Exome sequencing reveals distinct genetic architectures for syndromic and nonsyndromic congenital heart defects

Alejandro Sifrim[1*], Marc-Phillip Hitz[1,2*], Anna Wilsdon[5], Jeroen Breckpot[6], Saeed H. Al Turki[1,3,4], Bernard Thienpont[7,8], Jeremy McRae[1], Tomas W Fitzgerald[1], Tarjinder Singh[1], Ganesh Jawahar Swaminathan[1], Elena Prigmore[1], Diana Rajan[1], Hashim Abdul-Khaliq[9], Siddharth Banka[31,32], Ulrike Bauer[10], Jamie Bentham[43], Felix Berger[11], Shoumo Bhattacharya[43], Frances BuLock[12], Natalie Canham[22], Irina G. Colgiu[1], Catherine Cosgrove[43], Helen Cox[33] , Ingo Daehnert[13], Allan Daly[1], Alan Fryer[35], Marc Gewillig[14], Emma Hobson[20] , Kerstin Hoff[2], Tessa Homfray[25], Anne-Karin Kahlert[2], Ami Ketley[5], Hans-Heiner Kramer[2], Katherine Lachlan[26,27,28], Anne Lampe[34], Jacoba Louw[14], Ashok Kumar Manickara[15], Dorin Manase[15], Karen McCarthy[16], Kay Metcalfe[24], Ruth Newbury-Ecob[44] , Seham Omer[17], Soo-Mi Park[21], Michael Parker[29], Thomas Pickardt[10], Martin O. Pollard[1], Leema Robert[23], Kerry Setchfield[5], Brigitte Stiller[18], Chris Thornborough[12], Okan Toka[19], Hugh Watkins[43], Denise Williams[33], Michael Wright[30], Seema Mital[15], Piers Daubeney[40], Bernard Keavney[36], Judith Goodship[37], The UK10K Consortium, Riyadh Mahdi Abu-Sulaiman[17,38,39], Sabine Klaassen[11,41], Caroline F. Wright[1], Helen V. Firth[23], Jeffrey Barrett[1], Koenraad Devriendt[6], David R. Fitzpatrick[42], J. David Brook[5], The Deciphering Developmental Disorders Study, Matthew E. Hurles[1,£]

[*] These authors contributed equally to this study
[£] Corresponding author

1	Wellcome Trust Sanger Institute, Cambridge, United Kingdom
2	Department of Congenital Heart Disease and Pediatric Cardiology, UKSH Kiel, Kiel, Germany
3	Department of Pathology, King Abdulaziz Medical City, Riyadh, Saudi Arabia
4	Harvard Medical School Genetics Training Program, Boston, USA
5	School of Life Sciences, University of Nottingham, Queen's Medical Centre, Nottingham, United Kingdom
6	Center for Human Genetics, University Hospitals Leuven, Leuven, Belgium
7	Vesalius Research Center, VIB, Leuven, Belgium.
8	Department of Oncology, Laboratory for Translational Genetics, KU Leuven, Leuven, Belgium.

