1-15-2010

A Search For Genetic Modifiers Responsible For Congenital Heart Disease Variability In The Presence Of Nkx2-5 Haploinsufficiency

Julia Brandeis Winston Honold
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences

Human and Statistical Genetics

Dissertation Examination Committee:
Patrick Jay, Chair
James Cheverud
Robert Mecham
Kelle Moley
David Ornitz
D.C. Rao

A Search For Genetic Modifiers Responsible For Congenital Heart Disease Variability In The Presence Of Nkx2-5 Haploinsufficiency

by

Julia Brandeis Winston

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2010

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

A Search For Genetic Modifiers Responsible For Congenital Heart Disease Variability In The Presence Of Nkx2-5 Haploinsufficiency

By

Julia Brandeis Winston

Doctor of Philosophy in Biology and Biomedical Sciences (Human and Statistical Genetics)

Washington University in St. Louis, 2010

Assistant Professor Patrick Jay, Chairperson

While a clear heritable risk has been observed for congenital heart disease, there is considerable variation in penetrance and presentation likely due to multiple genetic and environmental risk factors. To identify causative factors and interactions responsible for variability in heart development, greater than 4,200 hearts from *Nkx2-5* heterozygous knockout mice have been collected and examined. *Nkx2-5*$^{+/-}$ mice in the inbred strain background C57Bl/6 frequently have atrial and ventricular septal defects. The incidences are substantially reduced in the *Nkx2-5*$^{+/-}$ progeny of first-generation (F1) outcrosses to the strains FVB/N or A/J. Defects recur in the second generation (F2) of the F1xF1 intercross or backcrosses to the parental strains. All 3 strains carry susceptibility alleles at different loci for atrial and ventricular septal defects. Relative to the other 2 strains, A/J carries polymorphisms that confer greater susceptibility to atrial septal defect (ASD) and atrioventricular septal defects and C57Bl/6 to muscular ventricular septal defects (VSD). Genome wide linkage analysis was conducted on 306 mice from the FVB/n F2 intercross and 80 mice from the A/J F2 intercross diagnosed with VSD to map main effect and interacting loci that correlate with risk. Additionally the possibility of environmental interactions was explored in this analysis by including maternal and paternal age and litter size as terms in the

models tested. Significantly linked genomic regions were identified from the FVB/n population on chromosomes 6, 8 and 10 implicating genes in these positions as important candidates for VSD risk. Linkage analysis on the A/J cross identified both shared and unique modifiers from the FVB/N cross scan. Maternal age was found to significantly correlate with VSD risk in FVB/N crosses but not A/J providing evidence for strain specific susceptibility to a non-heritable risk factor. The findings in this study implicate modifier genes as major factors in cardiac developmental pathways by buffering against genetic and environmental insults in the majority while directing the manifestation of disease in a few. Characterization of the genetic architecture of congenital heart disease in a mouse model will provide a deeper understanding of its multifactorial nature and possibly lead to novel strategies for prognosis and prevention.

Acknowledgements

Financial and academic support for this project was generously provided by a Ruth L. Kirschstein National Research Service Award from the Developmental Cardiology and Pulmonary Training Program (National Institutes of Health/National Heart, Lung, and Blood Institute T32 HL007873) and by the Lucille P. Markey Pathway in Human Pathobiology. Additional support from Dr. Jay's lab was provided by the American Heart Association, Edward Mallinckrodt Jr. Foundation, Hartwell Foundation, and March of Dimes.

It is my pleasure to thank my thesis mentor, Dr. Jay, and the members of my thesis committee who supported me with their wisdom and experience as this project progressed. I would also like to acknowledge the Human and Statistical Genetics and Genetic Epidemiology Master of Science programs which introduced me to this area of research and gave me the tools to contribute to it. Finally, I would like to mention my appreciation for the highly collaborative and knowledgeable faculty, staff and fellow graduate students of the human genetics community at Washington University who made my graduate experience a rich and productive one.

I would like to dedicate this thesis to my family, whose support has been constant and invaluable.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction and Background Review

Literature Review

Congenital Heart disease (CHD) is a complex disease occurring with an incidence of 1 in 100 live births and it is the leading cause of defect-related death among newborns in the United States [1,2]. Surveys from diverse locales and periods showed a remarkably consistent ~0.5% incidence of significant congenital heart disease in newborns. The distribution of defects assumes a regular pattern with ventricular septal defects (VSDs) being the most common, followed by atrial septal defects (ASDs) and atrioventricular canal defects (AVCs) [3]. It is suspected that this complex of diseases is caused by disruptions at various points along the cardiac developmental pathway resulting in a malformed heart. With a strong recurrence in 1st degree relatives, there is a clear genetic component to congenital heart disease risk, but its study is complicated by genetic heterogeneity as well as variable penetrance and expressivity [4,5]. Even when individuals are carrying a common genetic risk factor there is considerable variability in phenotypic outcome[6]. As early as 1968 Nora et al described congenital heart disease as having a multifactorial model of inheritance; or the result of multiple genetic and environmental factors and interactions[7].

The most important advances in this field have come from the discovery of human disease causing mutations in cardiac transcription factor genes such as Tbx5, GATA4 and Nkx2-5[8-10]. But, studies exemplifying syndromes with highly penetrant heart defects have resulted in an emphasis on monogenic causes that highlight the outdated genetic determinism paradigm where one gene is linked to one phenotype. The vast

majority of congenital heart disease occurs sporadically and is not due to obvious mutations in major genes [1,3]. Studies addressing sporadic incidences have identified numerous candidate loci that show correlation with cardiac malformations but mechanistic explanations remain incomplete [1,11]. There lacks a framework that relates the known major genes and candidate loci to each other in the context of functional developmental pathways. A coherent model would require a better understanding of the underlying variability including genes with smaller effects on the cardiac regulatory pathway and potential environmental risk factors that may interact with them.

One potential approach to a better understanding of congenital heart disease is dissecting the evolutionary basis for the heart's inherent complexity. Heart development occurs early in gestation with the first heart beats occurring 21-23 days into human embryonic development and the completion of the four chambers and outflow tract at 7 weeks [12]. The developmental sequence is replicated in birds and mammals and is governed by a highly conserved gene regulatory pathway. The first heart began as a simple tube with contractile functions in our distant aquatic ancestors, however millions of years of gene duplication and co-option of additional genetic networks have resulted in multiple chambers, valves, inflow and outflow tracks and a conduction system as seen in mammals. A series of transcription factors have been identified as playing a major role in morphogenesis but there still lacks a comprehensive model of upstream regulators and downstream activators to complete our understanding [13]. Multiple genetic factors affecting CHD makes sense as this is a vital and complex system that has been evolving for many millions of years. The polygenic risk of CHD may provide a clue to how the heart survives and remains a relatively robust organ system. Perhaps there is considerable

variation in small effect genes so that if a major genetic or environmental disruption to the system occurs, there will be a whole host of options for the system to evolve from[14]. It seems that this system would develop out of millions of years of stabilizing selection to maintain an organ that is robust to change. This presents an important problem in understanding points of failure in development underlying the epidemiology of congenital heart defects because different modifiers may direct disease in different populations.

NKX2-5 Cardiac Transcription Factor

Nkx2-5 is an ancient and well characterized gene that encodes for an NK2-type homeodomain cardiac transcription factor important in heart development. The NK2 homeobox is an ancient regulator of cardiac muscle cell lineages and has been reported to be co-expressed with MEF2, a central muscle gene regulator existing in Cniderians which lack a heart but contain the myoepithelial cells thought to be similar to the evolutionary precursors of myocytes [15]. The Nkx2-5 gene is well conserved because its integral role in heart development is ubiquitous across animal species. Homozygous Nkx2-5 knockout mice form an underdeveloped heart that is non-functional and therefore 100% embryonic lethal. Developing embryos show a heart tube that is only partially looped and has a bulbous atrium and ventricle, a wide, regurgitant atrioventricular canal, and a stenotic outflow tract. The embryo dies at E9.5 [16,17]. However in heterozygous knockout mice, animals normally survive to birth but their risk increases for a wide array of heart defects and thus, implicates other genetic and environmental elements as being important in heart development.

In human babies, mutations within have been linked to AV block, atrial and ventricular septal defects, tetralogy of Fallot, Ebstein's anomaly and other tricuspid valve abnormalities [18]. While it is understood that Nkx2-5 is one of a series of highly conserved transcription factors that make up a core set of regulators in heart development[13], the variation in congenital heart disease cannot be attributed to the heterogeneity of the many mutations described in these core genes alone. Nkx2-5 disruption has resulted in an increased risk for a variety of different cardiac malformations among individuals within the same pedigree indicating that a specific Nkx2-5 mutation does not condemn a person to a specific defect [10].

Mouse Model for Human Disease

The mouse serves as an excellent model for congenital heart disease because the human NKX2-5 haploinsufficient pleiotropy described above can be mimicked in inbred strains using heterozygous knockouts. These mice show an elevated risk for a variety of heart malformations and despite being in a more controlled environment than their human counterparts, mice show proportionately similar incidences of the different types of heart defects with VSDs being the most common followed by ASDs and then atrioventricular canal defects. Mating, living environment and diet can be controlled, and by choosing markers polymorphic between strains, there is no ambiguity of allelic ancestry so that outcrosses are 100% informative at each marker in linkage analysis. Different inbred strains can provide simple homogeneous models for the effects of variable genetic backgrounds seen in different individual human subjects. Human allelic background and

diversity can thus be modeled in mice in the context of genetic risk for human disease [19,20]. Strain-dependent variability of cardiac defects has been described in major heart genes such as Hey2, Tbx1 and Tbx5 [21-23]. This observation implies that while these genes have been identified as causative factors of human disease, genetic background plays an important role in directing the CHD phenotype. It is suspected that innate polymorphisms within modifier loci for the core genes mentioned are fixed in inbred lines resulting in strain specific variability. These loci can potentially affect alternative sub-networks of morphogenesis resulting in a propensity towards defects unique to the line.

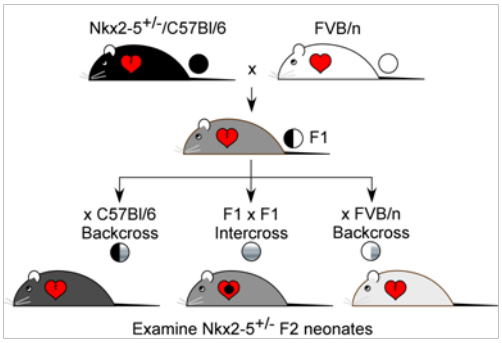
In this study we have crossed Nkx2-5$^{+/-}$ C57Bl/6 mice with wild type FVB/n and A/J to form an F1 hybrid generation which was then intercrossed to create two F2 generations in which each mouse has a unique combination of alleles from the parental lines. By outcrossing isogenic mice, the overall proportion of genomic heterozygosity can be experimentally varied. In general, genetic hybrids show an increased fitness when compared to their 99% homogeneous isogenic parents, a phenomenon called heterosis or hybrid vigor [24]. Heterosis has primarily been studied in the context of agriculture; particularly to investigate the origins of quantitative traits that affect biomass and fertility in plants and livestock [25]. However it has implications in human disease as well since the source of vigor or lowered risk could be the result of evolutionarily beneficial diversity at loci that have yet to be discovered. By elucidating the genetic differences among individuals that direct them towards or away from a specific defect, a more complete model of heart developmental pathways can be established.

BACKGROUND DATA

Homozygous Nkx2-5 knockout mice form an underdeveloped heart that is non-functional and therefore 100% embryonic lethal. Nkx2-$5^{+/-}$ C57Bl/6 isogenic mice are born with ASDs and/or VSDs ~55% of the time, but when these mice are outcrossed to FVB/n mice their offspring show a dramatic drop in incidence (~4%). The reduction in heart defects in the hybrid generation implicates genetic heterozygosity elsewhere in the genome as a major effecter on heart development.

Nkx2-$5^{+/-}$ F1 hybrid show decreased disease incidence

To generate the F2 animals, Nkx2-$5^{+/-}$ C57Bl/6 males are crossed to FVB/N females to generate the F1 hybrid mice. The F1 generation is genotyped through PCR for Nkx2-$5^{+/-}$ knockout and then F1 animals are intercrossed to generate F2 progeny. F2 neonates are collected within hours of birth to prevent cannibalization and in order to keep track of population genotype frequencies. The neonates are euthanized, dissected and genotyped for Nkx2-5. The mouse torso sections are fixed in 10% neutral buffered formalin for 2 nights and then transferred to 70% EtOH. The fixed hearts are dissected out, paraffin-embedded and serially sectioned completely in the frontal plane at 6 μm thickness. The cardiac tissue is stained with hemotoxylin and eosin for examination under a 5x microscope objective. Diagnoses of heart defects are independently validated for each heart.

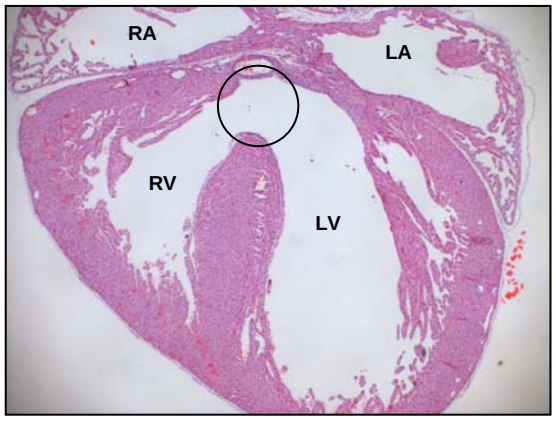


Crosses between the C57Bl/6 and FVB/N strains are shown. Pies depict the fraction of the genome from either strain with gray representing a random distribution in the population.

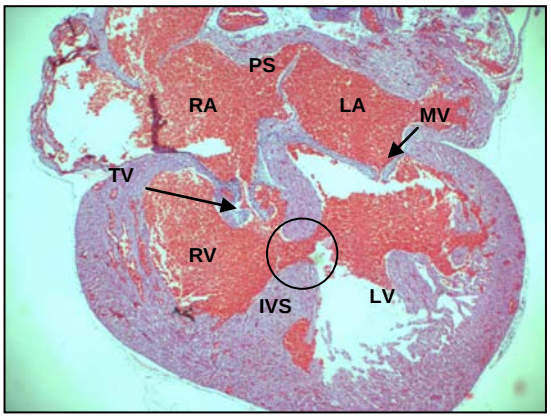
Figure 1. Breeding strategy.

Serial sectioned hearts from wild-type and Nkx2-5$^{+/-}$ neonatal pups in the isogenic C57Bl/6 background revealed a 17% incidence of ventricular septal defect and a 17% incidence of atrial septal defect (Fig. 2) which is significantly different ($P > 0.01$) from the 49 wild-type hearts diagnosed; all of which showed normal heart morphology.

Membranous VSD



Muscular VSD



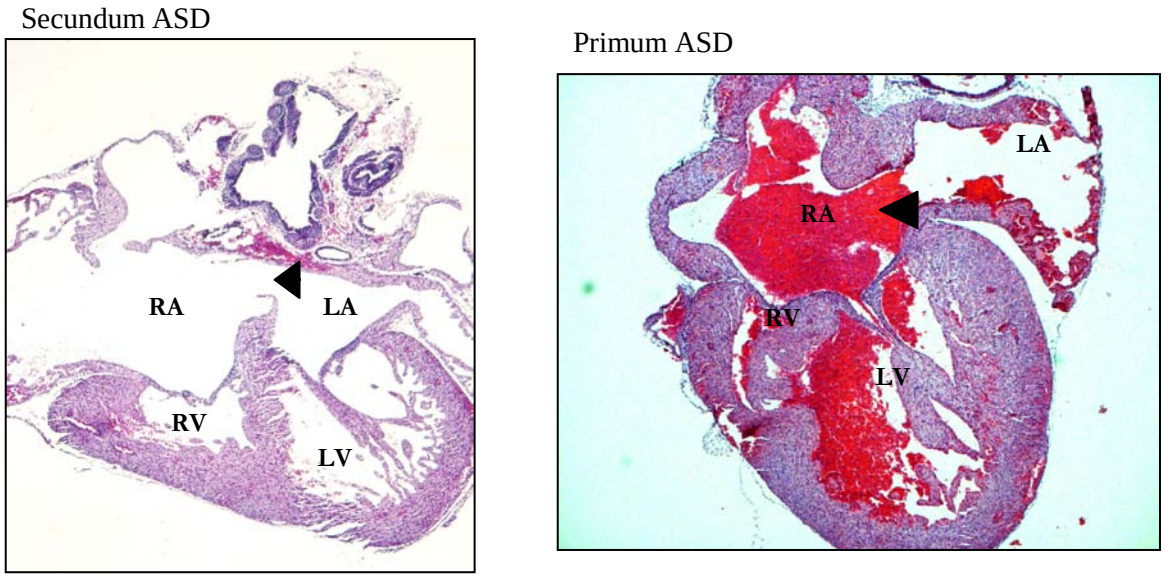


Figure 2. The most common defects observed in hearts with Congenital Heart Disease.

Sectioning and diagnosis of 52 WT and 57 Nkx2-5$^{+/-}$ F1 C57Bl/6-FVB/n newborn hearts showed a significant reduction is defects will all WT being structurally normal and one ASD and a small VSD detected in Nkx2-5$^{+/-}$ neonate hearts (Table 1). The reduction of ASD and VSD incidences in Nkx2-5$^{+/-}$ F1 progeny compared to isogenic Nkx2-5$^{+/-}$ C57Bl/6 mice was found to be significant ($P > 0.01$) for this outcross.

Because different inbred strains are likely fixed for their own distinct combinations of polymorphisms which can result in strain specific effects on defect incidence and presentation [26] another outcross was performed using the WT A/J mice. In the C57Bl/6-A/J outcross, no VSDs were observed in hybrid WT or Nkx2-5$^{+/-}$ hearts resulting in a significant reduction of VSD incidence compared to parental Nkx2-5$^{+/-}$

mice. Interestingly, ASD incidence was found not to be significantly affected by hybridization since 4 ASDs were diagnosed out of the 49 F1 C57Bl/6-A/J Nkx2-5$^{+/-}$ mice examined. While more complex mechanisms may underlie this finding, it is likely that one or more ASD susceptibility loci are not polymorphic between C57Bl/6 and A/J and thus risk reduction through increased heterozygosity is not observed as in the C57Bl/6-FVB/n hybrids. Another explanation may be the presence of dominant A/J polymorphism(s) conferring susceptibility in the heterozygous state as well. The F1 generational loss of VSDs in the presence of Nkx2-5$^{+/-}$ denotes a recovery from inbreeding depression and/or the introduction of hybrid vigor. This change in defect incidence through a reduction of homozygosity is indicative of a reliance of the cardiac phenotype on genetic variation outside of the Nkx2-5 locus.

F2 intercross show a recovery of defects

Given the observation that F1 hybrids showed a risk reduction for CHD in the presence of Nkx2-5$^{+/-}$, it was predicted that subsequent backcrosses to parental strains and intercrosses would increase the incidence of CHD since the F2 crosses would result in a proportion of loci reverting back to homozygosity. If recessive CHD susceptibility loci reestablish a homozygous state, disease risk should be recovered.

Mice from the Nkx2-5$^{+/-}$ C57Bl/6-FVB/n and C57Bl/6-A/J F1 hybrid generations were backcrossed to C57Bl/6 and separately intercrossed (C57Bl/6-FVB/n F1 x C57Bl/6-FVB/n F1, C57Bl/6-A/J F1 x C57Bl/6-A/J F1) to form F2 populations from which hearts were collected, sectioned and diagnosed.

Ventricular septal defects in the F2 C57Bl/6-FVB/n population remained significantly higher (P = 0.0053) in Nkx2-5$^{+/-}$ mice than in the WT mice (Fig. 3). Nkx2-5$^{+/-}$ atrial septal defect incidence did not reach significance compared to WT F2 progeny (P = 0.1199) however, this is likely due to ASDs being more rare and thus lowering the number of affected animals so that stochastic events have more influence. Intercross progeny showed a recovery of defects with a VSD incidence of 15.1% and an ASD incidence of 6.0% out of the 974 Nkx2-5$^{+/-}$ pups collected and diagnosed. VSD incidence was significantly increased (P = 0.0004) from the previous F1 hybrid generation. Since the status of the Nkx2-5 heterozygous genotype remained constant, an increase in defects from the F1 to F2 generations supports the existence of one or more genomic loci that modify the Nkx2-5$^{+/-}$ risk for heart defects.

