De novo mutations in Congenital Heart Disease with Neurodevelopmental and Other Birth Defects

Jason Homsy1,2,§, Samir Zaidi3,§, Yufeng Shen4,§, James S. Ware1,5,6,§, Kaitlin E. Samocha1,7, Konrad J. Karczewski1,7, Steven R. DePalma1,8, David McKean1, Hiroko Wakimoto1, Josh Gorham1, Sheng Chih Jin3, John Deanfield9, Alessandro Giardini9, George A. Porter Jr.10, Richard Kim11, Kaya Bilguvar3,12, Francesc Lopez12, Irina Tikhonova12, Shrikant Mane12, Angela Romano-Adesman13, Hongjian Qi4,14, Badri Vardarajan15, Lijiang Ma16, Mark Daly1,7, Amy E. Roberts17, Mark W. Russell18, Seema Mital19, Jane W. Newburger20, J. William Gaynor21, Roger E. Breitbart20, Ivan Iossifov22, Michael Ronemus22, Stephan J. Sanders23, Jonathan R. Kaltman24, Jonathan G. Seidman1, Martina Brueckner3,†, Bruce D. Gelb25,†, Elizabeth Goldmanz26,27,†, Richard P. Lifton3,28,†,* Christine E. Seidman1,8,29,†,* and Wendy K. Chung30,†,*

1Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA
2Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA
3Department of Genetics, Yale University School of Medicine, New Haven, Connecticut, USA
4Departments of Systems Biology and Biomedical Informatics, Columbia University Medical Center, New York, New York, USA
5NIHR Cardiovascular Biomedical Research Unit at Royal Brompton & Harefield NHS Foundation and Trust and Imperial College, London, London, UK
6National Heart & Lung Institute, Imperial College London, London, UK
7Analytical and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston Massachusetts, USA
8Howard Hughes Medical Institute, Harvard University, Boston, Massachusetts, USA
9Department of Cardiology, University College London and Great Ormond Street Hospital, London, UK
10Department of Pediatrics, University of Rochester Medical Center, The School of Medicine and Dentistry, Rochester, New York, USA
11Section of Cardiothoracic Surgery, University of Southern California Keck School of Medicine, Los Angeles, California, USA

†To whom correspondence should be addressed: ; Email: bruce.gelb@msm.edu, ; Email: goldmuntz@email.chop.edu, ; Email: martina.brueckner@yale.edu, ; Email: richard.lifton@yale.edu, ; Email: cseidman@genetics.med.harvard.edu, or ; Email: wkc15@cumc.columbia.edu
§These authors contributed equally to this work
*Co-senior authors
Editor correspondence to: cseidman@genetics.med.harvard.edu
Abstract

Congenital heart disease (CHD) patients have increased prevalence of extra-cardiac congenital anomalies (CA) and risk of neurodevelopmental disabilities (NDD). Exome sequencing of 1,213 CHD parent-offspring trios identified an excess of protein-damaging de novo mutations, especially in genes highly expressed in developing heart and brain. These mutations accounted for 20% of patients with CHD, NDD and CA but only 2% with isolated CHD. Mutations altered genes involved in morphogenesis, chromatin modification, and transcriptional regulation, including...
multiple mutations in RBFOX2, an mRNA splice regulator. Genes mutated in other cohorts ascertained for NDD were enriched in CHD cases, particularly those with coexisting NDD. These findings reveal shared genetic contributions to CHD, NDD, and CA and provide opportunities for improved prognostic assessment and early therapeutic intervention in CHD patients.

Extra-cardiac congenital anomalies (CA, structural or functional anomalies that arise in utero) occur in approximately 13% of newborns with congenital heart disease (CHD), including 2% with a genetic syndrome, almost twice the prevalence observed in infants without CHD (1). Newborns with CHD are also at risk for the emergence of neurodevelopmental disorders (NDD), including cognitive, motor, social and language impairments. NDD occurs in 10% of all children with CHD and in 50% with severe CHD (2). Explanations to account for the high frequency of CA and NDD in CHD patients include embryonic circulatory deficits and stresses associated with post-natal therapeutic interventions (3), but these hypotheses remain unproven.

