


RESEARCH

Open Access



ENU-based dominant genetic screen identifies contractile and neuronal gene mutations in congenital heart disease

Xiaoxi Luo^{1,2,3,4,5†}, Lifeng Liu^{2,3,4,5†}, Haowei Rong^{2,3,4,5}, Xiangyang Liu^{2,3,4,5}, Ling Yang⁶, Nan Li⁶ and Hongjun Shi^{2,3,4,5*} 

Abstract

Background Congenital heart disease (CHD) is the most prevalent congenital anomaly, but its underlying causes are still not fully understood. It is believed that multiple rare genetic mutations may contribute to the development of CHD.

Methods In this study, we aimed to identify novel genetic risk factors for CHD using an ENU-based dominant genetic screen in mice. We analyzed fetuses with malformed hearts and compared them to control littermates by whole exome or whole genome sequencing (WES/WGS). The differences in mutation rates between observed and expected values were tested using the Poisson and Binomial distribution. Additionally, we compared WES data from human CHD probands obtained from the Pediatric Cardiac Genomics Consortium with control subjects from the 1000 Genomes Project using Fisher's exact test to evaluate the burden of rare inherited damaging mutations in patients.

Results By screening 10,285 fetuses, we identified 1109 cases with various heart defects, with ventricular septal defects and bicuspid aortic valves being the most common types. WES/WGS analysis of 598 cases and 532 control littermates revealed a higher number of ENU-induced damaging mutations in cases compared to controls. GO term and KEGG pathway enrichment analysis showed that pathways related to cardiac contraction and neuronal development and functions were enriched in cases. Further analysis of 1457 human CHD probands and 2675 control subjects also revealed an enrichment of genes associated with muscle and nervous system development in patients. By combining the mice and human data, we identified a list of 101 candidate digenic genesets, from which each geneset was co-mutated in at least one mouse and two human probands with CHD but not in control mouse and control human subjects.

Conclusions Our findings suggest that gene mutations affecting early hemodynamic perturbations in the developing heart may play a significant role as a genetic risk factor for CHD. Further validation of the candidate gene set identified in this study could enhance our understanding of the complex genetics underlying CHD and potentially lead to the development of new diagnostic and therapeutic approaches.

Keywords Congenital heart disease, ENU, Genetic screen, Cardiac contraction, Genetic risk

[†]Xiaoxi Luo and Lifeng Liu contributed equally to this article.

*Correspondence:

Hongjun Shi
shihongjun@westlake.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Congenital heart disease (CHD) is the most common form of birth defect and affects approximately 1% of live-born children [1]. It is also among the top five causes of death in children younger than 1 year [2]. Despite improved prenatal care and increased awareness of risk factors, the global incidence of CHD has steadily increased over the last five decades [3]. With advancements in medical and surgical interventions, most CHD patients can now survive into adulthood, resulting in a doubling of the prevalence rate in the aged population between 1990 and 2017 [4]. The inability to control the incidence of CHD highlights the importance of etiological studies.

Currently, about 20–30% of CHD cases can be genetically diagnosed with known CHD-causing genetic changes. These include 8–10% gross chromosomal anomalies/aneuploidy, 3–25% copy number variations, and 3–5% single gene variants [5]. Positive genetic diagnoses are more likely to be achieved in syndromic cases at either the chromosomal [6] or gene levels [7], although syndromic cases account for only 25% of total CHD cases [8]. Large genetic CHD cohort studies with whole exome sequencing (WES) from the Pediatric Cardiac Genomics Consortium (PCGC) suggest that overall, 8% of cases can be attributed to de novo autosomal dominant mutations, with syndromic cases more likely to be explained (up to 28% of syndromic vs 3% of isolated CHD cases explainable); inherited rare variants are implicated in only 1.8% of cases [9, 10]. Thus far, the etiology of the vast majority of CHD cases, especially sporadic isolated cases, is still poorly understood.

Although classic Mendelian inheritance patterns have been identified in some familial clusters, the overall 2–6% sibling or offspring recurrence risk of isolated CHD suggests that the majority of CHD cases are multifactorial in origin, ranging from multiple genetic alterations to gene-environmental interactions [5]. Gifford et al. provided a compelling example of the oligogenic origin of congenital heart disease (CHD) by discovering that the combined inheritance of heterozygous mutations in three genes (MKL2, MYH7, and NKX2-5) resulted in left ventricular noncompaction (LVNC) in both humans and mice [11]. Priest et al. adopted a more systemic approach in their study of atrioventricular septal defects (AVSD). They analyzed the total rare variants with classical inheritance patterns (de novo, homozygous, compound heterozygous) from 59 AVSD trios using protein interaction network analysis. Their findings identified protein interaction networks, particularly a pair of interacting collagen genes (COL2A1, COL9A1), enriched in the AVSD trios, providing support for oligogenic inheritance [12]. Although numerous variant-level genome-wide

association studies have been conducted to identify genetic risk loci for CHD, there have been no reports on gene-level genome-wide assessments of rare inherited heterozygous mutations contributing to CHD.

ENU (N-ethyl-N-nitrosourea) is an alkylating agent commonly used to induce mutations in mice to mimic spontaneous mutations in humans for more than 30 years [13]. A large ENU-based recessive forward genetic screening study of CHD in mice identified double recessive mutations in *Sap130* and *Pcdha9*, previously not associated with CHD, as a novel digenic origin of HLHS [14]. The same genetic screen also revealed 91 recessive CHD mutations in 61 genes, half of which were related to cilia formation, with laterality defect syndromes being the most common manifestations [15]. This finding is consistent with that of a human CHD WES study in which among all major CHD subgroups, only laterality defects were significantly enriched for damaging recessive genes, particularly cilia-related genes [10]. These findings suggest that recessive damaging mutations may underlie a distinct minority class of CHD, largely in the syndromic form of laterality defects.

Since almost all known genes related to isolated CHD and more than half of the known genes related to syndromic CHD exhibit an autosomal dominant pattern [7], a reduced dosage or gene expression level, rather than complete loss of gene products is more likely the mechanism underlying most CHD cases. Individually, these mutations may not cause CHD, but their presence in combination can lead to the development of CHD. To identify such disease-contributing heterozygous mutations, we employed a large-scale ENU-based forward dominant screen in mice to explore potential novel genetic risk factors of CHD. This study identified a large number of mutations in genes regulating early embryonic heart contractility as a novel mechanism contributing to CHD.

Methods

Mouse breeding and ENU mutagenesis

All animal experiments were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee of Westlake University (approval 21-005-SHJ). Mice were maintained in a 12-h light/dark cycle and provided ad libitum access to food and water. ENU mutagenesis was performed as previously described [16]. Briefly, sexually mature (8 weeks) C57BL6/J males were administered an intraperitoneal injection with 90 mg/kg ENU (N-ethyl-N-nitrosourea, Sigma-Aldrich) once a week for two consecutive weeks. In total, 200 G0 mice were injected, and after a recovery period of 15 weeks to regain fertility, 88 mutagenized G0 males were mated with wild-type C57BL6/J female mice. At E18.5,

pregnant females were sacrificed by cervical dislocation. Their G1 fetuses were subsequently removed via cesarean section. Immediately thereafter, the fetuses were euthanized by decapitation, and their hearts were excised for imaging analysis. ENU-treated males exhibiting signs of morbidity and loss of reproductive function were euthanized with carbon dioxide [17].

Lightsheet fluorescence microscopy (LFM) and cardiac phenotyping

Fetal hearts were imaged using the Zeiss Lightsheet Z.1 microscope. Embryos were harvested at E18.5, and hearts were dissected in phosphate-buffered saline (PBS), fixed overnight in a mixed solution of 10% neutral buffered formalin and 2.5% glutaraldehyde, then rinsed twice in PBS, and dehydrated in 50%, 75%, and 100% ethanol for 30 min each at room temperature (RT). Samples were transferred into a specially designed glass tube containing 100 μ L of BABB solution (1:2 benzyl alcohol: benzyl benzoate) for 30 min to clear the sample. The glass tube was mounted into the sample chamber filled with 85% glycerol (RI~1.45). Whole Hearts were scanned for tissue autofluorescence using a 561-nm laser line with detection optics 5 \times /0.16. 3D reconstruction of the image stacks and morphological analyses were performed with Imaris 9.3 software. Heart morphology was assessed in 3D mode independently by two trained personnel.

Mouse WES, WGS, and data processing

One hundred seventy CHD G1 fetuses and 52 normal G1 fetuses were whole exome sequenced at Novogene. Genomic DNA was captured using Agilent SureSelect Mouse All Exon V1 and sequenced using the Illumina Novaseq 6000 platform, with a minimum average of 100 \times target sequence coverage. Additionally, 550 CHD G1 and 559 normal G1 fetuses were sequenced at BGI using the MGI DNBSEQ-T series platform, with a minimum average of 60 \times whole genome sequence coverage. The reads were aligned to the C57BL/6J mouse reference genome (mm10) using BWA v0.7.17 [18]. Duplicate reads were removed with sambalster v0.1.26 [19], and the data were sorted with sambamba v0.8.0 [20]. For the whole genome sequencing samples, the exon region BAM file was extracted using the sambamba application slice parameter. Local realignments were generated, and base quality scores were recalibrated using GATK 4.2.0.0 [21] following the GATK Best Practices. All samples were called together with platypus/0.8.1 [22] to obtain a single vcf file, and variants were annotated with ANNOVAR [23] and the SIFT-4G [24] database. To further control the sequence quality, variants with genotype score < 80 or allele balance < 0.2 (for heterozygous genotypes only) were filtered out.

