

1 **Genomic and transcriptomic data analyses highlight *KPNBI* and *MYL4* as novel risk genes for**  
2 **congenital heart disease**

3 Martin Broberg<sup>1,2</sup>, Minna Ampuja<sup>1</sup>, Samuel Jones<sup>2</sup>, Tiina Ojala<sup>3</sup>, Otto Rahkonen<sup>3</sup>, Riikka Kivelä<sup>1,4,5</sup>,  
4 James Priest<sup>6</sup>, FinnGen<sup>a</sup>, Hanna M. Ollila<sup>2,7,8</sup>, Emmi Helle<sup>1,2</sup>

5 <sup>1</sup>Stem Cells and Metabolism Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

6 <sup>2</sup>Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland

7 <sup>3</sup>New Children's Hospital, Pediatric Research Center, Helsinki University Hospital, Helsinki, Finland

8 <sup>4</sup>Wihuri Research Institute, Helsinki, Finland

9 <sup>5</sup>Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

10 <sup>6</sup>Stanford University School of Medicine, Stanford, CA, USA

11 <sup>7</sup>Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA., USA and Program in Medical  
12 and Population Genetics, Broad Institute, Cambridge, MA, USA

13 <sup>8</sup>Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital and Harvard  
14 Medical School, Boston, Massachusetts, USA

15 <sup>a</sup>Supplemental document 1

16

17 **Abstract**

18 Congenital heart defects (CHD) are structural defects of the heart affecting approximately 1% of  
19 newborns. CHDs exhibit a complex inheritance pattern. While genetic factors are known to play an  
20 important role in the development of CHD, relatively few variants have been discovered so far and  
21 very few genome-wide association studies (GWAS) have been conducted. We performed a GWAS  
22 of general CHD and five CHD subgroups in FinnGen followed by functional fine-mapping through  
23 eQTL analysis in the GTEx database, and target validation in human induced pluripotent stem cell -  
24 derived cardiomyocytes (hiPS-CM) from CHD patients. We discovered that the *MYL4-KPNBI* locus  
25 (rs11570508, beta = 0.24, P = 1.2x10<sup>-11</sup>) was associated with the general CHD group. An additional  
26 four variants were significantly associated with the different CHD subgroups. Two of these,

**NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.**

1 rs1342740627 associated with left ventricular outflow tract obstruction defects and rs1293973611  
2 associated with septal defects, were Finnish population enriched. The variant rs11570508 associated  
3 with the expression of *MYL4* (normalized expression score (NES) = 0.1, P = 0.0017, in the atrial  
4 appendage of the heart) and *KPNB1* (NES = -0.037, P = 0.039, in the left ventricle of the heart).  
5 Furthermore, lower expression levels of both genes were observed in human induced pluripotent stem  
6 cell derived cardiomyocytes (hiPSC-CM) from CHD patients compared to healthy controls. Together,  
7 the results demonstrate *KPNB1* and *MYL4* as in a potential genetic risk loci associated with the  
8 development of CHD.

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

## 1 **Introduction**

2 Congenital heart defects (CHD) are globally the most common birth defect, and include a wide range  
3 of structural malformations of the heart and great vessels, which have their origin in the first trimester  
4 of pregnancy (Huisenga et al., 2021; Liu et al., 2019). CHD may occur as an element of a syndrome  
5 or as an isolated defect. Although monogenic cases have been reported, most isolated CHD are  
6 thought to be complex traits with reduced penetrance and variable expressivity (Kerstjens-Frederikse  
7 et al., 2016). In addition to genetics, environmental exposures such as maternal obesity, diabetes, and  
8 certain medications increase the risk for CHD in the offspring (Helle and Priest, 2020; Øyen et al.,  
9 2016; Paterno et al., 2017; Persson et al., 2019). Next-generation sequencing studies have identified  
10 that the haploinsufficiency of NOTCH- and VEGF-signaling pathway genes associate with left  
11 ventricular outflow tract obstruction (LVOTO) defects and Tetralogy of Fallot (TOF), respectively  
12 (Garg, 2006; Helle et al., 2019, p. 1; Reuter et al., 2019). In most isolated CHD cases, however, the  
13 genetic etiology remains to be identified.

14  
15 As genome-wide sequence data are available for large populations, genome-wide association studies  
16 (GWAS) have been used as an additional tool to identify genetic determinants of CHD (Broberg et  
17 al., 2021). A number of studies have identified genome-wide significant loci for all CHD  
18 (rs185531658)(Lahm et al., 2021) as well as for certain CHD subgroups, such as rs11065987 in  
19 Tetralogy of Fallot (Cordell et al., 2013), rs72820264 in conotruncal heart defects and left ventricular  
20 obstructive tract defects (Agopian et al., 2017), and rs387906656 in coarctation of the aorta  
21 (Bjornsson et al., 2018). As genetic and phenotypic data for hundreds of thousands of individuals has  
22 become integrated into large biobank-based projects such as the FinnGen Biobank, UK Biobank or  
23 Japan Biobank, the statistical strength for population-wide GWAS has increased  
24 (<https://finngen.gitbook.io/documentation/>).

25

1 Here, we conduct a GWAS of 2436 CHD cases from FinnGen release 7 (R7). Our analyses identify  
2 one genome-wide significant signal for the total CHD group and a total of four genome-wide  
3 significant signals for the following subgroups; early aortic valve disease (aortic stenosis or  
4 insufficiency diagnosis at < 60 years), septal defects, and LVOTO defects. Using public tissue-  
5 specific expression Quantitative Trait Loci (eQTL) data, we detected significant impact of the CHD  
6 associated SNP on the expression of *MYL4* and *KPNB1* in the heart. In addition, we defined the  
7 expression of these genes in human induced pluripotent stem cell derived cardiomyocytes (hiPS-CM)  
8 from CHD patients and healthy controls.

9

## 10 **Methods**

### 11 **Study cohorts and phenotypes**

#### 12 **FinnGen**

13 FinnGen is a large-scale biobank study that aims to genotype 500,000 Finnish participants recruited  
14 from hospitals as well as prospective and retrospective epidemiological and disease-based cohorts.  
15 These data are combined with longitudinal registries that record phenotypes and health events  
16 (including ICD-based diagnosis) over the entire lifespan including the National Hospital Discharge  
17 Registry (inpatient and outpatient), Causes of Death Registry, the National Infectious Diseases  
18 Registry, Cancer Registry, Primary Health Care Registry (outpatient) and Medication Reimbursement  
19 Registry. This study used data from FinnGen Data Freeze 7, which consisted of 309,154 individuals.

20

21 From the FinnGen population we identified patients with a CHD diagnosis and from those excluded  
22 individuals with a diagnosis for syndromes that are known to include CHD phenotypes. For the  
23 general CHD category, we used 2436 cases and 306 718 controls. In addition, the following CHD  
24 subgroups were analyzed: septal defects (atrial septal defects (ASD) and ventricular septal defects  
25 (VSD), N=1370), conotruncal defects (N=307) LVOTO narrow (LVOTO type congenital heart

1 disease and coarctation of the aorta N=939), early aortic valve disease (including patients under 60  
2 years of age with non-congenital aortic valve stenosis and/or insufficiency diagnosis, N=2183), and  
3 LVOTO broad (including individuals included in the LVOTO narrow and early aortic valve disease  
4 groups N=2313). The diagnoses used for each subgroup and the diagnoses used for exclusion are  
5 listed in Supplemental Table 1.

## 7 **FinnGen ethics statement**

8 Patients and control subjects in FinnGen provided informed consent for biobank research, based on  
9 the Finnish Biobank Act. Alternatively, separate research cohorts, collected before the Finnish  
10 Biobank Act came into effect (in September 2013) and before start of FinnGen (August 2017), were  
11 collected based on study-specific consents and later transferred to the Finnish biobanks after approval  
12 by Fimea (Finnish Medicines Agency), the National Supervisory Authority for Welfare and Health.  
13 Recruitment protocols followed the biobank protocols approved by Fimea. The Coordinating Ethics  
14 Committee of the Hospital District of Helsinki and Uusimaa (HUS) statement number for the  
15 FinnGen study is Nr HUS/990/2017.

16  
17 The FinnGen study is approved by the Finnish Institute for Health and Welfare (permit numbers:  
18 THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018,  
19 THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019 and  
20 THL/1524/5.05.00/2020), Digital and population data service agency (permit numbers:  
21 VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit  
22 numbers: KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA 98/522/2019,  
23 KELA 134/522/2019, KELA 138/522/2019, KELA 2/522/2020, KELA 16/522/2020), Findata permit  
24 numbers THL/2364/14.02/2020, THL/4055/14.06.00/2020, THL/3433/14.06.00/2020,  
25 THL/4432/14.06/2020, THL/5189/14.06/2020, THL/5894/14.06.00/2020,

1 THL/6619/14.06.00/2020, THL/209/14.06.00/2021, THL/688/14.06.00/2021,  
2 THL/1284/14.06.00/2021, THL/1965/14.06.00/2021, THL/5546/14.02.00/2020 and Statistics  
3 Finland (permit numbers: TK-53-1041-17 and TK/143/07.03.00/2020 (earlier TK-53-90-20)).  
4

5 The Biobank Access Decisions for FinnGen samples and data utilized in FinnGen Data Freeze 7  
6 include: THL Biobank BB2017\_55, BB2017\_111, BB2018\_19, BB\_2018\_34, BB\_2018\_67,  
7 BB2018\_71, BB2019\_7, BB2019\_8, BB2019\_26, BB2020\_1, Finnish Red Cross Blood Service  
8 Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, Auria Biobank AB17-5154 and amendment  
9 #1 (August 17 2020), Biobank Borealis of Northern Finland\_2017\_1013, Biobank of Eastern Finland  
10 1186/2018 and amendment 22 § /2020, Finnish Clinical Biobank Tampere MH0004 and amendments  
11 (21.02.2020 & 06.10.2020), Central Finland Biobank 1-2017, and Terveystalo Biobank STB  
12 2018001.  
13

#### 14 **SAIGE GWAS analysis**

15 The GWAS for the CHD categories was performed using SAIGE on the FinnGen R7 cohort (N=309  
16 154) (Zhou et al., 2020). The primary settings for the SAIGE analysis were; ratio coefficient of  
17 variation (CV) cutoff: 0.001, trace CV cutoff: 0.0025, minmac: 5. The covariates used for the analysis  
18 included: sex, imputed age, age at death or end of follow up, 10 principal components and microarray  
19 genotyping batch (full list in Supplemental Table 2). The FinnGen variant filtering for PCA and  
20 GWAS analysis is: exclusion of variants with info score < 0.95, exclusion of variants with  
21 missingness > 0.01 (based on the GP; see conversion), exclusion of variants with MAF < 0.05 and  
22 LD pruning with window interval of 500kb and 50kb steps, and an  $r^2$  filter of 0.1  
23 (<https://finngen.gitbook.io/documentation/methods/phewas/quality-checks>).  
24

#### 25 **Sensitivity fine-mapping**

1 GWAS significant variants were re-tested for association using REGENIE (Mbatchou et al., 2021)  
2 fine-mapped using SuSiE and FINEMAP (Benner et al., 2016; Wang et al., 2020). The general  
3 FinnGen fine-mapping pipeline is described (<https://github.com/FINNGEN/finemapping-pipeline>).

