Genomic Insights into Athletic Performance and Immune Gene Diversity in Siberian Huskies

Honors Thesis
Presented to the Department of Biological Sciences
University of Maryland, Baltimore County
in Partial Fulfillment of the Requirements for the
Biological Sciences Departmental Honors in Research Program

by
Elizabeth M. Moser
May 2025

Dr. Tracy Smith

Abstract

As Arctic dog populations compete more frequently in warmer temperatures, sled dogs with physiological adaptations suited to these environments may gain a distinct performance advantage. Notably, research shows that approximately half of the racing Siberian Husky population carries recent genomic introgression from warm-adapted European breeds—an admixture that may enhance thermoregulation and improve racing performance in milder climates. To investigate the genetic basis of elite racing ability, I conducted a genome-wide association study (GWAS) comparing top-performing Siberian Huskies with documented European introgression to racing Siberians with minimal to no introgression. This analysis identified 1,304 significant single nucleotide polymorphisms (SNPs) associated with 87 genes. Enrichment analysis highlighted genes involved in neurodevelopmental regulation, cardiac muscle development, ion transport, and immune response mechanisms, suggesting a link between immune variation and performance. Given the critical role of the dog leukocyte antigen (*DLA-DRB1*) in immunity, I explored its possible association, hypothesizing that higher allelic diversity at this locus would confer an immunological advantage that might support enhanced performance. However, no significant association was found. Interestingly, the most compelling enrichment signal came from a strong overrepresentation of complement system regulators. This suggests that, rather than broad gene diversity at the *DLA-DRB1* locus, the presence of specific regulatory mechanisms may enhance immune resilience and neuroprotection during prolonged physical exertion in these dogs. Overall, identifying genetic markers associated with thermoregulation and elite racing performance has practical implications for selective breeding, but may also offer valuable translational insights into the genetics of athletic capacity and heat tolerance in humans.

Introduction

Working sled dogs in warm climates may struggle to adapt as human-accelerated climate change continues to make our planet warmer and shorten snow seasons. Natural selection is responsible for cold-climate adaptations such as thick dense coats, deep vasculature, and energy efficient fat metabolism in Arctic breeds, yet this evolutionary process is relaxed, or even reversed, in many modern European dog populations. Genomic introgression from warm-adapted European breeds may be one method of introducing heat-tolerant alleles to the Siberian Husky racing population. Comparing recently admixed racing Siberians with their relatively non-admixed counterparts may enable the identification of variants associated with heat tolerance and contribute to our understand of how these elite athletes differ from other racing Siberian Huskies. This could enable the development of DNA tests that allow owners to better understand their dog's performance and thermoregulatory characteristics. Ultimately, selective breeding for heat tolerance could assist cold-adapted breeds that have expanded to temperate regions to preserve their evolutionary niche in a rapidly changing climate.

As the planet warms, Heat Related Illnesses (HRI) are a growing concern to humans and animals. This range of progressive conditions occur when the body's heat-regulating mechanisms are unable to maintain a core temperature within the physiological normal range (Osilla et al. 2023). Intense exercise in hot and humid environments is a common cause of HRIs in humans, but sled dogs can overheat when temperatures surpass 40°F (Phillips et al. 1981). Less severe HRIs are reversible; however, hyperthermia results in cellular damage and can progress to systemic stress and dysfunction. If untreated, the onset of heat stroke may be fatal (Tripovich et al., 2023). While this poses a health threat to many dog breeds, it is a prominent issue that Siberian Huskies face when racing in warmer climates.

Investigating genes that contribute to elevated thermoregulation in humans is a field of continued interest. Previous work has identified associations between thermal and exertional heat tolerance and a heat shock protein polymorphism (Adele et al. 2022). Additional research has shown that gene expression is differentially regulated in people with or without a history of exertional heat stroke (EHS), suggesting that genetic screening may be a tool for predicting EHS risk level (Ren et al. 2019). Moreover, the varied effectiveness of heat acclimation among individuals indicates a potential genetic influence on thermal adaptation (Liu 2024). Beyond susceptibility to HRIs, heat tolerance is also a key factor in athletic performance. Efficient thermoregulation is necessary to complete high intensity exercises, and several variants associated with endurance and strength in humans have been identified (Appel et al. 2021). Therefore, it is reasonable to hypothesize that elite racing status is positively correlated with high heat tolerance.