Kinship analysis and variants filtering

Before conducting the kinship analysis, variants that occurred four times or more in all 1,331 mouse fetuses were removed to eliminate potential interference from background noise. This step aimed to focus on the variants most likely originating from ENU mutagenesis. The remaining variants were then used to quantify the inbreeding coefficient (IC) between pairs of samples in the GCTA v1.949 software [25]. An IC cutoff of 0.05 was applied to determine the relatedness between pairs of samples. Core families were identified, consisting of pairs of samples with an IC greater than 0.05. These core families were further condensed by merging families with any overlapping samples, resulting in the identification of 152 large families and 983 singletons out of a total of 1331 samples. To ensure diversity within the analysis, only one randomly selected sample from each large family was included. This process established a control group consisting of 532 normal fetuses and a case group consisting of 603 fetuses with malformed hearts. To focus specifically on ENU-induced variants, only variants that occurred once in the total 1135 samples were retained for the final analysis. This additional filtering step resulted in the removal of 5 samples. Consequently, the control group consisted of 532 samples, and the case group consisted of 598 samples. The variant filtering process is depicted in Additional file 1: Fig. S1.

Human WES data processing

For human CHD samples, raw sequencing files (.sra) were converted into FASTQ files, and the quality of the sequencing reads was assessed using fastp [26]. The reads were aligned to the human genome reference sequence (CRGh38/hg38) using Burrows-Wheeler Aligner (BWA) v0.7.17 [18], in particular the BWA-MEM algorithm. Duplicate reads were marked and removed after alignment using sambalster v0.1.26 [18], and BAM files were sorted using sambamba v0.8.0 [20]. The control WES files were downloaded from the 1000 Genomes Project [27] in CRAM format and converted to BAM format using samtools 1.14 [28]. After all BAM files were split by chromosomes using sambamba application according to The Exome-Agilent-V6.bed file, the GATK Best Practices workflows (4.2.0.0) were used to apply indel realignment and base quality recalibration [21, 29]. Single nucleotide variants and small indels were called with GATK HaplotypeCaller using “-ERC GVCF” parameter. Further processed using GenotypeGVCFs parameter to merge each chromosome of all samples. All mutations were annotated using ANNOVAR [23], dbSNP (v150), 1000 Genomes (August 2015), dbNSFP (41a), gnomAD (v3), and AlphaMissense [30].

Statistical analysis

All statistical analyses were conducted using Python 3.8.8.

Global variant burden test based on an expected mutational model

The expected mutation model was generated based on a hypothetical mutation model where each base of the whole exome was subject to an equal chance of mutation, adjusted by the ENU mutational bias. Briefly, GTF files for the main gene transcripts of the GRCm38/mm10 genome were obtained from the UCSC genome browser. In cases of existence of multiple transcript isoforms, only the longest transcripts were taken into account. Genes of olfactory receptor family, vomeronasal receptor family, KRTAP family, taste receptors *Tas1r* and *Tas2r* families, and *Ttn* gene were excluded from the analysis due to their hyperpolymorphic nature. We then created a mutation simulation dataset containing each of the three possible single nucleotide substitutions for each base in the exome (Additional file 2: Table S3). The occurrence of each simulated nucleotide change was adjusted by multiplying an ENU metagenesis bias factor as determined for each possible substitution (Additional file 1: Fig. S2b). The mutation simulation dataset was then annotated by ANNOVAR for variant classification (synonymous, missense, and LOF including nonsense, splicing, frameshift, start loss, and stop loss) and pathogenicity prediction based on the SIFT-4G score. The damaging missense (D-Mis) variants were determined as SIFT-4G score < 0.05. The total number of variants for each variant class (synonymous, LOF, D-Mis) was summed. The expected number of frameshift indels was estimated by multiplying the total number of simulated variants by the observed proportion of frameshift indels out of the total number of observed variants in the final control and case group included in the analysis which was determined to be 0.2%. The simulated number of frameshift indels was then added to the LOF class. In-frame indels were not considered in this analysis. The following formula was used to estimate the expected probability of each variant class per mouse (Pvc):

$$\text{Expected Pvc} = \bar{N}_{obs} \times \frac{Nvc}{Nt}$$

where \bar{N}_{obs} is the average number of variants observed in each mouse (which is determined to be 59 in this study); Nvc is the total number of variants in each variant class of the simulation dataset; Nt is the number of total variants of the whole simulation dataset.

Poisson statistics was used to test for an excess of mutations over expectation, from the expected probability of

each variant class Pvc, the total number of mice in control or case group, and the observed number of the variant class within the group.

Global variant burden test based on the control variant distribution

The total number of variants in each variant class observed in controls was summed (Nvc). The expected probability of each variant class per mouse (Pvc) was derived by dividing the Nvc by the number of control mice. Burden test was conducted using Poisson statistics as described above to test for an excess of mutation in cases over controls.

Gene burden test based on the expected mutational model

From the mutation simulation dataset, the total number of variants for each variant class was summed for each gene (Additional file 2: Table S3). The following formula was used to estimate the expected probability of each variant class of each gene per mouse (Pvcg):

$$\text{Expected Pvcg} = \bar{N}_{obs} \times \frac{Nvcg}{Nt}$$

where \bar{N}_{obs} is the average number of variants observed in each mouse (which is determined to be 59 in this study); Nvcg is the total number of variants of each gene in each variant class of the simulation dataset; Nt is the number of total variants of the whole simulation dataset.

To increase statistical power and minimize the risk of false positive discoveries, we removed genes from the analysis if the total number of mutation events, considering both controls and cases combined, was equal to or less than 5. A one-sided binomial test was then used to test for an excess of mutations over expectation for each gene, from expected probability of each variant class of each gene per mouse (Pvcg), the total number of mice in the control or case group, and the observed number of each variant classes of each gene within the group. After the generation of *p*-value for all genes, Storey's *q*-value procedure [31] was used to control the false discovery rate (FDR) under 0.05.

The expected number of mutations for each gene is defined as Pvcg multiplied by the total number of mice in the control or case group. Enrichment score is defined as the ratio of the observed number of mutations and the expected number of mutations.

Gene burden test based on human case-control comparison

After co-calling of variants from all control and case exomes, a filter was applied to obtain variants that were both rare (MAF < 0.001) and damaging (annotated as

LOF, or predicted to be pathogenic by AlphaMissense [30]). For each rare damaging variant, the total number of samples with identifiable genotype N_s (genotype score ≥ 20 and total reads ≥ 15 ; if the genotype is called as heterozygous, the variant reads ≥ 5) and the total number of samples with rare damaging variants from these identifiable genotypes N_v were determined within each sample group. All rare damaging variants were summed to the gene level within each sample group to give rise to the total number of samples with rare damaging variant per gene (N_{vg}). The median of the N_s of each gene (N_{ms}) represented the number of samples that were examined for the presence of N_{vg} . Based on N_{vg} and N_{ms} , one-sided Fisher's exact statistics was then used to test for an excess of mutation for each gene between the case and control group.

Gene sets used for mouse variant enrichment analysis

The known CHD genes set was adapted from the Knowledgebase for Congenital Heart Disease-related Genes and Clinical Manifestations (<http://chddb.fwgenetics.org/>) [32], which contains 1124 genes manually curated from multi-cohort analyses for CHD, among which 1044 mouse ortholog could be mapped. SysCilia genes [33], cilia genes [34], chromatin-modifying genes [34], high heart expression genes (HHE) [9, 35], and low heart expression genes (LHE) [9, 35] lists were adapted from previous reports. HHE were the top 25% of genes expressed in E14.5 mouse hearts and LHE were the bottom 25% of genes expressed in E14.5 mouse hearts.

Gene set enrichment analysis

Statistically significant gene sets were input into Metascape (<https://metascape.org/>) [36] to obtain enriched GO biological process and KEGG pathway with q -values calculated using the Benjamini-Hochberg procedure [37]. For gene set enrichment against MGI mammalian phenotype database, two files (All Genotypes and Mammalian Phenotype Annotations and Mammalian Phenotype Vocabulary in OBO v1.2) were obtained from Mouse Genome Informatics (<https://www.informatics.jax.org/>) [38] to compile a genotype-phenotype association file. Hypergeometric distribution test was then used to perform the term enrichment analysis.

Permutation test and principal component analysis

We randomly selected 148 genes from the pool of ENU-induced mutated genes that have served as the basis for our case-enriched geneset. After conducting 10,000 permutations, we subjected these random gene sets to enrichment analysis against the MGI mammalian phenotype database, yielding p -values for each phenotype term. Using the p -values < 0.01 as a cutoff, we assigned

the value of 1 to the significant term and 0 to the insignificant term for each geneset, and then performed principal component analysis (PCA) across 10,001 datasets (including our case-enriched gene set and 10,000 permutations) based on these assigned codes.

Results

ENU mutagenesis resulted in heart defects in G1

The process involved mating ENU-treated G0 males with wild-type females, resulting in the generation of G1 fetuses harboring multiple heterozygous de novo mutations derived from the mutagenized spermatogonial stem cells of the G0 males. At embryonic day 18.5 (E18.5), a total of 10,285 G1 fetal hearts were harvested and phenotyped using lightsheet fluorescence microscopy (LFM). The LFM technology enabled us to perform rapid scanning of the entire fetal heart in just 20 s per heart, achieving a three-dimensional resolution of $2.29 \times 2.29 \times 7.16 \mu\text{m}$. Leveraging this high-throughput and high-resolution imaging technique, we successfully identified 1109 G1 fetuses with diverse heart defects, leading to an overall defect rate of 10.8% (Fig. 1 and Table 1). The most frequently observed defect types were bicuspid aortic valve (BAV) and muscular ventricular septal defect (mVSD), each accounting for 30% of the total defects. Perimembranous ventricular septal defect (pmVSD) was the next most common defect and was observed in 11% of the fetuses. Outflow tract defects, including double outlet right ventricle (DORV), persistent truncus arteriosus (PTA), and transposition of the great arteries (TGA), accounted for 5% of the total defects. Atrial septal defect (ASD), in particular, secundum ASD which accounts for 70% of all ASD, is a common heart malformation in humans. However, the interatrial communication is normally present during fetal life, and consequently, the prenatal diagnosis of secundum ASD is rarely possible. We therefore did not identify any secundum ASD in this prenatal screen. Three primum ASDs were identified. Only a small percentage (2.4%) of heart defects were accompanied by visible external defects such as microcephaly and cleft palate (Additional file 1: Table S1). The distribution pattern of heart defect subtypes observed in this screen closely mirrored that of congenital heart defects observed in human patients [39, 40].