#### 4 5 **Calculating eQTL data**

6 To test whether the genome-wide significant SNPs from the FinnGen GWAS could have an effect on  
7 the expression of nearby genes, we utilized the eQTL calculator function of the GTEx portal database  
8 ([www.gtexportal.org](http://www.gtexportal.org)) (Lonsdale et al., 2013; Stanfill and Cao, 2021). We examined expression in the  
9 following tissues: ‘heart - atrial appendage’ and ‘heart – left ventricle’.

10

#### 11 **Defining endpoint associations using PheWeb**

12 In addition to CHD, FinnGen contains diagnosis codes and precomputed phenotypes. Therefore,  
13 instead of looking at each disease or GWAS alone, it is possible to examine a single significant variant  
14 horizontally across other diseases to understand which other diseases are associated with this variant.  
15 The FinnGen results website uses the PheWeb design interface for displaying SNP-endpoint  
16 associations for a convenient overview of GWAS results (Gagliano Taliun et al., 2020). This approach  
17 was used to identify other endpoints associated with the lead SNPs. The FinnGen R7 data will become  
18 available to the public similarly as release 5.

19

#### 20 **Generation of pluripotent stem cell lines**

21 Human induced pluripotent stem cells (hiPSCs) were reprogrammed at the Biomedicum Stem Cell  
22 Center Core Facility (University of Helsinki). The cell lines were derived from four male individuals  
23 with LVOTO defects, (three patients with HLHS and one patient with congenital aortic stenosis (AS)  
24 and coarctation of the aorta (COA)). The cell lines were created by using retroviral/Sendai virus  
25 transduction of Oct3/4, Sox2, Klf4, and c-Myc, as described previously (Trokovic et al., 2015). Three

1 control cell lines were obtained from The Biomedicum Stem Cell Center Core Facility (University of  
2 Helsinki) and one control line was a kind gift from Professor Anu Suomalainen-Wartiovaara. All  
3 control cell lines were male. The cell lines were differentiated to hiPS-CMs as previously described  
4 (Helle et al., 2021). Briefly, Wnt-pathway activation and inactivation were used to induce  
5 differentiation, and metabolic selection with media containing lactate instead of glucose was used to  
6 ensure purity of the CM culture. The cells were processed for single cell sequencing as described  
7 previously (Helle et al., 2021). The cells were detached using Multi Tissue Dissociation Kit 3 (130-  
8 110-204, Miltenyi Biotec), according to the manufacturer's instructions. The cells were washed once  
9 with PBS containing 0.04% BSA, and then resuspended in PBS with 0.04% BSA to a concentration  
10 of  $2.0 \times 10^6$  cells/ml. The cells were passed through a 35  $\mu\text{m}$  strainer (352235, Corning) and  
11 combined into two samples (one containing cells from healthy individuals, one containing patient  
12 cells) and placed on ice until continuation of the 10 $\times$  Genomics Single Cell Protocol at the Institute  
13 of Molecular Medicine Finland (FIMM).

14

#### 15 **Ethics statement of the hiPSC work**

16 The hiPSC-study protocol was reviewed and approved by Coordinating Ethics Committee of the HUS  
17 Hospital District. Written informed consent was obtained from the legal guardians of the participants  
18 and also directly from all participants aged over 6 years.

19

#### 20 **Single-cell RNA sequencing and analysis**

21 Single-cell gene expression profiles were studied using the 10X Genomics Chromium Single Cell  
22 3'RNAseq platform. The Chromium Single Cell 3'RNAseq run and library preparation were made  
23 using the Chromium Single Cell 3' Reagent version 3 chemistry. The sample libraries were sequenced  
24 on Illumina NovaSeq 6000 system. 12000 cells/combined sample and 50,000 PE/cell were analyzed.  
25 Data processing and analysis were performed using 10X Genomics Cell Ranger v2.1.1 pipelines. The



1 “cellranger mkfastq” pipeline was used to produce FASTQ (raw data) files. The “cellranger count”  
2 was used to perform alignment, filtering and UMI counting. Mkfastq was run using the Illumina  
3 bcl2fastq v2.2.0 and alignment was done against human genome GRCh38. Cellranger “aggr” pipeline  
4 was used to combine data from multiple samples into an experiment-wide gene-barcode matrix and  
5 analysis. The scRNAseq data was analyzed using the R package Seurat v3 (Satija et al., 2015). The  
6 data was filtered using a mitochondrial RNA percentage < 30, over 200 gene features per cell and at  
7 least 1000 identified RNA transcripts per cell. The control samples were pooled and used to compare  
8 against individual patients, and against the pooled patient data.

9  
10 **HLA allele fine-mapping association with early aortic valve disease**

11 We used the FinnGen imputed dosages of HLA alleles (HLA-A, B, C, DPB1, DQA1, DQB1, DRB1,  
12 DRB3, DRB4 and DRB5) for the R7 cohort. These alleles have been previously imputed using a  
13 Finnish specific imputation panel (Ritari et al., 2020). We then performed logistic regression using  
14 the binominal R generalized linear model (glm) function. As covariates, the principal components  
15 (PC) 1-20 for the cohort were used, and age, sex, age<sup>2</sup> and age\*sex.

16

## 1 **Results**

### 2 **FinnGen GWAS highlights SNPs significantly associated with CHD**

3 From the general CHD category GWAS we detected one lead SNP in chromosome 17, rs11570508  
4 (beta = 0.24, P =  $1.2 \times 10^{-11}$ , intronic region of CDC27) (Figure 1A,B, Table 1), and fine-mapping of  
5 the signal suggested this variant to be the primary causal variant (Supplemental Table 3).  
6 Furthermore, we detected GWAS significant lead SNPs on chromosome 17 associated with septal  
7 defects (rs1293973611, beta = 4.96, P =  $1.1 \times 10^{-8}$ , intronic region of RP11-1055B8.3), chromosome  
8 20 associated with LVOTO broad (rs1342740627, beta = 4.8, P =  $1.8 \times 10^{-8}$ , intronic region of  
9 *GGTLC1*) and two SNPs in chromosomes 6 and 20 associated with early aortic valve disease  
10 (rs17214148, beta = 3.2, P =  $3.4 \times 10^{-8}$ , 3 prime UTR-region of *HLA-DOA*; and rs1342740627, beta =  
11 4.9, P =  $3.4 \times 10^{-8}$ ) (Supplemental Figure 2, Table 1). Sensitivity analysis using fine-mapping  
12 confirmed that the lead SNPs rs1293973611 SNP associated with septal defects, and the  
13 rs1342740627 SNP associated with LVOTO broad. These two SNPs were also found to be greatly  
14 enriched in the Finnish population (Table 1). GWAS diagnostics were assayed with QQ-plots  
15 (Supplemental Figure 3). No GWAS significant SNPs were identified for LVOTO narrow or  
16 conotruncal defects. To fine map the association with the *HLA-DOA* region, we performed a glm  
17 regression showed a significant (Bonferroni corrected for multiple testing) association for 6 HLA  
18 alleles and early aortic valve disease (Supplemental Table 4). The most significant was DRB1\*01:02  
19 (beta = 1.14, P =  $6.9 \times 10^{-6}$ ).

20

21

### 22 **eQTL analysis on the lead SNPs**

23 The eQTL calculation demonstrated a potential effect by rs11570508 on the expression of two nearby  
24 genes: *MYL4* (eQTL P=0.0017 in heart - atrial appendage) and *KPNB1* (eQTL P=0.04 in heart - left  
25 ventricle). The eQTL calculation did not indicate significant effects on nearby genes for the other

1 four GWAS significant SNPs. Thus, we decided to further explore the expression of *KPNBI* and  
2 *MYL4* in different scRNAseq datasets. Furthermore, a FUMA-based analysis predicted chromatin  
3 interactions using FUMA and associated eQTL data (Watanabe et al., 2017), suggesting an interaction  
4 between rs11570508 and *KPNBI* (Supplemental Figure 3).

5

6

### 7 **Genetic association results of FinnGen endpoints for the lead SNPs**

8 In Figure 1C we show that rs11570508 is also significantly associated with other FinnGen  
9 cardiovascular endpoints, such as hypertensive diseases, atrial fibrillation and flutter, and use of  
10 antihypertensive medications (Figure 1C). The lead SNP for early aortic valve disease (rs17214148),  
11 is also significantly associated with the FinnGen endpoints for hypothyroidism disorders and  
12 disorders of the thyroid gland (Supplemental Figure 4). The other three lead SNPs did not have  
13 additional significant associations beyond CHD.