While related research has been conducted in ovine and cattle populations (Carabaño 2019), the genetic basis of performance-related traits in cold-adapted dog breeds is largely unexplored. Prior research is limited; however, single nucleotide polymorphisms (SNPs) in the myosin heavy chain 9 gene (*MYH9*) have been significantly associated with elevated heat tolerance in Alaskan sprint racing dogs (Huson et al. 2012). Here, I tested for *MYH9* variants in Siberian Huskies and conducted a genome-wide association study (GWAS) to detect genomic regions associated with elite performance in the introgressed Siberian population. I further characterized the functional enrichment of genes within these regions and examined genetic variation in key candidates to explore their potential roles in performance-related traits.

Materials and Methods

DNA Sample Collection

A total of 17 Siberian Husky DNA samples were analyzed for SNPs in *MYH9* and 31 samples were examined for variation in the *DLA-DRB1* gene. Buccal cell samples were obtained from 33 elite sprint racing Siberian Huskies at sled dog races in the US and Canada in 2024, and 13 additional samples were used from previous genotyping projects. A subset of samples were screened for SNPs in both genes, and the total number of samples analyzed for each gene depended on the availability of genetic material and time constraints.

GWAS and Enrichment Analysis

Raw genomic data from 111 racing Siberian Huskies was used from Smith et al. 2024. The 16 elite racing Siberian Huskies with European introgression, found outside of the confidence ellipse in Figure 1 in Smith et al. 2024, were labeled as cases and the other 95 racing Siberian Huskies with minimal to no introgression were labeled as controls. The Bonferroni-corrected significance threshold was made more conservative, adjusted from $P \leq 4.0 \times 10^{-7}$ to $P \leq 9.0 \times 10^{-11}$ SNPs. The NCBI Genome Data Viewer (Sayers et al. 2025, https://www.ncbi.nlm.nih.gov/gdv/) was used to identify genes that either contained the associated SNPs or were positioned immediately upstream of them. An enrichment analysis using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (Sherman et al. 2022) determined overrepresented gene ontology terms among the GWAS-identified candidates. R Studio was used to create visual representations of the GWAS and DAVID results.

Gene Sequencing of Key Candidate Genes

DNA was extracted and purified using a QIAamp DNA Mini Kit. *MYH9* primers were designed using the reference sequences from a published canine genome and *DLA-DRB1* was amplified using previously published primer sequences (DRB1-F: 5'-CCGTCCCCACAGCA CATTTC-3' and DRB1-R: 5'-TGTGTCACACACCTCAGCACCA-3'). 0.5µL of each respective primer was combined with 12.5µL of DreamTaq™ Hot Start Green PCR master mix, 5.0µL of genomic DNA at 10ng/µL, and 6.5µL of dH_2O. Genomic DNA was doubled for samples with concentrations lower than 7ng/µL and dH_2O was added to reach 25µL total. PCR amplification of *MYH9* was as follows: initial denaturation at 95°C for 3 minutes, followed by 35 cycles of 95°C/30 s, 50°C/30 s, 72°C/30 s, and final elongation for 3 minutes at 72°C. PCR protocol for *DLA-DRB1* was as follows: initial denaturation at 94°C for 3 minutes, followed by 35 cycles of 94°C/30 s, 57°C/30 s, and 72°C/1 min, with a final extension time of 10 minutes at 72°C. After PCR confirmation using a 1% agarose gel electrophoresis, gene sequencing was performed through Azenta Life Sciences.

Genetic variants segregating with high heat tolerance were explored using Benchling sequence alignment tool (Benchling, https://www.benchling.com/alignments). Diversity classifications were assigned by examining chromatograms and determining the number of heterozygous loci in *DLA-DRB1*. A count of zero to one was classified as none/low diversity, two to eight as moderate, and nine or more as high diversity (Table 1). Categories were loosely based on human leukocyte antigen diversity classifications (Tsai et al. 2023) and it was expected that low diversity would be functionally identical to no diversity. A chi-square test was used to determine the significance of an association between diversity and elite status.

