



HHS Public Access

Author manuscript

Hum Genet. Author manuscript; available in PMC 2015 September 01.

Published in final edited form as:

Hum Genet. 2014 September ; 133(9): 1127–1138. doi:10.1007/s00439-014-1453-1.

A genetic association study detects haplotypes associated with obstructive heart defects

Ming Li,

Department of Pediatrics, College of Medicine, University of Arkansas for Medical Sciences, 13 Children's Way Mail Slot 512-40, Little Rock, AR 72202, USA

Mario A. Cleves,

Department of Pediatrics, College of Medicine, University of Arkansas for Medical Sciences, 13 Children's Way Mail Slot 512-40, Little Rock, AR 72202, USA

Himel Mallick,

Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA

Stephen W. Erickson,

Department of Biostatistics, University of Arkansas for Medical Sciences, Little Rock, AR 72202, USA

Xinyu Tang,

Department of Pediatrics, College of Medicine, University of Arkansas for Medical Sciences, 13 Children's Way Mail Slot 512-40, Little Rock, AR 72202, USA

Todd G. Nick,

Department of Pediatrics, College of Medicine, University of Arkansas for Medical Sciences, 13 Children's Way Mail Slot 512-40, Little Rock, AR 72202, USA

Stewart L. Macleod,

Department of Pediatrics, College of Medicine, University of Arkansas for Medical Sciences, 13 Children's Way Mail Slot 512-40, Little Rock, AR 72202, USA

Charlotte A. Hobbs, and

Department of Pediatrics, College of Medicine, University of Arkansas for Medical Sciences, 13 Children's Way Mail Slot 512-40, Little Rock, AR 72202, USA

National Birth Defect Prevention Study

Ming Li: mli@uams.edu; Charlotte A. Hobbs: hobbscharlotte@uams.edu

Abstract

© Springer-Verlag Berlin Heidelberg 2014

Correspondence to: Charlotte A. Hobbs, hobbscharlotte@uams.edu.

Electronic supplementary material The online version of this article (doi:10.1007/s00439-014-1453-1) contains supplementary material, which is available to authorized users.

The contents are solely the responsibility of the authors and do not necessarily represent the official views of the Center of Disease Control and Prevention (CDC).

The development of congenital heart defects (CHDs) involves a complex interplay between genetic variants, epigenetic variants, and environmental exposures. Previous studies have suggested that susceptibility to CHDs is associated with maternal genotypes, fetal genotypes, and maternal–fetal genotype (MFG) interactions. We conducted a haplotype-based genetic association study of obstructive heart defects (OHDs), aiming to detect the genetic effects of 877 SNPs involved in the homocysteine, folate, and transsulfuration pathways. Genotypes were available for 285 mother–offspring pairs with OHD-affected pregnancies and 868 mother–offspring pairs with unaffected pregnancies. A penalized logistic regression model was applied with an adaptive least absolute shrinkage and selection operator (lasso), which dissects the maternal effect, fetal effect, and MFG interaction effects associated with OHDs. By examining the association between 140 haplotype blocks, we identified 9 blocks that are potentially associated with OHD occurrence. Four haplotype blocks, located in genes *MGMT*, *MTHFS*, *CBS*, and *DNMT3L*, were statistically significant using a Bayesian false-discovery probability threshold of 0.8. Two blocks in *MGMT* and *MTHFS* appear to have significant fetal effects, while the *CBS* and *DNMT3L* genes may have significant MFG interaction effects.

Introduction

Congenital heart defects (CHDs) are the most prevalent type of birth defect and the leading cause of infant mortality attributable to birth defects (Yang et al. 2006). Worldwide, 1.35 million infants are born with CHDs per year, with an estimated prevalence of 8.1 per 1,000 live births in North America, and 9.3 per 1,000 live births in Asia (Reller et al. 2008; van der Linde et al. 2011). CHDs have many subtypes, and obstructive heart defects (OHDs) are a large subgroup of CHDs with obstructive lesions in blood vessels or valves in the right or left heart, such as pulmonic stenosis, aortic stenosis, coarctation of the aorta, bicuspid aortic valve stenosis, and subaortic stenosis. OHDs account for approximately 25 % of all CHDs, and our previous study has shown that OHDs are the most lethal subtype of heart defect in infancy (Cleves et al. 2003). Almost 40 % of infants born with serious OHDs die in infancy, and many of those who survive may require repeated surgeries and lengthy hospitalizations (Cleves et al. 2003; Gilboa et al. 2010; Nembhard et al. 2001). Despite recent medical and surgical advances, substantial OHD-attributable morbidity and mortality continues to represent a global burden of disease (Botto and Correa 2003).

CHDs are believed to be caused by a complex interplay between environmental exposures, genetic, and epigenetic factors. Familial aggregation provides direct evidence for the heritability of CHDs (Hobbs et al. 2002), while twin studies have demonstrated that the concordance rates of CHD phenotypes are significantly higher among monozygotic twins (10.0 %) than dizygotic twins (2.5 %) (Berg et al. 1989). The estimated heritability may vary by the phenotypes of cardiovascular malformation; for example, 89 % for bicuspid aortic valve and 71–90 % for left ventricular outflow tract malformations (Cripe et al. 2004; McBride et al. 2005). In the past few years, we and others have identified multiple genetic variants associated with CHDs (Goldmuntz et al. 2008; Hobbs et al. 2006, 2011; Wessels and Willems 2010). Though these findings have provided valuable insights into the genetic etiology of CHDs, the disease-susceptibility genes identified so far only account for a small fraction of CHD prevalence (Fahed et al. 2013). Moreover, relatively few studies have

investigated the association between OHDs and genetic variants. The genetic architecture of OHDs has remained elusive. It is still unclear how many genetic variants are associated with OHDs, how they are distributed in the population, and how they interact with one another to cause OHDs.

Investigation of the genetic mechanisms underlying the development of CHDs can be especially challenging. CHDs develop during embryogenesis, with multiple maternal–fetal metabolic, genomic, and epigenomic interactions. As discussed by Haig, a maternal–fetal unit may have three distinct haplotypes at each locus: the maternally derived fetal haplotype that is shared by the mother and fetus, the paternally derived fetal haplotype, and the non-inherited maternal haplotype (Haig 2004). These three types of haplotypes do not always have identical effects during pregnancy, and may not always be mutually beneficial to the mother and her fetus. Therefore, interaction effects between the mother and her fetus may lead to either an opposing or mutually beneficial environment for fetal growth (Sinsheimer et al. 2010). Over the past decade, evidence has accumulated demonstrating that maternal–fetal genotype interactions may be a common mechanism for various complex human diseases and birth defects, such as neural tube defects (Relton et al. 2004), schizophrenia (Palmer et al. 2002), and autism (Zandi et al. 2006). However, relatively few studies have been conducted to detect the MFG interaction with regard to the development of CHDs (Lupo et al. 2010).

Here, we report a haplotype-based analysis to dissect the maternal, fetal, and MFG interaction effects associated with OHDs in a study population from the National Birth Defects Prevention Study (NBDPS; Yoon et al. 2001). The current study included 877 SNPs selected from 62 candidate genes. Haplotype blocks were defined by the SNPs in linkage disequilibrium (LD). Our method utilized the least absolute shrinkage and selection operator (lasso; Tibshirani 1996), a machine learning technique which allows simultaneous effect estimation and variable selection, providing an automatic inference for the underlying genetic mechanisms. No individual test is required to differentiate maternal, fetal, and MFG interaction effects. Finally, we explore possible biological mechanisms with respect to the MFG combinations that jointly alter OHD risk.