14

### 15 **Expression of *MYL4* and *KPNBI* in the cardiogenic region of early mouse embryos and in 16 organotypic hiPSCs**

17 Since our eQTL results suggested that rs11570508 (located in *CDC27*) could influence the expression  
18 of *MYL4* and *KPNBI* in the heart, we explored the expression of these two genes in published single-  
19 cell RNA seq data from the cardiogenic region of early mouse embryos (de Soysa et al., 2019). The  
20 data were visualized at the UCSC cell browser (Speir et al., 2021)(Figure 2A). The *KPNBI* homolog  
21 *Kpnb1* was expressed in all cells of the region, whereas the *MYL4* homolog *Myl4* was mainly  
22 expressed in the myocardium. Furthermore, we analyzed the expression of *MYL4* and *KPNBI* in  
23 hiPSC-derived endothelial cells (ECs) and cardiomyocytes (CMs) from a healthy individual grown  
24 in monoculture or in a coculture containing both cell types (Helle et al., 2021)(Figure 2B). The hiPS-  
25 ECs express the EC marker *KDR/VEGFR2*, and the hiPS-CMs express the cardiomyocyte marker

1 *TNNT2*, defining the different cell types in the scRNAseq data (Figure 2B). As shown in Figure 2B,  
2 *MYL4* was highly expressed in the hiPS-CMs, and low expression levels were also detected in cardiac  
3 like hiPS-ECs that had been cultured together with hiPS-CMs (EC-CM Coculture), while no  
4 expression was detected in hiPS-ECs cultured alone (EC Monoculture). *KPNB1* was expressed in  
5 both cell types at similar levels, similar to what was observed in the embryonic mouse heart (Figure  
6 2B).

7

### 8 **HiPS-CMs derived from CHD patients have lower expression of *MYL4* and *KPNB1***

9 Next, we compared the expression of *MYL4* and *KPNB1* in scRNAseq data from hiPS-CMs  
10 differentiated from hiPSCs of four LVOTO patients and four healthy controls. We observed that the  
11 expression levels of *KPNB1* and *MYL4* were lower in three of the four LVOTO patients as compared  
12 with healthy controls (Figure 3B). Comparing the pooled patient samples to controls also showed a  
13 significant downregulation of *KPNB1* as well: log2 fold change (log2FC) = -0.4, P = 5.2x10<sup>-49</sup>.

14

### 15 **Discussion**

16 Here we report the results of a GWAS of general CHD traits in the FinnGen R7 database, as well as  
17 of the CHD subgroups; septal defects, LVOTO (broad group) and early aortic valve disease (age <  
18 60 years). We detected one genome-wide significant SNP associated with CHD in general, and four  
19 genome-wide significant SNPs for the subgroups. According to eQTL analysis, the general CHD  
20 associated SNP rs11570508 associated with the expression of *KPNB1* and *MYL4* in heart tissue, and  
21 a chromatin interactions analysis suggested an interaction between rs11570508 and *KPNB1*. Further,  
22 we demonstrated that the expression of these genes was lower in CHD patient-derived iPSC-  
23 cardiomyocytes compared to healthy controls.

24

1 Two SNPs, rs1293973611 associated with septal defects, and rs1342740627 associated with LVOTO  
2 broad and early aortic valve disease, were highly enriched in the Finnish population. These loci have  
3 not been associated with CHD or aortic valve disease previously. The finding highlights the value of  
4 the FinnGen project in identifying novel rare or low-frequency disease variants in an isolated  
5 population with enrichment of variants due to recent population bottlenecks followed by rapid  
6 population growth (Table 1).

7 The SNP rs11570508 associated with general CHD is located in a region that contains several genes  
8 linked to heart development, CHD and other cardiac phenotypes; *MYLA*, *CDC27*, *GOSR2* and *WNT9B*  
9 (Goddard et al., 2017; Lahm et al., 2021; Orr et al., 2015; Sabour et al., 2018). A recent GWAS study  
10 on patients with anomalies of thoracic arteries and veins identified three significant SNPs within the  
11 *GOSR2* locus in chromosome 17 close to rs11570508 (Lahm et al., 2021), and the  
12 chr17:45013271\_T\_C (rs17608766) variant in *GOSR2* has been associated with a reduced aortic  
13 valve area (Córdova-Palomera et al., 2020) indicating importance of this locus in cardiac  
14 development.

15

16 In the eQTL database GTEx (Analysis Release V8 dbGaP Accession phs000424.v8.p2), rs11570508  
17 was predicted to influence the expression of *KPNB1* and *MYLA* in the heart. *KPNB1* encodes nuclear  
18 import receptor karyopherin  $\beta$ 1, which is a nuclear importer protein involved in the translocation of  
19 proteins and part of the regulation of mitosis (Zhu et al., 2018). It has not been linked to CHD  
20 previously, but has been reported as a potential oncogene (Kodama et al., 2017). Our analysis showed  
21 that *KPNB1* homolog is expressed in the developing cardiogenic region in the developing murine  
22 heart, and hiPSC-derived CMs and ECs, suggesting that it might have a role during embryonic  
23 development in CHD-relevant cell types.

24

1 *MYL4* encodes the atrial light chain-1 (ALC-1) protein, which is expressed in the human heart during  
2 development (Barton and Buckingham, 1985). After birth, the expression decreases in the ventricles  
3 but remains in the atria (Wang et al., 2018). ALC-1 has a role in sarcomere assembly and fine tuning  
4 of cardiac contractility. The expression of *MYL4* has shown to abnormally persist in ventricular tissue  
5 of individuals with CHD (England and Loughna, 2013), and the ventricular analog *MYL3* is replaced  
6 by re-expression of *MYL4* in failing and hypertrophied hearts, resembling a fetal remodeling pattern  
7 associated with ventricular dysfunction (Wang et al., 2019, 2018). *MYL4* loss-of-function variants  
8 have been associated with atrial cardiomyopathy and atrial arrhythmias such as atrial fibrillation  
9 (AF)(Gudbjartsson et al., 2017, 2015; Orr et al., 2015; Peng et al., 2017), demonstrating that although  
10 previously not associated with CHD directly, its role in cardiac development and function is  
11 important, and thus, could be a plausible candidate gene for CHD.

12

13 Lower expression levels of *MYL4* and *KPNB1* were seen in the scRNAseq data of three of four hiPS-  
14 CM cell lines derived from LVOTO patients as compared to healthy controls. Recent studies have  
15 indicated that despite the genetic heterogeneity, similar abnormal phenotypes and gene expression  
16 profiles in LVOTO patient derived hiPS-CMs can be observed, indicating for example diminished  
17 capacity for cardiomyocyte differentiation, more immature transcriptomic profile, and reduced  
18 expression of cardiac transcription factors (Jiang et al., 2014; Kobayashi et al., 2014; Krane et al.,  
19 2021; Theis et al., 2015; Yang et al., 2017). The lower expression levels of *MYL4* and *KPNB1* in  
20 patient-derived cells may, thus, indicate a role of these genes in the development of CHD.

21

22 Interestingly, in addition to CHD, the FinnGen Atrial fibrillation (AF) and flutter, and Hypertensive  
23 diseases endpoints were significantly associated with chr17:47151194:C\_A (rs11570508). Children  
24 and adults with CHD have an increased risk for AF (Labombarda et al., 2017; Mandalenakis et al.,

1 2018), and adults with CHD have increased cardiovascular morbidity relative to the general  
2 population (Saha et al., 2019). Although anatomical factors such as anomalous vessel anatomy and  
3 abnormal atrial hemodynamics, and disease related conditions in the heart such as progressive  
4 valvulopathy, residual shunts, and atrial scars from previous heart surgery are likely to be important  
5 predisposes for AF and other cardiovascular morbidity in CHD, it is tempting to speculate that shared  
6 genetic factors could potentially increase the risk for both developing CHD and later cardiovascular  
7 morbidity. In fact, atrial arrhythmias are a common phenotype in Holt-Oram syndrome, a  
8 developmental disorder that leads to heart and limb malformations. Holt-Oram syndrome is partially  
9 caused by variants in the *TBX5* gene (Basson et al., 1997; Vanlerberghe et al., 2019), and GWAS  
10 studies on AF have shown several significant loci with nearest genes known to be important heart  
11 developmental and causal for CHD, such as *NKX2-5*, *GATA4*, *GJA1*, and *GJA5* (Pierpont et al., 2018;  
12 Roselli et al., 2020). It has been also shown that women whose infants have congenital heart defects  
13 have an increased risk and earlier disease onset for cardiovascular morbidity, including  
14 atherosclerotic cardiovascular disease, ischemic heart disease, hospitalization due to cardiac diseases  
15 and cardiac transplantation (Auger et al., 2018). Thus, these associations could indicate shared genetic  
16 risks.

17

18 The eQTL analysis did not identify associations for genes expressed in adult cardiac tissues for the  
19 rs1293973611 variant associated with septal defects, the rs1342740627 variant associated with  
20 LVOTO broad, or the rs17214148 or rs1342740627 variants associated with early aortic valve  
21 disease. In addition, their nearest genes have not been associated with CHD before. Thus, at present,  
22 we do not have evidence for the functional causality for these variants. However, it is well known  
23 that the gene expression in the adult heart is significantly different from that of the development of  
24 cardiac structures, and thus absence of association with adult heart gene expression does not rule out  
25 potential association with gene expression in the embryonic heart. Interestingly, the beta values for

1 all these three very rare variants were quite big, potentially indicating large effect sizes. This finding  
2 has important implications for the genetic architecture of CHD suggesting that rare variants can be  
3 associated with a substantial risk for rare disease such as CHD.

4 Bicuspid aortic valve disease, and other less severe developmental aortic valve defects can be  
5 asymptomatic during childhood and cause aortic valve insufficiency or stenosis later in life. Calcific  
6 aortic valve stenosis manifests typically during the eighth decade of life (Martinsson et al., 2015), and  
7 it is likely that a developmental defect underlies cases diagnosed at younger ages suggesting two  
8 different disease entities. Thus, in addition to those with CHD diagnosis, we created a subgroup of  
9 those who received aortic valve stenosis or insufficiency diagnosis under the age of 60 years. GWAS  
10 loci in genes such as LPA and PALMD have been repeatedly associated with calcific aortic valve  
11 stenosis (Helgadottir et al., 2018; Thanassoulis et al., 2013; Thériault et al., 2018), however, those  
12 associations were not found in our study. Instead, two other significant loci were identified, of which  
13 one was significant even in the LVOTO broad group including patients with congenital LVOTO  
14 diagnoses. This finding could suggest a distinct genetic and pathophysiological etiology for early  
15 onset aortic valve disease in this population, which could be closer to that of CHD.

