Article

Gene–Folic Acid Interactions and Risk of Conotruncal Heart Defects: Results from the National Birth Defects Prevention Study

Daniel M. Webber [1,†], Ming Li [2,†,‡], Stewart L. MacLeod [3,‡], Xinyu Tang [4], Joseph W. Levy [5,†], Mohammad A. Karim [6,7,†], Stephen W. Erickson [8,†] , Charlotte A. Hobbs [9,*,†] and The National Birth Defects Prevention Study [§]

1   Department of Pathology and Immunology, Washington University School of Medicine,
    St. Louis, MO 63110, USA
2   Department of Epidemiology and Biostatistics, Indiana University at Bloomington,
    Bloomington, IN 47405, USA
3   Division of Birth Defects Research, Department of Pediatrics, University of Arkansas for Medical Sciences,
    Little Rock, AR 72205, USA
4   Biostatistics Program, Department of Pediatrics, University of Arkansas for Medical Sciences,
    Little Rock, AR 72205, USA
5   Department of Obstetrics and Gynecology, Wayne State University, Detroit, MI 48202, USA
6   Department of Child Health, College of Medicine, University of Arizona, Phoenix, AZ 85004, USA
7   Department of Neurology, Sections on Neurodevelopmental Disorders, Barrow Neurological Institute at
    Phoenix Children's Hospital, Phoenix, AZ 85016, USA
8   Center for Genomics in Public Health and Medicine, RTI International,
    Research Triangle Park, NC 27709, USA
9   Rady Children's Institute for Genomic Medicine, Rady Children's Hospital, San Diego, CA 92123, USA
*   Correspondence: chobbs@rchsd.org
†   Previous address: Division of Birth Defects Research, Department of Pediatrics, University of Arkansas for
    Medical Sciences and Arkansas Children's Research Institute, Little Rock, AR 72205, USA.
‡   These authors contributed equally to this work.
§   Membership of the National Birth Defects Prevention Study is provided in the Acknowledgments.

Abstract: Conotruncal heart defects (CTDs) are heart malformations that affect the cardiac outflow tract and typically cause significant morbidity and mortality. Evidence from epidemiological studies suggests that maternal folate intake is associated with a reduced risk of heart defects, including CTD. However, it is unclear if folate-related gene variants and maternal folate intake have an interactive effect on the risk of CTDs. In this study, we performed targeted sequencing of folate-related genes on DNA from 436 case families with CTDs who are enrolled in the National Birth Defects Prevention Study and then tested for common and rare variants associated with CTD. We identified risk alleles in maternal *MTHFS* (OR_{meta} = 1.34; 95% CI 1.07 to 1.67), maternal *NOS2* (OR_{meta} = 1.34; 95% CI 1.05 to 1.72), fetal *MTHFS* (OR_{meta} = 1.35; 95% CI 1.09 to 1.66), and fetal *TCN2* (OR_{meta} = 1.38; 95% CI 1.12 to 1.70) that are associated with an increased risk of CTD among cases without folic acid supplementation. We detected putative de novo mutations in genes from the folate, homocysteine, and transsulfuration pathways and identified a significant association between rare variants in *MGST1* and CTD risk. Results suggest that periconceptional folic acid supplementation is associated with decreased risk of CTD among individuals with susceptible genotypes.

Keywords: folate; congenital; heart; NBDPS; *MTHFS*; gene environment

1. Introduction

Congenital heart defects are the most-common congenital abnormality [1]. Both genetic and environmental factors have been implicated in the occurrence of congenital heart defects [2]. Evidence suggests that dietary folic acid fortification and use of a folate-containing prenatal multivitamin are associated with a decreased risk of heart defects, including conotruncal heart defects (CTDs) [3–5]. Research from our laboratory and

others has shown that maternal and fetal genetic variants in the folate, homocysteine, and transsulfuration pathways are associated with an increased risk of heart defects, specifically CTDs [6,7]. However, it is unclear if the effects of variants in these folate-related pathways on CTD risk are modified by maternal folate intake.

CTDs are life-threatening heart malformations that affect cardiac outflow tracts and great arteries. Our laboratory has previously reported single-nucleotide polymorphisms (SNPs) in the folate, homocysteine, and transsulfuration pathways that were independently associated with CTDs, and we identified SNPs that were only associated with CTD risk when folic acid supplementation was considered. These results highlighted the need for further study and replication. Therefore, this study identifies and characterizes variants associated with CTDs within candidate genes of the folate, homocysteine, and transsulfuration pathways. Specifically, we sought to (1) replicate significant gene–folic acid interactions among CTD susceptibility loci from our initial study, and (2) identify rare and de novo variants in CTD-susceptibility loci.

Here, we present data showing that periconceptional folic acid supplementation is consistently associated with a lower risk of CTD among families with risk alleles in *MTHFS*, *NOS2*, or *TCN2*. We also report putative de novo mutations in folate-related genes that are significantly enriched among cases, and we identify rare variants within *MGST1* that are associated with risk of CTD.

2. Materials and Methods

2.1. Compliance with Ethical Standards

This study was approved by the Institutional Review Boards of the Centers for Disease Control and Prevention (CDC), the University of Arkansas for Medical Sciences, and the other study sites in the National Birth Defect Prevention Study (NBDPS). Informed consent was obtained for all study participants. The collection of data and biologic samples was funded by the Centers for Disease Control and Prevention's National Center on Birth Defects and Developmental Disabilities and was coordinated by the National Birth Defects Prevention Study. The authors have no conflicts of interest, either financial or non-financial, in the subject matter discussed in this manuscript.

2.2. Study Population

We refer here to our initial work of identifying CTD risk alleles as the discovery phase and refer to the targeted resequencing of CTD-susceptibility loci as the replication phase. Participants in the discovery phase (*n* = 4648) were initially genotyped using a custom DNA microarray that included 1536 SNPs that were represented across 60 genes within the folate, homocysteine, and transsulfuration pathways. We sequenced buccal-cell DNA from 536 of the 4648 participants from the discovery phase and 740 additional participants in the replication phase. Both the discovery and replication phases used a family-based, case-control hybrid design that included case and control families from the National Birth Defect Prevention Study (NBDPS). Details of the two study phases are provided in Table 1.

The NBDPS represents one of the largest population-based studies of non-syndromic birth defects ever conducted, covering an annual birth population of 482,000, or 10% of US births [8–10]. Cases for this study were identified across 10 NBDPS sites using state birth defects surveillance systems. Cases included live births, stillbirths, and elective terminations with one or more of the following conotruncal heart defects: truncus arteriosus, interrupted aortic arch type B, transposition of great arteries, double outlet right ventricle, conoventricular septal heart defects, tetralogy of Fallot, and pulmonary atresia with ventricular septal defect. Controls were randomly selected at each NBDPS site from birth hospitals or birth certificates.

Table 1. Comparison of discovery and replication studies.

	Discovery Phase Microarray [a]	Discovery Phase Sequencing [b]	Replication Phase Sequencing [b]
Sample size, post-QC (*n*)			
Total participants	4648	536	740
Case families	616	186	251
Case parental trios	230	97	230
Control families	1645	37	6
Control parental trios	559	29	2
Genomic Variants			
Number of assayed genes	60	13	13
Number of variants genotyped	1536	9402	9402
Number of variants imputed (post-QC)	3505	9588	9588
Imputed variants shared between studies	2945	2945	2945
Genotype concordance [c]		99.0%	NA
Imputed concordance [c]		97.9%	NA
Analysis			
Hybrid log-linear model of gene-folate interactions	Yes	NA	Yes
Region-based analysis for rare variants	No	Yes	Yes
De novo variant detection	No	Yes	Yes

[a] DNA microarray; [b] Probe hybridization enrichment and sequencing; [c] Discovery microarray vs. replication sequencing.

We used DNA samples from case and control families enrolled in the NBDPS from 1998 to 2011 to test the hypothesis that genetic variants in folate-related pathways are risk factors for CTD. The majority of families included in this study were complete mother–father–infant trios; however, mother–infant dyads, father–infant dyads, mother–father dyads, mother-only, father-only, and infant-only samples were genotyped as well (Supplementary Table S1). Mothers participating in the NBDPS were asked by phone interview about vitamin use during and prior to pregnancy, including the brand, frequency, duration, and timeframe of vitamin use. Women with periconceptional folic acid supplementation were identified as those reporting use of folic-acid-containing vitamins or supplements during the month before pregnancy or the first three months of gestation.

Sample inclusion criteria from the replication phase mirrored those of the discovery phase, except the replication phase was limited to samples with at least 50 ng of available DNA and a call rate of at least 90% if previously genotyped as part of the discovery phase.

