

Article

The Needle in the Haystack—Searching for Genetic and Epigenetic Differences in Monozygotic Twins Discordant for Tetralogy of Fallot

Marcel Grunert ^{1,2,3}, Sandra Appelt ^{1,2,3}, Paul Grossfeld ⁴ and Silke R. Sperling ^{1,2,3,5,*} 

¹ Cardiovascular Genetics, Charité—Universitätsmedizin Berlin, 13125 Berlin, Germany; marcel.grunert@charite.de (M.G.); sandra.appelt@charite.de (S.A.)

² Berlin Institute of Health (BIH), 10178 Berlin, Germany

³ DZHK (German Centre for Cardiovascular Research), Partner Site Berlin, 10178 Berlin, Germany

⁴ Division of Cardiology, University of California San Diego, Rady's Hospital MC 5004, San Diego, CA 92123, USA; pgrossfeld@health.ucsd.edu

⁵ Department of Biology, Chemistry, and Pharmacy, Freie Universität Berlin, 14195 Berlin, Germany

* Correspondence: silke.sperling@charite.de; Tel.: +49-30450540123

Received: 3 November 2020; Accepted: 30 November 2020; Published: 2 December 2020



Abstract: Congenital heart defects (CHDs) are the most common birth defect in human with an incidence of almost 1% of all live births. Most cases have a multifactorial origin with both genetics and the environment playing a role in its development and progression. Adding an epigenetic component to this aspect is exemplified by monozygotic twins which share the same genetic background but have a different disease status. As a result, the interplay between the genetic, epigenetic and the environmental conditions might contribute to the etiology and phenotype. To date, the underlying causes of the majority of CHDs remain poorly understood. In this study, we performed genome-wide high-throughput sequencing to examine the genetic, structural genomic and epigenetic differences of two identical twin pairs discordant for Tetralogy of Fallot (TOF), representing the most common cyanotic form of CHDs. Our results show the almost identical genetic and structural genomic identity of the twins. In contrast, several epigenetic alterations could be observed given by DNA methylation changes in regulatory regions of known cardiac-relevant genes. Overall, this study provides first insights into the impact of genetic and especially epigenetic factors underlying monozygotic twins discordant for CHD like TOF.

Keywords: congenital heart disease; Tetralogy of Fallot; monozygotic twins; discordant phenotype; genetics; variations; epigenetics; DNA methylation; candidate genes

1. Introduction

The heart is the first functional organ during embryogenesis, and congenital heart disease (CHD) represents the most common birth defect in humans, affecting about 1% of all newborns [1]. Despite improvements in recognition and therapeutic opportunities based on anatomical, physiological and surgical considerations, CHD remains a leading cause of infant and child mortality. The majority of CHDs are most likely caused by the interplay of multiple subtle genetic, structural genomic or epigenetic alterations [2,3]. In addition, most of them are modulated by gen–environment interactions [2], of which epigenetic alterations are proposed to represent an important transmitter. The symptoms of CHD may vary from none to severe, which attests to its complex nature. This spread is well illustrated in Tetralogy of Fallot (TOF), the most common cyanotic form of CHDs. TOF is characterized by a ventricular septal defect (VSD) with an overriding aorta as well as an obstruction of the right ventricular outflow tract

(pulmonary stenosis) and right ventricular hypertrophy. These four anatomical features can all exhibit variable levels of severity.

Over recent decades and in particular with the advent of high-throughput DNA sequencing, several CHD-associated sequence variations have been identified by genetic studies of affected individuals and families [3–8]. Using a homogeneous TOF cohort with well-defined features, we recently identified a multigenic background of rare deleterious mutations in several genes, which discriminate TOF cases from controls and which are essential for apoptosis and cell growth, the assembly of the sarcomere as well as for the neural crest and secondary heart field [9]. One central mechanism of epigenetic control is DNA methylation. Thus, we also performed a genome-wide DNA methylation study on myocardial biopsies of TOF and VSD patients [10]. We found clear methylation differences between patients and controls and moreover, between patient groups. However, the underlying specific causes for the majority of CHDs including complex forms such as TOF remain poorly understood.

Monozygotic twins are also known as identical twins because they share nearly 100% of their genetic information stored in DNA. However, they are often discordant for complex diseases such as diabetes, neurodevelopmental disorders, cancer as well as CHD [11–14]. Considering that monozygotic twins are genetically identical, epigenetic mechanisms like DNA methylation might be modulators of the phenotypic discordance by mediating between the environment and phenotypic expression [13]. For CHD for example, Lyu et al. showed, in the case of an identical twin pair discordant for double outlet right ventricle, a high correlation between hypermethylated promoters based on reduced representation bisulfite sequencing and down-regulated gene expression levels in the patient compared to the healthy twin [14]. In this study, we will investigate for the first-time genome-wide genetic and epigenetic differences between two monozygotic twin pairs discordant for TOF, providing a deeper understanding of the interplay between genetic, structural genomic and epigenetic alterations involved in the etiology of this complex disease.

2. Materials and Methods

2.1. Study Participants and Ethics Statement

To obtain genomic DNA (gDNA) for whole genome sequencing (WGS) and whole genome bisulfite sequencing (WGBS), blood samples were taken from four individuals (two identical twins discordant for TOF, i.e., one affected case each (i.e., Twin1_TOF and Twin2_TOF) and one healthy monozygotic sibling each (i.e., Twin1_H and Twin2_H)). The first pair of twins are male individuals (i.e., Twin1) and the second pair are female individuals (i.e., Twin2) (Figure 1). Since a healthy co-twin in a discordant twin pair serves as a well-matched control with the same genetic background, no further controls were included in this study. The local institutional review board of University of California San Diego approved the study (UCSD IRB protocol #111523) and informed consent was obtained from all participants or guardians. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki.

2.2. Whole Genome Sequencing (WGS)

Genomic DNA from blood of TOF patients and healthy siblings was extracted using Puregene DNA purification kit (Gentra). DNA degradation and contamination were monitored on 1% agarose gels. Moreover, DNA concentration was measured using Qubit DNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). Library preparation and sequencing (2×150 bp paired-end Illumina sequencing; NextSeq PE150 with Q30 $\geq 80\%$) was performed by Novogene. Briefly, a total amount of 1 μ g DNA per sample was used as input material for the DNA sample preparations. Sequencing libraries were generated using NEBNext DNA Library Prep Kit following manufacturer's recommendations and indices were added to each sample. The genomic DNA was randomly fragmented to a size of 350 bp by shearing, then DNA fragments were end polished, A-tailed, and ligated with the NEBNext adapter for Illumina sequencing, and further PCR enriched by P5 and indexed P7 oligos. The PCR

products were purified (AMPure XP system) and resulting libraries were analyzed for size distribution by Agilent 2100 Bioanalyzer and quantified using real-time PCR.