Materials and methods

Ethics statement

The study was approved by University of Arkansas for Medical Sciences' Institutional Review Board and the National Birth Defects Prevention Study (NBDPS), with protocol oversight by the Centers for Disease Control and Prevention (CDC) Center for Birth Defects and Developmental Disabilities. All study subjects gave informed written consent. For minors, informed written consent was obtained from their legal guardian.

Study population

All subjects were participants of the National Birth Defects Prevention Study, an ongoing large-scale case control study covering an annual birth population of 482,000, or 10 % of U.S. births. OHD cases were ascertained from birth defect registries in ten participating

states that had similar inclusion criteria: Arkansas, California, Georgia, Iowa, Massachusetts, New Jersey, New York, North Carolina, Texas, and Utah. A detailed description of NBDPS methods can be found elsewhere (Gallagher et al. 2011; Rasmussen et al. 2002; Yoon et al. 2001). In the current study, we included all available mother-offspring pairs with estimated dates of delivery between October 1997 and August 2008. Case pairs were defined as those in which the child had at least one type of left or right OHDs, including hypoplastic left heart syndrome, tricuspid atresia, pulmonary valve atresia, coarctation of the aorta, interrupted aortic arch, aortic stenosis, valvar, pulmonic valve stenosis, and ebstein anomaly. Control pairs were defined as those in which the child had no structural birth defect. Control families were randomly selected from birth certificate and/or birth hospital records (Yoon et al. 2001) and thus represent a random sample from the general population. The study population included 294 case pairs and 874 control pairs. A comparison of maternal characteristics (Table 1) indicates that there was no significant difference between cases and controls with regard to maternal age, race, household income, education, smoking, drinking and folic acid exposure (all p values >0.05). However, the maternal BMI of case mothers was significantly higher than that of control mothers (p value <0.001). In order to adjust for the possible confounding effect of BMI, mother-offspring pairs with missing maternal BMI values were excluded from the analysis (9 case pairs and 26 control pairs), and the final analytical dataset includes 285 case pairs and 868 control pairs.

Genotyping and quality control

Our research team commissioned a custom 1,536 SNP panel covering 62 genes in the homocysteine, folate, and transsulfuration pathways potentially related to the development of CHDs, using the Illumina® GoldenGate custom genotyping platform, as described by Chowdhury et al. (2012). We found that the quality of genotype clustering varied substantially from SNP to SNP, which we attribute to the in silico design of the custom SNP panel without the subsequent quality checks that would be applied to a standard commercial array. The initial genotype calls, along with the raw intensity data, were used as inputs to SNPMClust, a bivariate Gaussian model-based genotype clustering and calling algorithm developed in-house, currently available as an R package on the Comprehensive R Archive Network (CRAN; <http://cran.r-project.org/>). After running SNPMClust, clustering and classification plots for all SNPs were visually inspected, leading to dropping a SNP from analysis or running SNPMClust under non-default settings in some cases. To ensure high-quality genotypes, we applied stringent quality control measures and excluded SNPs with obviously poor clustering behavior (60 SNPs), no-call rates $>10\%$ (389 SNPs), Mendelian error rates $>5\%$ (10 SNPs), minor allele frequencies (MAFs) $<5\%$ (192 SNPs), or significant deviation from Hardy–Weinberg Equilibrium in at least one racial group ($p < 10^{-4}$, 8 SNPs).

Haplotype blocks

The haplotype blocks were determined using software Haploview version 4.2 (Barrett et al. 2005). Linkage disequilibrium (LD) was first measured by the D' statistic between neighboring pairs of genetic variants. The Solid Spine of LD criterion, an internally developed method by Haploview, was used to determine the haplotype blocks using a

threshold of $D' > 0.6$. The LD structure was ignored if two SNPs were located more than 500 kb apart from each other, which is the default setting. A total of 140 haplotype blocks were identified for association analysis.

Statistical methods

Recently, we and others proposed using penalized logistic regression methods to dissect the maternal, fetal, and MFG interaction effects with mother-offspring pair data (Li et al. 2009, 2010). In this article, we adapt our method to detect main effects and interaction effects between haplotypes, the theoretical details of which can be found elsewhere (Li et al. 2010). The proposed method was implemented in R, and the source code is freely available upon request. We briefly describe our method below.

Denote H as a risk haplotype that potentially alters the likelihood of disease. The multi-locus genotypes within the haplotype block can then be mapped into three possible composite diplotypes, namely HH , $H\bar{H}$ and $\bar{H}\bar{H}$, where H represents all haplotypes that are different from the risk haplotype H . A logistic regression framework is then used to model the maternal effect, fetal effect, and MFG interaction effect:

$$\begin{aligned} \text{logit}(p(y_i=1)) = & \mu + a_m x_{i,m} + a_f x_{i,f} + d_m z_{i,m} + d_f z_{i,f} \\ & + i_{aa} x_{i,m} x_{i,f} + i_{ad} x_{i,m} z_{i,f} + i_{da} z_{i,m} x_{i,f} \\ & + i_{dd} z_{i,m} z_{i,f} + \gamma_1 UW_i + \gamma_2 OW_i + \gamma_3 OB_i \end{aligned} \quad (1)$$

where

$$x_{i,m} = \begin{cases} 1 & \text{for } G_{i,m} = HH \\ 0 & \text{for } G_{i,m} = H\bar{H} \\ -1 & \text{for } G_{i,m} = \bar{H}\bar{H} \end{cases} \quad \text{and } z_{i,m} = \begin{cases} -1/2 & \text{for } G_{i,m} = HH \\ 1/2 & \text{for } G_{i,m} = H\bar{H} \\ -1/2 & \text{for } G_{i,m} = \bar{H}\bar{H} \end{cases} .$$

and with $x_{i,f}$ and $z_{i,f}$ defined similarly. This coding strategy follows Cockerham's orthogonal partition method (Cockerham 1954; Kao and Zeng 2002), where $a_{m(f)}$ and $d_{m(f)}$ can be interpreted as the additive and dominance effects for the risk haplotype at a maternal (fetal) block, and i_{aa} , i_{ad} , i_{da} , i_{dd} can be interpreted as the additive \times additive, additive \times dominance, dominance \times additive, and dominance \times dominance interaction effects between the maternal and fetal blocks, respectively. UW , OW and OB are three dummy variables corresponding to three BMI categories: underweight, overweight, and obese.

The maternal-fetal genotype at a haplotype block has seven possible combinations, each corresponding to a likelihood of disease (Table 2), and coefficients were estimated by maximizing the penalized likelihood with an adaptive lasso penalty (Zou 2006). During parameter estimation, phase-ambiguous genotypes were treated as missing data, and phase was determined probabilistically via an expectation-maximization (EM) algorithm. The initial probability of haplotypes was estimated using the R package `haplo.stats`, which also allows for missing genotypes (Lake et al. 2003; Schaid et al. 2002). In our model, the haplotype block may have a large number of multi-locus genotypes. However, the number of composite diplotypes is always reduced to three after the haplotype configuration, which

significantly lessens data dimensionality. Such a modeling strategy has also been widely adopted in previous studies (Lin et al. 2007; Liu et al. 2004, 2011; Zhang et al. 2012). Further, adaptive lasso will simultaneously estimate parameters and perform model selection through shrinkage. Coefficients that do not significantly differ from 0 are expected to be shrunk to 0. Standard errors and confidence intervals of non-zero coefficients are empirically estimated using bootstrap resampling (Tibshirani 1996).

