Journal of Animal & Plant Sciences, 35(2): 2025, Page: 546-560
ISSN (print): 1018-7081; ISSN (online): 2309-8694
https://doi.org/10.36899/JAPS.2025.2.0045

GENOME-WIDE ASSOCIATION STUDY IDENTIFIED QTLS FOR BEHAVIORAL TRAITS IN KOREAN SAPSAREE DOGS

Md Azizul Haque[1,*], Na-Kuang Kim[1], Ryu Yeji[1], B. Lee[2], Ji-Hong Ha[2], Yun-Mi Lee[3], Han-Kook Il[4], and Jong-Joo Kim[1, *]

[1]Department of Biotechnology, Yeungnam University, Gyeongbuk 38541, Republic of Korea
[2]Sapsaree Breeding Research Institute, Gyeongsan, Gyeongbuk 38412, Republic of Korea
[3]Department of Veterinary Nursing, Daekyeung University, Gyeongbuk 38547, Republic of Korea
[4]HanCare School Cooperative, Seorabeol University, Gyeongbuk 38159, Republic of Korea
*Corresponding author's email: azizul@ynu.ac.kr; kimjj@ynu.ac.kr

ORCID
Md Azizul Haque: https://orcid.org/0000-0002-5158-4558; Na-Kuang Kim: https://orcid.org/0009-0009-4615-624X
Ryu Yeji: https://orcid.org/0009-0001-6932-0729; ; Bugeun Lee: https://orcid.org/0000-0003-3347-3739
Ji-Hong Ha: https://orcid.org/0000-0002-4805-8370 Yun-Mi Lee: https://orcid.org/0000-0003-3023-7474
Han-Kook Il: https://orcid.org/0000-0002-1399-2884Jong-Joo Kim: https://orcid.org/0000-0001-9687-0075

ABSTRACT

This study aimed to identify QTL and nearby candidate genes influencing behavioral traits within the Korean Sapsaree dog population. Phenotypic data for three key behavioral traits – emotion, companion, and training were collected from 378 dogs, of which 234 individuals were genotyped using the 170k Illumina CanineHD BeadChip. Genome-wide association studies (GWAS) was performed using a univariate linear mixed model (LMM) implemented in the genome-wide efficient mixed-model analysis (GEMMA) software. The analysis identified 40 significant SNPs associated with these behavioral traits. Specifically, 15 significant sites were identified within CFA27 and CFA38 for emotion, 17 significant genomic regions across CFA1, CFA19, and CFA38 were noted for companionship behavior, and training behavior was associated with 8 significant regions within CFA7 and CFA21. Interestingly, several associated sites were found in or near genes linked to neuropathology or neuronal functions, including *FGD4*, *OVCH1*, and *EIF2D*, which are associated with emotion. Additionally, four neuronal-related genes, *NMBR*, *ADGRG6*, *HIVEP2*, and *PTPN4*, have been identified as influencing companion behavior. Significant variants impacting trainability were located within the *GREM2* and *SWAP70* genes. This study significantly advances our knowledge of the genetic architecture of behavior in Korean Sapsaree dogs, offering valuable insights into canines.

Keywords: Behavioral traits, Candidate genes, Canine, Korean Sapsaree, SNP.

INTRODUCTION

Dogs were the first animals domesticated by humans, with a history dating back approximately 14,000 years (Galibert *et al*., 2011; Wang *et al*., 2016). The domestic dog (*Canis lupus familiaris*), due to their loyalty to humans, has always been regarded as a human's best friend since their domestication. Historical artifacts, paintings, and literature serve as compelling evidence of the enduring bond between humans and their beloved canine companions, portraying dogs as cherished pets and valued members of the family, actively participating in various familial and social activities. In today's modern era, dogs play significant roles in extending friendship, companionship, affection, and love with humans, which have become increasingly difficult to get in "our nuclear families living impersonal suburban lifestyles" (Salmon and Salmon, 1983). Over time, dog breeders have selectively bred different breeds for specialized tasks ranging from herding and rescue missions to guarding, hunting, sled pulling, retrieving, and even detective work. Presently, the global registry lists over 400 dog breeds (Ha *et al*., 2015), each tailored to fulfill specific roles, although some breeds are primarily valued for their companionship, with less emphasis on utilitarian functions (Stafford, 2006).

The domestic dog stands out as one of the most diverse mammalian domestic species in the animal kingdom, showcasing a remarkable array of both behavioral and morphological traits (Wayne and Ostrander, 1999). The Sapsaree, indigenous to the Korean peninsula, stands as a prominent example. This breed is easily identifiable by its characteristic droopy ears and shaggy coat, believed to mirror the personality traits of the Korean people. In terms of temperament, Sapsaree exhibits a friendly disposition, loyalty, docility, alertness, and a remarkable ability to form strong bonds with their owners (Gajaweera *et al*., 2019). Initially, this breed was kept only by royals and aristocrats, but nowadays the Sapsaree has become a family member and widespread household pet for all social levels within Korea. However, the breed faced a severe population decline during the Japanese colonial period (1910 – 1945), nearly facing extinction due to widespread hunting for their hides and fur. The soft, shaggy fur of the Sapsaree was highly sought after in the leather industry for crafting premium winter coats for the Japanese military. By the mid-80s, the population of Sapsaree in the Daegu area had dwindled to a critical level, almost disappearing entirely (Kang, 2010).

Studying genetic traits related to emotion, companionship, and training in dogs holds significant importance for several reasons. Firstly, understanding the genetic basis of these traits can provide insights into the underlying mechanisms that govern canine behavior. By identifying specific genes associated with traits such as empathy, loyalty, and trainability, researchers can unravel the genetic pathways involved in regulating these behaviors. Secondly, genetic studies on these traits can inform breeding practices aimed at producing dogs with desirable temperament and behavior characteristics. By selectively breeding dogs with genetic predispositions for positive traits such as sociability, trainability, and emotional resilience, breeders can work towards improving the overall well-being and suitability of dogs as companions and working animals (Bray *et al*., 2021). Moreover, genetic research on these traits can have implications for animal welfare and training practices. By understanding the genetic factors that influence a dog’s behavior and temperament, trainers and behaviorists can develop more effective training and behavior modification techniques tailored to individual dogs’ needs. This can ultimately lead to more successful outcomes in addressing behavior issues and improving the quality of life for both dogs and their owners. Furthermore, studying the genetic basis of these traits can have broader implications for human-animal interactions and the human-animal bond. Dogs serve as important companions and support animals for many people, including those with disabilities, mental health conditions, and special needs (Brooks *et al*., 2018). By understanding the genetic underpinnings of traits such as empathy and companionship in dogs, researchers can gain insights into how these animals contribute to human well-being and explore ways to enhance the positive impact of human-dog relationships. The genetic studies on emotion, companion, and training traits in dogs have far-reaching implications for animal welfare, breeding practices, training techniques, and the human-animal bond. By unraveling the genetic basis of these traits, researchers can pave the way for advancements in canine genetics, behavior science, and the responsible stewardship of dogs as beloved companions and working partners.

In this context, our research on Korean Sapsaree dogs aimed to identify significant SNPs associated with emotional, companion, and training behaviors. Our objectives were to map the genetic architecture and biological relevance of these markers at the genome-wide level and identify potential candidate genes. This research not only has the potential to enhance animal welfare but also to enrich the quality of human-dog relationships, making it a pivotal endeavor in the study of domestic animals.

