scientific reports

Check for updates

OPEN

A region on chromosome 16 is associated with Doberman Pinscher dilated cardiomyopathy

Alex Polenberg[1], Chad Grueter[2], Terry Braun[3] & Colleen Mitchell[1]✉

Doberman Pinschers are known for their increased susceptibility to dilated cardiomyopathy (DCM) relative to other domestic dogs. This makes the Doberman Pinscher a key model for gene-disease investigations. We conducted a genome-wide association study (GWAS) leveraging a database of genetic profiles obtained through collaboration with the Doberman Diversity Project (DDP). We worked in parallel to increase the depth and power of the DDP database. We exchanged direct correspondences with listed breeders and owners to establish health updates for participant dogs. In total, our study included data on 216, 184 single nucleotide polymorphisms (SNPs) in 46 cases and 3226 population control Doberman Pinschers. Using a generalized linear mixed model and saddlepoint approximation to correct for unbalanced group sizes, we identified a cluster of SNPs associated with DCM on chromosome 16.

Cardiomyopathies represent a heterogeneous family of diseases often leading to progressive heart failure with significant morbidity and mortality[1], and are classified as either primary or secondary. The former are predominantly confined to the heart muscle[1,2], and have genetic, acquired, or mixed causes, while the latter are characterized by myocardial damage resulting from systemic or multi-organ disease[2,3]. Members of this family include dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy, restrictive cardiomyopathy, and arrhythmogenic right ventricular cardiomyopathy[3].

DCM is the most common form of human cardiomyopathy, with a reported incidence of 5 to 7 in 100,000 adults and 0.57 in 100,000 children[4,5]. It is the third leading cause of heart failure in the United States behind coronary artery disease and hypertension[1,3]. DCM results in an enlarged and weakened heart, with symptoms including shortness of breath, fatigue, orthopnea, and edema[1–3]. Left ventricle dilation and contractile dysfunction characterize the disease. Consequently, the heart cannot pump sufficient blood throughout the rest of the body, further increasing susceptibility to complications like arrhythmias, blood clots, and sudden death.

Dogs share many hereditary diseases with humans[6]; as such, they provide a strong genetic model for human disease. The reduced heterogeneity amongst purebred dogs in particular provides a rich setting for causative mutation identification[7]. Moreover, DCM is reported to progress similarly between dogs and humans[8], developing across multiple phases. Beginning with an asymptomatic period before clinical signs start to manifest, there are no functional changes in cardiac tissue yet. But it is possible the underlying causes are initiating the disease[8]. Cardiovascular changes, both electrical and morphological, characterize the next stage; these changes are detectable via Holter monitors and echocardiography. Because subjects are still asymptomatic, this stage is often termed "occult DCM"[8,9]. Finally, subjects present clinical signs of heart failure during the "overt stage" of DCM.

Doberman Pinschers are one of the most susceptible breeds to DCM—the cumulative prevalence within a European cohort was estimated at 58.3%[10]. The disease's high prevalence, coupled with reduced heterogeneity within the breed, makes the Doberman Pinscher an ideal model for GWAS. Indeed, like the case for human DCM, contemporary literature suggests an important relationship between genetic predisposition and DCM pathology in Doberman Pinschers[11].

The first published GWAS for Doberman Pinscher DCM, conducted by Mausberg et al., identified an association on CFA5 in a cohort of German dogs, but did not identify specific candidate genes[12]. A subsequent GWAS by Meurs et al. reported an association with a gene encoding for PDK4 on CFA14[13]. Owczarek-Lipska et al. attempted to replicate the *PDK4* association in a European cohort, but were unsuccessful[14]. In a study of a family of Doberman Pinschers without the *PDK4* variant, Meurs et al. reported a variant in the *TTN* gene associated with the disease[15]. However, in a follow-up study of the *PDK4* and *TTN* associations, the authors