9	Department of Paediatric Cardiology, Saarland University, Homburg, Germany
10	Competence Network for Congenital Heart Defects, National Register for Congenital Heart Defects, Germany
11	German Heart Institute Berlin, Department of Congenital Heart Disease and Pediatric Cardiology, Berlin, Germany
12	East Midlands Congenital Heart Centre, Glenfield Hospital, Leicester, United Kingdom
13	Department of Pediatric Cardiology, Heart Center, University of Leipzig, Germany
14	Department of Pediatric Cardiology, University Hospitals Leuven, Leuven, Belgium
15	Hospital for Sick Children, Toronto, Canada
16	Royal Brompton Hospital and the National Heart and Lung Institute, Imperial College, United Kingdom
17	Division of Pediatric Cardiology, King Abdulaziz Cardiac Center, King Abdulaziz Medical City, Ministry of National Guard–Health Affairs, Riyadh, Saudi Arabia
18	Department of Congenital Heart Defects and Paediatric Cardiology, Heart Centre, University of Freiburg, Germany
19	Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Department of Pediatric Cardiology, Erlangen, Germany
20	Yorkshire Regional Genetics Service, Leeds Teaching Hospitals NHS Trust, Department of Clinical Genetics, Chapel Allerton Hospital, Leeds, United Kingdom
21	East Anglian Medical Genetics, Cambridge University Hospitals NHS Foundation Trust, Biomedical Campus, Cambridge, United Kingdom
22	North West Thames Regional Genetics Centre, North West London, Hospitals NHS Trust, The Kennedy Galton Centre, Northwick Park And St Mark's NHS Trust, Harrow, United Kingdom
23	South East Thames Regional Genetics Centre, Guy's and St Thomas' NHS Foundation Trust, Guy's Hospital, London, United Kingdom
24	Manchester Centre for Genomic Medicine, St Mary's Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom
25	South West Thames Regional Genetics, Centre, St George's Healthcare NHS Trust St George's, University of London, London, United Kingdom
26	Wessex Clinical Genetics Service, University Hospital Southampton, Princess Anne Hospital, Southampton, United Kingdom
27	Wessex Regional Genetics Laboratory, Salisbury NHS Foundation Trust, Salisbury District Hospital, Salisbury, Wiltshire, United Kingdom

28	Faculty of Medicine, University of Southampton, Southampton, United Kingdom
29	Sheffield Regional Genetics Services, SheffieldChildren's, NHS Trust, Sheffield, United Kingdom
30	Northern Genetics Service, Newcastle upon Tyne Hospitals NHS Foundation Trust, Institute of Human Genetics, International Centre for Life, Central Parkway, Newcastle upon Tyne, United Kingdom
31	Faculty of Medical and Human Sciences, Manchester Centre for Genomic Medicine, Institute of Human Development, University of Manchester, Manchester Academic Health Science Centre (MAHSC), Manchester, United Kingdom
32	Manchester Centre for Genomic Medicine, Central Manchester University Hospitals NHS Foundation Trust, MAHSC, Manchester, United Kingdom
33	West Midlands Regional Genetics Service, Birmingham Women's NHS Foundation Trust, Birmingham Women's Hospital, Birmingham, United Kingdom
34	Consultant in Clinical Genetics, Molecular Medicine Centre, Western General Hospital, Edinburgh, United Kingdom
35	Department of Clinical Genetic, Liverpool Women's NHS Foundation Trust, Alder Hey Children's Hospital, Eaton Road, West Derby, LIVERPOOL, L122AP, UNITED KINGDOM
36	Institute of Cardiovascular Sciences, University of Manchester, Manchester, UK
37	Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK
38	King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia.
39	King Abdullah International Medical Research Center (KAIMRC), Riyadh, Saudi Arabia
40	Division of Paediatric Cardiology, Royal Brompton Hospital; Reader in Paediatric Cardiology at Imperial College, London, United Kingdom
41	Experimental and Clinical Research Center (ECRC), Charite´ Medical Faculty and Max-Delbruck-Center for Molecular Medicine, Berlin, Germany
42	Medical Research Council (MRC) Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine (IGMM), University of Edinburgh, Western General Hospital, Edinburgh, United Kingdom
43	Department of Cardiovascular Medicine, University of Oxford, Oxford, United Kingdom
44	Department of Clinical Genetics, St Michael's Hospital, Bristol, United Kingdom

ACKNOWLEDGEMENTS:

We thank the families for their participation and patience. We are grateful to the Exome Aggregation Consortium for making their data available. The DDD study presents independent research commissioned by the Health Innovation Challenge Fund (grant HICF-1009-003), a parallel funding partnership between the Wellcome Trust and the UK Department of Health, and the Wellcome Trust Sanger Institute (grant WT098051). The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the UK Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South Research Ethics Committee and GEN/284/12, granted by the Republic of Ireland Research Ethics Committee). The research team acknowledges the support of the National Institutes for Health Research, through the Comprehensive Clinical Research Network. The authors wish to thank the Sanger Human Genetics Informatics team for their support in processing the data. We would like to thank the Pediatric Cardiac Genomics Consortium (PCGC) and dbGAP, for making the data publicly available.

D.R.F. is funded through an MRC Human Genetics Unit program grant to the University of Edinburgh.