2.3. Selection of Targeted Regions and Sequencing

In the discovery phase, we genotyped 1536 SNPs within 60 candidate genes in the homocysteine, folate, and transsulfuration pathways using Illumina's Golden Gate^{TM} custom platform (San Diego, CA, USA). SNPs were selected for genotyping in the discovery phase with the aim of providing a maximally informative set of haplotype-tagging SNPs that represented common variants (MAF > 0.1) across the 60 targeted genes [6]. In the replication phase, we sequenced exons, introns, and 1 kb upstream and downstream regions of 13 folate-related genes that were selected based on the rank order of gene variants from the discovery phase (prioritizing low Bayesian False Discovery Probability, BFDP) and the number of CTD-associated variants (BFDP < 0.8) within each gene. The 13 genes that were sequenced in the replication phase include (i) glutamate–cysteine ligase catalytic subunit (*GCLC*); (ii) serine hydroxymethyltransferase 1 (*SHMT1*); (iii) nitric oxide synthase 2 (*NOS2*); (iv) *5*-methyltetrahydrofolate-homocysteine methyltransferase reductase (*MTRR*); (v) 5,10-methenyltetrahydrofolate synthetase (*MTHFS*); (vi) glutaredoxin (*GLRX*); (vii) DNA methyltransferase 3 α (*DNMT3A*); (viii) superoxide dismutase 2, mitochondrial (*SOD2*); (ix) microsomal glutathione s-transferase 1 (*MGST1*); (x) glutathione peroxidase 4 (*GPX4*); (xi) transcobalamin 2 (*TCN2*); (xii) methylenetetrahydrofolate dehydrogenase (NADP + dependent) 2, methenyltetrahydrofolate cyclohydrolase (*MTHFD2*); (xiii) and

the enolase superfamily member 1 (*ENOSF1*)/thymidylate synthetase (*TYMS*) gene loci (*ENOSF1/TYMS*).

2.4. Sample Preparation

DNA was collected from the mother, father, and infant and was processed using well-established NBDPS and local Arkansas protocols [9]. Illumina Nextera hybridization enrichment was used to capture 13 candidate exons, introns, and 1 kb upstream and downstream regions from 50 ng per sample of genomic DNA. Target libraries were pooled at 48× and sequenced on an Illumina HiScan-SQ and NextSeq500.

2.5. Data Processing and Genotype Calling

Sequencing data were processed using Genome Analysis Toolkit (GATK) Best Practices for Germline SNPs and Indels in Whole Genomes and Exomes [11]. The basic steps are outlined here. Raw sequencing data were demultiplexed and converted to FASTQ format by CASAVA v1.8.2 and Illumina BaseSpace. Sequences were mapped to the reference genome (GRCh37/hg19) with the Burrows–Wheeler Aligner (BWA) and MEM algorithm [12]. Duplicate sequence reads were marked with Picard Tools. Base quality scores were recalibrated with GATK v3.3. GATK Haplotype Caller was used to generate per-sample Genomic VCF files, which were genotyped jointly to produce a multi-sample VCF. Resulting single-nucleotide variants were filtered out if any of the following criteria were met: quality by depth (QD) < 2.0, Fisher strand bias (FS) > 60.0, RMS mapping quality (MQ) < 40.0, haplotype score > 13.0, mapping quality rank sum < −12.5, read position rank sum test < −8.0. Indels were filtered out based on the following criteria: QD < 2.0, FS > 200.0, Read Position Rank Sum < −20.0. Filtered variants were phased among families using GATK PhaseByTransmission [13]. SnpEff v4.0 was used to annotate genomic variants with conservation and predictor scores from the dbNSFP v2.0 database [14].

2.6. Quality Control

Thirty-one samples from the present study were also genotyped with an Illumina HumanOmni 5M BeadChip as part of another project. Genotype calls were compared at 196 overlapping sites. The concordance rate across platforms was 99.0% before imputations, indicating high accuracy in variant calls. All samples were screened for Mendelian inconsistencies using PLINK [15]. Statistical analyses were independently replicated by a second statistician per NBDPS protocol. Sequencing methods were validated using a well-characterized sample from the 1000 Genomes project [16] as part of an NBDPS external quality assurance procedure.

In our replication phase, all families and SNPs showed Mendelian error rates of less than 5%. All genotypes with Mendelian inconsistency were set to missing. To ensure the genotyping quality, we further removed SNPs with a call rate of less than 95% and participants with a call rate of less than 70%.

2.7. Imputation

The targeted sequencing provided genotypes for 9402 variants in the replication phase. However, only 74 of those variants were genotyped in the discovery phase. To maximize the overlap of variants between the two phases, we conducted genotype imputation for each phase by using 1000 Genomes Project Phase 3 Genome Build b37 as reference panels. Software IMPUTE version 2 was used for the imputation with a "multi-population reference panels" option, which chooses a "custom" reference panel for each individual based on all available reference haplotypes [17,18]. The advantages of such a strategy have been discussed elsewhere [19]. After the imputation and quality control, the discovery phase and the replication phase included 3505 variants and 9588 variants, respectively. A total of 2945 variants overlapped between the two phases. We further compared the genotype concordance between two phases based on these 2945 variants from 536 overlapping

samples. The genotype concordance across platforms was estimated to be 97.9%, indicating a robust imputation result.

2.8. Gene–Folic Acid Interaction Analysis

There is strong evidence that periconceptional folic acid intake decreases the risk of conotruncal heart defects [20–27]. Folate-related genes, such as *MTHRF,* are known to influence the risk of heart defects via interactions with maternal folic acid intake [28]. We, therefore, hypothesized that genetic risk factors in CTD might be modified by periconceptional folic acid supplementation. We used an extended log-linear model, first proposed by Weinberg and Umbach, to investigate gene–folic acid interactions for each of the common variants [29]. This model simultaneously estimates the effects of maternal and fetal genotypes for a given SNP on disease risk [29]. It is also worthwhile to note that this model was initially proposed for analysis of case-parental trios and does not require samples from the control families [30]. Therefore, the estimated genetic relative risk does not rely on the contrast between case and control families. Additional details of the analysis are provided in the Supplementary File S1. We and others have previously used such modeling strategies for gene-by-environment analysis in studies with trio designs [6,7,31]. For the identified genetic variants, we report the relative genetic risk among 154 folate-supplemented families and 99 folate-unsupplemented families from the replication study phase.

We applied the log-linear model to the discovery phase and replication phase separately and further combined the results from both phases by fixed-effect meta-analysis with inverse variance weighting. The significance level of the interaction effect is evaluated by the Bayesian false-discovery probability (BFDP) for the meta-analysis results. When the associations between genetic variants and disease risk were consistent between discovery and replication phases, the BFDP for meta-analysis was expected to be smaller than those estimates from either the discovery or replication phases. Additional details of meta-analysis and BFDP derivatives are provided in the Supplementary File S1.

The BFDP assesses the noteworthiness of an association by balancing the costs of false discovery and non-discovery [32] and has been widely used in genetic association studies [33–37]. We used a widely accepted BFDP threshold of less than 0.8 for reporting noteworthy associations. A BFDP threshold of 0.8 implies that not reporting a noteworthy association (i.e., false non-discovery) is four-times as costly as reporting a non-noteworthy association (i.e., false discovery).

The statistical analysis was conducted in R 3.1.3 (R Development Core Team, Vienna, Austria) using the LEM package for log-linear modeling [38] and the METAL package for meta-analysis [39,40].

2.9. Detection and Testing of De Novo Variants

Phased genotyping calls from case and control trios were refined using GATK v3.3, based on the GATK genotype refinement workflow [11]. GATK VariantAnnotator was then used to annotate potential de novo mutations present within the set of refined, high-quality genotypes. This process resulted in 29 candidate de novo mutations among cases and 0 candidate mutations among controls. Further filtering of candidates based on quality and read depth yielded 12 likely de novo mutations among the cases. The observed frequency of likely de novo mutations among cases was compared to the predicted frequency within our targeted regions, based on the genome-wide human mutation rate map from Genome of the Netherlands Consortium [41,42].

2.10. Region-Based Analysis for Rare Variants

We conducted region-based association tests for rare variants (i.e., MAF < 5%) within each of the 13 targeted regions using the full replication dataset. The analysis was conducted by using two approaches: the rare-variant extensions of the transmission disequilibrium test (RV-TDT) and the family-based sequence kernel association test (FB-SKAT) [43,44]. Because these methods focus on the transmission of alleles from parents to offspring, they

are tolerant of population substructure and admixture. It was shown from previous studies that the two approaches may have their unique advantages under different disease etiology. The RV-TDT may have improved power over FB-SKAT when a relatively high proportion of variants is causal with the same direction of effect size (i.e., deleterious or protective), while FB-SKAT may have improved power when a relatively low proportion of variants is causal or they have bi-directional effects [45].

3. Results

3.1. Population

This study included 1276 participants from 476 families (436 case families and 40 control families) whose buccal-cell DNA underwent targeted sequencing, alignment of sequence reads to the human genome, variant calling, variant imputations, and quality control (see Materials and Methods for details on sample processing and Table 2 for demographic information). In total, 536 of these 1276 participants were previously genotyped in the discovery phase using a custom DNA microarray that covered 1536 SNPs within 60 genes in the folate, homocysteine, and transsulfuration pathways [6]. The remaining 740 of 1276 participants were not previously analyzed and were thus included in the replication phase. We utilized a hybrid log-linear model to test for interactions between prenatal folate supplementation and the genotype of 2945 SNPs that were represented across the 4684 participants in the discovery phase and 740 independent samples sequenced in the replication phase. Demographics for the discovery and replication sample sets are displayed in Supplementary Table S2.