[1]Department of Mathematics, University of Iowa, Iowa City 52246, USA. [2]Department of Internal Medicine and Cardiovascular Medicine, University of Iowa, Iowa City 52246, USA. [3]Department of Biomedical Engineering, University of Iowa, Iowa City 52246, USA. ✉email: colleen-mitchell@uiowa.edu

concluded that while the *TTN* variant was most common in a cohort of 48 affected Doberman Pinschers, 6 dogs had neither the *PDK4* nor the *TTN* variants[16]. In contrast, Niskanen et al. later refined and replicated Mausberg et al.'s findings using a discovery cohort of German dogs and identified two candidate genes: *RNF207* and *PRKAA2*, known for their involvement in cardiac action potentials[17]. The study reported risk allele frequencies of 50.5% in cases ($N = 178$) and 25.9% in controls ($N = 143$) for the chr5:53,109,178G>A variant. We could not find SNP array-based frequencies for the other reported variant, chr5:60,111,983G>A. Additionally, Niskanen et al.'s RNA analyses revealed alternatively spliced transcripts in *RNF207* and *PRKAA2*, suggesting that disrupted RNA processing might contribute to the molecular mechanisms underlying DCM.

Though associations between mutations and Doberman Pinscher DCM have been demonstrated already, there remains a clear need for further study. Since there are cases of Doberman Pinscher DCM with neither the *PDK4* nor *titin* mutations (and vice versa), these genes are not solely responsible for the high prevalence within the breed. However, the replication of findings on chromosome 5 in a cohort of German dogs strengthens the argument for this locus, though additional validation is needed. Furthermore, the observed allele frequencies in affected and unaffected dogs suggest that this variant is not the sole factor driving DCM in Doberman Pinschers. Thus, deeper exploration is imperative to improve diagnosis, treatment, and breeding strategies. Studying DCM in Doberman Pinschers also offers a valuable opportunity to uncover novel insights into the disease's manifestation and progression in humans.

Methods

Doberman diversity project database

The Doberman Diversity Project (DDP) provided collaborative access to their database of health profiles for more than 3000 Doberman Pinschers. The owners of these dogs have given permission for their dogs' phenotype information from DDP as well as their genotype information from Embark to be shared with researchers. Other studies have also focused on the DDP database; for instance, Wade et al. recently investigated the population structure in the DDP database, particularly in relation to the geography and purpose for which the dogs were bred[18]. For all of our analyses, we selected 46 dogs already in the DDP database with confirmed DCM diagnoses from veterinary cardiologists for the case group. Initially, in a preliminary experiment, we selected two control groups based on the following criteria:

Control Group I:	Lived at least 9 years and had healthy Holter/ECG reports at age 7 or older.
Control Group II:	Lived at least 9 years, but either had no available health reports or the Holter/ECG was performed before age 7.

Each of these control groups yielded 27 dogs, for a case-control study consisting of 100 Doberman Pinschers. De-identified data sets for the case, control 1, and control 2 groups as well as for the entire 3272 dog population were provided by the DDP in collaboration with Embark who, at the time of the study, housed the full genotype data. However, this comparison of cases to controls lacked sufficient statistical power to draw robust conclusions (see Supplementary File 1).

To address this limitation, we aimed to increase the database's depth and statistical power for future analyses, which may extend beyond the current focus on DCM. We exchanged direct correspondences with listed breeders and owners to establish health updates for participating dogs. In our communications, we discovered at least two of our controls later developed DCM. Due to the high prevalence of the disease, it is likely that our strict control groups were further compromised. Consequently, we opted to use a population control approach, treating unphenotyped samples as controls, and utilized the entire existing database for this study. The analysis that follows compares the original 46 cases with these population controls.

Ethics statement

The data analyzed in this paper is obtained with permission of the database owner, Doberman Diversity Project. Owners that participate in the Doberman Diversity Project do so after providing consent for de-identified data to be shared for research purposes. No animals were sampled directly for this study. All methods were carried out in accordance with relevant guidelines and regulations.

Genotyping

Embark[19] performed the genotyping for 3272 participating dogs. Originally based on the Illumina CanineHD BeadChip, Embark's custom microarray contains probes for 216,184 single nucleotide polymorphisms (SNPs). For maximum coverage, SNPs were placed throughout the genome according to the CanFam3[20] reference sequence.