Characteristics of ENU-induced de novo mutations

We performed WES/WGS on a randomly selected subset of fetuses. We included 720 fetuses with heart defects but without visible external defects (case group) and 611 litter-matched fetuses with normal hearts (control group) (Additional file 1: Fig. S1). The case and control groups exhibited similar sequencing metrics, ensuring a valid comparison (Additional file 1: Table S2). To

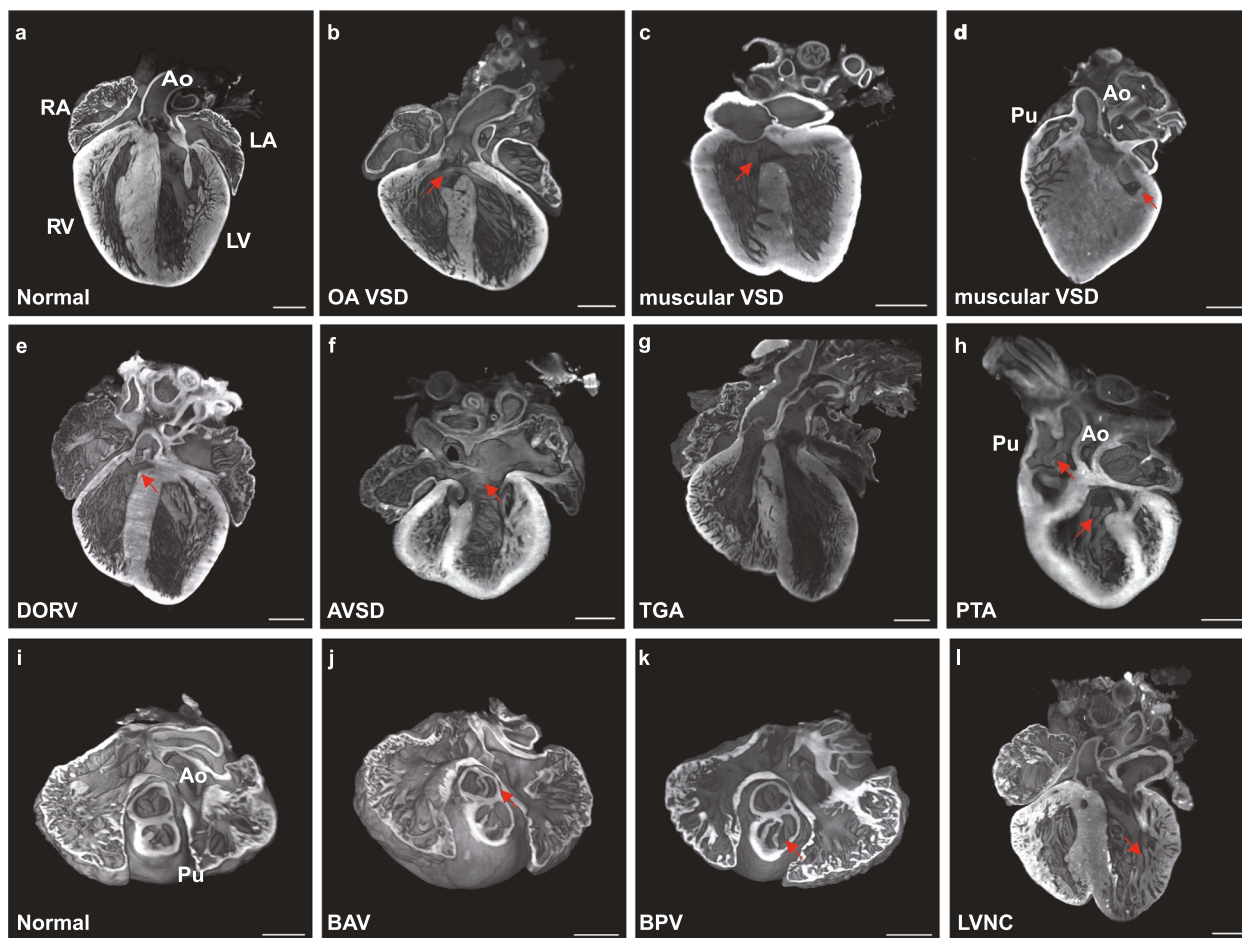


Fig. 1 Representative heart defects reconstructed by lightsheet fluorescence microscopy

Table 1 Heart defect rates in the G1 fetuses

Lightsheet diagnosis	G1 defect rate
Cardiac defect rate	1109/10,285 (10.8%)
BAV	325 (30%)
BPV	53 (5%)
Thickened aortic or pulmonary valve leaflets	66 (6%)
pmVSD/OA VSD	118 (11%)
Muscular VSD	328 (30%)
ASD (primum)	3 (0.3%)
AVSD	11 (1%)
DORV with VSD or AVSD	38 (3%)
TGA with VSD or AVSD	6 (0.5%)
PTA with VSD or AVSD	7 (0.6%)
Heterotaxy	1 (0.09%)
Coronary artery fistula	10 (0.9%)
Noncompaction	9 (0.8%)

distinguish rare ENU-induced variants from the background variations, variants that occurred four times or more in all 1331 samples were removed. Since there is a small chance of a clonal relationship among offspring, which would interfere with the burden analysis, a kinship analysis using an inbreeding coefficient cutoff of 0.05 was applied to remove related samples. We further filtered for the variants that occurred only once in all 1135 samples to ensure that we only analyzed purely ENU-induced de novo mutations. Finally, 598 cases and 532 controls were retained, each presumably derived from an independently edited single spermatogonial stem cell.

Regarding the variant classification, we found that non-synonymous single nucleotide variants (SNVs) were the most prevalent, accounting for 69% of all variants. Synonymous SNVs accounted for 24% of the variants. The remaining 7% were predominantly loss of function variants such as splicing, stop gain, and indels (Additional

file 1: Fig. S2a). We observed similar distributions of variants across all chromosomes in both the case and control groups, with the adenine and thymine being the predominant edited bases (Additional file 1: Fig. S2b and S2c). These findings suggest that there were no systemic differences in variant calling or annotation between the cases and controls. A total of 15807 coding genes were affected by ENU at least once in the 1130 case and control samples. This finding represents a coverage of approximately 76% of the whole mouse exome (Additional file 1: Fig. S2c). On average, the G1 progeny exhibited 59 (66,928/1130) exonic ENU-induced de novo variants on average per fetus. Notably, this number is 53 times greater than the observed de novo mutation rate in humans, as reported in previous studies [9]. This increased mutation load provides an opportunity to explore and identify risk genes associated with heart defects.

Increased mutation burden in mice with CHD

To assess the difference in mutation burden between the case and control groups, we developed an expected mutational model under the ENU treatment. This model involved simulating all possible nucleotide changes to the whole mouse protein coding sequence and deriving the expected frequency of each possible variant based on several factors, including the average of 59 exonic ENU-induced de novo variants per sample, transcript length, and ENU mutation bias (Additional file 2: Table S3).

All the variants were classified into distinct classes, such as synonymous, missense, and loss-of-function (LoF) variants, which included stop gain, stop loss, start loss, splicing, and frameshift variants. Damaging missense mutations (D-Mis) were defined as missense mutations predicted to be damaging by SIFT-4G, the only prediction algorithm available for mice. The expected and observed numbers of variants in each variant class were subsequently compared for the control and case groups using a one-tailed Poisson test [9] (Table 2). As expected, the mutation rates in all variant classes were accurately predicted in the control group. However, we observed a 1.13-fold excess of LoF mutations in the case group across all the genes, indicating a greater burden of LoF mutations in the cases compared to the controls ($p = 3.9 \times 10^{-10}$). Furthermore, damaging mutations, including LoF and D-Mis variants in genes known to be related to CHD [32], and LoF mutations in genes highly expressed in the developing heart [9, 35] were markedly enriched in CHD cases but not in controls. In contrast, genes with low expression in the heart [9, 35] were not enriched with any genetic variants in either the case or control groups. Interestingly, although previous studies have implicated recessive mutations in cilia genes in both mouse [15] and human CHD [34], we did not observe

significant enrichment of heterozygous mutations of cilia genes in this dominant screen. However, we found an increased incidence of LoF mutations in chromatin genes [34] in the cases but not in the controls (Additional file 1: Table S4).

The identical genetic backgrounds of the case-control mice and uniform sequence coverage allowed us to directly compare the mutation rates between the cases and controls using one-tailed Poisson tests [35] (Additional file 1: Table S5). In line with the previous analysis based on the expectation model, the cases had a significant excess of LoF mutations across all genes and genes with high heart expression. Additionally, CHD-related genes were enriched in cases for the damaging variants (LoF and D-Mis). These findings indicate that simultaneously disrupting single alleles of multiple genes in the germ line can lead to heart defects in the offspring.

Heart contraction genes enriched in mice with CHD

For each gene, we considered all qualifying variants within the specified variant classes and summed their allele counts in the case group, control group, and expected mutational model separately. We then performed a one-sided binomial test to determine whether there was a significant deviation in the frequency of damaging mutations in each gene from the expected distribution. After correcting for multiple testing ($FDR < 0.05$, Storey's q-value procedure [31]), a total of 148 and 25 genes were significantly enriched in the cases and controls, respectively (Additional file 2: Table S6 and S7). Among all these case-enriched genes, *Notch1* [41, 42], *Fbn2* [43], *Prr12* [44], and *Rere* [45] are established CHD genes with an autosomal dominant inheritance pattern in humans. No genes enriched in the control mice are known to be associated with CHD with an autosomal dominant inheritance pattern.

To gain further insights into the biological mechanisms underlying CHD, we conducted GO term and KEGG pathway enrichment analyses for the 148 genes overrepresented in mice with heart defects (Fig. 2 and Additional file 2: Table S8). Genes involved in regulating heart contraction, such as calcium ion transmembrane transport, action potential, and muscle structure development, were among the most enriched pathways. Notable examples include *Ryr2* and *Ryr3*, which encode ryanodine receptors involved in excitation-contraction coupling, *Atp2a2* and *Atp2b2*, which encode subunits of ATP-driven Ca^{2+} ion pumps critical for cardiac relaxation, and *Cacna1e* and *Cacna1s*, which encode calcium voltage-gated channel subunits required for calcium entry. *Abcc9* and *Kcnma1*, which encode subunits of potassium channels critical for regulating membrane potential, were also enriched.