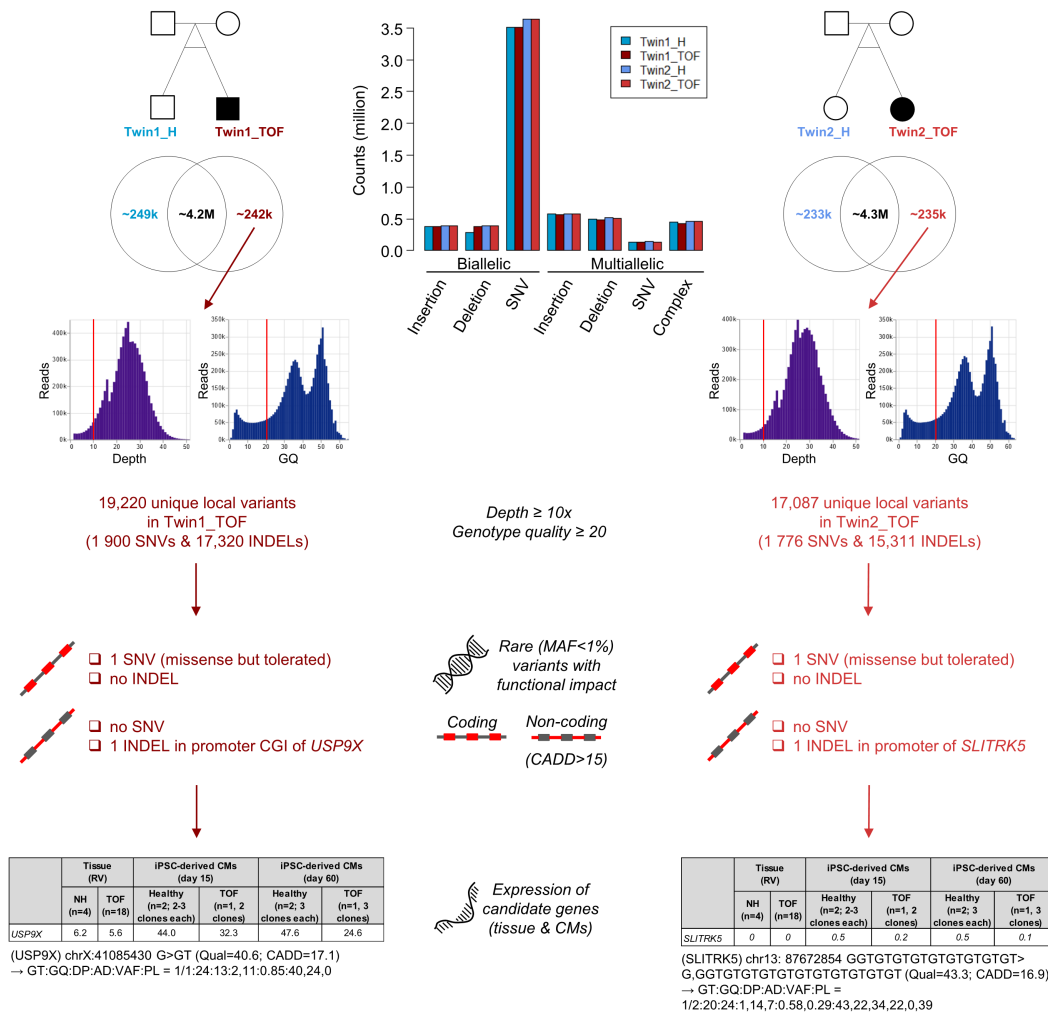


Figure 1. Filtering of local variations and candidate genes identified in affected Tetralogy of Fallot (TOF) twins using whole genome sequencing. Variant positions are based human reference genome (GRCh38.p13/hg38). AD, allelic depth; CADD, combined annotation dependent depletion; CMs, cardiomyocytes; DP, read depth; GT, genotype; GQ, genotype quality; H, healthy; INDEL, insertion and deletion; MAF, minor allele frequency; NH, normal heart; PL = Phred-scaled genotype likelihoods; Qual, base quality (Phred score); RV, right ventricle; SNV, single nucleotide variations; TOF, Tetralogy of Fallot; VAF, variant allele frequency.

Each sample was sequenced on four sequencing lanes (Table S1). On average, sequencing resulted in approximately 332 million read pairs per sample (Table S1). The quality of the sequencing samples was checked using FASTQC v1.7 [15] and Qualimap v2.2.1 [16]. All samples passed sequence quality. On average, a duplication rate of 10.4%, a GC content of 40.4% and mean base quality of 37.9 (Phred score) were observed over all samples (Table S1 and Figure S1A). After initial quality check, the reads were mapped to the human reference genome (GRCh38.p13/hg38) using Bowtie2 v2.4.1 with the “very-sensitive” parameter setting, which is more sensitive and more accurate [17]. On average, 98.8% of the reads per sample could be mapped. The average sequencing read depth over the human reference genome is 30.1x with a coverage of 99.5% (Table S2 and Figure S1B). Duplicate reads were marked using Picard tools v1.140 (<http://broadinstitute.github.io/picard>, v1.140).

Calling of local variations (single nucleotide variations (SNVs) as well as insertions and deletions; INDELS) was performed using DeepVariant v0.10.0 with default parameters [18]. Variants that passed DeepVariant default quality filters and with a genotype quality greater than or equal to 20 and a read depth greater than or equal to 10 were annotated using Ensembl Variant Effect Predictor (VEP; release 100) [19]. Local variations with a minor allele frequency greater or equal than 0.01 (1%) in the 1000 Genomes or Genome Aggregation Database (gnomAD) global populations [20] as well as variations with low functional impact based on Ensembl calculated variant consequences were filtered out [19]. Non-synonymous variants must be predicted to be damaging by either PolyPhen-2 [21], SIFT [22] or MutationTaster2 [23]. Moreover, variants in the non-coding region were filtered for those with a CADD (Combined Annotation Dependent Depletion) score greater than 15 (Phred-scaled) [24], a cutoff on deleteriousness. Structural variations (SVs) were called using Manta v1.6.0 with default parameters [25]. Moreover, common SVs based on gnomAD-SV were filtered out [20]. Copy number variations (CNVs) were called using Control-FREEC v11.5 with default parameters [26]. Both SVs and CNVs were further overlapped with genomic features obtained from VEP.

2.3. Whole Genome Bisulfite Sequencing (WGBS)

For WGBS, gDNA obtained from blood of the two twin pairs (in total four individuals) was extracted using Puregene DNA purification kit (Gentra) and further used to create bisulfite-treated DNA libraries. Library preparation and sequencing was performed by Novogene using EZ DNA Methylation Gold Kit from Zymo Research and 150 bp paired-end Illumina-Kit. Briefly, unmethylated λ 243 Sam7 Lambda DNA (48,502 bp, Promega D1521) was combined with gDNA to act as an internal control to monitor the bisulfite conversion rate. Afterwards, DNA was sheared into 200–400 bp fragments, sequence ends were repaired and 3' ends were adenylated. Next, methylation sequencing adapters were added and DNA was treated with bisulfite. Treatment of DNA with sodium bisulfite deaminates unmethylated cytosines to uracil while methylated cytosines are resistant to this conversion, allowing therefore for the discrimination between methylated and unmethylated CpG sites. Finally, bisulfite treated library was subjected to PCR amplification followed by standard DNA paired-end Illumina sequencing (HiSeq PE150 with Q30 \geq 80%). In general, sodium bisulfite pre-treatment of DNA coupled with high-throughput sequencing allows us to study DNA methylation quantitatively and genome-wide at single cytosine site resolution.

