Recent advances in understanding the genetics of congenital heart defects

Bruce D. Gelb, M.D.

Mindich Child Health and Development Institute and Departments of Pediatrics and Genetics and Genomics Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA

Abstract

Purpose of review—To review recent advances in our understanding of the genetic causes of congenital heart disease (CHD).

Recent findings—CHD behaves like a complex genetic trait in most instances. Recent advances in genomics have provided tools for uncovering genetic variants underlying complex traits that are now being applied to study CHD. Massively parallel DNA sequencing has shown that *de novo* mutations contribute to ~10% of severe CHD and implicated chromatin remodeling in pathogenesis. Genome scanning methods for copy number variants continue to identify lesions underlying CHD, some predisposing to it generally and others having lesion specificity. Gene-environment interactions are being explored, primarily related to the metabolism of folate and homocysteine. Finally, studies are addressing other aspects of complexity for CHD such as mutations in cis-regulatory elements and modifying genes.

Summary—The genetic architecture of CHD is being elaborated through the use of state-of-the-art genomic approaches. Through these scientific advances, new opportunities for preventing and ameliorating CHD and its co-morbidities are anticipated.

Keywords

congenital heart disease; copy number variants; exome sequencing; genetics

Introduction

Epidemiologic studies have strongly suggested that congenital heart defects (CHD) arise primarily through genetic abnormalities, but the identification of the precise mutations has been extremely challenging [1]. We have known for some time that a modest percentage of CHD (~5%) is attributable to aneuploidies such as trisomy 21. With the advent of molecular genetic approaches, point mutations with apparently strong effects have been identified in rare families inheriting CHD in Mendelian or near-Mendelian fashion. Unfortunately, the genes identified in this manner have not accounted for substantial percentages of CHD more

Correspondance to: Bruce D. Gelb, M.D., The Mindich Child Health and Development Institute, The Icahn School of Medicine at Mount Sinai, One Gustave Levy Place, Box 1040, New York, NY 10029, Tel: 212-824-8938, Fax: 212-241-3310, bruce.gelb@mssm.edu.

broadly. This review discusses recent advances in the application of the newest genomic technologies to elucidate the genetic causes of CHD.

De Novo Mutations

Through recent advances in molecular genetic technologies, it is now possible to sequence the roughly 1% of the human genome that contains the coding regions for all genes (called the exome), representing approximately 180,000 exons and 30 megabases (Mb), in a relatively rapid and affordable manner. While exome sequencing was initially used to discover mutations underlying Mendelian disorders, current efforts are increasingly focusing on unraveling complex genetic traits.

The Pediatric Cardiac Genomics Consortium (PCGC) [2], a National Heart, Lung, and Blood Institute-funded research enterprise, recently completed a first-of-kind study to determine the role of *de novo* mutations in the etiology of severe forms of CHD [3]. Exome sequencing was performed for 362 parent-offspring trios in which the offspring had a sporadic conotruncal defect, left ventricular outflow track obstructive lesion, or heterotaxy, and compared to comparable data from 264 control trios. While the overall rate of *de novo* point and small insertion/deletion (indel) changes was equivalent between CHD cases and controls, there was an excess burden of protein-altering mutations in genes highly expressed during heart development (odds ratio (OR) of 2.53). Excess mutations had a role in 10% of CHD cases and led to the estimate that ~400 genes underlie these birth defects. After filtering to retain variants most likely to be deleterious (nonsense, splice site and frameshift defects), the burden among CHD cases increased, attaining an OR of 7.50.

