Clinical Epigenetics

RESEARCH

Open Access

Epigenome-wide association study for dilated cardiomyopathy in left ventricular heart tissue identifies putative gene sets associated with cardiac pathology and early indicators of cardiac risk

Konstanze Tan[1], Darwin Tay[1], Wilson Tan[2], Hong Kiat Ng[1], Eleanor Wong[4], Michael P. Morley[5], Gurpreet K. Singhera[6], Chang Jie Mick Lee[3], Pritesh R. Jain[1], Fei Li Tai[1], Paul J. Hanson[6], Thomas P. Cappola[5], Kenneth B. Margulies[5], Roger Foo[2†] and Marie Loh[1*†]

Abstract

Background Methylation changes linked to dilated cardiomyopathy (DCM) affect cardiac gene expression. We investigate DCM mechanisms regulated by CpG methylation using multi-omics and causal analyses in the largest cohort of left ventricular tissues available.

Methods We mapped DNA methylation at ~850,000 CpG sites, performed array-based genotyping and conducted RNA sequencing on left ventricular tissue samples from failing and non-failing hearts across two independent DCM cohorts (discovery $n=329$, replication $n=85$). Summary-data-based Mendelian Randomisation (SMR) was applied to explore the causal contribution of sentinel CpGs to DCM. Fine-mapping of regions surrounding sentinel CpGs revealed additional signals for cardiovascular disease risk factors. Coordinated changes across multiple CpG sites were examined using weighted gene co-expression network analysis (WGCNA).

Results We identified 194 epigenome-wide significant CpGs associated with DCM (discovery $P<5.96E{-}08$), enriched in active chromatin states in heart tissue. Amongst these, 32 sentinel CpGs significantly influenced the expression of 30 unique proximal genes (± 1 Mb). SMR suggested the causal contribution of two sentinel CpGs to DCM and two other sentinel CpGs to the expression of two unique proximal genes ($P<0.05$). For one sentinel CpG, colocalisation analyses provided suggestive evidence for a single causal variant underlying the methylation-gene expression relationship. Fine-mapping revealed additional signals linked to cardiovascular disease-relevant traits, including creatinine levels and the Framingham Risk Score. Co-methylation modules were enriched in gene sets and transcriptional regulators related to cardiac physiological and pathological processes, as well as in transcriptional regulators whose cardiac relevance has yet to be determined.

†Roger Foo and Marie Loh have jointly supervised this work.

*Correspondence:
Marie Loh
marie_loh@ntu.edu.sg
Full list of author information is available at the end of the article

Conclusions Using the largest series of left ventricular tissue to date, this study investigates the causal role of cardiac methylation changes in DCM and suggests targets for experimental studies to probe DCM pathogenesis.

Introduction

Dilated cardiomyopathy (DCM) is the most common form of cardiomyopathy and a leading indication for heart transplantation. It affects 1 in 250–400 individuals and has an annual incidence of 5–7 cases per 100,000 persons per year [1]. DCM is characterised by thinning and weakening of the left ventricular heart walls, leading to contractile dysfunction [2–4]. There is marked heterogeneity in both the prognosis and age of onset for DCM, with most patients becoming symptomatic between 20–50 years of age [5]. DCM accounts for a substantial 2–3% of yearly sudden cardiac death (SCD) events which are characterised by an abrupt loss of cardiac function and death occurring within one hour of symptom onset [6, 7].

Over the past decade, the genetic basis for DCM has only been partially unravelled [8]. Beyond genetic predisposition, environmental factors could also influence the incidence and course of DCM [4]. DNA methylation, a key regulator of gene expression, as well as a molecular integrator of genetic and environmental influences, has been investigated in DCM to gain insights into its molecular pathogenesis. To date, multiple epigenome-wide association studies (EWAS) of DCM have identified a handful of differentially methylated loci [9–14]. These include CpGs at the natriuretic peptide A (*NPPA*) and natriuretic peptide B (*NPPB*) loci, where disease-linked methylation disturbances have been observed to be conserved between myocardial tissue and peripheral blood [9]. The *NPPA* and *NPPB* loci encode the cardiac stress markers atrial natriuretic peptide (ANP) and brain natriuretic peptide (BNP), which are commonly measured in clinical settings to establish the diagnosis and prognosis of heart failure [15]. These markers are considered gold standard indicators for heart-failure prognosis, underscoring the potential clinical relevance of DNA methylation in DCM.

Existing epigenome-wide methylation studies of DCM have largely been limited by sample size ($n=8$–72) owing to the scarcity of left ventricular tissue and the limited coverage of CpG sites on older methylation arrays [9–11, 14]. The largest existing EWAS of DCM (discovery $n=72$) detected only five loci at epigenome-wide significance ($P<5E{-}8$), highlighting the need for larger samples to uncover signals that are robustly associated with disease [9]. While previous DCM methylation studies utilised older Illumina arrays that covered up to 450,000 CpG sites, newer arrays now cover up to 850,000 sites and provide enhanced coverage of key gene regulatory regions, including enhancers [16]. Recent advancements in statistical methodologies, such as Mendelian Randomisation (MR), have allowed researchers to move beyond discovering associated signals to addressing causality [17]. This is particularly relevant for DNA methylation studies, as DNA methylation is affected by various disease-related and environmental factors.

In this study, we sought to expand our understanding of myocardial-specific methylation changes in DCM. We analysed intra-operative left ventricular tissue from a cohort of 329 individuals of White and African American ancestries ($n=159$ cases, 170 controls). This study builds upon existing research of DNA methylation in DCM by using a larger and more ancestrally-diverse myocardial tissue sample. Additionally, a suite of integrative omics, causal analyses and fine-mapping of putative methylation markers are employed to elucidate the putative causal involvement of methylation alterations to DCM pathogenesis.

Methods

Description of samples

Myocardial applied genomics network cohort (MAGNet)

Transmural biopsies of the left ventricular free wall (0.5–1.5 g per cryovial) were collected during cardiac surgery from subjects with heart failure attributed to DCM undergoing transplantation, and from donor hearts deemed unsuitable for transplantation but with apparently normal ventricular function. DCM was diagnosed based on the American Heart Association guidelines [18]. Specifically, all left ventricular tissue included in this investigation were obtained from subjects with an LVEF of $\leq 40\%$ and with a diagnostic workup that revealed no evidence of familial or pregnancy-associated DCM, cardiac ischaemia, significant toxin exposure or myocarditis. Hearts were perfused immediately with cold in-situ high-potassium cardioplegia before cardiectomy to arrest contraction and prevent ischemic damage. Tissue samples were promptly frozen in liquid nitrogen.

Bruce McManus cardiovascular biobank cohort (BMCB)

Biopsies of the left ventricular free wall (apical portion; 1- to 2-mm diameter) were obtained within five minutes post-cardiac surgery from subjects with DCM undergoing transplantation. Tissue samples were sourced from the Bruce McManus Cardiovascular Biobank (BMCB)

at the University of British Columbia (UBC). DCM was diagnosed based on the American Heart Association/American College of Cardiology (AHA/ACC) guidelines during pre-transplant admission, as well as during routine follow-ups and evaluation. Following the heart transplant, cardiac pathologists conducted detailed examinations of the explanted hearts to confirm the diagnosis of DCM. The pathology findings were documented in reports, which were subsequently utilised by the BMCB for DCM stratification. All left ventricular tissue included in this investigation were obtained from subjects with a diagnostic workup that revealed no evidence of familial DCM, cardiac ischaemia, significant toxin exposure or myocarditis. Biopsies were washed in ice-cold saline (0.9% NaCl) and stored in liquid nitrogen until DNA extraction. Control tissue samples were recovered within 0–2 h of circulatory death and were purchased by UBC from the International Institute for the Advancement of Medicine (IIAM).

The iHealth-T2D study

iHealth-T2D is a prospective study of Indian Asian males and females living in West London [19]. Participants were recruited in 2016. At enrolment, all participants completed a structured assessment of cardiovascular and metabolic health, which included (i) previous medical history of cardiovascular disease (CVD), specifically myocardial infarction, angina, and coronary heart disease (CHD); (ii) risk factors for CVD including hypertension, Framingham Coronary Heart Disease (FramCHD) score and creatinine; as well as (iii) high-sensitivity C-reactive protein (hsCRP). Methylation profiling was performed using genomic DNA from peripheral blood collected at enrolment. The study is approved by the National Research Ethics Service and all participants provided written informed consent.

Quantification of DNA methylation

Genomic DNA was bisulfite-converted using the EZ DNA methylation kit according to the manufacturer's instructions (Zymo Research). Case and control samples were randomised in all experiments. Bisulfite-converted genomic DNA was quantified using the Infinium MethylationEPIC BeadChip (EPIC array). Bead intensity retrieval and Illumina background correction were performed using the *minfi* R package (version 1.40.0). Markers were excluded from analyses if they met any of the following criteria: (i) they were non-CpG; (ii) they were present on sex chromosomes; or (iii) they had a low call-rate (< 95%). Samples were excluded from analyses if they met any of the following criteria: (i) they had a low call-rate (< 98%); (ii) they had mislabelled sex; or (iii) they were present in duplicate. After filtering, the discovery cohort (MAGNet) had 838,624 autosomal CpG probes and 329 samples, while the replication cohort (BMCB) had 841,904 autosomal CpG probes and 85 samples available for analysis.

Statistical analyses of epigenome-wide data

Epigenome-wide methylation data was analysed using the CPACOR pipeline (incorporating Control Probe Adjustment and reduction of global CORrelation) [20]. Quantile-normalised marker intensity values were used to calculate per-CpG methylation beta values. In the CPACOR pipeline, signal intensities from control probes built into the array are used to address technical biases in methylation data. These control probes assess key aspects of array chemistry, such as bisulfite conversion efficiency. A principal component analysis (PCA) is performed on the control probe intensities, and the resulting principal components (PCs) are included as linear predictors in regression analyses. This method was designed to minimise technical bias when comprehensive information on experimental factors that could cause data variations is lacking. In the current investigation, PCs accounting for ~ 95% of control probe variation were included as predictors in regression models to adjust for technical biases (MAGNet: first 10 PCs; BMCB: first 3 PCs). The association of each autosomal CpG site with DCM was tested using logistic regression, adjusting for age, sex and control probe PCs (Model: $DCM \sim Beta + Age + Sex + PCs_{control\text{-}probes}$). Within the mixed-ancestry MAGNet cohort, regression was performed separately for each ancestry group, followed by a trans-ancestry inverse-variance-weighted meta-analysis using METAL. Test-statistic bias and inflation in association results was adjusted pre- and post- meta-analysis, using a Bayesian algorithm implemented in the *bacon* R package (version 1.30.0) [21].

