Published in final edited form as:
Circ Res. 2014 October 24; 115(10): 821–823. doi:10.1161/CIRCRESAHA.114.305243.

Copy Number Variants and The Genetic Enigma of Congenital Heart Disease

AJ Marian, M.D.
Center for Cardiovascular Genetics, Institute of Molecular Medicine and Department of Medicine, University of Texas Health Sciences Center at Houston, and Texas Heart Institute, Houston, TX 77030

For years, many reasoned, perhaps not so naïvely that genetics of congenital heart diseases (CHD) would be the last frontier in efforts to elucidate the genetic etiology of cardiovascular disease. This impression was based on two main reasons. First, CHD seldom exhibit a clear familial inheritance pattern, as opposed to many single gene disorders. This is despite the strong evidence for familial aggregation of CHD and a higher risk of recurrence in the offspring, which denote a clear genetic etiology [1, 2]. CHD are often sporadic or part of the aneuploidy syndromes with pleomorphic non-cardiac phenotypes [1, 3]. Thus, an approach distinct from the genetic linkage analysis in large families, which was commonly applied to delineate the genetic basis of hereditary cardiovascular diseases such as cardiomyopathies and arrhythmia syndromes [4-8], was needed to define genetic etiology of CHD. The second reasoning was the extreme phenotypic assortment of CHD [3, 9-11], which is far beyond phenotypic variability of single gene disorders as well as the diversity of common complex diseases, such as coronary artery disease and systemic arterial hypertension. Consequently, it was difficult envisioning how the apparently the simple phenotype of atrial septal defects (ASD) and the complex phenotype of tetralogy of Fallot (TOF), which share no anatomical and physiological similarities, would etiologically share a common class of genetic networks, let alone arise from mutations in a single gene.

Recent discoveries are transforming the landscape of molecular genetics of CHD and diluting the aforementioned antediluvian impression. The pioneering work of Christine and Jon Seidman and colleague in late 1990s led to identification of loss-of-function mutations in *TBX5* and *NKX2-5* as causes of Holt-Oran syndrome and ASD, respectively [12, 13]. An intriguing finding was the assortment of clinical phenotypes in the mutation carriers of *NKX2-5*, ranging from ASD to TOF and hypoplastic left heart syndrome, often in conjunction with conduction defects, among others [13]. Over the course of the next several years, small-scale studies led to the identification of about three-dozen genes, encoding transcription factors, cell signaling molecules and structural proteins in patients with CHD (for review, please see Table 3 in [3]). Moreover, a genome-wide association study comprised of 1,995 cases with a variety of CHD and 5,159 controls was conducted with the goal of identifying susceptibility loci for ASD [14]. Despite these discoveries the genetic causes of

Address for Correspondence: AJ Marian, M.D. Center for Cardiovascular Genetics 6770 Bertner Street Suite C900A Houston, TX 77030 713 500 2350 Ali.J.Marian@uth.tmc.edu.

CONFLICT OF INTEREST: None

CHD in about 80% of the patients had remained unknown [3, 9]. The convergence of four sets of advances, indicated below, is gradually changing the landscape and accelerating the pace of new discoveries of genetic causes of CHD:

1. Initial genetic discoveries by pointing to mutations in cardiac transcription factors as causes of CHD provided the framework for subsequent genetic studies [12, 13, 15]
2. Delineation of the regulatory genetic networks involved in cardiac development offered a biological context for the genetic discoveries [16-21]
3. Availability of large repositories of patients with CHD, such as the Pediatric Cardiac Genomics Consortium (PCGC), afforded large-scale genetic studies [22-24]
4. Advent of newer genetic technologies enabled defining the transcriptional state of the genome as well as genome-wide sequence variations, including single nucleotide variants (SNVs) and copy number variants (CNVs) [25-30].