Next, the PCGC investigators asked whether the burden of *de novo* protein-altering mutations among the CHD cases preferentially targeted particular biologic processes [3]. Indeed, they observed a highly significant enrichment of mutation among genes encoding proteins relevant for chromatin biology, specifically the production, removal or reading of methylation of Lys4 of histone 3 (H3K4me)(Figure 1). The phenotypes of the eight subjects harboring H3K4me *de novo* mutations was diverse, both with respect to the form of CHD and the involvement of extracardiac tissues. In addition, two independent *de novo* mutations were found in *SMAD2*, which encodes a protein with relevance for demethylation of Lys27 of histone 3 (H3K27me). SMAD2 contributes to the development of the left-right body axis; both subjects harboring *SMAD2* mutations had dextrocardia with unbalanced complete atrioventricular canal defects with pulmonic stenosis. While the contribution of chromatin remodeling to cardiovascular development generally and certain rare genetic syndromes with CHD like Kabuki syndrome had been recognized previously, this study exposed a far broader role in CHD pathogenesis. The finding also suggests a fascinating potential link to other birth defects as *de novo* chromatin remodeling mutations have also been implicated in autism [4].

Copy Number Variants

Copy number variation (CNV), which are gains or losses of DNA ranging in size from 1 kb to several Mbs, affect roughly 10% of the human genome [5]. CNVs are typically detected on a genome-wide basis using SNP microarrays or array comparative genomic

hybridization. Although differentiating pathologic CNVs from benign polymorphic one remains challenging, it has become clear that pathologic CNVs contribute significantly to the pathogenesis of CHD [6].

Two new studies focused on CNVs in individuals with defects of the left ventricular outflow tract [7,8]. In a small study of children with hypoplastic left heart syndrome (HLHS), CNVs, predominantly small lesions, were more common among subjects than controls [8]. The frequency of rare CNVs, however, did not vary between the groups. In a study that included individuals with a broad range of left-sided cardiac lesions with a focus on those from families with more than one affected individual, the overall burden of CNVs was not increased in cases over controls but a number of rare CNVs were identified in cases only, of which none was recurrent within their study [7]. Of interest, these rare CNVs were found to harbor genes relevant for angiogenesis in a significant manner. After filtering, the authors identified 25 putative candidate genes for left-sided heart defects.

A large number of adults with tetralogy of Fallot (TOF) with or without pulmonary atresia but without known genetic abnormalities like 22q11 deletion were studied [9]. Large, rare CNVs (>500 kb) were more prevalent in cases compared to controls (OR 1.9) with most lesions being genomic gains. Especially when the CNVs overlapped with genes, the subjects harboring them tended to have extracardiac abnormalities. Several of the CNVs had been previously implicated in disease including duplications at 1q21.1, which were observed in four patients and were identified in a prior study of CNVs in TOF [10], a 16p11.2 duplication and a 22q11.21 duplication. Like the study of LVOT lesions, pathway analysis of genes in rare CNVs in TOF implicated vascular development.

Bernard Keavney's group focused on 1q21.1 CNVs, particularly for TOF [11]. Gain CNVs were observed in 8/948 cases compared to 3/10,910 controls, an OR of 30.9. Smaller gain CNVs (100-200 kb) at that locus altering *GJA5* were also increased among the TOF cases with an OR of 10.7. Although no 1q21.1 loss CNV was found among their TOF cohort, three such lesions were observed among nearly 1,500 cases of assorted other forms of CHD. That mixed CHD cohort harbored no gain CNV of this region. Taken as a whole, this study suggested that 1q21.1 gain CNVs account for approximately 1% of TOF and implicated altered expression of *GJA5* as the primary cause.

The Keavney group also performed a genome-wide CNV study with more than 2,000 CHD cases including ~800 with TOF and compared those to nearly 900 controls [12]. They found a significant burden of rare loss CNVs >100 kb in size that overlapped genes (OR 1.8; population-attributable risk of nearly 3.5%). There was trend towards greater ORs as progressively larger deletion CNVs were considered, and those CNVs affected more genes than did those found among the controls. Enrichment analysis with the rare loss CNVs revealed a significant association with WNT signaling, which was altered in a broad range of CHD. Examining CNVs that were recurrent, the authors confirmed prior observations of gain and loss CNVs altering *GATA4* at 8p23.1 and found deletion CNVs at 15q11.2 in 0.5% of cases, an OR of 8.2 compared to controls. Finally, the parent of origin for *de novo* CNVs was assessed in 11 cases, of which 10 were paternal.