Enrichment analysis of genomic regulatory features

Sentinel CpGs were evaluated for enrichment in genomic regulatory features. This evaluation was conducted against a background set of 1000 permutations of EPIC array CpGs. Permuted background sites were matched to sentinel CpGs based on similar methylation levels (± 0.025 difference in mean methylation beta) and variability (± 0.25 difference in standard deviation). Instead of using a fixed threshold for mean and standard deviation across all sentinel CpGs, a sliding-window approach was applied [22]. Using this approach, starting values of 0.025 for mean and 0.0025 for standard deviation were gradually incremented up to a maximum of 0.25 for mean and 0.025 for standard deviation, until 1000 permuted sites were identified for each sentinel CpG.

The assessed genomic features included 15 learned chromatin states, deoxyribonuclease I (DNAse I)

hotspots, regions marked by 5 core histone modifications, as well as 1210 transcription factor binding sites (TFBS). Information on chromatin states, DNAse I hotspots and regions enriched in histone modifications in various tissues and cell subsets were obtained from the Roadmap Epigenomic Consortium database, while known TFBS were sourced from the Remap 2022 database [23, 24]. Transcription Factors (TFs) corresponding to enriched TFBS were classified as follows: 'Cardiac TF' if expressed in humans and validated *in vivo* [25, 26]; 'Computational-based' if predicted as cardiac TFs computationally due to disruption of TFBS near heart-failure linked loci [27, 28], or 'Cardiac role unknown' if existing literature has not suggested any cardiac role in humans. Enrichment was determined through a one-tailed permutation test. Permutation *P* values were computed by comparing overlap counts in the sentinel CpG set (test set) with the background distribution.

RNA sequencing

RNA sequencing (RNA-Seq) data for gene expression was obtained from the same left ventricular tissue samples analysed for methylation, if available ($n=306$, MAGNet). Prior to downstream analyses, genes were removed if they met any of the following criteria: (i) they had low read counts (< 10 counts in the minimum group size of samples); or (ii) they were present on sex chromosomes. Gene expression levels were normalised using Trimmed Mean of M-values (TMM), recommended for quantitative trait loci analyses [29]. More details on the sequencing, alignment, filtering and normalisation of RNA-seq data are available in Additional file 3.

Expression quantitative trait methylation (cis-eQTM) analysis

Expression quantitative trait methylation (*cis*-eQTM) analysis was conducted to identify significant associations between sentinel CpG methylation and proximal (*cis*) gene expression (± 1 Mb of the sentinel CpG based on gene transcription start site (TSS)). Gene expression was modelled against CpG methylation with adjustment for age, sex, ethnicity, RNA Integrity number (RIN) and the top 5 probabilistic estimation of expression residuals (PEER) factors which represent latent sources of variability in the gene expression data, possibly including technical and unknown confounders [30]. Association testing was performed using a linear model implemented within the *MatrixEQTL* R package (version 2.3) [31]. Following *cis*-eQTM identification, physical interactions between sentinel CpGs and their target genes were further examined using left ventricular chromatin interaction data from an external dataset at 5 kb resolution [28]. Sentinel CpG enrichment in *cis*-eQTMs was evaluated against matched background CpGs, using the same permutation testing method earlier described in 'Methods: Enrichment analysis of genomic regulatory features'.

Methylation quantitative trait loci analysis (meQTL)

DNA samples were obtained from the same left ventricular tissue samples analysed for methylation, if available ($n=304$, MAGNet). Genotyping was performed using the Affymetrix Genome-Wide SNP Array 6.0 [32]. SNPs were removed from the current analyses if they met any of the following criteria: (i) they were multiallelic; (ii) they were present in duplicate; (iii) they had low imputation quality (R2 < 0.3); or (iv) they were rare, defined as having an ancestry-specific minor allele frequency (MAF) of less than 0.05. Methylation quantitative trait loci (meQTL) analysis was conducted to identify significant associations between SNPs and sentinel CpG methylation (± 500 kb). CpG methylation was regressed against SNP dosage values with adjustment for age, sex and methylation array control probe PCs capturing 95% of control probe variation. Association testing was performed using a linear model implemented within the *MatrixEQTL* R package (version 2.3) [31]. Regression was performed separately for each ancestry group, followed by a trans-ancestry inverse-variance-weighted meta-analysis using METAL.

Causal analyses (Mendelian Randomisation and Colocalisation)

Summary-data-based Mendelian randomisation (SMR) was performed to investigate the putative causal contribution of sentinel CpGs to DCM. SMR uses summary-level association statistics from independent association studies to compute a causal estimate for the influence of an exposure (e.g. sentinel CpG methylation) on an outcome (e.g. DCM disease status, or gene expression) [17]. The significance of the causal estimate is determined using the Wald Test. Separate SMR analyses were conducted to examine causal relationships between sentinel CpG methylation and (i) DCM, as well as (ii) proximal gene expression (± 1 Mb). For each sentinel CpG, the SNP that was most strongly associated with the CpG in meQTL analysis ($P<0.05$), that was also analysed for association with the respective outcome, was chosen to be the instrumental variable (IV) for assessing causal relationships. Genetic associations for DCM were obtained from the largest published GWAS of DCM ($n=355{,}381$, UKBioBank), while genetic associations for gene expression were obtained from the largest left ventricular tissue *cis*-expression quantitative trait loci (*cis*-eQTL) study ($n=386$; GTEx v8 release; SNPs within ± 1 Mb of gene transcription start sites) [33, 34]. We validated SMR-significant associations using one-sample Mendelian

Randomisation (one-sample MR), and also conducted colocalisation analyses to evaluate the posterior probabilities of a shared causal variant for the assessed traits. One-sample MR was performed with individual-level genotype, methylation, and outcome data (MAGNet) using the 2-stage least squares regression method implemented in the *AER* R package (version: 1.2–12). Colocalisation analysis was conducted using the *coloc.abf* function in the *coloc* R package (version 5.2.3), based on a ± 500 kb region around each sentinel CpG as this was the window size used for identifying meQTL [35].

Weighted gene co-expression network analysis (WGCNA)

We employed weighted gene co-expression network analysis (WGCNA) to construct co-methylation modules using DCM-associated CpGs from discovery-stage EWAS [36]. These CpGs were (i) associated with DCM (FDR $P < 0.05$); (ii) replicated with consistent association direction; and (iii) displayed variability (methylation beta SD > 0.02). Using 32,198 CpGs, we built a signed consensus co-methylation network using the *blockwiseConsensusModules* function in the *WGCNA* R package (version 1.72–5), using settings recommended for our sample size (soft thresholding power = 12, merge cut height = 0.25 and minimum module size = 30). Methylation levels within consensus modules were summarised as module eigengenes (ME) and tested for correlation with DCM separately by ancestry (Model: DCM ~ ME + Age + Sex + 10PCs). Ancestry-specific results were combined using inverse-variance-weighted meta-analysis to determine module associations with DCM. Additionally, a hypergeometric test was applied to assess module enrichment in DCM sentinel CpGs ($P < 0.05$) against all CpGs used for network construction. Genes mapped to CpGs within co-methylation modules were assessed for overrepresentation of Gene Ontology (GO) terms and biological processes from KEGG and REACTOME databases using the *clusterProfiler* R package (version 4.10.0). GO gene sets were obtained from *org.Hs.eg.db* (version 3.18.0), while KEGG and REACTOME datasets were accessed via the *msigdbr* R package (version 7.5.1). Co-methylation modules were also assessed for enrichment in known TFBS.

Targeted methylation sequencing

Targeted methylation sequencing was performed using genomic DNA from peripheral blood samples collected from participants of the iHealth-T2D study. Targeted methylation sequencing was performed for regions defined as ± 500 bp of DCM sentinel CpGs. The choice of window size is supported by previous publications demonstrating a decay in correlation between methylation sites beyond 1–2 kb [37, 38], and from our in-house whole-genome bisulfite sequencing dataset where we observed that $|r| > 0.2$ was within ± 500 bp for most of the assessed CpGs (data unpublished). Additional details on probe design, sample processing and library preparation, sequencing and quality control steps employed are available in Additional file 3.

Statistical analyses of targeted methylation sequencing data

Within each sequenced region, the association of each CpG site with CVD-related traits was tested using linear regression for continuous traits and logistic regression for binary traits, adjusting for age, sex and white blood cell proportion of six white blood cell sub-populations (CD8 T cells, CD4 T cells, Natural Killer Cells, B-cells, Monocytes, Granulocytes) estimated by the Houseman algorithm [39]. In parallel, pairwise correlation in methylation levels between CpGs within regions was calculated using the R function *cor*. For a given pair of CpGs, only samples where methylation data was available for both CpGs being compared were used to calculate correlation.

Regions represented by sentinel CpGs were assessed for enrichment in significant associations to CVD traits, compared regions represented by a background set of 1000 permutations of non-sentinel CpGs. The background set consisted of EPIC array CpGs matched by methylation levels and variability to sentinel CpGs using the same sliding-window approach that was described in 'Methods: Enrichment analysis of genomic regulatory features.'

Construction of methylation risk score (MRS)

We constructed a methylation risk score (MRS) from CpGs in fine-mapped regions and examined the association of these scores with CVD traits. Effect sizes of association between individual CpG loci and CVD traits were used as weights for constructing the methylation score. Specifically, we defined trait-specific MRS within each region as a linear combination of k CpG sites, beta values b and weights w:

$$\mathrm{MRS} = \sum_{i=1}^{k} w_i b_i$$

Within each region, the association of each MRS with CVD-related traits was assessed, adjusting for the same covariates as in single loci association testing described in 'Statistical analyses of targeted methylation sequencing data.' Permutation testing was conducted using

the same methodology as that applied to single CpG associations.

Results

Overview of study design

Our study design is summarised in Fig. 1. In brief, we first carried out an epigenome-wide association investigation of DCM using 414 left ventricular samples obtained from two repositories: the Myocardial Applied Genomics network (MAGNet; $n=329$ [discovery]) and the Bruce McManus Cardiovascular Biobank (BMCB; $n=85$ [replication]) (Additional file 1: Table S1). Integrative omics analyses were performed on the identified sentinel CpGs. To discover additional signals beyond CpG sites captured by the methylation array, we conducted fine-mapping of selected top-ranking loci in blood samples obtained from a population-based cohort (iHealth-T2D, $n=1974$).