S.H.A.T., S.O. and R.M.A-S. were supported by funding from King Abdullah International Medical Research Center (grant number RC12/037).

J.B. was supported by the Klinisch Onderzoeksfonds UZ; B.T. was supported by the CHAMELEO Marie Curie Career Integration Grant; J.L. and M.G. Eddy Merckx Research grant. K.D. was funded by the GOA/2012/015 grant.

A.K.M., D.M. and S.M. were supported by the Heart and Stroke Foundation of Ontario, Canadian Institutes of Health Research;

This study was supported by DZHK (German Center for Cardiovascular Research), partner sites: Berlin and Kiel.

This study was approved under the ethics approval (EA2/131/10) Berlin, Germany J.D.B is funded by British Heart Foundation Programme Grant RG/13/10/30376

The study was approved under East Midland Research Ethics Committee ref 6721

A.W is funded by a British Heart Foundation Clinical Fellowship FS/14/51/30879

ABBREVIATIONS:
CHD: Congenital Heart Defect
S-CHD: Syndromic CHD
NS-CHD: Non-Syndromic CHD
PTM: Protein-truncating mutation

PTV: Protein-truncating variant
PAM: Protein-altering mutation
PAV: Protein-altering variant
DNM: *De Novo* Mutation
TOF: Tetralogy of Fallot

AUTHOR CONTRIBUTIONS	
Patient recruitment:	A.W., J.B., H.A-K., U.B., F.B., F.B.L., K.H., A-K.K., J.L., K.P.M., S.O., T.P., B.S., C.T., O.T., I.D, H-H.K., M.G., B.K., J.G., R.M.A-S., S.M., P.D., S.K., D.F., A.L., A.F., H.V.F., K.D.E.M.,S.-M. P., N.C., L.R., K.M., T.H.,K.L., M.P., M.W., S.B., D. W., H.C.
Phenotyping:	M-P.H., A.W., J.B., H.A-K., U.B., F.B., F.B.L., A-K.K., J.L., A.K.M., D.M., K.P.M., S.O., B.S., C.T., O.T., I.D., H-H.K., M.G., B.K., J.G., R.M.A-S., S.M., P.D., S.K.,A.L., A.F, H.V.F.B.D.F., K.D. K.D.E.M.,S.-M. P., N.C., L.R., K.M., T.H.,K.L., M.P., M.W., S.B., D. W., H.C.
Sample preparation:	M-P.H., J.B., B.T., K.H., A-K.K., A.K., K.P.M., S.O., T.P., K.S., J.D.B.
Validation experiments:	E.P., D.R., M-P.H., S.H.A.T, K.H., A.K., K.S.,
Study design:	A.S., M-P.H., S.H.A.T, A.W., J.B., B.T., K.P.M., H-H.K., M.G., B.K., R.M.A-S., S.M., P.D., S.K.,H.V.F., C.F.W., J.C.B. D.F., K.D., J.D.B., M.E.H.
Method development and data analysis:	A.S., M-P.H., S.H.A.T, A.W., J.B., B.T., T.W.F.A.L., A.F., H.V.F., C.F.W.,J.C.B., G.J.S., J.M., T.S., M.E.H.
Data interpretation:	A.S., M-P.H., S.H.A.T, A.W., J.B., B.T., T.W.F., G.J.S., J.M., T.S.,H.V.F.,C.F.W.,J.C.B.P.D., S.K., D.F., K.D., J.D.B., M.E.H.
Writing:	A.S., M-P.H., S.H.A.T, A.W., J.B., B.T., T.W.F., M.E.H.
Supervision	M.E.H