F2 A/J intercross VSD incidence was significantly elevated from the previous F1 generation resulting in 10.8% having VSDs out of 465 diagnosed hearts (P = 0.0089). The 6.7% ASD incidence observed was not appreciably different than the F1 incidence lending support to the hypothesis that ASD susceptibility loci may not be polymorphic between C57Bl/6 and A/J or that there are dominant A/J risk allele(s) that confer similar risk in Aa and AA form. C57Bl/6 backcross mice showed a modest increase in VSD incidence of 10.9% while the ASD incidence of 7.9% was consistent with the corresponding intercross data as it was not dissimilar from the F1 generation. F2 A/J Nkx2-5 heterozygous mice showed significantly higher incidences for VSDs and ASDs than their WT counterparts (P = 0.0001 and P = 0.0012 respectively).

Conclusion

It is the goal of this study to characterize the genetic variation that is typical of congenital heart disease. An inbred C57Bl/6 mouse that has a heterozygous knockout for the cardiac transcription factor Nkx2-5 will serve as a model for risk. Controlled breeding to other inbred strains will allow for genetic variability to be re-introduced so that we may study its effects on the phenotype. Through observation, segregation, linkage and correlation analyses, this model will be systematically tested to determine the effects of yet unknown sources of genetic variation. The new information gathered from this study can then be applied to our current understanding of the heart developmental pathway so that a more accurate model of congenital heart disease may then be applied to improve standard of care for the afflicted population.

Reference List

(1) Gruber PJ, Epstein JA. Development gone awry: congenital heart disease. *Circ Res.* 2004;94:273-283.

(2) Hoffman JI, Kaplan S. The incidence of congenital heart disease. *J Am Coll Cardiol.* 2002;39:1890-1900.

(3) Ferencz C, Rubin JD, McCarter RJ et al. Congenital heart disease: prevalence at livebirth. The Baltimore-Washington Infant Study. *Am J Epidemiol.* 1985;121:31-36.

(4) Benson DW. Genetic origins of pediatric heart disease. *Pediatr Cardiol.* 2010;31:422-429.

(5) Oyen N, Poulsen G, Boyd HA, Wohlfahrt J, Jensen PK, Melbye M. Recurrence of congenital heart defects in families. *Circulation.* 2009;120:295-301.

(6) Nadeau JH, Topol EJ. The genetics of health. *Nat Genet.* 2006;38:1095-1098.

(7) Nora JJ. Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. *Circulation.* 1968;38:604-617.

(8) Basson CT, Bachinsky DR, Lin RC et al. Mutations in human TBX5 cause limb and cardiac malformation in Holt-Oram syndrome. *Nat Genet.* 1997;15:30-35.

(9) Li QY, Newbury-Ecob RA, Terrett JA et al. Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nat Genet.* 1997;15:21-29.

(10) Schott JJ, Benson DW, Basson CT et al. Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science.* 1998;281:108-111.

(11) Kirk EP, Hyun C, Thomson PC et al. Quantitative trait loci modifying cardiac atrial septal morphology and risk of patent foramen ovale in the mouse. *Circ Res.* 2006;98:651-658.

(12) Andelfinger G. Genetic factors in congenital heart malformation. *Clin Genet.* 2008;73:516-527.

(13) Olson EN. Gene regulatory networks in the evolution and development of the heart. *Science.* 2006;313:1922-1927.

(14) Rutherford SL, Lindquist S. Hsp90 as a capacitor for morphological evolution. *Nature.* 1998;396:336-342.

(15) Grens A, Gee L, Fisher DA, Bode HR. CnNK-2,an NK-2 Homeobox Gene, Has a Role in Patterning the Basal End of the Axis in Hydra. *Developmental Biology.* 1996;180:473-488.

(16) Lyons I, Parsons LM, Hartley L et al. Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev.* 1995;9:1654-1666.

(17) Tanaka M, Chen Z, Bartunkova S, Yamasaki N, Izumo S. The cardiac homeobox gene Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. *Development.* 1999;126:1269-1280.

(18) Benson DW, Silberbach GM, Kavanaugh-McHugh A et al. Mutations in the cardiac transcription factor NKX2.5 affect diverse cardiac developmental pathways. *J Clin Invest.* 1999;104:1567-1573.

(19) Campbell IG, Russell SE, Choong DY et al. Mutation of the PIK3CA gene in ovarian and breast cancer. *Cancer Res.* 2004;64:7678-7681.

(20) Campbell H, Carothers AD, Rudan I et al. Effects of genome-wide heterozygosity on a range of biomedically relevant human quantitative traits. *Hum Mol Genet.* 2007;16:233-241.

(21) Sakata Y, Koibuchi N, Xiang F, Youngblood JM, Kamei CN, Chin MT. The spectrum of cardiovascular anomalies in CHF1/Hey2 deficient mice reveals

roles in endocardial cushion, myocardial and vascular maturation. *J Mol Cell Cardiol.* 2006;40:267-273.

(22) Bruneau BG, Nemer G, Schmitt JP et al. A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell.* 2001;106:709-721.

(23) Taddei I, Morishima M, Huynh T, Lindsay EA. Genetic factors are major determinants of phenotypic variability in a mouse model of the DiGeorge/del22q11 syndromes. *Proc Natl Acad Sci U S A.* 2001;98:11428-11431.

(24) Lippman ZB, Zamir D. Heterosis: revisiting the magic. *Trends Genet.* 2007;23:60-66.

(25) Birchler JA, Yao H, Chudalayandi S. Unraveling the genetic basis of hybrid vigor. *Proc Natl Acad Sci U S A.* 2006;103:12957-12958.

(26) Winston JB, Erlich JM, Green CA et al. Heterogeneity of Genetic Modifiers Ensures Normal Cardiac Development. *Circulation.* 2010.

Chapter 2: Heterogeneity of Genetic Modifiers Ensures Normal Cardiac Development

Published in Circulation, March 2010

Abstract

Mutations of the transcription factor *Nkx2-5* cause pleiotropic heart defects with incomplete penetrance. This variability suggests that additional factors can affect or prevent the mutant phenotype. We assess here the role of genetic modifiers and their interactions. Heterozygous *Nkx2-5* knockout mice in the inbred strain background C57Bl/6 frequently have atrial and ventricular septal defects. The incidences are substantially reduced in the *Nkx2-5*$^{+/-}$ progeny of first-generation (F1) outcrosses to the strains FVB/N or A/J. Defects recur in the second generation (F2) of the F1xF1 intercross or backcrosses to the parental strains. Analysis of greater than 3000 *Nkx2-5*$^{+/-}$ hearts from 5 F2 crosses demonstrates the profound influence of genetic modifiers on disease presentation. On the basis of their incidences and coincidences, anatomically distinct malformations have shared and unique modifiers. All 3 strains carry susceptibility alleles at different loci for atrial and ventricular septal defects. Relative to the other 2 strains, A/J carries polymorphisms that confer greater susceptibility to atrial septal defect and atrioventricular septal defects and C57Bl/6 to muscular ventricular septal defects. Segregation analyses reveal that at least 2 loci influence membranous ventricular septal defect susceptibility, whereas at least 2 loci and at least 1 epistatic interaction affect muscular ventricular and atrial septal defects. Alleles of modifier genes can either buffer perturbations on cardiac development or direct the manifestation of a defect. In a genetically heterogeneous population, the predominant effect of modifier genes is health.

Introduction

Significant progress has been made toward defining genetic etiologies in congenital heart disease.[1,2] Current knowledge of the few dozen genes implicated in cardiac development does not explain, however, the basis of common epidemiologic and clinical observations. The incidence of heart defects in newborns is 0.5-1%, making it a leading cause of death in children. Still, the vast majority is born with a normal heart. Selection against fetuses that have cardiac defects does not cause the prevalence of the norm.[3,4] Insults on the embryonic heart may be rare or weak, or mechanisms may have evolved to ensure the robustness of development. Direct evidence for either hypothesis is minimal, but the latter can better explain the breadth of observations related to incomplete penetrance and phenotypic variability.[5,6]

Although family studies led to the discovery of mutations of prototypical cardiac developmental genes such as *TBX5*, *NKX2-5*, and *GATA4* [7-10], most cases of congenital heart disease have no identified cause or association. Such sporadic cases may be due to an unknown teratogen, spontaneous mutation or inconsistent expression of an inherited etiology. The last possibility is supported by the increased incidence of congenital heart disease in subsequent first-degree relatives, 2-11%, compared to the general population.[11-13] Some normal relatives of sporadic patients must somehow escape the manifestation of a genetic mutation. Furthermore, among the affected relatives the cardiac phenotypes are usually dissimilar and of wide-ranging severity.[11,13] Modifying factors likely influence the incomplete penetrance and phenotypic variability associated with the expression of a causative mutation. In a family, the activity of modifiers in

pathways involved in the development of an affected cardiac structure could determine whether a particular relative develops a heart defect, its type and severity. The unknown factors could be genetic, environmental or entirely stochastic. Of these, genetic modifiers are most amenable to investigation in a laboratory where environmental factors can be controlled between experimental groups.

Small studies have demonstrated the effect of the background strain on cardiac phenotypes related to embryonic viability and myocardial, valvar, or ventricular septal morphology in *Tbx5*, *Gata4* and *Hey2* mutations.[14-16] The extent to which genetic modifiers contribute to incomplete penetrance and phenotypic variability as observed in humans might only be appreciated, however, in a much larger study. In this regard, the diverse anatomic phenotypes described in association with *NKX2-5* mutation offer a foothold into the genetic modifiers that influence the susceptibility of developmental pathways. *NKX2-5* mutations were first reported in families with highly penetrant ASDs and atrioventricular block, but not every member of these and subsequent families who carried a mutation had an ASD or even a cardiac malformation. Among affected persons the various defects seem neither related to each other by any known developmental pathway nor correlated with *NKX2-5* genotype.[9,17,18] Thus, while mutations of *NKX2-5* or other cardiac transcription factors like *TBX5* and *GATA4* undoubtedly cause heart defects, modifying factors could have effects on phenotype as great as the causative mutation.

As the first step toward the elucidation of these factors, we examined the effect of genetic background variation on the incidence of specific heart defects in heterozygous

Nkx2-5 knockout animals in the inbred mouse strain C57Bl/6 and in crosses to the strains FVB/N and A/J. The survey of thousands of *Nkx2-5*$^{+/-}$ animals reveals the important role that genetic modifiers play in buffering cardiac developmental pathways against a major perturbation in a population while directing the manifestation of disease in a few. The results intertwine the genetic basis of health and congenital heart disease, providing a conceptual framework to understand common clinical observations.

Methods

Mouse Strains and Crosses

Nkx2-5$^{+/-}$ animals were backcrossed to C57Bl/6 for 10-13 generations. The knockout allele was generated and genotyped as described.[19] *Nkx2-5*$^{+/-}$ C57Bl/6 males were crossed to FVB/N or A/J females to generate F1 progeny. C57Bl/6 and FVB/N animals were obtained from Charles River Laboratories, and A/J from the Jackson Laboratory. *Nkx2-5*$^{+/-}$ F1 animals were intercrossed or backcrossed to the parental strains C57Bl/6 or FVB/N to generate F2 progeny. The backcross to A/J was not performed. Animals were housed under standard conditions in the same room and fed the same chow. The experiments were approved by the animal studies committee at Washington University School of Medicine.

Diagnosis of congenital heart defects

Neonatal pups were collected each morning within hours of birth to prevent cannibalization of animals that have serious congenital heart defects. The pups were euthanized. The thorax was fixed in 10% neutral buffered formalin overnight and then transferred to 70% ethanol. Fixed hearts were dissected, embedded in paraffin, and entirely sectioned in the frontal plane at 6 μm thickness. Each heart was inspected under a 5x microscope objective by at least two individuals. Defects were diagnosed by the appearance of the septae, valves, or blood in consecutive sections. Genomic DNA was isolated from every animal by phenol-chloroform extraction.

Statistical analyses

To compare the genotypic distribution in a particular strain or cross to the expected Mendelian ratio, all the pups from consecutive litters were counted up to at least 100 animals. Statistically significant deviation was defined in a χ^2-test by $P < 0.05$.

The incidence of specific types of heart defects in crosses were compared by a two-tailed Fisher's exact test. The co-incidence of heart defects in the same mouse was evaluated by comparison to the expected incidence, as determined by the product of their separate incidences, using a χ^2-test.

Correlation of muscular and membranous VSD incidence with the average number of C57Bl/6 alleles per locus as expected from the Mendelian distribution in the F2 crosses was evaluated by Pearson's product-moment correlation followed by a two-tailed t-test with significance defined by $P < 0.05$.

Segregation models were analysed by comparison of their predictions and observed incidences by a χ^2-test. The models and derivation of equations are detailed in the Supplemental Materials.

Results

Newborn Nkx2-5$^{+/-}$ animals in the C57Bl/6 background have a high incidence of ASD and VSD

Human *NKX2-5* mutations cause congenital heart defects, some of which are lethal if untreated in the newborn period. While such serious defects in *Nkx2-5$^{+/-}$* mice have not been described, we observed that ~17% of *Nkx2-5$^{+/-}$* pups in the C57Bl/6 background were missing at 2-3 weeks of age when they were weaned from their mothers. The attrition of *Nkx2-5$^{+/-}$* animals occurs after birth because newborn pups were obtained at the expected Mendelian ratio (Table 2.1). We therefore examined newborn hearts that were collected within hours of birth. A high incidence of ASD of the secundum type was found in *Nkx2-5$^{+/-}$* but not wild-type hearts (Figures 2.1, 2.2). A secundum ASD was diagnosed by an insufficiency in covering the fossa ovale due to a deficiency in its rim or of the septum primum. Hearts in which a potentially small ASD could not be distinguished from a patent foramen ovale were considered normal. The ASD incidence is greater than the ~20% previously reported by two different groups for adult *Nkx2-5$^{+/-}$* animals in the C57Bl/6 background possibly because of spontaneous closure or postnatal death.[20-22] VSDs of the membranous and muscular types were also found in *Nkx2-5$^{+/-}$* but not wild-type hearts (Figures 2.1, 2.2). VSDs were not previously reported in adult animals possibly because of spontaneous closure or death from pulmonary overcirculation in the newborn period, as seen in mouse mutants that have a persistent patent ductus arteriosus.[23]

Previous investigators have detected aortic stenosis by Doppler echocardiography in live animals and bicuspid aortic valves by gross examination of adult *Nkx2-5*$^{+/-}$ animals.[21] Given that we examined histologic sections, minor valvular abnormalities that might cause regurgitation or stenosis are difficult to assess. Abnormal blood flow patterns have also been described by fetal echocardiography in association with severe lesions like truncus arteriosus.[24] Severe defects and ones not easily detected by echocardiography like small VSDs because of physiologically elevated right-sided pressures in the fetus and newborn are easily recognized because of the greater spatial resolution of microscopy.

Nkx2-5$^{+/-}$ *animals in hybrid strain backgrounds generally have normal cardiac anatomy*

Nkx2-5$^{+/-}$ animals in the C57Bl/6 background were outcrossed to the FVB/N strain. The F1 progeny, which are C57Bl/6 X FVB/N hybrids, showed the expected Mendelian ratio of *Nkx2-5*$^{+/-}$ pups at weaning (Table 2.1). The absence of neonatal lethal defects was confirmed by direct inspection. Only one small VSD and 6 ASDs were found among 80 *Nkx2-5*$^{+/-}$ F1 newborn mouse hearts. All 52 wild-type hearts were normal. The incidence of either septal defect in the *Nkx2-5*$^{+/-}$ C57Bl/6 X FVB/N F1 hybrids is significantly lower compared to the C57Bl/6 background (Figure 2.2).

Nkx2-5$^{+/-}$ animals from the C57Bl/6 X A/J F1 hybrid cross similarly showed the expected Mendelian distribution of genotypes at weaning (Table 2.1). Direct inspection confirmed a reduced incidence of heart defects compared to the C57Bl/6 background, none of which was expected to be lethal. No VSDs and only 4 ASDs were found among

54 C57Bl/6 X A/J F1 *Nkx2-5*$^{+/-}$ newborn hearts. All 48 wild-type hearts were normal (Figure 2.2).

The C57Bl/6 X FVB/N F2 crosses reveal ASD and VSD susceptibility alleles in both strains

The C57Bl/6 X FVB/N F1 hybrid data suggest that C57Bl/6 carries one or more recessive alleles of modifier genes that increase susceptibility to heart defects in the presence of *Nkx2-5* mutation. The data do not exclude the possibility of recessive susceptibility alleles from FVB/N at other loci. To investigate these possibilities, *Nkx2-5*$^{+/-}$ F1 animals were backcrossed to C57Bl/6 and FVB/N or intercrossed to generate F2 progeny. Newborn F2 pups showed the expected Mendelian distribution of genotypes, which indicates the absence of selection against defects that could cause prenatal demise (Table 2.1).

VSD and ASD were more common in the *Nkx2-5*$^{+/-}$ animals from each of the C57Bl/6 X FVB/N F2 crosses than from the F1 but less common than in C57Bl/6 ($P < 0.001$ for comparisons to both the F1 and C57Bl/6). A VSD was present in 12-16% of the *Nkx2-5*$^{+/-}$ F2 hearts, and an ASD in 4-7% (Figure 2.3). These results indicate that at least two loci affect VSD and ASD susceptibility in the presence of *Nkx2-5* mutation. C57Bl/6 and FVB/N carry recessive, susceptibility alleles at different loci, given the increased incidence in the parental backcrosses compared to the F1. The *Nkx2-5*$^{+/-}$ F2 may have a lower incidence of particular defects compared to C57Bl/6 either because

FVB/N susceptibility genotypes have smaller effects or because interactions between C57Bl/6 and FVB/N alleles reduce risk.

The C57Bl/6 X A/J F2 crosses reveal A/J susceptibility alleles for ASD and atrioventricular septal defects

Similar to the C57Bl/6 X FVB/N F2 results, the *Nkx2-5*$^{+/-}$ F2 progeny from the C57Bl/6 X A/J intercross and C57Bl/6 backcross showed an incidence of VSD and ASD greater than the F1 and less than C57Bl/6 ($P < 0.001$ for each of the comparisons; Figure 2.3). The backcross to A/J was not performed. Comparing the C57Bl/6 X A/J to the C57Bl/6 X FVB/N F2 crosses revealed two significant differences.

First, ASDs occurred at a higher frequency in the C57Bl/6 X A/J F2 intercross than in the C57Bl/6 X FVB/N ($P < 0.001$). There was a similar trend between the F2 backcrosses to C57Bl/6, which was not significant possibly because of the smaller sample sizes ($P < 0.1$). A/J may carry polymorphisms that confer greater susceptibility specifically to ASDs in the presence of *Nkx2-5* mutation. The A/J polymorphisms do not increase susceptibility to septal defects in general because the VSD incidence was not different between the two intercrosses.

Second, atrioventricular septal defects (AVSD) occur at a low but significant incidence in the *Nkx2-5*$^{+/-}$ C57Bl/6 X A/J F2 intercross (Figures 2.3 and 2.4). *Nkx2-5* mutation causes the defect because none of 238 wild-type littermates from the same cross had an AVSD ($P < 0.05$). An A/J polymorphism enhances susceptibility because the defect is much rarer in the C57Bl/6 X FVB/N F2 intercross ($P < 0.001$). The A/J

polymorphism appears dominant with low penetrance because one AVSD was found among 121 *Nkx2-5*$^{+/-}$ hearts from the backcross to C57BL/6.

Double-outlet right ventricle occurs rarely in Nkx2-5$^{+/-}$ *F2 animals*

Double-outlet right ventricle (DORV), in which both great arteries arise from the right ventricle and the aortic and mitral valves are separated by a muscular conus, has been reported in association with human *NKX2-5* mutation. Two *Nkx2-5*$^{+/-}$ F2 animals were found to have DORV, one each from among 1552 C57Bl/6 X FVB/N and 1104 C57Bl/6 X A/J intercross hearts examined (Figure 2.4). The rarity of DORV precludes conclusions about the role of modifier genes in its pathogenesis. We note, though, that defects associated with human *NKX2-5* mutation that are more common than DORV in the general population like hypoplastic left heart syndrome have not been found in any *Nkx2-5*$^{+/-}$ mouse.[17,18,25] Such defects may have not been found either because the sample size remains too small or because the inbred strains examined do not carry the relevant susceptibility alleles.