Epigenome-wide association analysis

We performed EWAS of DCM using genomic DNA extracted from left ventricular free-wall tissue. Separate EWAS were carried out for Whites and African Americans within the discovery cohort (MAGNet), followed by inverse-variance meta-analysis (Methods; Additional file 2: Figure S1). From discovery-stage EWAS, 196 CpG sites were associated with DCM at a Bonferroni-corrected threshold of $P<5.96E{-}08$ (0.05/838,624) (Fig. 2A). Subsequent targeted replication testing in the BMCB cohort ($n=36,925$ CpGs, discovery FDR $P<0.05$) confirmed consistent directionality of effect size estimates for all 196 Bonferroni-significant discovery CpGs, as well as previously reported CpG associations with DCM (Additional file 1: Tables S2 and S3). The 196 CpG sites were distributed across 171 genetic loci, with 150/171(88%) genetic loci containing a single sentinel CpG and 21/171 (12%) loci containing two or more sentinel CpGs (Additional file 1: Table S2, Additional file 2: Figure S2). Conditional analyses at each locus identified a total of 194 robustly associated and conditionally independent signals ('sentinel CpGs'), which were further analysed for functional relevance and causal contribution to DCM pathogenesis (Fig. 2B).

Unsupervised hierarchical clustering based on the methylation levels of the 194 sentinel CpGs resulted

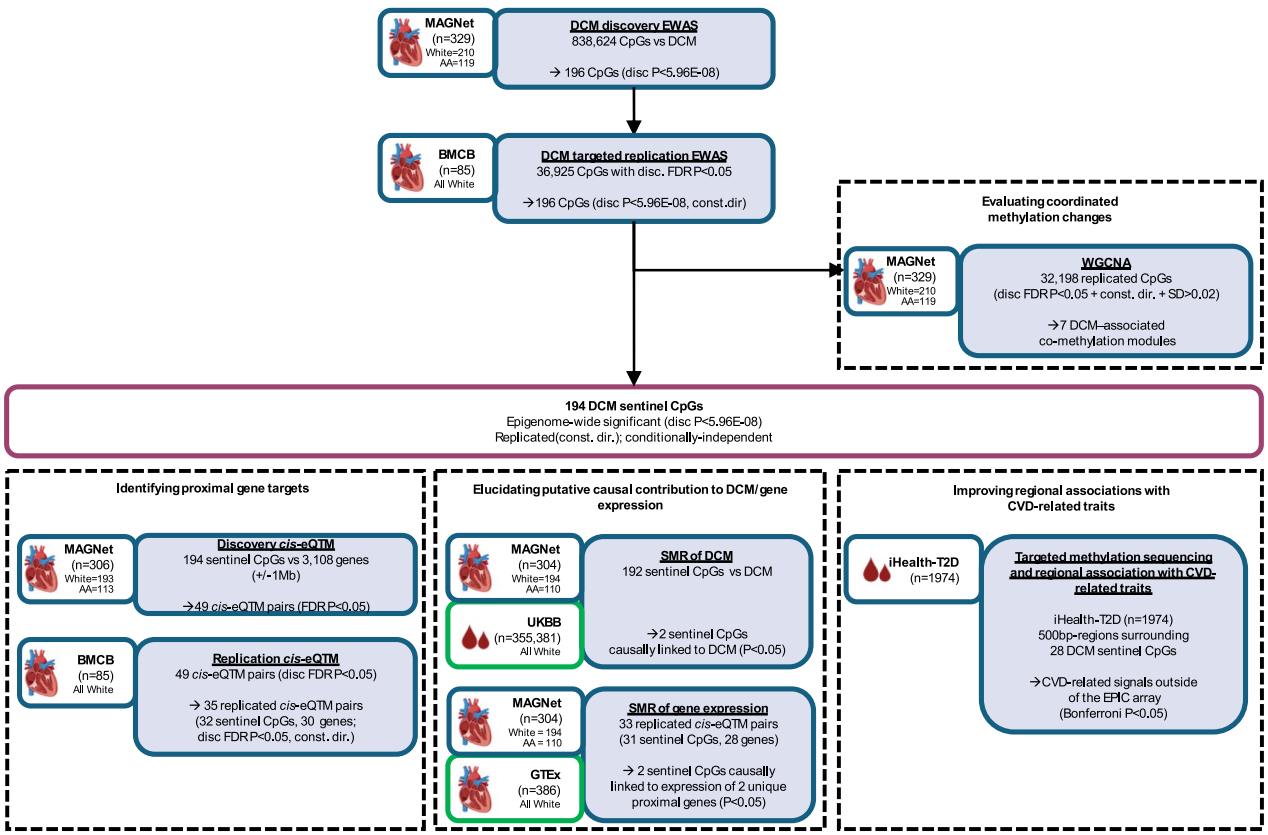

Fig. 1 Study design. Key abbreviations used: (Analyses) EWAS = epigenome-wide association study; eQTM = expression quantitative trait methylation analysis; WGCNA = weighted gene co-expression network analysis; (Ancestries) AA = African American; (Cohorts) MAGNet = Myocardial Applied Genomics Network; BMCB = Bruce McManus Cardiovascular Biobank; UKBB = UK Biobank; GTEx = Genotype-Tissue Expression project

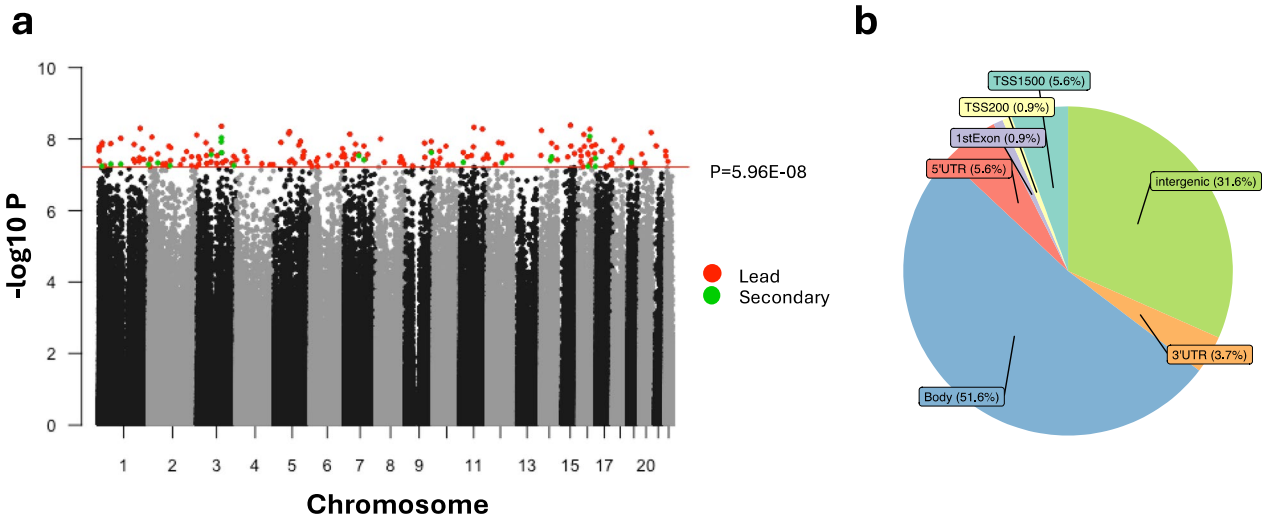

Fig. 2 DCM Sentinel CpGs. **a** Manhattan plot for discovery-stage EWAS of DCM. The horizontal significance line (red) corresponds to the epigenome-wide significance threshold ($P < 5.96E{-}08$, 0.05/838,624 tests). The 194 DCM sentinel CpGs are highlighted. At genomic loci with > 1 epigenome-wide significant signal, secondary signals (green) were identified by conditioning on the lead signal (lowest P in region; red). **b** Distribution of DCM Sentinel CpGs across various gene features. The 194 Sentinel CpGs were annotated to gene features based on the Illumina manifest file (version b2)

in two distinct clusters, segregating samples by their respective case and control status (binomial $P < 2.2E{-}16$) (Fig. 3). DCM cases comprised the majority of one cluster (128/145; 88%), while controls constituted the majority of the second cluster (153/184; 83%). This finding supports a perturbation of DNA methylation in DCM. Clustering by case and control status persisted in ancestry-specific unsupervised hierarchical clustering of methylation levels of the sentinel CpGs (Additional file 2: Figure S3).

Enrichment of Sentinel CpGs in active gene regulatory regions and impact on proximal gene expression

To understand the regulatory role of sentinel CpGs, we first examined the enrichment of chromatin states. Compared to a background of array CpGs matched by methylation levels and variability, sentinel CpGs were enriched in transcriptionally active chromatin states of left ventricular tissue, including weakly-transcribed regions (permutation test $P < 0.001$), actively transcribed regions and enhancers (permutation test $P < 0.05$; Additional file 2: Figure S4A). Conversely, sentinel CpGs exhibited depletion in polycomb-repressed regions relative to the background (permutation test $P < 0.001$). In addition to enrichment for transcriptionally active chromatin states, sentinel CpGs were also enriched in deoxyribonuclease I (DNase I) hotspots (Additional file 2: Figure S4B). DNase I hotspots are genomic regions that exhibit a significantly high frequency of cleavage by the enzyme DNase I, indicating areas of increased accessibility within the chromatin. Sentinel CpGs were not only enriched in cardiac tissue-specific DNase I hotspots, but also in DNAse I hotspots across various other tissue types and cell subsets, suggesting their gene regulatory role across multiple tissues. We also analysed the overlap between sentinel CpGs and regions marked by histone modifications associated with gene regulation. Sentinel CpGs were enriched in H3K4me1-marked regions indicative of primed enhancers and promoters. Conversely, sentinel CpGs were depleted in H3K4me3-marked regions associated with active promoters. This enrichment in primed, rather than active regulatory elements, suggests that sentinel CpGs could contribute to an epigenetic priming mechanism that facilitates changes in gene expression in response to pathological stressors.