Congenital Heart Defects (CHD) have a neonatal incidence of 0.8-1%[1,2]. Despite abundant examples of monogenic CHD in humans and mice, CHD has a low absolute sibling recurrence risk (~2.7%)[3], suggesting a considerable role for *de novo* mutations (DNM), and/or incomplete penetrance[4,5]. *De novo* protein-truncating mutations (PTM) have been shown to be enriched among the 10% of 'syndromic' patients with extra-cardiac manifestations[6,7]. We exome sequenced 4,593 individuals from 1,823 CHD families (1,371 trios, 32 multi-sibling families and 458 singletons), including both syndromic (S-CHD, n=604) and non-syndromic cases (NS-CHD, n=1,219). In S-CHD, we confirmed for known CHD-associated genes a significant enrichment of *de novo PTM*s, but not inherited protein truncating variants (PTVs), , consistent with recent findings[8]. Conversely, in NS-CHD we observed significant enrichment of PTVs in known CHD genes inherited from unaffected parents. We identified three novel genome-wide significant S-CHD disorders caused by DNMs in *CHD4*, *CDK13* and *PRKD*1. Our study reveals distinct genetic architectures underlying the low sibling recurrence risk in S-CHD and NS-CHD.

We evaluated the burden of DNMs within S-CHD and NS-CHD trios separately ($N_{S\text{-}CHD}$= 511, $N_{NS\text{-}CHD}$= 854). We classified DNMs into three distinct categories: PTMs (nonsense, frameshift and splice-site variants); protein-altering mutations (PAMs), encompassing missense and in-frame indels; and silent mutations. We compared the observed numbers of DNMs to those expected under a null mutational model[9], across a set of manually curated CHD-associated genes, non-CHD developmental disorder (DD) associated genes and all remaining protein coding genes (Supplementary Tables 1-3, Figure 1A). S-CHD probands exhibited the largest excess in PTMs (OR=81, $P=1.21\times10^{-43}$) and PAMs (OR=8.6, $P=7.35\times10^{-15}$) for autosomal dominant CHD genes. S-CHD probands also manifested a burden of PTMs in autosomal dominant DD-associated genes not currently associated with CHD (OR=18.4, $p=3.49\times10^{-13}$). In contrast, NS-CHD probands presented with a much lower burden of PTMs in CHD-associated genes (OR=7.3, $P=2.61\times10^{-4}$). Finally, we find a significant exome-wide excess of PAMs, but not silent mutations (after excluding CHD and DD genes) in both S-CHD and NS-CHD probands, suggesting additional undiscovered dominant CHD-associated genes. The excess of *de novo* PTMs in S-CHD cases reported here is of the same magnitude as that found in cases with severe developmental disorders without CHD and considerably higher than that found in Autism Spectrum Disorder (Figure 1D). Our finding of a marked difference in DNM burden between NS-CHD and S-CHD mirrors that observed in Autism between individuals with and without intellectual disability[10]. Neurodevelopmental deficits are by far the most common extra-cardiac manifestations observed in S-CHD[8].

To evaluate the contribution of incompletely penetrant inherited variants, we compared the burden of rare (MAF < 0.1%) inherited variants in the three previously described gene sets in the S-CHD and NS-CHD cases of European ancestry, relative to population-matched controls (n=12,031, Supplementary Note, Supplementary Figure PCA, Supplementary Table 3, Figure 1B). We observed a significant excess of rare inherited PTVs in autosomal dominant CHD genes in NS-CHD (OR=2.67,$p=1.1\times10^{-4}$), but not in S-CHD (p=0.3). The CHD-associated genes with

inherited PTVs in NS-CHD (Supplementary Table 4) have previously only been linked with non-syndromic or mild syndromic presentations, and were distinct from those with *de novo* PTMs in S-CHD (Figure 1C). Where possible, we re-evaluated available clinical information to confirm the unaffected status of the transmitting parent, and the non-syndromic presentation of the proband in NS-CHD families with these inherited PTVs. Non-syndromic presentations of inherited PTVs in several genes associated with mild CHD syndromes has previously been described (e.g. *ELN*[11], *JAG1*[12]). The opposing signals, for inherited PTVs and de novo PTMs in NS-CHD and S-CHD respectively, argues against systematic mis-classification of NS-CHD and S-CHD in our cohorts, and points instead to an appreciable role for incomplete penetrance for PTVs in NS-CHD. Moreover, we also observed an exome-wide excess of rare inherited PTVs (OR=1.08, $p=1.51 \times 10^{-5}$) in NS-CHD probands, even after excluding known CHD-associated and DD-associated genes, suggesting incomplete penetrance in additional, novel CHD-associated genes.