16

17 Our approach has some limitations. The discovery of novel loci in GWAS studies does not prove  
18 functional causality, and validation studies should be conducted. In addition, ICD10-based studies  
19 are limited by the fact that mild CHD such as BAV and hemodynamically insignificant ASD are often  
20 underdiagnosed as they may be asymptomatic. In addition, our association results are from the  
21 isolated Finnish population, which is genetically unique, and some of the identified variants were  
22 enriched in the Finnish population preventing replication in other populations. On the other hand, our  
23 study highlights the potential of FinnGen to identify disease associated high-impact variants that are  
24 very rare or absent in other populations.



1

2 In summary, our analysis of 2436 cases and 306718 controls identifies a robust risk locus for CHD  
3 near candidate genes with a known role in heart development, and *MYL4* and *KPNB1* were  
4 demonstrated as potentially causal genes for CHD. In addition, three risk loci for three CHD subtypes  
5 had a high beta value, suggesting that very rare variants can play a substantial role in CHD. The  
6 associations between the identified locus and other cardiovascular conditions, such as AF and flutter  
7 and hypertension suggest shared genetic etiology in developmental defects and adult cardiovascular  
8 morbidity.

9

## 10 **Data availability**

11 ScRNAseq data files used for gene expression analysis will be made available in a public repository  
12 prior to publication.

13

## 14 **Acknowledgements**

15 This work was funded by grants from the Finnish Cultural Foundation (E.H, T.O. and M.B.), Finnish  
16 Foundation for Pediatric Research (E.H, T.O and M.B.), the Finnish Medical Foundation (E.H),  
17 Helsinki University (E.H), Helsinki University Hospital (E.H), and The Academy of Finland (E.H).  
18 Furthermore, we would like to thank the FinnGen consortium for GWAS summary statistics. We also  
19 thank the FIMM Single-Cell Analytics unit supported by HiLIFE and Biocenter Finland for single  
20 cell services.

21

## 22 **References**

- 1 Agopian, A.J., Goldmuntz, E., Hakonarson, H., Sewda, A., Taylor, D., Mitchell, L.E., 2017. Genome-Wide  
2 Association Studies and Meta-Analyses for Congenital Heart Defects. *Circ. Cardiovasc. Genet.* 10.  
3 <https://doi.org/10.1161/CIRCGENETICS.116.001449>
- 4 Auger, N., Potter, B.J., Bilodeau-Bertrand, M., Paradis, G., 2018. Long-Term Risk of Cardiovascular Disease in  
5 Women Who Have Had Infants With Heart Defects. *Circulation* 137, 2321–2331.  
6 <https://doi.org/10.1161/CIRCULATIONAHA.117.030277>
- 7 Barton, P.J.R., Buckingham, M.E., 1985. The myosin alkali light chain proteins and their genes. *Biochem. J.*  
8 231, 249–261. <https://doi.org/10.1042/bj2310249>
- 9 Basson, C.T., Bachinsky, D.R., Lin, R.C., Levi, T., Elkins, J.A., Soultis, J., Grayzel, D., Kroumpouzou, E., Traill,  
10 T.A., Leblanc-Straceski, J., Renault, B., Kucherlapati, R., J.G. S., Seidman, C.E., 1997. Mutations in  
11 human cause limb and cardiac malformation in Holt-Oram syndrome. *Nat. Genet.* 15, 30–35.  
12 <https://doi.org/10.1038/ng0197-30>
- 13 Benner, C., Spencer, C.C.A., Havulinna, A.S., Salomaa, V., Ripatti, S., Pirinen, M., 2016. FINEMAP: efficient  
14 variable selection using summary data from genome-wide association studies. *Bioinformatics* 32,  
15 1493–1501. <https://doi.org/10.1093/bioinformatics/btw018>
- 16 Bjornsson, T., Thorolfsson, R.B., Sveinbjornsson, G., Sulem, P., Norddahl, G.L., Helgadóttir, A.,  
17 Gretarsdóttir, S., Magnusdóttir, A., Danielsen, R., Sigurdsson, E.L., Adalsteinsdóttir, B., Gunnarsson,  
18 S.I., Jonsdóttir, I., Arnar, D.O., Helgason, H., Gudbjartsson, T., Gudbjartsson, D.F., Thorsteinsdóttir,  
19 U., Holm, H., Stefansson, K., 2018. A rare missense mutation in MYH6 associates with non-  
20 syndromic coarctation of the aorta. *Eur. Heart J.* 39, 3243–3249.  
21 <https://doi.org/10.1093/eurheartj/ehy142>
- 22 Broberg, M., Hästbacka, J., Helle, E., 2021. From Stem Cells to Populations—Using hiPSC, Next-Generation  
23 Sequencing, and GWAS to Explore the Genetic and Molecular Mechanisms of Congenital Heart  
24 Defects. *Genes* 12, 921. <https://doi.org/10.3390/genes12060921>
- 25 Cordell, H.J., Topf, A., Mamasoula, C., Postma, A.V., Bentham, J., Zelenika, D., Heath, S., Blue, G., Cosgrove,  
26 C., Granados Riveron, J., Darlay, R., Soemedi, R., Wilson, I.J., Ayers, K.L., Rahman, T.J., Hall, D.,  
27 Mulder, B.J.M., Zwinderman, A.H., van Engelen, K., Brook, J.D., Setchfield, K., Bu’Lock, F.A.,  
28 Thornborough, C., O’Sullivan, J., Stuart, A.G., Parsons, J., Bhattacharya, S., Winlaw, D., Mital, S.,  
29 Gwillig, M., Breckpot, J., Devriendt, K., Moorman, A.F.M., Rauch, A., Lathrop, G.M., Keavney, B.D.,  
30 Goodship, J.A., 2013. Genome-wide association study identifies loci on 12q24 and 13q32 associated  
31 with Tetralogy of Fallot. *Hum. Mol. Genet.* 22, 1473–1481. <https://doi.org/10.1093/hmg/dd552>
- 32 Córdova-Palomera, A., Tcheandjieu, C., Fries, J.A., Varma, P., Chen, V.S., Fiterau, M., Xiao, K., Tejada, H.,  
33 Keavney, B.D., Cordell, H.J., Tanigawa, Y., Venkataraman, G., Rivas, M.A., Ré, C., Ashley, E., Priest,  
34 J.R., 2020. Cardiac Imaging of Aortic Valve Area From 34 287 UK Biobank Participants Reveals Novel  
35 Genetic Associations and Shared Genetic Comorbidity With Multiple Disease Phenotypes. *Circ.*  
36 *Genomic Precis. Med.* 13. <https://doi.org/10.1161/CIRCGEN.120.003014>
- 37 de Soysa, T.Y., Ranade, S.S., Okawa, S., Ravichandran, S., Huang, Y., Salunga, H.T., Schrick, A., del Sol, A.,  
38 Gifford, C.A., Srivastava, D., 2019. Single-cell analysis of cardiogenesis reveals basis for organ-level  
39 developmental defects. *Nature* 572, 120–124. <https://doi.org/10.1038/s41586-019-1414-x>
- 40 England, J., Loughna, S., 2013. Heavy and light roles: myosin in the morphogenesis of the heart. *Cell. Mol.*  
41 *Life Sci.* 70, 1221–1239. <https://doi.org/10.1007/s00018-012-1131-1>
- 42 Gagliano Taliun, S.A., VandeHaar, P., Boughton, A.P., Welch, R.P., Taliun, D., Schmidt, E.M., Zhou, W.,  
43 Nielsen, J.B., Willer, C.J., Lee, S., Fritsche, L.G., Boehnke, M., Abecasis, G.R., 2020. Exploring and  
44 visualizing large-scale genetic associations by using PheWeb. *Nat. Genet.* 52, 550–552.  
45 <https://doi.org/10.1038/s41588-020-0622-5>
- 46 Garg, V., 2006. Insights into the genetic basis of congenital heart disease. *Cell. Mol. Life Sci.* 63, 1141–1148.  
47 <https://doi.org/10.1007/s00018-005-5532-2>
- 48 Goddard, L.M., Duchemin, A.-L., Ramalingam, H., Wu, B., Chen, M., Bamezai, S., Yang, J., Li, L., Morley, M.P.,  
49 Wang, T., Scherrer-Crosbie, M., Frank, D.B., Engleka, K.A., Jameson, S.C., Morrissey, E.E., Carroll, T.J.,  
50 Zhou, B., Vermot, J., Kahn, M.L., 2017. Hemodynamic Forces Sculpt Developing Heart Valves  
51 through a KLF2-WNT9B Paracrine Signaling Axis. *Dev. Cell* 43, 274–289.e5.  
52 <https://doi.org/10.1016/j.devcel.2017.09.023>