To identify sentinel CpGs that impacted proximal gene expression, hereon referred to as *cis*-expression quantitative methylation loci (*cis*-eQTM), we examined the association between methylation levels of the 194 sentinel CpGs and expression of their proximal genes (< 1 Mb from the gene transcription start site (TSS)) present in the discovery (MAGNet) and replication (BMCB) RNA-seq datasets. The 194 sentinel CpGs were enriched for association with proximal gene expression (discovery FDR $P < 0.05$) in left ventricular tissue (4.20-fold compared to expectations under the null hypothesis; $P < 0.001$) (Additional file 2: Figure S5). Subsequent targeted replication testing on sentinel CpG-gene pairs reaching FDR $P < 0.05$ in discovery-stage association testing confirmed consistent directionality of effect size estimates between 32 sentinel CpGs and 30 unique proximal

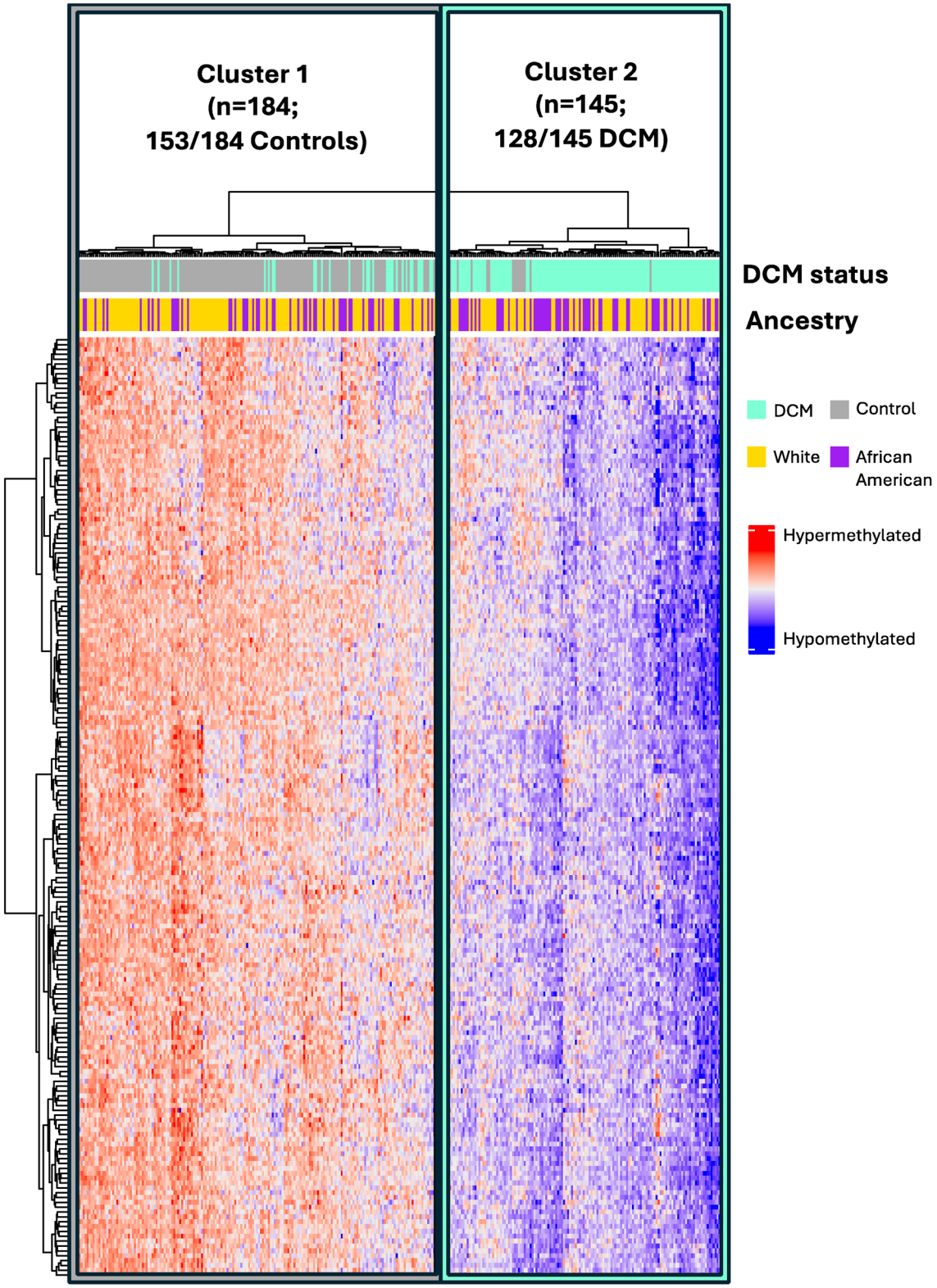

Fig. 3 Unsupervised clustering using methylation levels of 194 sentinel CpGs. DCM, dilated cardiomyopathy

genes (35 pairs; 'replicated *cis*-eQTMs') (Fig. 4, Additional file 1: Table S4). Bulk left ventricular Hi-C data further supported physical interactions between 29 sentinel CpGs and 27 unique proximal genes (31/35 pairs, 88.6%) (Fig. 4, Additional file 1: Table S4). For most replicated *cis*-eQTMs, sentinel CpGs were located 5' upstream of their target genes (Fig. 4A) and were inversely correlated with target gene expression (25/35 pairs, 71.4%) (Fig. 4B).

Summary-data-based Mendelian Randomisation to infer disease causation

Having obtained initial evidence of sentinel CpG contribution to transcriptional regulation, we next leveraged upon summary-data-based Mendelian randomisation (SMR) to further evaluate potential causal relationship of the sentinel CpGs with both DCM and proximal gene expression. Separate causal analyses were conducted for DCM and gene expression. We found two sentinel CpGs that were potentially causally linked to DCM (cg08140459 and cg12359658; $P < 0.05$), with subsequent validation testing via one-sample Mendelian Randomisation (one-sample MR) confirming consistent directionality of causal estimate for cg08140459-DCM ($P < 0.05$) (Additional file 1: Table S5).

To gain further insight into the molecular mechanisms underpinning the contribution of sentinel CpGs to DCM, we conducted a separate SMR of gene expression, focusing on replicated *cis*-eQTMs. After excluding sentinel CpGs without suitable instrumental variables (e.g. the SNP with the strongest association to CpG methylation was not assessed for association with proximal gene expression in the GTEx *cis*-eQTL analysis), 33 *cis*-eQTMs (31 unique sentinel CpGs, 28 unique genes) could be analysed in the SMR of gene expression. Of the 31 sentinel CpGs analysed, two sentinel CpGs showed putative causal relationships with two unique proximal genes ($P < 0.05$) (Additional file 1: Table S6).

We further integrated evidence from multi-omics analyses to evaluate the causal contribution of two sentinel CpGs to DCM, both of which had significant causal estimates for gene expression based on SMR. We examined the posterior probability for a shared causal variant influencing both CpG methylation and proximal gene expression (cg11793257-*ABHD12*, coloc.abf-PP.H4 = 0.47; cg01651169-*ATP5MF*, coloc.abf-PP.H4 = 0.020) (Additional file 1: Table S6). The causal estimates for both pairs were confirmed in one-sample MR, showing consistent direction of association (Additional file 1: Table S6). To illustrate our approach for causal analysis, we focused on cg11793257-*ABHD12*, which exhibited suggestive evidence of genetic colocalisation (Fig. 5). To further investigate the potential causal regulatory relationship between cg11793257 methylation and Abhydrolase Domain Containing 12 (*ABHD12*) expression, we additionally examined an external dataset of left ventricular Hi-C chromatin interactions [28]. This analysis revealed that the chromatin region containing cg11793257 interacts with several chromatin regions along *ABHD12*, including a putative promoter region at the 5' end of *ABHD12* that features H3K27ac peaks (Fig. 5C).

While the target genes identified by SMR of gene expression to be influenced by sentinel CpG methylation have not been directly evaluated for their impact on cardiac function, existing research suggests their relevance to cardiac pathogenesis. Abhydrolase Domain Containing 12 (*ABDH12*) is an enzyme that hydrolyses endocannabinoids, a class of lipids influencing various cardiac pathologies including inflammation and cell death [40]. ATP Synthase Membrane Subunit F (*ATP5MF*) encodes a subunit of the mitochondrial ATP Synthase complex that plays a crucial role in maintaining mitochondrial energy production, a fundamental process in cardiac cell contractility and function [41].

Genes mapped to CpG sites demonstrating coordinated changes in methylation patterns are enriched in disease-relevant pathways

Beyond analysing single CpG associations, examining coordinated methylation changes across multiple CpG sites and their linked genes could reveal disease-relevant pathways regulated by DCM methylation. To achieve this, we conducted weighted gene co-expression network analysis (WGCNA), constructing co-methylation modules using methylation levels of 32,918 DCM-associated CpG sites (discovery EWAS FDR $P < 0.05$, with consistent

(See figure on next page.)

Fig. 4 Association of sentinel CpG methylation with gene expression. **A** *cis*-eQTMs by genomic region. Sentinel CpGs are annotated by their genic locations: 5' upstream (the region from upstream to + 100 bp downstream of the gene TSS), gene body, and 3' downstream (the region following the gene body). Replicated *cis*-eQTMs were defined using two criteria: discovery FDR $P < 0.05$ (MAGNet) and confirmed directionality in replication testing (BMCB). **B** Directionality of replicated *cis*-eQTMs by genic location of sentinel CpGs. In each region-specific plot, replicated *cis*-eQTMs are ordered by chromosomal location of sentinel CpGs along the x-axis. The y-axis represents beta values of individual *cis*-eQTM relationships based on discovery-stage analysis. The table below each plot summarises the inverse and positive *cis*-eQTMs by genomic region, including their ratio (inverse/positive). Gene coordinates and TSS are based on the GENCODE version 19 annotation. eQTM = expression quantitative trait methylation loci

