Sidekli, O., Smedley, T., Raffle, J., Crean, A., Hinsu, A., Fuentes, V., Connolly, D., Psifidi, A. (2026) Integration of GWAS, long and short reads sequencing to identify genetic variants associated with hypertrophic cardiomyopathy in cats, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 24202: https://doi.org/10.31274/wcgalp.24202 ©2026 *Sidekli, et al.* CC BY-NC-ND

Integration of GWAS, long and short reads sequencing to identify genetic variants associated with hypertrophic cardiomyopathy in cats

O. Sidekli[1], T.A. Smedley[1], J. Raffle[1], A.M. Crean[1], A. Hinsu[1], V.L. Fuentes[1], D.J. Connolly[1], and A. Psifidi[*1]

[1]Royal Veterinary College, Hatfield, UK; [*]*Corresponding author:* apsifidi@rvc.ac.uk

Introduction

Hypertrophic cardiomyopathy (HCM) represents the most common inherited cardiovascular disease in both humans and cats. The estimated prevalence of HCM is approximately 0.2% in humans and up to 14.3% in cats, with disease incidence increasing markedly with age in feline populations (Mckenna et al., 2021; Paige et al., 2009). HCM is characterised by left ventricular hypertrophy in both species (Paige et al., 2009). HCM is a multifactorial disease influenced by both genetic and non-genetic factors, including breed predisposition, advancing age, male sex, and obesity (Joshua et al., 2023). These epidemiological patterns closely mirror those observed in human HCM, supporting the use of feline HCM as a naturally occurring translational model for studying disease mechanisms (Joshua et al., 2023; Maron et al., 2014). Despite extensive genetic characterization of human HCM, with over 1,500 identified causative variants (Maron et al., 2014), the genetic basis of feline HCM remains poorly defined. To date, only five genetic variants have been reported in cats, of which three are confirmed as causal (Meurs et al., 2007; Schipper et al., 2019; McNamara et al., 2020). In humans, genome-wide association studies (GWAS) and Whole Genome Sequencing (WGS) studies have revealed that HCM risk arises from a combination of rare, high-penetrance variants and more common variants with modest additive effects (Ning et al., 2023). However, there is a scarcity of relevant studies for feline HCM.

The present study builds upon a previous multi-breed GWAS in a cohort phenotyped pedigree and non-pedigree (British Shorthair (BSH), Birman, Sphynx, Bengal, and Domestic Shorthair (DSH)) cats. To refine GWAS-identified loci, WGS was performed in cats from the same breeds, targeting both GWAS candidate regions and 84 cardiomyopathy-related genes according to the human and feline literature. In addition, Oxford Nanopore Technologies (ONT) long-read sequencing was performed in a subset of BSH and Bengal cats to further characterise structural variants.

Materials & Methods

The study was approved by the Royal Veterinary College ethical review committee (URN 2019 1942-3, 2023 2192-3 and 2021 2051-3).

Phenotypic Classification. Echocardiographic phenotyping was performed in non-sedated cats using GE Vivid E9 or Vivid I systems by board-certified veterinary cardiologists or supervised residents. Measurements included interventricular septal thickness (IVSd) and left ventricular free wall thickness at end-diastole (LVFWd), averaged over three cardiac cycles. Left ventricular wall thickness (LVWT) was defined as the maximal measurement of either IVSd or LVFWd. Left atrial size and left ventricular systolic function were assessed using standard echocardiographic parameters. HCM cases were defined by LVWT $\geq$ 6 mm without secondary hypertrophy, while controls had LVWT $\leq$ 5 mm, normal cardiac findings, and were $\geq$ 9 years old.

GWAS Dataset. Genotyping data used in this study were generated using the Illumina Infinium Feline 63K iSelect DNA array in a multi-breed cohort of 429 cats. Genotype data underwent standard quality control using PLINK (Purcell and Chang, 2023), including filtering of samples and SNPs based on genotyping rate (%10), minor allele frequency (<0.05), and Hardy–Weinberg equilibrium ($P<10^{-6}$). GWAS was performed using a univariate linear mixed model implemented in GEMMA (Zhou and Stephens, 2014), incorporating a genomic relationship matrix to account for population structure and relatedness across breeds. Genome-wide significance was defined using a Bonferroni-corrected threshold ($P < 0.05 / N$, where N is the number of markers tested), and a suggestive genome-wide threshold corresponding to one expected false-positive per genome scan was also applied.