Table 2. Maternal demographics of discovery- and replication-phase samples used for DNA sequencing.

		CTD (*n* = 436)	Controls (*n* = 40) [a]	*p*-Value [b]
Age at delivery (years)		28.3 (5.79)	27.1 (5.71)	0.23
Race	Caucasian	298 (68.7%)	33 (86.8%)	0.05
	Hispanic	75 (17.3%)	1 (2.6%)	
	African American	29 (6.7%)	2 (5.3%)	
	Other	32 (7.3%)	2 (5.3%)	
Education	Less than 12 years	45 (10.3%)	1 (2.6%)	0.25
	High school or equivalent	100 (23.0%)	13 (34.2%)	
	1–3 years of college	133 (30.6%)	10 (26.3%)	
	Bachelors or 4 years college	156 (35.9%)	14 (36.8%)	
Household income	Less than $10,000	46 (24.1%)	3 (15%)	0.58
	$10,000–$30,000	57 (29.8%)	9 (45%)	
	$30,000–$50,000	39 (20.4%)	4 (20%)	
	Over $50,000	49 (25.7%)	4 (20%)	
Folic acid supplementation	Unsupplemented	176 (40.6%)	11 (28.9%)	0.17
	Supplemented	258 (59.4%)	27 (71.1%)	
Alcohol consumption	Unexposed	285 (65.7%)	24 (63.2%)	0.73
	Exposed	149 (34.3%)	14 (36.8%)	
Cigarette smoking	Unexposed	364 (84.1%)	30 (78.9%)	0.49
	Exposed	69 (15.9%)	8 (28.1%)	
	Missing	1	0	
Maternal BMI	Underweight (BMI < 18.5)	19 (4.5%)	2 (5.3%)	0.04
	Normal wt. (BMI 18.5–25)	190 (44.8%)	25 (65.8%)	
	Overweight (BMI 25–30)	116 (27.4%)	8 (21.1%)	
	Obese (BMI > 30)	99 (23.3%)	3 (7.9%)	
	Missing	10	0	

CTD, conotruncal defect; BMI, body mass index. [a] Demographic information was unavailable for 2 of the 40 control families included in this study. [b] Values based on Fisher's exact test.

Secondary aims for this study included (i) testing for an excess of de novo mutations within candidate genes and (ii) conducting region-based association tests for rare variants. These subsequent tests utilized the full set of 1276 participants from 476 families that were sequenced in this study and passed quality control, including 360 complete trios, 55 mother–infant dyads, 12 father–infant dyads, 13 mother–father dyads, 10 mother singlets, 2 father singlets, and 24 infant singlets (Supplementary Table S1). The majority of case and control mothers were non-Hispanic white, had some college education, earned less than USD 30,000, reported folic acid supplementation, and did not smoke or consume alcohol. Tetralogy of Fallot and Transposition of great arteries were the two most-common types of conotruncal heart defects included in the study, representing 175/422 (41.5%) and 138/422 (32.7%) of cases, respectively (see Supplementary Table S3 for phenotypic distribution).

3.2. Gene–Folic Acid Interaction

The fraction of case mothers receiving periconceptional folate supplementation was not significantly different between specific types of conotruncal heart defects (p-value = 0.59, Fisher's exact test), as is shown in Supplementary Table S3. We identified 54 significant interactions ($p_{meta} < 0.05$, $BFDP_{meta} < 0.8$) between SNPs and folate supplementation status (Supplementary Table S4), such that the minor allele of either the mother or the infant increased the risk of CTD among infants without folic acid supplementation. Thirty-three of these SNPs had significant infant gene-by-folate interactions. These 33 SNPs were part of two linkage disequilibrium (LD) blocks: one in *MTHFS* and one in *TCN2* (see Supplementary Figure S1A,B for LD plots of SNPs with significant infant gene-by-folate interactions). The remaining 21 SNPs had significant maternal gene-by-folate interactions and were part of an LD block in *NOS2* and one in *MTHFS* (see Supplementary Figure S1C,D for LD plots of SNPs with significant maternal gene-by-folate interactions). Supplementary Figure S2 depicts the workflow for gene–folic acid interaction analysis as it relates to SNP discovery and validation in the two phases of this study.

Table 3 displays the lead SNPs with the strongest interaction effect (lowest metanalysis p-value) for each LD block. The SNP, rs12438477, in *MTHFS* was notable for having the strongest fetal gene-by-folate interaction effect ($p_{meta} = 4.02 \times 10^{-5}$ and $BFDP_{meta} = 0.12$). Among mothers reporting no periconceptional folic acid supplementation, the fetal minor allele (A) of rs12438477 was estimated to increase the risk of CTDs in the meta-analysis by approximately 1.41-fold per copy, 95% CI [1.11, 1.79]. Among mothers reporting periconceptional folic acid supplementation, the fetal minor allele appears to be protective for CTD (meta-analysis relative risk = 0.74, 95% CI (0.61, 0.92)). The direction and magnitude of effects were consistent across the discovery and replication studies.

Table 3. Gene–folic acid interactions in the discovery and replication phases.

Gene	SNP	Location	p-Value			BFDP			Folate Supp.	RR and 95% CI		
			Disc.	Rep.	Meta	Disc.	Rep.	Meta		Discovery	Replication	Meta
SNPs with fetal gene-by-folate interactions												
MTHFS	rs12438477 [a]	chr15:80,178,283	1.38×10^{-3}	8.58×10^{-3}	4.02×10^{-5}	0.61	0.88	0.12	No	1.49 [1.12,1.97]	1.26 [0.81,1.95]	1.41 [1.11,1.79]
									Yes	0.83 [0.65,1.07]	0.59 [0.41,0.85]	0.74 [0.61,0.92]
TCN2	rs2301957 [b]	chr22:31,018,817	1.2×10^{-2}	1.04×10^{-2}	4.66×10^{-4}	0.83	0.88	0.39	No	1.14 [0.86,1.52]	1.55 [1.01,2.39]	1.25 [0.99,1.59]
									Yes	0.72 [0.56,0.92]	0.75 [0.52,1.08]	0.73 [0.59,0.89]
SNPs with maternal gene-by-folate interactions												
NOS2	rs2779248 [c]	chr17:26,127,832	1.61×10^{-3}	2.23×10^{-2}	1.02×10^{-4}	0.65	0.91	0.21	No	1.40 [1.05,1.87]	1.18 [0.76,1.82]	1.33 [1.04,1.69]
									Yes	0.77 [0.60,0.99]	0.59 [0.40,0.89]	0.72 [0.58,0.89]
MTHFS	rs7163338 [d]	chr15:80,198,408	1.57×10^{-2}	6.15×10^{-3}	3.97×10^{-4}	0.85	0.85	0.34	No	1.27 [0.96,1.67]	1.49 [1.02,2.17]	1.34 [1.07,1.67]
									Yes	0.82 [0.65,1.05]	0.73 [0.52,1.03]	0.79 [0.65,0.96]

Disc., discovery phase; Rep., replication phase; Meta, Meta-analysis; RR, relative risk; CI, Confidence Interval; BFDP, Bayesian Measure of the Probability; Folate supp., maternal periconceptional folic acid supplementation. [a] An additional 12 variants were identified within the same LD block with rs12438477 (i.e., Chr. 15, BP 80,157,583–80,179,847). [b] An additional 19 variants were identified within the same LD block with rs2301957 (i.e., Chr. 22, BP 31,004,127–31,024,856). [c] An additional 11 variants were identified within the same LD block with rs2779248 (i.e., Chr. 17, BP 26,103,034–26,129,996). [d] An additional 8 variants were identified within the same LD block with rs7163338 (i.e., Chr. 15, BP 80,166,597–80,198,417).

3.3. De Novo Mutations

Although de novo mutations do not contribute directly to association signals, these mutations can help identify the precise location of disease-associated variants. We identified 12 de novo mutations based on the targeted sequencing data of 365 case-parental trios. Table 4 shows likely de novo mutations identified in the replication phase. All but two of the mutations affected regulatory regions of introns, such as transcription factor binding sites (TFBSs) and DNase footprints. Mutations at chr2:25,460,216 and chr15:8,012,657 did not appear to affect regulatory DNA; instead, these variants were predicted to affect the strength of cryptic splice sites. The likely de novo mutation, rs568740763, was notable in that it affected the *ENOSF1* 3' un-translated (UTR) and appeared to provide a new Poly-A-signal for the *ENOSF1* protein. All 12 of these putative de novo mutations were at low allele frequency in the general population (AF < 1%) or were absent from the Genome Aggregation Database v3.1.2, which contains 76,156 whole genomes [46].

Table 4. Putative de novo mutations identified in the replication phase with target sequencing.