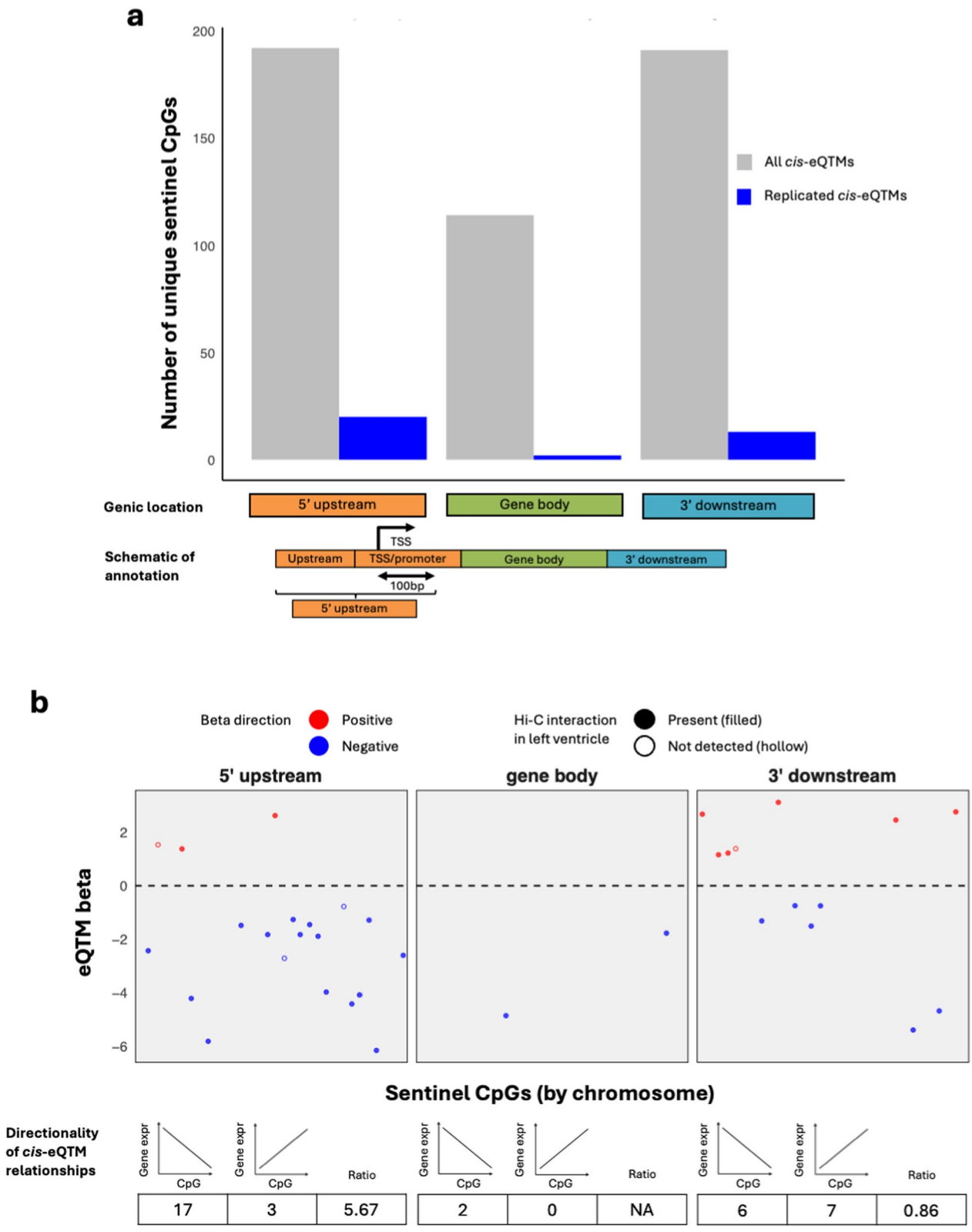

Fig. 4 (See legend on previous page.)

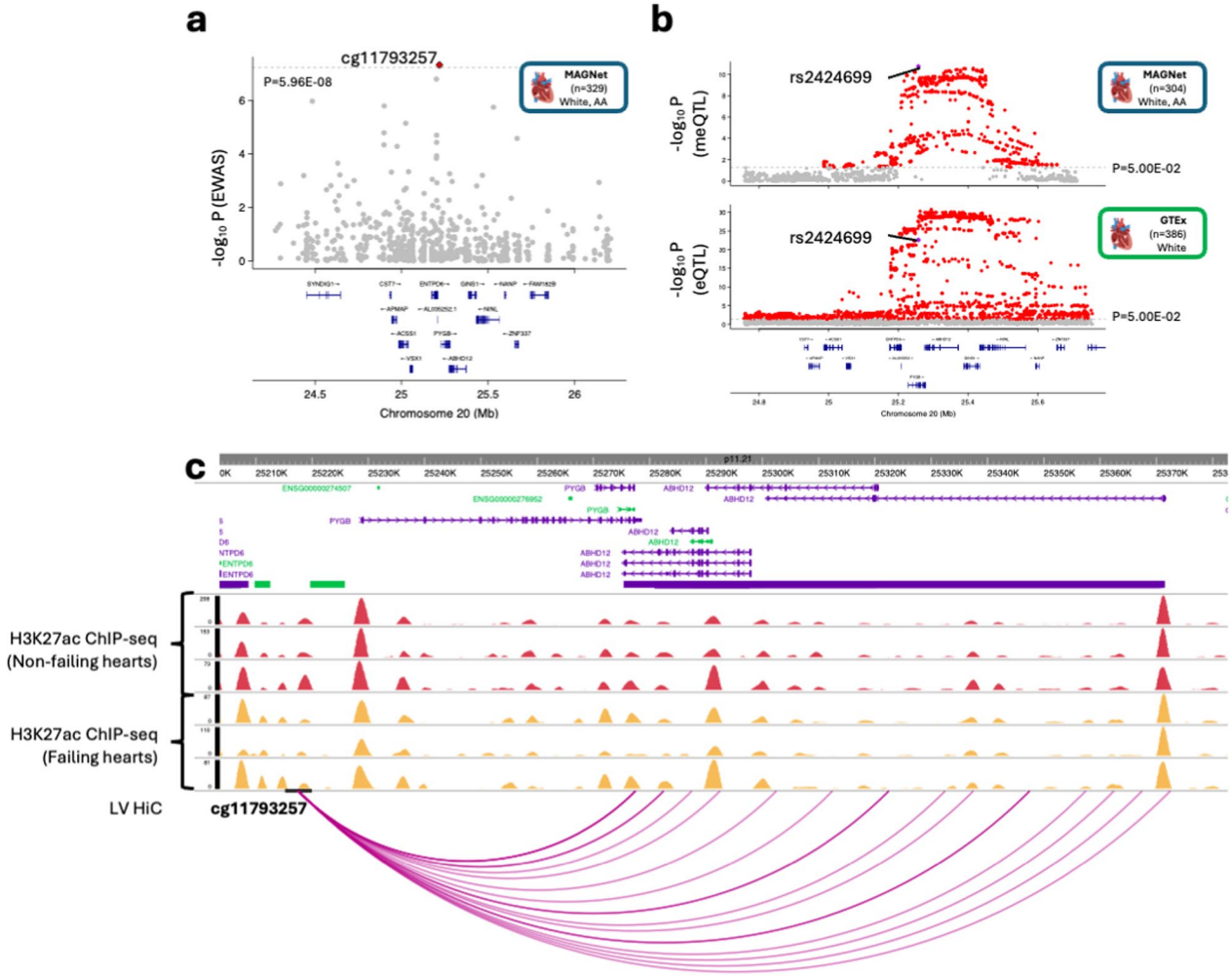

Fig. 5 Regional plots for a sentinel CpG gene pair with suggestive evidence for colocalisation. **A** Associations of CpGs surrounding cg11793257 with DCM from discovery-stage EWAS (MAGNet). The displayed region (2 Mb) is centred on cg11793257. **B** Regional genetic associations with cg11793257 and *ABHD12* Genetic associations were obtained for a region (± 500 kb) centred on the genetic variant (rs2424699) that was used as an instrumental variable in SMR to assess causal relationships between cg11793257 methylation and *ABHD12* gene expression. The posterior probability (coloc PP.H4) of colocalisation between cg11793257 meQTLs and *ABHD12* eQTLs was 0.47. Gene tracks for all regional plots consist of protein-coding genes obtained from Ensembl Release 75 (last release to be based on hg19 assembly). **C** Left ventricular chromatin interactions. At 5 kb resolution, multiple interactions were identified between the chromatin region containing cg11793257 and various regions along *ABHD12*. meQTL = methylation expression quantitative trait loci; eQTL = expression quantitative trait loci

direction of association in replication EWAS, SD > 0.02 across all samples). Seven co-methylation modules were identified (Fig. 6A). Regressing the module eigengene (the first principal component of module-specific methylation intensities) against DCM confirmed an association with DCM for all modules (Additional file 1: Table S7).

To investigate the biological relevance of co-methylation modules, we looked for over-represented gene sets and enrichment in transcription factor binding sites (TFBS). Gene set overrepresentation of gene ontology terms and pathways (KEGG/REACTOME) was analysed using genes belonging to replicated *cis*-eQTM pairs of module-specific CpGs (Fig. 6B–D, Additional file 1: Table S8). Three of the seven identified co-methylation modules had enriched gene ontology terms (FDR $P < 0.05$). None of the modules showed significant enrichment for KEGG or REACTOME pathways. The most enriched gene sets and pathways in each module reflected distinct aspects of DCM pathogenesis, including alterations in extracellular matrix composition (Module 1; Fig. 6B), cell motility (Module 3; Fig. 6C), and immune-mediated pathways (Module 6) (Fig. 6D).

Five of the seven identified modules showed significant enrichment in TFBS relative to background