Results and Discussion

Genome-wide Association Study

I found 1,304 SNPs to be significantly associated with introgressed elite racing Siberian Huskies (Bonferroni-corrected $P \leq 4.00E\text{-}7$, Fig. 1). For the scope of this study, I used a suggested P-value of $P \leq 9.00E\text{-}11$, resulting in 200 significant variants. A total of 87 genes were observed either surrounding or upstream of associated SNPs, while the remaining SNPs were located in intergenic regions or within uncharacterized genes. Investigating previous links to heat tolerance or athletic performance in humans, dogs, or other animals revealed a wide range of functions represented within the 87 genes. These include immune mechanisms, muscle development, metabolism, nervous system development, cancer progression, cell morphology and inflammatory response, among others.

Gene enrichment analysis

A variety of genes, protein functions, and biological pathways were significantly enriched among GWAS candidates associated with elite performance in introgressed dogs (Fig. 2). This analysis revealed genetic differences between the admixed racing Siberians and other racing Siberian Huskies, particularly in genes associated with cardiac, muscular, neural, immune, sensory, and respiratory systems. Most notably, I observed a large enrichment of complement system regulators in the admixed dogs, pointing to genetic adaptations that may improve immune regulation and resilience under intense physiological stress. The most overrepresented gene ontology term demonstrating high significance was “SEZ6_CSMD_C4BPB_Regulators”. SEZ6 (Seizure-related 6 homolog) is a gene that encodes a transmembrane protein highly expressed in the brain, and is thought to play a role in synaptic formation, dendritic branching, and

neurological signaling (Hidaka et al. 2022). CSMD (CUB and Sushi Multiple Domains) encodes a protein involved in immune system regulation, and C4BPB (C4 binding protein, beta) is also a part of the complement system that regulates the classical pathway (Ermis & Bell 2022; Criado et al. 1995). These complement regulators play key roles in controlling inflammation, promoting tissue repair, and protecting neural function (Cho 2015). Pathways related to cardiac muscle development, consistent with enhanced cardiovascular capacity, and in neuromuscular signaling, that likely contributes to speed, strength, and coordination, were also overrepresented. In addition, enrichment of sensory perception genes suggest improved sensory acuity and motor control, while motile cilium assembly genes may enhance respiratory efficiency, facilitating airway clearance and oxygen delivery during exercise. Immunoglobulin domains also appeared frequently among the top GO results, suggesting an association between immune diversity and elite performance. The dog leukocyte antigen (*DLA-DRB1*) contains many genes responsible for components of the major histocompatibility complex, which serves a fundamental role in immunity. This made *DLA-DRB1* a well-qualified candidate for further exploration.

Genetic variation assessment

Purification of DNA from buccal samples and PCR using both *MYH9* and *DLA-DRB1* primers was successful (Fig. 3). *DLA-DRB1* amplification failed for one sample, shown in lane 8, and was excluded from sequencing. Within the 17 samples screened for *MYH9* variants, sequence alignment revealed several polymorphisms; however, SNPs did not consistently segregate with high or low heat tolerance, suggesting that *MYH9* is unlikely to be a major contributor to heat adaptation in this population of dogs. Diversity in *DLA-DRB1* also presented no association with elite racing status based on the 31 Siberian Husky samples screened (χ^2 =

0.13512, p = 5.991). To assess variation, the number of heterozygous loci were counted in each chromatogram (Fig. 4) and samples were assigned to three diversity groups based on that count: none/low, moderate, or high diversity (Table 1). Despite the known importance of *DLA-DRB1* in immune function, these results suggest that heterozygosity at this locus does not differentiate elite performance from the broader Siberian Husky breed in this sample. Given the strong enrichment signal observed for complement system regulators, it is possible that specific immune regulatory mechanisms—rather than overall immune gene diversity at *DLA-DRB1*—contribute more directly to enhanced immune resilience and neuroprotection during sustained physical exertion in these dogs.