Restricting the search for CNVs to genes previously implicated in CHD, another study compared more than 800 CHD cases without known genomic defects (e.g., Down syndrome, 22q11del) to 3,000 controls [13]. Large, rare CNVs (>100 kb for losses and >200 kb for gains) were observed in 4.3% of cases compared to 1.8% of controls, a significant difference. The previously noted gain CNVs at 1q21.1 and loss CNVs at 8p23.1 that includes *GATA4* were detected in this study as well. Of interest, two CHD subjects without a clear syndrome were found to harbor gain CNVs at *HRAS*, which is the gene with gain-of-function mutations underlying Costello syndrome.

The role of CNVs in atrioventricular septal defects (AVSDs), both in patients with and without Down syndrome, was investigated [14]. None of the 29 individuals with AVSD without Down syndrome had a pathologic CNV on chromosome 21 and only two harbored rare CNVs elsewhere in the genome, deletions of 1-1.5 Mb that were deemed of uncertain significance. None of the 50 individuals with Down syndrome with AVSD had a CNV of clear significance. The authors concluded that larger rare CNVs altering chromosome 21 do not contribute significantly to the pathogenesis of AVSDs in patients without Down syndrome nor do large CNVs contribute to the AVSDs in Down syndrome.

Methyl Metabolism

The introduction of folate supplementation to women prior to conception, initially intended to reduce the incidence of neural tube defects, has also decreased CHD incidence [15]. Folate serves as a donor for 1-carbon groups for nucleotide synthesis and transfer reactions leading to the methylation of DNA and proteins. This has spurred interest in variants in genes relevant to this methyl metabolism, generally in the homocysteine pathway.

The most commonly studied allele to date is the C667T missense allele of the 5,10-methylenetetrahydrofolate reductase gene (*MTHFR*). The mutated form of MTHFR is thermolabile, resulting in reduced activity, which, in turn, decreases blood folate levels and increases plasma homocysteine concentrations. Prior studies attempting to associate the *MTHFR* C667T allele to CHD risk had produced conflicting results. Two recent meta analyses assessed those previously published studies [16,17]. One meta analysis, which included 20 published studies, detected no risk associated with maternal C667T alleles but did in their offspring [17]. Data from four family-based studies revealed no evidence for transmission distortion of the *MTHFR* C677T allele to offspring with CHD. The second meta analysis reviewed 29 prior studies, combining case-control and family-based ones [16]. This meta analysis showed a significant association between the maternal *MTHFR* C677T allele and offspring CHD risk with a pooled OR of 1.16. In contrast, meta analysis of children with CHD failed to reveal an association with the *MTHFR* C677T allele after adjusting for publication bias. Taken together, the two meta analyses gave conflicting results. As each had its strengths and limitations, one can only conclude that the status of *MTHFR* 6677T allele as a risk for CHD either in the mother or child remains unclear.

Two new studies examined possible roles of common polymorphisms in the methionine synthase reductase (*MTRR*) and cystationine β–synthase (*CBS*) genes in more than 2300 CHD cases and a like number of controls, both groups being Han Chinese [18,19]. For

MTRR, the minor c.56+781 A>C intronic allele engendered a 1.36-fold per allele increase in CHD risk such that individuals with the CC genotype had a 1.8-fold risk [18]. When stratified by CHD classification, the c.56+781 C allele was significantly associated with several types of CHD but most significantly with septation defects. Analysis for single CHD lesions showed particularly strong association with VSDs. Biologically, the c.56+781 C allele was associated with lower *MTRR* steady-state expression levels in cardiac tissues and higher plasma homocysteine levels in persons with the CC genotype compared to those with the reference AA genotype. An *in vitro* assay revealed that the c.56+781 C allele resulted in reduced binding affinity of the transcription factor CEBPα (Figure 2).