MATERIALS AND METHODS

Phenotypes and Genotypes: Phenotypic data were collected from 378 Korean Sapsaree dogs from the Korean Sapsaree Foundation. The phenotypic data included emotion, companion, and training traits. Among them, 234 animals were genotyped using an Illumina 170K CanineHD BeadChip (Illumina Inc., San Diego, CA, USA), containing 173,662 embedded SNPs. To ensure data quality, we initially removed SNPs located on sex chromosomes in duplicate or uncertain positions, resulting in the elimination of 6,478 SNPs. This left us with 167,184 SNPs for analysis. Subsequently, multiple quality control (QC) criteria were implemented to filter out low-quality SNPs. Specifically, SNPs with a minor allele frequency (MAF) of less than 5% (i.e., monomorphic), a SNP call rate below 90%, individuals with a genotyping call rate less than 90%, and SNPs showing a significant deviation from Hardy–Weinberg Equilibrium (HWE) with a p-value greater than 10^{-6} were excluded from the dataset. Additionally, the identity-by-state (IBS) test was conducted to identify genetically similar individuals or genotyping errors (Haque *et al*., 2023c). The pair of individuals showing a similarity rate >99% indicated either an identical animal or an error in genotyping. The IBS and entire QC process were performed using the PLINK v1.9 toolset (Purcell *et al*., 2007). Subsequently, all genotyped animals were imputed using Beagle v5.1 software (Browning *et al*., 2018). After IBS and QC procedures, a total of 205 animals with genotypes for 104,315 SNPs were available for further analysis.

GWAS analysis: The Sapsaree GWAS was conducted by applying a univariate linear mixed model (LMM) implemented in the genome-wide efficient mixed-model analysis (GEMMA) software v0.98.5 (Zhou and Stephens, 2012). GEMMA calculates a genomic relationship matrix (GRM) between individuals within each population to determine the population structure.

The univariate linear mixed model in GEMMA was described as follows:

$$y = W\alpha + X\beta + u + \varepsilon \qquad (5)$$

where y is the vector of phenotypes; W is the vector of the fixed effects of sex of the animals, measurement date, and age as a covariate; α is a vector of the corresponding coefficients, including the intercept; X represents the vector of all marker genotypes; β represents the effect size of the SNP; $u \sim MVN_n(0, \lambda\tau^{-1}K)$ is an n-vector of animal additive effects; and $\varepsilon \sim MVN_n(0, \tau^{-1}I_n)$ represents an n-vector of errors; τ^{-1} is the variance of the residual errors; λ is the ratio between the two variance components; K represents the genomic relationship matrix (GRM); I_n is an n x n identity matrix; and MVN_n represents the n-dimensional multivariate normal distribution. GEMMA calculates the GRM as follows (Zhou and Stephens, 2012):

$$G = \frac{1}{p}\sum_{i=1}^{p}(x_i - 1_n\bar{x}_i)(x_i - 1_n\bar{x}_i)^T \qquad (6)$$

where X represents the n × p matrix of the genotypes, x_i represents the genotypes of the i^{th} SNP, $\bar{x}_i$ is the sample mean, and 1_n is the n × 1 vector of 1.

Analysis of Haplotype block: GWAS often identifies significant SNPs associated with target traits within specific chromosomal regions (Haque *et al*., 2024). The clustering of these SNPs may arise from high linkage disequilibrium (LD) and the non-random association of alleles on the chromosome. To explore these genomic patterns in detail, we used PLINK v1.9 (Purcell *et al*., 2007) and LDBlockShow software (Dong *et al*., 2021). These tools allowed us to investigate regions where multiple SNPs clustered around the top SNP, providing insights into the LD patterns within these regions.

Identification of candidate genes: We identified potential candidate genes within the QTL regions and in the nearest upstream and downstream areas (500 kb) of our mapped significant SNPs (Haque *et al*., 2023a; Haque *et al*., 2024). For this analysis, we used the dog genome assembly (Canfam3.1) and online resources such as the NCBI Genome Data Viewer (https://www.ncbi.nlm.nih.gov/genome/gdv?org=canis-lupus-familiaris; accessed on 18 April 2024). We explored the functional roles of the genes located within and adjacent to significant SNPs associated with behavioral traits by reviewing published reports in the PMC and other databases from the National Center for Biotechnology Information (NCBI). We also obtained information on the functions of each gene from online resources, including human gene functions at GeneCards (www.genecards.org), the Mouse Genome Informatics (MGI) website (https://www.informatics.jax.org/), and Ensembl (www.ensembl.org/biomart/martview), accessed on 19 April 2024. Genes that demonstrated functional relevance to the traits of interest were considered promising candidates for further study. For the GWAS analysis of behavioral traits in the Sapsaree population, we employed the Wald test to assess the significance of associations (Zhou and Stephens, 2012). We determined the genome-wide significance threshold using the Bonferroni test (0.05/N, where N represents the number of SNPs analyzed). Despite this stringent criterion, the training trait did not yield any SNPs that met these stringent conditions. Consequently, we adopted a more lenient suggestive level threshold (1/N) to capture potential loci of interest, given the restrictive nature of the Bonferroni adjustment (Guo *et al*., 2017; Haque *et al*., 2023b; Haque *et al*., 2024). Nonetheless, even at this reduced threshold, the training trait did not meet the suggestive level. Therefore, we set a considered significance threshold at a $-\log_{10}P$ value of 4 for this trait (Alam *et al*., 2023). While this adjustment might increase the risk of false positives, it is intended to ensure that we do not overlook any plausible genetic loci associated with the trait under study in the Sapsaree population (Haque *et al*., 2023b; Haque *et al*., 2024).To visualize the distribution of significant SNPs across the genome, we generated Manhattan plots, with the significance levels depicted as the negative base-10 logarithm ($-\log_{10}$) of the p-value for each SNP. Additionally, we calculated the genomic inflation factor, lambda (λ), to assess potential population stratification. This was done by comparing the median of the observed chi-squared test statistics from the GWAS to the expected median of the chi-squared distribution, using the qchisq() function in R (R Core Team, 2023). Ideally, λ should approximate 1 after adjusting for population stratification (Haque *et al*., 2023b; Haque *et al*., 2024). However, a notably high λ value suggests that other factors such as strong linkage disequilibrium, significant phenotypic associations, or systematic biases might be influencing the observed inflation (Haque *et al*., 2023b; Haque *et al*., 2024). The QQ plots further illustrate this, showing the observed versus expected $-\log_{10}$ p-values; a straight line represents the null hypothesis distribution of markers, while a deviation at the upper end indicates markers more strongly associated with the traits than expected by chance.

RESULTS

Phenotypes: Table 1 presents the summary statistics for behavioral traits of Korean Sapsaree dogs that have both genotype and phenotype data available. For the trait emotion, the mean score is 8.88 with a standard deviation (SD) of 3.10, indicating a moderate variability among the scores. The coefficient of variation (CV) for this trait is 34.87%, reflecting a relatively high degree of variance relative to the mean. The scores for emotion ranged from a minimum of 2.00 to a maximum of 15.00. In terms of the companion trait, the average score is lower, at 4.66, with an SD of 1.03, suggesting less variability among individual scores compared to emotion. The CV for companion is 22.17%, indicating a more consistent behavior among the dogs regarding their companionship attributes. The scores varied from a minimum of 2.00 to a maximum of 6.00. The training trait exhibited the highest variability among the three traits, with an SD of 1.86 and a notably high CV of 53.26%. This suggests that training behavior varies widely among the dogs in the study. The mean score for training was 3.50, with values ranging from 0.00, indicating no training responsiveness, to a high of 8.00. These statistics provide a quantitative insight into the behavioral tendencies of Korean Sapsaree dogs, highlighting significant variations in training responsiveness compared to more consistent behaviors in emotional and companionship capacities.