We sequenced exomes in 1,213 CHD trios, (proband and their unaffected parents) enrolled in the Pediatric Cardiac Genetics Consortium (PCGC) (4) or the Pediatric Heart Network (PHN) (5) after excluding CHD cases with clinically recognized genetic syndromes. Analyses included 353 previously reported CHD trios (6). We compared de novo mutations identified in CHD that occurred in isolation, or accompanied by CA, NDD, or both (phenotypes in Tables S1 and Database S1). Previously sequenced trios (n=900) from the Simons Foundation Autism Research Initiative Simplex Collection, each consisting of the unaffected parents and sibling of a child with autism spectrum disorder, served as control trios (7–9).

CHD and control probands were analyzed for de novo mutations (Databases S2–3). To evaluate the significance of mutation frequencies, we adapted a recently reported de novo expectation model (10) to assess mutation rates by variant class (synonymous, loss of function [LoF; e.g., nonsense, frameshift, canonical splice disruptions], or missense). We derived gene-based rates of de novo mutation from local sequence context, and adjusted by per-base coverage separately in case and control cohorts (Databases S4–5). We extended the model by merging all possible transcripts to obtain transcript-independent probabilities and by adding rates for deleterious missense variants predicted by the Meta-SVM score (D-Mis) (11). This yielded an overall mean expected mutation rate of 1.1 de novo variant per proband.

The expected and observed numbers of de novo mutations in each variant class in all CHD and control subjects (Table 1) were compared using a Poisson distribution. De novo mutation rates per variant class were accurately predicted in controls, replicating previous model validations (10). However, among all CHD trios, we detected significant enrichment (i.e., observed divided by expected frequencies) of LoF and D-Mis variants of 1.3 (p=0.0016) and 1.6 (p=1.8×10⁻¹⁰), respectively, across all genes. The combination of LoF +D-Mis variants (hereafter denoted as “damaging”) was 1.4-fold enriched in CHD cases compared to expectation, similar to the observed case vs. control comparison (Table S2). This burden persisted after excluding 353 previously studied CHD trios (Table S3) and was
found in each CHD category (conotruncal defects, left ventricular outflow tract obstruction and “other”), except for heterotaxy, which showed no excess (Table S4).

Damaging de novo mutations were markedly increased in CHD cases (enrichment=2.4, p=5.1x10^{-24}) among 4,420 genes in the top quartile of expression during heart development (high heart expression, HHE (6)) (Table 1). Conversely, controls had no significant enrichment in de novo mutations in HHE genes. Neither cases nor controls were enriched in de novo mutations among genes within the lower three quartiles of developing heart expression (LHE) (Table 1). From the observed and expected values, we estimated that 58% of these damaging de novo mutations contributed to CHD.

Twenty-one genes had multiple damaging de novo mutations only in cases, an unlikely chance occurrence (Fig. 1A; median expected 7; p=1.1x10^{-5} by permutation), suggesting that these genes are likely to be pathogenic for CHD (Fig. 1B). Indeed, this list includes seven genes previously implicated in CHD (PTPN11, KMT2D, CHD7, MYH6, JAG1, NOTCH1, ZEB2). Enrichments were not observed among genes with multiple de novo synonymous variants in CHD cases or across any variant class among controls (Fig. 1A and Table S5). A variety of cardiac malformations were associated with mutations in each of these 21 genes (Fig 1B and Table S6). From simulations based on these data (12), we estimate that de novo mutations in ~392 HHE genes contribute to CHD pathogenesis (Fig. S1).