For *CBS*, a-4673C>G allele in the promoter reduced the risk of CHD such that child with the CG genotype had a 15% and those with the GG genotype had a 40% reduction in CHD risk compared to children with the reference CC genotype [19]. No epistatic relationship to the *MTRR* c.56+781 C allele was detected. When stratified by CHD classification, the *CBS* – 4673C>G allele was significantly associated with septation and conotruncal defects. Analysis for single CHD lesions showed particularly strong association with VSDs. Biologically, the *CBS* – 4673C>G allele was associated with higher *CBS* steady-state expression levels in cardiac tissues. Unlike the *MTRR* allele, this *CBS* allele did not significantly alter plasma homocysteine levels. An *in vitro* assay revealed that the *CBS*-4673C>G allele resulted in reduced binding affinity of the transcription repressor SP1.

These exciting new findings regarding *MTRR* and *CBS* await exploration in non-Han populations to determine if these CHD risk and protective alleles are similarly prevalent and whether they have comparable effects on CHD risk.

Other Aspects of Complexity

While exome sequencing approaches are beginning to uncover mutations underlying CHD that directly alter proteins qualitatively or quantitatively, it is likely that genetic alterations in promoters, enhancers and even intergenic regions are also relevant for CHD causality. Although ongoing large-scale genomics efforts such as ENCODE are facilitating such work [20], the task of identifying disease-causing variation in non-coding regions of the human genome remains daunting.

A recent study took up this challenge, focusing on the identification of cis-regulatory elements for *TBX5*, the gene mutated in Holt-Oram syndrome [21]. Using a series of constructs with LacZ reporters in transgenic mice, a 700-kb region around *TBX5* was scanned and three enhancer elements residing 380 kb, 140 kb, and 9 kb downstream of *TBX5* that drove cardiac expression were identified (Figure 3). The three enhancer regions were sequenced in 260 unrelated individuals with ASDs, VSDs, or AVSDs. Among the nine rare or novel variants identified, one altering a highly conserved nucleotide was present in homozygosity in a subject with a VSD. When tested *in vitro*, oligonucleotide with the variant bound the relevant transcription factor less avidly. Strikingly, use of the variant sequence with reporters in transgenic mice or zebrafish embryos resulted in dramatically reduced expression in the heart. Thus, the authors provided strong evidence that this single nucleotide variant in a *TBX5* enhancer alters expression of the gene during heart

development. The challenges, well illuminated through this study, are how to prove causality definitively and how to scale such work as large numbers of candidate variants are identified through whole genome sequencing.

Another study attempted to approach the genetic complexity underlying CHD using a systems biology approach with CHD risk genes and responder genes (those with altered expression in animal models of human CHD mutations or teratogenic exposures) [22]. There was statistical support for functional convergence around gene sets forming networks for particular cardiac developmental processes: ventricular septation, atrial septation and outflow tract formation. Evidence for direct convergence, CHD risk factors directly impacting the same genes, was not observed. These results support the idea that there is "a complex pattern of functional interactions between genomic variation and environmental exposures that modulate critical biological systems during heart development." [22]

A study of individuals with Down syndrome with AVSDs or without CHD attempted to uncover genetic factors underlying that phenotypic heterogeneity [23]. Using a candidate gene resequencing approach, the authors observed potentially damaging variants in one of the 26 genes in almost 20% of individuals with Down syndrome with AVSDs but only 3% of those with Down syndrome without CHD. Six genes were specifically implicated: *COL6A1, COL6A2, CRELD1, FBLN2, FRZB,* and *GATA5*. Pathway analysis with these six genes implicated VEGF-A signaling, which was biologically plausible due to VEGF-A's known role in atrioventricular valvuloseptal morphogenesis. The findings in this study provide an initial proof of principal for using individuals with a sensitized genetic background like trisomy 21 that predisposes to CHD to explore additional genetic variants mediating expression of CHD phenotypes.