Each sample was sequenced on six sequencing lanes (Table S3). On average, sequencing resulted in approximately 344 million read pairs per sample (Table S3). The quality of the sequencing samples was checked using FASTQC v1.7 and Qualimap v2.2.1. All samples passed sequence quality. On average, a duplication rate of 10%, a GC content of 21.1% and mean base quality of 36.95 (Phred score) were observed over all samples (Table S3 and Figure S2). The bisulfite conversion was efficient based on the base composition (<1–2% of C in the forward strand and G in the reverse strand) (Figure S3A) and based on the bisulfite conversion rate (>99.7%) (Figure S3B). The bisulfite conversion efficiency was calculated for each sample based on unmethylated control sequences (spike-ins) added to the library prior to fragmentation (i.e., 100—percentage of C methylated in CpG context when mapping against the control genome *Enterobacteria phage lambda*). After initial quality assessment, adapter trimming was performed by Trim Galore v0.4.4 [27]. Afterwards, sequencing reads were mapped to the human reference genome (GRCh38.p3) using Bismark v0.18.2 [28]. On average, 75% of the ~344 million reads per sample could be mapped (Table S4). After mapping, reads of all sequencing lanes were merged for each sample using SAMtools v1.2 [29] and subjected to deduplication using Bismark. Out of ~258 million mapped reads per sample, ~229 million left after deduplication (Tables S4 and S5). All mapped reads have a high mean base quality of ~35.7 (Phred score) with a GC content of ~21.3% (Figure S4A and Table S5). The average sequencing depth is 21.1 \times per base (Figure S4B and Table S5). To correct for methylation bias at 3' and 5' end of each read, MethylDackel v0.3.0 was used [30]. For forward reads, no cutoffs were suggested; however, there are clear drops in the methylation level at the beginning (15–18 bp) and end (149 bp) of the reads for which cutoffs were suggested by MethylDackel (Figure S5A)

and incorporated during methylation level extraction using Bismark. As expected, the samples have, on average, a CpG methylation rate of 82.5%. CHG and CHH methylation rates are very low at under 1% each (Figure S5B and Table S5). Differential methylation analyses of CpG sites between two samples are performed using methylSig v0.4.4 [31]. As methylSig only allows comparisons with at least two samples per group, each sample was duplicated. Furthermore, CpGs were filtered by coverage with a minimum coverage of 10 and a maximum of 500 (default values). Differential methylation is defined by a methylation difference of at least 25%. Differential methylated CpGs are further overlapped (at least 1 bp) or associated (i.e., nearest gene approach) with hg38 annotated genomic features as described in detail elsewhere [10]. Briefly, the overlap includes 50,497 promoters of UCSC RefSeq genes (−500 bp/+2 kb to TSS); 30,477 CpG islands (CGIs) based on UCSC track “cpgIslandsExt”; promoter CGIs (i.e., overlap of promoter and CGIs with at least 1 bp), CGI shores (i.e., regions outside CpG islands but within 2 kb of any CpG island), transcription factor binding sites (TFBS) predicted by the Transfac Matrix Database and conserved in the multiple alignment of human, mouse and rat, including cardiac transcription factors (TFs); and cardiac enhancers (p300 ChIP-seq data of human adult and fetal hearts [32]).

2.4. Filtering for Disease-Relevant Genetic and Epigenetic Alterations

Candidate genes with genetic, structural genomic or epigenetic alterations were overlapped with various datasets to filter for known or possible disease-relevant genes. This includes heart- and muscle relevant genes (865 genes based on several resources) [9]; cardiovascular-associated genes (list of 4275 genes created by the Cardiovascular Gene Annotation Initiative in collaboration with EMBL-EBI); genes which are differentially expressed in CHD patients (in particular TOF, VSD, atrial septal defect (ASD), hypoplastic left heart syndrome (HLHS), transposition of the great arteries combined with pulmonary artery (TGA-PS)) [9,33–37]; genes which are differentially methylated in the promoter, gene body or related enhancer of CHD patients (in particular TOF, VSD, LS-CHD) [10,38–41]; genes which overlap CNVs associated with CHD (including TOF, LS-CHD, HLHS and conotruncal defects) [42–56]; known human CHD genes [3,6,54]; and genes which are differentially expressed and targeted by differentially expressed microRNAs in CHD patients (TOF and HLHS) [57–59]. Moreover, candidate genes must be expressed (RPKM (reads per kilo base per million mapped reads) or TPM (transcript per million) value > 1) in the right ventricular tissue of normal or TOF hearts [9], or during cardiac differentiation (i.e., in cardiomyocytes of day 15 and/or day 60 derived from induced pluripotent stem cells of healthy, unaffected individuals and TOF patients) [60].

2.5. Statistics

General bioinformatics and statistical analyses were conducted using R (including Bioconductor packages) and Perl.

3. Results

3.1. Genomic Variations in Affected Twins

Whole genome sequencing of monozygotic twins resulted in approximately 4.4–4.6 million local variations (i.e., SNVs and INDELs) for each twin, with SNVs constituting the largest class of called genomic variations. Out of these raw variants, up to 6% are unique for either twin. Considering that monozygotic twins are genetically identical, this number seems very high; however, there are differences in terms of sequencing depth and quality, which may result in different calling results. Using a minimum sequencing depth of 10× and a genotype quality of at least 20, the number of local variants drops in Twin1_TOF from 4,167,698 to 19,220 (0.46%) and in Twin2_TOF from 4,331,595 to 17,087 (0.39%) unique ones compared to the respective unaffected twin (Figure 1). To identify possible disease-associated alterations, these unique variants were further annotated and filtered for rare variants (MAF < 1%) with functional impact on the coding or non-coding genomic sequence. For the coding part, this resulted in

no single INDEL and only one missense SNV for each affected twin, which are further predicted to be tolerated (Figure 1). For the non-coding sequence of Twin1_TOF, there is a homozygous INDEL (i.e., insertion) in the promoter region of *Ubiquitin Specific Peptidase 9 X-Linked (USP9X)*. The promoter also represents a CpG island (CGI). In Twin2_TOF, we also identified no SNV but one INDEL in the promoter of *SLIT and NTRK Like Family Member 5 (SLITRK5)*. Both non-coding variants are potentially pathogenic with a CADD score greater than 15. The related genes were further checked for being expressed in the human heart. Moreover, considering that TOF is a developmental disorder, these genes should be expressed during development. Thus, we examined their expression in the right ventricular tissue of normal and TOF hearts [9] as well as during cardiac differentiation using cardiomyocytes derived from induced pluripotent stem cells of healthy, unaffected individuals and TOF patients [60]. Out of these two genes, *SLITRK5* is not expressed in the heart or during cardiomyocyte differentiation (Figure 1). We further overlapped the two genes with known cardiovascular-associated genes and several CHD-related datasets, but without any overlap for both genes. In addition to unique variants in the two affected TOF twins (i.e., variants that do not occur in the healthy sibling), we also searched for variations with zygosity differences between the healthy and affected siblings. However, after filtering there is no SNV or INDEL in the coding and non-coding sequence of both twin pairs, which can explain the phenotype differences (Figure S6).