Conclusion

I conducted a genome-wide association study (GWAS) to detect candidate loci linked to elite performance traits in Siberian Huskies with documented European introgression. I then used DAVID to identify overrepresented biological themes within the set of 87 associated genes, providing insight into the molecular mechanisms underlying elite status in these dogs. This analysis highlighted distinct genetic adaptations in this population of admixed racing Siberians, with significant enrichment in genes related to neural, immune, cardiac, muscular, sensory, and respiratory functions. The most striking signal was a strong overrepresentation of complement system regulators—particularly those in the SEZ6_CSMD_C4BPB_Regulators ontology—suggesting enhanced immune resilience and neuroprotection under physiological stress. These findings point to coordinated adaptations across multiple systems that likely support endurance, speed, and recovery in the high-performance introgressed racing Siberian Huskies.

Based on preliminary research and the enrichment results, I chose to sequence *MYH9* and *DLA-DRB1* across a panel of Siberian Huskies to examine variation and explore their potential roles in heat tolerance and athletic performance; neither gene was found to contain SNPs that segregated with these traits. The lack of an association between *MHY9* variants and heat tolerance contradicts previous findings in Alaskan sprint racing dogs. While this result is notable, the sample size used in this project limits the strength of these conclusions. A larger sampling population is necessary to determine the significance of these results, as well as those for *DLA-DRB1*. Additional statistical tests may also reveal select associations that align with those in Alaskan populations or are unique to racing Siberian Huskies.

Exploring the genomic foundations of elite racing status in cold-adapted dog breeds represents a largely novel research area. The findings from this project may establish a framework to inform related studies, including further examination of *DLA-DRB1* allelic diversity in relation to athletic performance, heat tolerance, and immune resilience. Linking genotypes to physiological characteristics such as VO_2 max, lactate threshold, heart rate recovery, and core body temperature during exercise may provide alternative routes for identifying genetic variants associated with heat tolerance and other performance traits. Designing an assay to screen dogs for these mutations would enable the development of a targeted gene panel, offering owners insight into data about their animal's health and athletic capacity. Furthermore, these findings may also be relevant beyond canine research, contributing to the expanding knowledge of the genetic basis of athletic attributes in humans.

Acknowledgements

Thank you to Dr. Tracy Smith for her continuous support, guidance, and valuable mentorship throughout this project. Thank you to Sophie Berger for her laboratory and bioinformatics assistance, notably for constructing the Manhattan plot and DAVID functional analysis chart. This work was funded, in part, through an Undergraduate Research Award from the UMBC Office of Undergraduate Education.