Within this 21-gene set, PTPN11, KMT2D (MLL2), and RBFOX2, each had significantly more damaging de novo mutations than expected (Bonferroni corrected threshold for genome-wide significance = p<9x10^{-7}; Fig. 1B). RBFOX2, an RNA binding protein that regulates alternative splicing has not been previously implicated in CHD. RBFOX2 harbored three distinct de novo LoF mutations, a highly significant finding (Fig. 1B; p=3.4x10^{-8}). Additionally, we previously identified a de novo copy number loss that encompasses RBFOX2 in another CHD proband (13). Remarkably, these four probands all have hypoplastic left heart syndrome (HLHS). RBFOX2 is critical for zebrafish heart development (14), and regulates epithelial-mesenchymal transitions (EMT) (15). Disruption of EMT is felt to underlie HLHS pathogenesis (16). Notably, we observed significant enrichment of damaging mutations in RBFOX2 target genes (17) in CHD cases (1.9-fold, p=6.6x10^{-8}) but not controls (Table S7).

Gene ontology (GO) analysis revealed enrichment of damaging de novo mutations in genes involved in anatomic structure morphogenesis (GO:0009653; 2.4-fold; Bonferroni p=3.4x10^{-14}), cardiovascular system development (GO:0072358; 3.2-fold; Bonferroni p=7.5x10^{-9}), neurodevelopmental abnormality (HP:0012759; 2.6-fold; Bonferroni p=1.8x10^{-6}), and others (Database S6). We replicated the reported excess of de novo LoF mutations affecting chromatin modification (6), even after including only newly studied cases (GO:0016568, 5.1-fold enrichment, p-value=7.2x10^{-5}; Database S7). In the full CHD cohort, there were 25 de novo LoF mutations in chromatin modifying genes, a 5.3-fold enrichment over expectation (p=5.7x10^{-11}, Table S8; Fig. S2), strongly supporting the conclusion that these damaging de novo mutations impart large effects on CHD risk.
We examined the prevalence of damaging de novo mutations in CHD with or without NDD and/or CA (Fig 2A–C) after excluding 19 subjects found to have de novo mutations in known syndromic CHD genes and 279 subjects with uncertain NDD/CA status. Damaging de novo mutations in HHE genes were not significantly enriched in 356 subjects with isolated CHD nor in controls, but were ~3-fold enriched in 559 CHD cases with CA and/or NDD (CHD + Extra, \( p=1.1 \times 10^{-18} \)) including 97 probands diagnosed with either NDD or CA but unknown for the other phenotype. Excluding these 97 probands, we observed a 4.7-fold enrichment of damaging de novo mutations in HHE genes among 138 CHD cases with both NDD and CA (\( p=5.6 \times 10^{-15} \)), a 2-fold enrichment in CHD cases with only NDD (252 probands, \( p=3.8 \times 10^{-4} \)) and a 2.9-fold enrichment in CHD cases with only CA (72 probands, \( p=7.4 \times 10^{-4} \)). By comparing de novo rates in cases against expectation, we estimate that damaging de novo mutations in HHE genes contributed to 20% of CHD with both NDD and CA (95% confidence interval 12–30%), 10% (7–14%) of CHD with CA and/or NDD, 10% (2.5–23%) and 6% (2–11%) of CHD with CA or NDD only, respectively, and only 2% (0.5–5%) of isolated CHD (Fig. 2B). These results implied frequent pleiotropic effects of de novo mutations in CHD and raise the possibility that mutations in these same genes might also contribute to non-syndromic NDD and/or other CA. Indeed, we find that genes mutated in CHD are not only enriched for high expression in developing heart, they are also enriched for high expression in developing brain (Table S9).