Table 1. Summary statistics for behavior traits of Korean Sapsaree dogs.

Traits	Mean	SD	CV (%)	Maximum	Minimum
Emotion	8.88	3.10	34.87	15.00	2.00
Companion	4.66	1.03	22.17	6.00	2.00
Training	3.50	1.86	53.26	8.00	0.00

N, number of animals; SD, standard deviation; CV, coefficient of variation.

SNP genotyping: After conducting quality control measures that retained 62.39% of the initial SNPs, we selected a set of 104315 common SNPs across all 38 *Canis lupus familiaris* autosomes (CFA). The distribution of these genetic markers was not uniform, with a notable overrepresentation on certain chromosomes, as shown in Fig. 1. Notably, CFA1 contained the highest number of SNPs, totaling 5453 and spanning approximately 122 megabases (Mb). In contrast, CFA37 had the fewest SNPs, with 1226 SNPs covering about 30.73 Mb. Additionally, CFA3, CFA4, CFA5, and CFA7 each contained over 3000 SNPs, as detailed in Table S1.

Fig. 1. Distribution of SNPs across chromosomes after quality control.

Association analysis: In this GWAS of behavioral traits, we identified 40 significant SNPs across 38 CFAs. This analysis highlighted specific chromosomes with a higher concentration of significant genetic markers, particularly CFA27, which emerged as a hotspot with 12 significant SNPs. Following closely, CFA1 displayed 11 significant SNPs, while CFA7, CFA19, CFA21, and CFA38 each had notable counts of significant SNPs, underscoring their potential relevance in behavioral genetics. The comprehensive summary of these results includes significant SNP IDs, their positions on respective chromosomes, allele types (minor/major), MAF, p-values, and nearby candidate genes, providing a detailed map of the genetic architecture underlying behavioral traits in the Korean Sapsaree dog.

Marker loci associated with emotion trait: The GWAS on emotional behavior in Korean Sapsaree dogs revealed several SNPs associated with this trait. We identified a total of 15 SNPs across all 38 CFA, as detailed in Table 2 and illustrated in Fig. 2a. Notably, CFA27 was highlighted as a potential genomic hotspot, hosting the highest number of SNPs (12 variants) related to emotion, which suggests a significant influence on this behavioral trait. These SNPs are distributed from 16.34 Mb to 18.99 Mb, indicating a substantial genetic influence in this region. All identified SNPs on CFA27 exhibited complete LD and were clustered within a 2.64 Mb haplotype block (27: 16344940 – 18989641bp). This clustering suggests that mutations near this region, which likely represent a QTL, could markedly impact emotional behaviors, as depicted in Fig. 3. The regional association plots and a heatmap of LD for these findings are also presented in Fig. 3, which includes a specific focus on BTA38. Among these SNPs, the one located at 16.76 Mb on CFA27 (BICF2G630146733) showed the highest significance, with a p-value of 1.17×10^{-10}. Other noteworthy significant SNPs on CFA27, which met the p-value threshold, included BICF2P268720, $p = 3.14 \times 10^{-10}$; BICF2G630146755, $p = 5.77 \times 10^{-10}$; BICF2P652058, $p = 7.45 \times 10^{-10}$; BICF2P906575, $p = 9.45 \times 10^{-10}$; and BICF2G630146263, $p = 1.40 \times 10^{-9}$, located between 16.34 Mb to 18.99 Mb. To assess potential population stratification, we calculated the λ, which resulted in a value of 0.995 (Fig. 2b). We also generated a QQ plot to visually represent the observed versus expected p-values (-$\log_{10}$P), as displayed in Fig. 2b. This QQ plot indicated a close alignment between observed and expected values, confirming that the p-values were normally distributed. This alignment suggests that population stratification was effectively managed using the appropriate model, thereby enhancing the probability of identifying true genetic associations.

Table 2. Genome-wide significant SNPs underlying emotion behavior in Sapsaree dogs.

SNP ID	CFA	Position (bp)	Allele	MAF	P	Gene
BICF2G630146263	27	16,344,940	[G/A]	0.613	1.40×10^{-9}	*FGD4*
BICF2P866415	27	16,423,267	[G/A]	0.613	1.40×10^{-8}	*FGD4*
BICF2G630146347	27	16,441,261	[A/C]	0.613	3.40×10^{-8}	*FGD4, BICD1*
BICF2G630146668	27	16,732,436	[G/A]	0.613	6.40×10^{-8}	*BICD1, KIAA1551*
BICF2G630146733	27	16,761,901	[A/G]	0.718	1.17×10^{-10}	*BICD1, KIAA1551*
BICF2G630146755	27	16,771,403	[G/A]	0.718	5.77×10^{-10}	*BICD1, KIAA1551*
BICF2G630146879	27	16,863,513	[A/G]	0.704	2.26×10^{-7}	*KIAA1551, LOC106557885*
BICF2P268720	27	17,955,918	[A/G]	0.549	3.14×10^{-10}	*LOC100685338, LOC111092883*
BICF2P963607	27	18,339,272	[C/A]	0.563	8.26×10^{-9}	*TMTC1, OVCH1*
BICF2P652058	27	18,485,690	[G/A]	0.521	7.45×10^{-10}	*TMTC1, OVCH1*
TIGRP2P350000_rs8670902	27	18,587,720	[C/A]	0.619	4.49×10^{-7}	*OVCH1, LOC111092860*
BICF2P906575	27	18,989,641	[A/C]	0.521	9.45×10^{-10}	*LOC102156548, LOC102156632*
BICF2S2361184	38	2,415,363	[G/A]	0.613	1.33×10^{-7}	*LOC111094591, C38H1orf186*
BICF2P58157	38	2,589,332	[A/C]	0.584	2.72×10^{-8}	*SRGAP2, FAM72A*
BICF2S23650480	38	2,864,169	[A/G]	0.606	4.26×10^{-8}	*RASSF5, EIF2D, CFH*

CFA, *Canis lupus familiaris* autosomes; bp, base pairs; Allele, minor/major; MAF, minor allele frequency.