Table 2 Variant class enrichment by expectation analysis

	Case N = 598						Control N = 532					
	Observed		Expected		Enrichment	P	Observed		Expected		Enrichment	P
	n	Rate	n	Rate			n	Rate	n	Rate		
All genes												
Total	36,148	60.4	35,418.5	59.2	1.02	5.6E−05	30,780	57.9	31,509.5	59.2	0.98	1.00
Synonymous	8446	14.1	9006.2	15.1	0.94	1.00	7392	13.9	8012.2	15.1	0.92	1.00
Missense	25,029	41.9	24,164.4	40.4	1.04	1.6E−08	21,219	39.9	21,497.4	40.4	0.99	0.97
D-mis	13,729	23.0	13,427.6	22.5	1.02	0.005	11,723	22.0	11,945.6	22.5	0.98	0.98
LoF	2518	4.2	2221.6	3.7	1.13	3.9E−10	2008	3.8	1976.4	3.7	1.02	0.24
Damaging	16,247	27.2	15,649.2	26.2	1.04	1.0E−06	13,731	25.8	13,922.0	26.2	0.99	0.95
CHD genes 1044												
Total	3367	5.6	3205.4	5.4	1.05	0.002	2690	5.1	2851.6	5.4	0.94	1.00
Synonymous	770	1.3	813.8	1.4	0.95	0.94	647	1.2	723.9	1.4	0.89	1.00
Missense	2341	3.9	2189.4	3.7	1.07	0.0007	1866	3.5	1947.8	3.7	0.96	0.97
D-mis	1415	2.4	1270.5	2.1	1.11	3.6E−05	1073	2.0	1130.3	2.1	0.95	0.96
LoF	234	0.4	199.8	0.3	1.17	0.01	163	0.3	177.8	0.3	0.92	0.87
Damaging	1649	2.8	1470.3	2.5	1.12	2.5E−06	1236	2.3	1308.0	2.5	0.94	0.98
HHE genes 4420												
Total	9025	15.1	8777.9	14.7	1.03	0.004	7562	14.2	7809.1	14.7	0.97	1.00
Synonymous	2137	3.6	2249.8	3.8	0.95	0.99	1881	3.5	2001.5	3.8	0.94	1.00
Missense	6164	10.3	5947.2	9.9	1.04	0.003	5150	9.7	5290.8	9.9	0.97	0.97
D-mis	3623	6.1	3494.8	5.8	1.04	0.02	3030	5.7	3109.1	5.8	0.97	0.92
LoF	694	1.2	574.3	1.0	1.21	7.1E−07	504	0.9	510.9	1.0	0.99	0.63
Damaging	4317	7.2	4069.1	6.8	1.06	6.1E−05	3534	6.6	3620.0	6.8	0.98	0.93
LHE genes 3145												
Total	5736	9.6	5635.0	9.4	1.02	0.09	4912	9.2	5013.0	9.4	0.98	0.92
Synonymous	1259	2.1	1409.6	2.4	0.89	1.00	1130	2.1	1254.1	2.4	0.90	1.00
Missense	4108	6.9	3876.3	6.5	1.06	0.0001	3477	6.5	3448.5	6.5	1.01	0.32
D-mis	2193	3.7	2112.3	3.5	1.04	0.04	1889	3.6	1879.2	3.5	1.01	0.41
LoF	343	0.6	344.8	0.6	0.99	0.55	280	0.5	306.7	0.6	0.91	0.94
Damaging	2536	4.2	2457.1	4.1	1.03	0.06	2169	4.1	2185.9	4.1	0.99	0.64

n, number of ENU-induced mutations; *rate*, number of ENU-induced mutations divided by the number of fetuses in the group (*N*); *Enrichment*, the ratio of observed to expected rate of mutations; *HHE*, high heart-expressed genes (top quartile of genes expressed in E14.5 mouse hearts as previously reported) [9, 35]; *LHE*, lower heart-expressed genes (bottom quartiles of genes expressed in E14.5 mouse hearts as previously reported) [9, 35]; *D-Mis*, damaging missense variants predicted by SIFT-4G; *Damaging*, *D-Mis*+LoF variants. *P*-values were calculated to compare the number of variants in each category between the observed and expected values using a one-tailed Poisson test. The bold numbers indicate Enrichment > 1.1 and *P* < 0.05

Interestingly, many genes involved in neuronal function and development were also highly overrepresented in the cases. These included genes involved in neurite growth and axon guidance (*Slit2*, *Chl1*, *Celsr2*), neuronal migration (*Wdr47*, *Kif26a*), and neurotransmitter secretion (*Stxbp5*, *Stxbp5l*). In contrast, our analysis did not identify any significant enrichment of pathways or processes for genes with increased mutations in the control group. These findings highlight the importance of genes involved in heart contractility and potential neuronal regulation of early cardiac functions in the development of CHD.

To further characterize the genes overrepresented in mice with heart defects, we submitted the case-enriched

148 genes to the MGI mammalian phenotype term enrichment analysis. The results confirm a significant enrichment for abnormal channel response, abnormal cardiovascular morphology, impaired muscle contractility, and abnormal neurological response (Additional file 1: Fig. S3a). Furthermore, to ascertain that the observed association between the case-enriched geneset and these phenotypes is statistically robust and not merely a reflection of the overall mutation characteristics in our screen, we performed a permutation test from the pool of ENU-induced mutations from which the 148 case-enriched genes were derived. The analysis revealed that our case-enriched geneset was significantly

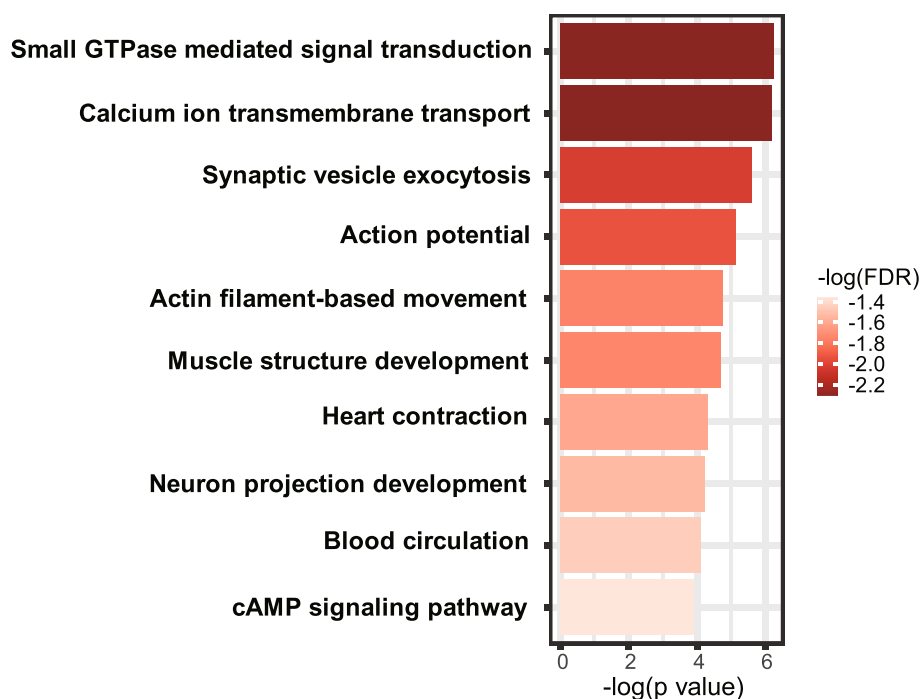


Fig. 2 GO term enrichment analysis for genes enriched in mice with heart defects. Pathways with q -value < 0.05 (Benjamini-Hochberg method) were considered significant

distinct from the random 10,000 permutation datasets ($p < 2.2e-16$, Hotelling's T -squared test [46]) (Additional file 1: Fig. S3b). Lastly, to further substantiate that the case-enriched geneset is specifically associated with cardiovascular system and nervous system phenotypes, we plotted a density map of all p -values for these two terms across the 10,001 datasets (Additional file 1: Fig. S3c and S3d). The findings demonstrate that the p -value for our case-enriched gene set falls within the top 0.01% of all p -values for the nervous system phenotype, and top 0.17% for the cardiovascular system phenotype, when ranked from smallest to largest. Taken together, the association of case-enriched genes with cardiac contraction and neuronal function and development is specific.

Heart contraction-related genes enriched in human CHD

We obtained WES/WGS data from 3406 CHD probands from the US National Heart, Lung, and Blood Institute (NHLBI) Pediatric Cardiac Genomics Consortium (PCGC). After excluding aortic arch patterning defects that were not found in our mouse screen, we obtained a sample size of 1457 probands with CHD. Out of these probands, 1333 also had WES or WGS data available for their parents. The defect types included VSD (15%), pulmonary stenosis (15%), TGA (12%), ASD (10%), Tetralogy of Fallot (9%), aortic stenosis (9%), BAV (4%), and others. WES of 2675 control subjects were obtained from the

1000 Genomes Project [27]. Variants were co-called from all bam files of cases and controls by GATK as described in the Methods.