To further explore these pleiotropic effects, we considered whether genes with damaging de novo mutations in CHD with NDD overlapped with 1,161 genes (Database S8) found to contain damaging de novo mutations in seven cohorts ascertained for NDD phenotypes excluding CHD (published NDD, P-NDD gene set) \((7, 18–23)\). Sixty-nine genes (Table S10) with damaging de novo mutations (n=85 mutations) were shared in CHD and P-NDD cohorts, far more than expected by chance (expected = 32 mutations; 2.6-fold enrichment; \( p=8.9 \times 10^{-15} \), Fig. 3A and Table S19). HHE genes were particularly enriched among P-NDD genes that were mutated in CHD (4.4-fold for all CHD cases, \( p=1.2 \times 10^{-23} \); Fig. 3A and Table S11). Moreover, genes mutated both in P-NDD and CHD cohorts are in the top quartile of both developmental heart and brain expression far more than expected by chance (observed = 38, expected = 11, binomial test, Fig. 3B). Input of these 69 overlapping genes into GO ontology analysis revealed significant terms that were broadly involved in the regulation of developmental transcription programs. These included 19 chromatin modifiers (GO:0016568 9.3-fold, \( p=8.5 \times 10^{-10} \); Database S9 and Fig. S2) including genes responsible for altering the methylation, acetylation or ubiquitination status of numerous regulatory lysine residues on the nucleosome. Additionally, there were 32 transcriptional regulators (GO:0006355 2.8-fold \( p=1.5 \times 10^{-4} \); Database S9), including genes involved in the Wnt (CTNNB1, DVL3, LRP5) and Notch (NOTCH1, EP300) signaling, important pathways in cardiac development. These findings demonstrate shared genetic etiologies for CHD and NDD patients, and confirm pleiotropic effects of mutations in these genes. Because it is unlikely that many P-NDD-ascertained patients with these mutations had clinically important CHD, and not all CHD patients with these mutations have NDD (see below), our findings also indicate that these mutations have variable expressivity, including isolated CHD, isolated NDD, or both.
Consistent with these observations, CHD subjects with damaging de novo mutations in these 69 overlapping genes (Fig. 3A) had a significantly increased risk of NDD (absolute risk of 73%, odds ratio 3.1, p=7.9×10^{-5}, Fisher exact test). Damaging mutations (28 mutations in 27 subjects) in chromatin modifiers showed the highest risk of NDD (19 subjects with NDD, 8 with unknown NDD status due to age <1 year at evaluation). Moreover, the marked enrichments in damaging de novo mutations among P-NDD genes with HHE (Fig. 3A and Table S11), 7.4-fold in 413 CHD cases with NDD (p=3.9×10^{-22}), 4.1-fold in 362 CHD infants with unknown NDD status (p=2.2×10^{-7}) and no significant enrichment in 438 CHD cases without NDD (p=0.075), strongly implies a future risk of NDD among CHD infants with these variants. These observations suggest that genotype is a strong predictor for future development of NDD in CHD infants. Despite these highly significant findings, our estimates are based purely on statistical grounds and limited to in silico predictions of damaging variants, a caveat that should be considered when extrapolating these results to identify causative/predictive mutations in individual patients.

Contemporary therapeutic interventions have substantially improved survival among newborns with serious CHD. Despite these advances, many life-long medical issues remain. The demonstration that damaging de novo gene mutations cause CHD, particularly when associate with NDD and other birth defects has both clinical and research implications. First, clinical genotyping may help stratify CHD patients and identify those at high risk for NDD, enabling surveillance and early interventions to improve school performance, employability and quality of life. Second, the pleiotropic consequence of these mutations implies that further study of these genes may uncover critical regulation of broad developmental programs. Finally, the high frequency of mutation in transcriptional regulators suggests that mutations in regulatory elements (promoters and enhancers) may be additional causes of CHD, particularly isolated CHD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to the patients and families who participated in this research and team members who supported subject recruitment and sequencing: Danielle Awad, Carlos Breton, Katrina Celia, Charina Duarte, Davina Etwaru, Nathan Fishman, Mereurt Kaspakova, Jennie Kline, Rosalind Korsin, Alyssa Lanz, Emma Marquez, Dawn Queen, Ashley Rodriguez, Janine Rose, Jaswinder K. Sond, Dorothy Warburton, Abigail Wilpers, and Roslyn Yee (Columbia Medical School), B. McDonough, A. Monafo, J. Stryker (Harvard Medical School); N. Cross (Yale School of Medicine); S. M. Edman, J. L. Garbarini, J. E. Tusi, S. H. Woyciechowski (Children’s Hospital of Philadelphia); J. Ellashek, N. Tran (Children’s Hospital of Los Angeles); K. Flack L. Panesar, N. Taylor (University College London); D. Gruber, N. Stellato (Steve and Alexandra Cohen Children’s Medical Center of New York); D. Guevara, A. Julian, M. MacNeal, C. Mintz (Icahn School of Medicine at Mount Sinai); E. Taillie (University of Rochester School of Medicine and Dentistry). We thank P. Candrea, E. Mazaika, K. Pavlik, V. Spotlow, and M. Sotiropoulos for production exome sequences and variant confirmation.