Using a previously described null mutation model[6,9], we evaluated individual genes for an excess of *de novo* PTMs and PAMs separately, defining genome-wide significance as $p < 1.3 \times 10^{-6}$ (Methods). We identified 11 genome-wide significant genes considering all CHD trios, zero genes when considering only NS-CHD but 12 when considering only S-CHD (Supplementary Table 5, Table 1, Figure 2A), despite the smaller sample size in the latter, due to their concentrated burden of DNMs. Ten of the 12 genome-wide significant genes were known to be DD-associated, although not all had previously been implicated in CHD (Supplementary Table X). To maximise power to detect novel causative genes, we focused on S-CHD trios without a plausible genetic cause among known DD- and CHD-associated genes (n=398) and identified three novel genes: *CDK13*, *CHD4* and *PRKD1*, at genome-wide significance (Supplementary Table 6, Table 2, Figure 2B), elaborated below.

We identified six S-CHD individuals (Figure 4A) with clustered de novo PAMs in the highly conserved serine/threonine protein kinase domain of cyclin-dependent kinase 13 (*CDK13*), which shows a marked depletion of PAVs in the European population (Figure 4B). Three probands carry an identical missense mutation (p.Asn842Ser), also observed in an additional S-CHD singleton (Figure 4B).
These seven S-CHD cases were characterised by septal defects (VSD n= 2, ASD N= 4), with two also presenting with pulmonary stenosis, as well as global developmental delay, slight to moderate microcephaly, clinodactyly of the 5th finger and a recognizable facial gestalt (Figure 4C, Supplementary Table X). Modelling of the kinase domain indicates that the observed mutations impair: ATP-binding, binding of the magnesium ion that is essential for enzymatic activity, or interactions with Cyclin K, with which *CDK13* forms a cyclin-dependent kinase complex (Figure 4D). This Cyclin K/CDK13 complex phosphorylates RNA polymerase II and is necessary for alternative splicing of RNA[13,14]. The knockout mice for *Cdk12*, the closest homolog for *Cdk13*, both of which have ubiquitous developmental expression patterns, die at post-implantation (E5.5) suggesting a strong developmental effect[15].

We observed five S-CHD individuals with DNMs in *CHD4* (4 PAMs and 1 PTM), which encodes a chromodomain containing protein that catalyses ATP-dependent chromatin remodelling as a core component of the nucleosome remodeling and histone deacetylase (NuRD) repressor complex[16]. Three patients manifested TOF or TOF-like features, while the remaining two had CoA and a septal defect (Supplementary Figure X, Supplementary Table X). The majority of patients presented with motor (N=3) and global developmental delay (N=3), with two of them showing severe structural brain defects and four had various genitourinary abnormalities. Haploinsufficiency of another component of the NuRD complex, *GATAD2B*, has been identified as causing a recognisable intellectual disability syndrome, although associated CHD has not been reported[17]. More generally, several components of other ATP-dependent chromatin remodelling complexes have been associated with dominant developmental syndromes, including CHD in some patients[6,7]. A possible explanation for the observed cardiac phenotype comes from a recent study showing, that mice with endothelial knockdown of *CHD4* die during midgestation of vascular rupture, due to a dysfunctional NuRD-complex[18].

We detected three S-CHD individuals with *de novo* PAMs in *PRKD1*, with two having identical DNMs, a mutational pattern suggestive of gain of function (Supplementary Figure X, Supplementary Table X). These three individuals had different types of CHD (PS, AVSD and ???), severe developmental delay, ectodermal (dry skin, teeth and nail defects) and limb abnormalities. A homozygous PTV in *PRKD1* has recently been associated with truncus arteriousus through autozygosity mapping[19]. *PRKD1* encodes a serine/threonine protein kinase that regulates diverse cellular functions, including the transcriptional response to cardiac hypertrophy[20]. Homozygous knockout of PRKD1 in mice is embryonic lethal and tissue-specific knockout results in abnormal cardiac remodelling[20].