Statistical analysis

Data cleaning and quality control

Of the 3272 dogs, 46 were cases and 3226 were missing phenotype information. We treated those dogs with no phenotype data as controls for this study. We performed stringent preliminary quality control to assess and improve data integrity prior to association analysis. Filtering measures were conducted for both samples and SNPs to ensure robust association testing. We used KING[21] to identify duplicate samples and PLINK v1.90[22] for all remaining filtration steps.

Our analysis dealt exclusively with autosomal SNPs; that is, we removed data corresponding to sex chromosomes and mitochondrial DNA. Then we removed the case dogs necessarily duplicated in the unphenotyped group (a list of these duplicates is provided in Supplementary Table S3). After, we discarded

another 14 duplicates identified within the unphenotyped subjects. We did not filter any further on relatedness; instead, we computed and utilized the genetic relatedness matrix (GRM) to fit the association model.

Next, we removed 13,678 SNPs and 8 dogs because of missing values using both relaxed (+–geno 0.2+; +–mind 0.2+) and strict (+–geno 0.02+; +–mind 0.02+) thresholds[23]. An additional 87,611 SNPs failed to meet the minor allele frequency (MAF) cutoff of 5% and were therefore excluded from analysis (+–maf 0.05+). We removed another 5 dogs due to excessive heterozygosity and 74 SNPs which strongly deviated from Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$).

In total, 3199 dogs and 103,849 SNPs remained for association analysis.

Association testing

To account for sample structure, we used a generalized linear mixed model for the association testing. Mixed models are statistical models consisting of both fixed effects and random effects. They are most useful in settings involving clusters of related statistical units[24]. As such, mixed models are effective in preventing type I errors (i.e., false positive associations) resulting from sample structure. In these models, each SNP is tested to determine whether the variance of the genetic effect at that locus significantly deviates from zero[25]. We used SAIGE[26] for the association testing because of the included saddlepoint approximation (SPA), which adjusts test statistics to correct for unbalanced case-control ratios.

We analyzed the results of the association test graphically using both quantile–quantile (Q–Q) and Manhattan plots. Q–Q plots allow for the comparison of two probability distributions. Specifically, they quantify the extent to which the distribution of observed *P*-values deviates from the expected distribution under the null hypothesis of no association. Early deviation from the expected distribution is indicative of population stratification. We plotted each SNP against its $-\log_{10}(P)$-value to obtain a Manhattan plot. In addition, we calculated the genomic inflation factor λ_{GC} to measure systematic bias and stratification in our study.

Candidate gene identification

Associated regions identified during testing were further explored in the UCSC genome browser according to the canFam3 reference genome[27]. To identify nearby genes, we extended the analysis by including 100kb flanking regions on either side of the identified loci.

Results

Results from our analysis are given in Table 1 and Fig. 1. Specifically, details for the 10 most significant markers are provided in Table 1, along with locations, raw and adjusted *P*-values, reference and alternate alleles, and allele frequencies in cases and controls. We focused subsequent analysis on chromosome 16 because it provided the strongest signal.

Figure 1 details a more complete view of the results, including Q–Q and Manhattan plots in the top panels. Note the Q–Q plot in the top-left panel of Fig. 1 confirms our analysis accurately controlled for population stratification ($\lambda_{GC} \approx 1.058$). The most compelling signal in the top-right panel of Fig. 1 implicates a cluster of loci near CFA16:31,161,738 bp–31,408,726 bp. The leading marker in this region—SNP BICF2G630113368—has saddlepoint adjusted $P_{spa} = 3.687 \times 10^{-6}$ ($P_{raw} = 4.02 \times 10^{-7}$). We identified this variant in approximately 60.8% of cases and in nearly 30.3% of population controls.