Genotypes at multiple modifier loci determine the VSD phenotype

Membranous and muscular VSDs are found in *Nkx2-5*$^{+/-}$ hearts, but muscular VSDs are more common in the C57Bl/6 background ($P < 0.0005$; Figure 2.5). Which type of VSD develops appears to be determined by the cumulative effect of susceptibility genotypes at multiple modifier loci for either defect. We compared the incidence of each VSD type to the fraction of the C57Bl/6 genome in the genetically heterogeneous F2 crosses, as expected from a Mendelian distribution (Figure 2.5). There was a strong, positive correlation for muscular VSD ($r = 0.91$, $P < 0.04$) and a negative correlation for

membranous VSD ($r = -0.97$, $P < 0.006$). Thus, C57Bl/6, FVB/N and A/J all carry recessive, susceptibility alleles at multiple loci for either VSD phenotype, but in the mixed strain backgrounds the additive effect of the C57Bl/6 susceptibility genotypes is less or greater than the other two strains for membranous and muscular VSD, respectively. Additional epistatic effects between loci may also influence the incidence of specific congenital heart defects, as shown by segregation analyses below.

ASD and VSD co-occur more frequently than expected by chance in Nkx2-5$^{+/-}$ hearts of mixed genetic background

Polymorphic modifier genes clearly influence the development of specific types of defects in *Nkx2-5* mutation. To test the hypothesis that the same polymorphism could increase susceptibility to two different defects, the incidence of hearts having both defects was compared to the product of the individual defect incidences in the population, i.e., the chance or null hypothesis. VSD and ASD co-occur more often than expected by chance in the C57Bl/6 X FVB/N intercross and the C57Bl/6 X A/J intercross and backcross to C57Bl/6 (Figure 2.6A). In the two F2 intercrosses where the sample sizes are large, ASD co-occurred with either muscular or membranous VSD more often than expected (Figure 2.6B). On the other hand, muscular and membranous VSD did not co-occur more often than expected. The co-occurrence of AVSD with other heart defects was not evaluated because of its low incidence. Therefore, some modifiers may increase the susceptibility to several types of heart defects in the presence of *Nkx2-5* mutation whereas others may affect the development of just one.

Segregation analysis of the genetic architecture of common septal defects

To characterize the genetic architecture of modifiers that influence ASD and VSD susceptibility in *Nkx2-5*$^{+/-}$ animals, we examine two segregation models. The models focus on the C57Bl/6 X FVB/N crosses because data from F2 backcrosses to both parental strains are available. How well the predictions of a model fit the observed data set lower bounds on the number of modifiers and interactions that are involved in the pathogenesis of particular defects.

The first model, M1, postulates two loci *A* and *B* that do not interact. C57Bl/6 and FVB/N carry the recessive, susceptibility alleles of *A* and *B*, respectively. The effects of susceptibility genotypes at *A* and *B* are estimated from the incidences in the C57Bl/6 strain and the F2 backcross to FVB/N. The expected incidences in the F2 intercross and C57Bl/6 backcross are then calculated from the model. Inclusion in the model of additional loci does not alter its predictions because the total effect of susceptibility alleles from one strain reduces to a single term. The predictions deviate significantly from the observed incidences of ASD and muscular VSD ($P < 0.0001$) but approximate the membranous VSD incidences ($P = 0.14$. Figure 2.7A-C). Therefore, at least two loci affect all three defect types.

Model M2 permits interaction between two modifier loci. In the simplest case, C57Bl/6 carries the recessive susceptibility alleles at two modifier loci *A* and *B*, which interact synergistically. FVB/N carries the recessive susceptibility allele at a third locus *C*. If the effect of *A* or *B* alone is negligible compared to their interaction, then the incidence of the defect in C57Bl/6 provides an estimate of the epistatic effect. Based on this model, the incidences attributed to the epistatic effect and locus *C*, as estimated from

the F2 FVB/N backcross, match the observed incidences of muscular VSD in the F2 intercross and C57Bl/6 backcross (Figure 2.7A).

M2 underestimates the incidences of membranous VSD in the F2 intercross and C57Bl/6 backcross ($P < 0.0001$, Figure 2.7B). M2 may not model the genetic architecture of membranous VSD because either epistatic interactions have no role (i.e., M1 is correct) or there are unaccounted interactions between C57Bl/6 and FVB/N alleles at other loci.

M2 poorly fits the ASD incidences in the F2 intercross and C57Bl/6 backcross ($P < 0.0001$; Figure 2.7C). ASDs appear more genetically complex than VSD. First, additional loci and interactions must be postulated to account for the >4-fold difference in the incidence of ASD between the C57Bl/6 backcross and C57Bl/6. Inspection of the incidences in the crosses indicates that more than two modifier loci in each of the lines are likely be involved. Second, M1 and M2 predict the absence of defects in the F1 because homozygosity of a susceptibility allele at a modifier locus is assumed necessary for a defect. The assumption appears reasonable for muscular and membranous VSDs but not for ASD. The presence of ASDs in the F1 suggests that heterozygosity at some loci confer susceptibility.

Taken together, the two segregation models indicate that three or more modifier loci and at least one epistatic interaction influence the susceptibility to muscular VSD and ASD in *Nkx2-5*$^{+/-}$ animals. At least two loci influence membranous VSD susceptibility. The loci and interactions inferred from the C57Bl/6 X FVB/N crosses probably represent only a subset of all that exist among inbred mouse strains or the human population. A

more precise estimate of the number of modifier loci in the mouse model, whether independent or epistatic, involved in any of the defects requires additional information, such as genetic linkage analyses that are underway.

Discussion

Normal cardiac development is critical for the survival of an individual. Any mechanism that could ensure normal development would increase the fitness of a species. In the short term, natural selection would eliminate major deleterious mutations. In contrast, stabilizing selection could promote the evolution of a versatile system that buffers the effects of genetic and environmental perturbations to the developing heart. The invariance of the normal heart form would then be enhanced.[6] Waddington described the general mechanism metaphorically as "canalization" in which development flows down channels to produce a highly invariant wild-type form. He considered the depth of the channels analogous to the degree of buffering, which he deduced to have a genetic basis.[26] Crucial phenotypes are maintained about a stable optimum. Sufficiently great insults, however, decanalize an individual and reveal the influence of cryptic variants of modifier genes that do not affect the phenotype of genetically wild-type individuals.[5] Investigation of modifier genes thus provides an alternative approach to the genetic pathways that shape a normal heart. Cryptic polymorphisms within them may give rise to much of the complexity of congenital heart disease.

The results of inbred strain crosses indicate that modifier genes contribute to the canalization and decanalization of cardiac development in the presence of *Nkx2-5* mutation. A large fraction of *Nkx2-5*$^{+/-}$ animals in the C57Bl/6 background has atrial and ventricular septal defects, whereas few among F1 hybrid crosses to FVB/N or A/J do. Defects recur in F2 crosses at a lower incidence than in C57Bl/6. We interpret this to mean that strain-specific polymorphisms of modifier genes alter the susceptibility of

certain cardiac developmental pathways to *Nkx2-5* mutation but do not affect the wild-type form. Complementation of susceptibility alleles in the F1 and homozygosity at some but not all modifier loci in the F2 contribute to the relative differences in defect incidence in the successive crosses. Furthermore, analysis of the incidences of ASD and VSD in the F2 crosses and C57Bl/6 suggest that modifier loci act independently of or epistatically with other loci. Thus, a protective allele at a modifier locus might directly reduce the risk of a defect or abrogate a synergistic interaction between loci that increases risk.

The surprisingly normal hearts of *Nkx2-5*$^{+/-}$ animals from the F1 hybrid crosses illustrate the important role of polymorphic modifier genes in ensuring the robustness of cardiac development. Protective polymorphisms likely exist against a variety of insults, as reductions in heart defect incidence have also been observed in *Tbx5* or *Gata4* mutants in mixed compared to isogenic strain backgrounds.[14,15] The greater heterogeneity in humans could provide even more genetic material to buffer against numerous, potentially common insults like nutrient deprivation or congenital infection. For well-canalized traits, cryptic variation can accumulate in a population in the absence of a major perturbation because the phenotype is buffered against change.[5] Natural selection due to diverse pressures in the wild could also maintain heterogeneity because normally cryptic polymorphisms may be protective or detrimental depending upon the context. The resulting genetic architecture could ensure that a fraction of the population survives almost any particular perturbation.

Recessive susceptibility alleles of modifier genes exist in each of the three inbred strains examined, given the presence of defects in the *Nkx2-5*$^{+/-}$ F2 progeny from all the parental backcrosses and intercrosses. Our statistical analysis of thousands of hearts reveals that modifier genes fixed for different variants in the inbred strains influence the development of specific types of defects in the presence of *Nkx2-5* mutation. Relative to the other strain backgrounds, C57Bl/6 carries polymorphisms that increase the susceptibility to muscular VSD, and A/J carries polymorphisms that increase the susceptibility to ASD and AVSD. Modifier genes reside in pathways that affect the development of individual or multiple anatomic structures like the atrial and ventricular septum, based upon the higher than expected co-incidence of some defects. The observations inform how modifier loci and interactions could be mapped to define the genetic pathways leading from *Nkx2-5* mutation to various types of defects.

Even genetic diseases thought to exhibit simple Mendelian inheritance manifest with incomplete penetrance and varying phenotypes. The properties of genetic modifiers of the *Nkx2-5* mutant phenotype help to explain the basis of the variable presentations of congenital heart disease. Ostensibly sporadic cases of congenital heart disease could result from germline or somatic mutations, teratogens, or purely stochastic events. Our results offer an alternative, but not mutually exclusive possibility that polymorphisms of genetic modifiers can either suppress or promote the effect of a major mutation. Maximum genetic heterogeneity in the F1 is associated with near complete suppression of deleterious phenotypes, leading to normal carriers. On the other hand, homozygosity at some loci in the F2 contributes to the manifestation of a specific defect among multiple

potential phenotypes, which in a heterogeneous human population could appear like sporadic disease or pleiotropic defects.

This large study, which systematically varies the genetic background of a mouse model, offers insights into how genetic modifiers ensure normal cardiac development and influence the manifestation of congenital heart disease. Knowledge resulting from the characterization and identification of the modifiers in the mouse will almost certainly be relevant to human disease because pathways in cardiovascular development and physiology are strongly conserved. For example, dozens of quantitative trait loci that affect blood pressure are common between rodent models and human populations.[27,28] Cardiac development is robust. Appreciating how, one could conceivably mimic nature to reduce the heavy burden of congenital heart disease on children and their families.

Reference List

(1) Bruneau BG. Transcriptional regulation of vertebrate cardiac morphogenesis. *Circ Res.* 2002;90:509-519.

(2) Gruber PJ, Epstein JA. Development gone awry: congenital heart disease. *Circ Res.* 2004;94:273-283.

(3) Ursell PC, Byrne JM, Strobino BA. Significance of cardiac defects in the developing fetus: a study of spontaneous abortuses. *Circulation.* 1985;72:1232-1236.

(4) Chinn A, Fitzsimmons J, Shepard TH, Fantel AG. Congenital heart disease among spontaneous abortuses and stillborn fetuses: prevalence and associations. *Teratology.* 1989;40:475-482.

(5) Gibson G, Dworkin I. Uncovering cryptic genetic variation. *Nat Rev Genet.* 2004;5:681-690.

(6) Gibson G. Decanalization and the origin of complex disease. *Nat Rev Genet.* 2009;10:134-140.

(7) Basson CT, Bachinsky DR, Lin RC et al. Mutations in human TBX5 cause limb and cardiac malformation in Holt-Oram syndrome. *Nat Genet.* 1997;15:30-35.

(8) Li QY, Newbury-Ecob RA, Terrett JA et al. Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nat Genet.* 1997;15:21-29.

(9) Schott JJ, Benson DW, Basson CT et al. Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science.* 1998;281:108-111.

(10) Garg V, Kathiriya IS, Barnes R et al. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature.* 2003;424:443-447.

(11) Gill HK, Splitt M, Sharland GK, Simpson JM. Patterns of recurrence of congenital heart disease: an analysis of 6,640 consecutive pregnancies evaluated by detailed fetal echocardiography. *J Am Coll Cardiol.* 2003;42:923-929.

(12) Nora JJ. From generational studies to a multilevel genetic-environmental interaction. *J Am Coll Cardiol.* 1994;23:1468-1471.

(13) Whittemore R, Wells JA, Castellsague X. A second-generation study of 427 probands with congenital heart defects and their 837 children. *J Am Coll Cardiol.* 1994;23:1459-1467.

(14) Bruneau BG, Nemer G, Schmitt JP et al. A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell.* 2001;106:709-721.

(15) Rajagopal SK, Ma Q, Obler D et al. Spectrum of heart disease associated with murine and human GATA4 mutation. *J Mol Cell Cardiol.* 2007;43:677-685.

(16) Sakata Y, Koibuchi N, Xiang F, Youngblood JM, Kamei CN, Chin MT. The spectrum of cardiovascular anomalies in CHF1/Hey2 deficient mice reveals roles in endocardial cushion, myocardial and vascular maturation. *J Mol Cell Cardiol.* 2006;40:267-273.

(17) Benson DW, Silberbach GM, Kavanaugh-McHugh A et al. Mutations in the cardiac transcription factor NKX2.5 affect diverse cardiac developmental pathways. *J Clin Invest.* 1999;104:1567-1573.

(18) McElhinney DB, Geiger E, Blinder J, Benson DW, Goldmuntz E. NKX2.5 mutations in patients with congenital heart disease. *J Am Coll Cardiol.* 2003;42:1650-1655.

(19) Tanaka M, Chen Z, Bartunkova S, Yamasaki N, Izumo S. The cardiac homeobox gene Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. *Development.* 1999;126:1269-1280.

(20) Azhari N, Shihata MS, Al-Fatani A. Spontaneous closure of atrial septal defects within the oval fossa. *Cardiol Young.* 2004;14:148-155.

(21) Biben C, Weber R, Kesteven S et al. Cardiac septal and valvular dysmorphogenesis in mice heterozygous for mutations in the homeobox gene Nkx2-5. *Circ Res.* 2000;87:888-895.

(22) Tanaka M, Berul CI, Ishii M et al. A mouse model of congenital heart disease: cardiac arrhythmias and atrial septal defect caused by haploinsufficiency of the cardiac transcription factor Csx/Nkx2.5. *Cold Spring Harb Symp Quant Biol.* 2002;67:317-325.

(23) Nguyen M, Camenisch T, Snouwaert JN et al. The prostaglandin receptor EP4 triggers remodelling of the cardiovascular system at birth. *Nature.* 1997;390:78-81.

(24) Shen Y, Leatherbury L, Rosenthal J et al. Cardiovascular phenotyping of fetal mice by noninvasive high-frequency ultrasound facilitates recovery of ENU-induced mutations causing congenital cardiac and extracardiac defects. *Physiol Genomics.* 2005;24:23-36.

(25) Elliott DA, Kirk EP, Yeoh T et al. Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J Am Coll Cardiol.* 2003;41:2072-2076.

(26) Waddington CH. Canalization of development and the inheritance of acquired characters. *Nature.* 1942;150:563-565.

(27) Cowley AW, Jr. The genetic dissection of essential hypertension. *Nat Rev Genet.* 2006;7:829-840.

(28) Feng M, Deerhake ME, Keating R et al. Genetic Analysis of Blood Pressure in 8 Mouse Intercross Populations. *Hypertension.* 2009;54:802-809.

Figures:

Figure 2.1. Atrial and ventricular septal defects in newborn *Nkx2-5*$^{+/-}$ animals. An intact ventricular septum (A) and a normal atrial septum with a patent foramen ovale (B) in wild-type hearts. (C) Membranous VSD and (D) muscular VSD in *Nkx2-5*$^{+/-}$ hearts. (E) Secundum ASD in an *Nkx2-5*$^{+/-}$ heart. Arrowheads point to the relevant defect in each section. Ao, Aorta. RA, LA, right, left atrium. RV, LV, right, left ventricle.

Figure 2.2. The incidence of septal defects in wild-type and *Nkx2-5*$^{+/-}$ animals in the C57Bl/6 background and in F1 crosses to FVB/N and A/J. The incidence of ASD and VSD in *Nkx2-5*$^{+/-}$ hearts is substantially reduced in the F1 hybrids compared to the C57Bl/6 parental strain ($P < 0.001$).

Figure 2.3. The incidence of specific heart defects in *Nkx2-5*$^{+/-}$ animals from F2 intercrosses and parental backcrosses of C57Bl/6 and FVB/N or A/J. Cross-specific differences in the incidence of ASD and AVSD are significant between the C57Bl/6 X FVB/N and C57Bl/6 X A/J F2 intercrosses. The incidence of VSD is comparable between the C57Bl/6 X FVB/N and X A/J F2 crosses.

Figure 2.4. Rarer heart defects are found in *Nkx2-5*$^{+/-}$ F2 progeny. Sections through an *Nkx2-5*$^{+/-}$ heart from the C57Bl/6 X A/J F2 intercross demonstrate an atrioventricular

septal defect with its VSD (A), common valve (B) and primum ASD (C) marked by arrowheads in each section. Sections through an *Nkx2-5*$^{+/-}$ heart from the C57Bl/6 X FVB/N F2 intercross demonstrate double outlet right ventricle with the pulmonary artery (D, PA) and aorta (F, Ao) both arising from the right ventricle. An intermediate section (E) demonstrates the side-by-side arrangement of the great arteries.

Figure 2.5. The incidence of muscular and membranous VSD among the F2 crosses and C57Bl/6 background is determined by the cumulative effect of multiple, strain-specific susceptibility genotypes. Muscular VSD is more common in *Nkx2-5*$^{+/-}$ animals in the C57Bl/6 background than membranous ($P < 0.0005$). In the F2 crosses, muscular VSD incidence is positively correlated with the average fraction of the C57Bl/6 genome in the progeny ($R = 0.91$, $P < 0.02$). Membranous VSD is negatively correlated ($R = -0.97$, $P < 0.003$). The average number of C57Bl/6 or non-C57Bl/6 alleles per locus based on a Mendelian distribution in each F2 cross is noted. MBR, membranous VSD. MUSC, muscular VSD.

Figure 2.6. Some modifier loci may contribute to the development of both ASD and VSD. (A) ASD and VSD co-occur in the F2 intercrosses and C57Bl/6 backcross of the C57Bl/6 X A/J F1 more often than expected by the null hypothesis of complete independence between the defects. (B) The co-incidence rates are subdivided by VSD types in the two intercross populations, which have sufficiently large sample sizes for statistical comparison.

Figure 2.7. Segregation analyses define the minimum number of modifier loci and interactions involved in the susceptibility to ASD and VSD. The observed and expected incidences of (A) muscular VSD, (B) membranous VSD, and (C) ASD are predicted in the *Nkx2-5*$^{+/-}$ C57Bl/6 X FVB/N F1 and F2 progeny under two segregation models. Model M1 postulates two loci that do not interact. M2 postulates three loci, two of which interact. M1 approximates the observed incidences of membranous VSD but fits muscular VSD and ASD poorly. M2 fits muscular VSD well but not ASD or membranous VSD. Both models predict no defects in the F1, which contrasts with the ASDs observed.

Table 2.1

Table. Observed and Expected Distributions of *Nkx2-5*$^{+/-}$ and Wild-Type Pups

				Nkx2-5$^{+/-}$		Wild Type		
Background	*Nkx2-5* Genotype	Age	n	Observed	Expected	Observed	Expected	*P*
C57Bl/6	HET×HET	P21	505	62.5	66.7	37.4	33.3	<0.05
C57Bl/6	HET×WT	P21	854	45.3	50	54.7	50	<0.01
C57Bl/6	HET×WT	P0	100	56	50	44	50	NS
C57Bl/6×FVB/N F1	HET×WT	P21	214	46.7	50	53.3	50	NS
C57Bl/6×A/J F1	HET×WT	P21	156	51.9	50	48.1	50	NS
C57Bl/6×FVB/N F2	HET×HET	P0	212	66.5	66.7	33.5	33.3	NS
C57Bl/6×A/J F2	HET×HET	P0	562	64.4	66.7	35.6	33.3	NS

Shown are the observed and expected distributions of *Nkx2-5*$^{+/-}$ and wild-type pups in various crosses either at birth (P0) or at weaning (P21). Pups were collected within hours of birth or at weaning.