It is worthwhile to note that a risk haplotype is defined here for the purpose of dimension reduction. A risk haplotype may actually have a protective effect that corresponds to a lower likelihood of disease. In the real data application, all haplotypes with a frequency greater than 5 % were examined as potential risk haplotypes, and the haplotype with a minimum BIC was selected as the optimal risk haplotype. It should be noted that both the adaptive lasso estimator and the BIC criteria are asymptotically consistent in terms of model selection, which means that the probability of the selected model being the true model converges to 1 as the sample size increases (Yang 2005; Zou 2006).

Significance level of an association

The Bayesian false-discovery probability (BFDP) was used to evaluate the significance of associations for each haplotype block using the estimated odds ratios (OR) and corresponding 95 % confidence intervals (Wakefield 2007). BFDP has become a popular strategy to assess the noteworthiness of an association by balancing the costs of false discovery and non-discovery. In our analyses, we used the most widely accepted practice of pre-setting the BFDP threshold at 0.80, (Liu et al. 2010; Oh et al. 2010; Park et al. 2010; Spitz et al. 2012; Wakefield 2007; Zienolddiny et al. 2013) which means that missing a true association is considered four times more costly than falsely reporting an association. We also assume a 5 % prior probability for the existence of an association between OHD outcome and haplotype blocks. The upper bound of the 95 % CI of prior OR was assumed to be 1.5, meaning that the prior probability of an OR being greater than 1.5 is 2.5 %.

Results

Application of our method to 140 haplotype blocks identified 9 blocks with non-zero coefficients, indicating a potential genotype–phenotype association. Information for the identified haplotypes is summarized in Table 3. The frequencies of risk haplotypes were estimated based on the entire study population, including both cases and controls.

Based on the non-zero coefficients obtained by adaptive lasso estimator, the identified blocks fell into three possible categories: maternal main effect (i.e. $a_m, d_m = 0$), fetal main effect ($a_f, d_f = 0$), or MFG interaction effect (i.e. $i_{aa}, i_{ad}, i_{da}, i_{dd} = 0$). Among the nine identified haplotype blocks, four were found to have a BFDP less than 0.8, and were located in the genes *MGMT*, *MTHFS*, *CBS*, and *DNMT3L*. We further investigated the likelihood of disease for each MFG combination, and those 9 blocks exhibit three possible genetic mechanisms. In Figs. 1, 2, and 3, we plot one example for each possible mechanism. Supplementary materials (Figure S1–S9) provide LD plots and plots of the inferred genetic mechanisms for all identified blocks.

1. Five blocks exhibited only a fetal main effect

The results are summarized in Table 4. A haplotype block within the *MGMT* gene showed a significant association with the disease. This haplotype includes six SNPs covering a 14.5 kb region on chromosome 10. The seven MFG combinations at this region correspond to three levels of disease likelihood. Two MFG combinations had the lowest likelihood of disease, and were used as reference group. We denoted the maternal–fetal genotype combination in the reference group as $R1 = \{HH/HH; HH/\bar{H}H\}$. Compared to the reference group, the other MFG combinations formed two groups with increasing likelihoods of diseases, denoted as $R2 = \{HH/HH; HH/HH; HH/HH\}$ and $R3 = \{HH/HH; HH/HH\}$. The OR between R1 and R2 was estimated to be 1.47 (95 % CI 1.11, 1.96), and the OR between R1 and R3 was 2.16 (95 % CI 1.22, 3.83). As illustrated in Fig. 1, the fetal haplotype *H* showed an additive effect that decreases the risk of disease, while the risk of disease was unchanged by maternal genotypes. Similarly, our results show that the other four blocks may have a dominance effect (Figure S2–S5). However, only one block within gene *MTHFS* had a BFDP less than 0.8.

2. One block exhibited only a maternal main effect

One block was located within the *GSTM4* gene, comprising five SNPs on chromosome 1. The result is summarized in Table 4. For this block, the MFG combinations can be partitioned into three risk groups, according to the maternal genotypes. We illustrate the pattern in Fig. 2. However, this block had a BFDP greater than 0.8.

3. Three blocks exhibited MFG interaction effect

Three blocks were identified with MFG interaction effect (i.e. $i_{aa}, i_{ad}, i_{da}, i_{dd} \neq 0$). These three blocks were located within genes *DNMT3A*, *CBS*, *DNMT3L*, respectively, on chromosome 2, 21, and 21. The results are summarized in Table 4. The block within gene *DNMT3L* has the smallest size, comprising two highly linked SNPs 0.6 kb apart. Based on the estimated coefficients, the MFG combinations were partitioned into three risk groups. As illustrated in Fig. 3, when the maternal genotype is *HH*, the disease risk shows a decreasing pattern with the fetal genotypes. However, when the maternal genotype is *HH* (*HH*), the risk of disease shows an increasing pattern with the fetal genotype. This pattern of “cross-over” was an indication of the potential MFG interaction effect. Similarly, the interactive pattern of the blocks in gene *CBS* and *DNMT3A* are illustrated in Figure S8–S9. Among the three blocks, only *CBS* and *DNMT3L* had BFDP values less than 0.8.

Discussion

Our study builds on previous publications that have reported that the genetic susceptibility of CHDs may be associated with maternal genotypes, fetal genotypes, and MFG interactions (Goldmuntz et al. 2008; Hobbs et al. 2006, 2011; Lupo et al. 2010; Wessels and Willems 2010). Differentiating those genetic effects from each other, however, remains a challenge in maternal and perinatal research. This challenge is partly due to the correlation between maternal and fetal genomes. Maternal or fetal effects are likely to confound each other if they are tested separately. Therefore, a single model that simultaneously includes both maternal and fetal effects is preferred (Shi et al. 2008). In addition, although MFG

interactions are thought to exist pervasively in the development of birth defects, genetic studies of MFG interactions and our understanding of the complex interactive mechanisms are still in their infancy (Sinsheimer et al. 2010). The current study is motivated by the importance of and challenges faced by maternal and perinatal research, aiming to (1) dissect the genetic effects (i.e. maternal, fetal and MFG interactions) associated with OHDs, and (2) investigate the underlying genetic mechanisms for each identified gene (i.e. how a MFG combination influences the risk of disease). By employing an innovative logistic regression framework with adaptive lasso, we identified four genes potentially associated with OHDs. Further analyses of these results suggest that the identified genes may influence the phenotype through various genetic mechanisms.

In the current study, we have conducted a haplotype- or region-based association test, which is likely to improve the power to detect a causal association by aggregating collective effects of small to moderate size, offer an opportunity to capture complex interactions among variants, and reduce the burden of multiple testing (Li et al. 2011; Wang et al. 2013). However, compared to the conventional single-variant analysis, it is also less convenient to assess which particular variants might be driving the associations. In supplementary materials (Table S1), we also provide the single-variant analysis results for each SNP within the identified haplotype block, using a conventional logistic regression for the corresponding genetic effect model inferred by the adaptive lasso estimator. Further extension of the haplotype-based approach to a gene-based approach is likely to share the benefit of the single-variant approach, and avoid inference of phase-ambiguous haplotypes.