In the gene burden analysis, we summed the number of all qualifying variants of each gene in each variant class for the 1457 cases and 2675 controls, respectively. Damaging variants were defined as LOF or predicted to be pathogenic by AlphaMissense [30]. We used a one-sided Fisher's exact test to identify genes with significant differences in the frequency of rare damaging mutation (LoF+D-Mis, $\text{MAF} < 0.001$) between the cases and controls. Due to the small sample size and small number of mutation events in most genes being analyzed, which resulted in generally modest p -values in Fisher's exact test, we have chosen not to apply multiple testing corrections and directly used p -value < 0.05 for the significance test in this circumstance. Since this approach may increase the potential for false positive findings, we have conducted a comparison of the mutation tolerability of the 373 genes identified to be overrepresented in the cases and 432 genes overrepresented in the controls (Additional file 3: Table S9 and S10). The comparison revealed that the genes associated with CHD were functionally less tolerant to damaging mutations, indicating their potential role in the development of CHD. As shown in Fig. 3a from the gnomAD database of constraint scores, we found that the observed/expected scores of

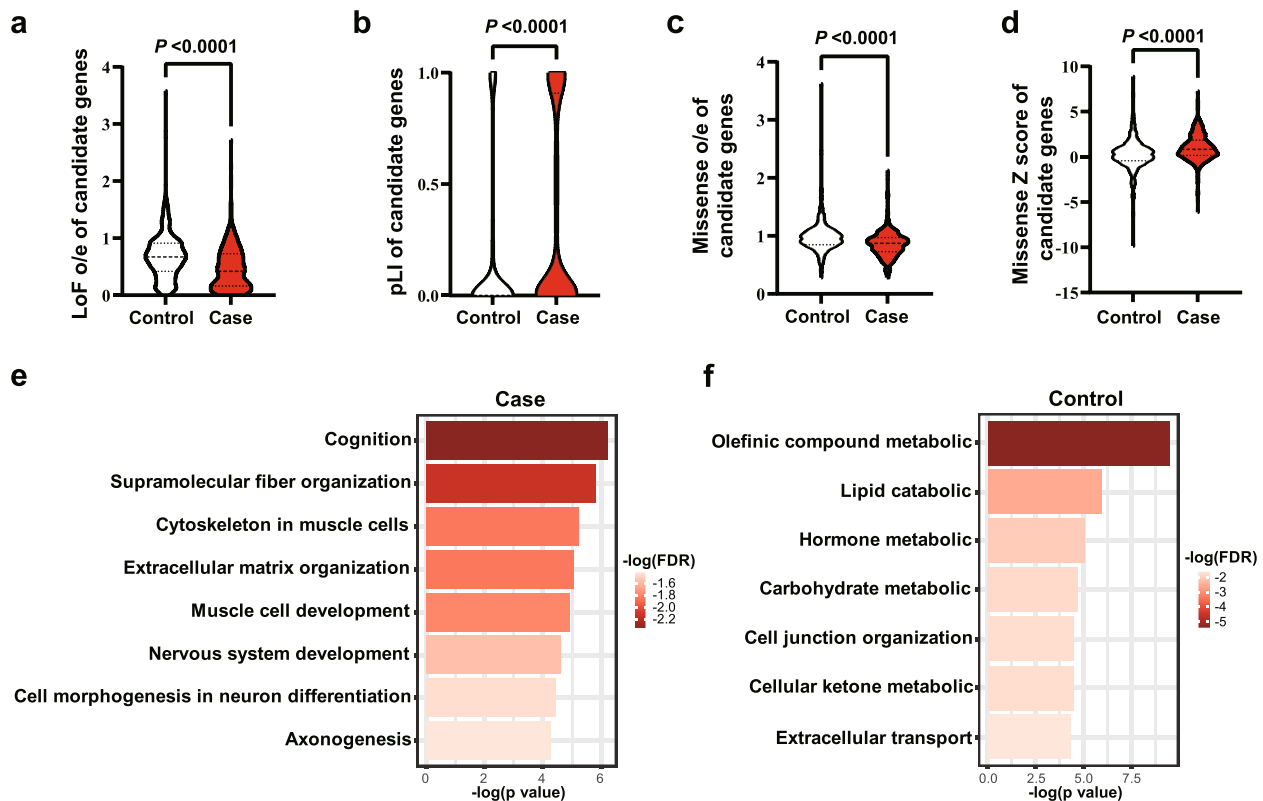


Fig. 3 Pathway and process enrichment analysis for human CHD. **a** The ratio of the observed/expected (oe) number of loss-of-function variants in genes enriched in cases and controls. **b** pLI score for genes enriched in cases and controls. **c** The ratio of the observed/expected (oe) number of missense variants in genes enriched in cases and controls. **d** Missense Z score for genes enriched in cases and controls. **e** Pathway and process enrichment analysis for genes enriched in the cases. **f** Pathway and process enrichment analysis for genes enriched in the controls

LoF variants of these case-enriched genes were statistically significantly lower than those of genes enriched in controls. Accordingly, the probability of loss-of-function intolerant scores (pLI) was higher in the case-enriched genes (Fig. 3b). Missense variants were also significantly depleted for genes enriched in the case compared to that in control-enriched genes (Fig. 3c, d). These results indicate that the case-enriched genes are less tolerant to damaging mutations than other genes, further supporting their potential role in CHD pathogenesis.

Enrichment analysis of these genes revealed that cognition and nervous development were among the top enriched cellular processes (Fig. 3e, Additional file 3: Table S9 and Table S11). These included genes involved in axon growth and guidance (*PLXND1*, *SEMA3B*, *SEMA3D*, *SEMA6A*, *ULK2*), neurotrophic factors, and transcription factors critical for neuron differentiation (*ATOH1*, *NOTCH1*, *LHX2*, *NTF4*) and synaptogenesis and neurotransmission (*SLITRK1*, *LRTM2*, *CLSTN2*). Genes involved in muscle contraction and development, such as the calcium voltage-gated channel subunit *CACNAIS* and genes required for myofibril assembly (*MYH11*,

TNNT1, *NRAP*, and *OBSCN*), were also enriched in cases. No genes related to heart contraction or nervous system development were enriched in the control group (Fig. 3f, Additional file 3: Table S10 and S12).

To explore potential causal genetic factors for individual CHD probands, we conducted a search for digenic gene sets characterized by the concurrent occurrence of mutations in the same probands or in the same mice exhibiting heart defects. The following criteria were applied to identify these digenic gene sets: (1) both genes in the pair must have rare damaging mutations, as previously defined, in at least one mouse and at least two human CHD probands from the 1333 trios; (2) the gene pair should not have concurrent rare damaging mutations in the control mice; (3) the gene pair should not have concurrent rare damaging mutations in either parent of the proband; (4) the gene pair should not have concurrent rare damaging mutations in the 2675 control subjects from the 1000 Genomes Project. Based on these criteria, 101 candidate digenic gene sets were identified for future validation. (Additional file 3: Table S13).

Discussion

The molecular genetics of CHD has long been a challenging puzzle to solve. Based on the total CHD recurrence rate of 5% and an incidence rate of 1% in the general population [5], a simple calculation suggests that genetic factors may contribute to approximately 75% of CHD cases. However, currently known genetic factors such as copy number variations, aneuploidy, de novo mutations, and transmitted variants only account for approximately 30% of CHD cases overall [47].

Genome-wide association studies (GWAS) have been conducted to identify common susceptibility loci for CHD. However, these GWAS studies often suffer from small cohort sizes, low reproducibility, and small effect sizes of identified loci [47, 48]. A recent meta-analysis by Yu et al [49] addressed some of these issues by conducting a large-scale analysis of 4597 cases and 50,745 control individuals from four CHD cohorts. Sixteen novel loci, including 12 rare noncoding variants with moderate or large effect sizes were identified. These loci were found to disrupt transcription factor binding sites involved in cardiac development or disrupt physical contact with key genes in cardiac development [49]. This study reproduced the well-known phenomenon of rare variants with large effect sizes and common variants with small effect sizes. While these discovered loci provide valuable insights into the molecular mechanisms of CHD and confer overall risk for CHD, the percentage of CHD cases substantially influenced by each specific variant is likely to be small. As an alternative to GWAS, gene-level burden testing has been developed to condense multiple rare inherited variants into individual genes before conducting association studies [50, 51]. However, to the best of our knowledge, there have been no reports of whole exome-wide gene burden testing in either animal genetic screens or human CHD studies.

In this study, a saturating ENU mutagenesis screen in mice led to the identification of a large number of genes that were previously unknown to be associated with CHD. Notably, all these variants were present in a heterozygous state in the affected animals, consistent with the oligogenic inheritance model. GO term and KEGG pathway enrichment analysis revealed that heart contraction and neuronal genes were the most significantly enriched in mice with CHD. This finding was consistent with subsequent findings from a human case-control gene burden test.

This study employed LFM as a valuable tool for screening heart defects in mice at a large scale. LFM provides high resolution and rapid imaging speeds (less than 20 s per heart), making it highly effective for 3D reconstructions, particularly when compared to more traditional methods such as micro-CT, MRI, and stacked histology

images. This technique has enabled us to identify subtle heart defects, including BAV, which comprises a significant 40% of the detected anomalies. The Zeiss Z1 LFM model we utilized is optimized for imaging mouse hearts from E14 to postnatal day 7. Specimens outside this developmental window may not benefit from the same level of resolution at 5× objective or may not fit within the imaging chamber. Similar to other imaging techniques that focus on anatomical features, LFM does not directly provide insights into functional deficits. Conditions such as aortic or pulmonary stenosis require additional hemodynamic data for a more complete assessment. Defects like VSD or outflow tract (OFT) misalignment are more straightforward to identify, while assessing the extent of left ventricular non-compaction (LVNC) can be more subjective. In our study, we applied a non-compacted to compacted myocardium ratio greater than 4 as a criterion for diagnosing LVNC. This stringent criterion resulted in an incidence rate of 0.09% for LVNC in our sample which may be somewhat underestimated. Given the overall 10× incidence rate of heart defects observed in this study compared to human CHD rates, this incidence rate of LVNC is disproportionately lower than the 0.076% prevalence rate of LVNC in human newborns, as reported by Kock et al. [52].

In this screen, the average litter size at E18.5 was recorded at 7, a figure that aligns closely with the standard litter size of 6–8 pups observed in wild-type, untreated C57BL/6J mice [53]. Only 0.8% of fetuses were found nonviable at the time of harvest at E18.5. Given that all identified mutations are paternally inherited and present in a heterozygous state, embryonic lethality was infrequent. However, it is noteworthy that fetuses exhibiting severe cardiac malformations may succumb postnatally.