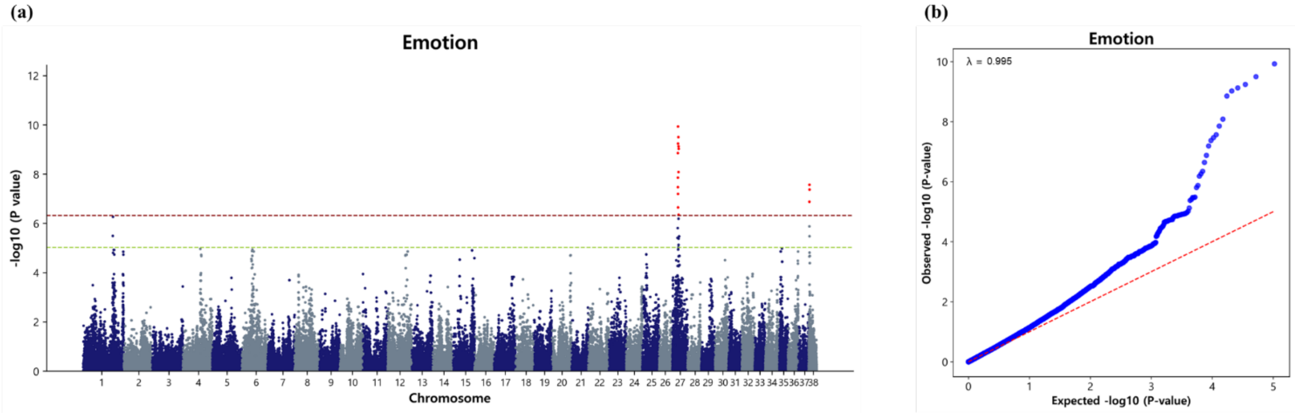

Fig. 2. Association of 104,315 SNPs with the emotional behavior in the Sapsaree. (a) Manhattan plot. The y-axis represents -$\log_{10}$ (observed) p-values for genome-wide SNPs against their respective positions on each chromosome (x-axis). The horizontal dashed yellow-green line indicated the suggestive ($p = 9.59 \times 10^{-6}$) threshold level, and the dashed red line indicated the Bonferroni corrected ($p = 4.79 \times 10^{-7}$) threshold level. (b) Quantile – quantile plot. The red line represents the 95% confidence interval under the null hypothesis of no association. The blue dot represents the p-values of the entire study.

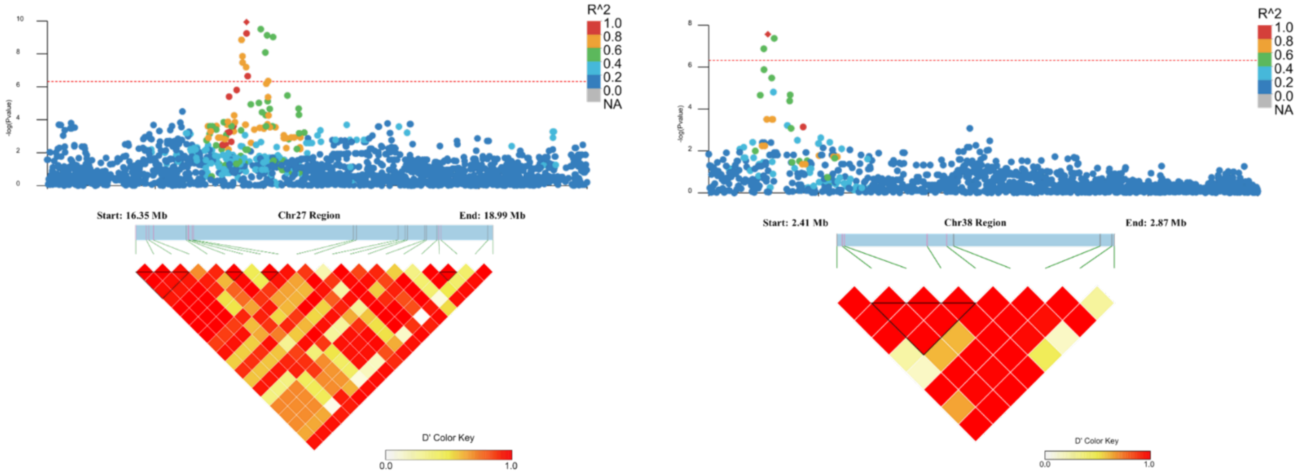

Fig. 3. Regional association plot (top) illustrating the distribution of significant loci associated with emotional behavior and heatmap of LD (bottom) at CFA27 and CFA38. The dashed red horizontal line indicates -log10P = 6.32 (Bonferroni threshold).

Marker loci associated with companion: The GWAS revealed significant chromosomal regions related to companion behavior in Sapsaree dogs, with notable associations on CFA1, CFA19, and CFA38 (Table 3 and Fig. 4a). CFA1 was associated with eleven SNPs spanning from 31.89 Mb to 35.60 Mb while CFA19 showed three SNPs between 30.26 Mb and 43.12 Mb. CFA38 exhibited three significant SNPs located at 5.64 Mb and 6.99 Mb related to companion behavior. The SNPs with the strongest associations were BICF2P1426908 ($p = 1.58 \times 10^{-9}$), BICF2G630719299 ($p = 2.19 \times 10^{-9}$), BICF2G630719350 ($p = 2.19 \times 10^{-9}$), BICF2G630719572 ($p = 2.19 \times 10^{-9}$), and BICF2G630719628 ($p = 2.19 \times 10^{-9}$). The regional association plots and LD heatmaps for these regions on CFA1, CFA19, and CFA38 are illustrated in Fig. S1. A QQ plot was also generated, indicating a normal distribution of the test statistics, and the λ value was calculated at 1.033, confirming effective management of potential confounding factors, as shown in Fig. 4b.

Table 3. Genome-wide significant SNPs underlying companion behavior in Sapsaree dogs.

SNP ID	CFA	Position (bp)	Allele	MAF	P	Gene
BICF2G630717812	1	31,894,464	[A/G]	0.859	4.55×10^{-7}	*LOC111095494, LOC111090654*
BICF2G630719010	1	33,777,244	[G/A]	0.831	1.81×10^{-7}	*NMBR, VTA1*
BICF2G630719299	1	33,967,724	[G/A]	0.866	2.19×10^{-9}	*VTA1, ADGRG6*
BICF2G630719350	1	34,024,044	[A/G]	0.866	2.19×10^{-9}	*ADGRG6*
BICF2G630719469	1	34,339,914	[C/G]	0.831	1.81×10^{-7}	*ADGRG6, LOC100686509*
BICF2P1099766	1	34,344,589	[A/C]	0.831	1.81×10^{-7}	*LOC100686509, HIVEP2*
BICF2G630719572	1	34,563,040	[A/G]	0.866	2.19×10^{-9}	*HIVEP2, LOC100856282*
BICF2G630719628	1	34,599,217	[G/A]	0.866	2.19×10^{-9}	*LOC111090656*
BICF2G630719767	1	34,826,842	[G/A]	0.852	6.87×10^{-8}	*AIG1*
BICF2G630720123	1	35,289,223	[A/G]	0.866	2.19×10^{-9}	*PHACTR2, LTV1*
BICF2S23633648	1	35,601,699	[A/G]	0.880	2.69×10^{-7}	*STX11*
BICF2G63050675	19	30,262,180	[A/T]	0.718	8.36×10^{-9}	*PTPN4, CFAP221*
BICF2P1426908	19	39,228,170	[G/C]	0.669	1.58×10^{-9}	*LOC111091069*
TIGRP2P268055_rs8830561	19	43,123,608	[G/A]	0.570	4.65×10^{-7}	*LRP1B*
BICF2S23143320	38	5,637,193	[G/A]	0.739	4.03×10^{-8}	*LOC111094538*
BICF2P1182782	38	5,656,982	[G/A]	0.739	4.03×10^{-8}	*LOC111094538, CDC73*
BICF2P1179733	38	6,990,679	[G/A]	0.930	7.70×10^{-8}	*LOC102155162*

CFA, *Canis lupus familiaris* autosomes; bp, base pairs; Allele, minor/major; MAF, minor allele frequency.