In our results, two haplotypes within gene *MGMT* and *MTHFS* exhibited fetal main effects only, which are significantly associated with the occurrence of OHDs (BFDP = 0.77 and 0.68, respectively). *MGMT* is thought to be involved in the prevention of DNA damage and oxidative stress, and previous studies have shown that *MGMT* gene expression is related to antioxidant mechanisms (Niture et al. 2007). Previous work by our research group has also found associations between multiple maternal SNPs of *MGMT* and maternal metabolites, such as levels of glutamylcysteine (GluCys) and plasma folate (Chowdhury et al. 2012). In particular, the plasma folate levels significantly varied by the maternal genotypes of rs10764896 (located in *MGMT*) in both cases and controls, indicating a potential involvement of rs10764896 in folate-related metabolite plasma concentrations. This SNP (rs10764896) is within the haplotype block identified in the current study. The *MTHFS* gene encodes a key enzyme involved in the folate pathway that metabolizes formyltetrahydrofolate. The folate pathway is interconnected with methionine metabolism by providing the methyl groups for the remethylation of homocysteine back to methionine. Our previous study also found maternal SNPs within *MTHFS* that are related to total plasma glutathione (GSH) levels (Chowdhury et al. 2012). Though experimental validation is needed to elucidate the functional mechanism of SNPs in *MGMT* and *MTHFS*, our current finding has provided additional evidence for their potential involvement in the development of CHDs.

Two haplotypes within the *CBS* and *DNMT3L* genes exhibited MFG interaction effects which are significantly associated with the occurrence of OHDs (BFDP = 0.46 and 0.49, respectively). The *CBS* gene is located on chromosome 21q22.3, and *CBS* deficiency is a

common cause of an inherited metabolic disorder, classical homocystinuria. The *CBS* gene catalyzes the conversion of homocysteine to cystathionine, the first step in the transsulfuration pathway. Studies have found that the expression of fetal *CBS* is concentrated in the neural and cardiac tissues, supporting a hypothesis of its involvement in embryo cardiac development (Quere et al. 1999; Robert et al. 2003). In particular, a recent study in Han Chinese population identified a functional variant (rs2850144) in *CBS* that is significantly associated with the genetic susceptibility to CHDs (Zhao et al. 2012). This SNP (rs2850144, BP: 43370045) was not genotyped in our study, but it is located approximately 5.5 kb from the identified haplotype block (BP 43354960–43364494). It is very likely that rs2850144 is in strong LD with the haplotype block identified in our study, and our results also suggest that the maternal *CBS* gene may interact with fetal *CBS* to jointly alter genetic susceptibility to the disease. The *DNMT3L* gene encodes an enzyme that stimulates de novo methylation by DNA cytosine methyltransferase 3 alpha, which is thought to be required for the establishment of maternal genomic imprintings (Hata et al. 2002). This enzyme also mediates transcriptional repression through interaction with histone deacetylase 1 (Aapola et al. 2002; Deplus et al. 2002). Studies have reported association of *DNMT3L* with multiple conditions, such as ovarian endometriosis (Borghese et al. 2012), embryonal carcinoma (Minami et al. 2010), and cervical cancer (Gokul et al. 2007). Further studies will be necessary to validate the potential involvement of *DNMT3L* in the development of CHDs.

A few limitations of our study should be noted. First, although all of the identified haplotypes had *p* values less than the nominal significant threshold of 0.05, none of them would remain significant with the stringent Bonferroni correction. However, Bonferroni correction is usually overly conservative, and BFDP has been suggested as a good alternative, especially in candidate gene-based studies when there is strong *priori* probability that the genes are likely to be associated with the disease. Second, two types of gene–gene interactions are possible during pregnancy: intra-generational interaction *within* either the maternal or fetal genome, and inter-generational interaction *between* maternal and fetal genomes. Our current analysis only considers the inter-generational interactions between maternal and fetal genes from the same genomic region (LD block). Intra-generational interaction and MFG interaction between genes from different regions are not within the scope of the current study. Third, the current study only included common SNPs that have minor allele frequencies (MAFs) of 5 % or higher. Evidence from Phase III of the International HapMap Project and 1,000 Genomes Project have supported that rare variants with lower MAFs may contribute considerably to the development of complex human diseases (Abecasis et al. 2012; Altshuler et al. 2010). However, our current analysis tests each haplotype block, which is less easy to take into account the rare variants. Fourth, the focus of our study is to investigate the genetic susceptibility of OHDs. We are able to control the potential confounding effect of environmental factors such as BMI, but considering potential gene-by-environment interactions would substantially increase the model complexity and are not considered in the current study.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We want to thank two anonymous reviewers for their insightful comments and suggestions. We also wish to thank numerous families for their generous participation in the National Birth Defects Prevention Study that made this research possible. We also thank the Centers for Birth Defects Research and Prevention in Arkansas, California, Georgia, Iowa, Massachusetts, New Jersey, New York, North Carolina, Texas, and Utah for their contribution of data and manuscript review. The authors also wish to thank Jingyun Li and Ashley S. Block for assistance in the preparation of this manuscript. This work is supported by the National Institute of Child Health and Human Development (NICHD) under award number 5R01HD039054-12 and the National Center on Birth Defects and Developmental Disabilities (NCBDDD) under award number 5U01DD000491-05.

References

- Aapola U, Liiv I, Peterson P. Imprinting regulator DNMT3L is a transcriptional repressor associated with histone deacetylase activity. *Nucleic Acids Res.* 2002; 30:3602–3608. [PubMed: 12177302]
- Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. *Nature.* 2012; 491:56–65.10.1038/nature11632 [PubMed: 23128226]
- Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Peltonen L, Dermitzakis E, Bonnen PE, Altshuler DM, Gibbs RA, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Yu F, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Gibbs RA, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, Lee C, McCarroll SA, Nemes J, Dermitzakis E, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo N, Bonnen PE, Gibbs RA, Gonzaga-Jauregui C, Keinan A, Price AL, Yu F, Anttila V, Brodeur W, Daly MJ, Leslie S, McVean G, Moutsianas L, Nguyen H, Schaffner SF, Zhang Q, Ghorri MJ, McGinnis R, McLaren W, Pollack S, Price AL, Schaffner SF, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti PC, Adebamowo CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare D, Wang VO, Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD, McEwen JE. Integrating common and rare genetic variation in diverse human populations. *Nature.* 2010; 467:52–58.10.1038/nature09298 [PubMed: 20811451]
- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005; 21:263–265.10.1093/bioinformatics/bth457 [PubMed: 15297300]
- Berg KA, Astemborski JA, Boughman JA, Ferencz C. Congenital cardiovascular malformations in twins and triplets from a population-based study. *Am J Dis Child.* 1989; 143:1461–1463. [PubMed: 2589279]
- Borghese B, Santulli P, Hequet D, Pierre G, de Ziegler D, Vaiman D, Chapron C. Genetic polymorphisms of DNMT3L involved in hypermethylation of chromosomal ends are associated with greater risk of developing ovarian endometriosis. *Am J Pathol.* 2012; 180:1781–1786.10.1016/j.ajpath.2012.01.009 [PubMed: 22401780]
- Botto LD, Correa A. Decreasing the burden of congenital heart anomalies: an epidemiologic evaluation of risk factors and survival. *Prog Pediatr Cardiol.* 2003; 18:111–121.
- Chowdhury S, Hobbs CA, MacLeod SL, Cleves MA, Melnyk S, James SJ, Hu P, Erickson SW. Associations between maternal genotypes and metabolites implicated in congenital heart defects. *Mol Genet Metab.* 2012; 107:596–604.10.1016/j.ymgme.2012.09.022 [PubMed: 23059056]
- Cleves MA, Ghaffar S, Zhao W, Mosley BS, Hobbs CA. First-year survival of infants born with congenital heart defects in Arkansas (1993–1998): a survival analysis using registry data. *Birth Defects Res A Clin Mol Teratol.* 2003; 67:662–668.10.1002/bdra.10119 [PubMed: 14703791]
- Cockerham CC. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics.* 1954; 39:859–882. [PubMed: 17247525]
- Cripe L, Andelfinger G, Martin LJ, Shooner K, Benson DW. Bicuspid aortic valve is heritable. *J Am Coll Cardiol.* 2004; 44:138–143.10.1016/j.jacc.2004.03.050 [PubMed: 15234422]
- Deplus R, Brenner C, Burgers WA, Putmans P, Kouzarides T, de Launoit Y, Fuks F. Dnmt3L is a transcriptional repressor that recruits histone deacetylase. *Nucleic Acids Res.* 2002; 30:3831–3838. [PubMed: 12202768]