Figure 2.1

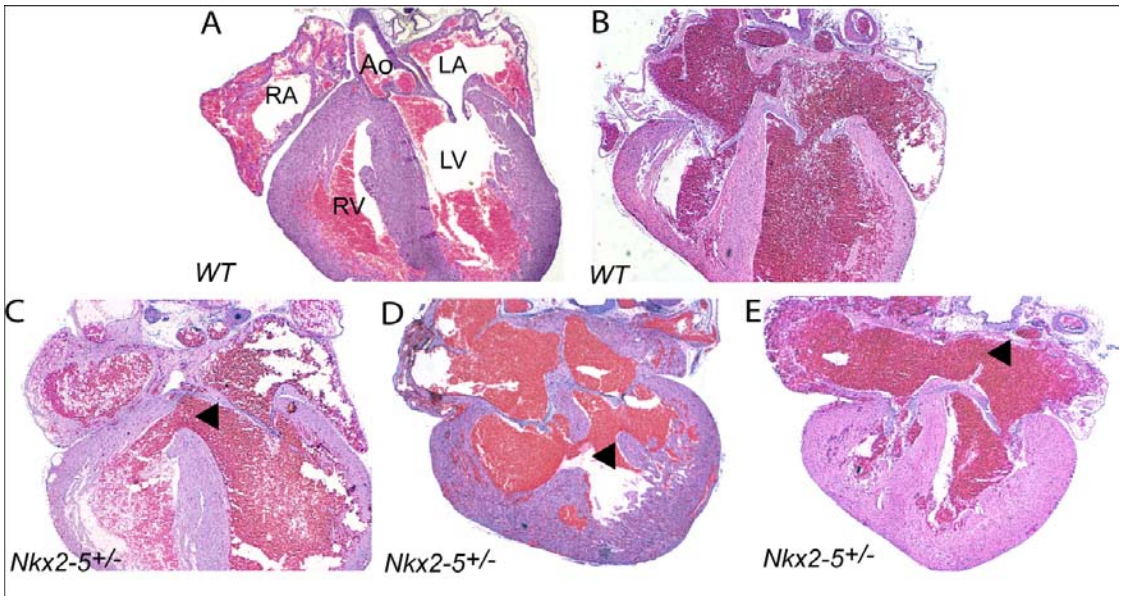


Figure 2.2

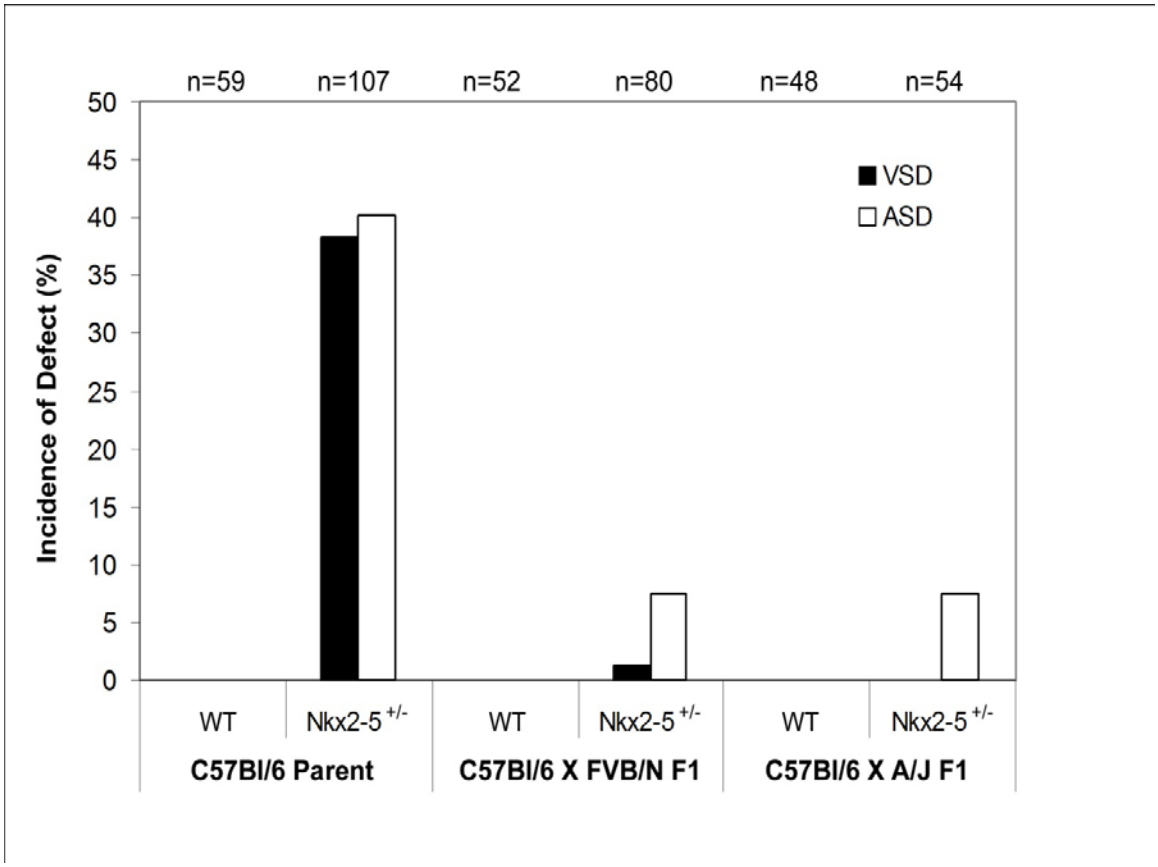

n=59
n=107
n=52
n=80
n=48
n=54
50
45
40
35
30
25
20
15
10
5
0
Incidence of Defect (%)
VSD
ASD
WT
Nkx2-5 +/-
WT
Nkx2-5 +/-
WT
Nkx2-5 +/-
C57Bl/6 Parent
C57Bl/6 X FVB/N F1
C57Bl/6 X A/J F1

Figure 2.3

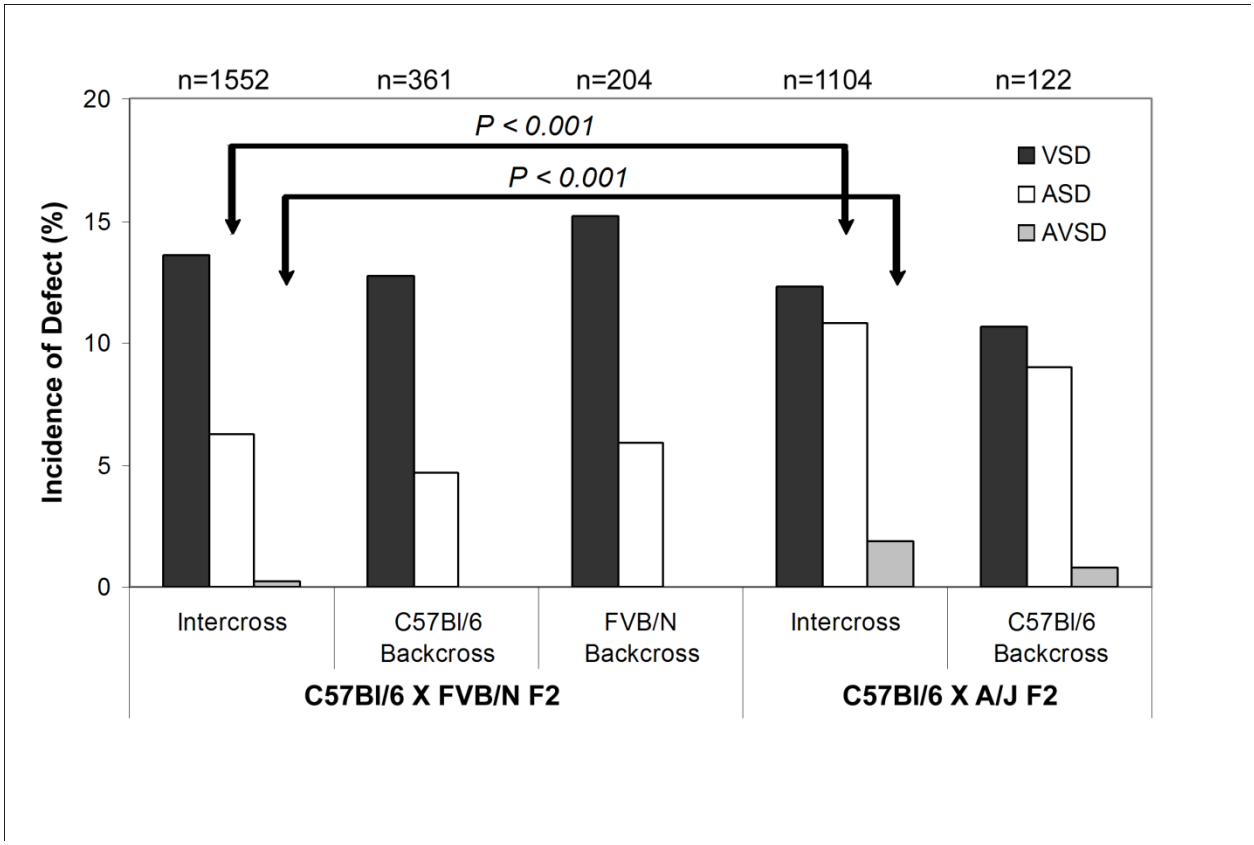

n=1552
n=361
n=204
n=1104
n=122
P < 0.001
P < 0.001
VSD
ASD
AVSD
Incidence of Defect (%)
20
15
10
5
0
Intercross
C57Bl/6 Backcross
FVB/N Backcross
Intercross
C57Bl/6 Backcross
C57Bl/6 X FVB/N F2
C57Bl/6 X A/J F2

Figure 2.4

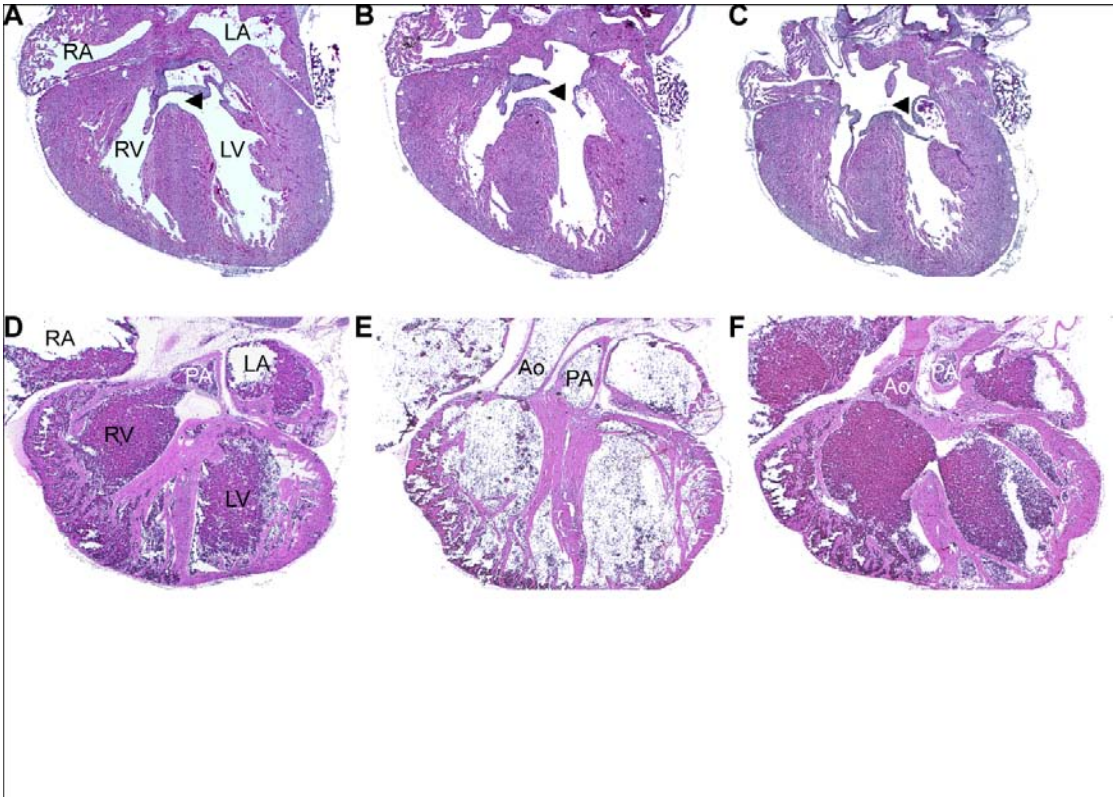


Figure 2.5

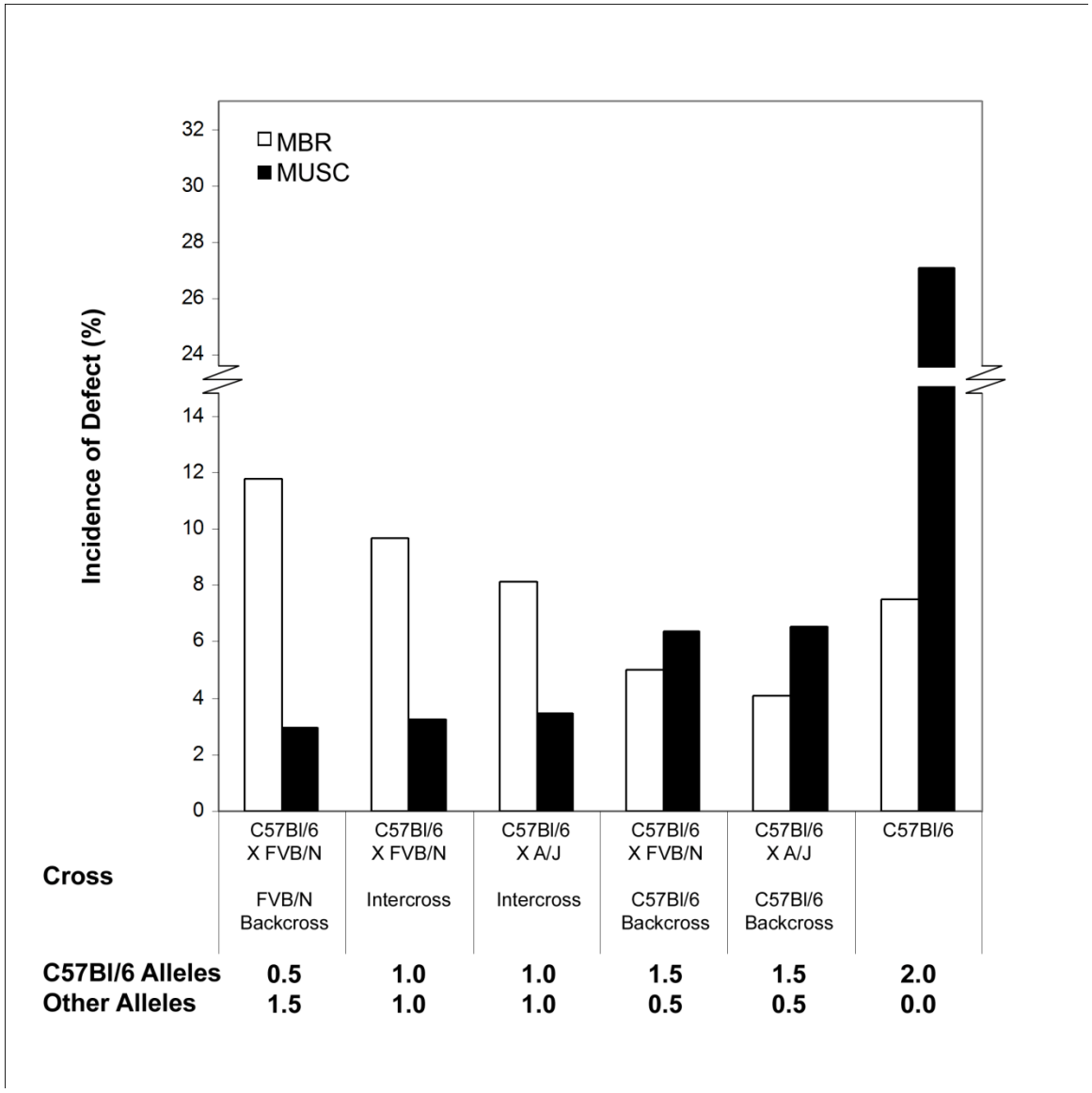

MBR
MUSC
Incidence of Defect (%)
32
30
28
26
24
14
12
10
8
6
4
2
0
Cross
C57Bl/6 X FVB/N
FVB/N Backcross
C57Bl/6 X FVB/N
Intercross
C57Bl/6 X A/J
Intercross
C57Bl/6 X FVB/N
C57Bl/6 Backcross
C57Bl/6 X A/J
C57Bl/6 Backcross
C57Bl/6
C57Bl/6 Alleles 0.5 1.0 1.0 1.5 1.5 2.0
Other Alleles 1.5 1.0 1.0 0.5 0.5 0.0