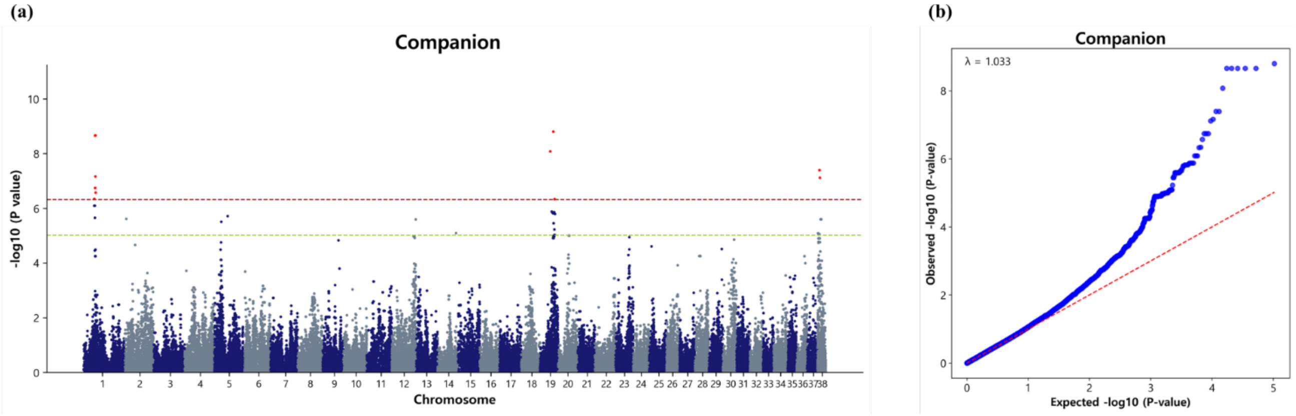

Fig. 4. Association of 104,315 SNPs with the companion behavior in the Sapsaree. (a) Manhattan plot. The y-axis represents $-\log_{10}$ (observed) p-values for genome-wide SNPs against their respective positions on each chromosome (x-axis). The horizontal dashed yellow-green line indicated the suggestive ($p = 9.59 \times 10^{-6}$) threshold level, and the dashed red line indicated the Bonferroni corrected ($p = 4.79 \times 10^{-7}$) threshold level. (b) Quantile – quantile plot. The red line represents the 95% confidence interval under the null hypothesis of no association. The blue dot represents the p-values of the entire study.

Marker loci associated with training: SNPs with a p-value ($-\log_{10}P$) of 4.0 or higher were selected as significant markers for the training trait in Sapsaree dogs. This threshold was set to capture a broader spectrum of genetic variations, thus providing a more comprehensive understanding of the genetic architecture influencing training behavior. Eight significant marker loci were identified across two CFAs, with CFA21 containing six loci and CFA7 harboring two (Fig. 5a). The SNPs on CFA21 clustered in a genomic region between 31.65 Mb and 33.90 Mb (Table 4). The most significant SNPs related to training included BICF2P786475, BICF2P313736, BICF2S23334127, BICF2S2347653, and BICF2P581743, each with a p-value of 2.51×10^{-5}. The regional association plot and LD heatmap for these regions on CFA7 and CFA21 are illustrated in Fig. S2. The QQ plot for this genetic locus exhibited a normal distribution of p-values, with a λ value of 1.057, indicating effective control of spurious results and a high likelihood of identifying true associations (Fig. 5b).

Table 4. Genome-wide significant SNPs underlying training behavior in Sapsaree dogs.

SNP ID	CFA	Position (bp)	Allele	MAF	P	Gene
BICF2P1015105	7	30915911	[A/G]	0.852	9.12×10^{-5}	*CREG1, CD247*
BICF2P46284	7	32098697	[A/G]	0.774	3.94×10^{-5}	*GREM2*
BICF2P786475	21	32987753	[A/G]	0.788	2.51×10^{-5}	*SWAP70, SBF2*
TIGRP2P284427_rs9140654	21	31656691	[G/A]	0.556	7.84×10^{-5}	*RIC3*
BICF2P313736	21	33079977	[A/G]	0.788	2.51×10^{-5}	*SBF2*
BICF2S23334127	21	33524208	[A/C]	0.788	2.51×10^{-5}	*LOC106560154*
BICF2S2347653	21	33828382	[A/G]	0.787	2.51×10^{-5}	*MRVI1*
BICF2P581743	21	33906373	[A/G]	0.788	2.51×10^{-5}	*LOC111091556*

CFA, *Canis lupus familiaris* autosomes; bp, base pairs; Allele, minor/major; MAF, minor allele frequency.

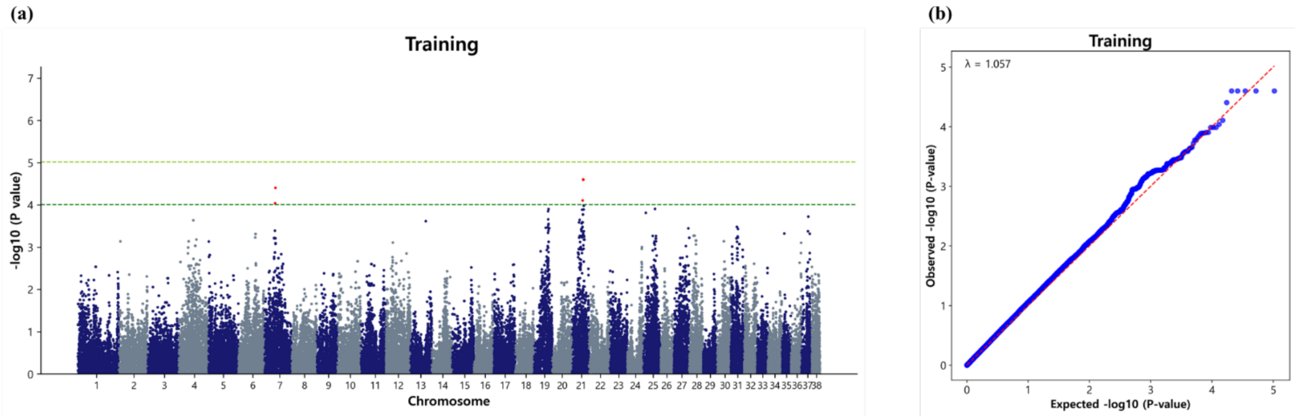

Fig. 5. Association of 104,315 SNPs with the training behavior in the Sapsaree. (a) Manhattan plot. The y-axis represents -log$_{10}$ (observed) p-values for genome-wide SNPs against their respective positions on each chromosome (x-axis). The horizontal dashed green line indicated the considered ($p = 1 \times 10^{-4}$) threshold level and the dashed yellow green line indicated the suggestive ($p = 9.59 \times 10^{-6}$) threshold level. (b) Quantile – quantile plot. The red line represents the 95% confidence interval under the null hypothesis of no association. The blue dot represents the p-values of the entire study.

Identification of candidate genes: We searched the National Center for Biotechnology Information (NCBI) database using the CanFam3.1 genome assembly to identify potential candidate genes associated with the behavioral traits of sapsaree dogs. We focused on a ±500 Kb window around the identified significant marker loci and found 47 unique positional candidate genes situated within a 1 Mb radius of these loci (Tables 2, 3, 4). Specifically, 19 genes were associated with emotion, 20 with companion, and 8 with training. These findings provide insights into the genetic contributors to these behavioral traits in Sapsaree dogs and suggest several candidate genes that could play key roles in behavior regulation. For the emotion trait, the potential candidate genes include *FGD4*, *OVCH1*, and *EIF2D*. The companion trait was associated with the genes *NMBR*, *ADGRG6*, *HIVEP2*, and *PTPN4*. Additionally, two genes, *GREM2* and *SWAP70*, were identified as influential for training behavior (Table 5).