Figures

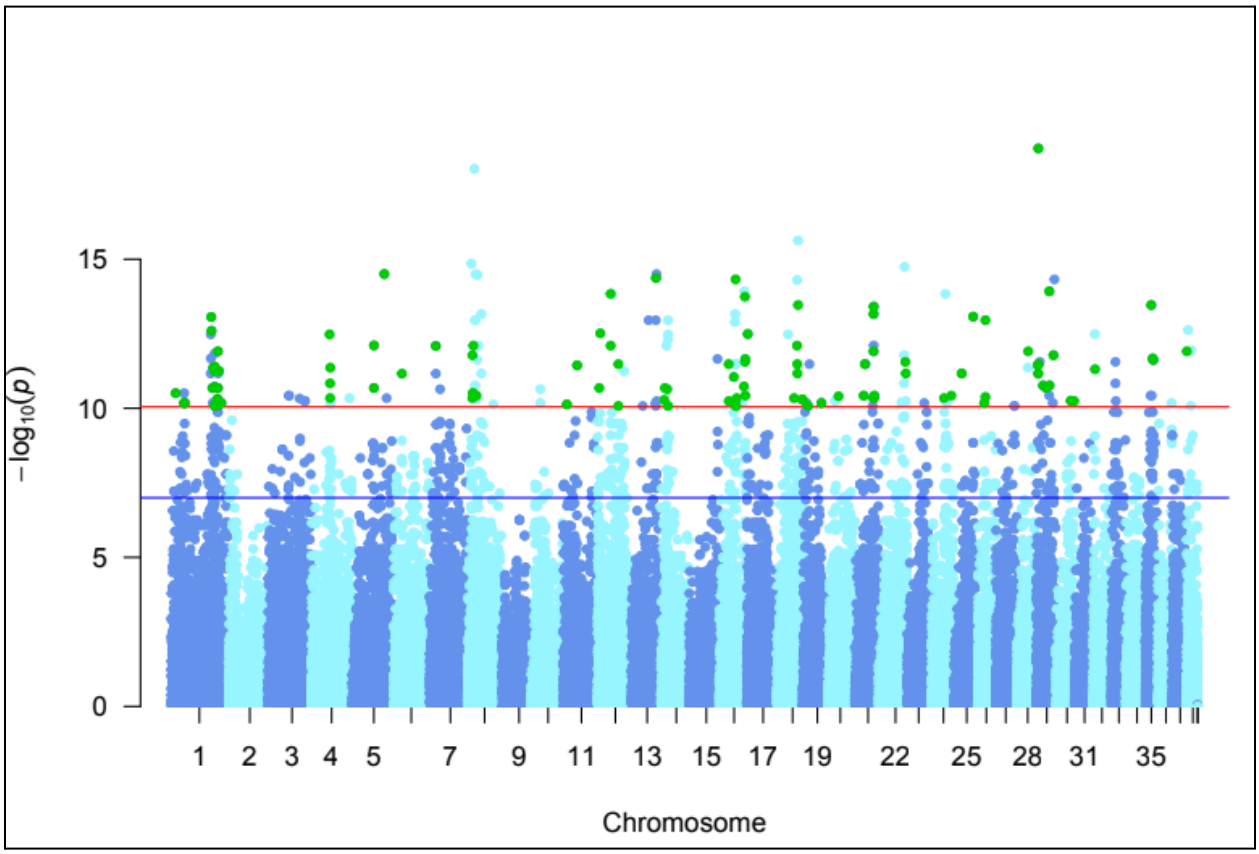

Figure 1: Manhattan plot showing 1,304 significant SNPs (Bonferroni- corrected P ≤ 4.00E-7, blue line) identified by the GWAS associated with 87 genes (highlighted in green). Non-highlighted SNPs above the suggested gene-level (P ≤ 9.00E-11, red line) are intergenic or within uncharacterized genes.