Figure 2.6

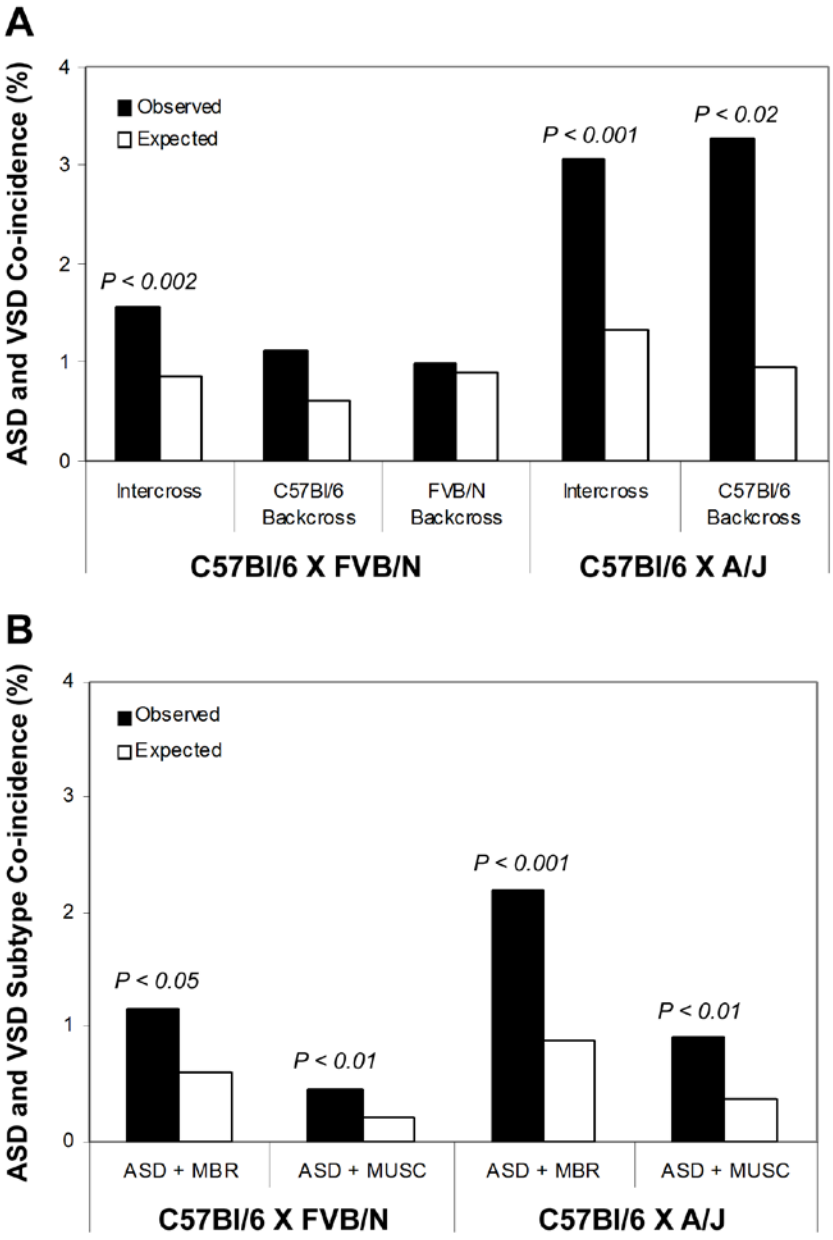
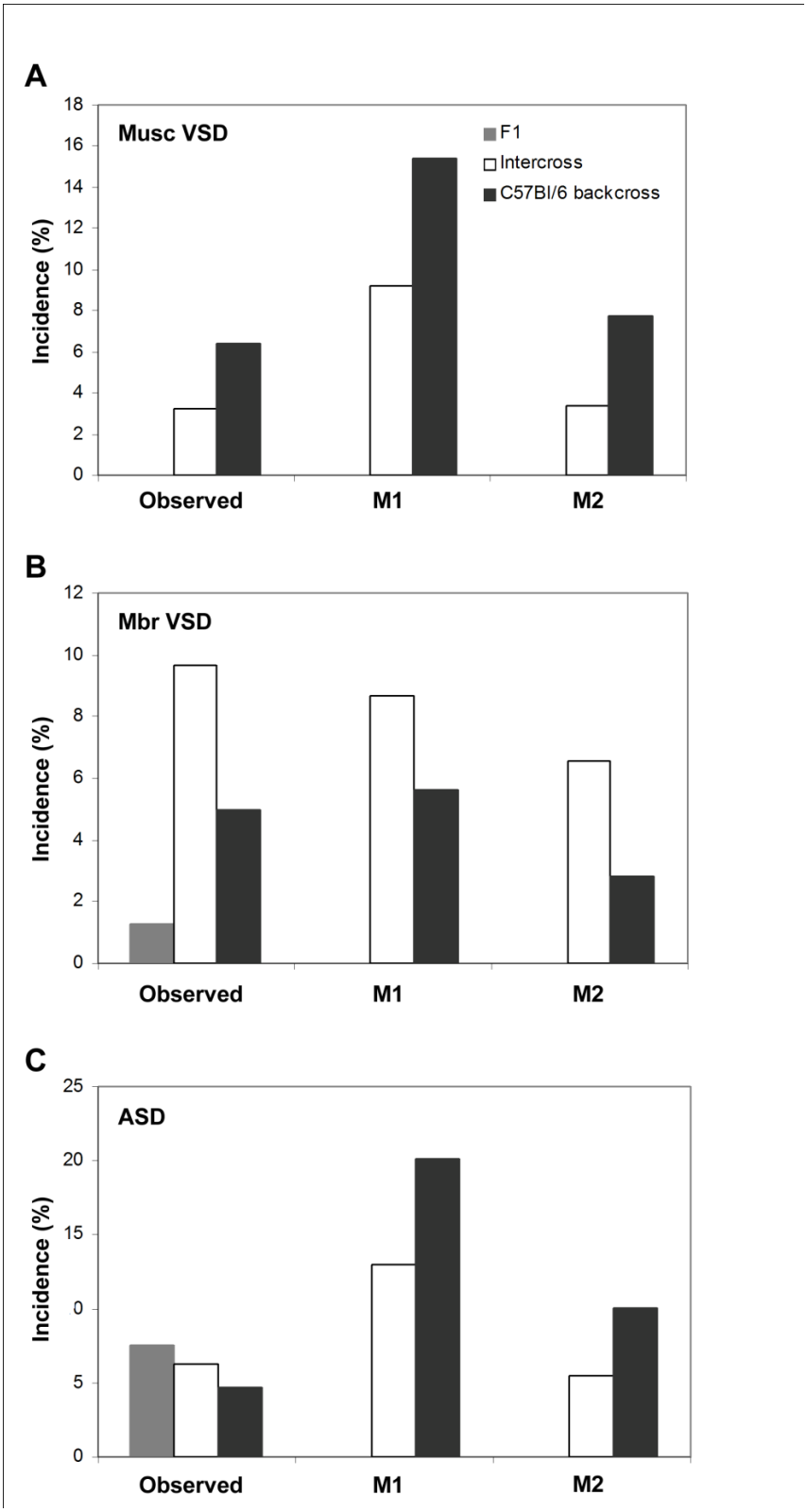

A
ASD and VSD Co-incidence (%)
Observed
Expected
P < 0.002
P < 0.001
P < 0.02
Intercross
C57Bl/6 Backcross
FVB/N Backcross
Intercross
C57Bl/6 Backcross
C57Bl/6 X FVB/N
C57Bl/6 X A/J
B
ASD and VSD Subtype Co-incidence (%)
Observed
Expected
P < 0.05
P < 0.01
P < 0.001
P < 0.01
ASD + MBR
ASD + MUSC
ASD + MBR
ASD + MUSC
C57Bl/6 X FVB/N
C57Bl/6 X A/J

Figure 2.7


A
Musc VSD
F1
Intercross
C57Bl/6 backcross
Incidence (%)
18
16
14
12
10
8
6
4
2
0
Observed
M1
M2
B
Mbr VSD
Incidence (%)
12
10
8
6
4
2
0
Observed
M1
M2
C
ASD
Incidence (%)
25
20
15
0
5
0
Observed
M1
M2

Chapter 3: Characterization of Modifiers of Congenital Heart Disease in an Nkx2-5$^{+/-}$ risk Model

Abstract

While a clear heritable risk has been observed for congenital heart disease, there is considerable variation in penetrance and presentation likely due to multiple genetic and environmental risk factors. To identify causative genetic and environmental factors and interactions responsible for variability in heart development, greater than 4,200 hearts from Nkx2-5 heterozygous knockout mice have been collected and examined. Genome wide linkage analysis was conducted on 306 mice from the FVB/n F2 intercross and 80 mice from the A/J F2 intercross diagnosed with VSDs to map main effect and interacting loci throughout the mouse genome that correlate with risk. Additionally the possibility of environmental interactions was explored in this analysis by including maternal and paternal age and litter size as additive and interactive terms in the models tested. Significantly linked genomic regions were identified from the FVB/n population on chromosomes 6, 8 and 10 implicating genes in these positions as important candidates for VSD risk. Linkage analysis on the A/J cross identified both shared and unique modifiers from the FVB/N cross scan. Maternal age was found to significantly correlate with VSD risk in FVB/N crosses but not A/J providing evidence for strain specific susceptibility to a non-heritable risk factor. The findings in this study implicate modifier genes as major players in cardiac developmental pathways by buffering against genetic and environmental insults in the majority of a population while directing the manifestation of disease in a few. Characterization of the genetic architecture of congenital heart disease in a mouse model will provide a deeper

understanding of its multifactorial nature and possibly lead to novel strategies for prognosis and prevention.

Introduction

Congenital heart disease is the top source of birth defect related death and a major burden on our health care system as 10% of these children require 1 or more surgeries in their lifetime [1]. With a 16% incidence in offspring of affected patients, it is clear that a strong genetic component exists and thus epidemiological patterns have been studied extensively over the last 50 years [2]. Up until the past decade, syndromes with highly penetrant heart defects have resulted in an emphasis on monogenic causes that highlight the outdated genetic determinism paradigm where one gene is linked to one phenotype. Pleiotropy further complicates the disease model as most disease causing mutations described increase risk for a number of different types of defects. Congenital heart disease has thus been reported as having a multifactorial model of inheritance indicating that multiple genes and environmental factors are to blame for the variability in disease manifestation [3].

Despite heart defects being a major health concern for 1 in every 120 babies born in the United States, the phenotype for the other 119 babies is a normal heart. The complexity of the system may reflect redundancy, and compensatory regulatory interactions that buffer the system to genetic and environmental perturbations [4]. The range of phenotypes segregating within human families and among different inbred mouse backgrounds with a common disease gene reflects the influence of an

underlying genetic distribution [5]. The effects of mutations in major heart development genes are likely modified by other factors which may be genetic or environmental.

Investigating modifiers is challenging because their phenotypic effects might only become observable in the context of a dramatic perturbation of the regulatory network that controls heart development. Gibson et al has coined the term cryptic variation to describe this phenomenon which is likely the rule rather than the exception in complex disease etiology. Important information lies in this genetic variance because while specific alleles of genetic modifiers may cause congenital malformations, it is likely that other variants are responsible for buffering the effects of mutations in major heart genes and thus responsible for how robust heart development is the majority of the time.

In this study we investigate potential sources of variability in risk using a simple inbred mouse model system in which much of the variation that occurs in natural populations can be controlled for. Both heritable and non heritable modifiers may affect penetrance, dominance deviation, expressivity and pleiotropy, but only in the presence of a major heart gene mutation [6]. The Nkx2-5 heterozygous knockout mouse in the C57Bl/6 inbred strain served as a near homogeneous, disease vulnerable population in which we could add in genetic sources of variability to observe the affect on heart disease risk. Outcrossing this mouse to two other inbred strains provided for variability in genetic background that could then be mapped to identify genetic modifiers. Non heritable risk factors such as parental age and litter size were also noted as potential sources of variability in disease risk.

In this study, we seek to parse out the underlying variation that affects heart defect incidence and presentation through genetic linkage mapping. In addition, we explore how environmental variance due to non heritable factors such as parental age of the F2 pups and litter size might modify disease risk. Once we are able to isolate the different sources of variance responsible for the complex nature of congenital heart disease, the medical community will be considerably better equipped to prevent and protect against this debilitating health problem.

Methods

Mouse Strains and Breeding

Nkx2-5$^{+/-}$ mice were backcrossed to C57Bl/6 for 10-13 generations. The knockout allele was generated and genotyped as previously described [7]. The C57Bl/6 and FVB/N inbred strains were obtained from Charles River Laboratories and the A/J strains from the Jackson Laboratories. Nkx2-5$^{+/-}$ C57Bl/6 males were crossed to FVB/N females to generate the F1 hybrid mice. *Nkx2-5*$^{+/-}$ F1 animals were intercrossed to generate F2 progeny as described [8]. Animals were housed under standard conditions in the same room and fed the same chow. The experiments were approved by the animal studies committee at Washington University School of Medicine.

Approximately 4,200 F2 neonates were collected within hours of birth to prevent cannibalization and to keep track of population genotype frequencies. The neonates were euthanized, dissected and genotyped for the Nkx2-5 knockout through

standard PCR protocol. The mouse torso sections were fixed in 10% neutral buffered formalin for 2 nights and then transferred to 70% ethanol. The fixed hearts were dissected out, paraffin-embedded and serially sectioned completely in the frontal plane at 6 μm thickness. The cardiac tissue was stained with hematoxylin and eosin for examination under a 5x microscope objective. Diagnoses of heart defects were independently validated by at least two people for each heart and recorded (Table 3.1). Genomic DNA was isolated from every animal by standard phenol-chloroform extraction.

SNP Genotyping

Approximately 120 SNPs polymorphic between C57Bl/6 and FVB/n or A/J mouse strains were chosen for genome wide coverage of the 19 autosomes at an average density of 15-20cM. SNPs with common alleles between FVB/n and A/J were chosen whenever possible. High throughput genotyping was accomplished using the Sequenom MassARRAY system at the Human Genetics Division Genotyping Core at Washington University.

Statistical Analysis

Linear dependence between non heritable factors; litter size, maternal and paternal age was evaluated using the Pearson Product Moment correlation. Logistic regression using the R statistical software package was then carried out to determine correlation of the binary heart phenotype with each of the continuous non heritable variables. Maternal and Paternal age were log transformed prior to analysis to adjust

for skewness and kurtosis. Statistical significance was defined by $P < 0.05$. Because correlation was evident between the predictors, multiple logistic regression was performed anytime multiple predictors for a heart phenotype were found to be significant. Odds ratios and adjusted odds ratios where determined from the regression coefficients along with 95% confidence intervals.

To detect risk loci, genome-wide linkage analysis was conducted with Interval mapping using the R/QTL software package. A total of 251 C57Bl/6 x FVB/n F2 mice with membranous VSDs and 80 mice muscular VSDs as well as 285 control mice were included in the genome scan. For the C57Bl/6 x A/J cross, 87 F2 mice with membranous VSDs and 116 controls were included in the analysis. The heart defects were found to occur evenly in both sexes indicating a lack of sex specific effects. The Expectation-Maximization algorithm with a binary model was specified to generate LOD scores representing the likelihood of linkage at a particular genomic region. Main effect scans evaluated the probability of VSD risk linkage at each marker genotyped as well as imputed marker genotypes at 5.0 cM intervals between true markers. The scan is designed to detect the presence of single QTL effects throughout the mouse genetic map.

Both litter size and maternal age were added to single effect scans as additive and then interactive covariates separately. The additive covariates account for potential residual variance and the interactive covariates detect gene by covariate specific interactions. A LOD score exceeding a 0.05 threshold alpha value was considered significant.

The Genome wide threshold of significance was determined through the use of permutation analysis as described by Sen and Churchill [9]. Greater than 5000 permutations were conducted for each main effect scan run. This method has the strength of taking the data structure into account as it randomizes the genotypes with respect to phenotypes for each permutation. The weakness of this method is that it can give inflated significance to smaller sample sized runs (n < 100 affected) [10].

A two-dimensional (two-way) scan was conducted to account for the potential effects of multiple genetic loci, some of which may interact. The interaction may be additive, synergistic or more complex as when the expression of one locus depends on the genotype of another (an epistatic interaction). Each marker (including imputed markers) is tested in combination with every other marker for correlation with phenotype. The following models are considered in this analysis for each of the locus pair combinations as described in Broman et al[11,12].

Full: $(y) = \beta_0 + \beta_1(qtl_1) + \beta_2(qtl_2) + \beta_3(qtl_1,qtl_2) + \varepsilon$

Additive: $(y) = \beta_0 + \beta_1(qtl_1) + \beta_2(qtl_2) + \varepsilon$

Single QTL: $(y) = \beta_0 + \beta_1(qtl_1) + \varepsilon$

Null: $(y) = \beta_0 + \varepsilon$

The models are then compared to generate five LOD scores:

LOD_f = Full – Null

LOD_a = Additive – Null

LOD_{fv1} = Full – Single

LOD_{av1} = Additive – Single

LOD_i = Full - Additive

Genome wide thresholds for significance for all five LOD scores tested were generated with 1000 permutations for the two-way scans.

$Full_T$	$Fv1_T$	Int_T	Add_T	$Av1_T$	One_T
8.51	6.69	6.1	5.85	2.79	3.44

Loci of interest were identified by two dimensional scan results based on recommendations by Broman at all in which multiple LOD score thresholds are interpreted in combination:

(1) $LOD_f > Full_T$ and either $LOD_{fv1} > Fv1_T$ or $LOD_i > Int_T$

(2) $LOD_a > Add_T$ and $LOD_{av1} > Av1_T$

Satisfaction of condition (1) indicates evidence for two loci in which there is potential for epistatic interactions. In contrast, condition (2) solely supports evidence that there are two loci interacting additively.

Two-way genome wide scans have high and sometimes stifling thresholds of significance to account for the large number of tests being conducted. Larger sample sizes are required since there are 9 possible biallelic genotypes that need to be represented for each marker locus. Single effect genome scans were also conducted

with significant candidate peaks as covariates. Scans were conducted with a candidate locus set as an additive covariate and then compared to scans with the candidate locus as an interactive covariate. Any regions which show a significant jump in LOD score would be considered to be evidence of an interaction, however only subject to single genome wide scan thresholds because of prior knowledge of the candidate.

Models including significant main effect loci and non heritable variables were tested using the *fitqtl* function provided in R/qtl in order to estimate the percentage of phenotypic variance explained [11]. A drop analysis was also conducted with this function to test models in which each term was successively dropped and then compared to the full model.

Results

Heritable Risk Factors

Linkage analysis on the binary phenotype of (VSD presence/No defects) was performed to detect significant main effect loci across the genome indicating locations of potential risk modifying candidates (Figure 3.1). The F2 intercross generation of the C57Bl/6 x FVB/n cross yielded 3 loci located on chromosomes 6, 8 and 10 with logarithm of odds (LOD) scores exceeding the 3.44 threshold of significance ($\alpha = 0.05$) defined by permutation analysis.

Based on the physiological differences between membranous and muscular VSDs, linkage scans were also conducted on these phenotypes separately. Removal of the muscular VSDs from the membranous data results in a reduction in false

positives for any main effect loci that are not associated with muscular VSD risk. The membranous VSD specific scan showed increases in LOD scores for loci on chromosomes 8 and 10 despite the reduction of sample size of affected mice (Figure 3.2). The increase is evidence of distinct genetic etiologies of these two defects specifically suggesting that loci on chromosomes 8 and 10 are only important for membranous VSD risk and not muscular. In contrast, the chromosome 6 candidate peak showed a modest decrease in significance suggesting that this peak may be a shared risk factor for the two types of VSDs. The muscular VSD phenotype showed considerably reduced signals without any genomic regions exceeding the α threshold of 0.05 as well as greater background noise (Figure 3.3). A reduced sample size (n=80 affecteds) probably is a major contributor to the lack of significance in this scan. However, the data may also be indicative of a more genetically heterogeneous risk profile in which the presence of a large number of small effect genes directs penetrance of the phenotype. The peak on chromosome 6 appears to be common between the two types of VSDs indicating some commonality in the disease pathways of the two defects.

Inheritance patterns of each significant peak show risk alleles originating from both C57Bl/6 and FVB/n strains as hypothesized in Winston et al through segregation analysis on this data (Figure 3.4)[8]. Both chromosome 8 and 10 follow a semi recessive inheritance which is consistent with the lack of VSDs in F1 hybrid mice. Chromosome 6 shows the risk allele being inherited semi-recessively from the C57Bl/6 parent were as chromosome 8 has the risk allele inherited semi-recessively

from the FVB/n parent. In the case of chromosome 10, the risk allele is inherited from the C57Bl/6 parent with no dominance effects detected.

Genome wide scans for the F2 C57Bl/6 x A/J mice with membranous VSDs yielded loci on chromosomes 3 and 4 that exceeded the genome wide $\alpha = 0.05$ threshold of significance (Figure 3.5). The effect plots show that both of these candidate regions have their risk allele being inherited from the A/J inbred line (Figure 3.6). In the case of the chromosome 3 peak there appears to be no dominance effects whereas for peak on chromosome 4, risk is inherited recessively. Both of these peaks are not present in the FVB/n F2 genome scans indicating that A/J mice have risk factors specific to their strain. Of particular interest are peaks present on chromosomes 8 and 10 in very close proximity to the significant loci on chromosomes 8 and 10 mapped in the FVB/n membranous VSD scan although in the A/J scan, they do not achieve genome wide significance. The A/J chromosome 8 and 10 loci also share similar inheritance patterns which is further evidence that these are truly common risk factors between FVB/n and A/J F2 intercross mice. Given the differences in VSD incidence between F2 FVB/n and A/J mice (Table 3.1) it is expected that at least some of the risk loci are strain specific as is the case here.

Of particular interest is the lack of major heart development genes within the three candidate gene intervals on chromosomes 6, 8 and 10. Only chromosome 8 has genes implicated in heart defect risk, Hand2 and Tll1, located very close to the region of peak LOD score. Hand2 has been implicated as being in a common pathway with Nkx2-5 and involved in ventricular septal development[13,14]. It has been reported that Tll1 is activated by Nkx2-5 and also important in cardiac septation[15,16]. Likely, the

genetic modifiers at the other two loci will prove to be genes that have novel relationships to the cardiac developmental pathway.

Non-heritable Risk Factors

Paternal and maternal ages and litter size were evaluated as potential risk factors for increased membranous VSD incidence in C57Bl/6 x FVB/n and A/J F2 intercross mice. Analysis of 2,400 F2 intercross heterozygous Nkx2-5$^{+/-}$ knockout mice collected showed a significant positive correlation for maternal age with membranous VSD defect incidence (P = 0.006, OR = 4.07) in the FVB/n intercross (Table 3.2). Litter size was found to negatively correlate with VSD risk in FVB/n (P= 0.002, OR=0.92) although the effect size is considerably smaller (Figure 3.7). These two non-heritable risk factors are significantly correlated (P < 0.001, R = -0.35) independently of VSD incidence since older mice have smaller litters, however this is adjusted for in the multiple logistic regression model. Paternal age did not show a significant correlation with membranous VSD risk. In the case of muscular VSD risk, there were considerably fewer affected mice but a significant maternal age affect was detected (P=0.03, OR=2.99).