Table 5. Promising candidate genes associated with studied behavioral traits in Sapsaree populations.

Traits	Genes	Name	CFA	QTL Position (Mb)
Emotion	*FGD4*	FYVE, RhoGEF and PH domain containing 4	27	16.23 – 16.49
	OVCH1	Ovochymase 1	27	18.52 – 18.60
	EIF2D	Eukaryotic translation initiation factor 2D	38	2.85 – 2.87
Companion	*NMBR*	Neuromedin B receptor	1	33.76 – 33.87
	ADGRG6	Adhesion G protein-coupled receptor G6	1	33.95 – 34.12
	HIVEP2	HIVEP zinc finger 2	1	34.35 – 34.58
	PTPN4	Protein tyrosine phosphatase non-receptor type 4	19	30.06 – 30.31
Training	*GREM2*	Gremlin 2, DAN family BMP antagonist	7	32.04 – 32.17
	SWAP70	Switching B cell complex subunit SWAP70	21	32.90 – 32.99

CFA, *Canis lupus familiaris* autosomes; QTL, quantitative trait loci; Mb, megabases.

DISCUSSION

Candidate gene functions associated with emotion: The emotion trait was analyzed to understand the underlying genetic variance influencing emotional responsiveness in Korean Sapsaree dogs. The GWAS results identified specific loci associated with emotion, suggesting that certain genetic configurations may predispose these dogs to more expressive or subdued emotional behaviors. Such insights are invaluable for breeders aiming to select temperament traits that suit service or therapy roles. In this study, we identified the presence of the *FGD4* and *OVCH1* genes on CFA27, while the *EIF2D* gene was located on CFA38 near the significant SNPs associated with emotion in Sapsaree dogs. Specifically, the FYVE, RhoGEF, and PH domain containing 4 (*FGD4*) was located within the genomic range of 16.23 Mb to 16.49 Mb. The *FGD4* gene is recognized for its role in cellular signaling and cytoskeletal organization, primarily through its function as a guanine nucleotide exchange factor (GEF) for Rho family GTPases (Martin-Vilchez *et al*., 2017). Although research has traditionally focused on *FGD4*'s physiological and pathological roles in cellular processes, there is growing interest in its implications for behavioral traits, particularly due to its involvement in neural development and function. *FGD4* is crucial for the development and maintenance of myelin, which insulates nerve fibers and is essential for efficient neural signal transmission. Abnormalities in myelin, which can result from mutations in *FGD4*, can disrupt neural connectivity and affect behavior, as seen in disorders like Charcot-Marie-Tooth disease type 4H (Horn *et al*., 2012). Furthermore, *FGD4*'s role in cytoskeletal dynamics may influence neural plasticity, a critical element of learning, memory, and behavioral adaptation (Priel *et al*., 2010). While direct studies connecting *FGD4* to specific behavioral outcomes are currently limited, the gene's impact on neural development and synaptic plasticity suggests a potential role in cognitive functions, motor skills, and emotional responses (Marzola *et al*., 2023). Additionally, the ovochymase 1 (*OVCH1*) gene, is primarily known for encoding a protease enzyme related to the reproductive system (Kiyozumi and Ikawa, 2022) and has intrigued researchers regarding its potential indirect influence on behavioral traits. Although not as directly connected to neurological pathways as other genes, *OVCH1* might impact behavior through mechanisms such as hormonal regulation and developmental processes. For instance, given its role in protease activity, *OVCH1* could affect the processing and metabolism of key hormones like estrogen and testosterone (Snyder *et al*., 2009), which are known to significantly influence mood and behavior. Additionally, the expression of *OVCH1* during critical developmental periods might impact the maturation of brain tissues, subtly influencing an individual's behavioral patterns. The complexity of behavioral genetics suggests that *OVCH1* could interact with multiple genes and environmental factors, complicating the task of understanding its precise role without extensive multidisciplinary research. The eukaryotic translation initiation factor 2D (*EIF2D*), is traditionally recognized for its role in protein synthesis (Gilabert-Juan *et al*., 2019) but is increasingly studied for its potential impact on behavioral traits. This interest stems from the fact that protein synthesis is crucial for neurobiological processes such as neuroplasticity, memory formation, and stress responses in the brain (Hernandez and Abel, 2008). Abnormalities in *EIF2D* could affect these processes, influencing behaviors related to learning, stress resilience, and even neurodevelopmental or neurodegenerative disorders. Given its central role in cellular function and response to environmental stimuli, variations in *EIF2D* activity or expression could potentially lead to significant behavioral phenotypes.

Candidate gene functions associated with companion: In our study, we identified three significant candidate genes, *NMBR, ADGRG6,* and *HIVEP2*, associated with companionship. These genes play an important role in various behavioral traits through their involvement in neural and physiological processes. The neuromedin B receptor (*NMBR*) is activated by neuromedin B, a neuropeptide belonging to the bombesin-like peptide family (Cikes *et al*., 2022). This receptor influences physiological processes including stress response, thermoregulation, and feeding behavior. Its activation in specific brain regions has been linked to anxiety and stress-related behaviors, mediated through intricate neural signaling pathways (Martin *et al*., 2009). *ADGRG6*, a member of the adhesion GPCR family, is essential for synaptic organization and neuron-neuron communication, impacting cognitive traits such as learning and memory (Baxendale *et al*., 2021). This receptor's involvement in neuronal development suggests its role in behavioral adaptation and plasticity. *HIVEP2* encodes a zinc finger protein that regulates transcription, influencing memory and learning (Srivastava *et al*., 2016). Alterations in *HIVEP2* expression have been connected to intellectual disabilities and cognitive disorders, underscoring its importance in neural function and development (Steinfeld *et al*., 2016). Protein tyrosine phosphatase non-receptor type 4 (*PTPN4*) is involved in cell growth and synaptic scaling, affects neuroplasticity and thus influences learning processes and behavioral responses (Hendriks *et al*., 2022). Its dysregulation has been associated with neurological disorders, highlighting its potential impact on related behaviors. These genes collectively contribute to the neural architecture that underpins companion traits in dogs, offering insights into their complex behavioral patterns.

Candidate gene functions associated with training: In our GWAS focusing on the training trait in dogs, we identified several genomic regions significantly associated with trainability, indicating a genetic foundation for this behavior. The loci identified suggest that specific genetic variations influence a dog's ability to respond to training cues and learn new tasks, which is critical for roles in service, therapy, and obedience. In exploring canine genetics, particularly concerning training behaviors, we identified two key candidate genes in Korean Sapsaree dogs, located on CFA7 (*GREM2*) and CFA21 (*SWAP70*). *GREM2* plays a vital role in developmental processes by acting as a BMP (bone morphogenetic protein) antagonist, potentially affecting neural development and plasticity (Wang *et al*., 2017) – both of which are crucial for a dog's learning and memory capabilities. Though primarily associated with embryonic development and tissue differentiation (Tanwar *et al*., 2014), *GREM2*'s involvement in neural pathways (Sanders *et al*., 2016) suggests that it might also affect cognitive abilities and trainability. Conversely, *SWAP70*, known for its role in actin rearrangement and cellular signaling, particularly within the immune system, could indirectly influence training behaviors through its impact on neural signaling (Baranov *et al*., 2016). Effective neural signaling underpins cognitive function and the ability to learn and execute complex behaviors, making *SWAP70* another gene of interest in the study of canine trainability.

