey, *Pit Bull* (Vintage Books, 2017).
95. P. M. Visscher *et al.*, 10 years of GWAS discovery: Biology, function, and translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
96. M. E. Goddard, K. E. Kemper, I. M. MacLeod, A. J. Chamberlain, B. J. Hayes, Genetics of complex traits: Prediction of phenotype, identification of causal polymorphisms and genetic architecture. *Proc. Biol. Sci.* **283**, 20160569 (2016).
97. A. J. Sams, A. R. Boyko, Fine-scale resolution of runs of homozygosity reveal patterns of inbreeding and substantial overlap with recessive disease genotypes in domestic dogs. *G3 Genes/Genomes/Genetics* **9**, 117–123 (2019).
98. T. W. Lewis, B. M. Abhayaratne, S. C. Blott, Trends in genetic diversity for all Kennel Club registered pedigree dog breeds. *Canine Genet. Epidemiol.* **2**, 13 (2015).
99. R. J. Edwards *et al.*, Chromosome-length genome assembly and structural variations of the primal Basenji dog (*Canis lupus familiaris*) genome. *BMC Genomics* **22**, 188 (2021).
100. K. Lord, L. Coppinger, R. Coppinger, “Differences in the behavior of landraces and breeds of dogs” in *Genetics and the Behavior of Domestic Animals*, (Elsevier, 2014), pp. 195–235.
101. D. Coutinho-Lima *et al.*, Multiple ancestries and shared gene flow among modern livestock guarding dogs. *iScience* **27**, 110396 (2024).
102. E. A. Ostrander, R. K. Wayne, A. H. Freedman, B. W. Davis, Demographic history, selection and functional diversity of the canine genome. *Nat. Rev. Genet.* **18**, 705–720 (2017).
103. J. G. Pepper, *The Golden Retriever: An Authoritative Look at the Breed’s Past, Present, and Future* (Kennel Club Books, 2012).
104. K. Lord, M. Feinstein, B. Smith, R. Coppinger, Variation in reproductive traits of members of the genus *Canis* with special attention to the domestic dog (*Canis familiaris*). *Behav. Processes* **92**, 131–142 (2013).
105. E. S. Lander, N. J. Schork, Genetic dissection of complex traits. *Science* **265**, 2037–2048 (1994).
106. A. V. Khera *et al.*, Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).
107. D. M. Jennewein *et al.*, *The Sol Supercomputer at Arizona State University in Practice and Experience in Advanced Research Computing* (ACM, 2023).
108. S. Rubinacci, R. J. Hofmeister, B. Sousa da Mota, O. Delaneau, Imputation of low-coverage sequencing data from 150,119 UK Biobank genomes. *Nat. Genet.* **55**, 1088–1090 (2023).
109. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).

110. J. Yang, S. H. Lee, N. R. Wray, M. E. Goddard, P. M. Visscher, GCTA-GREML accounts for linkage disequilibrium when estimating genetic variance from genome-wide SNPs. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E4579–80 (2016).
111. J. Yang, J. Zeng, M. E. Goddard, N. R. Wray, P. M. Visscher, Concepts, estimation and interpretation of SNP-based heritability. *Nat. Genet.* **49**, 1304–1310 (2017).
112. J. Yang, N. A. Zaitlen, M. E. Goddard, P. M. Visscher, A. L. Price, Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* **46**, 100–106 (2014).
113. D. H. Alexander, K. Lange, Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**, 246 (2011).
114. W. McKinney, "Data Structures for Statistical Computing in Python" in *Proceedings of the Python in Science Conference*, (SciPy, 2010), pp. 56–61.
115. K. Bryc *et al.*, Low coverage whole genome sequencing of pet dog samples collected by the Darwin's Ark project. NCBI. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA675863/>. Deposited 23 August 2024.
116. K. Bryc, dog_behavioral_gwas_paper. Github. https://github.com/broadinstitute/dog_behavioral_gwas_paper. Accessed 8 June 2025.
117. V. Sohrab, VistaSohrab/dog-gwas-heritability-nextflow: dog-gwas-heritability-nextflow-v1.0.0 (v1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.15392359>. Deposited 12 May 2025.
118. V. Sohrab, GWAS software tools dependencies, 2023, Docker Hub, v1.0. <https://hub.docker.com/r/vistasohrab/terra-dogagingproject-gwas>. Accessed 20 October 2024.