A total of 1,370 A/J F2 intercross mice were analyzed for correlation with the three non-heritable risk factors. C57Bl/6 x A/J F2 intercross mice show a significant maternal age affect on membranous VSD risk (P=0.02, OR=3.96) but not muscular VSDs. Additionally, litter size was found to significantly affect muscular VSD incidence in these mice although with a marginal effect size (P=0.02, OR=1.14).

Interaction Scan

A two-way genome scan for interacting loci was performed to assess potential interactions between loci as might be expected for complex disease susceptibility. The interaction scan was only run on membranous VSD data since larger sample sizes are required to exceed the very stringent genome wide thresholds of significance. An initial genome wide scan was run in which every marker locus was tested in combination with every other marker locus (Figure 3.8). The threshold for significance was generated by 1000 permutations:

FullT	Fv1T	IntT	AddT	Av1T	OneT
8.51	**6.69**	**6.1**	**5.85**	**2.79**	**3.44**

The two-way genome wide scan yielded three pairs of loci that satisfied condition 1 (Table 3.3). Each pair consists of genomic regions that exceeded genome wide thresholds in the previous main effect scan, loci on chromosomes 6, 8 and 10, indicating that the effect of these regions may be enhanced by the genotypes of one another. The interactive LOD score for each pair did not exceed genome wide thresholds indicating that this may be a purely additive or synergistic effect and likely not a true case of epistasis. These locus pairs also satisfied the condition 2 threshold confirming at least an additive interaction between them.

To further check whether yet undetected loci are interacting with the main effect regions mapped on chromosomes 6, 8 and 10, each locus was tested as an additive and then interactive covariate in order to compare the results of the two scans. There were no major jumps in LOD scores between these scans which supports the conclusion that in this model, no epistatic interactions have been detected (data not included).

Covariate addition

Since maternal age is a risk factor for VSD in FVB/n F2 intercross mice, genome-wide linkage scans were run including maternal age as an additive and then as an interactive covariate. The additive covariate accounts for residual variance in the population of mice due to maternal age. A negligible increase in power resulted from this adjustment. Maternal age was also included as an interactive covariate to scan for regions in the genome that are specifically affected by variance in maternal age. No significant jumps in LOD score were detected. Inclusion of litter size as an additive covariate showed marginal improvement in LODs scores of peak regions. There was no evidence of litter size specific interactions when litter size was added to the model as an interactive term (data not included).

Discussion

The pleiotropy, genetic heterogeneity, variable expressivity and incomplete penetrance of congenital heart disease have been described exhaustively in the literature as major hurdles in the characterization of risk [1,3,17]. Individual risk is difficult to assess even in the presence of a known mutation segregating in a family or a knockout in an inbred animal strain due to yet undefined risk factors. While Nkx2-5 undoubtedly plays a major role in heart development [18-20], haploinsufficiency of the gene is not sufficient to cause disease. In the Nkx2-5 knockout inbred mouse model, it has been shown that the disease phenotype is heavily influenced by genetic background. Each of the three strains studied C57Bl/6, FVB/n and A/J carry susceptibility alleles at different loci for ventricular septal defects demonstrating the influence of alternative genomic variability [8]. Additionally, non-heritable sources of variation such as litter size and maternal age further modify risk but only in specific genetic backgrounds. Understanding why individuals with known predisposing factors do not manifest the disease may provide insight into therapeutics that can circumvent risk pathways[21].

Variability within the VSD Phenotype

Congenital heart disease is a heterogeneous group of malformations in which clinical taxonomy is not necessarily based on developmental etiology [22]. Intellectually there is a disconnect between the fields of cardiac development and congenital heart disease which contrasts with fields like cancer biology and oncology or basic and clinical immunology where one can feasibly envision how a discovery in

the lab could be practically relevant to their respective medical fields. In this study we have presented evidence that muscular and membranous VSD subtypes have distinct genetic etiologies in this risk model. VSDs are generally described based on physiologic location in the heart and not developmental etiology. Membranous defects are found at the point where the septum reaches just under the aorta and pulmonary arteries. In cross section, the defect shows a septum that takes on a characteristic clean and rounded shape where the hole forms. In contrast, muscular VSDs also called trabecular VSDs occur anywhere within the muscular portion of the ventricular septum and vary in appearance from smooth to jagged edged. Additionally, muscular VSDs often occur multiple times within one heart. While the etiologies of these two defects remain unclear, it has been hypothesized that muscular defects may be the result of incomplete proliferation of cardiomyocytes or excessive resorbtion of the muscular septal wall during ventricular remodeling [23,24]. This contrasts with membranous VSDs which are more likely the result of insufficient growth of the ventricular septal wall as it invaginates towards the endocardial cushion during development.

The results of the linkage scans give evidence to support separate developmental pathways and thus distinct genetic etiologies of these two defects. However, there is potentially a common risk factor on chromosome 6. Given that Nkx2-5 increases risk for both types of phenotypes, it is not unexpected that another gene may direct a heart towards risk for either or both subtypes of VSDs. Fine mapping regions of interest in the muscular VSD scan may reveal genes involved in programmed cell death and remodeling whereas, the membranous candidates may

reveal growth factors and genes that regulate them. The noisier appearance and lack of significance of loci in the muscular scan may be due to a more heterogeneous genetic pathway which is consistent with the increased variability observed in the cross sections. However dissimilar these VSD subtypes may be in their origin, both are considerably more likely to occur in an Nkx2-5 haploinsufficient individual indicating that there is some commonality in their disease pathways. The disruption caused by the knockout mutation is adequate to unlock the importance of underlying buffering genetic variation that directs towards a particular disease phenotype. In families with a segregating mutation in a major cardiac gene, risk may be governed by a number of common variants that offer alternative pathways around the initial disruption.

Variability within Membranous VSD Risk

Genome wide linkage mapped three significant risk loci for the membranous VSD phenotype in the presence of Nkx2-5 haploinsufficiency. The presence of multiple main effect modifiers is consistent with the polygenic models that have been predicted in our lab and others for congenital heart malformations. We have reported evidence through segregation analysis for at least 2 FVB/n specific risk alleles and 2 C57Bl/6 risk alleles which is further corroborated by the linkage scan results [8]. In an inbred mouse model in which ~99% of the alleles are fixed, it is intriguing that polymorphisms between strains that have only recently diverged and come from a limited gene pool have a substantial effect on VSD risk. These strain specific differences are evidence that a greater source of underlying genetic variation is

present in natural populations that is only subject to selection in the presence of a major failure in the developmental pathway such as a major gene knockout or drastic environmental change. An old system such as the heart is ideal for this variation structure as it has been subject to many generations of stabilizing selection [25].

The Effect of Non heritable risk Factors

The concept of context specific genetic risk factors that are only important under certain genetic or environmental circumstances is consistent with the lack of reproducibility in many large genome association studies [25]. The effect of maternal age in the FVB/n F2 intercross and A/J exemplifies that the normally robust cardiac developmental pathway can become vulnerable to non-heritable factors in the presence of a major haploinsufficiency and the right (or wrong) genetic background.

Evolutionary Benefits of Variability

The Nkx2-5 heterozygous knockout model serves as an example of how a major perturbation to the normally robust cardiac development pathway can result predominantly in a healthy phenotype. The mosaic backgrounds of the F2 intercross mice are a closed pool of variation in which bottled up genetic potential serves as the major governing force of risk. G. Gibson described this as the effect of cryptic genetic variation which can be present in any well canalized system which is evolutionarily old and has experienced many generations of stabilizing selection [5]. We propose that the complexity of congenital heart disease is the result of important underlying variability that is selectively beneficial in populations. Alleles of modifier genes have the potential to increase risk or protect against it as well as direct presentation. A

system in which underlying variation that is potentially protective is better equipped to maintain developmental homeostasis than a single-variant determined system [26]. The fitness of the species is enhanced by the range of options presented by underlying variation in modifier genes that normally have little or no detectable phenotypic effect. A major perturbation of the system however can reveal an entire distribution of important risk modifying alleles. This utilitarian mechanism provides a robust developmental system to the majority while predisposing a minority to unfortunate susceptibilities. In this study we have identified multiple sources of variability in a simple model that any one of which represents an interesting pathway for scientists to study or for clinicians to target.

Reference List

(1) Andelfinger G. Genetic factors in congenital heart malformation. *Clin Genet.* 2008;73:516-527.

(2) Hoffman JI, Kaplan S. The incidence of congenital heart disease. *J Am Coll Cardiol.* 2002;39:1890-1900.

(3) Nora JJ, Nora AH. Maternal transmission of congenital heart diseases: new recurrence risk figures and the questions of cytoplasmic inheritance and vulnerability to teratogens. *Am J Cardiol.* 1987;59:459-463.

(4) Olson EN. Gene regulatory networks in the evolution and development of the heart. *Science.* 2006;313:1922-1927.

(5) Gibson G, Reed LK. Cryptic genetic variation. *Curr Biol.* 2008;18:R989-R990.

(6) Nadeau JH. Modifier genes in mice and humans. *Nat Rev Genet.* 2001;2:165-174.

(7) Tanaka M, Chen Z, Bartunkova S, Yamasaki N, Izumo S. The cardiac homeobox gene Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. *Development.* 1999;126:1269-1280.

(8) Winston JB, Erlich JM, Green CA et al. Heterogeneity of Genetic Modifiers Ensures Normal Cardiac Development. *Circulation.* 2010.

(9) Sen S, Churchill GA. A statistical framework for quantitative trait mapping. *Genetics.* 2001;159:371-387.

(10) Xu S. Theoretical basis of the Beavis effect. *Genetics.* 2003;165:2259-2268.

(11) Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses. *Bioinformatics.* 2003;19:889-890.

(12) Broman, K. W. and Sen, S. A Guide to QTL Mapping with R/qtl: Online complements. 7-20-2009. Springer.
Ref Type: Generic

(13) Togi K, Yoshida Y, Matsumae H, Nakashima Y, Kita T, Tanaka M. Essential role of Hand2 in interventricular septum formation and trabeculation during cardiac development. *Biochem Biophys Res Commun.* 2006;343:144-151.

(14) Yamagishi H, Yamagishi C, Nakagawa O, Harvey RP, Olson EN, Srivastava D. The combinatorial activities of Nkx2.5 and dHAND are essential for cardiac ventricle formation. *Dev Biol.* 2001;239:190-203.

(15) Clark TG, Conway SJ, Scott IC et al. The mammalian Tolloid-like 1 gene, Tll1, is necessary for normal septation and positioning of the heart. *Development.* 1999;126:2631-2642.

(16) Sabirzhanova I, Sabirzhanov B, Bjordahl J, Brandt J, Jay PY, Clark TG. Activation of Tolloid-like 1 gene expression by the cardiac specific homeobox gene Nkx2-5. *Dev Growth Differ.* 2009;51:403-410.

(17) Gruber PJ, Epstein JA. Development gone awry: congenital heart disease. *Circ Res.* 2004;94:273-283.

(18) Elliott DA, Kirk EP, Yeoh T et al. Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J Am Coll Cardiol.* 2003;41:2072-2076.

(19) Biben C, Weber R, Kesteven S et al. Cardiac septal and valvular dysmorphogenesis in mice heterozygous for mutations in the homeobox gene Nkx2-5. *Circ Res.* 2000;87:888-895.

(20) Jay PY, Rozhitskaya O, Tarnavski O et al. Haploinsufficiency of the cardiac transcription factor Nkx2-5 variably affects the expression of putative target genes. *FASEB J.* 2005;19:1495-1497.

(21) Nadeau JH. Modifier genes in mice and humans. *Nat Rev Genet.* 2001;2:165-174.

(22) Benson DW, Martin LJ. Complex story of the genetic origins of pediatric heart disease. *Circulation.* 2010;121:1277-1279.

(23) Ransom J, Srivastava D. The genetics of cardiac birth defects. *Semin Cell Dev Biol.* 2007;18:132-139.

(24) Fisher SA, Langille BL, Srivastava D. Apoptosis during cardiovascular development. *Circ Res.* 2000;87:856-864.

(25) Gibson G. Decanalization and the origin of complex disease. *Nat Rev Genet.* 2009;10:134-140.

(26) Rutherford SL, Lindquist S. Hsp90 as a capacitor for morphological evolution. *Nature.* 1998;396:336-342.

Figures:

Figure 3.1. Genome wide linkage scan result to map VSD risk loci in FVB/n F2 intercross mice. A total of 306 mice with VSDs and 284 controls were included in this scan.

Figure 3.2. Genome wide linkage scan result to map membranous VSD specific risk loci in FVB/n F2 intercross mice. A total of 251 mice with membranous VSDs and 284 controls were included in this scan.

Figure 3.3. Genome wide linkage scan result to map muscular VSD specific risk loci in FVB/n F2 intercross mice. A total of 80 mice with muscular VSDs and 284 controls were included in this scan.

Figure 3.4. Effect plots for significant loci mapped in the FVB/n F2 intercross membranous VSD linkage scan. C/C = homozygous for C57Bl/6 alleles, C/F = heterozygous for C57Bl6 and FVB/n alleles and F/F = homozygous for FVB/n alleles. The y-axis represents that incidence of mice with membranous VSDs that have the respective genotype indicated on the x-axis. Error bars are derived from the imputed genotypes in interval mapping analysis.

Figure 3.5. Genome wide linkage scan result to map VSD risk loci in A/J F2 intercross mice. A total of 87 mice with membranous VSDs and 116 controls were included in this scan.

Figure 3.6. Effect plots for significant loci mapped in the A/J F2 intercross membranous VSD linkage scan. C/C = homozygous for C57Bl/6 alleles, C/A = heterozygous for C57Bl6 and A/J alleles and A/A = homozygous for A/J alleles. The y-axis represents that incidence of mice with membranous VSDs that have the respective genotype indicated on the x-axis. Error bars are derived from the imputed genotypes in interval mapping analysis.

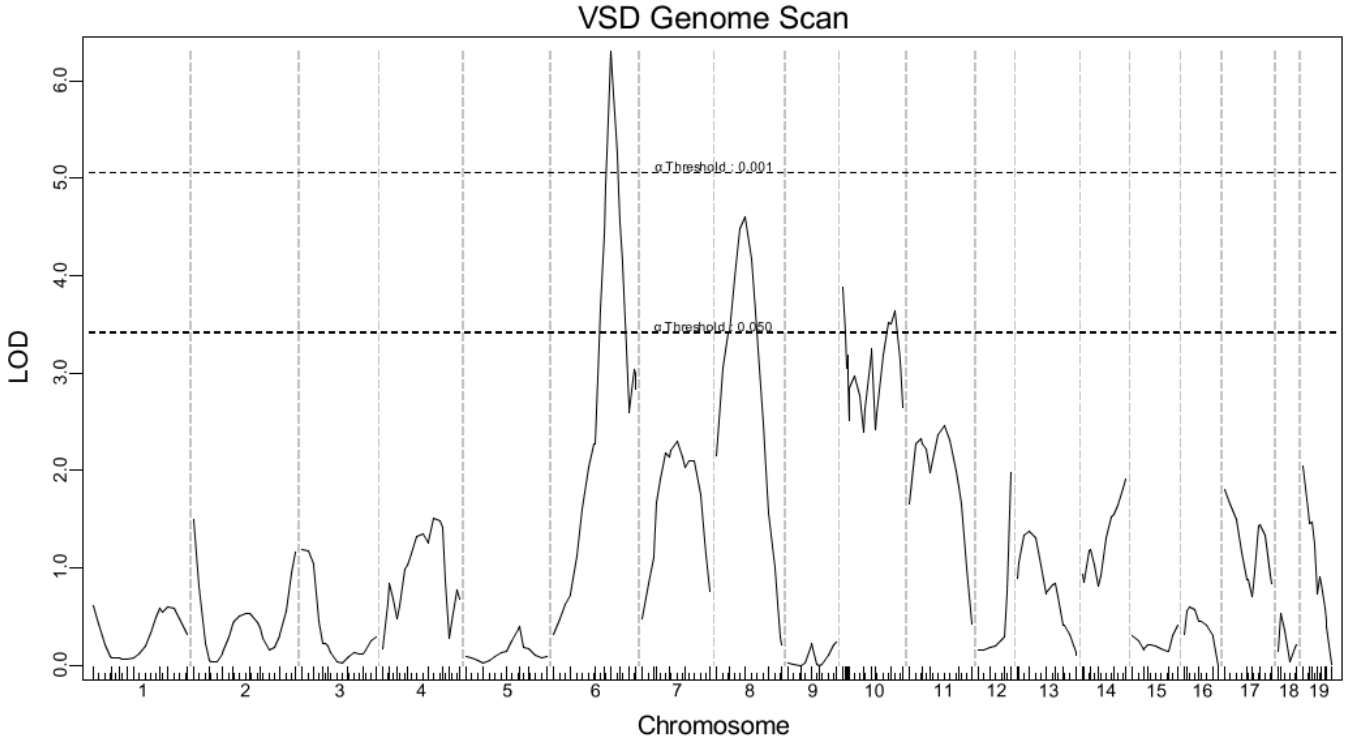
Figure 3.7. Incidence of VSD is plotted against maternal age (days) and litter size.

Figure 3.8. Two-way scan for genetic interactions that correlate with membranous VSD risk. The x-axis includes every marker location tested and the resulting full model LOD scores color coded according the right hand key. The y-axis includes every marker location tested and the resulting epistatic LOD scores color coded according to the left hand side of the key.