- 1 Gudbjartsson, D.F., Helgason, H., Gudjonsson, S.A., Zink, F., Oddson, A., Gylfason, A., Besenbacher, S.,  
2 Magnusson, G., Halldorsson, B.V., Hjartarson, E., Sigurdsson, G.T., Stacey, S.N., Frigge, M.L., Holm,  
3 H., Saemundsdottir, J., Helgadottir, H.T., Johannsdottir, H., Sigfusson, G., Thorgeirsson, G.,  
4 Sverrisson, J.T., Gretarsdottir, S., Walters, G.B., Rafnar, T., Thjodleifsson, B., Bjornsson, E.S.,  
5 Olafsson, S., Thorarinsdottir, H., Steingrimsdottir, T., Gudmundsdottir, T.S., Theodors, A., Jonasson,  
6 J.G., Sigurdsson, A., Bjornsdottir, G., Jonsson, J.J., Thorarensen, O., Ludvigsson, P., Gudbjartsson, H.,  
7 Eyjolfsson, G.I., Sigurdardottir, O., Olafsson, I., Arnar, D.O., Magnusson, O.T., Kong, A., Masson, G.,  
8 Thorsteinsdottir, U., Helgason, A., Sulem, P., Stefansson, K., 2015. Large-scale whole-genome  
9 sequencing of the Icelandic population. *Nat. Genet.* 47, 435–444. <https://doi.org/10.1038/ng.3247>
- 10 Gudbjartsson, D.F., Holm, H., Sulem, P., Masson, G., Oddsson, A., Magnusson, O.Th., Saemundsdottir, J.,  
11 Helgadottir, H.Th., Helgason, H., Johannsdottir, H., Gretarsdottir, S., Gudjonsson, S.A., Njølstad, I.,  
12 Løchen, M.-L., Baum, L., Ma, R.C.W., Sigfusson, G., Kong, A., Thorgeirsson, G., Sverrisson, J.Th.,  
13 Thorsteinsdottir, U., Stefansson, K., Arnar, D.O., 2017. A frameshift deletion in the sarcomere gene  
14 *MYL4* causes early-onset familial atrial fibrillation. *Eur. Heart J.* 38, 27–34.  
15 <https://doi.org/10.1093/eurheartj/ehw379>
- 16 Helgadottir, A., Thorleifsson, G., Gretarsdottir, S., Stefansson, O.A., Tragante, V., Thorolfsson, R.B.,  
17 Jonsdottir, I., Bjornsson, T., Steinthorsdottir, V., Verweij, N., Nielsen, J.B., Zhou, W., Folkersen, L.,  
18 Martinsson, A., Heydarpour, M., Prakash, S., Oskarsson, G., Gudbjartsson, T., Geirsson, A., Olafsson,  
19 I., Sigurdsson, E.L., Almgren, P., Melander, O., Franco-Cereceda, A., Hamsten, A., Fritsche, L., Lin,  
20 M., Yang, B., Hornsby, W., Guo, D., Brummett, C.M., Abecasis, G., Mathis, M., Milewicz, D., Body,  
21 S.C., Eriksson, P., Willer, C.J., Hveem, K., Newton-Cheh, C., Smith, J.G., Danielsen, R., Thorgeirsson,  
22 G., Thorsteinsdottir, U., Gudbjartsson, D.F., Holm, H., Stefansson, K., 2018. Genome-wide analysis  
23 yields new loci associating with aortic valve stenosis. *Nat. Commun.* 9, 987.  
24 <https://doi.org/10.1038/s41467-018-03252-6>
- 25 Helle, E., Ampuja, M., Dainis, A., Antola, L., Temmes, E., Tolvanen, E., Mervaala, E., Kivelä, R., 2021. HiPS-  
26 Endothelial Cells Acquire Cardiac Endothelial Phenotype in Co-culture With hiPS-Cardiomyocytes.  
27 *Front. Cell Dev. Biol.* 9, 715093. <https://doi.org/10.3389/fcell.2021.715093>
- 28 Helle, E., Córdova-Palomera, A., Ojala, T., Saha, P., Potiny, P., Gustafsson, S., Ingelsson, E., Bamshad, M.,  
29 Nickerson, D., Chong, J.X., University of Washington Center for Mendelian Genomics, Ashley, E.,  
30 Priest, J.R., 2019. Loss of function, missense, and intronic variants in *NOTCH1* confer different risks  
31 for left ventricular outflow tract obstructive heart defects in two European cohorts: HELLE ET AL.  
32 *Genet. Epidemiol.* 43, 215–226. <https://doi.org/10.1002/gepi.22176>
- 33 Helle, E., Priest, J.R., 2020. Maternal Obesity and Diabetes Mellitus as Risk Factors for Congenital Heart  
34 Disease in the Offspring. *J. Am. Heart Assoc.* 9. <https://doi.org/10.1161/JAHA.119.011541>
- 35 Huisenga, D., La Bastide-Van Gemert, S., Van Bergen, A., Sweeney, J., Hadders-Algra, M., 2021.  
36 Developmental outcomes after early surgery for complex congenital heart disease: a systematic  
37 review and meta-analysis. *Dev. Med. Child Neurol.* 63, 29–46. <https://doi.org/10.1111/dmnc.14512>
- 38 Jiang, Y., Habibollah, S., Tilgner, K., Collin, J., Barta, T., Al-Aama, J.Y., Tesarov, L., Hussain, R., Trafford, A.W.,  
39 Kirkwood, G., Sernagor, E., Eleftheriou, C.G., Przyborski, S., Stojković, M., Lako, M., Keavney, B.,  
40 Armstrong, L., 2014. An Induced Pluripotent Stem Cell Model of Hypoplastic Left Heart Syndrome  
41 (HLHS) Reveals Multiple Expression and Functional Differences in HLHS-Derived Cardiac Myocytes.  
42 *STEM CELLS Transl. Med.* 3, 416–423. <https://doi.org/10.5966/sctm.2013-0105>
- 43 Kerstjens-Frederikse, W.S., van de Laar, I.M.B.H., Vos, Y.J., Verhagen, J.M.A., Berger, R.M.F., Lichtenbelt,  
44 K.D., Klein Wassink-Ruiter, J.S., van der Zwaag, P.A., du Marchie Sarvaas, G.J., Bergman, K.A.,  
45 Bilardo, C.M., Roos-Hesselink, J.W., Janssen, J.H.P., Frohn-Mulder, I.M., van Spaendonck-Zwarts,  
46 K.Y., van Melle, J.P., Hofstra, R.M.W., Wessels, M.W., 2016. Cardiovascular malformations caused  
47 by NOTCH1 mutations do not keep left: data on 428 probands with left-sided CHD and their  
48 families. *Genet. Med.* 18, 914–923. <https://doi.org/10.1038/gim.2015.193>
- 49 Kobayashi, J., Yoshida, M., Tarui, S., Hirata, M., Nagai, Y., Kasahara, S., Naruse, K., Ito, H., Sano, S., Oh, H.,  
50 2014. Directed Differentiation of Patient-Specific Induced Pluripotent Stem Cells Identifies the  
51 Transcriptional Repression and Epigenetic Modification of NKX2-5, HAND1, and NOTCH1 in

- 1 Hypoplastic Left Heart Syndrome. *PLoS ONE* 9, e102796.  
2 <https://doi.org/10.1371/journal.pone.0102796>
- 3 Kodama, M., Kodama, T., Newberg, J.Y., Katayama, H., Kobayashi, M., Hanash, S.M., Yoshihara, K., Wei, Z.,  
4 Tien, J.C., Rangel, R., Hashimoto, K., Mabuchi, S., Sawada, K., Kimura, T., Copeland, N.G., Jenkins,  
5 N.A., 2017. In vivo loss-of-function screens identify KPNB1 as a new druggable oncogene in  
6 epithelial ovarian cancer. *Proc. Natl. Acad. Sci.* 114, E7301–E7310.  
7 <https://doi.org/10.1073/pnas.1705441114>
- 8 Krane, M., Dreßen, M., Santamaria, G., My, I., Schneider, C.M., Dorn, T., Laue, S., Mastantuono, E., Berutti,  
9 R., Rawat, H., Gilsbach, R., Schneider, P., Lahm, H., Schwarz, S., Doppler, S.A., Paige, S., Puluca, N.,  
10 Doll, S., Neb, I., Brade, T., Zhang, Z., Abou-Ajram, C., Northoff, B., Holdt, L.M., Sudhop, S., Sahara,  
11 M., Goedel, A., Dendorfer, A., Tjong, F.V.Y., Rijlaarsdam, M.E., Cleuziou, J., Lang, N., Kupatt, C.,  
12 Bezzina, C., Lange, R., Bowles, N.E., Mann, M., Gelb, B.D., Crotti, L., Hein, L., Meitinger, T., Wu, S.,  
13 Sinnecker, D., Gruber, P.J., Laugwitz, K.-L., Moretti, A., 2021. Sequential Defects in Cardiac Lineage  
14 Commitment and Maturation Cause Hypoplastic Left Heart Syndrome. *Circulation* 144, 1409–1428.  
15 <https://doi.org/10.1161/CIRCULATIONAHA.121.056198>
- 16 Labombarda, F., Hamilton, R., Shohoudi, A., Aboulhosn, J., Broberg, C.S., Chaix, M.A., Cohen, S., Cook, S.,  
17 Dore, A., Fernandes, S.M., Fournier, A., Kay, J., Macle, L., Mondésert, B., Mongeon, F.-P.,  
18 Opotowsky, A.R., Proietti, A., Rivard, L., Ting, J., Thibault, B., Zaidi, A., Khairy, P., 2017. Increasing  
19 Prevalence of Atrial Fibrillation and Permanent Atrial Arrhythmias in Congenital Heart Disease. *J.*  
20 *Am. Coll. Cardiol.* 70, 857–865. <https://doi.org/10.1016/j.jacc.2017.06.034>
- 21 Lahm, H., Jia, M., Dreßen, M., Wirth, F., Puluca, N., Gilsbach, R., Keavney, B.D., Cleuziou, J., Beck, N.,  
22 Bondareva, O., Dzilic, E., Burri, M., König, K.C., Ziegelmeüller, J.A., Abou-Ajram, C., Neb, I., Zhang, Z.,  
23 Doppler, S.A., Mastantuono, E., Lichtner, P., Eckstein, G., Hörer, J., Ewert, P., Priest, J.R., Hein, L.,  
24 Lange, R., Meitinger, T., Cordell, H.J., Müller-Myhsok, B., Krane, M., 2021. Congenital heart disease  
25 risk loci identified by genome-wide association study in European patients. *J. Clin. Invest.* 131,  
26 e141837. <https://doi.org/10.1172/JCI141837>
- 27 Liu, Y., Chen, S., Zühlke, L., Black, G.C., Choy, M., Li, N., Keavney, B.D., 2019. Global birth prevalence of  
28 congenital heart defects 1970–2017: updated systematic review and meta-analysis of 260 studies.  
29 *Int. J. Epidemiol.* 48, 455–463. <https://doi.org/10.1093/ije/dyz009>
- 30 Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N.,  
31 Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron,  
32 J., Fleming, J., Siminoff, L., Traino, H., Mosavel, M., Barker, L., Jewell, S., Rohrer, D., Maxim, D.,  
33 Filkins, D., Harbach, P., Cortadillo, E., Berghuis, B., Turner, L., Hudson, E., Feenstra, K., Sobin, L.,  
34 Robb, J., Branton, P., Korzeniewski, G., Shive, C., Tabor, D., Qi, L., Groch, K., Nampally, S., Buia, S.,  
35 Zimmerman, A., Smith, A., Burges, R., Robinson, K., Valentino, K., Bradbury, D., Cosentino, M., Diaz-  
36 Mayoral, N., Kennedy, M., Engel, T., Williams, P., Erickson, K., Ardlie, K., Winckler, W., Getz, G.,  
37 DeLuca, D., MacArthur, D., Kellis, M., Thomson, A., Young, T., Gelfand, E., Donovan, M., Meng, Y.,  
38 Grant, G., Mash, D., Marcus, Y., Basile, M., Liu, J., Zhu, J., Tu, Z., Cox, N.J., Nicolae, D.L., Gamazon,  
39 E.R., Im, H.K., Konkashbaev, A., Pritchard, J., Stevens, M., Flutre, T., Wen, X., Dermitzakis, E.T.,  
40 Lappalainen, T., Guigo, R., Monlong, J., Sammeth, M., Koller, D., Battle, A., Mostafavi, S., McCarthy,  
41 M., Rivas, M., Maller, J., Rusyn, I., Nobel, A., Wright, F., Shabalina, A., Feolo, M., Sharopova, N.,  
42 Sturcke, A., Paschal, J., Anderson, J.M., Wilder, E.L., Derr, L.K., Green, E.D., Struwing, J.P., Temple,  
43 G., Volpi, S., Boyer, J.T., Thomson, E.J., Guyer, M.S., Ng, C., Abdallah, A., Colantuoni, D., Insel, T.R.,  
44 Koester, S.E., Little, A.R., Bender, P.K., Lehner, T., Yao, Y., Compton, C.C., Vaught, J.B., Sawyer, S.,  
45 Lockhart, N.C., Demchok, J., Moore, H.F., 2013. The Genotype-Tissue Expression (GTEx) project.  
46 *Nat. Genet.* 45, 580–585. <https://doi.org/10.1038/ng.2653>
- 47 Mandalenakis, Z., Rosengren, A., Lappas, G., Eriksson, P., Gilljam, T., Hansson, P.-O., Skoglund, K.,  
48 Fedchenko, M., Dellborg, M., 2018. Atrial Fibrillation Burden in Young Patients With Congenital  
49 Heart Disease. *Circulation* 137, 928–937. <https://doi.org/10.1161/CIRCULATIONAHA.117.029590>
- 50 Martinsson, A., Li, X., Andersson, C., Nilsson, J., Smith, J.G., Sundquist, K., 2015. Temporal Trends in the  
51 Incidence and Prognosis of Aortic Stenosis: A Nationwide Study of the Swedish Population.  
52 *Circulation* 131, 988–994. <https://doi.org/10.1161/CIRCULATIONAHA.114.012906>