Among the first fruits of these advances was a whole exome sequencing (WES) project that involved 362 parent-offspring trios with severe CHD in the PCGC population [31]. It led to the identification of premature truncation, frameshift and splice site *de novo* mutations in 28 gene encoding histone modifying proteins [31]. The *de novo* variants collectively contributed to about 10% of severe cases of CHD in the PCGC population. The findings not only broadened the spectrum of the disease-causing mechanisms to include the epigenetic machinery, but also advocated for the gene dosage mechanism, both increased as well as reduced gene expression, in the pathogenesis of CHD.

In a recent issue of *Circulation Research*, Glessner et al. report another large-scale whole-genome study designed to delineate the causal role of CNVs in CHD in the PCGC population [32]. CNVs are structural genomic variations, typically larger than 1,000 base pairs that through duplication or deletion lead to gain or loss of chromosomal segments that often contain multiple contiguous genes. The results of the study by Glessner et al. are notable for a four-fold increase in the frequency of all *de novo* CNVs and a two-fold increase in the frequency of novel *de novo* CNVs in trios with CHD, as compared to controls [32]. Approximately 10% of the CHD population had rare *de novo* CNVs, resulting from deletions or duplication, the former being modestly more common. Combining the sequencing data and CNVs, the authors identified *ETS1*, encoding the ETS1 transcription factor [33], and *CTBP2*, which codes for a transcriptional co-repressor [34], as the likely pathogenic genes affected in the 11q24.2-q25 (Jacobsen syndrome) and 10q sub-telomeric deletions, respectively [32]. The findings are in accord with the prevailing gene dosage mechanism and the pathogenic role of CNVs, particularly *de novo* CNVs, in CHD [3, 9, 35, 36].

The study by Glessner et al. benefits from a robust family-based trio study design comprised of probands who did not have the known cytogenetic anomalies and pathogenic CNVs. It also utilizes state-of-the-art technology that includes detection of CNVs by two independent and complementary methods of SNP arrays and WES in a subset of 233 trios, and validation of the CNVs by digital droplet PCR (ddPCR). The use of two independent CNVs detection platforms also afforded the opportunity to compare detection sensitivity of each platform, which was calculated to be about 65 to 70% (30-35% false-negative rate), upon the

condition of ≥10 adjacent SNPs for calling CNVs by the SNP arrays and involvement of 3 or more adjacent exons in the WES approach. This finding suggests considerable under-detection of the CNVs, if only one of the detection methods (SNP arrays and WES) is used. It also has direct implications in the design of future studies.

The findings of the study by Glessner et al. and the existing data identify CNVs as important causes of CHD and imply that CNVs are likely to contribute to a larger fraction of CHD that has been demonstrated so far. Several CNVs, identified by Glessner et al. impacted only a single gene, hence, rendering them potentially causal genes. However, the pathogenic role of the individual CNVs identified in the study by Glessner et al, with the exception of a few whose causality in CHD already has been established, such as CNVs impacting *NKX2-5* and *GATA4*, cannot be ascertained and will require additional experimentation. The majority of identified CNVs affected large chromosomal segments involving up to several million base pairs of DNA and multiple contiguous genes, rendering the identification of the specific causal gene more tedious. It also merits noting that singleton *de novo* CNVs identified by Glessnert et al. as well as those identified previously cannot be considered causal pending replication of the findings in independent populations and validation through experimentation, as clearly stated by the authors.

The collective results of two large-scale studies on the PCGC population support the gene dosage hypothesis in the pathogenesis of a subset of CHD, as mutations in genes encoding histone modifying proteins as well as the CNVs, are expected to change expression levels of the affected proteins. As for the pleiotropic phenotypic expression of CHD that also includes the extra-cardiac phenotypes, one could speculate several plausible explanations that might operate in isolation or cooperatively as follows:

1. Altered expression levels of multiple proteins resulting from CNVs impacting multiple contiguous genes
2. Altered regulatory elements in the non-protein coding regions, including non-coding RNAs and enhancers, impacted by the CNVs
3. Altered expression of multiple gene targets of the mutant transcription factors
4. Key position of the mutant protein in the regulatory networks, resulting in altered biological functions of multiple target proteins
5. Multiple hit (digenic or mulitgenic) hypothesis, with one mutation serving as the variant with the largest effect size (causal) and multiple others exerting a gradient of effect sizes (modifiers)
6. Concomitant epigenetic variants, modified partly by the environmental factors, exerting additional changes (modifiers) in the background of the main causal mutation. Recent discoveries have hastened elucidation of the molecular genetic basis of CHD.