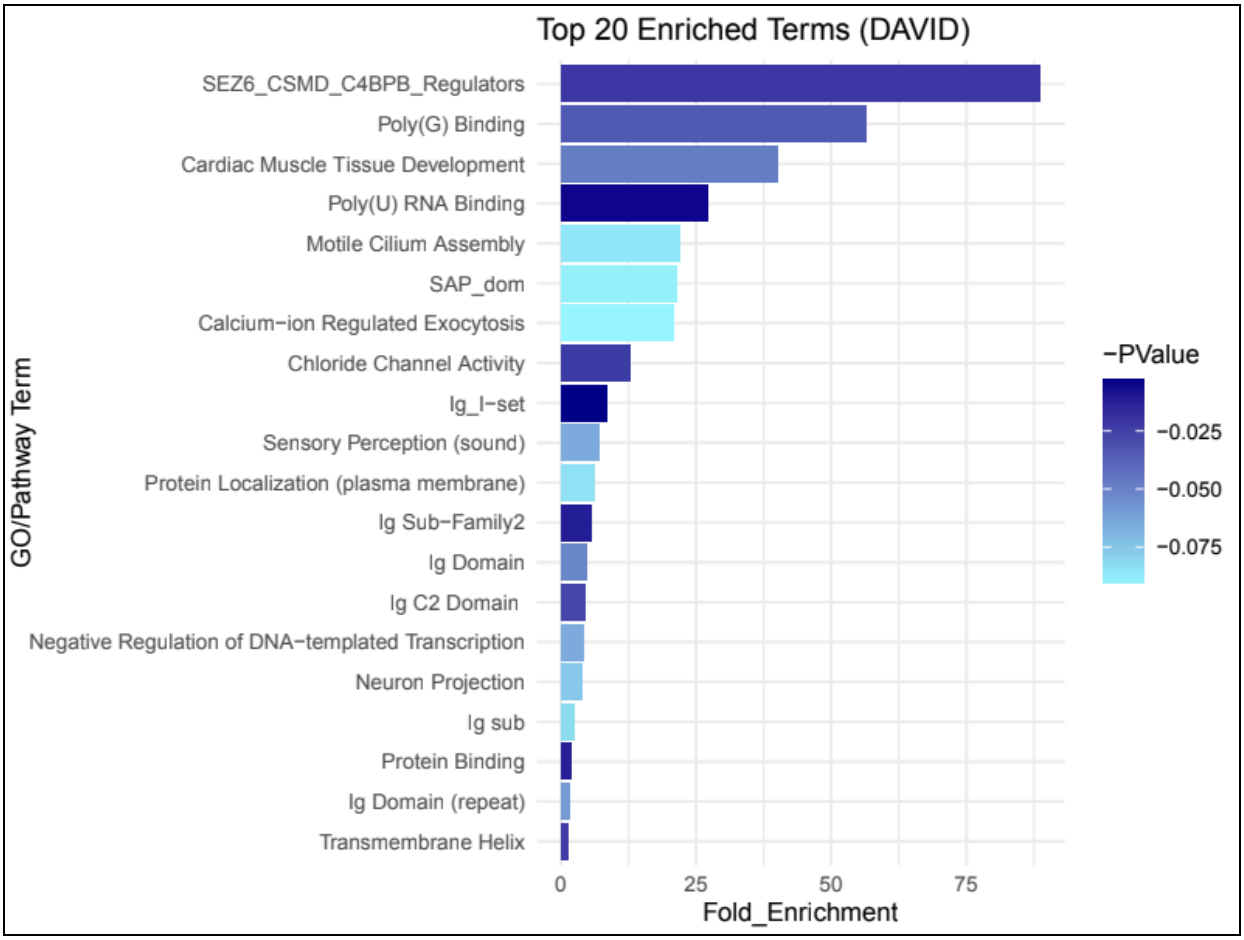

Figure 2: DAVID functional enrichment analysis results ranked by fold enrichment (X-axis). Gene ontology and pathway terms are notated along the Y-axis with a color gradient indicating P-value. Long, dark blue bars indicate a biological function that is highly overrepresented and significant within the list of 87 genes.

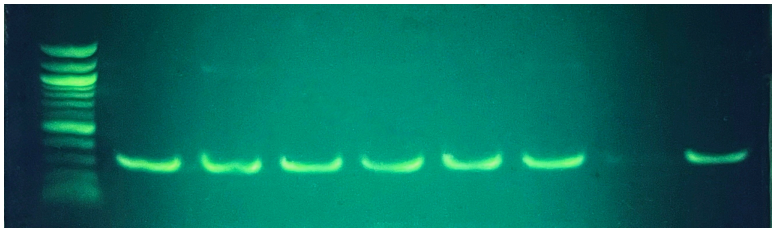

Figure 3: A confirmation gel of PCR products from 8 samples using *DLA-DRB1* primers and a 100bp ladder. All bands fall around 270bp, which is the expected size. Amplification using the sample in lane 8 was unsuccessful and was not used for further analysis. A negative control was confirmed, although not shown in this image.