Conclusion

The genetic architecture of CHD is become clearer, although much remains to be done. We anticipate that continuing advances in genomics, which are rendering studies of CHD more robust and affordable, will facilitate further clarification of CHD genetics. The challenge will be translating discoveries of causative genetic variation underlying CHD into actionable clinical approaches. Improved prognostication, both with respect to the cardiac defects and extra-cardiac co-morbidities, will likely represent the low-hanging fruit. Finding strategies that reduce CHD incidence or alter its natural history will require the elucidation of pathogeneses and a careful balancing of potential benefits with adverse effects if fundamental biological process like chromatin remodeling predominate. Finding the next folate will be an exciting challenge!

- *De novo* mutations account for approximately 10% of severe forms of CHD, disproportionately affecting chromatin remodeling.

- Larger CNVs account for 3-5% of CHD with recurrence of lesions associated with more complex developmental phenotypes.

- Ongoing association studies suggest a link between CHD pathogenesis and mildy perturbed homocysteine metabolism.

Acknowledgments

References and recommended reading

1. Fahed AC, Gelb BD, Seidman JG, Seidman CE. Genetics of congenital heart disease: the glass half empty. Circ Res. 2013; 112:707–720. [PubMed: 23410880]

2. Gelb B, Brueckner M, Chung W, et al. The Congenital Heart Disease Genetic Network Study: rationale, design, and early results. Circ Res. 2013; 112:698–706. [PubMed: 23410879]

3**. Zaidi S, Choi M, Wakimoto H, et al. De novo mutations in histone-modifying genes in congenital heart disease. Nature. 2013 This study describes exome sequencing of probands with CHD and their unaffected parents to discover *de novo* mutations.

4. O'Roak BJ, Vives L, Fu W, et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science. 2012; 338:1619–1622. [PubMed: 23160955]

5. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature. 2006; 444:444–454. [PubMed: 17122850]

6. Southard AE, Edelmann LJ, Gelb BD. Role of copy number variants in structural birth defects. Pediatrics. 2012; 129:755–763. [PubMed: 22430448]

7. Hitz MP, Lemieux-Perreault LP, Marshall C, et al. Rare copy number variants contribute to congenital left-sided heart disease. PLoS Genet. 2012; 8:e1002903. [PubMed: 22969434]

8. Payne AR, Chang SW, Koenig SN, et al. Submicroscopic chromosomal copy number variations identified in children with hypoplastic left heart syndrome. Pediatr Cardiol. 2012; 33:757–763. [PubMed: 22349727]

9. Silversides CK, Lionel AC, Costain G, et al. Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. PLoS Genet. 2012; 8:e1002843. [PubMed: 22912587]

10. Greenway SC, Pereira AC, Lin JC, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. Nat Genet. 2009; 41:931–935. [PubMed: 19597493]

11. Soemedi R, Topf A, Wilson IJ, et al. Phenotype-specific effect of chromosome 1q21.1 rearrangements and GJA5 duplications in 2436 congenital heart disease patients and 6760 controls. Hum Mol Genet. 2012; 21:1513–1520. [PubMed: 22199024]

12*. Soemedi R, Wilson IJ, Bentham J, et al. Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. Am J Hum Genet. 2012; 91:489–501. [PubMed: 22939634] Largest study to date of CNVs in subjects with CHD.