It is interesting to note that except a very small number of genes such as *Notch1*, the majority of these genes identified through the mouse screen and human CHD burden test are not traditionally considered classical monogenic CHD-causing genes based on family linkage analysis. Furthermore, most of these genes do not cause heart defects or severely impact adult heart function in mice when only one copy of the gene is deleted (Additional file 2: Table S6). One possible explanation for the involvement of these heart contraction genes in CHD pathogenesis is that they may exhibit haploinsufficiency, specifically during a critical window of heart development when the developing heart is most vulnerable to hemodynamic disruption. Doppler ultrasound imaging of early-stage mouse embryos has revealed that from the onset of heartbeat (E8.0) through E14.5, which is a critical period of cardiac morphogenesis, there is a progressive increase in heart rate, peak velocities, and

cardiac output [54, 55]. This contractile change is accompanied by a steady increase in the expression of cardiac contractile proteins and subunits of various sarcolemma or sarcoplasmic reticulum ion pumps and channels [56], leading to changes in electrophysiology and calcium handling [57]. Based on these observations, it is possible that a decreased dosage of a specific contraction-related gene during a certain stage of heart development may transiently disrupt the balance between gradual myocyte maturation and the increasing metabolic demands, resulting in a temporary disruption of cardiac function, such as heart rate, rhythm, or contractility. This functional impairment might subside eventually after full maturation of the heart. This theory needs to be tested by knocking out one copy of individual candidate cardiac contraction genes and studying its impact on early cardiac function and their risk to CHD.

Another interesting finding is the association of neuronal genes with CHD. The intricate relationship between neural regulation and early cardiac function and morphogenesis is not yet fully elucidated. Innervation of the cardiac conduction system, primarily originating from neural crest cells of neuroectodermal lineage, follows intricate migratory pathways guided by a complex interplay of factors, including neurotrophic molecules, axon guidance proteins, differentiation signals, and survival cues [58]. Notably, many of the genes involved in these processes have been identified through our studies in both mouse models and human subjects. Anatomical studies in mouse embryos indicate that parasympathetic neurons expressing the vesicular acetylcholine transporter (VACHT) and sympathetic neurons expressing tyrosine hydroxylase (TH) are present near the venous pole and the dorsal meso-cardial connection, respectively, by E12.5 [59]. These positions align closely with the developing atrioventricular and sinus nodes. However, the functionality of these early neural elements is uncertain, with conventional wisdom suggesting that functional cardiac innervation emerges later in fetal development, post-morphogenesis [60, 61]. A meticulous investigation of the early developmental status of the cardiac conduction system, in conjunction with an analysis of early cardiac rhythm, would be indispensable in genetic models of compromised NCC migration, differentiation and innervation.

Contrary to prevailing beliefs, both human and mouse embryonic heart rates have been demonstrated to react to cholinergic and adrenergic stimulation during the morphogenesis period, at mouse E12.5 and prior to human week 8 of gestation [61–63]. Mice deficient in catecholamines, specifically those lacking dopamine β -hydroxylase [64] or tyrosine hydroxylase [65], exhibit cardiovascular failure and begin dying as early as E11.5. Consistent

with this, cardiac intrinsic catecholamine-producing cells have been localized predominantly to the dorsal venous valve and atrioventricular canal regions from E11.5 [66]. This highlights the critical role of β 1-adrenergic receptor signaling in maintaining fetal heart rate during morphogenesis, particularly in response to hypoxia-induced bradycardia [62]. Given that the final closure of inter-ventricular communications occurs around E13.5 [67] and the tri-leaflet semilunar valve formation and remodeling continue beyond this stage [68], it is plausible that disruptions to cardiac contraction due to neuronal or cardiac intrinsic adrenergic inadequacy during this critical window could perturb the normal hemodynamics which is known to be an important risk factor for cardiac malformation [69].

Another possible explanation for these observations is that certain genes annotated to have a function in nerve conduction may also regulate action potential in cardiac muscles. For example, the R-type calcium channel *Cacna1e* is expressed in both the heart and the central nervous system. Ablation of *Cacna1e* has been shown to cause arrhythmia in isolated prenatal mouse hearts [70], and mutations in *CACNA1E* have been associated with developmental and epileptic encephalopathy in humans [71]. Similarly, reduced activity of *Atp2a2* (SERCA Ca^{2+} -ATPases that are responsible for translocation of calcium from the cytosol into the sarcoplasmic reticulum lumen) led to a significant reduction in single action potential-driven Ca^{2+} signals and synaptic exocytosis [72] and impairs cardiac contractility and relaxation in mice [73]. Mutation of *ATP2A2* is associated with skin disorder as well as neuropsychiatric disorders and heart failure [74]. This finding may potentially explain why neurodevelopmental disorders and arrhythmias are prevalent comorbidities affecting survivors with CHD from a genetic perspective [75, 76].

It remains uncertain as to how mutations in genes regulating cardiac contraction contribute to the risk of CHD. Mutations in genes encoding cardiac ion channels and intracellular calcium handling are associated with congenital arrhythmia syndromes [77]. Although arrhythmia is a common comorbidity in patients with structural heart defects [78], definitive monogenic causes of CHD attributable to calcium handling genes remain less well-established. We hypothesize that heterozygous loss-of-function mutations in ion channel and calcium handling genes may induce cardiac arrhythmias during certain stage of fetal heart development. However, hemodynamic disturbance alone may not be sufficient to cause morphological malformations unless it occurs in the context of mutations in other susceptibility genes, such as *Notch1*, *Klf2*, and *Yap* which are known to be involved in the endocardial response to flow shear stress and mediate the

process of endocardial-to-mesenchymal transition [79]. For instance, transiently inducing bradycardia in mouse embryos at E9.5 through pharmacological inhibition of the rapid component of the delayed rectifier potassium current (I_{Kr}) with dofetilide, or blocking L-type calcium channels with verapamil, does not result in heart defects on its own. However, when these pharmacological agents are combined with a heterozygous *Notch1* mutation, over 50% of fetuses exhibit various heart defects due to abnormal endocardial-to-mesenchymal transition, highlighting the multifactorial nature of CHD [80]. These additional factors are yet to be found. Thus, simultaneous hemi-knockout of the candidate digenic gene sets identified in this study would be necessary to test this hypothesis in future experiments.

One limitation of this mouse screen study is the relatively small sample size. While the study managed to cover 76% of all mouse coding genes, the burden-based testing results may be biased toward larger genes, potentially leading to an underpowered statistical analysis for smaller yet functionally important genes. Based on our expectation model and an average of 59 hits per mouse, it is estimated that 10,000 independent G1 mice would be required to cover 80% of all coding genes with at least 5 hits per gene. This suggests the need for larger-scale studies to achieve a more robust statistical analysis and capture the full spectrum of genetic variations associated with CHD.

Conclusions

The mouse forward genetic screen resulted in the identification of numerous genes that contribute to an increased risk of CHD when present in the heterozygous mutant state. Notably, genes involved in regulating cardiac contraction and nervous system development and functions were significantly enriched in CHD cases. These findings align with the results obtained from a gene-based burden testing on human CHD probands, which further emphasized impaired heart contraction as a previously underappreciated risk factor for CHD. The identification of candidate digenic gene sets in this study holds promise for shedding light on the complex genetics underlying CHD and should be further investigated in future studies.

Abbreviations

CHD	Congenital heart disease
ENU	N-Ethyl-N-nitrosourea
BAV	Bicuspid aortic valve
BPV	Bicuspid pulmonary valve
VSD	Ventricular septal defect
pmVSA	Perimembranous VSD
OA VSD	Overriding aortic ventricular septal defect
ASD	Atrium septal defect
AVSD	Atrioventricular septal defect
DORV	Double outlet right ventricle
TGA	Transposition of great arteries

PTA	Persistent truncus arteriosus.
LVNC	Left ventricular noncompaction
WES	Whole-exome sequencing
WGS	Whole-genome sequencing
PCGC	Pediatric Cardiac Genetics Consortium
HHE	High heart expression genes
LHE	Low heart expression genes
MAF	Minor allele frequency
PCA	Principal component analysis
FDR	False discovery rate

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-024-01372-x>.

Additional file 1: Table S1. The incidence of isolated and syndromic heart defects in the G1 fetuses; Fig S1. Summary of the ENU dominant screening procedures and variants filtering processes; Table S2. Summary sequencing statistics for the mouse cases and controls; Fig S2. Characterization of ENU-induced mutations; Table S4. Variant class enrichment by expectation analysis; Table S5. Variant class enrichment in cases versus controls; Fig S3. Characterization of ENU-induced mutations against the MGI mammalian phenotype database

Additional file 2: Table S3. The mutation simulation dataset of ENU treatment; Table S6. 148 genes with rare damaging variants that were exceeded than expected in CHD mice; Table S7. 25 genes with rare damaging variants that were exceeded than expected in normal mice; Table S8. Top terms from enrichment analysis among 148 genes with rare damaging variants that were exceeded than expected in CHD mice

Additional file 3: Table S9. 373 genes with rare damaging variants that were overrepresented in CHD probands; Table S10. 432 genes with rare damaging variants that were overrepresented in normal subjects; Table S11. Top terms from enrichment analysis among 373 genes with rare damaging variants overrepresented in CHD humans; Table S12. Top terms from enrichment analysis among 432 genes with rare damaging variants overrepresented in normal humans; Table S13. 101 gene groups that had multiple concurrent damaging mutations in mouse and human subjects with CHD

Acknowledgements

The human CHD WES data were generated by the Pediatric Cardiac Genomics Consortium (PCGC), under the auspices of the National Heart, Lung, and Blood Institute's Bench to Bassinet Program <https://benchtoBassinet.com> (dbGaP Study Accession: phs001194.v3.p2). The Pediatric Cardiac Genomics Consortium (PCGC) program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through grants UM1HL128711, UM1HL098162, UM1HL098147, UM1HL098123, UM1HL128761, and U01HL131003. This manuscript was not prepared in collaboration with investigators of the PCGC, has not been reviewed and/or approved by the PCGC, and does not necessarily reflect the opinions of the PCGC investigators or the NHLBI.