Despite these advances, however, genetic etiology of CHD has remained inadequately defined, eloquently described by the leading scientists in the field as “the glass half empty” [3] (Figure 1). To shift the paradigm, elaborate studies would be required to further define the genetic etiology of CHD, elucidate the underpinning mechanisms and delineate the

molecular basis of phenotypic plasticity. Perhaps, not so naïvely, many still consider elucidation of the molecular genetic basis of CHD as the last frontier in human cardiovascular genetics.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The author wishes to acknowledge of funding support from NIH/NHLBI (R01 HL088498 and R34 HL105563), Leducq Foundation, Roderick MacDonald Foundation (13RDM005), TexGen Fund from Greater Houston Community Foundation and George and Mary Josephine Hamman Foundation.

REFERENCES

1. Oyen N, Poulsen G, Boyd HA, Wohlfahrt J, Jensen PK, Melbye M. Recurrence of congenital heart defects in families. Circulation. 2009; 120:295–301. [PubMed: 19597048]
2. Whittemore R, Hobbins JC, Engle MA. Pregnancy and its outcome in women with and without surgical treatment of congenital heart disease. The American Journal of Cardiology. 1982; 50:641–651. [PubMed: 7113941]
3. Fahed AC, Gelb BD, Seidman JG, Seidman CE. Genetics of congenital heart disease: The glass half empty. Circulation Research. 2013; 112:707–720. [PubMed: 23410880]
4. Seidman CE, Seidman JG. Identifying sarcomere gene mutations in hypertrophic cardiomyopathy: A personal history. Circulation Research. 2011; 108:743–750. [PubMed: 21415408]
5. Cahill TJ, Ashrafian H, Watkins H. Genetic cardiomyopathies causing heart failure. Circulation Research. 2013; 113:660–675. [PubMed: 23989711]
6. George AL Jr. Molecular and genetic basis of sudden cardiac death. The Journal Of Clinical Investigation. 2013; 123:75–83. [PubMed: 23281413]
7. Cerrone M, Napolitano C, Priori SG. Genetics of ion-channel disorders. Current Opinion in Cardiology. 2012; 27:242–252. [PubMed: 22450718]
8. Moore JR, Leinwand L, Warshaw DM. Understanding cardiomyopathy phenotypes based on the functional impact of mutations in the myosin motor. Circulation Research. 2012; 111:375–385. [PubMed: 22821910]
9. Gelb BD, Chung WK. Complex genetics and the etiology of human congenital heart disease. Cold Spring Harb Perspect Med. 2014; 4:a013953. [PubMed: 24985128]
10. French VM, van de Laar IM, Wessels MW, Rohe C, Roos-Hesselink JW, Wang G, Frohn-Mulder IM, Severijnen LA, de Graaf BM, Schot R, Breedveld G, Mientjes E, van Tienhoven M, Jadot E, Jiang Z, Verkerk A, Swagemakers S, Venselaar H, Rahimi Z, Najmabadi H, Meijers-Heijboer H, de Graaff E, Helbing WA, Willemsen R, Devriendt K, Belmont JW, Oostra BA, Amack JD, Bertoli-Avella AM. Nphp4 variants are associated with pleiotropic heart malformations. Circulation Research. 2012; 110:1564–1574. [PubMed: 22550138]
11. D'Alessandro LC, Latney BC, Paluru PC, Goldmuntz E. The phenotypic spectrum of zic3 mutations includes isolated d-transposition of the great arteries and double outlet right ventricle. American journal of medical genetics. Part A. 2013; 161A:792–802. [PubMed: 23427188]
12. Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soults J, Grayzel D, Kroumpouzou E, Traill TA, Leblanc-Straceski J, Renault B, Kucherlapati R, Seidman JG, Seidman CE. Mutations in human tbx5 cause limb and cardiac malformation in holt-oram syndrome. Nat.Genet. 1997; 15:30–35. [PubMed: 8988165]
13. Schott JJ, Benson DW, Basson CT, Pease W, Silberbach GM, Moak JP, Maron BJ, Seidman CE, Seidman JG. Congenital heart disease caused by mutations in the transcription factor nkx2-5. Science. 1998; 281:108–111. [PubMed: 9651244]