The potential implications of identifying genetic variants in *GREM2* and *SWAP70* that correlate with training behaviors are vast. For breeders and trainers, such insights could lead to more informed selective breeding practices aimed at enhancing desirable traits. Additionally, understanding these genetic underpinnings could help in developing tailored training programs that align with the innate abilities and limitations of individual dogs, thereby optimizing training efficiency and effectiveness. This is particularly beneficial in roles that demand high levels of trainability and cognitive function, such as service or therapy dogs, and working dogs in security positions. While the definitive impact of *GREM2* and *SWAP70* on canine training behaviors requires further empirical validation, their established roles in developmental and signaling pathways provide a solid foundation for hypothesizing their influence. Continued genetic research is crucial to untangling the complex genetic networks that dictate these sophisticated animal behaviors, potentially revolutionizing our approach to canine training and breeding.

Conclusions: This study identified multiple QTL and candidate genes associated with behavioral traits in Korean Sapsaree dogs. A total of 40 significant SNPs were detected across different genomic regions. Genes such as *FGD4*, *OVCH1*, and *EIF2D* were implicated in emotional regulation, while *NMBR*, *ADGRG6*, *HIVEP2*, and *PTPN4* were associated with companionship behavior. Additionally, trainability was associated with variants in *GREM2* and *SWAP70*. By identifying genetic markers related to emotion, companionship, and training, this research offers valuable insights into canine behavioral genetics. The identified markers may improve breeding programs, enabling the selection of dogs with desirable traits for specific roles or as family pets.

Authorship contribution statement: M.A.H.: Conceptualization, Methodology, Formal analysis, Visualization, Investigation, Writing - Original draft, Writing - Review and Editing. N.K.K.: Methodology, Writing - Review and Editing. R.Y.: Investigation, Writing - Review and Editing. B.L.: Data curation, Investigation, Writing - Review and Editing. J.H.H.: Data curation, Investigation, Writing - Review and Editing. Y.M.L.: Investigation, Writing - Review and Editing. J.J.K.: Supervision, Project administration, Funding acquisition Writing - Review and Editing.

Funding: This work was supported by the 2023 Yeungnam University Research Grant.

Data availability statement: The original genotypic data (SNPs) used in this study are accessible on Figshare (https://doi.org/10.6084/m9.figshare.25769994.v2). To access the data for reproducing the results, an application must be submitted to, and approved by, the authors. The phenotypic data analyzed in this study are not currently publicly available, as further analytical studies are planned. However, they can be obtained from the corresponding author upon reasonable request.

Ethics statement: The care and management of all animals used in this study were approved by the Animal Care and Use Committee of the National Institute of Animal Science (NIAS), Rural Development Administrations (RDA), South Korea (Approval No. 2016-177). Ear tissue samples were collected by veterinarians following ethical guidelines for animal health and welfare. The experimental animals were not anesthetized or euthanized in this study. We confirmed that all methods are reported in accordance with ARRIVE guidelines (https://arriveguidelines.org) for the reporting of animal experiments.

Conflict of interest statement: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

Alam, M.Z., M.A. Haque, A. Iqbal, Y.M. Lee, J.J. Ha, S. Jin, B. Park, N.Y. Kim, J.I. Won and J.J. Kim (2023). Genome-wide association study to identify QTL for carcass traits in Korean Hanwoo cattle. Animals (Basel) 13, 2737. https://doi.org/10.3390/ani13172737.

Baranov, M.V., N.H. Revelo, I. Dingjan, R. Maraspini, M. Ter Beest, A. Honigmann and G. van den Bogaart (2016). SWAP70 organizes the actin cytoskeleton and is essential for phagocytosis. Cell reports 17, 1518-1531. https://doi.org/10.1016/j.celrep.2016.10.021.

Baxendale, S., A. Asad, N.O. Shahidan, G.R. Wiggin and T.T. Whitfield (2021). The adhesion GPCR Adgrg6 (Gpr126): insights from the zebrafish model. Genesis (New York, N.Y.: 2000) 59, e23417. https://doi.org/10.1002/dvg.23417.

Bray, E.E., C.M. Otto, M.A.R. Udell, N.J. Hall, A.M. Johnston and E.L. MacLean (2021). Enhancing the selection and performance of working dogs. Frontiers in Veterinary Science 8, 644431. https://doi.org/10.3389/fvets.2021.644431.

Brooks, H.L., K. Rushton, K. Lovell, P. Bee, L. Walker, L. Grant and A. Rogers (2018). The power of support from companion animals for people living with mental health problems: a systematic review and narrative synthesis of the evidence. BMC psychiatry 18, 31. https://doi.org/10.1186/s12888-018-1613-2.

Browning, B.L., Y. Zhou and S.R. Browning (2018). A one-penny imputed genome from next-generation reference panels. Am. J. Hum. Genet. 103, 338-348. https://doi.org/10.1016/j.ajhg.2018.07.015.

Cikes, D., P. Atanes, S.J.F. Cronin, A. Hagelkruys, G.C. Huang, S.J. Persaud and J.M. Penninger (2022). Neuropeptide neuromedin B does not alter body weight and glucose homeostasis nor does it act as an insulin-releasing peptide. Sci Rep 12, 9383. https://doi.org/10.1038/s41598-022-13060-0.

Dong, S.S., W.M. He, J.J. Ji, C. Zhang, Y. Guo and T.L. Yang (2021). LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. Brief Bioinform. 22. https://doi.org/10.1093/bib/bbaa227.

Gajaweera, C., J.M. Kang, D.H. Lee, S.H. Lee, Y.K. Kim, H.I. Wijayananda, J.J. Kim, J.H. Ha, B.H. Choi and S.H. Lee (2019). Genetic diversity and population structure of the Sapsaree, a native Korean dog breed. BMC Genet 20, 66. https://doi.org/10.1186/s12863-019-0757-5.

Galibert, F., P. Quignon, C. Hitte and C. Andre (2011). Toward understanding dog evolutionary and domestication history. C R Biol 334, 190-196. https://doi.org/10.1016/j.crvi.2010.12.011.

Gilabert-Juan, J., G. Lopez-Campos, N. Sebastia-Ortega, S. Guara-Ciurana, F. Ruso-Julve, C. Prieto, B. Crespo-Facorro, J. Sanjuan and M.D. Molto (2019). Time dependent expression of the blood biomarkers EIF2D and TOX in patients with schizophrenia. Brain Behav Immun 80, 909-915. https://doi.org/10.1016/j.bbi.2019.05.015.

Guo, Y., Y. Huang, L. Hou, J. Ma, C. Chen, H. Ai, L. Huang and J. Ren, J., (2017). Genome-wide detection of genetic markers associated with growth and fatness in four pig populations using four approaches. Genet Sel Evol 49, 21. https://doi.org/10.1186/s12711-017-0295-4.

Ha, J.H., M. Alam, D.H. Lee and J.J. Kim (2015). Whole Genome association study to detect single nucleotide polymorphisms for behavior in Sapsaree dog (Canis familiaris). Asian-Australas J Anim Sci 28, 936-942. https://doi.org/10.5713/ajas.14.0941.

Haque, M.A., M.Z. Alam, A. Iqbal, Y.M. Lee, C.G. Dang and J.J. Kim (2023b). Genome-wide association studies for body conformation traits in Korean Holstein population. Animals (Basel) 13, 2964. https://doi.org/10.3390/ani13182964.