WGS -Short reads. WGS was performed in 53 cats, including HCM cases and controls from BSH (8 cases, 4 controls), DSH (6, 6), Birman (6, 6), Sphynx (6, 6), and Bengal (4, 1). Genomic DNA was extracted using the Qiagen DNeasy kit. Library preparation and sequencing were outsourced to Novogene (Cambridge, UK). Libraries were generated using standard Illumina protocols and sequenced on the NovaSeq 6000 platform to a mean depth of ~30. Raw reads were quality-checked using FastQC (Andrew, 2010), followed by adapter trimming and quality filtering with Trimmomatic (Blger, Lohse and Usadel, 2014). Filtered reads were aligned to the Felis catus 9.0 reference genome using BWA-MEM2 (Vasimuddin e al., 2019). Variant calling was performed with GATK Best Practices (Van der Auwera and O'Connor, 2020). Variants were annotated using Ensembl Variant Effect Predictor (VEP; McLaren et al., 2016). Candidate regions were defined based on previous GWAS significant associations using 1 Mb windows upstream and downstream of the significant markers. In addition, the sequencing of feline orthologues of 84 cardiomyopathy-associated genes (Pua et al., 2016) was interrogated. Variants within these regions were extracted and prioritised for downstream analysis. Case–control association testing using the chi-square test was performed in PLINK (Purcell and Chang, 2023), with multiple testing correction applied.

WGS -Long reads. Long-read sequencing was conducted in a subset of eight cats comprising BSH and Bengal breeds (2 HCM cases and 2 controls per breed). High molecular Weight DNA (New England Biolabs, UK) was extracted from cats using residual blood samples. Library preparation was performed following the manufacturer's instructions. Sequencing was carried out on the Oxford Nanopore PromethION platform, achieving approximately 40× genome coverage per sample. Long-read sequences were aligned to the Felis catus 9.0 reference genome using a Nextflow-based alignment (Talenti and Prendergast, 2021) workflow. Structural variants (SVs) were identified from ONT alignments using Sniffles2 (Smolka et al., 2014). SV calls were generated per individual and subsequently aggregated within each breed for case–control comparisons. As an initial prioritization step, SVs were filtered based on a minimum variant length of ≥1 kb and differences in SV frequencies between HCM cases and controls were assessed descriptively based on carrier counts within each breed. Genes disrupted by case-enriched exonic SVs were summarised per breed and subjected to exploratory functional enrichment analysis using Panther.

Results & Discussion

This study integrates GWAS, short-read WGS, and long-read SV analyses to study the genetic architecture of feline HCM across multiple breeds. GWAS identified 35 genomic loci with significant or suggestive associations with HCM. Within these regions, WGS of the BSH population revealed ~545,000 variants, of which 1,612 remained significant after multiple-testing correction (Padj < 0.05). The majority of these variants were more frequent in controls, indicating a predominantly control-enriched signal, while only 17 variants showed

higher allele frequencies in HCM cases. These case-enriched variants were all non-coding, mapping to intronic (overlapping *AGAP1, SPACA3, ARHGAP28, ENSFCAG00000005384,* and *ENSFCAG00000047115*) or intergenic regions. Notably, *ARHGAP28* gene has been linked to cardiac disease–associated remodelling in human (Simonson et al., 2023), supporting a potential regulatory involvement of this locus in cardiac pathology. Mouse data also implicates this gene in cardiac structural abnormalities (Chi et al., 2025). Across all significant variants, only two protein-altering variants were detected, comprising one frameshift and one missense variant in *SH3BP4* gene (Table 1). The intronic association signal within *NFIA* in control group, a transcription factor with reported myocardial expression (Bastian et al., 2021), further supports a regulatory disease model. Pathway enrichment analysis highlighted processes related to cardiac development and cellular signalling.