The bottom panel of Fig. 1 offers a closer view of this region on chromosome 16. Specifically, we overlaid nearby genes relative to the reported SNPs in the UCSC genome browser according to the canFam3 reference genome. We included flanking regions of 100 kb on either end of the signaled region and noted a total of five nearby genes: *DUSP26* (31,115,381–31,118,532 bp), *RNF122* (31,135,967–31,144,921 bp), *TTI2* (31,161,076–31,168,142 bp), *MAK16* (31,166,051–31,179,981 bp), and *FUT10* (31,194,629–31,266,964 bp).

Chr	Pos. (bp)	SNP	Ref	Alt	P_{spa}	P_{raw}	AF_{case}	AF_{ctrl}
16	31,408,726	BICF2G630113368	A	G	3.7×10^{-6}	4×10^{-7}	0.39	0.17
16	31,399,042	BICF2S2344623	C	A	7.2×10^{-6}	1×10^{-6}	0.38	0.16
16	31,324,250	BICF2P1079773	G	A	8.8×10^{-6}	1.3×10^{-6}	0.38	0.16
16	31,310,441	BICF2S23048638	C	A	9.6×10^{-6}	1.5×10^{-6}	0.38	0.17
16	31,201,462	chr16_31201462	G	A	1.3×10^{-5}	2.4×10^{-6}	0.38	0.17
16	31,161,738	chr16_31161738	G	C	1.4×10^{-5}	2.4×10^{-6}	0.37	0.16
9	21,398,863	chr9_21398863	A	G	4.8×10^{-5}	2.3×10^{-5}	0.59	0.25
17	19,893,878	BICF2S22910212	G	A	4.9×10^{-5}	5.1×10^{-6}	0.29	0.096
32	7,105,969	BICF2S22942598	A	G	5.6×10^{-5}	2.8×10^{-5}	0.59	0.25
17	15,791,076	BICF2S22925658	G	A	5.7×10^{-5}	3×10^{-5}	0.6	0.26

Table 1. The 10 most significant SNPs from our analysis sorted by P_{spa}. Note the cluster of SNPs on CFA16 represent the only signals with P_{spa} on the order of 10^{-6}. SNP positions aligned according to the canFam3 reference panel.