This work was supported by grants from the National Heart, Lung, and Blood Institute (Pediatric Cardiac Genomics Consortium, Pediatric Heart Network and Cardiovascular Development Consortium) and the National Human Genome Research Institute of the National Institutes of Health, Howard Hughes Medical Institute, Simons Foundation for Autism Research, John S. LaDue Fellowship at Harvard Medical School, Medical Scientist Training Program and National Research Science Award, Academy of Medical Sciences, British Heart Foundation, Wellcome Trust, Arthritis Research UK and the NHHR Cardiovascular Biomedical Research Unit at Royal Brompton and Harefield NHS Foundation Trust and Imperial College London, Leducq Foundation, Heart and
References and Notes


Fig. 1.
Genes with multiple *de novo* mutations are candidate CHD risk genes. A: Histograms show the expected distribution of the number of genes containing multiple *de novo* mutations (empirically derived using 1M permutations, black), and the observed number of genes with multiple mutations in cases (red line) for each class. P-values were calculated by permutation. B: Twenty-one genes with multiple damaging *de novo* mutations in cases. P-values are from Poisson test against expectation, with significance threshold < 9x10^{-7}. Further details are shown in Table S6.
Fig. 2.
Burden of damaging *de novo* mutations in HHE genes among CHD cases with extra-cardiac phenotypes. A: The enrichment (ratio of observed/expected) of damaging *de novo* mutations in HHE genes is shown for each phenotype (± 95% confidence interval). Case probands were excluded if they carried *de novo* mutations in known CHD syndrome genes (n=19), had unknown extracardiac phenotype for both NDD and CA (n=6), or had one unknown phenotype and were negative for the other (n=273). Cases with either CA or NDD and unknown status for the other phenotype (n=97) were included in the “Extra” category but...
excluded from the “only” categories. B: Percent excess of individuals carrying damaging de novo mutations in HHE genes by indicated phenotype (± 95% confidence interval). Explanation of calculation is provided (12). C: Table of observed and expected de novo rates for the indicated variant classes by phenotype.
Fig. 3.
Genes containing de novo mutations in CHD cases show pleotropic developmental effects. 
A: Individuals with CHD carry an excess of damaging de novo mutations among 1,161 genes identified by containing damaging de novo mutations in 7 published studies of NDD (P-ND cohort) (7, 18–23). All CHD cases were subdivided by NDD status (CHD + NDD, n=417 subjects, CHD − NDD n=440, Unknown NDD, n=363). P-values ≤0.005 as indicated (stars) were calculated from Poisson test against model-derived distribution (values in Table S11). P-ND gene set (blue) was further filtered for HHE genes (P-ND/HHE, red, 564 genes). Enrichments are shown ± 95% confidence intervals. B: Percentile gene expression ranks (100=high) are shown for all genes (grey) in the developing brain and heart, highlighting 69 genes with damaging de novo mutations in both CHD cases and the P-ND cohort (blue or purple). Genes with multiple de novo mutations in CHD (red or purple) are shown. Point size represents numbers of de novo events.