Table 1. Nine representative significant variants in the BSH population. AF_AFF and AF_UNAFF denote allele frequencies in affected and unaffected individuals, respectively.

Chr:Pos	Type	Ref	Alt	Chisq	Pvalue	Prev	Gene	Class
A1:55463619	SNP	C	T	12.63	0.0003793	Control	*ENSFCAG00000045833*	intronic
D3:38947691	SNP	A	G	12.2	0.0004775	Control	-	intergenic
C1:50587131	SNP	*	G	11.92	0.0005563	Control	*NFIA*	intronic
C1:217324691	SNP	T	A	14	0.0001828	Control	*SH3BP4*	missense
C1:50613268	Indel	GTTAC	G	10.37	0.00128	Control	*NA*	frameshift
E1:20967800	SNP	C	G	10.91	0.0009569	Case	*SPACA3*	intronic
D3:39570298	SNP	G	C	10.48	0.001209	Case	*ARHGAP28*	intronic
D3:37878685	SNP	A	T	10.48	0.001209	Case	-	intergenic
B1:24760910	SNP	C	G	10.48	0.001209	Case	*ENSFCAG00000047115*	intronic

[1]Chr:Pos, chromosome and genomic position; [2]SNP, single nucleotide polymorphism; [3]Ref, reference allele; [4]Alt, alternative allele; [5]Chisq, chi-square statistic; [6]Prev, prevalence group

Targeted analysis of 84 cardiomyopathy-associated genes identified 105,978 variants in the DSH population, including 22 predicted high-impact and 323 missense variants. Although 11 high-impact variants were more frequent in HCM cases, none reached statistical significance. In contrast, nine missense variants with a significantly ($P<0.05$) higher frequency in HCM affected cats, mapping to *TTN*, a novel gene overlapping *TTN* locus *(ENSFCAG00000049768), SCN5A, GAA,* and *LAMA2*. Notably, the variant in *LAMA2* gene occurred exclusively in HCM cases, consistent with evidence linking laminin-α2 dysfunction to cardiac abnormalities in humans (Wang et al., 2022). Variants in *TTN* and *SCN5A* genes implicate cytoskeletal and ion channel pathways similar to those described in human HCM (Herman et al., 2022; Li et al., 2018). Two missense variants within *CSRP3* and *PRDM16* had a significantly higher frequency in controls, suggesting a possible protective or modifier effects.

In the BSH population, 49,838 single-nucleotide variants were identified, the majority of which were intronic. Functional annotation identified 1,302 missense variants, 385 synonymous variants, and 289 predicted high-impact variants. Chi-square testing revealed 205 variants with significant case–control differences after multiple-testing correction. Most significant variants were intronic and enriched in controls, particularly within *MYH7* and *LAMA2*, suggesting tolerance of common non-coding variation in unaffected cats. In contrast, only three variants showed higher allele frequencies in HCM cases, all of which were intronic variants within the *LAMA2* gene. Together, these findings indicate that variation within the predefined HCM candidate genes in BSH cats is largely non-coding, with limited evidence for direct protein-

altering effects contributing to disease status. Functional enrichment analysis of implicated genes highlighted metabolic and signalling-related biological processes.
Oxford Nanopore long-read sequencing identified 3,820 SVs in Bengal cats and 2,741 SVs in BSH cats, predominantly deletions and insertions. Among SVs ≥1 kb, case-enriched exonic SVs mapped to 591 genes in Bengal cats and 83 genes in BSH cats, indicating a higher burden of gene-overlapping SVs in the Bengal population. An inversion affecting the *SLC6A7* gene was identified in HCM cases of both breeds. Serotonin transporter genes have been implicated in human cardiovascular disease (Fumeron et al., 2002), with functional variation in *SLC6A4* associated with increased myocardial infarction risk; therefore, serotonin-related pathways may also be involved in cardiomyopathy. Functional enrichment analysis of genes affected by case-enriched SVs highlighted breed-specific biological themes, with BSH cats showing enrichment of metabolic and cellular-homeostasis processes (including calcium ion homeostasis and transmembrane transport), whereas Bengal cats were enriched for cell-cycle regulation and inflammatory signalling (e.g., G2/M transition and TNF-mediated pathways).