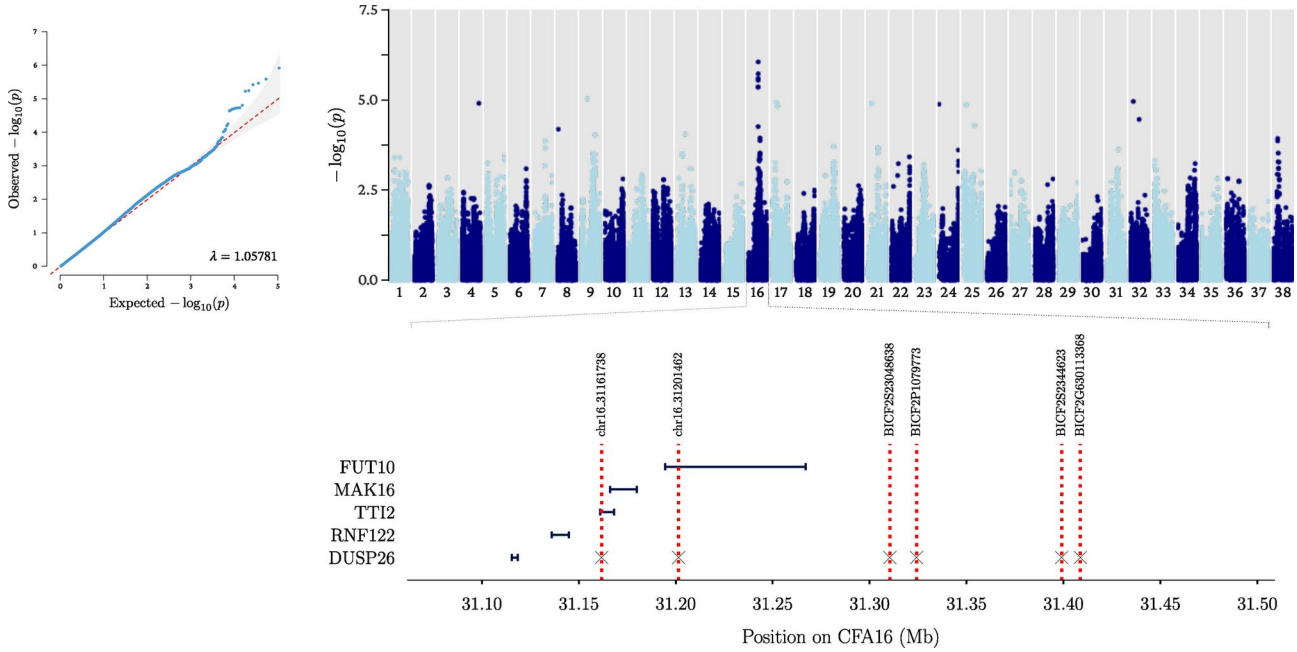

Fig. 1. Results from case group vs. population controls. The Q–Q plot in the top left indicates adequate control for population stratification ($\lambda_{GC} \approx 1.05$). Manhattan plot reveals a signal on chromosome 16 ($P_{spa} = 3.687 \times 10^{-6}$, $P_{raw} = 4.02 \times 10^{-7}$). Bottom panel highlights the window around this signal, including 100kb bands on either end. Signaled SNPs are shown in red. Genes within this window according to the canFam3 reference genome include: *DUSP26* (31,115,381–31,118,532 bp), *RNF122* (31,135,967–31,144,921 bp), *TTI2* (31,161,076–31,168,142 bp), *MAK16* (31,166,051–31,179,981 bp), and *FUT10* (31,194,629–31,266,964 bp).

Discussion

We identified a region of suggestive significance near CFA16:31,161,738–31,408,726 bp associated with DCM in a cohort of 3272 Doberman Pinschers. To the best of our knowledge, no prior associations between Doberman Pinscher DCM and CFA16 have been demonstrated. Importantly, we did not replicate the associations previously identified on chromosomes 14[13], 36[15], or 5[12,17]. Moreover, we presented a unique analytical approach, utilizing a broad population control group for exploratory analysis. This method offers a pragmatic means of uncovering potential associations that may have otherwise remained undetected.

Using this approach, we identified 5 candidate genes for further analysis (Fig. 1). A review of the existing literature did not expose a connection between cardiovascular function or disorder and the *RNF122* and *MAK16* genes. Amongst the remaining candidates, *DUSP26* may offer particular insight relative to the others. In a 2021 study, Zhao et al. concluded DUSP26 protected against pressure overload induced cardiac hypertrophy in mice. Specifically, the authors noted "cardiac-specific overexpression of DUSP26 mice showed attenuated cardiac hypertrophy and fibrosis, while deficiency of DUSP26 in mouse hearts resulted in increased cardiac hypertrophy and deteriorated cardiac function"[28].

Continuing our review of existing literature surrounding these genes, the TTT complex—formed when TTI2 interacts with TELO2 and TTI1—also plays a role in related conditions. Notably, congenital heart disease has been observed in patients with mutations in components of this complex. These interactions are essential for the maturation of phosphatidylinositol 3-kinase-related protein kinases[29].

Similarly, genes involved in cellular responses under stress conditions, like *FUT10*, may also be relevant. FUT10 is a regulator of cell proliferation and is upregulated in ischemic heart disease. Mulari et al. demonstrated that, in patients undergoing bypass surgery, an increased expression of *FUT10* is correlated with coronary artery obstruction complexity and with NT-proBNP levels[30].

We initially restricted our search to the associated region itself and the 100 kb bands on either end of it. But extending our window of investigation out to 500 kb reveals one additional candidate, Neuroregulin-1 (*NRG1*). NRG1 is secreted by endothelial vascular cells in response to injury or stress and activates ErbB family receptors. In adults, the NRG1/ErbB pathway is viewed as compensatory and attenuates cardiac remodeling, fibrosis, and inflammation[31,32]. Additionally, loss of function of NRG1, ErbB2 or ErbB4 in mouse models results in a dilated cardiomyopathy phenotype[33].