- Fahed AC, Gelb BD, Seidman JG, Seidman CE. Genetics of congenital heart disease: the glass half empty. *Circ Res*. 2013; 112:707–720.10.1161/CIRCRESAHA.112.300853 [PubMed: 23410880]
- Gallagher ML, Sturchio C, Smith A, Koontz D, Jenkins MM, Honein MA, Rasmussen SA. Evaluation of mailed pediatric buccal cytobrushes for use in a case-control study of birth defects. *Birth Defects Res A Clin Mol Teratol*. 2011; 91:642–648.10.1002/bdra.20829 [PubMed: 21630425]
- Gilboa SM, Salemi JL, Nembhard WN, Fixler DE, Correa A. Mortality resulting from congenital heart disease among children and adults in the United States, 1999 to 2006. *Circulation*. 2010; 122:2254–2263. [PubMed: 21098447]
- Gokul G, Gautami B, Malathi S, Sowjanya AP, Poli UR, Jain M, Ramakrishna G, Khosla S. DNA methylation profile at the DNMT3L promoter: a potential biomarker for cervical cancer. *Epigenetics*. 2007; 2:80–85. [PubMed: 17965599]
- Goldmuntz E, Woyciechowski S, Renstrom D, Lupo PJ, Mitchell LE. Variants of folate metabolism genes and the risk of conotruncal cardiac defects. *Circ Cardiovasc Genet*. 2008; 1:126–132.10.1161/CIRCGENETICS.108.796342 [PubMed: 20031554]
- Haig D. Evolutionary conflicts in pregnancy and calcium metabolism—a review. *Placenta*. 2004; 25(Suppl A):S10–5.10.1016/j.placenta.2004.01.006 [PubMed: 15033301]
- Hata K, Okano M, Lei H, Li E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development*. 2002; 129:1983–1993. [PubMed: 11934864]
- Hobbs CA, Cleves MA, Simmons CJ. Genetic epidemiology and congenital malformations: from the chromosome to the crib. *Arch Pediatr Adolesc Med*. 2002; 156:315–320. [PubMed: 11929362]
- Hobbs CA, James SJ, Parsian A, Krakowiak PA, Jernigan S, Greenhaw JJ, Lu Y, Cleves MA. Congenital heart defects and genetic variants in the methylenetetrahydrofolate reductase gene. *J Med Genet*. 2006; 43:162–166.10.1136/jmg.2005.032656 [PubMed: 15951337]
- Hobbs CA, MacLeod SL, Jill James S, Cleves MA. Congenital heart defects and maternal genetic, metabolic, and lifestyle factors. *Birth Defects Res A Clin Mol Teratol*. 2011; 91:195–203.10.1002/bdra.20784 [PubMed: 21384532]
- Kao CH, Zeng ZB. Modeling epistasis of quantitative trait loci using Cockerham’s model. *Genetics*. 2002; 160:1243–1261. [PubMed: 11901137]
- Lake SL, Lyon H, Tantisira K, Silverman EK, Weiss ST, Laird NM, Schaid DJ. Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. *Hum Hered*. 2003; 55:56–65.10.1159/000071811 [PubMed: 12890927]
- Li S, Lu Q, Fu W, Romero R, Cui Y. A regularized regression approach for dissecting genetic conflicts that increase disease risk in pregnancy. *Stat Appl Genet Mol Biol*. 2009; 8:45.10.2202/1544-6115.1474
- Li M, Romero R, Fu WJ, Cui Y. Mapping haplotype–haplotype interactions with adaptive LASSO. *BMC Genet*. 2010; 11:79.10.1186/1471-2156-11-79 [PubMed: 20799953]
- Li MX, Gui HS, Kwan JS, Sham PC. GATES: a rapid and powerful gene-based association test using extended Simes procedure. *Am J Hum Genet*. 2011; 88:283–293.10.1016/j.ajhg.2011.01.019 [PubMed: 21397060]
- Lin M, Li H, Hou W, Johnson JA, Wu R. Modeling sequence–sequence interactions for drug response. *Bioinformatics*. 2007; 23:1251–1257.10.1093/bioinformatics/btm110 [PubMed: 17392331]
- Liu T, Johnson JA, Casella G, Wu R. Sequencing complex diseases With HapMap. *Genetics*. 2004; 168:503–511.10.1534/genetics.104.029603 [PubMed: 15454560]
- Liu Y, Shete S, Etzel CJ, Scheurer M, Alexiou G, Armstrong G, Tsavachidis S, Liang FW, Gilbert M, Aldape K, Armstrong T, Houlston R, Hosking F, Robertson L, Xiao Y, Wiencke J, Wrensch M, Andersson U, Melin BS, Bondy M. Polymorphisms of LIG4, BTBD2, HMGA2, and RTEL1 genes involved in the double-strand break repair pathway predict glioblastoma survival. *J Clin Oncol*. 2010; 28:2467–2474.10.1200/JCO.2009.26.6213 [PubMed: 20368557]
- Liu T, Thalamuthu A, Liu JJ, Chen C, Wang Z, Wu R. Asymptotic distribution for epistatic tests in case-control studies. *Genomics*. 2011; 98:145–151.10.1016/j.ygeno.2011.05.001 [PubMed: 21620949]