Conclusions

According to the results of this study, HCM susceptibility is a complex polygenic trait with a different genetic architecture depending on cat breed. Single nucleotide polymorphisms and structural variation seem to play a role. Ongoing, integration of WGS with total RNA-Seq data will further elucidate the underlying mechanisms of HCM susceptibility.

Acknowledgements

This study was supported by funding from Petplan Charitable Trust

References

Andrews, S. (2010) Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. visited on: 23-09-2025.
Bastian, F.B. et al. (2021) Nucleic Acids Res 49:831–847. doi: 10.1093/nar/gkaa793
Bolger, A.M., Lohse, M., Usadel, B. (2014) Bioinformatics, 30(15): 2114–2120.doi: 10.1093/bioinformatics/btu170
Chi, D. et al. (2025) Genes Dis, 7;12(4):101404. doi: 10.1016/j.gendis.2024.101404
Fumeron, F. et al (2002) J Am Heart Assoc, 105:25. doi.org/10.1161/01.CIR.0000022603.92986.99
Geraldine van der A., and O'Connor B.D. (2020) Genomics in the Cloud : Using Docker, GATK, and WDL in Terra. First edition. Sebastopol, CA: O'Reilly Media, Print.
Herman, D.S. et al. (2012) N Engl J Med 366(7). doi: 10.1056/NEJMoa1110186.
Joshua J., Caswell J., O'Sullivan M.L., Wood G., Fonfara S. (2023) PLOS ONE 18(3). doi: 10.1371/journal.pone.0283244
Li, W. et al. (2018) Frontiers in Physiology, 9(10):1372. doi: 10.3389/fphys.2018.01372
Luis Fuentes, V. et al. (2020) J Vet Intern Med 34(3):1062–1077. doi: 10.1111/jvim.15745
Maron, B.J. et al. (2014) J AM Coll Cardiol 64(1): 83–99. doi.org/10.1016/j.jacc.2014.05.003
McKenna W.J., Judge D.P. (2021) Nat Rev Cardiol. 18(1):22-36. doi: 10.1038/s41569-020-0428-2
McLaren, W. et al. (2016) Genome Biology 17:1–14. doi: 10.1186/s13059-016-0974-4
McNamara, J.W. et al. (2020) Frontiers in Physiology 11. doi: 10.3389/fphys.2020.608473
Meurs, K.M. et al. (2007) Genomics, 90(2):261–264. doi: 10.1016/j.ygeno.2007.04.007
Ning, C. et al. (2023) Nat Commun 14:11-15. doi: 10.1038/s41467-023-43771-5
Paige, C.F.E. et al. (2009) J Am Vet Assoc 234(11):1398-403. doi: 10.2460/javma.234.11.1398
Pua, C.J. et al. (2016) J Cardiovasc Transl Res 9(1):3–11. doi: 10.1007/s12265-016-9673-5
Purcell, S., Chang, C. (2023) PLINK version 1.90 COG-genomics. Available at: www.cog-genomics.org/plink/1.9/.
Simonson B. et al (2023) Cell Rep 28;42(2):112086. doi: 10.1016/j.celrep.2023.112086.
Smolka, M., Paulin, L.F., Grochowski, C.M. et al. (2024) Nat Biotech 42, 1571 doi:10.1038/s41587-023-02024-y
Talenti A., Prendergast J. (2021) Genome Biol Evol 13(9). doi: 10.1093/gbe/evab183
Vasimuddin, M., Misra, S., Li, H., & Aluru, S. (2019) IEEE 33rd IPDPS, Rio de Janeiro, Brazil 314-324.
Wang, Y. et al. (2022) Frontiers in Medicine, 9: 922347. doi: 10.3389/fmed.2022.922347
Zhou, X. and Stephens, M. (2014) Nature Methods 11(4):407– 409. doi: 10.1038/nmeth.2848