Variant	Gene	Region	MAF [a]	Affected Regulatory Features [b]	Predicted Effects
chr2:25460216C > T	*DNMT3A*	Intron	6.6×10^{-6}	None	Enhances cryptic SSA
chr2:25532226T > C	*DNMT3A*	Intron	3.1×10^{-4}	TFBS (YY1), DNase peak	
chr2:25548310C > T	*DNMT3A*	Intron	1.5×10^{-4}	TFBS (GATA2, REST, ZNF263), DNase peak	Enhances cryptic SSD
chr2:25549757C > G	*DNMT3A*	Intron	None	TFBS (SETDB1), TF motif (EBF and EBF1), DNase Footprint, DNase peak	
chr2:25550547A > C	*DNMT3A*	Intron	None	TFBS (Freac-4)	
chr5:7875826A > G	*MTRR*	Intron	4.5×10^{-4}	TFBS (Bhlhb2)	
chr12:16503384C > A	*MGST1*	Intron	None	TFBS (SMARCA4)	
chr15:80162657C > T	*MTHFS*	Intron	3.2×10^{-3}	None	Enhances cryptic SSD
chr17:26101117A > G	*NOS2*	Intron	None	TFBS (NRSE, NRSF, NFAT5)	
chr17:26110823G > A	*NOS2*	Intron	2.0×10^{-3}	TFBS (ESR1)	
chr18:671738C > T	*ENOSF1- TYMS*	Exon- Intron	6.2×10^{-4}	TFBS (GATA1, POLR2A)	May provide new Poly-A signal for ENOSF1 3'UTR
chr22:31009950C > T	*TCN2*	Intron	None	DNase peak	Enhances cryptic SSA

[a] MAF, minor allele frequency provided by the Genome Aggregation Database (gnomAD v3.1.2 using liftover from GRCh37 to GRCh38); [b] Regulatory features from RegulomeDB; TFBS, transcription factor binding site; SSA, splice site acceptor; SSD, splice site donor.

The normal mutation rate for the selected sequenced regions was estimated to be $p_0 = 3.90408 \times 10^{-8}$, based on the genome-wide human mutation rate [41,42]. We tested for enrichment of de novo mutations within our case samples by using an exact test of binomial distribution. Under the null hypothesis that the mutation rate among our samples

was no different from the normal rate (i.e., $H_0 : p = p_0$), the probability of observing at least 12 de novo mutations among 365 trios is 1.22×10^{-67}, as shown below:

$$\sum_{k=12}^{k=365} \binom{365}{k} p_0^k (1 - p_0)^{365-k} = 1.22 \times 10^{-67}$$

This result indicates that the actual mutation rate among infant cases is higher than expected, assuming that the baseline rate of de novo mutations in the case population is similar to that of healthy individuals represented by the genome-wide human mutation rate map from Genome of the Netherlands Consortium [41,42].

3.4. Region-Based Analysis for Rare Variants

We used the FB-SKAT and RV-TDT methods to test for unexpectedly high transmission of genetic variants from parents to offspring. We applied these methods to each of the 13 genes and restricted the analysis to variants with a low allele frequency (MAF < 5%). Whereas the analysis described above (Gene–Folic Acid Interaction section) is well-suited to identifying common, disease-associated variants, these two methods are well-suited to identifying disease-associated regions/genes. These two region-based methods are complementary and their performance and results depend on disease mechanisms as discussed in the Materials and Methods section.

The results of FB-SKAT and the RV-TDT analyses are shown in Table 5. With a Bonferroni-corrected threshold for 13 regions (i.e., 0.05/13 = 0.0038), only *MGST1* showed statistical significance when analyzed by FB-SKAT. It is likely that this result was discordant with RV-TDT because of how these methods account for differences in the effect size and direction of variants within a region. *GLRX* and *GPX4* displayed nominal statistical significance ($p < 0.05$) that did not reach the Bonferroni-corrected threshold.

Table 5. Region-based analysis for rare variants.

Gene	Genomic Range	FB-SKAT	RV-TDT Burden Test
DNMT3A	chr2:25,454,844–25,566,444	0.628	0.755
MRHFD2	chr2:74,424,787–74,443,469	0.436	0.833
MTRR	chr5:7,868,246–7,902,214	0.727	0.867
GLRX	chr5:95,148,542–95,159,615	0.0147	0.392
GCLC	chr6:53,361,276–53,411,009	0.274	0.228
SOD2	chr6:160,099,857–160,115,421	0.217	0.434
MGST1	chr12:16,499,138–16,517,690	0.00298	0.67
MTHFS	chr15:80,134,820–80,190,536	0.24	0.605
SHMT1	chr17:18,230,107–18,267,686	0.066	0.212
NOS2	chr17:26,082,704–26,128,435	0.942	0.884
TYMS	chr18:657,140–674,589	0.238	0.322
GPX4	chr19:1,103,255–1,107,846	0.308	0.0384
TCN2	chr22:31,001,992–31,024,193	0.603	0.445

FB-SKAT, family-based sequence kernel association test; RV-TDT, rare-variant transmission disequilibrium test.

4. Discussion

4.1. Gene–Folic Acid Interaction

MTHFS polymorphisms. Testing for gene–folic acid interactions identified one maternal (rs7163338) and one fetal SNP (rs12438477) in the folate pathway gene *MTHFS* that were associated with CTD among infants with low folic acid supplementation ($p_{meta} < 0.05$, $BFDP_{meta} < 0.8$). MTHFS maintains the supply of the methyl donor 5,10-methenyltetrahydrofolate (5,10-methenyl-THF), which is a critical cofactor in the de novo synthesis of nucleotides thymidine, adenine, and guanine [47–49]. Of note, the lead *MTHFS* SNP in infants, rs12438477, was previously shown to increase the risk of congenital heart defects by 6.76-fold (95% CI [2.56,17.89]) among infants with the AA genotype and SSRI exposure [50].

TCN2 polymorphisms. Our meta-analysis identified fetal polymorphisms in *TCN2* that were associated with an increased risk of a CTD-affected pregnancy among women

with no periconceptional folic acid supplementation. The *TCN2* gene encodes transcobal-amin 2, which mediates transport of B12 to peripheral tissues and enables cellular B12 uptake [51]. Both B12 and circulating folate (5-methyltetrahydrofolate) are necessary for recycling homocysteine into methionine [52]. As a result, homocysteine increases when these two compounds are deficient, and it decreases when folic acid supplementation is initiated [53]. This likely explains why polymorphisms that affect the B12 transport protein *TCN2* have been associated with higher circulating homocysteine levels in the meta-analysis [54]. Of particular importance, our lead SNP within the *TCN2* locus, rs2301957, is associated with elevated blood homocysteine levels in prior research [55]. Elevated homocysteine is a risk factor for congenital heart defects, both in humans and in model systems. The risk of congenital heart defects can be over 4-fold higher among women with elevated homocysteine levels [56], whereas studies using the avian embryo model have demonstrated that homocysteine can induce CTDs in a manner that is dependent upon dose and time of exposure [57–59].

Our data suggest that folic-acid-containing vitamins mitigate the risk of CTD among infants with *TCN2* polymorphisms. Folate deficiency and *TCN2* polymorphisms are both associated with increased homocysteine; thus, hyperhomocysteinemia represents a potential mechanism for congenital heart defects. However, homocysteine was not measured in this study and further research is needed to fully resolve this question.

NOS2 polymorphisms. *NOS2* (inducible NOS) is an isoform of nitric oxide synthase (NOS) that produces nitric oxide in response to various environmental stimuli. Nitric oxide synthase is expressed and found in high abundance in heart tissue [60,61] and has multiple cardiovascular effects, including vasodilation, regulation of vessel permeability, anti-coagulation, angiogenesis, and formation of cardiac cushions [62–65].

Our results demonstrated that mothers with no folic acid supplementation and minor alleles at the *NOS2* risk loci have an increased risk of having CTD-affected offspring. The absence of folic acid supplementation predisposes these individuals to hyperhomocysteine-mia [53]. Hyperhomocysteinemia is known to inhibit angiogenesis [66,67] and increase the risk of heart malformations [56,68]. Prior studies have shown that the inhibitory effects of homocysteine on angiogenesis are dependent upon *NOS2* expression [66,67,69,70]. Thus, low folic acid intake and resulting hyperhomocysteinemia may interact with *NOS2* polymorphisms to influence the risk of CTD.

NOS2 also plays a major role in placental blood flow [71], response to infection [72], and development of diabetic embryopathy [73]. *NOS2* expression is dramatically increased in diabetic mice, and evidence suggests that excess *NOS2* contributes to diabetes-induced heart defects by increasing nitrosative stress and cell apoptosis [73–75]. It is unclear what effect *NOS2* risk alleles (Table 3 and Supplementary Table S4) might have on gene expression or activity within the developing heart. However, there are examples from the literature of *NOS2* gene variants within and near our risk loci, which are associated with increased NO production (e.g., rs4795067, rs2297518) [76,77].

This study replicated a maternal *NOS2* gene–folic acid interaction from the discovery phase. *NOS2* SNPs have not been well studied in the context of CTD. However, *NOS2* associations have been reported for various other conditions, including a GWAS association between rs4795067 and psoriasis [78]. This particular SNP (rs4795067) was identified as a CTD-risk allele in our study and has also been associated with neural tube defects in a large candidate gene study [79].