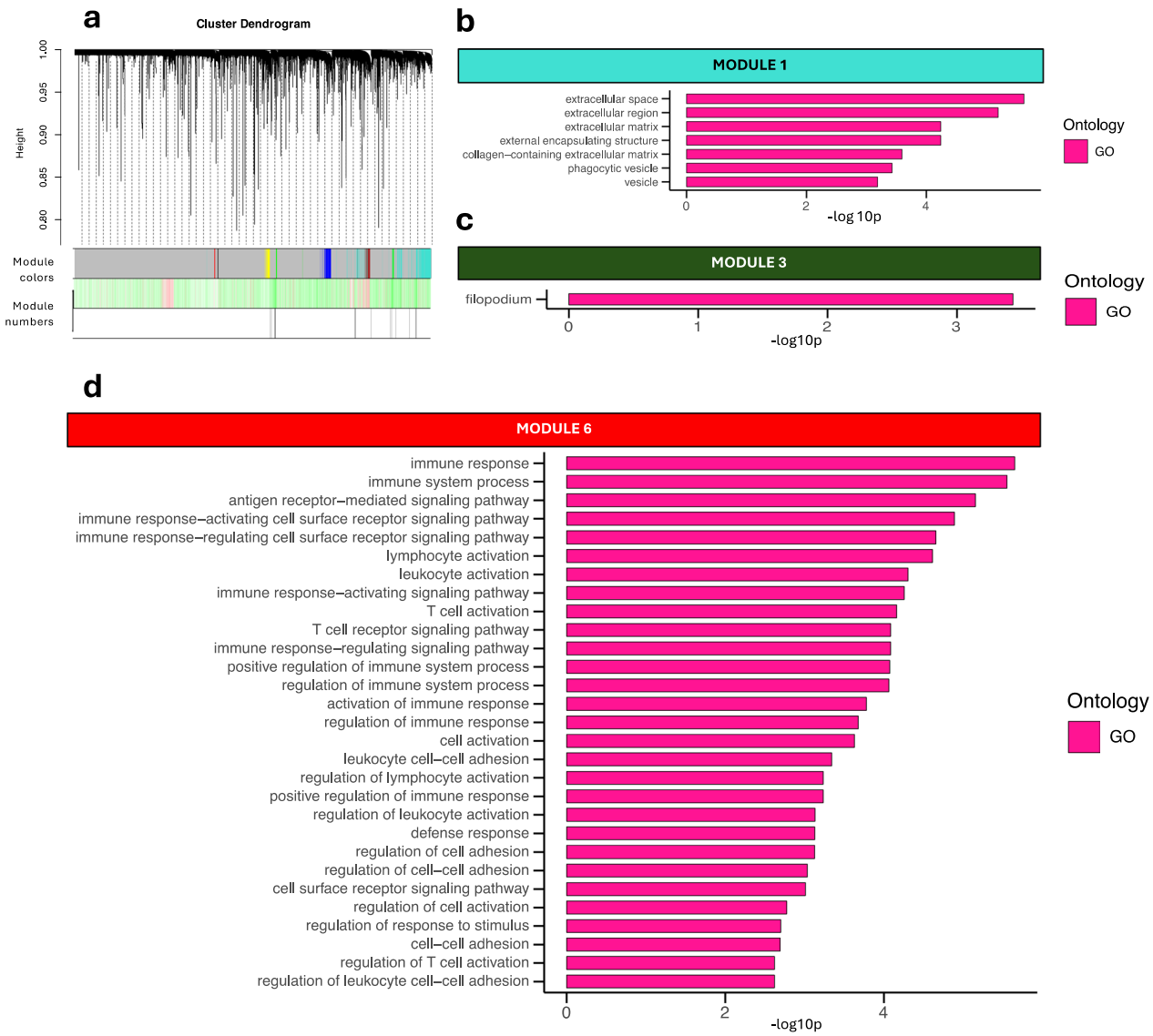

Fig. 6 Co-methylation modules identified using DCM-associated CpGs. **A** Hierarchical cluster tree (dendogram) of co-methylation modules. WGCNA identified seven modules that exhibit conservation between co-methylation networks constructed independently within the White ($n=209$) and African American ($n=118$) ancestries. The colour band underneath the tree indicates distinct modules (grey = CpG sites that are not clustered into any module). **B**–**D** Bar plots of over-represented gene sets (FDR $P<0.05$) in co-methylation modules. Genes were assigned to CpGs based on replicated *cis*-eQTM relationships. Gene set enrichment within a module was assessed against a background set consisting of all genes mapped to CpG sites that were used to identify co-methylation patterns (855 unique genes). Enriched gene ontology terms are displayed by decreasing order of significance (–log10p value)

CpGs (permutation test $P<0.001$) (Additional file 1: Table S9, Additional file 2: Figure S6). Examining the most enriched TFBS for each module, two modules featured TFBS with known or likely cardiac roles, namely Thyroid Hormone Receptor Alpha (*THRA*) (Module 1) and Homeobox protein Hox-B8 (*HOXB8*) (Module 2). *THRA* contributes to cardiac function and contractility and *HOXB8* belongs to the *HOX* family of genes that contribute to cardiac development [42, 43]. The top enriched TFBS for the other modules corresponded to TFs which are not currently known to specifically contribute to cardiac pathways, including a component of the RNA polymerase II transcription machinery (*GTF2A2*; Module 4), a zinc finger protein (*ZNF224;* Module 5), and an oncogenesis-linked TF (*LMO1*; Module 7) [44]. Taken together, findings from gene

set and TFBS enrichment analyses highlight diverse aspects of DCM pathogenesis driven by coordinated changes in methylation patterns.

Fine-mapping sentinel CpGs to investigate regional associations with traits related to cardiac disease and disease risk

As the methylation array covers only 2–3% of CpG sites in the epigenome, we sought to improve our investigation of regional associations using an existing target methylation sequencing dataset to increase coverage of CpG sites (± 500 bp) surrounding the top-performing DCM sentinel CpGs. Targeted methylation sequencing was performed using blood samples of individuals from the iHealth-T2D study, which is well phenotyped for various traits relevant to cardiovascular disease (CVD) ($n=1974$). We assessed regional associations with: (i) previous medical history of CVD, specifically myocardial infarction, angina, and coronary heart disease (CHD); (ii) risk factors for CVD including hypertension, Framingham Coronary Heart Disease (FramCHD) score and the renal marker creatinine, which has been associated with a greater risk of heart disease and early death in the general population [45]; as well as (iii) high-sensitivity C-reactive protein (hsCRP), an inflammatory marker that has been independently associated with increased risk of CVD in asymptomatic individuals [46].

Targeted sequencing was performed on regions surrounding 28 DCM sentinel CpGs (28 regions) (Additional file 1: Table S10). A total of 293 CpG sites were captured, with two to 24 CpG sites captured in each region. Seventeen regions had significant associations with at least one of the 10 unique CVD traits (Bonferroni-corrected $P<0.05$) (Additional file 1: Table S11). At 14 regions, the CpGs with the strongest association with our CVD traits were not CpGs on the EPIC array, further illustrating the added value brought upon by targeted sequencing. Multiple independent signals were also found for two regions, whereby conditioning on the lead signal revealed secondary signals in both regions. Compared to a background of non-sentinel CpGs matched by methylation levels and variability to sentinel CpGs and with ± 500 bp regions captured in the targeted sequencing experiment, the 28 sentinel CpGs had a greater number of regions containing significant associations with creatinine (6/28 regions; permutation test $P<2.20E{-}02$) and FramCHD score (5/28 regions; permutation test $P<3.80E{-}02$) (Additional file 1: Table S13).

To illustrate the utility of targeted sequencing to improve regional associations with DCM, we highlight a region surrounding the DCM sentinel CpG, cg1179325 (EWAS $P=1.94E{-}09$). Targeted sequencing of this region revealed stronger signals for creatinine and FramCHD score (Bonferroni-corrected $P<0.05$) from CpGs that were not present on the EPIC array (chr20_25218304 with creatinine, $P=1.26E{-}03$; chr20_25218276 with FramCHD score, $P=4.08E{-}03$) (Fig. 7A, B; Additional file 1: Table S11.) The newly identified CpGs exhibited methylation levels that correlated with cg11793257 (chr20_25218304 with cg11793257 $|r|=0.42$; chr20_25218276 with cg11793257 $|r|=0.46$) (Additional file 1: Table S11.)

We next combined methylation information from multiple CpGs in each region into a weighted methylation risk score (MRS) to investigate associations with CVD. Of the 28 sequenced regions, 25 regions had significant MRS (Bonferroni $P<0.05$) for at least one investigated CVD trait (Additional file 1: Table S12). In 23 of these 25 regions, individual CpGs were not significantly associated with CVD-related traits (Bonferroni $P<0.05$). However, combining the methylation patterns of multiple CpGs within these regions into an MRS revealed significant associations with at least one of the investigated CVD-related traits, indicating that the combined effect of multiple CpGs provides stronger association with CVD traits than individual CpGs alone. Nonetheless, unlike regional single CpG association tests, the regional MRS of sentinel CpGs did not show enrichment for significant associations with specific CVD traits when compared to the regions surrounding permuted CpG sites (Additional file 1: Table S13).

Discussion

We perform the largest EWAS of DCM in cardiac tissues to date (discovery $n=159$ DCM, 170 control), extending on previous EWAS of DCM in terms of sample size and coverage of CpG sites. In discovery-stage EWAS, we identified 36,925 CpG associations with DCM (FDR $P<0.05$), which we assessed for replication in an independent DCM cohort. We further performed comprehensive multi-omics and causal analyses on the top signals: 194 independent CpG signals for DCM that reached epigenome-wide significance ($P<5.96E{-}08$). These integrative omics analyses suggested the causal contribution of a subset of the 194 sentinel CpGs to DCM pathogenesis and transcriptional regulation. Fine-mapping of putative methylation markers supported the relevance of regions containing DCM-linked methylation changes to early indicators of CVD risk. Network analysis of coordinated changes across multiple DCM-associated CpG loci highlighted pathways relevant to DCM pathogenesis.

In addition to identifying novel CpG associations with DCM, our study confirmed strong associations reported by the most comprehensive existing EWAS of DCM conducted by Meder et al. on a predominantly White cohort (discovery $n=41$ cases, $n=31$ controls) utilising the

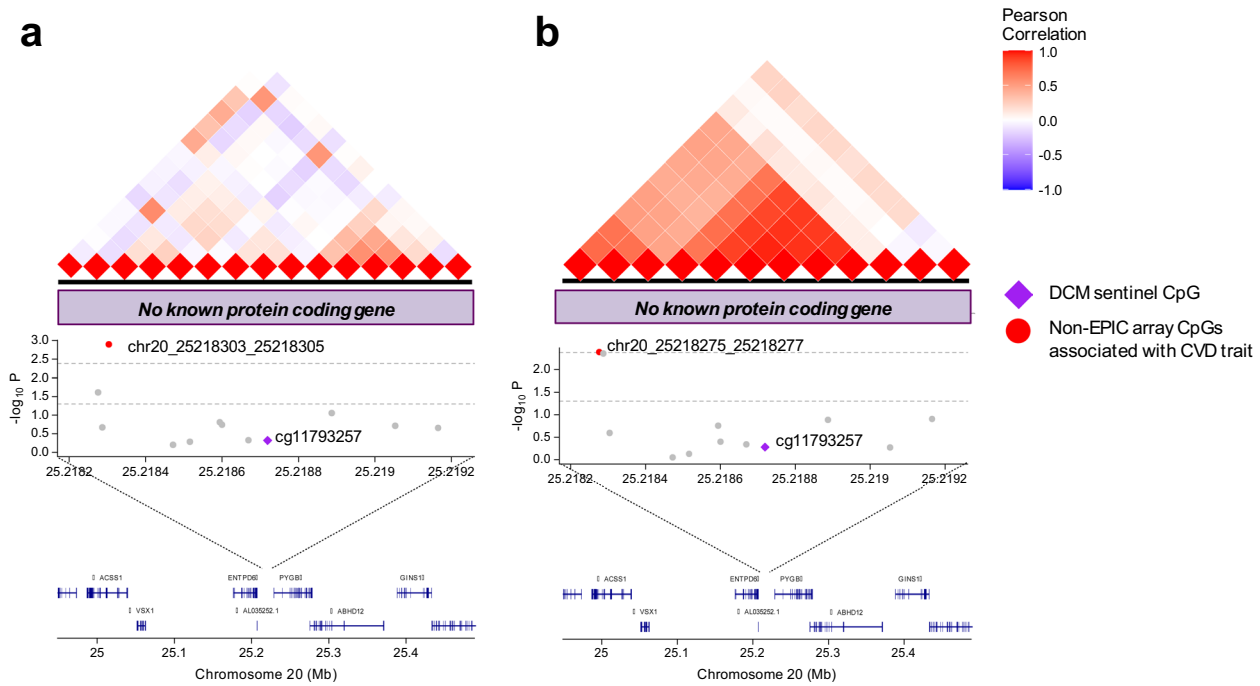

Fig. 7 Regional associations of a DCM sentinel CpG with CVD-related traits. This plot shows the regional associations of the cg11793257 region with creatinine and FramCHD score. 12 CpGs were sequenced in this region. Three types of plots are included in each panel: (top) pairwise correlation in methylation levels between CpGs within the region; (middle) CpG-CVD trait associations; (bottom) protein-coding genes in the region. **A** Regional associations with creatinine. The lead signal is chr20_25218304 ($P = 1.26E{-}03$). **B** Regional associations with FramCHD score. The lead signal is chr20_25218276 ($P = 4.08E{-}03$). CVD = cardiovascular disease

older 450 k methylation array [9]. Among Meder et al.'s reported associations that were confirmed in the current investigation (discovery FDR $P < 0.05$, consistent directionality of effect between MAGNet and BMCB cohorts) were two CpG sites (cg25838968 and cg16254946) that had reached epigenome-wide significance in Meder et al.'s investigation, as well as an additional CpG site (cg24884140) that had been singled out as a promising DCM diagnostic biomarker demonstrating consistent hypomethylation in both myocardial tissue and blood samples from individuals with DCM compared to control subjects, as well as superior classification accuracy for DCM compared to the clinical gold standard biomarker NT-proBNP.