- 1 Mbatchou, J., Barnard, L., Backman, J., Marcketta, A., Kosmicki, J.A., Ziyatdinov, A., Benner, C., O'Dushlaine,  
2 C., Barber, M., Boutkov, B., Habegger, L., Ferreira, M., Baras, A., Reid, J., Abecasis, G., Maxwell, E.,  
3 Marchini, J., 2021. Computationally efficient whole-genome regression for quantitative and binary  
4 traits. *Nat. Genet.* 53, 1097–1103. <https://doi.org/10.1038/s41588-021-00870-7>
- 5 Orr, Y., Leclair, K., Jacobe, S., Badawi, N., Nicholson, I.A., Chard, R.B., Sholler, G.F., Winlaw, D.S., 2015. Early  
6 outcomes from a new regional programme for the surgical management of hypoplastic left heart  
7 syndrome: Outcomes after surgery for HLHS. *ANZ J. Surg.* 85, 466–471.  
8 <https://doi.org/10.1111/ans.12103>
- 9 Øyen, N., Diaz, L.J., Leirgul, E., Boyd, H.A., Priest, J., Mathiesen, E.R., Quertermous, T., Wohlfahrt, J.,  
10 Melbye, M., 2016. Prepregnancy Diabetes and Offspring Risk of Congenital Heart Disease: A  
11 Nationwide Cohort Study. *Circulation* 133, 2243–2253.  
12 <https://doi.org/10.1161/CIRCULATIONAHA.115.017465>
- 13 Patorno, E., Huybrechts, K.F., Bateman, B.T., Cohen, J.M., Desai, R.J., Mogun, H., Cohen, L.S., Hernandez-  
14 Diaz, S., 2017. Lithium Use in Pregnancy and the Risk of Cardiac Malformations. *N. Engl. J. Med.*  
15 376, 2245–2254. <https://doi.org/10.1056/NEJMoa1612222>
- 16 Peng, G., Barro-Soria, R., Sampson, K.J., Larsson, H.P., Kass, R.S., 2017. Gating mechanisms underlying  
17 deactivation slowing by two KCNQ1 atrial fibrillation mutations. *Sci. Rep.* 7, 45911.  
18 <https://doi.org/10.1038/srep45911>
- 19 Persson, M., Razaz, N., Edstedt Bonamy, A.-K., Villamor, E., Cnattingius, S., 2019. Maternal Overweight and  
20 Obesity and Risk of Congenital Heart Defects. *J. Am. Coll. Cardiol.* 73, 44–53.  
21 <https://doi.org/10.1016/j.jacc.2018.10.050>
- 22 Pierpont, M.E., Brueckner, M., Chung, W.K., Garg, V., Lacro, R.V., McGuire, A.L., Mital, S., Priest, J.R., Pu,  
23 W.T., Roberts, A., Ware, S.M., Gelb, B.D., Russell, M.W., On behalf of the American Heart  
24 Association Council on Cardiovascular Disease in the Young; Council on Cardiovascular and Stroke  
25 Nursing; and Council on Genomic and Precision Medicine, 2018. Genetic Basis for Congenital Heart  
26 Disease: Revisited: A Scientific Statement From the American Heart Association. *Circulation* 138.  
27 <https://doi.org/10.1161/CIR.0000000000000606>
- 28 Reuter, M.S., Jobling, R., Chaturvedi, R.R., Manshaei, R., Costain, G., Heung, T., Curtis, M., Hosseini, S.M.,  
29 Liston, E., Lowther, C., Oechslin, E., Sticht, H., Thiruvahindrapuram, B., Mil, S. van, Wald, R.M.,  
30 Walker, S., Marshall, C.R., Silversides, C.K., Scherer, S.W., Kim, R.H., Bassett, A.S., 2019.  
31 Haploinsufficiency of vascular endothelial growth factor related signaling genes is associated with  
32 tetralogy of Fallot. *Genet. Med.* 21, 1001–1007. <https://doi.org/10.1038/s41436-018-0260-9>
- 33 Ritari, J., Hyvärinen, K., Clancy, J., FinnGen, Partanen, J., Koskela, S., 2020. Increasing accuracy of HLA  
34 imputation by a population-specific reference panel in a FinnGen biobank cohort. *NAR Genomics*  
35 *Bioinforma.* 2, lqaa030. <https://doi.org/10.1093/nargab/lqaa030>
- 36 Roselli, C., Rienstra, M., Ellinor, P.T., 2020. Genetics of Atrial Fibrillation in 2020: GWAS, Genome  
37 Sequencing, Polygenic Risk, and Beyond. *Circ. Res.* 127, 21–33.  
38 <https://doi.org/10.1161/CIRCRESAHA.120.316575>
- 39 Sabour, D., Machado, R.S.R., Pinto, J.P., Rohani, S., Sahito, R.G.A., Hescheler, J., Futschik, M.E., Sachinidis,  
40 A., 2018. Parallel Genome-wide Profiling of Coding and Non-coding RNAs to Identify Novel  
41 Regulatory Elements in Embryonic and Matured Heart. *Mol. Ther. - Nucleic Acids* 12, 158–173.  
42 <https://doi.org/10.1016/j.omtn.2018.04.018>
- 43 Saha, P., Potiny, P., Rigdon, J., Morello, M., Tcheandjieu, C., Romfh, A., Fernandes, S.M., McElhinney, D.B.,  
44 Bernstein, D., Lui, G.K., Shaw, G.M., Ingelsson, E., Priest, J.R., 2019. Substantial Cardiovascular  
45 Morbidity in Adults With Lower-Complexity Congenital Heart Disease. *Circulation* 139, 1889–1899.  
46 <https://doi.org/10.1161/CIRCULATIONAHA.118.037064>
- 47 Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., Regev, A., 2015. Spatial reconstruction of single-cell gene  
48 expression data. *Nat. Biotechnol.* 33, 495–502. <https://doi.org/10.1038/nbt.3192>
- 49 Speir, M.L., Bhaduri, A., Markov, N.S., Moreno, P., Nowakowski, T.J., Papatheodorou, I., Pollen, A.A., Raney,  
50 B.J., Seninge, L., Kent, W.J., Haeussler, M., 2021. UCSC Cell Browser: visualize your single-cell data.  
51 *Bioinformatics* btab503. <https://doi.org/10.1093/bioinformatics/btab503>