3.2. Structural Genomic Variations in Affected Twins

Besides local variations, we are interested in copy number variations as well as structural variations, which are unique for the affected TOF twins. For Twin1_TOF, we identified 11 CNVs, which are associated with 9 genes (Table S6). The majority of these CNVs are copy number gains (i.e., 9 out of 11) and only two represent a copy number loss. For Twin2_TOF, we found four CNVs (two gains and two losses), of which only one copy number gain affects a genomic feature, namely the rRNA *RNA5SP19*. The very few genomic features affected by CNVs in Twin1_TOF are either protein-coding or T cell receptors, where the latter represent an adaptive immune response (Table S6). The protein-coding genes were checked for cardiac expression and further overlapped with a list of cardiovascular-associated genes and other CHD-related datasets. The latter revealed only for *UDP Glucuronosyltransferase Family 2 Member B17 (UGT2B17)* an overlap with cardiovascular-associated genes and known CNVs in CHD patients; however, the gene is not expressed in the normal or TOF heart or during cardiomyocyte differentiation. The only expressed gene associated with a CNV in twin1_TOF is *Fc Fragment of IgA Receptor (FCAR)*, but without any overlap to other cardiac-related annotations or datasets. In summary, there is no real candidate gene with a CNV in both affected twins.

To find possible disease-relevant SVs, we filtered for unique ones in the affected twins, which are rare and have a high or modifier functional impact. In total, we found 30 rare SVs in Twin1_TOF and 20 rare SVs in Twin2_TOF with modifier impact only, affecting 20 and 13 genes, respectively (Tables S7 and S8 and Figure S7). Both affected twins share five SVs. Overall, 83% of the genes associated with SVs are expressed in the normal or TOF heart or during cardiac differentiation. Moreover, the majority of these expressed genes overlap with several cardiac related genes or datasets. For example, *ANTXR Cell Adhesion Molecule 1 (ANTXR1)* and *L3MBTL Histone Methyl-Lysine Binding Protein 4 (L3MBTL4)* with SVs in Twin1_TOF are also differentially expressed (significantly up-regulated) in right ventricle of TOF patients compared to normal hearts [9]. Moreover, both are differentially methylated in VSD patients compared to normal hearts (i.e., significantly hypermethylated cardiac enhancer near *ANTXR1* and hypermethylated gene body of *L3MBTL4*) [10]. The genes with SVs in Twin2_TOF, encoding for *Bardet-Biedl Syndrome 2 (BBS2)* and *Teneurin Transmembrane Protein 4 (TENM4)*, are also associated with differential DNA methylation in VSD cases versus normal hearts (i.e., significantly hypermethylated gene body of *BBS2* and hypermethylated promoter of *TENM4*) [10].

3.3. DNA Methylation Differences between Discordant Twins

To investigate the impact of DNA methylation changes between discordant twins, we performed WGBS and studied the methylation level of approximately 28 million CpG sites in the human genome. In general, the global CpG methylation shows no obvious differences between each affected and healthy twin (Figures S8 and S9A), with a somewhat higher coverage in Twin1_TOF compared to Twin1_H and a similar one for Twin2_TOF and Twin2_H (Figure S9B). Moreover, genomic features associated to genes have a higher methylation level (Figure 2A).

Next, we searched for differentially methylated CpG sites (DMCs) with at least 25% differences between the discordant twins. We found 299,643 DMCs in Twin1 and ~150,000 more DMCs (in total 457,108) in Twin2 (Table 1). For Twin1, there are more hypomethylated CpGs, while Twin2 is more balanced regarding hypo- and hyper-methylation (Figure 2B,C, left each). However, the overlap of DMCs with genomic features is similar for both twin pairs, with the majority of DMCs in non-coding regions (intergenic and intronic; Figure 2B,C, right each). Nevertheless, there are DMCs in coding and even more interestingly in regulatory regions such as promoters, CGIs, promoter CGIs, CGI shores, TFBS (cardiac and/or located in promoters) and cardiac enhancers (Table 1).

Table 1. Differentially methylated CpG sites with $\geq 25\%$ methylation difference between the healthy and affected twin and their overlap (≥ 1 bp) or association (i.e., nearest gene approach) with different genomic features. CGIs, CpG islands; DMCs, differentially methylated CpGs; H, healthy; TFBS, transcription factor binding site; TOF, Tetralogy of Fallot.

	Twin1_TOF vs. Twin1_H	Twin2_TOF vs. Twin2_H
<i>Differentially methylated CpGs (DMCs) with $\geq 25\%$ methylation difference</i>	299,643	457,108
DMCs in promoters	18,052	22,537
DMCs in CpG islands (CGIs)	397	713
DMCs in promoter CGIs	131	429
DMCs in CGI shores	17,337	21,455
DMCs at TFBS	23,716	31,789
DMCs at cardiac TFBS	2066	2751
DMCs at TFBS in promoters	2125	2929
DMCs at cardiac TFBS in promoters	215	264
DMCs in cardiac enhancers (p300)	2042	2518

Of particular interest is the methylation pattern of dense regions of CpGs, the CGIs. For Twin1, we found 397 DMCs in CGIs and 131 of these are located in promoters. For Twin2, there are 713 DMCs in CGIs and the majority (in total 429) overlap promoter regions. Next, we overlapped (promoter) CGI-associated genes with cardiovascular-associated genes including known CHD genes and found several cardiac-related genes (Figure 3A,B). In total, there are 4 and 20 genes with DMCs in promoter CGIs for Twin1 and Twin2, respectively. Interestingly, 3 out of 4 genes in Twin1 also harbor DMCs in Twin2. These genes are *BARX Homeobox 2 (BARX2)*, *Kinesin Family Member C3 (KIFC3)* and *Nuclear Factor of Activated T Cells 1 (NFATC1)*. The latter is a transcription factor required for valve formation [61] and thus, a known CHD gene. In addition to *NFACT1*, there is another transcription factor and well-known CHD gene in Twin2 with DMCs in its promoter CGI, namely *T-Box Transcription Factor 20 (TBX20)* [62,63].