Further to describing robust CpG associations with DCM, our study expands previous investigations with causal analyses of sentinel CpG contribution to DCM pathogenesis. While individual-level genotype data avails the option of conducting one-sample MR as the primary analysis to avoid issues arising from population heterogeneity, we opted instead to conduct SMR as our primary analysis to leverage genetic association data from a large-scale external GWAS of DCM. Additionally, utilising two samples minimises the risk of false positives. For CpG-DCM or CpG-gene pairs with SMR causal estimates that reached nominal significance ($P < 0.05$), one-sample MR was subsequently performed for validation. We identified a putative causal relationship between cg08140459 and DCM. In our investigation, cg08140459 was also robustly linked to the expression of *LTBP2*, a recently discovered prognostic biomarker for DCM, in independent cohorts [47]. To gain insight into the molecular pathways in DCM pathogenesis that involve sentinel CpGs, we conducted a separate SMR analysis for gene expression, revealing the putative contribution of two sentinel CpGs to nearby genes (cg11793257-*ABHD12* and cg01651169-*ATP5MF*) with functions relevant to DCM pathogenesis, thus warranting further investigation in a cardiac context.

Using targeted methylation sequencing data from a population-based cohort phenotyped for multiple cardiac traits, we discovered independent CVD risk factor signals near DCM sentinel CpGs that were not captured by the EPIC array, particularly for creatinine and FramCHD. We presented an example of new signals identified for both creatinine and FramCHD in the cg11793257 locus. Notably, this locus was also highlighted in our SMR of gene expression, which suggested a putative causal relationship between cg11793257 methylation and

ABHD12 expression. *ABHD12* plays a role in metabolic processes potentially related to cardiac pathology and cardiovascular disease, making it a relevant candidate for further investigation. While we found that aggregating methylation data from CpGs on a regional level improved regional associations with CVD traits, permutation analysis did not reveal enrichment for specific CVD traits among sentinel CpG regions. Further validation is required to ascertain the biological relevance of regions with MRS that are associated with CVD traits.

Our study has some limitations. Firstly, the cell-type heterogeneity of left ventricular tissue makes it challenging to delineate the specific cell types driving the associations between CpG methylation and DCM and to investigate cell composition bias. Although cell-type deconvolution algorithms exist for this task, the lack of reference methylation profiles for heart cell types means that only reference-free approaches can be applied. While reference-free deconvolution algorithms can predict distinct cell classes using major variations in methylation profiles, finding a biological basis to justify assigning these output classes to specific heart cell types (e.g. cardiomyocytes or cardiac fibroblasts) currently poses a significant challenge. As single-cell profiling techniques for methylation and gene expression in cardiac cell types advance, future methylation studies of DCM should prioritise elucidating the cellular context of DCM-linked CpG methylation and their gene regulatory roles. Regarding our MR analyses, one limitation is the potential bias in the causal relationships estimated by two-sample MR due to population heterogeneity between the multi-ancestry MAGNet cohort (which includes White and African American participants) used to generate meQTL, compared to the predominantly White cohorts used in the GWAS of DCM and left ventricular eQTL analyses. Despite this, we did not restrict the meQTL analysis to only samples from subjects of White ancestry in our cohort with methylation and genotype data available ($n = 194$) to maximise power for detecting left ventricular meQTL associations. A second limitation of the MR analyses would be the limited sample size of the cohort used to generate meQTL, despite the current meQTL analysis being the largest to be conducted in left ventricular tissue to date.

Nonetheless, our study has important strengths. Besides being the largest existing study of DCM-linked methylation disturbances in left ventricular tissue to date, the current investigation extends previous methylation studies of DCM by seeking evidence for the causal contribution of DCM-linked sentinel CpGs and by investigating regional associations with traits indicative of CVD risk.

Conclusion

This is the largest investigation of perturbed CpG methylation in DCM to be conducted in disease-relevant left ventricular tissue obtained from patients and controls. We identify CpGs independently and robustly associated with DCM and suggest molecular players in putative causal mechanisms by which DNA methylation may impact DCM. We also provide preliminary indications of the prognostic potential of regions containing DCM-linked methylation alterations that are associated with CVD-relevant traits in the general population.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13148-025-01854-8.

Additional file 1.

Additional file 2.

Additional file 3.

Acknowledgements

None.

Author contributions

This study was conceived, designed, and interpreted by KT and ML. KT performed the statistical analysis. KT drafted the manuscript, and ML contributed to the manuscript writing. DT, WT, HN, PJ and CJM pre-processed the data used for key analyses. EW, MM, GS, CJM, FT, PH, TC, KM, RF supplied and/or processed samples and provided phenotype data. All authors read and approved the final manuscript.

Funding

This work was funded by a start-up grant awarded to Asst Prof Marie Loh [PI] by the Ministry of Education, Singapore (MOE; Grant ID: #002751-00001**).**

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files. Original R scripts are available in GitHub (https://github.com/KonstanzeTan/tan_etal_DCM). Gene regulatory features for enrichment analysis were downloaded from the Roadmap Epigenomics database (https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/). Transcription factor binding sites are available at the ReMap2022 database (https://remap.univ-amu.fr/download_page). Raw RNA-seq counts and accompanying metadata from the MAGNet cohort are available at the following URL: https://github.com/mpmorley/MAGNet?tab=readme-ov-file. Full summary statistics for the GWAS of DCM can be accessed from the NBDC Human Database (https://humandbs.dbcls.jp/en/hum0197-v3-220; dataset ID: hum0197.v3.EUR.DC.v1). The full set of left ventricular eQTL associations (v8) were downloaded from Google Cloud's requester pay buckets (https://console.cloud.google.com/storage/browser/gtex-resources;tab=objects?pli=1&prefix=&forceOnObjectsSortingFiltering=false; dataset ID: Heart_Left_Ventricle.v8.EUR.allpairs.chr*.parquet), using Google's command-line tool, gsutil. Left ventricular chromatin interaction data (H3K27ac-ChIPseq, bulk tissue Hi-C and H3K27ac-based HiChIP data) were requested from authors of Tan et al. (https://doi.org/10.1161/CIRCRESAHA.120.317254). Other datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

For each cohort, written informed consent for the research use of donated left ventricular tissue was obtained. For heart transplant recipients, consent was

obtained from the transplant recipient. For brain-dead organ donors, consent was obtained from the next-of-kin. All analyses and study protocols were approved by the relevant institutional review boards.

Competing interests
The authors declare no competing interests.

Author details
[1]Lee Kong Chian School of Medicine, Nanyang Technological University, Clinical Sciences Building, 11 Mandalay Road, Singapore 308232, Singapore. [2]Cardiovascular Research Institute, National University Health System, Singapore, Singapore. [3]Cardiovascular-Metabolic Disease Translational Research Programme, National University of Singapore, Singapore, Singapore. [4]Genome Institute of Singapore (GIS), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore. [5]Penn Cardiovascular Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [6]Bruce McManus Cardiovascular Biobank, UBC- Centre for Heart Lung Innovation, Vancouver, BC, Canada.

Received: 24 September 2024 Accepted: 28 February 2025
Published online: 08 March 2025