4.2. De Novo Mutations

De novo mutations were over-represented among our cases compared to reference mutation rates from the Genome of the Netherlands Consortium [41,42]. No de novo mutations were detected among 40 control families. Mutations appear to be especially common within *DNMT3A*; however, this is due, in part, to the comparatively large size of this gene (111.6 kb). Nonetheless, three of the *DNMT3A* de novo mutations are within significant regulatory regions. The variants at chr2:25,532,226 and chr2:25,548,310 affect

transcription factor binding sites (TFBS) within fetal heart enhancers, which represent important regulatory regions during fetal heart development [80], whereas the mutation at chr2:25,549,757 affects a region with abundant regulatory features and modest transcription in fetal heart tissue.

4.3. Region-Based Association Tests

MGST1 gene variants. Region-based association tests identified three regions that were nominally associated with CTD; however, only *MGST1* remained associated with CTD (Bonferroni-corrected $p < 0.0038$) after accounting for multiple testing. *MGST1* encodes a glutathione S-transferases enzyme, which protects the endoplasmic reticulum and outer mitochondrial membranes from xenobiotics, reactive metabolites, and oxidative stress [81]. Prior work in our lab identified rs7294985 as an *MGST1* polymorphism that influences the risk of heart defects among infants with prenatal exposure to selective serotonin reuptake inhibitors (SSRIs) [50]. Chowdhury et al. [82] identified five SNPs within *MGST1* that were associated with the ratio between reduced and oxidized glutathione (GSH/GSSG) among mothers. Cardiogenesis is sensitive to oxidative stress and xenobiotics [83,84]. It is, therefore, possible that harmful variants within *MGST1* sensitize the fetal heart to xenobiotics and reactive oxygen species.

4.4. Overview of Gene–Environment Interactions

GWAS studies of CTDs have identified several risk loci [85,86], including risk alleles at 12q24 and 13q32. However, prior GWAS of CTDs have not identified significant associations for the candidate genes included in this study. This may be due to the high significance threshold set by GWAS studies (1×10^{-8}) or may be a result of unmeasured gene–folic acid interactions in GWAS that have the potential to mask significant associations [87]. As the cost of sequencing decreases, one can envision a near future with routine genetic screening. Such screening will enable targeted prevention of heart defects based on gene–environmental interactions detailed in this and other studies.

The importance of gene–environment interactions is further illustrated by the results of this study. Folic acid supplementation was included as a covariate in this analysis based on evidence of a gene–folic acid interaction in the discovery phase. Inclusion of folic acid supplementation in the model unmasked relevant genetic associations and provides useful data for targeted interventions.

4.5. Strengths and Limitations

Limitations. This study was limited in the number of covariates in the analysis. The log-linear model used in the analysis has many advantages, including tolerance to population stratification for fetal effect [88]. However, one of its limitations is that it is quite inflexible to the inclusion of covariates. We chose to evaluate the interaction between gene variants and folic acid supplementation rather than other potential covariates because of evidence for this interaction in the discovery phase [6]. Although this approach was well-reasoned, it does not account for potential gene–environment interactions with factors, such as maternal obesity, diabetes, and medication exposure [50,75,89–93]. The study also has certain biases that are inherent to targeted gene analyses. For instance, it is limited to the detection of associations within the targeted genes and is not powered to detect associations at the genome-wide significance threshold of 5×10^{-8}. Although 1276 participants were sequenced as part of the replication study, 536 individuals were excluded from the meta-analysis due to an overlap between study phases. Nonetheless, the overlap in samples improved confidence in genotype calls and increased the density of genetic data. Furthermore, the full set of 1276 participants was utilized in rare-variant testing and de novo analysis.

Strengths. A major strength of this research was that it utilized samples from the NBDPS. The NBDPS has the advantage of being a large, multi-center, population-based, case-control study with consistency across centers afforded by well-validated protocols for

case ascertainment, maternal interview, sample collection, DNA extraction, case review, and phenotype classification [94]. This analysis benefited from a large sample size consisting of 4648 participants in the discovery phase and 1319 participants in the replication phase. It also benefited from a two-phase design, which generally reduces false positives and improves the detection of genetic associations [95]. The study was strengthened by the use of complementary methods, including microarray and sequencing. The use of both methods allowed us to confirm the accuracy of genotyping calls across multiple platforms and to conduct rare-variant burden testing among sequenced samples. An additional strength of the study was its trio family design, which provided a means of detecting de novo mutations and conducting family-based analyses [38]. By testing candidate genes with clear biologic relevance, this study was able to identify low-penetrant genetic variants that influence the risk of CTD.

5. Conclusions

Results suggest that folic acid supplementation reduces the risk of CTD among families with a maternal-risk allele in *MTHFS* (rs7163338) or *NOS2* (rs2779248) or a fetal-risk allele in *MTHFS* (rs12438477) or *TCN2* (rs2301957). We observed an excess in likely de novo mutations in folate, homocysteine, and transsulfuration genes among case families. We also found an excess of rare variants in *MGST1* among cases with CTD, suggesting that rare variants within this gene may also be linked to CTD. These results support current recommendations by the U. S. Public Health Service and CDC that women of reproductive age take a supplement containing 400–800 μg folic acid daily [96].

References

1. Zaidi, S.; Brueckner, M. Genetics and Genomics of Congenital Heart Disease. *Circ. Res.* **2017**, *120*, 923–940. [CrossRef] [PubMed]
2. Diab, N.S.; Barish, S.; Dong, W.; Zhao, S.; Allington, G.; Yu, X.; Kahle, K.T.; Brueckner, M.; Jin, S.C. Molecular Genetics and Complex Inheritance of Congenital Heart Disease. *Genes* **2021**, *12*, 1020. [CrossRef] [PubMed]
3. Wang, D.; Jin, L.; Zhang, J.; Meng, W.; Ren, A. Maternal Periconceptional Folic Acid Supplementation and Risk for Fetal Congenital Heart Defects. *J. Pediatr.* **2022**, *240*, 72–78. [CrossRef]
4. Feng, Y.; Wang, S.; Chen, R.; Tong, X.; Wu, Z.; Mo, X. Maternal Folic Acid Supplementation and the Risk of Congenital Heart Defects in Offspring: A Meta-Analysis of Epidemiological Observational Studies. *Sci. Rep.* **2015**, *5*, 8506. [CrossRef] [PubMed]
5. Canfield, M.A.; Collins, J.S.; Botto, L.D.; Williams, L.J.; Mai, C.T.; Kirby, R.S.; Pearson, K.; Devine, O.; Mulinare, J. National Birth Defects Prevention Network Changes in the Birth Prevalence of Selected Birth Defects after Grain Fortification with Folic Acid in the United States: Findings from a Multi-State Population-Based Study. *Birth Defects Res. A Clin. Mol. Teratol.* **2005**, *73*, 679–689. [CrossRef] [PubMed]
6. Hobbs, C.A.; Cleves, M.A.; MacLeod, S.L.; Erickson, S.W.; Tang, X.; Li, J.; Li, M.; Nick, T.; Malik, S. Conotruncal Heart Defects and Common Variants in Maternal and Fetal Genes in Folate, Homocysteine, and Transsulfuration Pathways. *Birth Defects Res. A Clin. Mol. Teratol.* **2014**, *100*, 116–126. [CrossRef]
7. Li, M.; Erickson, S.W.; Hobbs, C.A.; Li, J.; Tang, X.; Nick, T.G.; Macleod, S.L.; Cleves, M.A. Detecting Maternal-Fetal Genotype Interactions Associated with Conotruncal Heart Defects: A Haplotype-Based Analysis with Penalized Logistic Regression. *Genet. Epidemiol.* **2014**, *38*, 198–208. [CrossRef]
8. Rasmussen, S.A.; Lammer, E.J.; Shaw, G.M.; Finnell, R.H.; McGehee, R.E., Jr.; Gallagher, M.; Romitti, P.A.; Murray, J.C.; McGehee, R.E.; Gallagher, M.; et al. Integration of DNA Sample Collection into a Multi-site Birth Defects Case-control Study. *Teratology* **2002**, *66*, 177–184. [CrossRef]
9. Yoon, P.W.; Rasmussen, S.A.; Lynberg, M.C.; Moore, C.A.; Anderka, M.; Carmichael, S.L.; Costa, P.; Druschel, C.; Hobbs, C.A.; Romitti, P.A. The National Birth Defects Prevention Study. *Public Health Rep.* **2001**, *116*, 32–40. [CrossRef]
10. Gallagher, M.L.; Sturchio, C.; Smith, A.; Koontz, D.; Jenkins, M.M.; Honein, M.A.; Rasmussen, S.A. 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. [CrossRef]
11. Van der Auwera, G.A.; Carneiro, M.O.; Hartl, C.; Poplin, R.; del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J.; et al. From FastQ Data to High Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr. Protoc. Bioinformatics* **2013**, *43*, 11.10.1–11.10.33. [CrossRef] [PubMed]
12. Li, H.; Durbin, R. Fast and Accurate Long-Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **2010**, *26*, 589–595. [CrossRef] [PubMed]
13. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; del Angel, G.; Rivas, M.A.; Hanna, M. A Framework for Variation Discovery and Genotyping Using Next-Generation DNA Sequencing Data. *Nat. Genet.* **2011**, *43*, 491–498. [CrossRef] [PubMed]
14. Liu, X.; Jian, X.; Boerwinkle, E. DbNSFP v2.0: A Database of Human Non-synonymous SNVs and Their Functional Predictions and Annotations. *Hum. Mutat.* **2013**, *34*, E2393–E2402. [CrossRef] [PubMed]
15. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]
16. Stewart, C.; Kural, D.; Strömberg, M.P.; Walker, J.A.; Konkel, M.K.; Stütz, A.M.; Urban, A.E.; Grubert, F.; Lam, H.Y.K.; Lee, W.-P. A Comprehensive Map of Mobile Element Insertion Polymorphisms in Humans. *PLoS Genet.* **2011**, *7*, e1002236. [CrossRef]
17. Marchini, J.; Howie, B.; Myers, S.; McVean, G.; Donnelly, P. A New Multipoint Method for Genome-Wide Association Studies by Imputation of Genotypes. *Nat. Genet.* **2007**, *39*, 906–913. [CrossRef]
18. Howie, B.N.; Donnelly, P.; Marchini, J. A Flexible and Accurate Genotype Imputation Method for the next Generation of Genome-Wide Association Studies. *PLoS Genet.* **2009**, *5*, e1000529. [CrossRef]