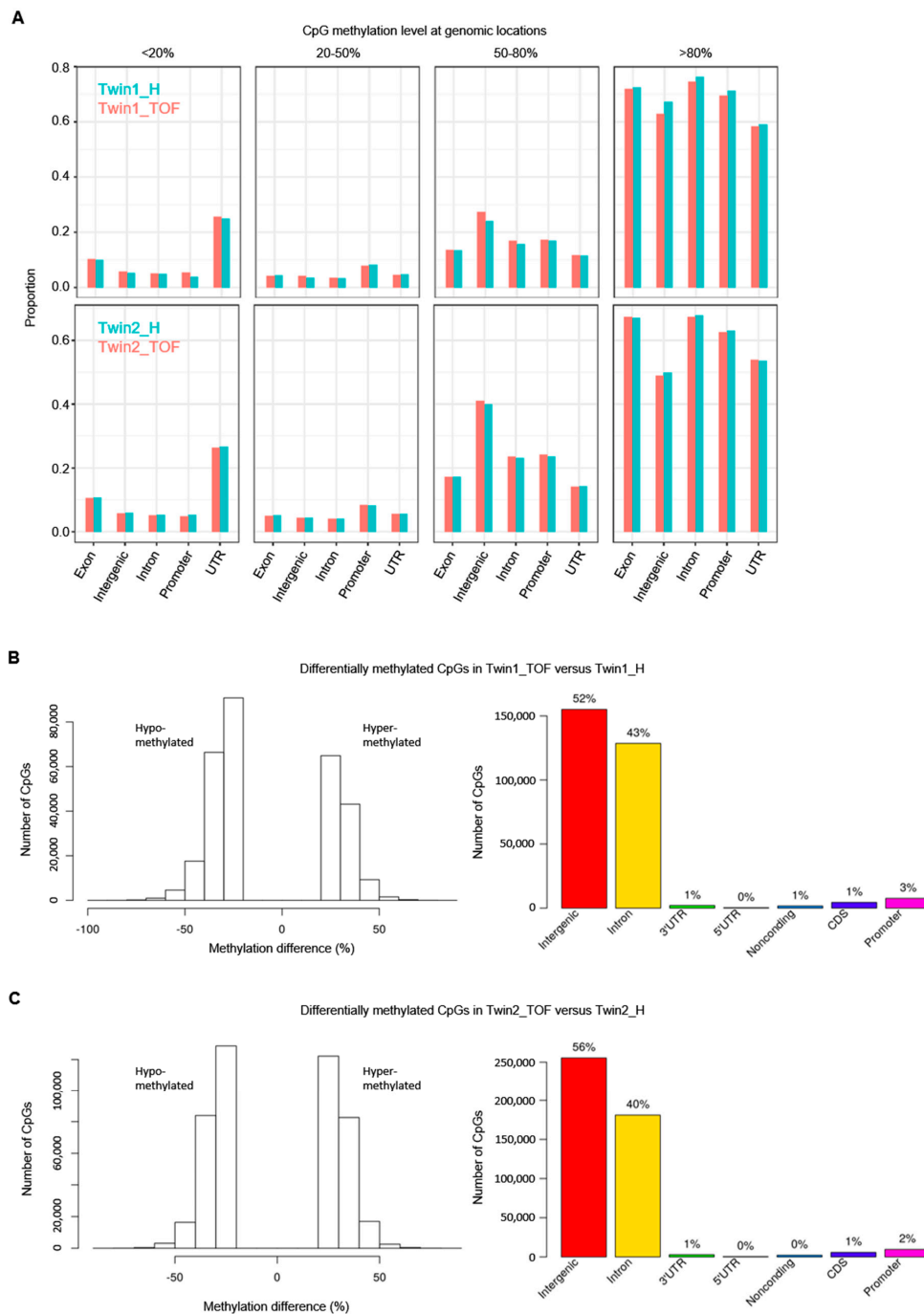


Figure 2. DNA methylation levels and differences of CpG sites in twins and their genomic locations. (A) Overview of methylation levels of CpGs at genomic locations. Proportions of CpG are given over the total count for each feature (no pairwise comparison of single CpG sites). (B) Differentially methylated (hypo- and hyper-methylated) CpG sites in Twin1_TOF compared to Twin1_H and their genomic locations. (C) Differentially methylated (hypo- and hyper-methylated) CpG sites in Twin2_TOF compared to Twin2_H and their genomic locations. Differentially methylated CpGs with at least 25% methylation difference between affected (*_TOF) and healthy (*_H) twin.

Besides the regulatory impact of CpG Islands, we also examined genes that have DMCs at TBFS in their promoter. In total, 27 and 45 genes in Twin1 and Twin2, respectively, harbor such DMCs and overlap with cardiovascular-associated genes (Figure 3C,D). Four of these genes were found between both twin pairs (i.e., *CPNE1*, *NEDD4L*, *PPFIA2*, *RBM12*). Moreover, there are four different known

CHD genes with DMCs at TBFS in their promoter in Twin1 (i.e., *DNAH5*, *FOXP1*, *LAMA4*, *NIPBL*) and Twin2 (i.e., *CACNA1D*, *CBL*, *DMD*, *NSD1*).