- Lupo PJ, Goldmuntz E, Mitchell LE. Gene-gene interactions in the folate metabolic pathway and the risk of conotruncal heart defects. *J Biomed Biotechnol.* 2010; 2010:630940.10.1155/2010/630940 [PubMed: 20111745]
- McBride KL, Pignatelli R, Lewin M, Ho T, Fernbach S, Menesses A, Lam W, Leal SM, Kaplan N, Schliekelman P, Towbin JA, Belmont JW. Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: segregation, multiplex relative risk, and heritability. *Am J Med Genet A.* 2005; 134A:180–186.10.1002/ajmg.a.30602 [PubMed: 15690347]
- Minami K, Chano T, Kawakami T, Ushida H, Kushima R, Okabe H, Okada Y, Okamoto K. DNMT3L is a novel marker and is essential for the growth of human embryonal carcinoma. *Clin Cancer Res.* 2010; 16:2751–2759.10.1158/1078-0432.CCR-09-3338 [PubMed: 20460473]
- Nembhard WN, Waller DK, Sever LE, Canfield MA. Patterns of first-year survival among infants with selected congenital anomalies in Texas, 1995–1997. *Teratology.* 2001; 64:267–275. [PubMed: 11745833]
- Niture SK, Velu CS, Smith QR, Bhat GJ, Srivenugopal KS. Increased expression of the MGMT repair protein mediated by cysteine prodrugs and chemopreventative natural products in human lymphocytes and tumor cell lines. *Carcinogenesis.* 2007; 28:378–389.10.1093/carcin/bgl155 [PubMed: 16950796]
- Oh SS, Chang SC, Cai L, Cordon-Cardo C, Ding BG, Greenland S, He N, Jiang Q, Kheifets L, Le A, Lee YC, Liu S, Lu ML, Mao JT, Morgenstern H, Mu LN, Pantuck A, Papp JC, Park SL, Rao JY, Reuter VE, Tashkin DP, Wang H, You NC, Yu SZ, Zhao JK, Belldegrin A, Zhang ZF. Single nucleotide polymorphisms of 8 inflammation-related genes and their associations with smoking-related cancers. *Int J Cancer.* 2010; 127:2169–2182.10.1002/ijc.25214 [PubMed: 20112337]
- Palmer CG, Turunen JA, Sinsheimer JS, Minassian S, Paunio T, Lonnqvist J, Peltonen L, Woodward JA. RHD maternal-fetal genotype incompatibility increases schizophrenia susceptibility. *Am J Hum Genet.* 2002; 71:1312–1319.10.1086/344659 [PubMed: 12439825]
- Park SL, Bastani D, Goldstein BY, Chang SC, Cozen W, Cai L, Cordon-Cardo C, Ding B, Greenland S, He N, Hussain SK, Jiang Q, Lee YC, Liu S, Lu ML, Mack TM, Mao JT, Morgenstern H, Mu LN, Oh SS, Pantuck A, Papp JC, Rao J, Reuter VE, Tashkin DP, Wang H, You NC, Yu SZ, Zhao JK, Zhang ZF. Associations between NBS1 polymorphisms, haplotypes and smoking-related cancers. *Carcinogenesis.* 2010; 31:1264–1271.10.1093/carcin/bgq096 [PubMed: 20478923]
- Quere I, Paul V, Rouillac C, Janbon C, London J, Demaille J, Kamoun P, Dufier JL, Abitbol M, Chasse JF. Spatial and temporal expression of the cystathionine beta-synthase gene during early human development. *Biochem Biophys Res Commun.* 1999; 254:127–137.10.1006/bbrc.1998.9079 [PubMed: 9920745]
- Rasmussen SA, Lammer EJ, Shaw GM, Finnell RH, McGehee RE Jr, Gallagher M, Romitti PA, Murray JC. National Birth Defects Prevention Study. Integration of DNA sample collection into a multi-site birth defects case-control study. *Teratology.* 2002; 66:177–184.10.1002/tera.10086 [PubMed: 12353214]
- Reller MD, Strickland MJ, Riehle-Colarusso T, Mahle WT, Correa A. Prevalence of congenital heart defects in metropolitan Atlanta, 1998–2005. *J Pediatr.* 2008; 153:807–813.10.1016/j.jpeds.2008.05.059 [PubMed: 18657826]
- Relton CL, Wilding CS, Pearce MS, Laffling AJ, Jonas PA, Lynch SA, Tawn EJ, Burn J. Gene-gene interaction in folate-related genes and risk of neural tube defects in a UK population. *J Med Genet.* 2004; 41:256–260. [PubMed: 15060097]
- Robert K, Vialard F, Thiery E, Toyama K, Sinet PM, Janel N, London J. Expression of the cystathionine beta synthase (CBS) gene during mouse development and immunolocalization in adult brain. *J Histochem Cytochem.* 2003; 51:363–371. [PubMed: 12588964]
- Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet.* 2002; 70:425–434.10.1086/338688 [PubMed: 11791212]
- Shi M, Umbach DM, Vermeulen SH, Weinberg CR. Making the most of case-mother/control-mother studies. *Am J Epidemiol.* 2008; 168:541–547.10.1093/aje/kwn149 [PubMed: 18650222]
- Sinsheimer JS, Elston RC, Fu WJ. Gene-gene interaction in maternal and perinatal research. *J Biomed Biotechnol* 2010. 201010.1155/2010/853612

- Spitz MR, Gorlov IP, Dong Q, Wu X, Chen W, Chang DW, Etzel CJ, Caporaso NE, Zhao Y, Christiani DC, Brennan P, Albanes D, Shi J, Thun M, Landi MT, Amos CI. Multistage analysis of variants in the inflammation pathway and lung cancer risk in smokers. *Cancer Epidemiol Biomarkers Prev.* 2012; 21:1213–1221.10.1158/1055-9965.EPI-12-0352-T [PubMed: 22573796]
- Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B.* 1996; 58:267–288.
- van der Linde D, Konings EE, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJ, Roos-Hesselink JW. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J Am Coll Cardiol.* 2011; 58:2241–2247.10.1016/j.jacc.2011.08.025 [PubMed: 22078432]
- Wakefield J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am J Hum Genet.* 2007; 81:208–227.10.1086/519024 [PubMed: 17668372]
- Wang X, Lee S, Zhu X, Redline S, Lin X. GEE-based SNP set association test for continuous and discrete traits in family-based association studies. *Genet Epidemiol.* 2013; 37:778–786.10.1002/gepi.21763 [PubMed: 24166731]
- Wessels MW, Willems PJ. Genetic factors in non-syndromic congenital heart malformations. *Clin Genet.* 2010; 78:103–123.10.1111/j.1399-0004.2010.01435.x [PubMed: 20497191]
- Yang Y. Can the strengths of AIC and BIC be shared? A conflict between model identification and regression estimation. *Biometrika.* 2005; 92:937–950.
- Yang Q, Chen H, Correa A, Devine O, Mathews TJ, Honein MA. Racial differences in infant mortality attributable to birth defects in the United States, 1989–2002. *Birth Defects Res A Clin Mol Teratol.* 2006; 76:706–713.10.1002/bdra.20308 [PubMed: 17022030]
- Yoon PW, Rasmussen SA, Lynberg MC, Moore CA, Anderka M, Carmichael SL, Costa P, Druschel C, Hobbs CA, Romitti PA, Langlois PH, Edmonds LD. The National Birth Defects Prevention Study. *Public Health Rep.* 2001; 116(Suppl 1):32–40. [PubMed: 11889273]
- Zandi PP, Kalaydjian A, Avramopoulos D, Shao H, Fallin MD, Newschaffer CJ. Rh and ABO maternal-fetal incompatibility and risk of autism. *Am J Med Genet B Neuropsychiatr Genet.* 2006; 141B:643–647.10.1002/ajmg.b.30391 [PubMed: 16856119]
- Zhang L, Liu R, Wang Z, Culver DA, Wu R. Modeling haplotype–haplotype interactions in case-control genetic association studies. *Front Genet.* 2012; 3:2.10.3389/fgene.2012.00002 [PubMed: 22303409]
- Zhao JY, Yang XY, Shi KH, Sun SN, Hou J, Ye ZZ, Wang J, Duan WY, Qiao B, Chen YJ, Shen HB, Huang GY, Jin L, Wang HY. A functional variant in the cystathionine beta-synthase gene promoter significantly reduces congenital heart disease susceptibility in a Han Chinese population. *Cell Res.* 2012.10.1038/cr.2012.135
- Zienolddiny S, Haugen A, Lie JA, Kjuus H, Anmarkrud KH, Kjaerheim K. Analysis of polymorphisms in the circadian-related genes and breast cancer risk in Norwegian nurses working night shifts. *Breast Cancer Res.* 2013; 15:R53.10.1186/bcr3445 [PubMed: 23822714]
- Zou H. The adaptive LASSO and its oracle properties. *JASA.* 2006; 101:1418–1429.