Table 3.1. Incidence of VSDs in inbred mouse crosses

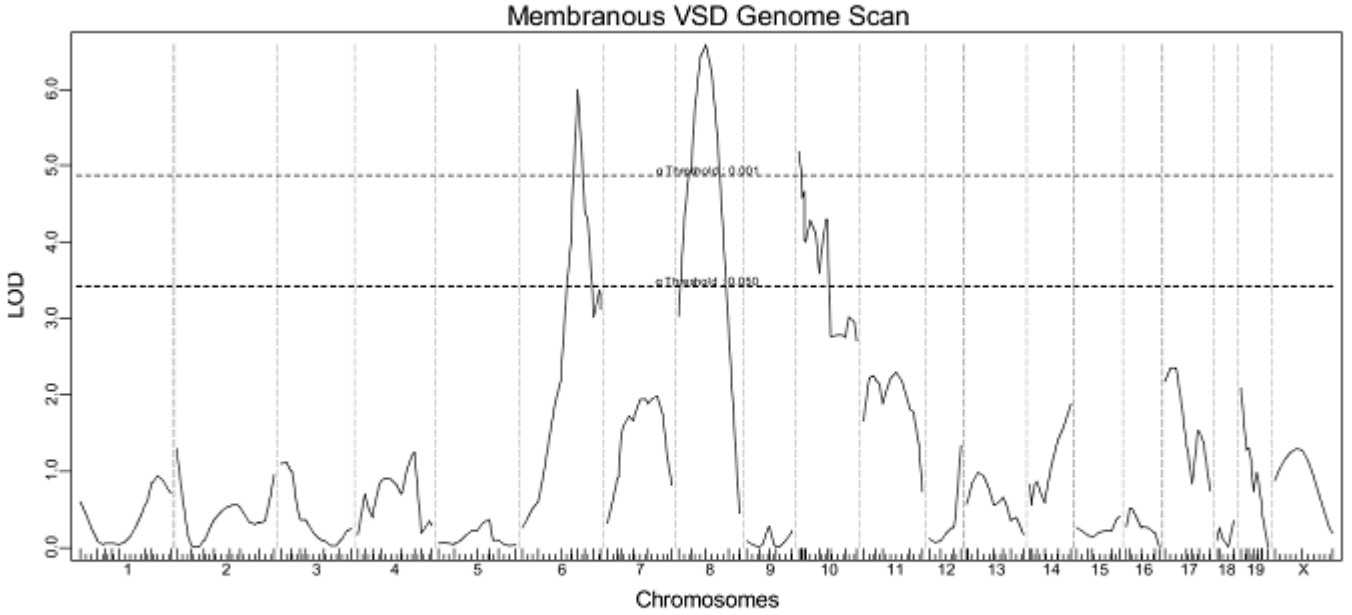
	VSD Incidence	
Cross	Membranous	Muscular
C57 Parent	7.5%	27.1%
F1 FVB Hybrid	1.3%	0.0%
F1 AJ Hybrid	0.0%	0.0%
F2 FVB Intercross	9.7%	3.2%
F2 AJ Intercross	8.2%	3.4%

Figure 3.1



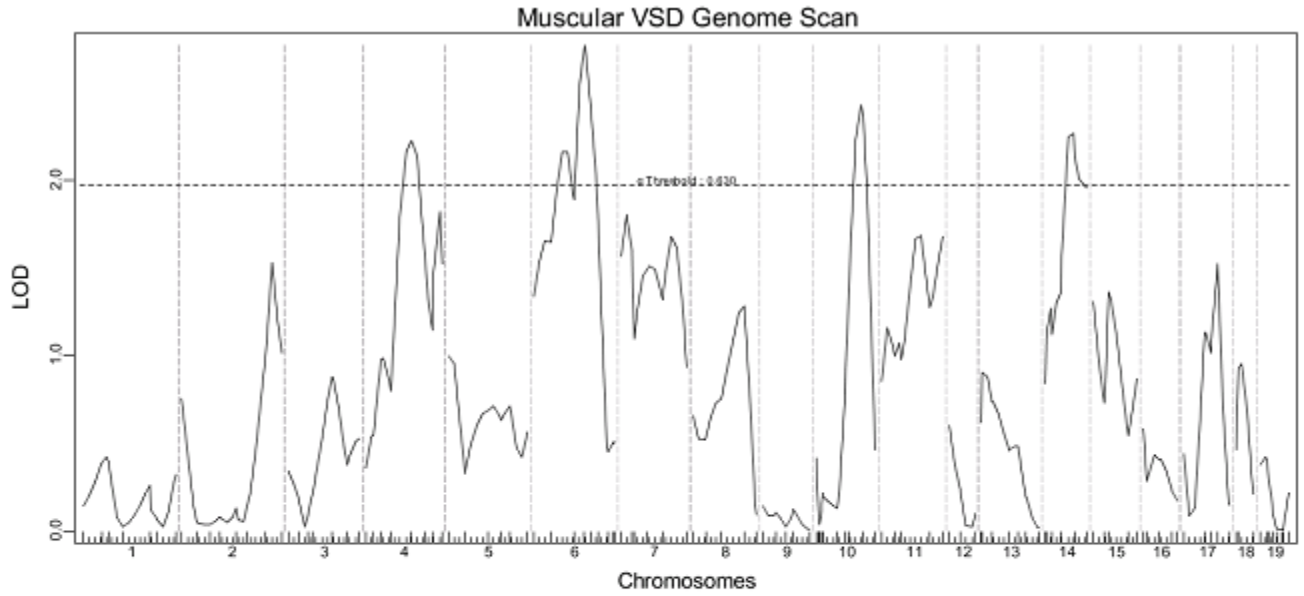
Chrom6 LOD=6.29	Actual Location	Lower	Upper	Chrom8 LOD=4.60	Actual Location	Lower	Upper	Chrom10 LOD=3.88	Actual Location	Lower	Upper
cM	52.3	52.3	57.3	cM	32.3	22.4	42.4	cM	4.5	4.5	56.9
Mb	104,600,000	104,600,000	114,600,000	Mb	64,600,000	44,800,000	84,800,000	Mb	9,000,000	9,000,000	113,800,000

Figure 3.2



Chrom6 LOD=5.996	*Actual Location*	*Lower*	*Upper*	*Chrom8 LOD=6.95*	*Actual Location*	*Lower*	*Upper*	*Chrom10 LOD=5.188*	*Actual Location*	*Lower*	*Upper*
cM	52.3	47.3	57.3	cM	30.1	22.4	37.4	cM	4.5	4.5	29.9
Mb	104,600,000	94,600,000	114,600,000	Mb	60,200,000	44,800,000	74,800,000	Mb	9,000,000	9,000,000	59,800,000

Figure 3.3



Chrom4 LOD=2.22	*Actual Location*	*Lower*	*Upper*	*Chrom6 LOD= 2.76*	*Actual Location*	*Lower*	*Upper*
cM	49.8	34.8	76.9	cM	47.3	12.3	59.2
Mb	99,600,000	69,600,000	153,800,000	Mb	94,600,000	24,600,000	118,400,000

Chrom10 LOD=2.42	*Actual Location*	*Lower*	*Upper*	*Chrom14 LOD=2.26*	*Actual Location*	*Lower*	*Upper*
cM	44.5	34.5	49.5	cM	34.2	19.2	46.4
Mb	89,000,000	69,000,000	99,000,000	Mb	68,400,000	38,400,000	92,800,000

Figure 3.4

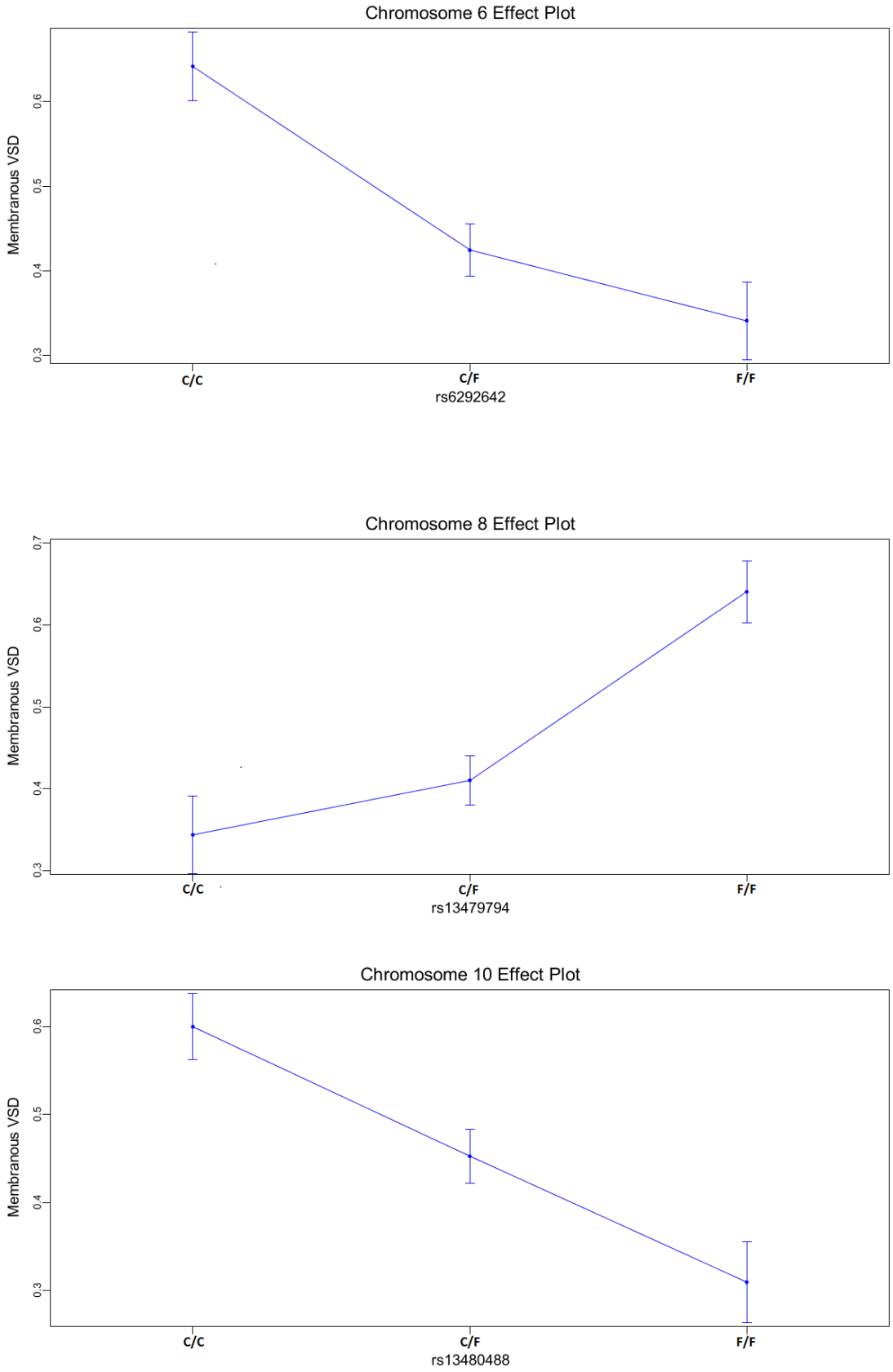
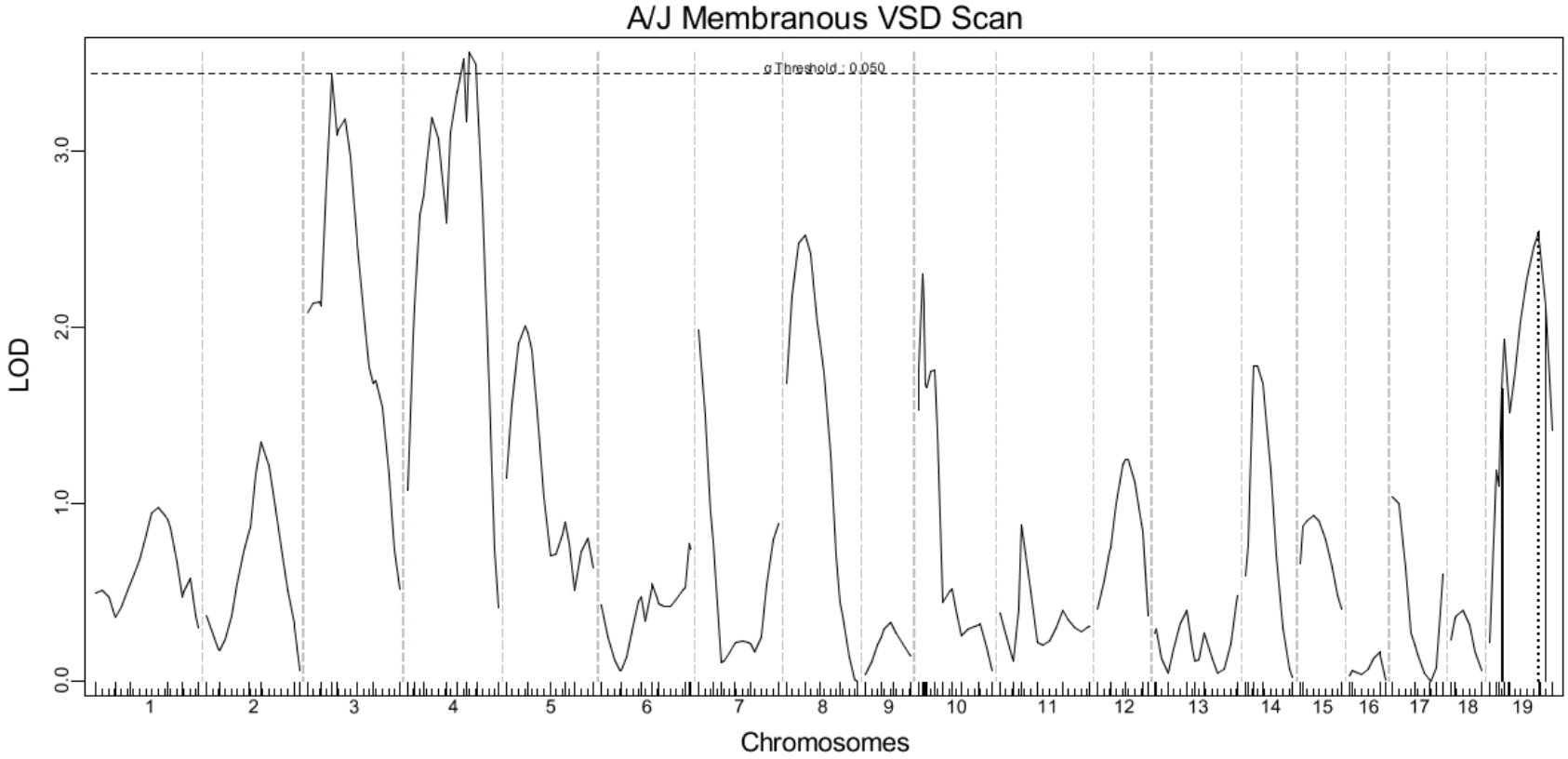

Chromosome 6 Effect Plot
Membranous VSD
0.3
0.4
0.5
0.6
C/C
C/F
F/F
rs6292642
Chromosome 8 Effect Plot
Membranous VSD
0.3
0.4
0.5
0.6
0.7
C/C
C/F
F/F
rs13479794
Chromosome 10 Effect Plot
Membranous VSD
0.3
0.4
0.5
0.6
C/C
C/F
F/F
rs13480488

Figure 3.5



Chrom3 LOD=3.43	Actual Location	Lower	Upper	Chrom4 LOD=3.55	Actual Location	Lower	Upper
cM	33.8	18.8	54	cM	59.8	23.1	69.8
Mb	67,600,000	37,600,000	108,000,000	Mb	119,600,000	46,200,000	139,600,000

Figure 3.6

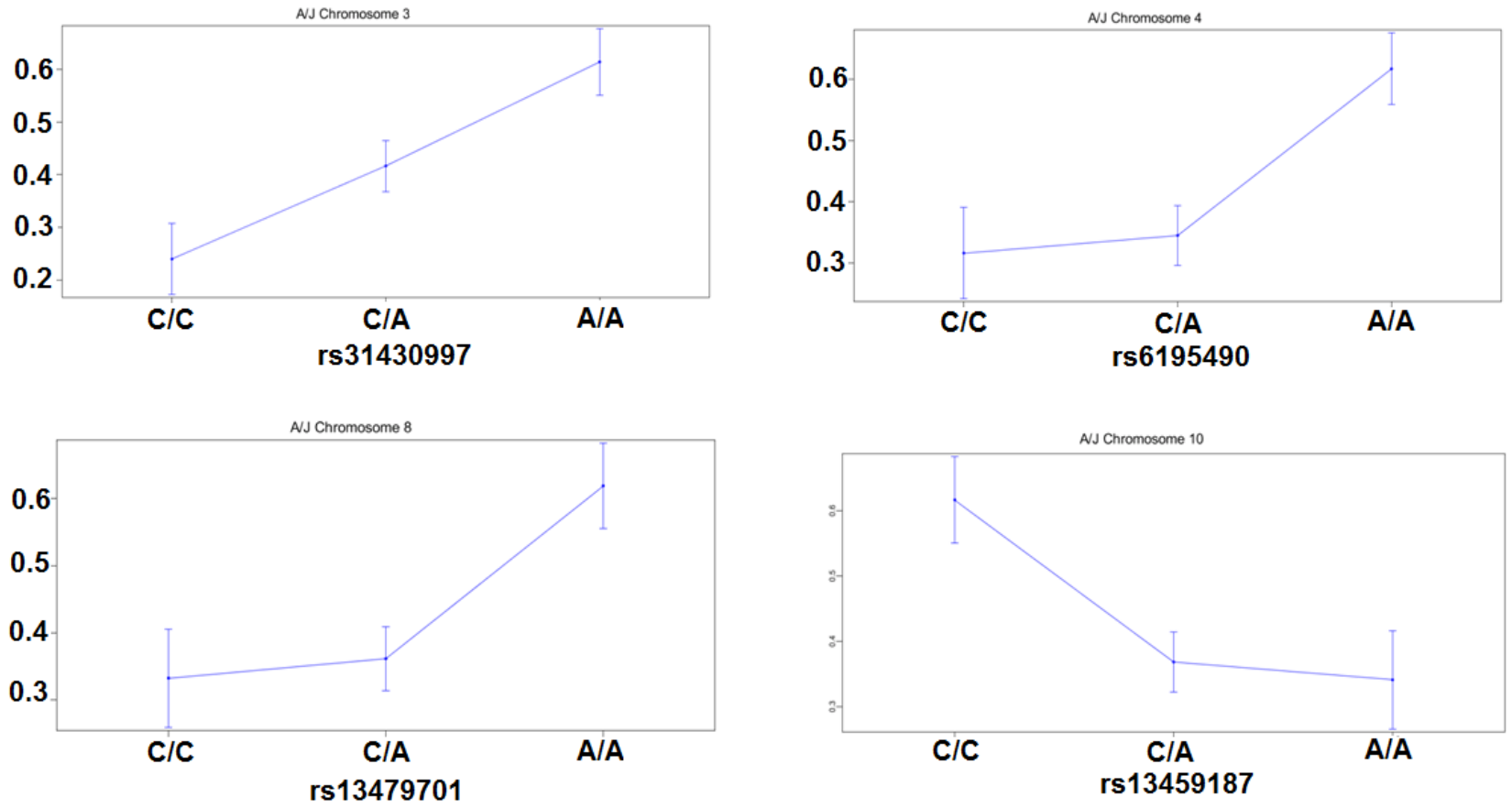
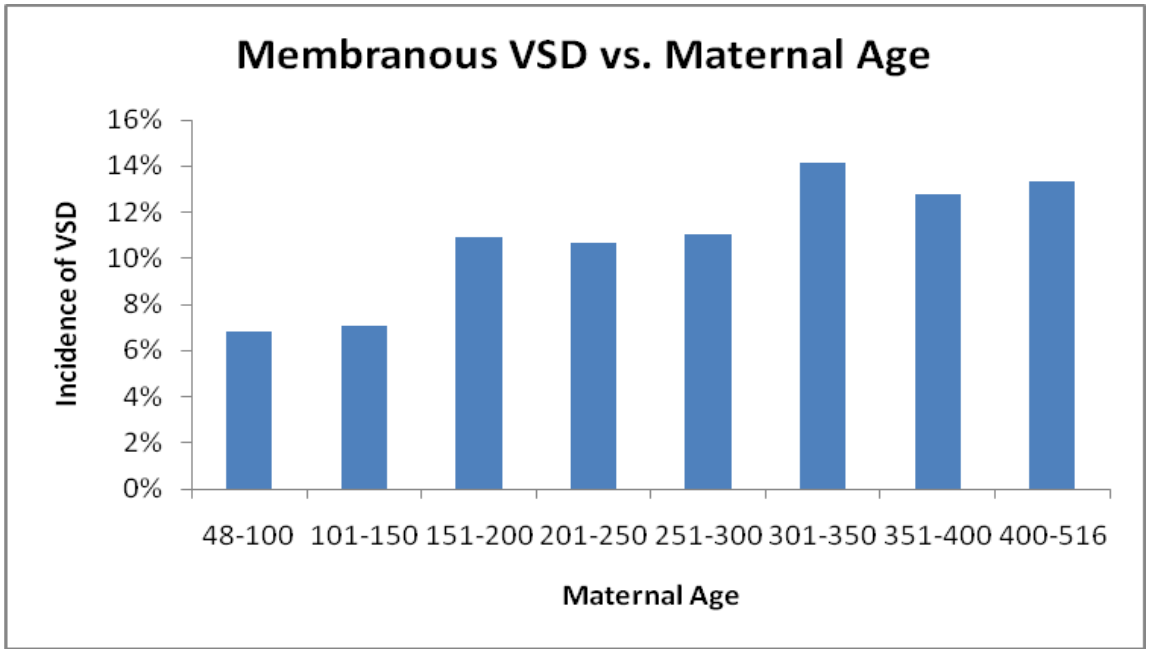

A/J Chromosome 3
0.6
0.5
0.4
0.3
0.2
C/C
C/A
A/A
rs31430997
A/J Chromosome 4
0.6
0.5
0.4
0.3
C/C
C/A
A/A
rs6195490
A/J Chromosome 8
0.6
0.5
0.4
0.3
C/C
C/A
A/A
rs13479701
A/J Chromosome 10
C/C
C/A
A/A
rs13459187

Table 3.2. Regression of Non-heritable predictors and VSD phenotypes

F2 Intercross		**Membranous VSD**			**Muscular VSD**		
		P-value	*OR*	*95% CI*	*P-value*	*OR*	*95% CI*
FVB/n	Mother	**0.006**	4.07	1.52-11.42	**0.030**	2.99	1.12-8.18
	Father	0.097	0.45	0.17-1.12	0.642	1.24	0.50-3.03
	Litter	**0.002**	0.92	0.86-0.96	0.890	1.01	0.93-1.07
A/J	Mother	**0.021**	3.96	1.26-13.18	0.383	0.53	0.13-2.27
	Father	0.131	2.23	0.79-6.44	0.240	0.45	0.11-1.74
	Litter	0.636	0.98	0.89-1.07	**0.022**	1.14	1.07-1.28