19. Howie, B.; Marchini, J.; Stephens, M. Genotype Imputation with Thousands of Genomes. *G3 Genes|Genomes|Genet.* **2011**, *1*, 457–470. [CrossRef]

20. Botto, L.D.; Mulinare, J.; Erickson, J.D. Occurrence of Congenital Heart Defects in Relation to Maternal Multivitamin Use. *Am. J. Epidemiol.* **2000**, *151*, 878–884. [CrossRef]

21. Botto, L.D.; Khoury, M.J.; Mulinare, J.; Erickson, J.D. Periconceptional Multivitamin Use and the Occurrence of Conotruncal Heart Defects: Results from a Population-Based, Case-Control Study. *Pediatrics* **1996**, *98*, 911–917. [CrossRef] [PubMed]

22. Shaw, G.M.; O'Malley, C.D.; Wasserman, C.R.; Tolarova, M.M.; Lammer, E.J. Maternal Periconceptional Use of Multivitamins and Reduced Risk for Conotruncal Heart Defects and Limb Deficiencies among Offspring. *Am. J. Med. Genet.* **1995**, *59*, 536–545. [CrossRef] [PubMed]

23. Czeizel, A.E. Reduction of Urinary Tract and Cardiovascular Defects by Periconceptional Multivitamin Supplementation. *Am. J. Med. Genet.* **1996**, *62*, 179–183. [CrossRef]

24. Czeizel, A.E.; Toth, M.; Rockenbauer, M. Population-Based Case Control Study of Folic Acid Supplementation during Pregnancy. *Teratology* **1996**, *53*, 345–351. [CrossRef]

25. Van Beynum, I.M.; Kapusta, L.; Bakker, M.K.; den Heijer, M.; Blom, H.J.; de Walle, H.E.K.; van Beynum, I.M.; Kapusta, L.; Bakker, M.K.; den Heijer, M.; et al. Protective Effect of Periconceptional Folic Acid Supplements on the Risk of Congenital Heart Defects: A Registry-Based Case-Control Study in the Northern Netherlands. *Eur. Heart J.* **2010**, *31*, 464–471. [CrossRef]

26. Ionescu-Ittu, R.; Marelli, A.J.; Mackie, A.S.; Pilote, L. Prevalence of Severe Congenital Heart Disease after Folic Acid Fortification of Grain Products: Time Trend Analysis in Quebec, Canada. *BMJ* **2009**, *338*, b1673. [CrossRef]

27. Goh, Y.I.; Bollano, E.; Einarson, T.R.; Koren, G. Prenatal Multivitamin Supplementation and Rates of Congenital Anomalies: A Meta-Analysis. *J. Obstet. Gynaecol. Can.* **2006**, *28*, 680–689.

28. Van Beynum, I.M.; den Heijer, M.; Blom, H.J.; Kapusta, L. The MTHFR 677C → T Polymorphism and the Risk of Congenital Heart Defects: A Literature Review and Meta-Analysis. *QJM Int. J. Med.* **2007**, *100*, 743–753. [CrossRef]

29. Weinberg, C.R.; Umbach, D.M. A Hybrid Design for Studying Genetic Influences on Risk of Diseases with Onset Early in Life. *Am. J. Hum. Genet.* **2005**, *77*, 627–636. [CrossRef]

30. Weinberg, C.R.; Wilcox, A.J.; Lie, R.T. A Log-Linear Approach to Case-Parent-Triad Data: Assessing Effects of Disease Genes That Act Either Directly or through Maternal Effects and That May Be Subject to Parental Imprinting. *Am. J. Hum. Genet.* **1998**, *62*, 969–978. [CrossRef]

31. Tang, X.; Cleves, M.A.; Nick, T.G.; Li, M.; MacLeod, S.L.; Erickson, S.W.; Li, J.; Shaw, G.M.; Mosley, B.S.; Hobbs, C.A. Obstructive Heart Defects Associated with Candidate Genes, Maternal Obesity, and Folic Acid Supplementation. *Am. J. Med. Genet. A* **2015**, *167*, 1231–1242. [CrossRef] [PubMed]

32. Wakefield, J. A Bayesian Measure of the Probability of False Discovery in Genetic Epidemiology Studies. *Am. J. Hum. Genet.* **2007**, *81*, 208–227. [CrossRef] [PubMed]

33. Liu, Y.; Shete, S.; Etzel, C.J.; Scheurer, M.; Alexiou, G.; Armstrong, G.; Tsavachidis, S.; Liang, F.-W.; Gilbert, M.; Aldape, K. 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. [CrossRef] [PubMed]

34. Oh, S.S.; Chang, S.; Cai, L.; Cordon-Cardo, C.; Ding, B.; Greenland, S.; He, N.; Jiang, Q.; Kheifets, L.; Le, A. Single Nucleotide Polymorphisms of 8 Inflammation-related Genes and Their Associations with Smoking-related Cancers. *Int. J. Cancer* **2010**, *127*, 2169–2182. [CrossRef]

35. Park, S.L.; Bastani, D.; Goldstein, B.Y.; Chang, S.-C.; Cozen, W.; Cai, L.; Cordon-Cardo, C.; Ding, B.; Greenland, S.; He, N. Associations between NBS1 Polymorphisms, Haplotypes, and Smoking-Related Cancers. *Carcinogenesis* **2010**, *31*, 1264–1271. [CrossRef]

36. Spitz, M.R.; Gorlov, I.P.; Dong, Q.; Wu, X.; Chen, W.; Chang, D.W.; Etzel, C.J.; Caporaso, N.E.; Zhao, Y.; Christiani, D.C.; et al. Multistage Analysis of Variants in the Inflammation Pathway and Lung Cancer Risk in Smokers. *Cancer Epidemiol. Biomark. Prev.* **2012**, *21*, 1213–1221. [CrossRef]

37. Zienolddiny, S.; Haugen, A.; Lie, J.-A.S.; Kjuus, H.; Anmarkrud, K.H.; Kjærheim, 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. [CrossRef]

38. Vermunt, J.K. *Log-Linear Event History Analysis: A General Approach with Missing Data, Latent Variables, and Unobserved Heterogeneity*; Tilburg University Press: Tilburg, The Netherlands, 1996; Volume 8.

39. Willer, C.J.; Li, Y.; Abecasis, G.R. METAL: Fast and Efficient Meta-Analysis of Genomewide Association Scans. *Bioinformatics* **2010**, *26*, 2190–2191. [CrossRef]

40. Willer, C.J.; Sanna, S.; Jackson, A.U.; Scuteri, A.; Bonnycastle, L.L.; Clarke, R.; Heath, S.C.; Timpson, N.J.; Najjar, S.S.; Stringham, H.M. Newly Identified Loci That Influence Lipid Concentrations and Risk of Coronary Artery Disease. *Nat. Genet.* **2008**, *40*, 161–169. [CrossRef]

41. The Genome of the Netherlands Consortium. Whole-Genome Sequence Variation, Population Structure and Demographic History of the Dutch Population. *Nat. Genet.* **2014**, *46*, 818–825. [CrossRef]

42. Francioli, L.C.; Polak, P.P.; Koren, A.; Menelaou, A.; Chun, S.; Renkens, I.; van Duijn, C.M.; Swertz, M.; Wijmenga, C.; van Ommen, G. Genome-Wide Patterns and Properties of de Novo Mutations in Humans. *Nat. Genet.* **2015**, *47*, 822–826. [CrossRef] [PubMed]

43. He, Z.; O'Roak, B.J.; Smith, J.D.; Wang, G.; Hooker, S.; Santos-Cortez, R.L.P.; Li, B.; Kan, M.; Krumm, N.; Nickerson, D.A. Rare-Variant Extensions of the Transmission Disequilibrium Test: Application to Autism Exome Sequence Data. *Am. J. Hum. Genet.* **2014**, *94*, 33–46. [CrossRef] [PubMed]