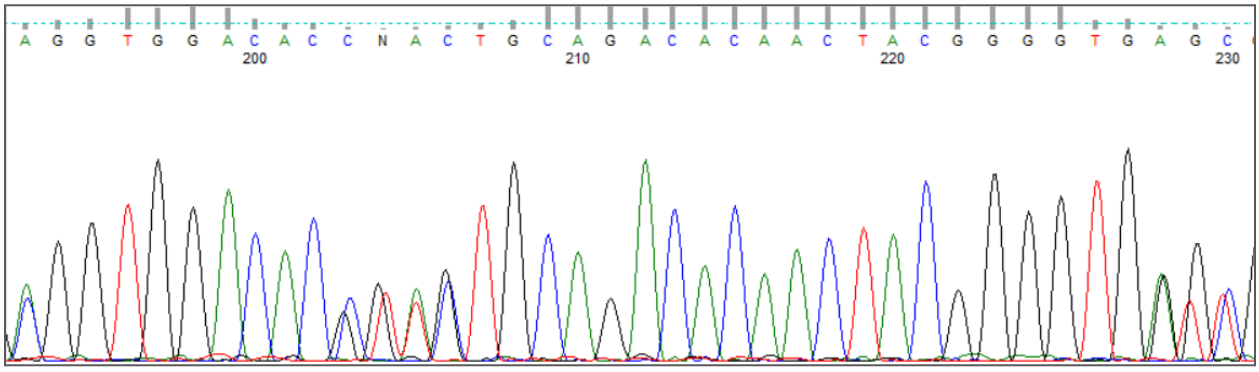

Figure 4: An example of a portion of a chromatogram showing high diversity in *DLA-DRB1*. Green peaks are designated as adenosine (A), blue as cytosine (C), black as guanine (G), and red as thymine (T). N represents an ambiguous base call. The quality of a call is represented by the height of the bars above each peak; surpassing the light blue dashed line indicates a confident call. Peaks of two overlapping colors represent heterozygosity.

Diversity level	**None or low (0-1)**	**Moderate (2-8)**	**High (9+)**
Number of total dogs	13	8	10
Number of elite dogs	9	5	7

Table 1: Diversity classification (n=31) based on number of heterozygous loci in *DLA-DRB1* of total dogs and only elite racing dogs. A chi-squared test revealed an a nonsignificant association between diversity in this gene and elite status ($\chi^2 = 0.13512$, $p = 5.991$).

Citations

Alele F.O., Otto J.R., Malau-Aduli B.S., Malau-Aduli A.E.O. Next Generation Sequencing of Genotype Variants and Genetic Association between Heat Shock Proteins HSPA1B Single Nucleotide Polymorphism at the g.31829044 Locus and Heat Tolerance: A Pilot Quasi-Experimental Study. Biomolecules. 2022 Oct 12;12(10):1465. doi: 10.3390/biom12101465.

Angles J.M., Kennedy L.J., Pedersen N.C. Frequency and distribution of alleles of canine MHC-II DLA-DQB1, DLA-DQA1 and DLA-DRB1 in 25 representative American Kennel Club breeds. Tissue Antigens. 2005 Sep;66(3):173-84. doi: 10.1111/j.1399-0039.2005.00461.x.

Appel M., Zentgraf K., Krüger K., Alack K. Effects of Genetic Variation on Endurance Performance, Muscle Strength, and Injury Susceptibility in Sports: A Systematic Review. Front Physiol. 2021 Jul 21;12:694411. doi: 10.3389/fphys.2021.694411.

Benchling [Biology Software]. (2025). Retrieved from https://benchling.com

Carabaño M.J., Ramón M., Menéndez-Buxadera A., Molina A, Díaz C. Selecting for heat tolerance. Anim Front. 2019 Jan 3;9(1):62-68. doi: 10.1093/af/vfy033.

Cho H. Complement regulation: physiology and disease relevance. Korean J Pediatr. 2015 Jul;58(7):239-44. doi: 10.3345/kjp.2015.58.7.239. Epub 2015 Jul 22.

Criado García O, Sánchez-Corral P, Rodríguez de Córdoba S. Isoforms of human C4b-binding protein. II. Differential modulation of the C4BPA and C4BPB genes by acute phase cytokines. J Immunol. 1995 Oct 15;155(8):4037-43.

Ermis Akyuz E., Bell S.M. The Diverse Role of CUB and Sushi Multiple Domains 1 (CSMD1) in Human Diseases. Genes (Basel). 2022 Dec 10;13(12):2332. doi: 10.3390/genes13122332.

Hidaka C., Mitsui S., Osako Y., Takahashi K., Tanaka K., Yuri K. Differential brain expression pattern of Sez6 alternative splicing isoform with deleted transmembrane domain. Biochem Biophys Res Commun. 2022 Dec 25;636(Pt 2):128-132. doi: 10.1016/j.bbrc.2022.10.103. Epub 2022 Nov 1.

Huang Da W., Sherman B.T., Lempicki R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57. doi: 10.1038/nprot.2008.211.

Huson H.J., vonHoldt B.M., Rimbault M., Byers A.M., Runstadler J.A., Parker H.G., Ostrander E.A. Breed-specific ancestry studies and genome-wide association analysis highlight an association between the MYH9 gene and heat tolerance in Alaskan sprint racing sled dogs. Mamm Genome. 2012 Feb;23(1-2):178-94. doi: 10.1007/s00335-011-9374-y.