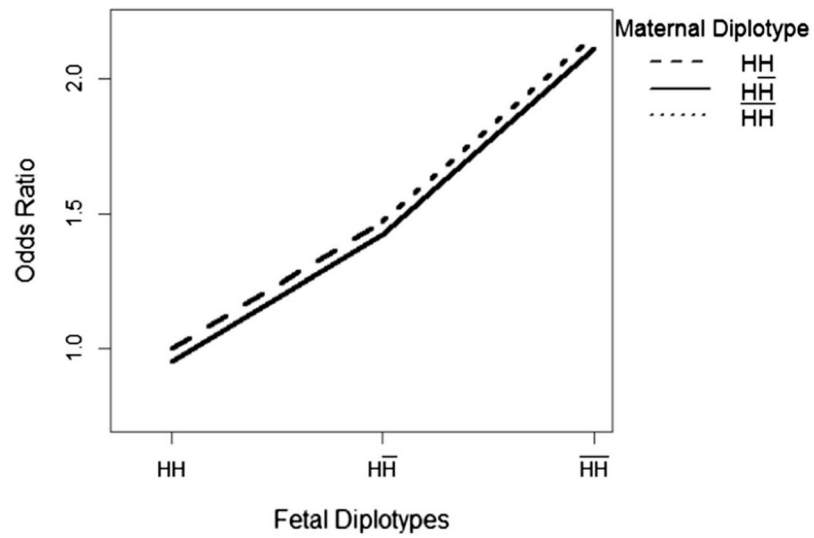


Fig. 1. *MGMT* (Block 4)—Fetal main effect only. Fetal haplotype *H* showed an additive effect that was protective of the disease, while the risk of disease was unchanged by maternal genotypes

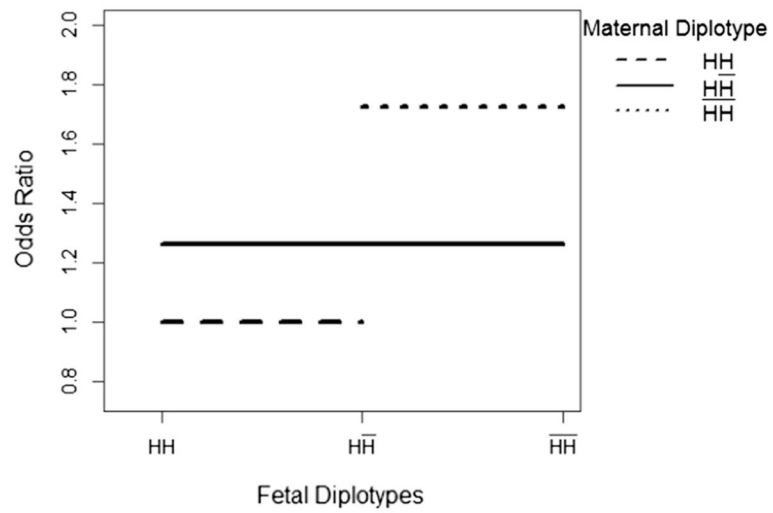


Fig. 2. *GSTM4* (Block 1)—maternal main effect only. Maternal haplotype *H* showed an additive effect that was protective of the disease, while the risk of disease was unchanged by fetal genotypes

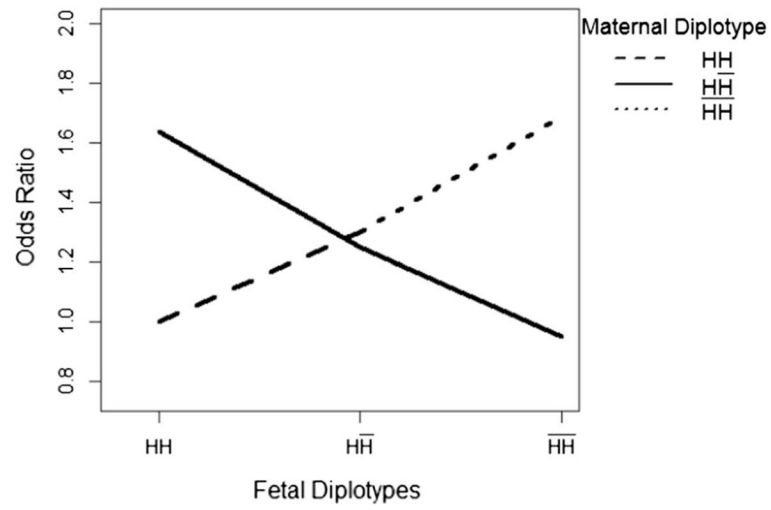


Fig. 3. *DNMT3L* (Block 9)—MFG interaction effect. Maternal and fetal genotypes showed interactive pattern in terms of disease risk

Table 1

Maternal characteristics: NBDPS 1997–2007

	Cases (N = 294)	Controls (N = 874)	<i>p</i> value
Mother's age at conception [mean (SD)]	28.0 (5.8 %)	27.7 (5.9 %)	0.401
Mother's BMI at conception [<i>n</i> (%)]			<0.001
Underweight	2 (0.7 %)	35 (4.1 %)	
Normal	126 (44.2 %)	460 (54.3 %)	
Overweight	82 (28.8 %)	195 (23.0 %)	
Obese	75 (26.3 %)	158 (18.6 %)	
Missing	9	26	
Mother's race [<i>n</i> (%)]			0.330
African American	30 (10.2 %)	88 (10.1 %)	
Caucasian	221 (75.2 %)	618 (70.7 %)	
Hispanic	30 (10.2 %)	124 (14.2 %)	
Others	13 (4.4 %)	44 (5.0 %)	
Household income [<i>n</i> (%)]			0.076
\$0–\$10,000	38 (13.4 %)	112 (13.6 %)	
\$10,000–\$30,000	89 (31.4 %)	236 (28.6 %)	
\$30,000–\$50,000	80 (28.3 %)	190 (23.1 %)	
>\$50,000	76 (26.9 %)	286 (34.7 %)	
Missing	11 (3.7 %)	50 (5.7 %)	
Mother's education [<i>n</i> (%)]			0.942
<12 years	40 (13.6 %)	117 (13.4 %)	
High school degree or equivalent	71 (24.1 %)	208 (23.8 %)	
1–3 years of college	86 (29.3 %)	244 (27.9 %)	
4 years of college or Bachelor degree	97 (33.0 %)	305 (34.9 %)	
Smoking [<i>n</i> (%)]			0.459
Yes	46 (15.6)	153 (17.5 %)	
No	248 (84.4)	720 (82.5 %)	
Missing	0	1	
Drinking [<i>n</i> (%)]			0.518
Yes	59 (20.1 %)	191 (21.9 %)	
No	234 (79.9 %)	680 (78.1 %)	
Missing	1	3	
Folic acid exposure [<i>n</i> (%)]			0.572
Yes	164 (55.8 %)	504 (57.7 %)	
No	130 (44.2 %)	370 (42.3 %)	

p value less than 0.05 (in bold) indicates a significant difference between case mothers and control mothers

Table 2

Maternal–fetal genotype combinations and numerical coding for disease risk

MFG	Fetal diplotype		
	<i>HH</i>	<i>HH⁻</i>	<i>H\bar{H}⁻</i>
Maternal diplotype			
<i>HH</i>	(1, 1, -1/2, -1/2) ^a	(1, 0, -1/2, 1/2)	<i>_b</i>
<i>HH⁻</i>	(0, 1, 1/2, -1/2)	(0, 0, 1/2, 1/2)	(0, -1, 1/2, -1/2)
<i>H\bar{H}⁻</i>	<i>_b</i>	(-1, 0, -1/2, 1/2)	(-1, -1, -1/2, -1/2)