44. Ionita-Laza, I.; Lee, S.; Makarov, V.; Buxbaum, J.D.; Lin, X. Family-Based Association Tests for Sequence Data, and Comparisons with Population-Based Association Tests. *Eur. J. Hum. Genet.* **2013**, *21*, 1158–1162. [CrossRef]

45. Wu, M.C.; Lee, S.; Cai, T.; Li, Y.; Boehnke, M.; Lin, X. Rare-Variant Association Testing for Sequencing Data with the Sequence Kernel Association Test. *Am. J. Hum. Genet.* **2011**, *89*, 82–93. [CrossRef] [PubMed]

46. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alföldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The Mutational Constraint Spectrum Quantified from Variation in 141,456 Humans. *Nature* **2020**, *581*, 434–443. [CrossRef] [PubMed]

47. Dayan, A.; Bertrand, R.; Beauchemin, M.; Chahla, D.; Mamo, A.; Filion, M.; Skup, D.; Massie, B.; Jolivet, J. Cloning and Characterization of the Human 5, 10-Methenyltetrahydrofolate Synthetase-Encoding CDNA. *Gene* **1995**, *165*, 307–311. [CrossRef] [PubMed]

48. Girgis, S.; Nasrallah, I.M.; Suh, J.R.; Oppenheim, E.; Zanetti, K.A.; Mastri, M.G.; Stover, P.J. Molecular Cloning, Characterization and Alternative Splicing of the Human Cytoplasmic Serine Hydroxymethyltransferase Gene. *Gene* **1998**, *210*, 315–324. [CrossRef] [PubMed]

49. Field, M.S.; Anderson, D.D.; Stover, P.J. Mthfs Is an Essential Gene in Mice and a Component of the Purinosome. *Front. Genet.* **2011**, *2*, 36. [CrossRef]

50. Nembhard, W.N.; Tang, X.; Hu, Z.; Macleod, S.; Stowe, Z.; Webber, D. Maternal and Infant Genetic Variants, Maternal Periconceptional Use of Selective Serotonin Reuptake Inhibitors, and Risk of Congenital Heart Defects in Offspring: Population Based Study. *BMJ* **2017**, *356*, j832. [CrossRef]

51. Alam, A.; Woo, J.-S.; Schmitz, J.; Prinz, B.; Root, K.; Chen, F.; Bloch, J.S.; Zenobi, R.; Locher, K.P. Structural Basis of Transcobalamin Recognition by Human CD320 Receptor. *Nat Commun* **2016**, *7*, 12100. [CrossRef]

52. Selhub, J. Homocysteine Metabolism. *Annu. Rev. Nutr.* **1999**, *19*, 217–246. [CrossRef] [PubMed]

53. Homocysteine Lowering Trialists' Collaboration. Dose-Dependent Effects of Folic Acid on Blood Concentrations of Homocysteine: A Meta-Analysis of the Randomized Trials. *Am. J. Clin. Nutr.* **2005**, *82*, 806–812. [CrossRef] [PubMed]

54. Oussalah, A.; Levy, J.; Filhine-Tresarrieu, P.; Namour, F.; Gueant, J.L. Association of TCN2 Rs1801198 c.776G>C Polymorphism with Markers of One-Carbon Metabolism and Related Diseases: A Systematic Review and Meta-Analysis of Genetic Association Studies. *Am. J. Clin. Nutr.* **2017**, *106*, 1142–1156. [CrossRef] [PubMed]

55. Hsu, F.C.; Sides, E.G.; Mychaleckyj, J.C.; Worrall, B.B.; Elias, G.A.; Liu, Y.; Chen, W.M.; Coull, B.M.; Toole, J.F.; Rich, S.S.; et al. Transcobalamin 2 Variant Associated with Poststroke Homocysteine Modifies Recurrent Stroke Risk. *Neurology* **2011**, *77*, 1543–1550. [CrossRef] [PubMed]

56. Verkleij-Hagoort, A.; Bliek, J.; Sayed-Tabatabaei, F.; Ursem, N.; Steegers, E.; Steegers-Theunissen, R. Hyperhomocysteinemia and MTHFR Polymorphisms in Association with Orofacial Clefts and Congenital Heart Defects: A Meta-Analysis. *Am. J. Med. Genet. A* **2007**, *143*, 952–960. [CrossRef]

57. Rosenquist, T.H.; Finnell, R.H. Genes, Folate and Homocysteine in Embryonic Development. *Proc. Nutr. Soc.* **2001**, *60*, 53–61. [CrossRef]

58. Boot, M.J.; Steegers-Theunissen, R.P.; Poelmann, R.E.; van Iperen, L.; Groot, A.C.G. Cardiac Outflow Tract Malformations in Chick Embryos Exposed to Homocysteine. *Cardiovasc. Res.* **2004**, *64*, 365–373. [CrossRef]

59. Rosenquist, T.H.; Ratashak, S.A.; Selhub, J. Homocysteine Induces Congenital Defects of the Heart and Neural Tube: Effect of Folic Acid. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 15227–15232. [CrossRef]

60. Kolker, E.; Higdon, R.; Haynes, W.; Welch, D.; Broomall, W.; Lancet, D.; Stanberry, L.; Kolker, N. MOPED: Model Organism Protein Expression Database. *Nucleic Acids Res.* **2011**, *40*, D1093–D1099. [CrossRef]

61. Wilhelm, M.; Schlegl, J.; Hahne, H.; Gholami, A.M.; Lieberenz, M.; Savitski, M.M.; Ziegler, E.; Butzmann, L.; Gessulat, S.; Marx, H. Mass-Spectrometry-Based Draft of the Human Proteome. *Nature* **2014**, *509*, 582–587. [CrossRef]

62. Aicher, D.; Urbich, C.; Zeiher, A.; Dimmeler, S.; Schäfers, H.-J. Endothelial Nitric Oxide Synthase in Bicuspid Aortic Valve Disease. *Ann. Thorac. Surg.* **2007**, *83*, 1290–1294. [CrossRef] [PubMed]

63. Dimmeler, S.; Fleming, I.; Fisslthaler, B.; Hermann, C.; Busse, R.; Zeiher, A.M. Activation of Nitric Oxide Synthase in Endothelial Cells by Akt-Dependent Phosphorylation. *Nature* **1999**, *399*, 601–605. [CrossRef] [PubMed]

64. Feng, Q.; Song, W.; Lu, X.; Hamilton, J.A.; Lei, M.; Peng, T.; Yee, S.P. Development of Heart Failure and Congenital Septal Defects in Mice Lacking Endothelial Nitric Oxide Synthase. *Circulation* **2002**, *106*, 873–879. [CrossRef] [PubMed]

65. Lee, T.C.; Zhao, Y.D.; Courtman, D.W.; Stewart, D.J. Abnormal Aortic Valve Development in Mice Lacking Endothelial Nitric Oxide Synthase. *Circulation* **2000**, *101*, 2345–2348. [CrossRef]

66. Duan, J.; Murohara, T.; Ikeda, H.; Sasaki, K.; Shintani, S.; Akita, T.; Shimada, T.; Imaizumi, T. Hyperhomocysteinemia Impairs Angiogenesis in Response to Hindlimb Ischemia. *Arterioscler. Thromb. Vasc. Biol.* **2000**, *20*, 2579–2585. [CrossRef]

67. Nagai, Y.; Tasaki, H.; Takatsu, H.; Nihei, S.; Yamashita, K.; Toyokawa, T.; Nakashima, Y. Homocysteine Inhibits Angiogenesis in Vitro and in Vivo. *Biochem. Biophys. Res. Commun.* **2001**, *281*, 726–731. [CrossRef]

68. Kang, S.-S.; Wong, P.W.K.; Norusis, M. Homocysteinemia Due to Folate Deficiency. *Metabolism* **1987**, *36*, 458–462. [CrossRef]

69. Mayo, J.N.; Chen, C.; Liao, F.; Bearden, S.E. Homocysteine Disrupts Outgrowth of Microvascular Endothelium by an INOS-Dependent Mechanism. *Microcirculation* **2014**, *21*, 541–550. [CrossRef]

70. Mayo, J.N.; Chen, C.-H.; Bearden, S.E. Homocysteine Disrupts Actin Organization and Vascular Outgrowth by an INOS-Dependent Mechanism during Angiogenesis. *FASEB J.* **2013**, *27*, 685.5. [CrossRef]

71. Matsubara, K.; Matsubara, Y.; Hyodo, S.; Katayama, T.; Ito, M. Role of Nitric Oxide and Reactive Oxygen Species in the Pathogenesis of Preeclampsia. *J. Obstet. Gynaecol. Res.* **2010**, *36*, 239–247. [CrossRef]

72. Bogdan, C. Nitric Oxide Synthase in Innate and Adaptive Immunity: An Update. *Trends Immunol.* **2015**, *36*, 161–178. [CrossRef] [PubMed]