- 1 Stanfill, A.G., Cao, X., 2021. Enhancing Research Through the Use of the Genotype-Tissue Expression (GTEx)  
2 Database. *Biol. Res. Nurs.* 23, 533–540. <https://doi.org/10.1177/1099800421994186>
- 3 Thanassoulis, G., Campbell, C.Y., Owens, D.S., Smith, J.G., Smith, A.V., Peloso, G.M., Kerr, K.F., Pechlivanis,  
4 S., Budoff, M.J., Harris, T.B., Malhotra, R., O'Brien, K.D., Kamstrup, P.R., Nordestgaard, B.G.,  
5 Tybjaerg-Hansen, A., Allison, M.A., Aspelund, T., Criqui, M.H., Heckbert, S.R., Hwang, S.-J., Liu, Y.,  
6 Sjogren, M., van der Pals, J., Kälsch, H., Mühleisen, T.W., Nöthen, M.M., Cupples, L.A., Caslake, M.,  
7 Di Angelantonio, E., Danesh, J., Rotter, J.I., Sigurdsson, S., Wong, Q., Erbel, R., Kathiresan, S.,  
8 Melander, O., Gudnason, V., O'Donnell, C.J., Post, W.S., 2013. Genetic Associations with Valvular  
9 Calcification and Aortic Stenosis. *N. Engl. J. Med.* 368, 503–512.  
10 <https://doi.org/10.1056/NEJMoa1109034>
- 11 Theis, J.L., Hrstka, S.C.L., Evans, J.M., O'Byrne, M.M., de Andrade, M., O'Leary, P.W., Nelson, T.J., Olson,  
12 T.M., 2015. Compound heterozygous NOTCH1 mutations underlie impaired cardiogenesis in a  
13 patient with hypoplastic left heart syndrome. *Hum. Genet.* 134, 1003–1011.  
14 <https://doi.org/10.1007/s00439-015-1582-1>
- 15 Thériault, S., Gaudreault, N., Lamontagne, M., Rosa, M., Boulanger, M.-C., Messika-Zeitoun, D., Clavel, M.-  
16 A., Capoulade, R., Dagenais, F., Pibarot, P., Mathieu, P., Bossé, Y., 2018. A transcriptome-wide  
17 association study identifies PALMD as a susceptibility gene for calcific aortic valve stenosis. *Nat.*  
18 *Commun.* 9, 988. <https://doi.org/10.1038/s41467-018-03260-6>
- 19 Trokovic, R., Weltner, J., Otonkoski, T., 2015. Generation of iPSC line HEL24.3 from human neonatal  
20 foreskin fibroblasts. *Stem Cell Res.* 15, 266–268. <https://doi.org/10.1016/j.scr.2015.05.012>
- 21 Vanlerberghe, C., Jourdain, A.-S., Ghoumid, J., Frenois, F., Mezel, A., Vaksman, G., Lenne, B., Delobel, B.,  
22 Porchet, N., Cormier-Daire, V., Smol, T., Escande, F., Manouvrier-Hanu, S., Petit, F., 2019. Holt-  
23 Oram syndrome: clinical and molecular description of 78 patients with TBX5 variants. *Eur. J. Hum.*  
24 *Genet.* 27, 360–368. <https://doi.org/10.1038/s41431-018-0303-3>
- 25 Wang, G., Sarkar, A., Carbonetto, P., Stephens, M., 2020. A simple new approach to variable selection in  
26 regression, with application to genetic fine mapping. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 82, 1273–  
27 1300. <https://doi.org/10.1111/rssb.12388>
- 28 Wang, T.Y., Arking, D.E., Maleszewski, J.J., Fox-Talbot, K., Nieuwenhuis, T.O., Santhanam, L., Virmani, R.,  
29 Rosenberg, A.Z., Halushka, M.K., 2019. Human cardiac myosin light chain 4 (MYL4) mosaic  
30 expression patterns vary by sex. *Sci. Rep.* 9, 12681. <https://doi.org/10.1038/s41598-019-49191-0>
- 31 Wang, T.Y., Lee, D., Fox-Talbot, K., Arking, D.E., Chakravarti, A., Halushka, M.K., 2018. Cardiomyocytes have  
32 mosaic patterns of protein expression. *Cardiovasc. Pathol.* 34, 50–57.  
33 <https://doi.org/10.1016/j.carpath.2018.03.002>
- 34 Watanabe, K., Taskesen, E., van Bochoven, A., Posthuma, D., 2017. Functional mapping and annotation of  
35 genetic associations with FUMA. *Nat. Commun.* 8, 1826. <https://doi.org/10.1038/s41467-017-01261-5>
- 36
- 37 Yang, C., Xu, Y., Yu, M., Lee, D., Alharti, S., Hellen, N., Ahmad Shaik, N., Banaganapalli, B., Sheikh Ali  
38 Mohamoud, H., Elango, R., Przyborski, S., Tenin, G., Williams, S., O'Sullivan, J., Al-Radi, O.O., Atta, J.,  
39 Harding, S.E., Keavney, B., Lako, M., Armstrong, L., 2017. Induced pluripotent stem cell modelling of  
40 HLHS underlines the contribution of dysfunctional NOTCH signalling to impaired cardiogenesis.  
41 *Hum. Mol. Genet.* 26, 3031–3045. <https://doi.org/10.1093/hmg/ddx140>
- 42 Zhou, W., Zhao, Z., Nielsen, J.B., Fritsche, L.G., LeFaive, J., Gagliano Taliun, S.A., Bi, W., Gabrielsen, M.E.,  
43 Daly, M.J., Neale, B.M., Hveem, K., Abecasis, G.R., Willer, C.J., Lee, S., 2020. Scalable generalized  
44 linear mixed model for region-based association tests in large biobanks and cohorts. *Nat. Genet.*  
45 52, 634–639. <https://doi.org/10.1038/s41588-020-0621-6>
- 46 Zhu, Z.-C., Liu, J.-W., Li, K., Zheng, J., Xiong, Z.-Q., 2018. KPNB1 inhibition disrupts proteostasis and triggers  
47 unfolded protein response-mediated apoptosis in glioblastoma cells. *Oncogene* 37, 2936–2952.  
48 <https://doi.org/10.1038/s41388-018-0180-9>
- 49

1 Table 1. GWAS significant lead SNPs found in FinnGen release 7 for several CHD groups.

Subgroup	Variant	Location (hg38)	Finnish Enrichment	AF cases	AF controls	beta	SE	P-value
CHD general	rs11570508	17:47151194	1.190	0.29	0.24	0.24	0.035	$1.2 \times 10^{-11}$
LVOTO broad	rs1342740627	20:23988232	1000000.000	0.0030	0.00054	4.79	0.85	$1.78 \times 10^{-8}$
Early aortic valve disease	rs17214148	6:33004385	0.0770	0.0051	0.0013	3.23	0.59	$3.38 \times 10^{-8}$
Early aortic valve disease	rs1342740627	20:23988232	1000000.000	0.0037	0.00074	4.92	0.89	$3.44 \times 10^{-8}$
Septal defects	rs1293973611	17:81382139	1000000.000	0.0061	0.00087	4.96	0.87	$1.1 \times 10^{-8}$

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

1 **Supplemental Files**

2

3 **Supplemental Document 1. Expanded FinnGen contributor list.**

4

5 **Supplemental Table 1. ICD-10 codes used as end-points for the main CHD categories in this**  
6 **study.**

7

8 **Supplemental Table 2. List of covariates used in the SAIGE GWAS analysis.**

9

10 **Supplemental Table 3. FinnGen SuSiE and FINEMAP results for the lead SNPs.**

11

12 **Supplemental Table 4. Most significant HLA alleles from the generalized linear model**  
13 **regression analysis against Early aortic valve disease.**

14

15 **Supplemental Figure 1. Locus zoom plots (v0.13.3) of the lead SNPs for LVOTO broad, aortic**  
16 **valve disease early age (<60 years) and septal defects.**

17

18 **Supplemental Figure 2. QQ plots from the GWAS analyses of the different CHD groups.**

19

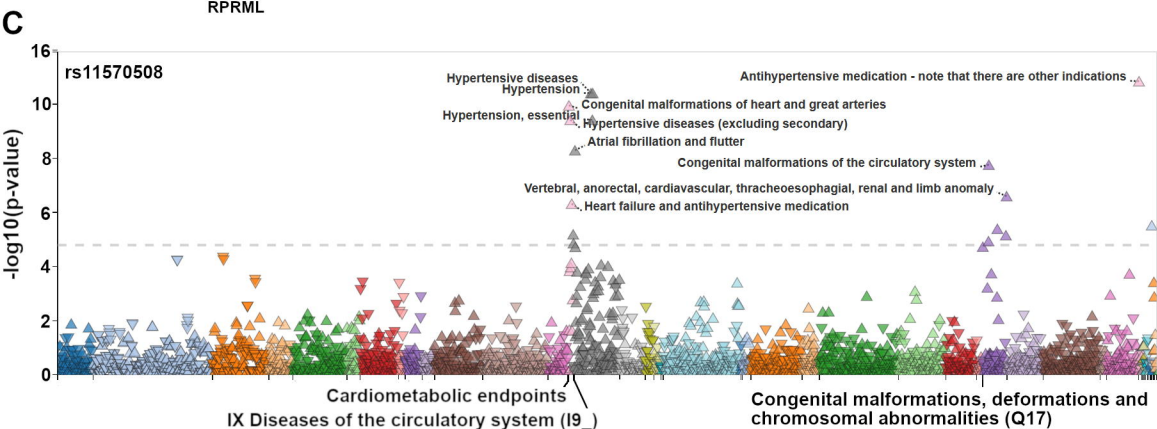
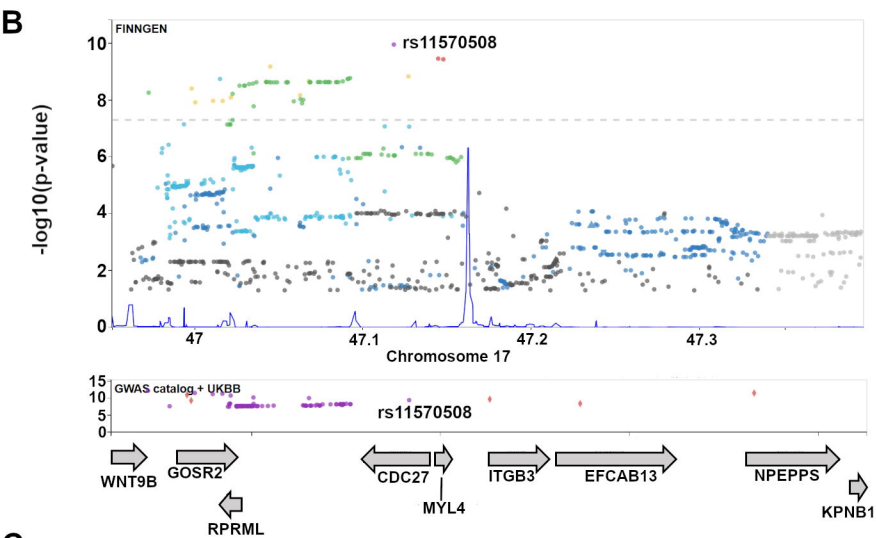
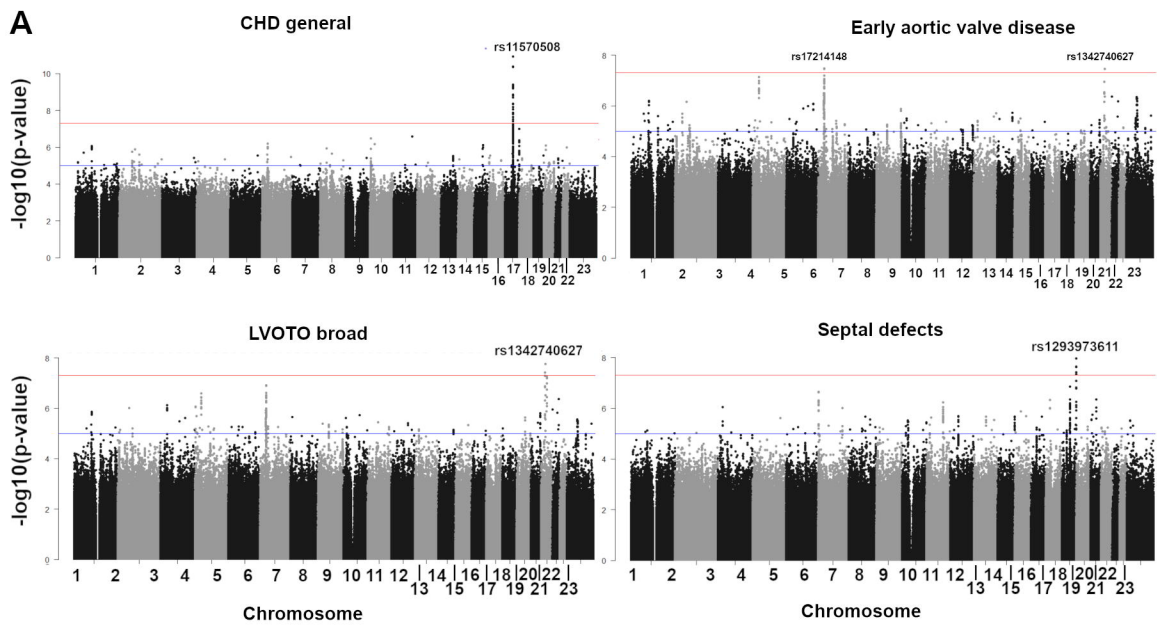
20 **Supplemental Figure 3. Circos plot displaying a putative chromatin interaction (orange link)**  
21 **between KPNB1 and rs11570508**