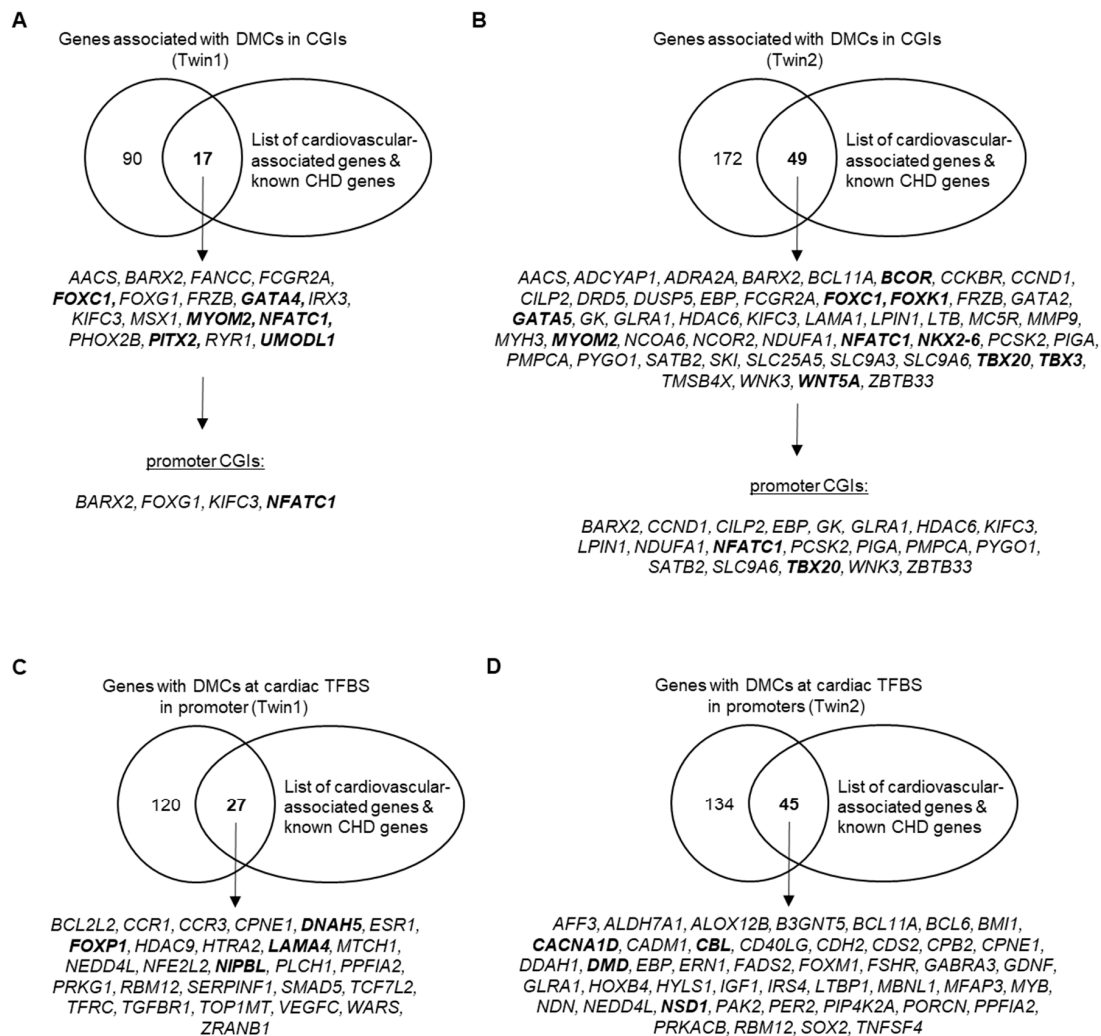


Figure 3. Overlap of protein-coding genes associated with differentially methylated CpGs in CGIs and promoter CGIs or at cardiac TFBS in promoters with cardiovascular-associated genes and known CHD genes. Known CHD genes are marked in bold. (A) Genes associated with DMCs of Twin1_TOF versus Twin1_H (i.e., Twin1) in CGIs or promoter CGIs. (B) Genes associated with DMCs of Twin2_TOF versus Twin2_H (i.e., Twin2) in CGIs or promoter CGIs. (C) Genes with DMCs of Twin1 at cardiac TFBS in their promoter. (D) Genes with DMCs of Twin2 at cardiac TFBS in their promoter. CGIs; CpG islands; CHD, congenital heart disease; DMCs, differentially methylated CpG sites; TFBS, transcription factor binding sites.

Figure 4 shows the difference in the methylation level of both twins at the promoter of four selected known CHD genes, including TFBS in these regions. In addition, the mean methylation level obtained from methyl-CpG binding domain (MBD) protein-enriched genome sequencing (MBD-seq) of tissues from normal hearts as well as from a homogenous group of TOF and VSD patients is provided [10]. Figure 4 above shows the methylation level at the promoter CGI of *NFATC1* and *TBX20*. The methylation level at the promoter of two essential cardiac transcription factors, *NKX2-5* and *GATA4*, which both harbor single DMCs in their promoter region with more than 25% methylation difference between healthy and affected twin, is shown in the bottom of Figure 4. In general, the observations in WGBS partially differ with the methylation pattern observed in MBD-seq, which may be due to the resolution (1 bp in WGBS vs. ~150 bp in MBD-seq) and coverage (i.e., single CpGs in WGBS $\geq 5\times$)

or simply based on the homogeneity of the affected twin compared to other TOF and VSD patients. However, for both WGBS and MBD-seq one can observe methylation alterations in the promoter of these selected genes.