We would also like to extend our thanks to Dr. Jian Yang for his advice on statistical analysis. We are grateful to the Westlake Animal Facility for their excellent care of the mice, and to the microscopy core facility of Westlake University for their assistance with microscopy. Special thanks go to Youshi Chen and Jianfeng Wang for their assistance with mouse breeding.

Authors' contributions

H.S. raised hypothesis, designed experiments, interpreted data, and drafted the work. X.L. performed ENU screening and statistical analysis. L.L. performed WES and WGS variant calling and QC. X.L. and H.R. conducted fetal heart phenotyping. N.L. and L.Y. advised on the bioinformatic analysis. All authors read and approved the final manuscript.

Funding

This work was supported by the Westlake Education Foundation, China National GeneBank (CNGB), and Natural Science Foundation of Zhejiang Province of China (LZ19H040001).

All mouse WES/WGS Bam files can be accessed from NCBI SRA under the accession number SRP467869 (<https://www.ncbi.nlm.nih.gov/sra/?term=SRP467869>) [81]. Code for human WGS/WES and mouse WES/WGS data analysis can be found at <https://github.com/ShiLabNGS/CHDWGS> [82]. The human CHD WES data were generated by the Pediatric Cardiac Genomics Consortium (PCGC), were available in dbGaP (database of Genotypes and Phenotypes) under accession phs001194.v3.p2, (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001194.v3.p2) [83]. The datasets supporting the conclusions of this article are included within the article and its additional files.

Declarations

Ethics approval and consent to participate

All animal experiments were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee of Westlake University (approval 21-005-SHJ). This study was carried out in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹College of Life Sciences, Zhejiang University, Hangzhou, Zhejiang, China. ²School of Medicine, Westlake University, Hangzhou, Zhejiang, China. ³Key Laboratory of Growth Regulation and Translational Research of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou 310024, China. ⁴Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China. ⁵Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China. ⁶Westlake University High-Performance Computing Center, Westlake University, Hangzhou, Zhejiang, China.

Received: 20 January 2024 Accepted: 5 August 2024

Published online: 12 August 2024

References

- van der Linde D, Konings EEM, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJM, Roos-Hesselink JW. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J Am Coll Cardiol*. 2011;58(21):2241–7.
- Zimmerman MS, Smith AGC, Sable CA, Echko MM, Wilner LB, Olsen HE, Atalay HT, Awasthi A, Bhutta ZA, Boucher JL, et al. Global, regional, and national burden of congenital heart disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Child Adolesc Health*. 2020;4(3):185–200.
- Liu Y, Chen S, Zuhlke L, Black GC, Choy MK, Li N, Keavney BD. Global birth prevalence of congenital heart defects 1970–2017: updated systematic review and meta-analysis of 260 studies. *Int J Epidemiol*. 2019;48(2):455–63.
- Collaborators GCHD. Global, regional, and national burden of congenital heart disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Child Adolesc Health*. 2020;4(3):185–200.
- Pierpont ME, Brueckner M, Chung WK, Garg V, Lacro RV, McGuire AL, Mital S, Priest JR, Pu WT, Roberts A, et al. Genetic basis for congenital heart disease: revisited: a scientific statement from the American Heart Association. *Circulation*. 2018;138(21):e653–711.
- Helm BM, Landis BJ, Ware SM. Genetic evaluation of inpatient neonatal and infantile congenital heart defects: new findings and review of the literature. *Genes*. 2021;12(8):1244.
- Griffin EL, Nees SN, Morton SU, Wynn J, Patel N, Jobanputra V, Robinson S, Kochav SM, Tao A, Andrews C, et al. Evidence-based assessment of congenital heart disease genes to enable returning results in a genomic study. *Circ Genom Precis Med*. 2023;16(2):e003791.
- De Backer J, Callewaert B, Muino Mosquera L. Genetics in congenital heart disease. Are we ready for it? *Rev Esp Cardiol (Engl Ed)*. 2020;73(11):937–47.
- Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, DePalma SR, McKean D, Wakimoto H, Gorham J, et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science*. 2015;350(6265):1262–6.
- Jin SC, Homsy J, Zaidi S, Lu Q, Morton S, DePalma SR, Zeng X, Qi H, Chang W, Sierant MC, et al. Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. *Nat Genet*. 2017;49(11):1593–601.
- Gifford CA, Ranade SS, Samarakoon R, Salunga HT, de Soysa TY, Huang Y, Zhou P, Eifenbein A, Wyman SK, Bui YK, et al. Oligogenic inheritance of a human heart disease involving a genetic modifier. *Science*. 2019;364(6443):865–70.
- Priest JR, Osoegawa K, Mohammed N, Nanda V, Kundu R, Schultz K, Lamer EJ, Girirajan S, Scheetz T, Waggott D, et al. De novo and rare variants at multiple loci support the oligogenic origins of atrioventricular septal heart defects. *PLoS Genet*. 2016;12(4):e1005963.
- Russell WL, Kelly EM, Hunsicker PR, Bangham JW, Maddux SC, Phipps EL. Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. *Proc Natl Acad Sci U S A*. 1979;76(11):5818–9.
- Liu X, Yagi H, Saeed S, Bais AS, Gabriel GC, Chen Z, Peterson KA, Li Y, Schwartz MC, Reynolds WT, et al. The complex genetics of hypoplastic left heart syndrome. *Nat Genet*. 2017;49(7):1152–9.
- Li Y, Klena NT, Gabriel GC, Liu X, Kim AJ, Lemke K, Chen Y, Chatterjee B, Devine W, Damerla RR, et al. Global genetic analysis in mice unveils central role for cilia in congenital heart disease. *Nature*. 2015;521(7553):520–4.
- Salinger AP, Justice MJ. Mouse mutagenesis using N-ethyl-N-nitrosourea (ENU). *CSH Protoc*. 2008;2008:pdb.prot4985.
- Shomer NH, Allen-Worthington KH, Hickman DL, Jonnalagadda M, Newsome JT, Slate AR, Valentine H, Williams AM, Wilkinson M. Review of rodent euthanasia methods. *J Am Assoc Lab Anim Sci*. 2020;59(3):242–53.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60.
- Faust GG, Hall IM. SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics*. 2014;30(17):2503–5.
- Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P, Sambamba: fast processing of NGS alignment formats. *Bioinformatics*. 2015;31(12):2032–4.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–303.
- Rimmer A, Phan H, Mathieson I, Iqbal Z, Twigg SRF, Consortium WGS, Wilkie AOM, McVean G, Lunter G. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet*. 2014;46(8):912–8.
- Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
- Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res*. 2001;11(5):863–74.
- Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88(1):76–82.
- Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884–90.
- Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68–74.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome Project Data Processing S: The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
- Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir R, Roazen D, Thibault J, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43(1110):11.10.11-11.10.33.
- Cheng J, Novati G, Pan J, Bycroft C, Zemgulyte A, Applebaum T, Pritzel A, Wong LH, Zielinski M, Sargeant T, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*. 2023;381(6664):eadg7492.
- Storey JD. A direct approach to false discovery rates. *J Royal Stat Soc Series B (Stat Methodol)*. 2002;64(3):479–98.