14. Cordell HJ, Bentham J, Topf A, Zelenika D, Heath S, Mamasoula C, Cosgrove C, Blue G, Granados-Riveron J, Setchfield K, Thornborough C, Breckpot J, Soemedi R, Martin R, Rahman TJ, Hall D, van Engelen K, Moorman AF, Zwinderman AH, Barnett P, Koopmann TT, Adriaens ME, Varro A, George AL Jr. dos Remedios C, Bishopric NH, Bezzina CR, O'Sullivan J, Gewillig M, Bu'Lock FA, Winlaw D, Bhattacharya S, Devriendt K, Brook JD, Mulder BJ, Mital S, Postma AV, Lathrop GM, Farrall M, Goodship JA, Keavney BD. Genome-wide association study of multiple congenital heart disease phenotypes identifies a susceptibility locus for atrial septal defect at chromosome 4p16. Nature Genetics. 2013; 45:822–824. [PubMed: 23708191]

15. Li QY, Newbury-Ecob RA, Terrett JA, Wilson DI, Curtis AR, Yi CH, Gebuhr T, Bullen PJ, Robson SC, Strachan T, Bonnet D, Lyonnet S, Young ID, Raeburn JA, Buckler AJ, Law DJ, Brook JD. Holt-oram syndrome is caused by mutations in tbx5, a member of the brachyury (t) gene family. Nature Genetics. 1997; 15:21–29. [PubMed: 8988164]

16. Komuro I, Izumo S. Csx: A murine homeobox-containing gene specifically expressed in the developing heart. Proceedings Of The National Academy Of Sciences Of The United States Of America. 1993; 90:8145–8149. [PubMed: 7690144]

17. de la Pompa JL, Timmerman LA, Takimoto H, Yoshida H, Elia AJ, Samper E, Potter J, Wakeham A, Marengere L, Langille BL, Crabtree GR, Mak TW. Role of the nf-atc transcription factor in morphogenesis of cardiac valves and septum. Nature. 1998; 392:182–186. [PubMed: 9515963]

18. Horb ME, Thomsen GH. Tbx5 is essential for heart development. Development. 1999; 126:1739–1751. [PubMed: 10079235]

19. Kim KH, Rosen A, Bruneau BG, Hui CC, Backx PH. Iroquois homeodomain transcription factors in heart development and function. Circulation Research. 2012; 110:1513–1524. [PubMed: 22628575]

20. Mommersteeg MT, Andrews WD, Ypsilanti AR, Zelina P, Yeh ML, Norden J, Kispert A, Chedotal A, Christoffels VM, Parnavelas JG. Slit-roundabout signaling regulates the development of the cardiac systemic venous return and pericardium. Circulation Research. 2013; 112:465–475. [PubMed: 23255421]

21. Risebro CA, Petchey LK, Smart N, Gomes J, Clark J, Vieira JM, Yanni J, Dobrzynski H, Davidson S, Zuberi Z, Tinker A, Shui B, Tallini YI, Kotlikoff MI, Miquerol L, Schwartz RJ, Riley PR. Epistatic rescue of nkx2.5 adult cardiac conduction disease phenotypes by prospero-related homeobox protein 1 and hdac3. Circulation Research. 2012; 111:e19–31. [PubMed: 22647876]