While this region represents a novel association, we emphasize the need for replication studies and further analysis to corroborate our findings. The candidate genes introduced require in-depth investigation to establish their role in Doberman Pinscher DCM pathogenesis. Gene expression studies, for example, may help ascertain their true relevance.

It is crucial to acknowledge the inherent limitations of this study. Primarily, our broad definition of a "control" group, necessitated by our analytical approach of utilizing unphenotyped samples, introduces a degree

of uncertainty that must be considered. In addition, the relatively small pool of 46 cases restricts alternative approaches that might be considered with a larger case cohort. For instance, with an expanded case group, fitting a Bayesian sparse linear mixed model to predict phenotypes may offer more robust insights[34]. Moreover, variability in diagnostic criteria used by different veterinarians could affect the consistency of our case data. Finally, although we controlled for population stratification to the best of our ability, the de-identification process prevented us from accessing information on the ancestral origins of our case group, which may limit the precision of our findings. These limitations underscore the importance of validation studies through independent cohorts.

In conclusion, we have identified a region containing five genes on CFA16 that may be associated with DCM in the Doberman Pinscher. Future analysis, including replication and functional studies in independent cohorts, is essential to assess the strength and overall relevance of these candidates. Importantly, this research represents another step in bridging the gap between canine and human cardiac genetics, offering potential insights that could inform and improve our understanding of DCM in both species.

Data availability

The data that support the findings of this study are available from the Doberman Diversity Project, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available upon reasonable request from the Doberman Diversity Project. Requests can be made online at https://www.dobermandiversityproject.org.

Received: 28 May 2024; Accepted: 31 October 2024
Published online: 08 November 2024