Haque, M.A., A. Iqbal, H. Bae, S.E. Lee, S. Park, Y.M. Lee and J.J. Kim (2023c). Assessment of genomic breeding values and their accuracies for carcass traits in Jeju Black cattle using whole-genome SNP chip panels. J Anim Breed Genet 140, 519-531. https://doi.org/10.1111/jbg.12776.

Haque, M.A., Y.M. Lee, J.J. Ha, S. Jin, B. Park, N.Y. Kim, J.I. Won and J.J. Kim (2024). Genome-wide association study identifies genomic regions associated with key reproductive traits in Korean Hanwoo cows. BMC Genomics 25, 496. https://doi.org/10.1186/s12864-024-10401-3.

Hendriks, W., R.T.P. van Cruchten and R. Pulido (2022). Hereditable variants of classical protein tyrosine phosphatase genes: Will they prove innocent or guilty? Front Cell Dev Biol 10, 1051311. https://doi.org/10.3389/fcell.2022.1051311.

Hernandez, P.J. and T. Abel (2008). The role of protein synthesis in memory consolidation: progress amid decades of debate. Neurobiology of learning and memory 89, 293-311. https://doi.org/10.1016/j.nlm.2007.09.010.

Horn, M., R. Baumann, J.A. Pereira, P.N. Sidiropoulos, C. Somandin, H. Welzl, C. Stendel, T. Luhmann, C. Wessig, K.V. Toyka, J.B. Relvas, J. Senderek and U. Suter (2012). Myelin is dependent on the Charcot-Marie-Tooth Type 4H disease culprit protein FRABIN/FGD4 in Schwann cells. Brain : A J. Neurology 135, 3567-3583. https://doi.org/10.1093/brain/aws275.

Kang, H. (2010). Korean Sapsaree dogs bounce back from the brink. Reuters. Thomson Reuters, London, UK.

Kiyozumi, D. and M. Ikawa (2022). Proteolysis in reproduction: lessons from gene-modified organism studies. Frontiers in endocrinology 13, 876370. https://doi.org/10.3389/fendo.2022.876370.

Martin-Vilchez, S., L. Whitmore, H. Asmussen, J. Zareno, R. Horwitz and K. Newell-Litwa (2017). RhoGTPase regulators orchestrate distinct stages of synaptic development. PLoS One 12, e0170464. https://doi.org/10.1371/journal.pone.0170464.

Martin, E.I., K.J. Ressler, E. Binder and C.B. Nemeroff (2009). The neurobiology of anxiety disorders: brain imaging, genetics, and psychoneuroendocrinology. The Psychiatric clinics of North America 32, 549-575. https://doi.org/10.1016/j.psc.2009.05.004.

Marzola, P., T. Melzer, E. Pavesi, J. Gil-Mohapel and P.S. Brocardo (2023). Exploring the role of neuroplasticity in development, aging, and neurodegeneration. Brain sciences 13. https://doi.org/10.3390/brainsci13121610.

Priel, A., J.A. Tuszynski and N.J. Woolf (2010). Neural cytoskeleton capabilities for learning and memory. J. Biological Physics 36, 3-21. https://doi.org/10.1007/s10867-009-9153-0.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, J. Maller, P. Sklar, P.I. de Bakker, M.J. Daly and P.C. Sham (2007). Plink: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559-575. https://doi.org/10.1086/519795.

R Core Team (2023). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Salmon, P.W. and I.M. Salmon (1983). Society with animals: who owns who? Psychological research into the human-pet bond in Australia. University of Pennsylvania Press, Philadelphia, PA.

Sanders, L.N., J.A. Schoenhard, M.A. Saleh, A. Mukherjee, S. Ryzhov, W.G. McMaster, W.G., K. Nolan, R.J. Gumina, T.B. Thompson, M.A. Magnuson, D.G. Harrison and A.K. Hatzopoulos (2016). BMP antagonist Gremlin 2 Limits inflammation after myocardial infarction. Circ. Res. 119, 434-449. https://doi.org/10.1161/circresaha.116.308700.

Snyder, E.M., C.L. Small, Y. Li and M.D. Griswold (2009). Regulation of gene expression by estrogen and testosterone in the proximal mouse reproductive tract. Biol. Reprod. 81, 707-716. https://doi.org/10.1095/biolreprod.109.079053.

Srivastava, S., H. Engels, I. Schanze, K. Cremer, T. Wieland, M. Menzel, M. Schubach, S. Biskup, M. Kreiss, S. Endele, T.M. Strom, D. Wieczorek, M. Zenker, S. Gupta, J. Cohen, A.M. Zink and S. Naidu (2016). Loss-of-function variants in HIVEP2 are a cause of intellectual disability. Eur. J. Hum. Genet. 24, 556-561. https://doi.org/10.1038/ejhg.2015.151.

Stafford, K. (2006). The welfare of dogs. Springer Netherlands. https://doi.org/10.1007/978-1-4020-4362-8.

Steinfeld, H., M.T. Cho, K. Retterer, R. Person, G.B. Schaefer, N. Danylchuk, S. Malik, S.B. Wechsler, P.G. Wheeler, K.L. van Gassen, P.A. Terhal, V.J. Verhoeven, M.A. van Slegtenhorst, K.G. Monaghan, L.B. Henderson and W.K. Chung (2016). Mutations in HIVEP2 are associated with developmental delay, intellectual disability, and dysmorphic features. Neurogenetics 17, 159-164. https://doi.org/10.1007/s10048-016-0479-z.

Tanwar, V., J.B. Bylund, J. Hu, J. Yan, J.M. Walthall, A. Mukherjee, W.H. Heaton, W.D. Wang, F. Potet, M. Rai, S. Kupershmidt, E.W. Knapik and A.K. Hatzopoulos (2014). Gremlin 2 promotes differentiation of embryonic stem cells to atrial fate by activation of the JNK signaling pathway. Stem Cells 32, 1774-1788. https://doi.org/10.1002/stem.1703.

Wang, C.L., F. Xiao, C.D. Wang, J.F. Zhu, C. Shen, B. Zuo, H. Wang, D. Li, X.Y. Wang, W.J. Feng, Z.K. Li, G.L. Hu, X. Zhang and X.D. Chen (2017). Gremlin2 suppression increases the BMP-2-induced osteogenesis of human bone marrow-derived mesenchymal stem cells via the BMP-2/Smad/Runx2 Signaling Pathway. J. Cellular Biochemistry 118, 286-297. https://doi.org/10.1002/jcb.25635.

Wang, G.D., W. Zhai, H.C. Yang, L. Wang, L. Zhong, Y.H. Liu, R.X. Fan, T.T. Yin, C.L. Zhu, A.D. Poyarkov, D.M. Irwin, M.K. Hytonen, H. Lohi, C.I. Wu, P. Savolainen and Y.P. Zhang (2016). Out of southern East Asia: the natural history of domestic dogs across the world. Cell Res 26, 21-33. https://doi.org/10.1038/cr.2015.147.

Wayne, R.K. and E.A. Ostrander (1999). Origin, genetic diversity, and genome structure of the domestic dog. Bioessays 21, 247-257. https://doi.org/10.1002/(SICI)1521-1878(199903)21:3%3C247::AID-BIES9%3E3.0.CO;2-Z.

Zhou, X. and M. Stephens (2012). Genome-wide efficient mixed-model analysis for association studies. Nat Genet 44, 821-824. https://doi.org/10.1038/ng.2310.