Liu T. The roles of ACE I/D and ACTN3 R577X gene variants in heat acclimation. Heliyon. 2024 Jun 15;10(12):e33172. doi: 10.1016/j.heliyon.2024.e33172.

Miyamae J., Okano M., Katakura F., Kulski J.K., Moritomo T., Shiina T. Large-Scale Polymorphism Analysis of Dog Leukocyte Antigen Class I and Class II Genes (*DLA-88, DLA-12/88L* and *DLA-DRB1*) and Comparison of the Haplotype Diversity between Breeds in Japan. Cells. 2023 Mar 6;12(5):809. doi: 10.3390/cells12050809.

Osilla E.V., Marsidi J.L., Shumway K.R., Sharma S. Physiology, Temperature Regulation. 2023 Jul 30. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan–.

Phillips, C.J., Coppinger, R.P. and Schimel, D.S., 1981. Hyperthermia in running sled dogs. Journal of Applied Physiology, 51(1), pp.135-142. doi: 10.1152/jappl.1981.51.1.135.

Ren M.Q., Kazman J.B., Abraham P.A., Atias-Varon D., Heled Y., Deuster P.A. Gene expression profiling of humans under exertional heat stress: Comparisons between persons with and without exertional heat stroke. J Therm Biol. 2019 Oct;85:102423. doi: 10.1016/j.jtherbio.2019.102423.

Safra N., Pedersen N.C., Wolf Z., Johnson E.G., Liu H.W., Hughes A.M., Young A., Bannasch D.L. Expanded dog leukocyte antigen (DLA) single nucleotide polymorphism (SNP) genotyping reveals spurious class II associations. Vet J. 2011 Aug;189(2):220-6. doi: 10.1016/j.tvjl.2011.06.023.

Sayers E.W., Beck J., Bolton E.E., Brister J.R., Chan J., Connor R., Feldgarden M., Fine A.M., Funk K., Hoffman J., Kannan S., Kelly C., Klimke W., Kim S., Lathrop S., Marchler-Bauer A., Murphy T.D., O'Sullivan C., Schmieder E., Skripchenko Y., Stine A., Thibaud-Nissen F., Wang J., Ye J., Zellers E., Schneider V.A., Pruitt K.D. Database resources of the National Center for Biotechnology Information in 2025. Nucleic Acids Res. 2025 Jan 6;53(D1):D20-D29. doi: 10.1093/nar/gkae979.

Sherman B.T., Hao M., Qiu J., Jiao X., Baseler M.W., Lane H.C., Imamichi T., Chang W. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 2022 Jul 5;50(W1):W216-W221.

Smith T.A., Srikanth K., Huson H.J. Comparative Population Genomics of Arctic Sled Dogs Reveals a Deep and Complex History. Genome Biol Evol. 2024 Sep 3;16(9):evae190. doi: 10.1093/nar/gkac194.

Tripovich J.S., Wilson B., McGreevy P.D., Quain A. Incidence and risk factors of heat-related illness in dogs from New South Wales, Australia (1997-2017). Aust Vet J. 2023 Dec;101(12):490-501. doi: 10.1111/avj.13296.

Tsai Y.Y., Qu C., Bonner J.D., Sanz-Pamplona R., Lindsey S.S., Melas M., McDonnell K.J., Idos G.E., Walker C.P., Tsang K.K., Da Silva D.M., Moratalla-Navarro F., Maoz A., Rennert H.S., Kast W.M., Greenson J.K., Moreno V., Rennert G., Gruber S.B., Schmit S.L.. Heterozygote advantage at HLA class I and II loci and reduced risk of colorectal cancer. Front Immunol. 2023 Oct 24;14:1268117. doi: 10.3389/fimmu.2023.1268117.

Veríssimo, T.N.S., Saraiva, E.P., Sant'Anna, A.C., Loureiro, B.A., Nascimento, P.V.D., Lima, L.A.D.A., Maia, M.I.L., Morais, L.K.D.C., Santos, S.G.C.G.D., Lima, E.H.S.D. and Fonseca, V.D.F.C., 2023. Heat Tolerance of Siberian Husky Dogs Living in Brazil: A Case Study on the Perceptions and Attitudes of Their Owners. Animals, 13(17), p.2774. doi: 10.3390/ani13172774.