References

1. Wexler, R., Elton, T., Pleister, A. & Feldman, D. Cardiomyopathy: an overview. *Am. Fam. Physician* **79**, 778 (2009).
2. Sisakian, H. Cardiomyopathies: Evolution of pathogenesis concepts and potential for new therapies. *World J. Cardiol.* **6**, 478 (2014).
3. Maron, B. J. et al. Contemporary definitions and classification of the cardiomyopathies: An American heart association scientific statement from the council on clinical cardiology, heart failure and transplantation committee; quality of care and outcomes research and functional genomics and translational biology interdisciplinary working groups; and council on epidemiology and prevention. *Circulation* **113**, 1807–1816 (2006).
4. Dec, G. W. & Fuster, V. Idiopathic dilated cardiomyopathy. *N. Engl. J. Med.* **331**, 1564–1575 (1994).
5. Towbin, J. A. et al. Incidence, causes, and outcomes of dilated cardiomyopathy in children. *JAMA* **296**, 1867–1876 (2006).
6. Shearin, A. L. & Ostrander, E. A. Leading the way: canine models of genomics and disease. *Dis. Models Mech.* **3**, 27–34 (2010).
7. Karlsson, E. K. & Lindblad-Toh, K. Leader of the pack: gene mapping in dogs and other model organisms. *Nat. Rev. Genet.* **9**, 713–725 (2008).
8. Simpson, S. *et al.* Genetics of human and canine dilated cardiomyopathy. *Int. J. Genom.* **2015** (2015).
9. Dukes-McEwan, J. et al. Proposed guidelines for the diagnosis of canine idiopathic dilated cardiomyopathy. *J. Vet. Cardiol.* **5**, 7–19 (2003).
10. Wess, G. et al. Prevalence of dilated cardiomyopathy in doberman pinschers in various age groups. *J. Vet. Intern. Med.* **24**, 533–538 (2010).
11. Tidholm, A., Häggström, J., Borgarelli, M. & Tarducci, A. Canine idiopathic dilated cardiomyopathy. Part I: Aetiology, clinical characteristics, epidemiology and pathology. *Vet. J.* **162**, 92–107 (2001).
12. Mausberg, T.-B. et al. A locus on chromosome 5 is associated with dilated cardiomyopathy in doberman pinschers. *PLoS One* **6**, e20042 (2011).
13. Meurs, K. M. et al. A splice site mutation in a gene encoding for pdk4, a mitochondrial protein, is associated with the development of dilated cardiomyopathy in the doberman pinscher. *Hum. Genet.* **131**, 1319–1325 (2012).
14. Owczarek-Lipska, M. et al. A 16-bp deletion in the canine pdk4 gene is not associated with dilated cardiomyopathy in a european cohort of doberman pinschers. *Anim. Genet.* **44**, 239–239 (2012).
15. Meurs, K. M. et al. A missense variant in the titin gene in doberman pinscher dogs with familial dilated cardiomyopathy and sudden cardiac death. *Hum. Genet.* **138**, 515–524 (2019).
16. Meurs, K. M. et al. Assessment of pdk4 and ttn gene variants in 48 doberman pinschers with dilated cardiomyopathy. *J. Am. Vet. Med. Assoc.* **257**, 1041–1044 (2020).
17. Niskanen, J. E. et al. Identification of novel genetic risk factors of dilated cardiomyopathy: from canine to human. *Genome Med.* **15**, 73 (2023).
18. Wade, C. M., Nuttall, R. & Liu, S. Comprehensive analysis of geographic and breed-purpose influences on genetic diversity and inherited disease risk in the doberman dog breed. *Canine Med. Genet.* **10**, 7 (2023).
19. Embark. Embark dog DNA test: Most accurate breed identification (2021).
20. Hoeppner, M. P. et al. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One* **9**, e91172 (2014).
21. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
22. Chang, C. C. et al. Second-generation plink: rising to the challenge of larger and richer datasets. *Gigascience* **4**, s13742-015 (2015).
23. Marees, A. T. et al. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *Int. J. Methods Psychiatr. Res.* **27**, e1608 (2018).
24. Gomes, D. G. Should i use fixed effects or random effects when i have fewer than five levels of a grouping factor in a mixed-effects model?. *PeerJ* **10**, e12794 (2022).
25. Li, G. & Zhu, H. Genetic studies: the linear mixed models in genome-wide association studies. *Open Bioinform. J.* **7** (2013).
26. Zhou, W. et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* **50**, 1335–1341 (2018).
27. Kent, W. J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
28. Zhao, J. et al. Dual-specificity phosphatase 26 protects against cardiac hypertrophy through tak1. *J. Am. Heart Assoc.* **10**, e014311 (2021).
29. You, J. et al. A syndromic intellectual disability disorder caused by variants in telo2, a gene encoding a component of the ttt complex. *Am. J. Hum. Genet.* **98**, 909–918 (2016).
30. Mulari, S. et al. Ischemic heart disease selectively modifies the right atrial appendage transcriptome. *Front. Cardiovasc. Med.* **8**, 728198 (2021).
31. De Keulenaer, G. W. et al. Mechanisms of the multitasking endothelial protein nrg-1 as a compensatory factor during chronic heart failure. *Circ. Heart Fail.* **12**, e006288 (2019).
32. Wang, Y. et al. Neuregulin-1, a potential therapeutic target for cardiac repair. *Front. Pharmacol.* **13**, 945206 (2022).

33. Geissler, A., Ryzhov, S. & Sawyer, D. B. Neuregulins: protective and reparative growth factors in multiple forms of cardiovascular disease. *Clin. Sci.* **134**, 2623–2643 (2020).
34. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012).

Acknowledgements
The authors acknowledge and thank the many breeders and owners who participated by sharing the data analyzed throughout. We thank the Doberman Pinscher Club of America for their support and the Doberman Diversity Project for providing access to their database of participant data.

Author contributions
A.P. and C.M. designed the study and analyzed the results. T.B. lent expertise to advise on all aspects related to computational genomics. C.M. facilitated initial communications with the Doberman Diversity Project; A.P. and C.M. coordinated with study participants to incorporate updated data in the DDP database. C.G. provided insight on genetics of cardiac disease and helped edit manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-78511-2.

Correspondence and requests for materials should be addressed to C.M.

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

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