1

2 **Supplemental Figure 4. FinnGen PheWeb plot of rs17214148 with most significantly associated**  
3 **end points highlighted.**

4

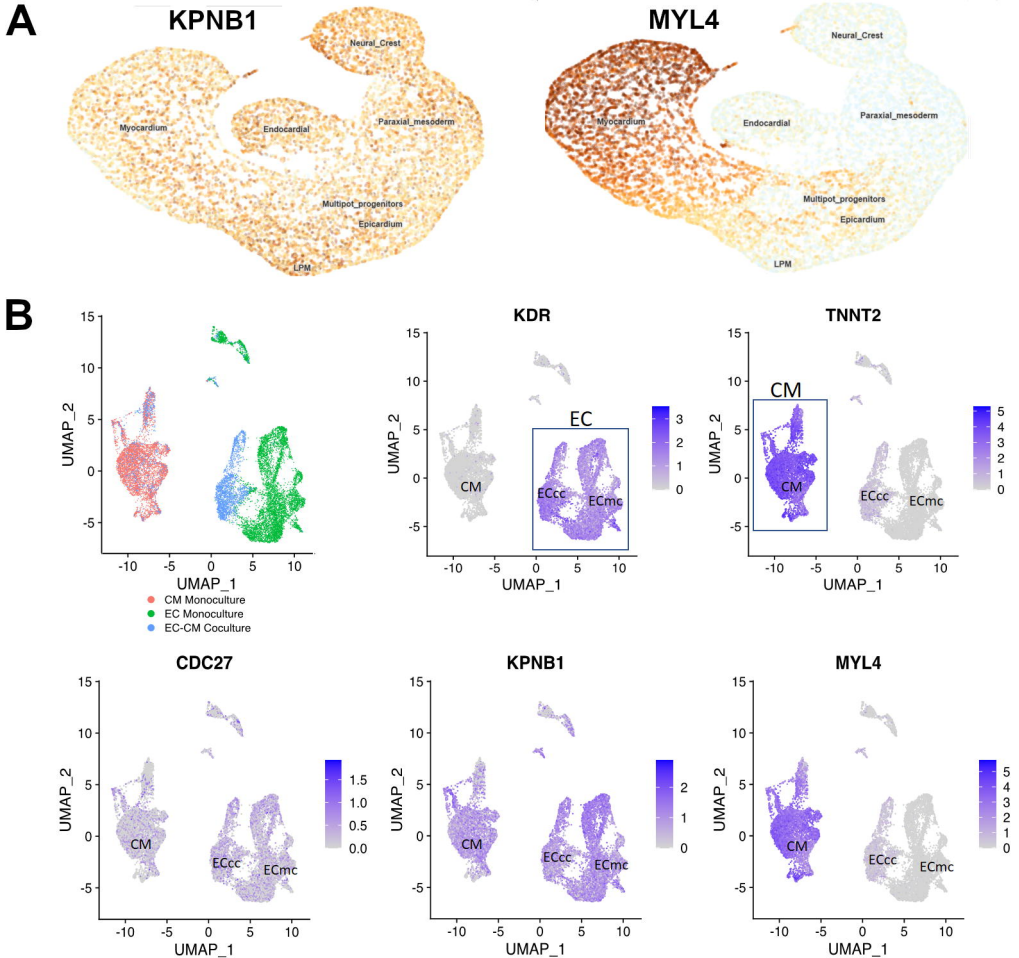


**Figure 1. FinnGen R7 GWAS for general congenital heart disease (congenital malformations of the heart and great arteries) highlights rs11570508 as significantly associated.**

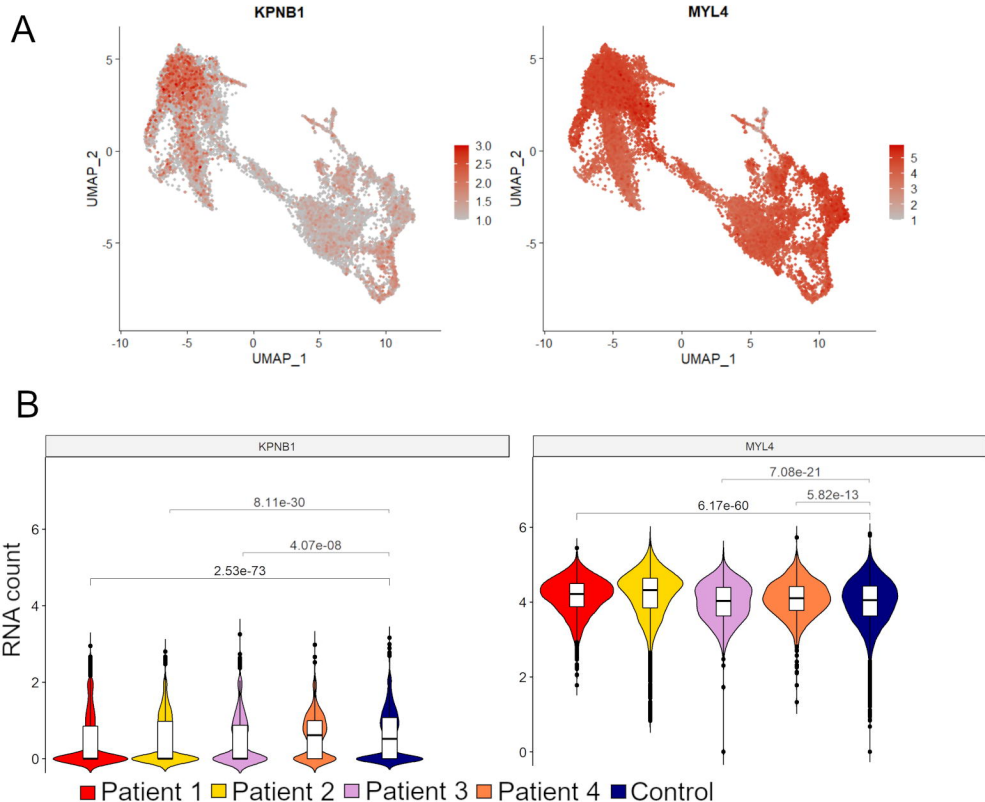
A: Genome-wide manhattan plots of CHD general, early aortic valve disease, LVOTO broad and septal defects with lead SNPs highlighted.

B: Locus zoom plot with rs11570508 and nearby SNPs and genes.

C: FinnGen PheWeb plot demonstrating rs11570508 associations with heart-based endpoints.



**Figure 2. MYL4 and KPNB1 are expressed differently in cardiac cell types in mice and humans.**  
 A: early embryonic mouse cardiogenic map of KPNB1 and MYL4 cell expression based on the study by de Soysa et al. (2019), as visualized on the UCSC cell browser website.  
 B: scRNAseq gene expression UMAP plots from cardiomyocyte-endothelial cell cultures (Helle et al. 2019) for KDR, TNNT2, MYL4, KPNB1 and CDC27.



**Figure 3. KPNB1 and MYL4 are significantly down-regulated LVOTO patients' hiPSC-CMs compared to healthy controls.**

A; UMAP plot of the KPNB1 and MYL4 expression across all cells.

B; violin plot and box plots displaying the scaled RNA counts of KPNB1 and MYL4 in all 4 patient samples and the control pool sample.