* P-value adjusted for correlation with other significant predictors

Figure 3.7


Membranous VSD vs. Maternal Age
Incidence of VSD
16%
14%
12%
10%
8%
6%
4%
2%
0%
48-100
101-150
151-200
201-250
251-300
301-350
351-400
400-516
Maternal Age

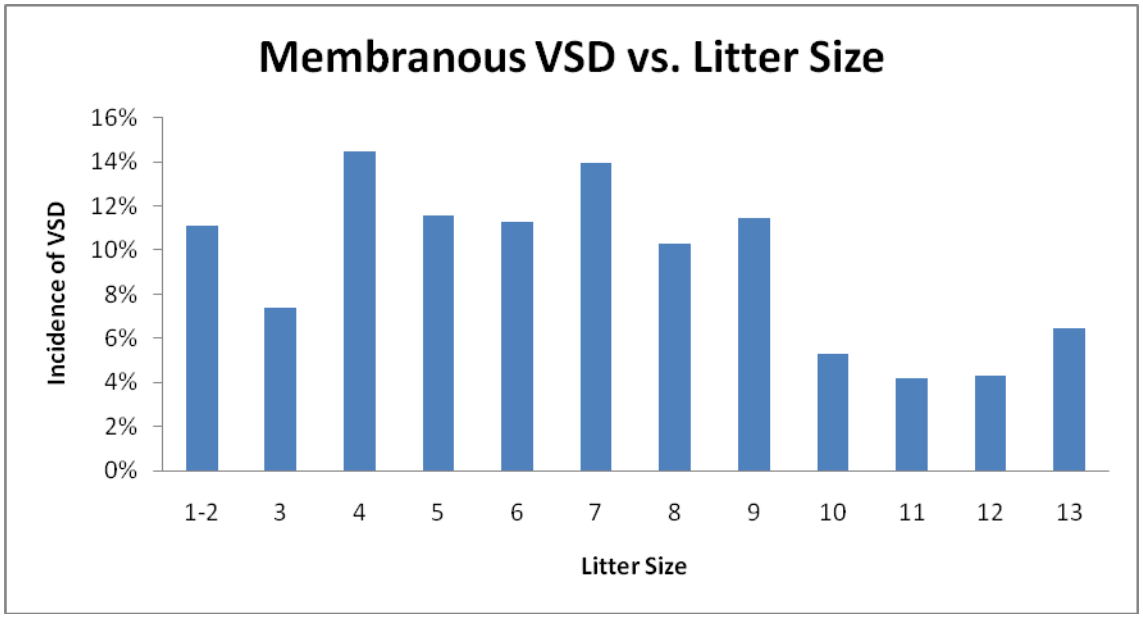

Membranous VSD vs. Litter Size
Incidence of VSD
16%
14%
12%
10%
8%
6%
4%
2%
0%
1-2
3
4
5
6
7
8
9
10
11
12
13
Litter Size

Figure 3.8

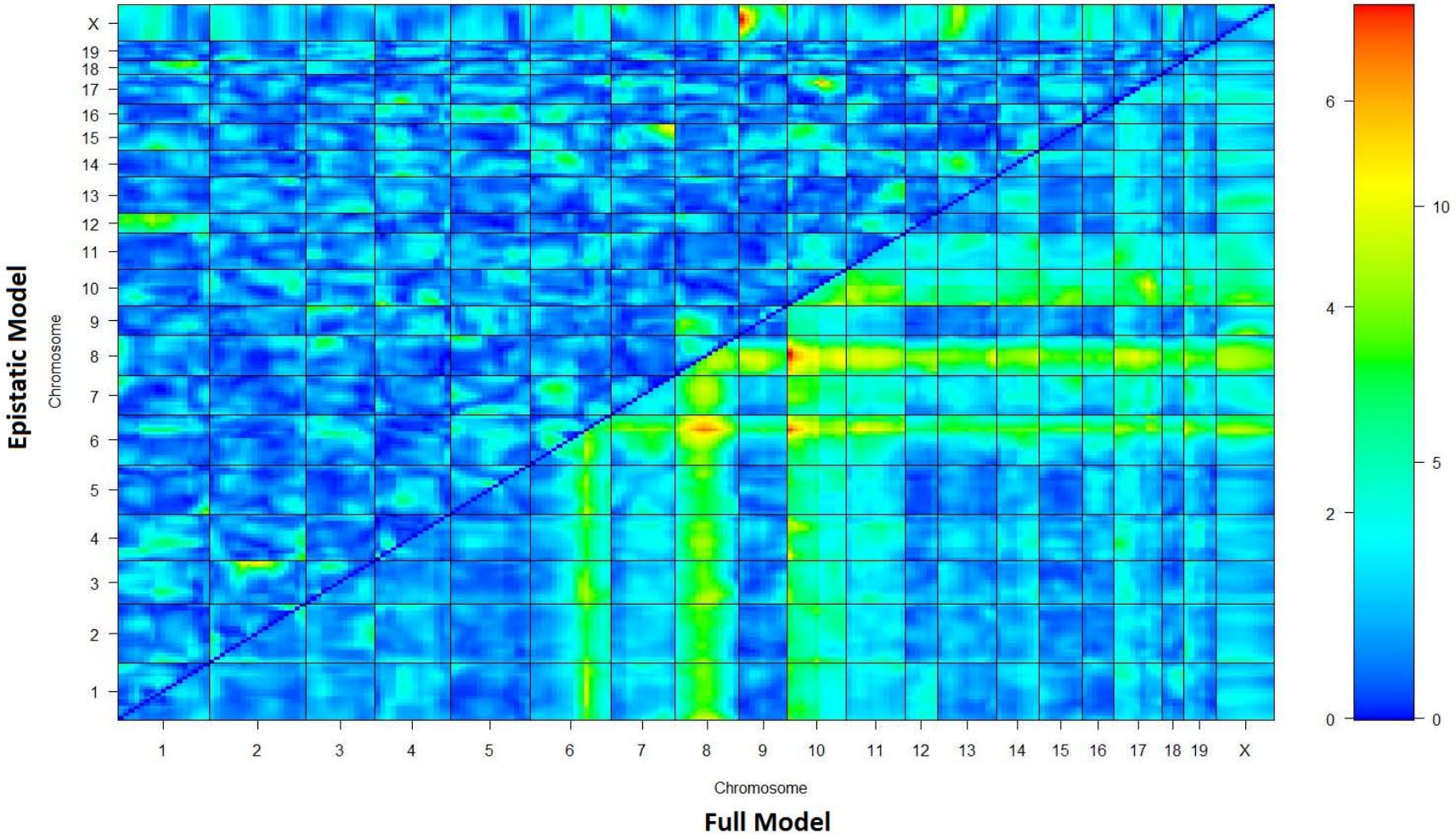

Epistatic Model
Chromosome
X
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
Chromosome
Full Model
6
4
2
0
10
5
0

Table 3.3. Two-way scan candidate regions that exceeded threshold conditions.

Condition 1

Epistatic Model				
Locus 1 (cM)	**Locus 2 (cM)**	**Full**	**Fv1**	**Int**
Chr8@37.4	Chr10@4.5	13.901	7.324	2.378
Chr6@52.3	Chr10@4.5	13.509	7.513	1.689
Chr6@52.3	Chr8@32.4	12.926	6.35	0.254

Condition 2

Additive Model			
Locus 1 (cM)	**Locus 2 (cM)**	**Add**	**Av1**
Chr6@52.3	Chr8@32.4	12.672	6.095
Chr8@32.4	Chr10@4.5	11.875	5.299
Chr6@52.3	Chr10@4.5	11.821	5.824

Chapter 4: Conclusion and Future Directions

Genomic Mapping of Genetic Modifiers

In this study, an Nkx2-5 haploinsufficient mouse has served as a risk model for the phenotypic variability of congenital heart disease risk that is commonly observed in human studies. Through the examination of the incidences of congenital heart defects in C57Bl/6 heterozygous knockout mice outcrossed to FVB/n and A/J as described in **Chapter 1**, it has been determined that while Nkx2-5 mutation is a major risk factor for heart defects, genetic background is crucial for directing penetrance and presentation of those defects. Segregation analysis has implicated three or more modifier loci and at least one epistatic interaction that influence susceptibility to muscular VSD and ASD in *Nkx2-5*$^{+/-}$ animals. At least two loci influence membranous VSD susceptibility (**Chapter 2**).

Genomic mapping using F2 intercross strains has proven to be an effective method for detecting and mapping candidate risk loci and interactions[1]. Significant regions were mapped for membranous VSD risk on chromosomes 6, 8 and 10. Additionally, the genome scan for muscular VSD risk loci revealed a potentially common risk factor on chromosome 6 as well as a number of completely different suggestive candidate regions elsewhere in the genome. Unfortunately, a caveat of using F2 generation outcrosses for gene mapping is that the candidate intervals are rather large (20-40Mb) because there are only two generations of recombination. Lists of genes located in these regions can be narrowed based on gene functionality and expression data.

However, there is a risk that using filters based on known information will rule out small effect genes that have novel roles in heart development. Given that only one of the significant regions (chromosome 8) yielded genes that have been previously implicated in heart development, this seems like a valid concern.

Also noted is the potential for multiple QTLs on a single chromosome. The multiple peaks on chromosome 10 likely represent two and possibly three separate candidate regions. It is possible that the middle peak in the 30cM region may simply be an artifact of the two surrounding peaks being in linkage disequilibrium, however finer resolution would be preferred over speculation.

Fine Mapping Strategy

An important future direction for this research project should include implementing a strategy to narrow QTL intervals and resolve multiple peaks in order to facilitate the identification of risk affecting loci. Fine-mapping with advanced intercross lines (AIL) has been shown to be an effective method to accomplish this goal [2-4]. AILs are designed to increase the number of informative meioses in the mapping population by continued intercrossing of a population to reduce linkage disequilibrium and cause the proportion of recombinants between linked loci to asymptotically approach 0.5 [5,6]. Animals from the F2 generation are continually intercrossed through successive generations. The breeding is controlled so that siblings and first cousins are not mated to avoid excessive inbreeding. Below is a schematic describing a semi-circular breeding scheme based on the circular AIL screen described by Kimura and Crow[7]:

Figure 4.1

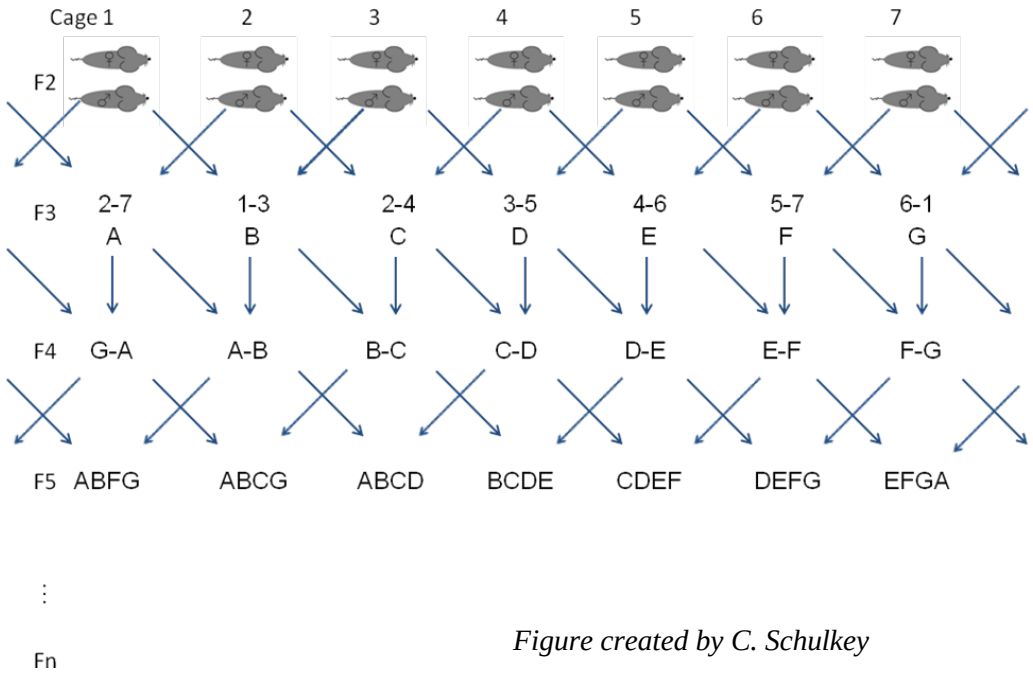


Figure created by C. Schulkey

Animals from the F10 generation will be phenotyped and affected animals will be genotyped for SNPs spaced 1-2 Mb across the genome to get high resolution mapping information[8]. Linkage analysis on this new population will likely reveal considerably greater resolution around candidate regions as up to a 5 fold reduction in confidence interval compared to the F2 map has been reported using this method[9]. It is expected that fine mapping with AILs will take 2-3 years for the crosses to be complete and the adequate number of mice to be collected and diagnosed. At that point candidate regions will be narrowed enough to allow for the search for functional variants. As of January of 2010, A/J has been sequenced with 23x coverage by the Mouse Genome Sequencing Consortium at the Sanger Institute. While the FVB/n strain is not listed as a priority for this consortium, hopefully it will be the next in line of mouse genomes to be sequenced and therefore complete when the AIL analysis is. Clearly, direct alignment within candidate regions would be an excellent source of information especially in non-coding

regions. Other new avenues to explore include the potential for microRNAs being a source of functional variance. The miRBase is continually being populated with more information which will be particularly relevant when the candidate gene regions are more finely mapped.

Further examination of Maternal Age Effect

The collection and diagnosis of 2,400 FVB/n and 1,370 A/J intercross mice at high risk for congenital heart defects allowed for the unique opportunity to conduct an epidemiological experiment in which a number of environmental factors could be controlled for. For example, mice were kept in the same physical environment with the same chow diet their whole lives. On the other hand, some non-heritable variables were not controlled for such as maternal and paternal age and litter size. These variables were recorded and later tested in a regression model as predictors of VSD risk. A significant effect on membranous VSD risk was detected for maternal age and litter size in both crosses.

Evidence for a maternal age affect in this congenital heart disease risk model is an exciting find as it mimics findings in epidemiologic studies on human populations. It is well documented that with advanced maternal age comes the increased risk for chromosomal abnormalities in the offspring [10]. Most notably, these copy number variants (CNVs) are responsible for an increased risk of Down's syndrome which is strongly associated with congenital heart defects[11]. Therefore, one important way that the maternal

age effect should be explored further is to check the F2 mice for increased CNVs in pups with older mothers.

However, numerous studies have also been published reporting that incidence of heart defects increase with advanced maternal age (35 yrs+) even when controlling for chromosomal abnormalities such as Down's syndrome[12-14]. Given that increased risk with advanced maternal are may not be due to CNVs alone, the question arises as to whether increased risk is the result of a change in older oocytes or with the physiologic differences in the uterine environment that come with age. For example, it is clear that fecundity decreases with advanced age which in mice is reportedly due to increased resorption rates, morphological abnormalities of the embryos and delayed development and not because of reduced ovulation or implantation rates[15].

Potential changes in uterine environment include the effects of obesity and diabetes. C57Bl/6 x FVB/n mice experience clear weight gain as they age which can lead to changes in hormonal levels. Type II diabetes in these mothers could further affect uterine environment as it does in humans which show increased risk for birth defects[16]. It has been reported that C57Bl/6 mice are at risk for obesity and diabetes, however, A/J mice are only at risk for obesity and do not show significant insulin resistance[17]. The strain specific susceptibility to heart defects observed in this study would follow that pattern since C57Bl/6 x A/J F2 mice did not show a significant maternal age effect on incidence. Two potential experiments to address these hypotheses are 1) glucose tolerance testing and 2) high fat diet for young mothers. The first experiments would simply determine whether insulin insensitivity is increasing with age in our mice. Ideally,

this study could be done prospectively throughout the life of a group of female F1 hybrid mice, starting from weaning and ending with menopause. The second experiment would address the obesity factor by attempting to accelerate the age/ risk correlation. Are the younger, fatter mice having more pups with heart defects? Are they having pups with comparable incidences to their older counterparts when they weigh the same?

Studies on human fertility report the existence of an opposing effect where oocytes donated by young women to older women result in restored embryo implantation and pregnancy rates indicating the problem lies with the aged oocytes and not the uterine environment[18]. Therefore it seems important to evaluate pre-implantation and post-implantation changes as risk modifying factors. Experiments which separate the oocyte from its environment could potentially distinguish which factors are causing heart defects, the mother or the egg. Oocytes from young donor mothers would be transferred by *in vitro* fertilization procedures into older recipients (and young for control). Another experiment would involve actually transferring ovaries from old mice (and young for control) into young recipients. These experiments would be conducted in the Nkx2-5 heterozygous knockout background. F2 pups produced in this manner would be collected and diagnosed using standard procedures to determine the incidence of defects. Even if maternal age is secondary to the real effect underlying the increase in congenital heart defects, these experiments could potentially rule out a number of physiologic causes of risk to further understand the basis of variance in congenital heart disease.

Reference List

(1) Silver LM. *Mouse Genetics: Concepts and Applications*. New York: Oxford University Press; 1995.

(2) Darvasi A, Soller M. Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics.* 1995;141:1199-1207.

(3) Wang X, Le R, I, Nicodeme E et al. Using advanced intercross lines for high-resolution mapping of HDL cholesterol quantitative trait loci. *Genome Res.* 2003;13:1654-1664.

(4) Fawcett GL, Jarvis JP, Roseman CC, Wang B, Wolf JB, Cheverud JM. Fine-mapping of Obesity-related Quantitative Trait Loci in an F(9/10) Advanced Intercross Line. *Obesity (Silver Spring).* 2009.

(5) Rockman MV, Kruglyak L. Breeding designs for recombinant inbred advanced intercross lines. *Genetics.* 2008;179:1069-1078.

(6) Darvasi A, Soller M. Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics.* 1995;141:1199-1207.

(7) Motoo Kimura, James F.Crow. *Genetical Research.* 1963;4:399-415.

(8) Fawcett GL, Roseman CC, Jarvis JP, Wang B, Wolf JB, Cheverud JM. Genetic architecture of adiposity and organ weight using combined generation QTL analysis. *Obesity (Silver Spring).* 2008;16:1861-1868.

(9) Darvasi A, Soller M. Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics.* 1995;141:1199-1207.

(10) Kornafel KL, Sauer MV. Increased rates of aneuploidy in older women. Increased rates of aneuploidy do not occur in gestations of older embryo recipients. *Hum Reprod.* 1994;9:1981-1982.

(11) Marino B. Congenital heart disease in patients with Down's syndrome: anatomic and genetic aspects. *Biomed Pharmacother.* 1993;47:197-200.

(12) Reefhuis J, Honein MA. Maternal age and non-chromosomal birth defects, Atlanta--1968-2000: teenager or thirty-something, who is at risk? *Birth Defects Res A Clin Mol Teratol.* 2004;70:572-579.

(13) Hook EB. Risk of Down syndrome in relation to maternal age. *Lancet.* 1976;2:465.

(14) Hollier LM, Leveno KJ, Kelly MA, McIntire DD, Cunningham FG. Maternal age and malformations in singleton births. *Obstet Gynecol.* 2000;96:701-706.

(15) Lopes FL, Fortier AL, Darricarrere N, Chan D, Arnold DR, Trasler JM. Reproductive and epigenetic outcomes associated with aging mouse oocytes. *Hum Mol Genet.* 2009;18:2032-2044.

(16) Becerra JE, Khoury MJ, Cordero JF, Erickson JD. Diabetes mellitus during pregnancy and the risks for specific birth defects: a population-based case-control study. *Pediatrics.* 1990;85:1-9.

(17) Surwit RS, Kuhn CM, Cochrane C, McCubbin JA, Feinglos MN. Diet-induced type II diabetes in C57BL/6J mice. *Diabetes.* 1988;37:1163-1167.

(18) Sauer MV. The impact of age on reproductive potential: lessons learned from oocyte donation. *Maturitas.* 1998;30:221-225.