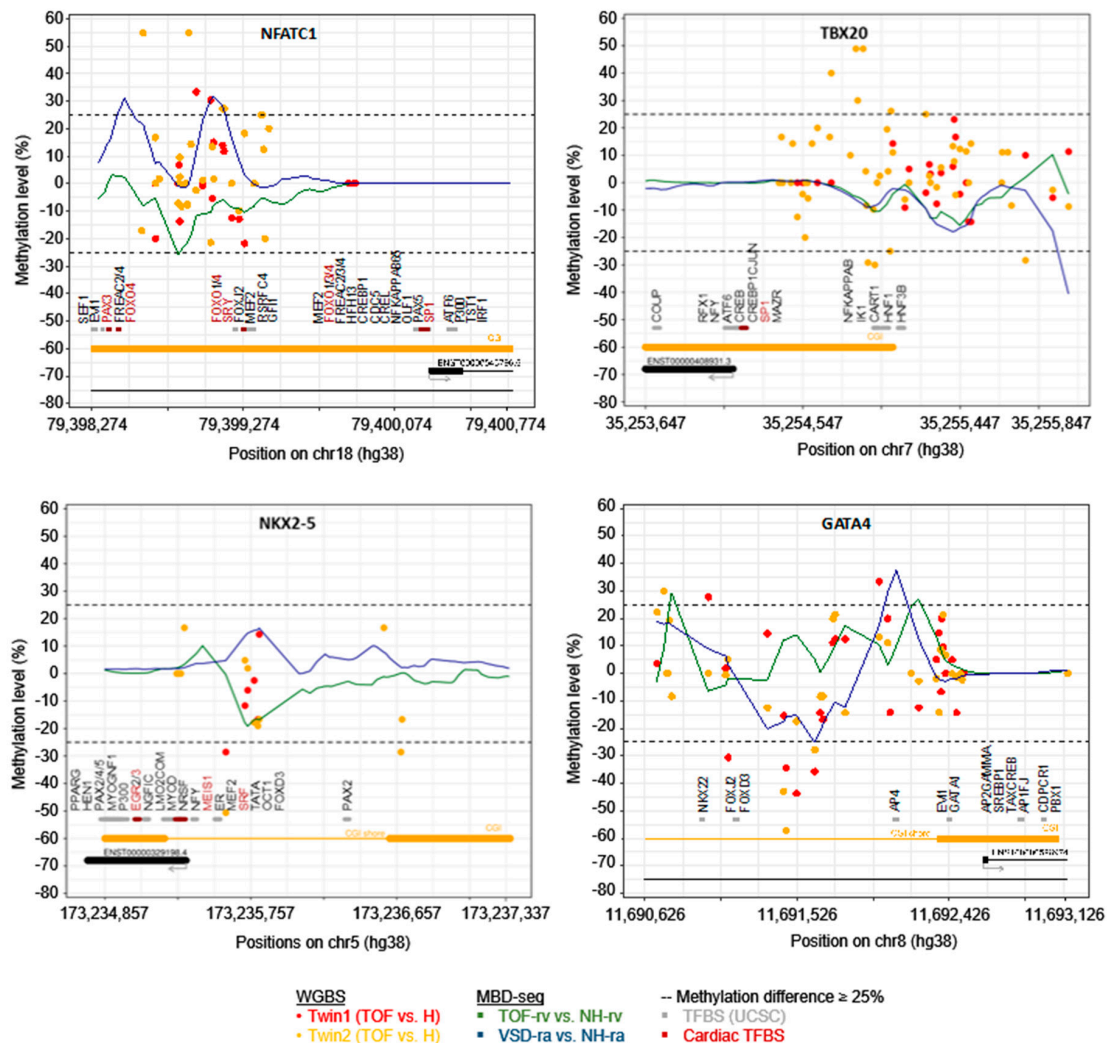


Figure 4. DNA methylation level at promoter regions of *NFATC1*, *TBX20*, *NKX2-5* and *GATA4* in healthy and affected twins obtained by WGBS ($n = 1$ each; coverage at CpGs $\geq 5\times$) as well as in right ventricular or atrial tissue of normal hearts and patients with TOF or VSD obtained by methyl-CpG binding domain protein-enriched genome sequencing (MBD-seq) ($n = 7$ for TOF-rv; $n = 8$ for VSD-ra; $n = 3$ for NH-rv; $n = 4$ for NH-ra). Promoters are defined as -2 kb/ $+500$ bp of transcriptional start site. CGIs and CGI shores are indicated by orange line at the bottom. Moreover, binding sites of transcription factors (obtained from UCSC) are provided in gray or for cardiac factors in red. For WGBS, methylation at single CpG is provided. CGIs, CpG islands; NH, normal heart; H, healthy; RA, right atrium; RV, right ventricle; TFBS, transcription factor binding site; TOF, Tetralogy of Fallot; VSD, ventricular septal defect; WGBS, whole genome bisulfite sequencing.

4. Discussion

Out of millions of unique local variations (i.e., SNVs and INDELs) for each affected twin obtained from WGS, we identified no single candidate gene in one of the two monozygotic discordant twin pairs, which might explain the difference regarding TOF (Figure 1). Searching for candidate genes affected by structural genomic variations, we observed in both TOF twins several SV candidates but no relevant CNVs. For SVs, there are several candidates in both affected twins who are also expressed in the adult heart and during cardiac differentiation (Table S9). Moreover, most of these expressed genes

also overlap with cardiovascular-associated genes and CHD-related datasets. Most interestingly for Twin1_TOF, the affected genes *ANTXR* and *L3MBTL* are also found to be differentially expressed in right ventricle of TOF versus normal hearts [9] and in addition, differentially methylated in right atrium of VSD versus normal hearts [10]. *L3MBTL* is involved in transcriptional repression [64]. Its MBT domain binds to methylated histone residues, which is linked to the formation and maintenance of heterochromatin [65]. For Twin2_TOF, the affected genes *BBS2* and *TENM4* are also differentially methylated in VSD versus normal hearts [10]. However, none of the affected genes is known to be involved in TOF or CHD in general so far. Moreover, since we do not have any tissue available from the twins (only blood samples), we unfortunately cannot make any statement regarding the regulatory impact of these SVs on the expression of related genes. Due to ethical restrictions, it is further practically impossible to obtain such tissue from the healthy, unaffected twin; however, follow-up studies based on induced pluripotent stem cells and derived cardiomyocytes might be an option to verify the evidence for their causative impact.

As mentioned above, some of the genes with SVs in the TOF twins are already known to be associated with differential DNA methylation in cardiac tissue of patients with TOF and in its sub-feature, VSD [10]. To study the epigenetic differences based on DNA methylation between the monozygotic discordant twin pairs, we used blood sample for WGBS as for WES analysis of genetic and structural genomic alteration. CpG methylation can persist in steady-state from early embryonic development throughout a lifetime, and only a fraction of 22% of autosomal CpGs shows dynamic methylation in the normal developmental context [66,67]. With respect to the study of altered DNA methylation pattern, this enables the analysis of up to 80% of CpG methylations in cells independent from the affected cell type. Overall, we observed no differences based on the global CpG methylation between each of the two twin pairs. However, we found numerous single CpGs differentially methylated with more than 25% difference between the twins (Table 1). Most interesting are those DMCs in promoters that also overlap with CGIs or TFBS (Figure 3). Usually, CGIs are largely unmethylated [68] and frequently located in promoters [69]. Multiple methylated CpGs in promoter CGIs cause stable silencing of genes and thus, hypermethylated promoter CGIs are often known to be disease-related [70]. For both affected twins, we found DMCs in the promoter CGI of a gene encoding for the TF NFATC1 (top left of Figure 4), which is required for normal valve formation [61]. Moreover, genetic and structural genomic alterations in *NFATC1* have already been identified in patients with CHD such as VSD or tricuspid atresia [71,72]. For Twin2_TOF, we identified several DMCs in the promoter CGI of *TBX20* (top right of Figure 4), a cardiac TF [63]. Mutations in *TBX20* have been identified in different types of cardiac defects regarding septation, chamber growth, and valvulogenesis [62]. Tbx20 interacts with the TF Gata4 to activate both *Mef2c* and *Nkx2-5* enhancers [73], which all are core cardiac TFs [63]. These factors regulate each other's expression, partly with combinatorial impact of downstream targets. Disturbances of their associated cardiac regulatory network have been associated with various congenital heart malformations [63]. For both cardiac TFs GATA4 and *NKX2-5*, we found DMCs in the promoter region of both affected twins, who further represent CGI shores (bottom of Figure 4). These regions are within 2 kb of a CGI and they are known to occur most of the tissue-specific methylation [74].

There are many other candidate genes associated or overlapping DMCs with different regulatory features (Table 1); however, it clearly shows that the difference between the two monozygotic discordant twin pairs is rather observed in epigenetic alterations than on the genetic or structural genomic level. Nevertheless, the impact of these epigenetic variations on candidate gene expression and thus, their causality for the phenotypic discordance needs to be verified by further studies, in functional respect but also in terms of cases and comparisons (i.e., this study comprises just two cases with each one affected vs. one healthy twin comparison). However, our study confirms other studies on identical twins that discordances in a disease such as CHD cannot be explained by genetic or structural genomic differences [14,75–79]. Moreover, blood-derived DNA might be chimeric between the identical twins and as shared blood circulations have been found during embryogenesis in most monozygotic

twin pregnancies [80], hematopoietic stem cells can be transferred between twins. Such a created hematopoietic system might mask the underlying genetic, structural genomic and also epigenetic differences between the monozygotic twin discordant for a disease like TOF [75].