The burden analyses described above clearly show an enrichment for both PTMs, PAMs and inherited PTVs within our CHD dataset, therefore we hypothesised that some genes might be enriched for both *de novo* and rare inherited variants and that integrating both classes of variation, using a previously described hierarchical Bayesian model[21] (Supplementary Note), may improve power to detect novel CHD-associated genes. We analysed PTMs and PAMs separately (Supplementary Figure) and considered candidate CHD-associated genes at strong (FDR < 1%), intermediate (1%< FDR < 5%) and weak (5%< FDR < 10%) levels of confidence (Supplementary Table X and X, Figure 3). We found 16 genes at the strongest level of confidence, 12 were known DD-associated genes (2 of which have not been previously associated with CHD: *DYRK1A*, *PACS1*), 1 gene was only associated with CHD but not DD (*MYH6*), and 3 are novel candidate genes (*CHD4*, *CDK13*, *DIAPH3*). Most high confidence genes, exhibited enrichment for either DNMs or inherited variants, only two genes, *NOTCH1* and *KAT6A* exhibited appreciable enrichment for both. *NOTCH1* was notable as being the only high confidence gene for which the evidence from inherited PTVs exceeds that from DNMs (Figure 3A). Due to the likely concentration of false discovery signals in novel gene associations, we believe the evidence presented here to be insufficient to conclusively assert novel CHD associations based on this analysis alone. Additional functional evidence can prioritise genes for future follow-

up studies (Supplementary Table X). We evaluated the over-representation of particular gene functions and pathways among the top 374 genes with (FDR < 50%). We observed a significant over-representation (FDR < 10%) of genes associated with Gene Ontology terms related to chromatin modification, protein phosphorylation, neural tube and cardiac development (Supplementary Table X). Over-represented pathways included: NOTCH1-, IGF1-, HDAC Class II-, ERBB- and NFKB- signalling (Supplementary Table X). In addition, the set of top-ranking genes showed a significant overrepresentation of high quality (STRING Score > 0.9) protein-protein interactions within the gene set ($p=5.84x10^{-3}$, Supplementary Figure X), with key nodes being *NOTCH1*, *SOS1*, *EP300* and *SMAD4*.

Several mechanisms have been proposed to explain the low sibling recurrence risk of CHD, ranging from a major role for DNMs[7], incomplete penetrance of high risk variants, and a polygenic and/or multifactorial aetiology[22]. Our analyses show that the relative contributions of DNMs and incomplete penetrance differ markedly between NS-CHD and S-CHD, with a major role for *de novo* mutations in the latter, and inherited high risk variants in the former. By focusing on S-CHD cases with no clear known diagnosis, we discovered three novel causes for S-CHD (*PRKD1*, *CHD4* and *CDK13*). CHD is often not fully penetrant in syndromic CHD disorders (e.g. *KMT2D*[23], *NSD1*[24]), and as all patients in our study were ascertained for CHD, further studies are necessary to quantify the penetrance of CHD in these three new syndromes.

Current sample sizes provide limited statistical power to detect novel S-CHD disorders, and given the observed burden of *de novo* PTMs in S-CHD we estimate that data sets at least 20-fold larger will be needed to discover most dominant CHD-associated genes (Supplementary Figure). This challenge is likely to be even greater for identifying most genes harbouring incompletely penetrant variation in NS-CHD[25]. Our data motivate different future study design strategies for S-CHD (trios) and NS-CHD (case/control), nonetheless international collaboration and data sharing will be essential to achieve a deeper understanding of the genetic architecture of CHD.

1. Hoffman, J. I. E. & Kaplan, S. The incidence of congenital heart disease. *J. Am. Coll. Cardiol.* **39,** 1890–900 (2002).

2. Øyen, N. *et al.* Recurrence of congenital heart defects in families. *Circulation* **120,** 295–301 (2009).

3. Gill, H. K., Splitt, M., Sharland, G. K. & Simpson, J. M. Patterns of recurrence of congenital heart disease: an analysis of 6,640 consecutive pregnancies evaluated by detailed fetal echocardiography. *J. Am. Coll. Cardiol.* **42,** 923–9 (2003).

4. Schulkey, C. E. *et al.* The maternal-age-associated risk of congenital heart disease is modifiable. *Nature* **520,** 230–3 (2015).