22. Gelb B, Brueckner M, Chung W, Goldmuntz E, Kaltman J, Kaski JP, Kim R, Kline J, Mercer-Rosa L, Porter G, Roberts A, Rosenberg E, Seiden H, Seidman C, Sleeper L, Tennstedt S, Schramm C, Burns K, Pearson G. The congenital heart disease genetic network study: Rationale, design, and early results. Circulation Research. 2013; 112:698–706. [PubMed: 23410879]

23. van der Velde ET, Vriend JW, Mannens MM, Uiterwaal CS, Brand R, Mulder BJ. Concor, an initiative towards a national registry and DNA-bank of patients with congenital heart disease in the netherlands: Rationale, design, and first results. Eur J Epidemiol. 2005; 20:549–557. [PubMed: 16121765]

24. Escot S, Blavet C, Hartle S, Duband JL, Fournier-Thibault C. Misregulation of sdf1-cxcr4 signaling impairs early cardiac neural crest cell migration leading to conotruncal defects. Circulation Research. 2013; 113:505–516. [PubMed: 23838132]

25. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, Macdonald JR, Pang AW, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, Venter JC. The diploid genome sequence of an individual human. PLoS.Biol. 2007; 5:e254. [PubMed: 17803354]

26. Wang J, Wang W, Li R, Li Y, Tian G, Goodman L, Fan W, Zhang J, Li J, Guo Y, Feng B, Li H, Lu Y, Fang X, Liang H, Du Z, Li D, Zhao Y, Hu Y, Yang Z, Zheng H, Hellmann I, Inouye M, Pool J, Yi X, Zhao J, Duan J, Zhou Y, Qin J, Ma L, Li G, Zhang G, Yang B, Yu C, Liang F, Li W, Li S, Ni P, Ruan J, Li Q, Zhu H, Liu D, Lu Z, Li N, Guo G, Ye J, Fang L, Hao Q, Chen Q, Liang Y, Su Y, San A, Ping C, Yang S, Chen F, Li L, Zhou K, Ren Y, Yang L, Gao Y, Yang G, Li Z, Feng X, Kristiansen K, Wong GK, Nielsen R, Durbin R, Bolund L, Zhang X, Yang H. The diploid genome sequence of an asian individual. Nature. 2008; 456:60–65. [PubMed: 18987735]

27. Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. Nature. 2010; 467:1061–1073. [PubMed: 20981092]
28. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song Xz, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM. The complete genome of an individual by massively parallel DNA sequencing. Nature. 2008; 452:872–876. [PubMed: 18421352]
29. Itsara A, Cooper GM, Baker C, Girirajan S, Li J, Absher D, Krauss RM, Myers RM, Ridker PM, Chasman DI, Mefford H, Ying P, Nickerson DA, Eichler EE. Population analysis of large copy number variants and hotspots of human genetic disease. Am J Hum Genet. 2009; 84:148–161. [PubMed: 19166990]
30. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nature Genetics. 1998; 20:207–211. [PubMed: 9771718]
31. Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, Romano-Adesman A, Bjornson RD, Breitbart RE, Brown KK, Carriero NJ, Cheung YH, Deanfield J, DePalma S, Fakhro KA, Glessner J, Hakonarson H, Italia MJ, Kaltman JR, Kaski J, Kim R, Kline JK, Lee T, Leipzig J, Lopez A, Mane SM, Mitchell LE, Newburger JW, Parfenov M, Pe'er I, Porter G, Roberts AE, Sachidanandam R, Sanders SJ, Seiden HS, State MW, Subramanian S, Tikhonova IR, Wang W, Warburton D, White PS, Williams IA, Zhao H, Seidman JG, Brueckner M, Chung WK, Gelb BD, Goldmuntz E, Seidman CE, Lifton RP. De novo mutations in histone-modifying genes in congenital heart disease. Nature. 2013; 498:220–223. [PubMed: 23665959]
32. Glessner JT, Bick AG, Ito K, Homsy J, Rodriguez-Murrillo A, Fromer M, Mazaika E, Vardarajan B, Italia MJ, Leipzig J, DePalma S, Golhar R, Sanders SJ, Yamrom B, Ronemus M, Iossifov I, Willsey AJ, State MW, Kaltman JR, White PS, Shen Y, Warburton D, Brueckner M, Seidman C, Goldmuntz E, Gelb BD, Lifton RP, Seidman J, Hakonarson H, Chung WK. Increased frequency of de novo copy number variations in congenital heart disease by integrative analysis of snp array and exome sequence data. Circulation Research. 2014 (insert pages).
33. Findlay VJ, LaRue AC, Turner DP, Watson PM, Watson DK. Understanding the role of ets-mediated gene regulation in complex biological processes. Adv Cancer Res. 2013; 119:1–61. [PubMed: 23870508]
34. Chinnadurai G. Ctbp, an unconventional transcriptional corepressor in development and oncogenesis. Molecular Cell. 2002; 9:213–224. [PubMed: 11864595]
35. Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, Ergul E, Conta JH, Korn JM, McCarroll SA, Gorham JM, Gabriel S, Altshuler DM, Quintanilla-Dieck Mde L, Artunduaga MA, Eavey RD, Plenge RM, Shadick NA, Weinblatt ME, De Jager PL, Hafler DA, Breitbart RE, Seidman JG, Seidman CE. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of fallot. Nature Genetics. 2009; 41:931–935. [PubMed: 19597493]
36. Silversides CK, Lionel AC, Costain G, Merico D, Migita O, Liu B, Yuen T, Rickaby J, Thiruvahindrapuram B, Marshall CR, Scherer SW, Bassett AS. Rare copy number variations in adults with tetralogy of fallot implicate novel risk gene pathways. PLoS genetics. 2012; 8:e1002843. [PubMed: 22912587]

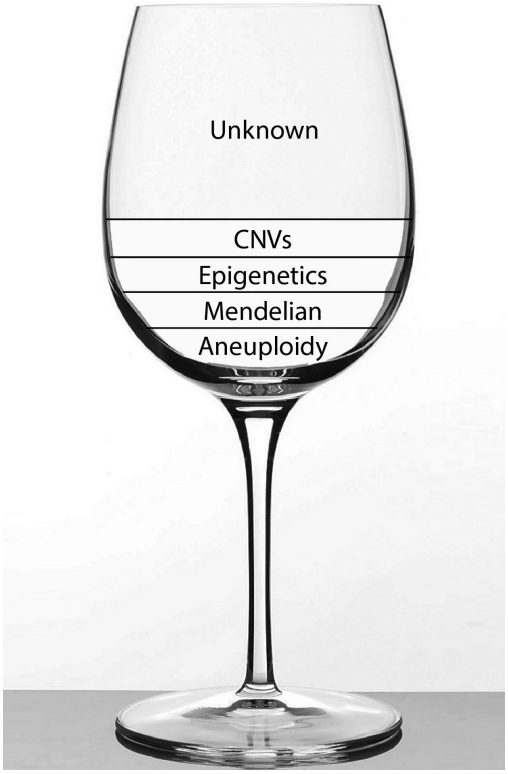

Figure 1. Genetic Etiology of Congenital Heart Disease (CHD): "The glass is half empty"[3]

The known genetic etiology of CHD, namely chromosomal aneuploidy, rare familial CHD with Mendelian patterns of inheritance, mutations affecting proteins involved in histone modifiers (epigenetics) and Copy Number Variants (CNVs) contribute to minority of the CHD cases. The genetic etiology of CHD in the majority of the cases is unknown.