73. Yang, P.; Cao, Y.; Li, H. Hyperglycemia Induces Inducible Nitric Oxide Synthase Gene Expression and Consequent Nitrosative Stress via C-Jun N-Terminal Kinase Activation. *Am. J. Obstet. Gynecol.* **2010**, *203*, 185.e5–185.e11. [CrossRef] [PubMed]

74. Jawerbaum, A.; Gonzalez, E.T.; Novaro, V.; Faletti, A.; Sinner, D.; Gimeno, M.A.F. Increased Prostaglandin E Generation and Enhanced Nitric Oxide Synthase Activity in the Non-Insulin-Dependent Diabetic Embryo during Organogenesis. *Reprod. Fertil. Dev.* **1998**, *10*, 191–196. [CrossRef] [PubMed]

75. Correa, A.; Gilboa, S.M.; Besser, L.M.; Botto, L.D.; Moore, C.A.; Hobbs, C.A.; Cleves, M.A.; Riehle-Colarusso, T.J.; Waller, D.K.; Reece, E.A. Diabetes Mellitus and Birth Defects. *Am. J. Obstet. Gynecol.* **2008**, *199*, 237.e1–237.e9. [CrossRef]

76. Bouzigon, E.; Monier, F.; Boussaha, M.; le Moual, N.; Huyvaert, H.; Matran, R.; Letort, S.; Bousquet, J.; Pin, I.; Lathrop, M. Associations between Nitric Oxide Synthase Genes and Exhaled NO-Related Phenotypes According to Asthma Status. *PLoS ONE* **2012**, *7*, e36672. [CrossRef]

77. Dhillon, S.S.; Mastropaolo, L.A.; Murchie, R.; Griffiths, C.; Thöni, C.; Elkadri, A.; Xu, W.; Mack, A.; Walters, T.; Guo, C. Higher Activity of the Inducible Nitric Oxide Synthase Contributes to Very Early Onset Inflammatory Bowel Disease. *Clin. Transl. Gastroenterol.* **2014**, *5*, e46. [CrossRef]

78. Stuart, P.E.; Nair, R.P.; Ellinghaus, E.; Ding, J.; Tejasvi, T.; Gudjonsson, J.E.; Li, Y.; Weidinger, S.; Eberlein, B.; Gieger, C. Genome-Wide Association Analysis Identifies Three Psoriasis Susceptibility Loci. *Nat. Genet.* **2010**, *42*, 1000–1004. [CrossRef]

79. Soldano, K.L.; Garrett, M.E.; Cope, H.L.; Rusnak, J.M.; Ellis, N.J.; Dunlap, K.L.; Speer, M.C.; Gregory, S.G.; Ashley-Koch, A.E. Genetic Association Analyses of Nitric Oxide Synthase Genes and Neural Tube Defects Vary by Phenotype. *Birth Defects Res. B Dev. Reprod. Toxicol.* **2013**, *98*, 365–373. [CrossRef]

80. May, D.; Blow, M.J.; Kaplan, T.; McCulley, D.J.; Jensen, B.C.; Akiyama, J.A.; Holt, A.; Plajzer-Frick, I.; Shoukry, M.; Wright, C. Large-Scale Discovery of Enhancers from Human Heart Tissue. *Nat. Genet.* **2012**, *44*, 89–93. [CrossRef]

81. Kelner, M.J.; Bagnell, R.D.; Montoya, M.A.; Estes, L.A.; Forsberg, L.; Morgenstern, R. Structural Organization of the Microsomal Glutathione S-Transferase Gene (MGST1) on Chromosome 12p13.1-13.2. Identification of the Correct Promoter Region and Demonstration of Transcriptional Regulation in Response to Oxidative Stress. *J. Biol. Chem.* **2000**, *275*, 13000–13006. [CrossRef]

82. Chowdhury, S.; Hobbs, C.A.; MacLeod, S.L.; Cleves, M.A.; Melnyk, S.; James, S.J.; Hu, P.; Erickson, S.W. Associations between Maternal Genotypes and Metabolites Implicated in Congenital Heart Defects. *Mol. Genet. Metab.* **2012**, *107*, 596–604. [CrossRef]

83. Morgan, S.C.; Relaix, F.; Sandell, L.L.; Loeken, M.R. Oxidative Stress during Diabetic Pregnancy Disrupts Cardiac Neural Crest Migration and Causes Outflow Tract Defects. *Birth Defects Res. A Clin. Mol. Teratol.* **2008**, *82*, 453–463. [CrossRef]

84. Wells, P.G.; McCallum, G.P.; Chen, C.S.; Henderson, J.T.; Lee, C.J.; Perstin, J.; Preston, T.J.; Wiley, M.J.; Wong, A.W. Oxidative Stress in Developmental Origins of Disease: Teratogenesis, Neurodevelopmental Deficits, and Cancer. *Toxicol. Sci.* **2009**, *108*, 4–18. [CrossRef] [PubMed]

85. Agopian, A.J.; Mitchell, L.E.; Glessner, J.; Bhalla, A.D.; Sewda, A.; Hakonarson, H.; Goldmuntz, E. Genome-Wide Association Study of Maternal and Inherited Loci for Conotruncal Heart Defects. *PLoS ONE* **2014**, *9*, e96057. [CrossRef] [PubMed]

86. Cordell, H.J.; Topf, A.; Mamasoula, C.; Postma, A.V.; Bentham, J.; Zelenika, D.; Heath, S.; Blue, G.; Cosgrove, C.; Riveron, J.G.; et al. Genome-Wide Association Study Identifies Loci on 12q24 and 13q32 Associated with Tetralogy of Fallot. *Hum. Mol. Genet.* **2013**, *22*, 1473–1481. [CrossRef]

87. McCarthy, M.I.; Hirschhorn, J.N. Genome-Wide Association Studies: Potential next Steps on a Genetic Journey. *Hum. Mol. Genet.* **2008**, *17*, R156–R165. [CrossRef]

88. Price, A.L.; Zaitlen, N.A.; Reich, D.; Patterson, N. New Approaches to Population Stratification in Genome-Wide Association Studies. *Nat. Rev. Genet.* **2010**, *11*, 459–463. [CrossRef]

89. Andrade, S.E.; Raebel, M.A.; Brown, J.; Lane, K.; Livingston, J.; Boudreau, D.; Rolnick, S.J.; Roblin, D.; Smith, D.H.; Willy, M.E. Use of Antidepressant Medications during Pregnancy: A Multisite Study. *Am. J. Obstet. Gynecol.* **2008**, *198*, 194.e1–194.e5. [CrossRef]

90. Waller, D.K.; Shaw, G.M.; Rasmussen, S.A.; Hobbs, C.A.; Canfield, M.A.; Siega-Riz, A.-M.; Gallaway, M.S.; Correa, A. Prepregnancy Obesity as a Risk Factor for Structural Birth Defects. *Arch. Pediatr. Adolesc. Med.* **2007**, *161*, 745–750. [CrossRef] [PubMed]

91. Crider, K.S.; Cleves, M.A.; Reefhuis, J.; Berry, R.J.; Hobbs, C.A.; Hu, D.J. Antibacterial Medication Use during Pregnancy and Risk of Birth Defects: National Birth Defects Prevention Study. *Arch. Pediatr. Adolesc. Med.* **2009**, *163*, 978. [CrossRef]

92. Malik, S.; Cleves, M.A.; Honein, M.A.; Romitti, P.A.; Botto, L.D.; Yang, S.; Hobbs, C.A. Maternal Smoking and Congenital Heart Defects. *Pediatrics* **2008**, *121*, e810–e816. [CrossRef]

93. Ethen, M.K.; Ramadhani, T.A.; Scheuerle, A.E.; Canfield, M.A.; Wyszynski, D.F.; Druschel, C.M.; Romitti, P.A. Alcohol Consumption by Women before and during Pregnancy. *Matern. Child Health J.* **2009**, *13*, 274–285. [CrossRef] [PubMed]

94. Reefhuis, J.; Gilboa, S.M.; Anderka, M.; Browne, M.L.; Feldkamp, M.L.; Hobbs, C.A.; Jenkins, M.M.; Langlois, P.H.; Newsome, K.B.; Olshan, A.F. The National Birth Defects Prevention Study: A Review of the Methods. *Birth Defects Res. A Clin. Mol. Teratol.* **2015**, *103*, 656–669. [CrossRef] [PubMed]

95. Skol, A.D.; Scott, L.J.; Abecasis, G.R.; Boehnke, M. Joint Analysis Is More Efficient than Replication-Based Analysis for Two-Stage Genome-Wide Association Studies. *Nat. Genet.* **2006**, *38*, 209–213. [CrossRef] [PubMed]

96. Bibbins-Domingo, K.; Grossman, D.C.; Curry, S.J.; Davidson, K.W.; Epling, J.W.; García, F.A.R.; Kemper, A.R.; Krist, A.H.; Kurth, A.E.; Landefeld, C.S. Folic Acid Supplementation for the Prevention of Neural Tube Defects: US Preventive Services Task Force Recommendation Statement. *JAMA* **2017**, *317*, 183–189.