32. Zhou WZ, Li W, Shen H, Wang RW, Chen W, Zhang Y, Zeng Q, Wang H, Yuan M, Zeng Z, et al. CHDbase: a comprehensive knowledgebase for congenital heart disease-related genes and clinical manifestations. *Genomics Proteomics Bioinformatics*. 2023;21(1):216–27.
33. van Dam TJ, Wheway G, Slaats GG, Group SS, Huynen MA, Giles RH. The SYSCILIA gold standard (SCGSv1) of known ciliary components and its applications within a systems biology consortium. *Cilia*. 2013;2(1):7.
34. Watkins WS, Hernandez EJ, Wesolowski S, Bigsrove BW, Sunderland RT, Lin E, Lemmon G, Demarest BL, Miller TA, Bernstein D, et al. De novo and recessive forms of congenital heart disease have distinct genetic and phenotypic landscapes. *Nat Commun*. 2019;10(1):4722.
35. Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, Romano-Adesman A, Bjornson RD, Breitbart RE, Brown KK, et al. De novo mutations in histone-modifying genes in congenital heart disease. *Nature*. 2013;498(7453):220–3.
36. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523.
37. Benjamini Y, Hochberg Y. Controlling the false discovery rate— a practical and powerful approach to multiple testing. *J R Stat Soc Series B (Methodological)*. 1995;57(1):289–300.
38. Baldarelli RM, Smith CL, Ringwald M, Richardson JE, Bult CJ. Mouse Genome Informatics G: Mouse Genome Informatics: an integrated knowledgebase system for the laboratory mouse. *Genetics*. 2024;227(1):iyae031.
39. Hoffman JJ, Kaplan S. The incidence of congenital heart disease. *J Am Coll Cardiol*. 2002;39(12):1890–900.
40. Wu W, He J, Shao X. Incidence and mortality trend of congenital heart disease at the global, regional, and national level, 1990–2017. *Medicine (Baltimore)*. 2020;99(23):e20593.
41. Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, Grossfeld PD, Srivastava D. Mutations in NOTCH1 cause aortic valve disease. *Nature*. 2005;437(7056):270–4.
42. Durbin MD, Cadar AG, Williams CH, Guo Y, Bichell DP, Su YR, Hong CC. Hypoplastic left heart syndrome sequencing reveals a novel NOTCH1 mutation in a family with single ventricle defects. *Pediatr Cardiol*. 2017;38(6):1232–40.
43. Wang M, Clericuzio CL, Godfrey M. Familial occurrence of typical and severe lethal congenital contractural arachnodactyly caused by missplicing of exon 34 of fibrillin-2. *Am J Hum Genet*. 1996;59(5):1027–34.
44. Chowdhury F, Wang L, Al-Raqad M, Amor DJ, Baxova A, Bendova S, Biamino E, Brusco A, Caluseriu O, Cox NJ, et al. Haploinsufficiency of PRR12 causes a spectrum of neurodevelopmental, eye, and multisystem abnormalities. *Genet Med*. 2021;23(7):1234–45.
45. Fregeau B, Kim BJ, Hernandez-Garcia A, Jordan VK, Cho MT, Schnur RE, Monaghan KG, Juusola J, Rosenfeld JA, Bhoj E, et al. De novo mutations of RERE cause a genetic syndrome with features that overlap those associated with proximal 1p36 deletions. *Am J Hum Genet*. 2016;98(5):963–70.
46. Harold H. The generalization of student's ratio. *Ann Math Stat*. 1931;2(3):360–78.
47. Diab NS, Barish S, Dong W, Zhao S, Allington G, Yu X, Kahle KT, Brueckner M, Jin SC. Molecular genetics and complex inheritance of congenital heart disease. *Genes*. 2021;12(7):1020.
48. Blue GM, Kirk EP, Giannoulitou E, Sholler GF, Dunwoodie SL, Harvey RP, Winlaw DS. Advances in the genetics of congenital heart disease: a clinician's guide. *J Am Coll Cardiol*. 2017;69(7):859–70.
49. Yu M, Aguirre M, Jia M, Gjoni K, Cordova-Palomera A, Munger C, Amgala D, Rosa Ma X, Pereira A, Tcheandjieu C, et al. Oligogenic architecture of rare noncoding variants distinguishes 4 congenital heart disease phenotypes. *Circ Genom Precis Med*. 2023;16(3):258–66.
50. Hui D, Mehrabi S, Quimby AE, Chen T, Chen S, Park J, Li B, Ruckenstein MJ, Rader DJ, Ritchie MD, et al. Gene burden analysis identifies genes associated with increased risk and severity of adult-onset hearing loss in a diverse hospital-based cohort. *PLOS Genetics*. 2023;19(1):e1010584.
51. Guo MH, Plummer L, Chan Y-M, Hirschhorn JN, Lippincott MF. Burden testing of rare variants identified through exome sequencing via publicly available control data. *Am J Human Gene*. 2018;103(4):522–34.
52. Borresen MF, Blixenkroner-Moller E, Kock TO, Sillesen AS, Vogg ROB, Pihl CA, Norsk JB, Vejlstup NG, Christensen AH, Iversen KK, et al. Prevalence of left ventricular noncompaction in newborns. *Circ Cardiovasc Imaging*. 2022;15(6):e014159.
53. Flurkey K, Curren JM, Leiter EH, Witham B, Laboratory J. The Jackson Laboratory Handbook on Genetically Standardized Mice. Bar Harbor: Jackson Laboratory; 2009.
54. Phoon CK, Aristizabal O, Turnbull DH. 40 MHz Doppler characterization of umbilical and dorsal aortic blood flow in the early mouse embryo. *Ultrasound Med Biol*. 2000;26(8):1275–83.
55. Ji RP, Phoon CKL, Aristizabal O, McGrath KE, Palis J, Turnbull DH. Onset of cardiac function during early mouse embryogenesis coincides with entry of primitive erythroblasts into the embryo proper. *Circ Res*. 2003;92(2):133–5.
56. Edwards W, Greco TM, Miner GE, Barker NK, Herring L, Cohen S, Cristea IM, Conlon FL. Quantitative proteomic profiling identifies global protein network dynamics in murine embryonic heart development. *Dev Cell*. 2023;58(12):1087–1105.e1084.
57. Guo Y, Pu WT. Cardiomyocyte maturation. *Circ Res*. 2020;126(8):1086–106.
58. Vegh AMD, Duim SN, Smits AM, Poelmann RE, Ten Harkel ADJ, DeRuiter MC, Goumans MJ, Jongbloed MRM. Part and parcel of the cardiac autonomic nerve system: unravelling its cellular building blocks during development. *J Cardiovasc Dev Dis*. 2016;3(3):28.
59. Hildreth V, Webb S, Bradshaw L, Brown NA, Anderson RH, Henderson DJ. Cells migrating from the neural crest contribute to the innervation of the venous pole of the heart. *J Anat*. 2008;212(1):1–11.
60. Marvin WJ, Hermsmeyer K, McDonald RI, Roskoski LM, Roskoski R. Ontogenesis of cholinergic innervation in the rat heart. *Circ Res*. 1980;46(5):690–5.
61. Papp JG. Autonomic responses and neurohumoral control in the human early antenatal heart. *Basic Res Cardiol*. 1988;83(1):2–9.
62. Chandra R, Portbury AL, Ray A, Ream M, Groelle M, Chikaraishi DM. Beta1-adrenergic receptors maintain fetal heart rate and survival. *Biol Neonate*. 2006;89(3):147–58.
63. Shigenobu K, Tanaka H, Kasuya Y. Changes in sensitivity of rat heart to norepinephrine and isoproterenol during pre- and postnatal development and its relation to sympathetic innervation. *Dev Pharmacol Ther*. 1988;11(4):226–36.
64. Thomas SA, Matsumoto AM, Palmiter RD. Noradrenaline is essential for mouse fetal development. *Nature*. 1995;374(6523):643–6.
65. Zhou QY, Quafe CJ, Palmiter RD. Targeted disruption of the tyrosine hydroxylase gene reveals that catecholamines are required for mouse fetal development. *Nature*. 1995;374(6523):640–3.
66. Ebert SN, Thompson RP. Embryonic epinephrine synthesis in the rat heart before innervation: association with pacemaker and conduction tissue development. *Circ Res*. 2001;88(1):117–24.
67. Liu X, Li C, Wang J, Jin Y, Zhu J, Li S, Shi H. The developmental processes of ventricular septal defects with outflow tract malalignment. *Ann Anat*. 2024;255:152293.
68. Odelin G, Faure E, Couplier F, Di Bonito M, Bajolle F, Studer M, Avierinos JF, Charnay P, Topilko P, Zaffran S. Krox20 defines a subpopulation of cardiac neural crest cells contributing to arterial valves and bicuspid aortic valve. *Development*. 2018;145(1):dev151944.
69. Courchaine K, Rykiel G, Rugonyi S. Influence of blood flow on cardiac development. *Prog Biophys Mol Biol*. 2018;137:95–110.
70. Lu ZJ, Pereverzev A, Liu HL, Weiergraber M, Henry M, Krieger A, Smyth N, Hescheler J, Schneider T. Arrhythmia in isolated prenatal hearts after ablation of the Cav2.3 (alpha1E) subunit of voltage-gated Ca2+ channels. *Cell Physiol Biochem*. 2004;14(1–2):11–22.
71. Helbig KL, Lauerer RJ, Bahr JC, Souza IA, Myers CT, Uysal B, Schwarz N, Gandini MA, Huang S, Keren B, et al. De novo pathogenic variants in CACNA1E cause developmental and epileptic encephalopathy with contractures, macrocephaly, and dyskinesias. *Am J Human Genet*. 2018;103(5):666–78.
72. de Juan-Sanz J, Holt GT, Schreiter ER, de Juan F, Kim DS, Ryan TA. Axonal endoplasmic reticulum Ca(2+) content controls release probability in CNS nerve terminals. *Neuron*. 2017;93(4):867–88 1866.
73. Periasamy M, Reed TD, Liu LH, Ji Y, Loukianov E, Paul RJ, Nieman ML, Riddle T, Duffy JJ, Doetschman T, et al. Impaired cardiac performance in heterozygous mice with a null mutation in the sarco(endo)plasmic reticulum Ca2+-ATPase isoform 2 (SERCA2) gene. *J Biol Chem*. 1999;274(4):2556–62.

74. Bachar-Wikstrom E, Wikstrom JD. Darier Disease - A Multi-organ Condition? *Acta Derm Venereol.* 2021;101(4):adv00430.
75. Downing KF, Oster ME, Klewer SE, Rose CE, Nembhard WN, Andrews JG, Farr SL. Disability among young adults with congenital heart defects: congenital heart survey to recognize outcomes, needs, and well-being 2016–2019. *J Am Heart Assoc.* 2021;10(21):e022440.
76. Hernández-Madrid A, Paul T, Abrams D, Aziz PF, Blom NA, Chen J, Chessa M, Combes N, Dagres N, Diller G, et al. Arrhythmias in congenital heart disease: a position paper of the European Heart Rhythm Association (EHRA), Association for European Paediatric and Congenital Cardiology (AEPC), and the European Society of Cardiology (ESC) Working Group on Grown-up Congenital heart disease, endorsed by HRS, PACES, APHRS, and SOLAECE. *EP Europace.* 2018;20(11):1719–53.
77. Glazer AM. Genetics of congenital arrhythmia syndromes: the challenge of variant interpretation. *Curr Opin Genet Dev.* 2022;77:102004.
78. Hernandez-Madrid A, Paul T, Abrams D, Aziz PF, Blom NA, Chen J, Chessa M, Combes N, Dagres N, Diller G, et al. Arrhythmias in congenital heart disease: a position paper of the European Heart Rhythm Association (EHRA), Association for European Paediatric and Congenital Cardiology (AEPC), and the European Society of Cardiology (ESC) Working Group on Grown-up Congenital heart disease, endorsed by HRS, PACES, APHRS, and SOLAECE. *Europace.* 2018;20(11):1719–53.
79. Duchemin AL, Vignes H, Vermot J, Chow R. Mechanotransduction in cardiovascular morphogenesis and tissue engineering. *Curr Opin Genet Dev.* 2019;57:106–16.
80. Mu Y, Hu S, Liu X, Tang X, Shi H. Mechanical forces pattern endocardial Notch activation via mTORC2-PKC pathway. *eLife.* 2024;13:RP97268.
81. Luo X, Liu L, Rong H, Liu X, Yang L, Li N, Shi H. (2024). ENU screening for genes related to congenital heart disease. Retrieved from <https://www.ncbi.nlm.nih.gov/sra/?term=SRP467869>.
82. Luo X, Liu L, Rong H, Liu X, Yang L, Li N, Shi H. (2024). ENU-based dominant genetic screen identifies contractile and neuronal gene mutations in congenital heart disease. Retrieved from <https://github.com/ShiLaBNGS/CHDWGS>.
83. Investigators PGC. National Heart, Lung, and Blood Institute (NHLBI) Bench to Bassinet Program: The Pediatric Cardiac Genetics Consortium (PCGC) Study. Retrieved from https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001194.v3.p2.

Publisher's Note

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