13. Tomita-Mitchell A, Mahnke DK, Struble CA, et al. Human gene copy number spectra analysis in congenital heart malformations. Physiol Genomics. 2012; 44:518–541. [PubMed: 22318994]

14. Priest JR, Girirajan S, Vu TH, et al. Rare copy number variants in isolated sporadic and syndromic atrioventricular septal defects. Am J Med Genet A. 2012; 158A:1279–1284. [PubMed: 22529060]

15. Bailey LB, Berry RJ. Folic acid supplementation and the occurrence of congenital heart defects, orofacial clefts, multiple births, and miscarriage. Am J Clin Nutr. 2005; 81:1213S–1217S. [PubMed: 15883454]

16. Wang W, Wang Y, Gong F, et al. MTHFR C677T polymorphism and risk of congenital heart defects: evidence from 29 case-control and TDT studies. PLoS One. 2013; 8:e58041. [PubMed: 23536781]

17. Yin M, Dong L, Zheng J, et al. Meta analysis of the association between MTHFR C677T polymorphism and the risk of congenital heart defects. Ann Hum Genet. 2012; 76:9–16. [PubMed: 22175539]

18**. Zhao JY, Yang XY, Gong XH, et al. Functional variant in methionine synthase reductase intron-1 significantly increases the risk of congenital heart disease in the Han Chinese population. Circulation. 2012; 125:482–490. [PubMed: 22179537] Association of a novel single nucleotide polymorphism in homocysteine metabolism with risk for CHD.

19. Zhao JY, Yang XY, Shi KH, et al. A functional variant in the cystathionine beta-synthase gene promoter significantly reduces congenital heart disease susceptibility in a Han Chinese population. Cell Res. 2012

20. Dunham I, Kundaje A, Aldred SF, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]

21**. Smemo S, Campos LC, Moskowitz IP, et al. Regulatory variation in a TBX5 enhancer leads to isolated congenital heart disease. Hum Mol Genet. 2012; 21:3255–3263. [PubMed: 22543974] Analysis of cis-regulatory elements of the *TBX5* gene for a role in CHD.

22*. Lage K, Greenway SC, Rosenfeld JA, et al. Genetic and environmental risk factors in congenital heart disease functionally converge in protein networks driving heart development. Proc Natl Acad Sci U S A. 2012; 109:14035–14040. [PubMed: 22904188] Application of systems biology approach to understanding the genetics of CHD.