5. Li, Y. *et al.* Global genetic analysis in mice unveils central role for cilia in congenital heart disease. *Nature* (2015). doi:10.1038/nature14269

6. The Deciphering Developmental Disorders Consortium. Large-scale discovery of novel genetic causes of developmental disorders. *Nature* **10,** (2014).

7. Zaidi, S. *et al.* De novo mutations in histone-modifying genes in congenital heart disease. *Nature* **498,** 220–3 (2013).

8. Homsy, J. *et al.* De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science (80-.).* **350,** 1262–1266 (2015).

9. Samocha, K. E. *et al.* A framework for the interpretation of de novo mutation in human disease. *Nat. Genet.* **46,** 944–50 (2014).

10. Iossifov, I. *et al.* The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **515,** 216–21 (2014).

11. Metcalfe, K. *et al.* Elastin: mutational spectrum in supravalvular aortic stenosis. *Eur. J. Hum. Genet.* **8,** 955–63 (2000).

12. Bauer, R. C. *et al.* Jagged1 (JAG1) mutations in patients with tetralogy of Fallot or pulmonic stenosis. *Hum. Mutat.* **31,** 594–601 (2010).

13. Liang, K. *et al.* Characterization of human cyclin-dependent kinase 12 (CDK12) and CDK13 complexes in C-terminal domain phosphorylation, gene transcription, and RNA processing. *Mol. Cell. Biol.* **35,** 928–38 (2015).

14. Blazek, D. *et al.* The Cyclin K/Cdk12 complex maintains genomic stability via regulation of expression of DNA damage response genes. *Genes Dev.* **25,** 2158–72 (2011).

15. Chen, H.-H., Wang, Y.-C. & Fann, M.-J. Identification and characterization of the CDK12/cyclin L1 complex involved in alternative splicing regulation. *Mol. Cell. Biol.* **26,** 2736–45 (2006).

16. Polo, S. E., Kaidi, A., Baskcomb, L., Galanty, Y. & Jackson, S. P. Regulation of DNA-damage responses and cell-cycle progression by the chromatin remodelling factor CHD4. *EMBO J.* **29,** 3130–3139 (2010).

17. de Ligt, J. *et al.* Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. *N. Engl. J. Med.* **367,** 1921–1929 (2012).

18. Ingram, K. G., Curtis, C. D., Silasi-Mansat, R., Lupu, F. & Griffin, C. T. The NuRD chromatin-remodeling enzyme CHD4 promotes embryonic vascular integrity by transcriptionally regulating extracellular matrix proteolysis. *PLoS Genet.* **9,** e1004031 (2013).

19. Shaheen, R. *et al.* Positional mapping of PRKD1, NRP1 and PRDM1 as novel candidate disease genes in truncus arteriosus. *J. Med. Genet.* **52,** 322–9 (2015).

20. Fielitz, J. *et al.* Requirement of protein kinase D1 for pathological cardiac remodeling. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 3059–63 (2008).

21. He, X. *et al.* Integrated model of de novo and inherited genetic variants yields greater power to identify risk genes. *PLoS Genet.* **9,** e1003671 (2013).

22. Pierpont, M. E. *et al.* Genetic Basis for Congenital Heart Defects : Current Knowledge A Scientific Statement From the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular. *Circulation* **115,** 3015–3038 (2007).

23. Miyake, N. *et al.* MLL2 and KDM6A mutations in patients with Kabuki syndrome. *Am. J. Med. Genet. A* **161A,** 2234–43 (2013).

24. Tatton-Brown, K. *et al.* Genotype-phenotype associations in Sotos syndrome: an analysis of 266 individuals with NSD1 aberrations. *Am. J. Hum. Genet.* **77,** 193–204 (2005).

25. Kiezun, A. *et al.* Exome sequencing and the genetic basis of complex traits. *Nat. Genet.* **44,** 623–30 (2012).

26. Consortium, E. A. *et al. Analysis of protein-coding genetic variation in 60,706 humans*. *bioRxiv* (Cold Spring Harbor Labs Journals, 2015). doi:10.1101/030338