^aCoding values for ($x_{i,m}$, $z_{i,m}$, $x_{i,f}$, $z_{i,f}$). According to Eq. (1), the corresponding log-odds of being a case pair with a normal maternal weight is:
 $a_m + a_f - \frac{1}{2}d_m - \frac{1}{2}d_f + i_{aa} - \frac{1}{2}i_{ad} - \frac{1}{2}i_{da} + \frac{1}{4}i_{dd}$

^bCombination not possible under Mendelian transmission

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3
 Nine haplotype blocks identified with non-zero genetic effects: NBDPS participants 1997–2007

Block ^a	gene ^b	Chro.	Block size	Ave. r^2	SNP in block	Position	Allele	“Risk” haplotype
Block 1		<i>GSTM4</i>	1	14.7 kb	0.301	rs542338	109994289	A/G
						rs668413	109997467	A/C
						rs560018	110001883	A/G
						rs535537	110002649	A/G
						rs670439	110009005	A/G
						Among mothers: 19.4 %		
						Among offspring: 18.4 %		
Block 2		<i>DNMT3A</i>	2	0.4 kb	0.380	rs7578575	25342323	A/T
						rs7590760	25342687	C/G
						Among mothers: 21.6 %		
						Among offspring: 22.7 %		
Block 3		<i>PGDS</i>	4	61.5 kb	0.503	rs8336	95430633	A/G
						rs11097411	95441649	A/G
						rs2016483	95448062	A/T
						rs11932130	95448646	A/G
						rs10033662	95458485	A/G
						rs11097413	95468309	A/G
						rs2865352	95472286	A/G
						rs10084984	95473598	A/G
						rs10516950	95473753	A/G
						rs2059605	95474235	A/G
						rs2289186	95474660	A/C
						rs724260	95479708	A/G
						rs11727030	95481688	A/G
						rs1991316	95487295	A/C
						rs10856909	95487771	C/G
						rs4282187	95492168	A/G
						Among mothers: 37.7 %		
						Among offspring: 36.7 %		
Block 4		<i>MGMT</i>	10	14.5 kb	0.215	rs11813363	131359916	A/G
						rs7068306	131360027	C/G

Block <i>a</i>	gene ^p	Chro.	Block size	Ave. <i>r</i> ²	SNP in block	Position	Allele	"Risk" haplotype
Block 5	<i>MGMT</i>	10	3.4 kb	0.047	rs10829615	131361169	A/G	G
					rs10764896	131361712	A/G	A
					rs11016866	131361958	A/G	A
					rs7910123	131363699	A/C	A
					Among mothers: 12.8 %		Among offspring: 13.8 %	
rs10829616	131370984	A/G	A					
rs2039374	131373272	A/G	G					
rs11016875	131374400	A/G	G					
Frequency of "Risk" haplotype		Among mothers: 12.0 %		Among offspring: 12.3 %				
Block 6	<i>MTHFS</i>	15	23.5 kb	0.246	rs17175723	77918543	C/G	C
					rs685487	77923184	A/G	A
					rs11854561	77923827	A/G	A
					rs2733103	77925626	A/G	G
					rs622506	77926250	A/C	A
rs2054287	77930090	A/G	G					
rs17284990	77931252	A/G	A					
rs7164897	77934635	A/T	A					
rs12912711	77938591	A/G	G					
rs16971449	77940052	A/G	A					
rs6495446	77942037	A/G	A					
Frequency of "Risk" haplotype		Among mothers: 15.1 %		Among offspring: 15.1 %				
Block 7	<i>MTHFS</i>	15	7.8 kb	0.397	rs12440609	77978543	A/G	G
					rs8039272	77983805	A/G	A
					rs8030396	77986304	A/G	G
Frequency of "Risk" haplotype		Among mothers: 46.6 %		Among offspring: 45.9 %				
Block 8	<i>CBS</i>	21	9.5 kb	0.253	rs4920037	43354960	A/G	G
					rs12329764	43356317	A/G	G
					rs234705	43356841	A/G	G
					rs9974224	43357811	A/G	G
					rs2851391	43360473	A/G	G
rs11701048	43364494	A/G	G					

Block ^a	gene ^b	Chro.	Block size	Ave. r^2	SNP in block	Position	Allele	“Risk” haplotype
Frequency of “Risk” haplotype Among mothers: 22.9 %								
Block 9	<i>DNMT3L</i>	21	0.6 kb	0.04	rs2838540	44512009	A/G	G
Frequency of “Risk” haplotype Among mothers: 63.7 %								
					rs2838541	44512643	A/G	G
Among offspring: 22.0 %								
Among offspring: 63.2 %								

^aThe blocks are numbered by chromosome and physical location, and do not represent the level of association

^b*GSTM4* glutathione S-transferase mu 4, *DNMT3A* DNA (cytosine-5-)-methyltransferase 3 alpha, *PGDS* hematopoietic prostaglandin D synthase, *MGMT* O-6-methylguanine-DNA methyltransferase, *MTHFS* 5,10-methylenetetrahydrofolate synthetase (5-formyltetrahydrofolate cycloligase), *CBS* cystathionine-beta-synthase, *DNMT3L* DNA (cytosine-5-)-methyltransferase 3-like

Table 4

Haplotype blocks with non-zero genetic effects: NBDPS 1997–2007

Block gene	MFG combination α	Maternal/fetal	OR (95 % CI)	p value	BFDP
Haplotype blocks with fetal main effect only					
Block 4	<i>MGMT</i>	R1: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	Ref	0.0082	0.77
		R2: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.47 [1.11, 1.96]		
		R3: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	2.16 [1.22, 3.83]		
Block 7	<i>MTHFS</i>	R1: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	Ref	0.0043	0.68
		R2: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.48 [1.13, 1.94]		
Block 3	<i>PGDS</i>	R1: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	Ref	0.017	0.83
		R2: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.41 [1.06, 1.87]		
Block 5	<i>MGMT</i>	R1: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	Ref	0.0015	0.89
		R2: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.36 [1.01, 1.83]		
		R3: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.86 [1.03, 3.35]		
Block 6	<i>MTHFS</i>	R1: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	Ref	0.031	0.88
		R2: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.42 [1.03, 1.95]		
Haplotype blocks with maternal main effect only					
Block 7	<i>GSTM4</i>	R1: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	Ref	0.038	0.89
		R2: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.31 [1.02, 1.70]		
		R3: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.72 [1.03, 2.88]		
Haplotype blocks with MFG interaction effect					
Block 8	<i>CBS</i>	R1: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	Ref	0.0040	0.46
		R2: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.36 [1.13, 1.65]		
		R3: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.86 [1.27, 2.72]		
Block 9	<i>DNMT3L</i>	R1: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	Ref	0.0032	0.49
		R2: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.30 [1.09, 1.54]		
		R3: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.69 [1.19, 2.38]		
Block 2	<i>DNMT3A</i>	R1: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	Ref	0.028	0.86
		R2: $\overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}; \overline{HH}/\overline{HH}$	1.27 [1.03, 1.57]		

Block gene	MFG combination ^a	Maternal/fetal	OR (95 % CI)	p value	BFDP
	<i>R3</i> : HH/HH ; HH/HH		1.61 [1.05, 2.45]		

BFDP values less than 0.8 (in bold) indicate a significant genotype–phenotype association

^aPartition of MFG combinations into various risk groups according to their likelihoods of disease. *R1* is the reference group with the lowest likelihood of disease

^bBased on 100 bootstrap samples

^cBased on 100 bootstrap samples. Null hypothesis assumes all MFG combinations have the same likelihood of disease