References

1. Reichart D, Magnussen C, Zeller T, Blankenberg S. Dilated cardiomyopathy: from epidemiologic to genetic phenotypes: a translational review of current literature. J Intern Med. 2019;286:362–72. https://doi.org/10.1111/joim.12944.
2. Cuenca S, Ruiz-Cano MJ, Gimeno-Blanes JR, Jurado A, Salas C, Gomez-Diaz I, Padron-Barthe L, Grillo JJ, Vilches C, Segovia J, et al. Genetic basis of familial dilated cardiomyopathy patients undergoing heart transplantation. J Heart Lung Transplant. 2016;35:625–35. https://doi.org/10.1016/j.healun.2015.12.014.
3. Alraies MC, Eckman P. Adult heart transplant: indications and outcomes. J Thorac Dis. 2014;6:1120–8. https://doi.org/10.3978/j.issn.2072-1439.2014.06.44.
4. Schultheiss HP, Fairweather D, Caforio ALP, Escher F, Hershberger RE, Lipshultz SE, Liu PP, Matsumori A, Mazzanti A, McMurray J, et al. Dilated cardiomyopathy. Nat Rev Dis Primers. 2019;5:32. https://doi.org/10.1038/s41572-019-0084-1.
5. Towbin JA, Lowe AM, Colan SD, Sleeper LA, Orav EJ, Clunie S, Messere J, Cox GF, Lurie PR, Hsu D, et al. Incidence, causes, and outcomes of dilated cardiomyopathy in children. JAMA. 2006;296:1867–76. https://doi.org/10.1001/jama.296.15.1867.
6. Halliday BP, Cleland JGF, Goldberger JJ, Prasad SK. Personalizing risk stratification for sudden death in dilated cardiomyopathy: the past, present, and future. Circulation. 2017;136:215–31. https://doi.org/10.1161/CIRCULATIONAHA.116.027134.
7. Al-Khatib SM, Stevenson WG, Ackerman MJ, Bryant WJ, Callans DJ, Curtis AB, Deal BJ, Dickfeld T, Field ME, Fonarow GC, et al. 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. J Am Coll Cardiol. 2018;72:e91–220. https://doi.org/10.1016/j.jacc.2017.10.054.
8. Garnier S, Harakalova M, Weiss S, Mokry M, Regitz-Zagrosek V, Hengstenberg C, Cappola TP, Isnard R, Arbustini E, Cook SA, et al. Genome-wide association analysis in dilated cardiomyopathy reveals two new players in systolic heart failure on chromosomes 3p25.1 and 22q11.23. Eur Heart J. 2021;42:2000–11. https://doi.org/10.1093/eurheartj/ehab030.
9. Meder B, Haas J, Sedaghat-Hamedani F, Kayvanpour E, Frese K, Lai A, Nietsch R, Scheiner C, Mester S, Bordalo DM, et al. Epigenome-wide association study identifies cardiac gene patterning and a novel class of biomarkers for heart failure. Circulation. 2017;136:1528–44. https://doi.org/10.1161/CIRCULATIONAHA.117.027355.
10. Haas J, Frese KS, Park YJ, Keller A, Vogel B, Lindroth AM, Weichenhan D, Franke J, Fischer S, Bauer A, et al. Alterations in cardiac DNA methylation in human dilated cardiomyopathy. EMBO Mol Med. 2013;5:413–29. https://doi.org/10.1002/emmm.201201553.
11. Koczor CA, Lee EK, Torres RA, Boyd A, Vega JD, Uppal K, Yuan F, Fields EJ, Samarel AM, Lewis W. Detection of differentially methylated gene promoters in failing and nonfailing human left ventricle myocardium using computation analysis. Physiol Genomics. 2013;45:597–605. https://doi.org/10.1152/physiolgenomics.00013.2013.
12. Jo BS, Koh IU, Bae JB, Yu HY, Jeon ES, Lee HY, Kim JJ, Choi M, Choi SS. Methylome analysis reveals alterations in DNA methylation in the regulatory regions of left ventricle development genes in human dilated cardiomyopathy. Genomics. 2016;108:84–92. https://doi.org/10.1016/j.ygeno.2016.07.001.
13. Movassagh M, Choy MK, Goddard M, Bennett MR, Down TA, Foo RS. Differential DNA methylation correlates with differential expression of angiogenic factors in human heart failure. PLoS ONE. 2010;5: e8564. https://doi.org/10.1371/journal.pone.0008564.
14. Movassagh M, Choy MK, Knowles DA, Cordeddu L, Haider S, Down T, Siggens L, Vujic A, Simeoni I, Penkett C, et al. Distinct epigenomic features in end-stage failing human hearts. Circulation. 2011;124:2411–22. https://doi.org/10.1161/CIRCULATIONAHA.111.040071.
15. Troughton R, Michael Felker G, Januzzi JL Jr. Natriuretic peptide-guided heart failure management. Eur Heart J. 2014;35:16–24. https://doi.org/10.1093/eurheartj/eht463.
16. Bizet M, Defrance M, Calonne E, Bontempi G, Sotiriou C, Fuks F, Jeschke J. Improving Infinium MethylationEPIC data processing: re-annotation of enhancers and long noncoding RNA genes and benchmarking of normalization methods. Epigenetics. 2022;17:2434–54. https://doi.org/10.1080/15592294.2022.2135201.
17. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48:481–7. https://doi.org/10.1038/ng.3538.
18. Bozkurt B, Colvin M, Cook J, Cooper LT, Deswal A, Fonarow GC, Francis GS, Lenihan D, Lewis EF, McNamara DM, et al. Current diagnostic and treatment strategies for specific dilated cardiomyopathies: a scientific statement from the American Heart Association. Circulation. 2016;134:e579–646. https://doi.org/10.1161/CIR.0000000000000455.
19. Kasturiratne A, Khawaja KI, Ahmad S, Siddiqui S, Shahzad K, Athauda LK, Jayawardena R, Mahmood S, Muilwijk M, Batool T, et al. The iHealth-T2D study, prevention of type 2 diabetes amongst South Asians with central obesity and prediabetes: study protocol for a randomised controlled trial. Trials. 2021;22:928. https://doi.org/10.1186/s13063-021-05803-7.
20. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, Afzal U, Scott J, Jarvelin MR, Elliott P, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 2015;16:37. https://doi.org/10.1186/s13059-015-0600-x.
21. van Iterson M, van Zwet EW, Consortium B, Heijmans BT. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. Genome Biol. 2017;18:19. https://doi.org/10.1186/s13059-016-1131-9.
22. McAllan L, Baranasic D, Villicana S, Brown S, Zhang W, Lehne B, Adamo M, Jenkinson A, Elkalaawy M, Mohammadi B, et al. Integrative genomic analyses in adipocytes implicate DNA methylation in human obesity and diabetes. Nat Commun. 2023;14:2784. https://doi.org/10.1038/s41467-023-38439-z.
23. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518:317–30. https://doi.org/10.1038/nature14248.
24. Hammal F, de Langen P, Bergon A, Lopez F, Ballester B. ReMap 2022: a database of Human, Mouse, Drosophila and Arabidopsis regulatory regions from an integrative analysis of DNA-binding sequencing experiments. Nucleic Acids Res. 2022;50:D316–25. https://doi.org/10.1093/nar/gkab996.
25. Akazawa H, Komuro I. Roles of cardiac transcription factors in cardiac hypertrophy. Circ Res. 2003;92:1079–88. https://doi.org/10.1161/01.RES.0000072977.86706.23.
26. Meganathan K, Sotiriadou I, Natarajan K, Hescheler J, Sachinidis A. Signaling molecules, transcription growth factors and other regulators revealed from in-vivo and in-vitro models for the regulation of cardiac development. Int J Cardiol. 2015;183:117–28. https://doi.org/10.1016/j.ijcard.2015.01.049.

27. Hannenhalli S, Putt ME, Gilmore JM, Wang J, Parmacek MS, Epstein JA, Morrisey EE, Margulies KB, Cappola TP. Transcriptional genomics associates FOX transcription factors with human heart failure. Circulation. 2006;114:1269–76. https://doi.org/10.1161/CIRCULATIONAHA.106.632430.
28. Tan WLW, Anene-Nzelu CG, Wong E, Lee CJM, Tan HS, Tang SJ, Perrin A, Wu KX, Zheng W, Ashburn RJ, et al. Epigenomes of human hearts reveal new genetic variants relevant for cardiac disease and phenotype. Circ Res. 2020;127:761–77. https://doi.org/10.1161/CIRCRESAHA.120.317254.
29. Yang J, Wang D, Yang Y, Yang W, Jin W, Niu X, Gong J. A systematic comparison of normalization methods for eQTL analysis. Brief Bioinform. 2021;22:bbab193. https://doi.org/10.1093/bib/bbab193.
30. Stegle O, Parts L, Piipari M, Winn J, Durbin R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat Protoc. 2012;7:500–7. https://doi.org/10.1038/nprot.2011.457.
31. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics. 2012;28:1353–8. https://doi.org/10.1093/bioinformatics/bts163.
32. Hu R, Morley MP, Brandimarto J, Tucker NR, Parsons VA, Zhao SD, Meder B, Katus HA, Ruhle F, Stoll M, et al. Genetic reduction in left ventricular protein kinase C-alpha and adverse ventricular remodeling in human subjects. Circ Genom Precis Med. 2018;11: e001901. https://doi.org/10.1161/CIRCGEN.117.001901.
33. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, Narita A, Konuma T, Yamamoto K, Akiyama M, et al. A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet. 2021;53:1415–24. https://doi.org/10.1038/s41588-021-00931-x.
34. Consortium GT. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020;369:1318–30. https://doi.org/10.1126/science.aaz1776.
35. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014;10: e1004383. https://doi.org/10.1371/journal.pgen.1004383.
36. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinform. 2008;9:559. https://doi.org/10.1186/1471-2105-9-559.
37. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, Gilad Y, Pritchard JK. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol. 2011;12:R10. https://doi.org/10.1186/gb-2011-12-1-r10.
38. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet. 2006;38:1378–85. https://doi.org/10.1038/ng1909.
39. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinform. 2012;13:86. https://doi.org/10.1186/1471-2105-13-86.
40. Alfulaij N, Meiners F, Michalek J, Small-Howard AL, Turner HC, Stokes AJ. Cannabinoids, the Heart of the Matter. J Am Heart Assoc. 2018;7:e009099. https://doi.org/10.1161/JAHA.118.009099.
41. Long Q, Yang K, Yang Q. Regulation of mitochondrial ATP synthase in cardiac pathophysiology. Am J Cardiovasc Dis. 2015;5:19–32.
42. Han CR, Wang H, Hoffmann V, Zerfas P, Kruhlak M, Cheng SY. Thyroid hormone receptor alpha mutations cause heart defects in zebrafish. Thyroid. 2021;31:315–26. https://doi.org/10.1089/thy.2020.0332.
43. Roux M, Zaffran S. Hox genes in cardiovascular development and diseases. J Dev Biol. 2016;4:14. https://doi.org/10.3390/jdb4020014.
44. Wang K, Diskin SJ, Zhang H, Attiyeh EF, Winter C, Hou C, Schnepp RW, Diamond M, Bosse K, Mayes PA, et al. Integrative genomics identifies LMO1 as a neuroblastoma oncogene. Nature. 2011;469:216–20. https://doi.org/10.1038/nature09609.
45. Sibilitz KL, Benn M, Nordestgaard BG. Creatinine, eGFR and association with myocardial infarction, ischemic heart disease and early death in the general population. Atherosclerosis. 2014;237:67–75. https://doi.org/10.1016/j.atherosclerosis.2014.08.040.
46. Musunuru K, Kral BG, Blumenthal RS, Fuster V, Campbell CY, Gluckman TJ, Lange RA, Topol EJ, Willerson JT, Desai MY, et al. The use of high-sensitivity assays for C-reactive protein in clinical practice. Nat Clin Pract Cardiovasc Med. 2008;5:621–35. https://doi.org/10.1038/ncpcardio1322.
47. Nishiura K, Yokokawa T, Misaka T, Ichimura S, Tomita Y, Miura S, Shimizu T, Sato T, Kaneshiro T, Oikawa M, et al. Prognostic role of circulating LTBP-2 in patients with dilated cardiomyopathy: a novel biomarker reflecting extracellular matrix LTBP-2 accumulation. Can J Cardiol. 2023;39:1436–45. https://doi.org/10.1016/j.cjca.2023.05.015.

Publisher's Note