23. Ackerman C, Locke AE, Feingold E, et al. An excess of deleterious variants in VEGF-A pathway genes in Down-syndrome-associated atrioventricular septal defects. Am J Hum Genet. 2012; 91:646–659. [PubMed: 23040494]
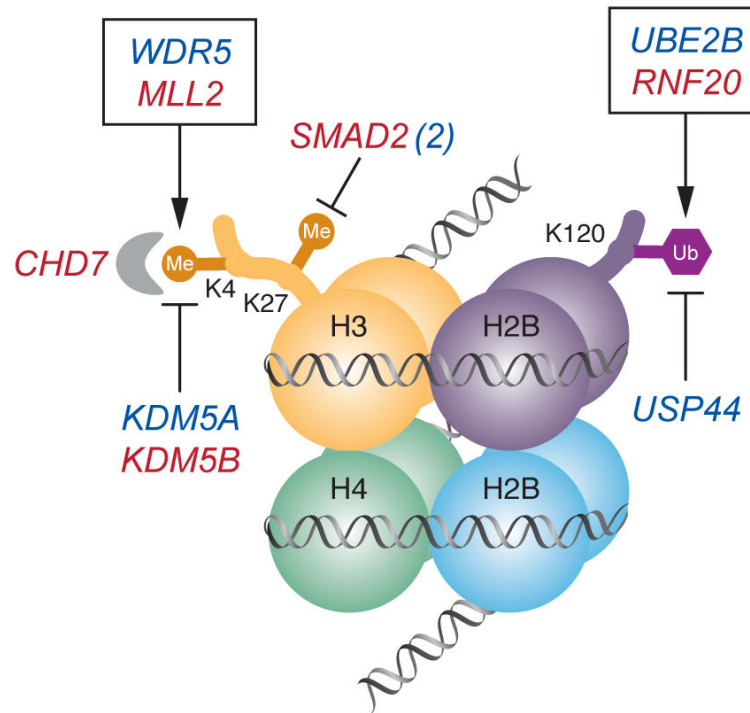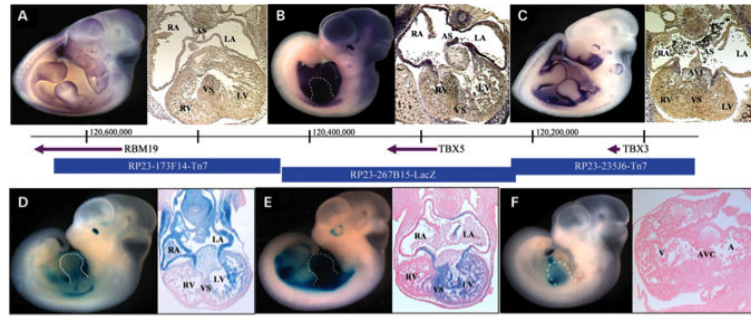
Figure 1.
De novo mutations in the H3K4 and H3K27 methylation pathways. Nucleosome with histone octamer and DNA, with H3K4 methylation bound by CHD7, H3K27 methylation and H2BK120 ubiquitination is shown. Genes mutated in CHD that affect the production, removal and reading of these histone modifications are shown; genes with damaging mutations are shown in red, those with missense mutations are shown in blue. SMAD2 (2) indicates there are two patients with a mutation in this gene. Genes whose products are found together in a complex are enclosed in a box. Reprinted without modification from Ref. 3.
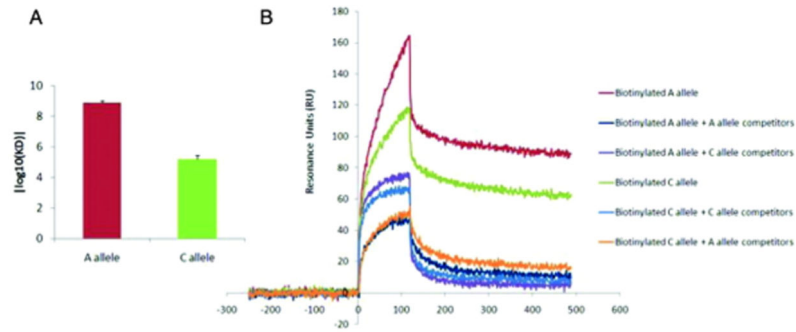
Figure 2.
The *MTRR* c.56_781 C allele attenuates transcription binding affinity. Data shown are mean + SE. **A**, Surface plasmon resonance (SPR) analysis comparing the binding affinity of nuclear extracts to DNA probes containing either the c.56+781 A or C allele. The SPR binding activity of the A allele (log10 (KD) = 8.37 + 0.23) was >1000-fold higher than that of the C allele (log10 (KD) = 5.18 + 0.42). Each value represents the mean +SD of 3 experiments. **B**, Competition SPR assays performed in the presence of 5-fold excess nonbiotinylated A or C allele probes. Reprinted with modification from Ref. 18 with permission from the American Heart Association.

Figure 3.
Regulatory landscape of the *TBX5* locus. Whole mount and histological sections through the heart of embryonic day 11.5 (E11.5) embryos. Top row: *In situ* hybridizations showing endogenous expression (purple) of (**A**) *RBM19*, (**B**) *TBX5* and (**C**) *TBX3*. Bottom row: β-Galactosidase staining (blue) captures the regulatory landscape of enhancers within BACs (**D**) RP23-173F14-Tn7, (**E**) RP23-267B15-LacZ and (**F**) RP23-235J6-Tn7. The genes probed for in A, B and C are contained, respectively, within the BACs tested in D, E and F. Panels B, D-F, forelimb is outlined. A-E, transverse sections; F, sagittal section. A, atrium; RA, right atrium; LA, left atrium; AS, atrial septum; V, ventricle; RV, right ventricle; LV, left ventricle; VS, ventricular septum; AVC, atrioventricular canal. Reprinted from Ref. 21 with permission from Oxford University Press.