Recent large-scale studies with up to thousands of CHD cases and of controls only show possible genetic or structural genomic causes for a minority of cases [4,7,44,53,81]. The reasons for this might be found in the multifactoriality paired with complex inheritance patterns and low prevalence [82]. Moreover, the majority of cases occur sporadic, possibly triggered by stochastic or environmental events during maternal pregnancy [83,84]. The latter is accompanied by epigenetic changes such as those related to histone modifications and DNA methylation, important modulators of gene–environment interactions. Moreover, deregulated microRNA abundance levels also seem to play a role in the development of CHD in twins [85]. Stochastic events such as an unequal division of the inner mass cells during twinning or an unequal allocation of developmental markers and precursor cells might be responsible for a discordance in monozygotic twins [86], which might be of interest with respect to a possible premature stop of cellular growth in TOF. Moreover, this may result, for example, in altered sub-populations of ventricular and atrial cardiomyocytes [87], which in turn might be disease-causing. In general, monozygotic twins are up to seven times more likely to develop a congenital heart malformation, which further usually occurs in just one of the twins [86,88]. Reasons for this increased risk include the twin–twin competition for maternal resources resulting in, for example, unequal placental sharing, unequal sharing of the vascular system or other placental mechanisms like diffusion and endocrine function, which increases the probability of a skewed in utero environment affecting the twins [86,89–91]. Interestingly, a study based on the Danish Twin Registry showed (unlike several other studies) that there is no higher risk of CHD in monozygotic twins compared to dizygotic ones [92]. This indicates that the intrauterine environment of twin gestation is predisposed to the development of CHD in a certain way. Moreover, it is suggested that twin gestation changes maternal nutrition to the fetus and nutritional deficiencies might increase the sensitivity of the fetus to other teratogenic factors such as prenatal exposure or poor maternal health [92].

The search for the underlying mechanisms of CHD in twins is like searching for the needle in the haystack. The four pieces of the TOF pathology can all exhibit variable levels of severity and as a result, no two TOF cases are the same. However, studying the genetic and epigenetic mechanisms in identical discordant twins can provide important insights into complex diseases such as congenital heart malformations. Therefore, we studied here the genetic, structural genomic and epigenetic differences of two identical twin pairs discordant for TOF. The two cases show common and unique patterns of epigenetic modifications, which might also be related to a common and unique level of pathological features. Understanding the mechanisms that trigger the molecular plasticity, namely the underlying genetic composition, the epigenetic make-up and environmental insults, is quite challenging but will enhance our current knowledge of CHD.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2308-3425/7/4/55/s1>, Figure S1: Statistics of reads obtained from whole genome sequencing. Figure S2: Statistics of input reads obtained from whole genome bisulfite sequencing. Figure S3: Bisulfite conversion efficiency. Figure S4: Quality measurements of deduplicated mapped reads obtained from whole genome bisulfite sequencing. Figure S5: Methylation bias and rates. Figure S6: Filtering of possible disease-relevant local variations identified in affected TOF twins based on zygosity differences between healthy and affected sibling. Figure S7: Filtering of possible disease-relevant structural variations (SVs) identified in affected TOF twins using whole genome sequencing. Figure S8: Global DNA methylation levels of twins. Line plots show the DNA methylation along the chromosomal length (GRCh38.p13/hg38). Figure S9: DNA methylation rate and coverage of CpGs. Table S1: Overview of sequencing reads obtained from whole genome sequencing. Table S2: Statistics of read mapping, coverage and read depth in each sample obtained from whole genome sequencing. Table S3: Overview of sequencing reads obtained from whole genome bisulfite sequencing. Table S4: Mapping result of reads obtained from whole genome bisulfite sequencing. Table S5: Read statistics after mapping and deduplication as well as methylation rates over CpGs, CHGs and CHHs in each sample obtained from whole genome bisulfite sequencing. Table S6: Copy number variations identified in Twin1_TOF and Twin2_TOF based on whole genome sequencing data. Table S7: Structural variations uniquely identified in Twin1_TOF based on whole genome sequencing data. Table S8: Structural variations uniquely identified in Twin2_TOF based on whole genome sequencing data. Table S9: Candidate genes with structural variations in Twin1_TOF and/or Twin2_TOF.

Author Contributions: Conceptualization, M.G. and S.R.S.; resources, P.G.; formal analysis, M.G. and S.A.; data curation, M.G. and S.A.; writing—original draft preparation, M.G.; writing—review and editing, M.G., S.R.S. and P.G.; visualization, M.G. and S.A.; supervision, S.R.S. and P.G.; funding acquisition, S.R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Berlin Institute of Health [BIH-CRG2-ConDi] (S.R.S.).

Acknowledgments: We are deeply grateful to the patients and families for their cooperation. We thank Andreas Perrot for review and discussion of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. van der Linde, D.; Konings, E.E.M.; Slager, M.A.; Witsenburg, M.; Helbing, W.A.; Takkenberg, J.J.M.; Roos-Hesselink, J.W. Birth prevalence of congenital heart disease worldwide: A systematic review and meta-analysis. *J. Am. Coll. Cardiol.* **2011**, *58*, 2241–2247. [[CrossRef](#)]
2. Nora, J.J. Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. *Circulation* **1968**, *38*, 604–617. [[CrossRef](#)] [[PubMed](#)]
3. Rickert-Sperling, S.; Kelly, R.G.; Driscoll, D.J. (Eds.) *Congenital Heart Diseases: The Broken Heart; Clinical Features, Human Genetics and Molecular Pathways*; Springer: Wien, Austria, 2016; ISBN 978-3-7091-1883-2.
4. Page, D.J.; Miossec, M.J.; Williams, S.G.; Monaghan, R.M.; Fotiou, E.; Cordell, H.J.; Sutcliffe, L.; Topf, A.; Bourgey, M.; Bourque, G.; et al. Whole Exome Sequencing Reveals the Major Genetic Contributors to Nonsyndromic Tetralogy of Fallot. *Circ. Res.* **2019**, *124*, 553–563. [[CrossRef](#)] [[PubMed](#)]
5. Fahed, A.C.; Gelb, B.D.; Seidman, J.G.; Seidman, C.E. Genetics of congenital heart disease: The glass half empty. *Circ. Res.* **2013**, *112*, 707–720. [[CrossRef](#)] [[PubMed](#)]
6. Andersen, T.A.; Troelsen, K.d.L.L.; Larsen, L.A. Of mice and men: Molecular genetics of congenital heart disease. *Cell. Mol. Life Sci.* **2014**, *71*, 1327–1352. [[CrossRef](#)] [[PubMed](#)]
7. Sifrim, A.; Hitz, M.-P.; Wilsdon, A.; Breckpot, J.; Turki, S.H.A.; Thienpont, B.; McRae, J.; Fitzgerald, T.W.; Singh, T.; Swaminathan, G.J.; et al. Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. *Nat. Genet.* **2016**, *48*, 1060–1065. [[CrossRef](#)] [[PubMed](#)]
8. Jin, S.C.; Homsy, J.; Zaidi, S.; Lu, Q.; Morton, S.; DePalma, S.R.; Zeng, X.; Qi, H.; Chang, W.; Sierant, M.C.; et al. Contribution of rare inherited and de novo variants in 2871 congenital heart disease probands. *Nat. Genet.* **2017**, *49*, 1593–1601. [[CrossRef](#)]
9. Grunert, M.; Dorn, C.; Schueler, M.; Dunkel, I.; Schlesinger, J.; Mebus, S.; Alexi-Meskishvili, V.; Perrot, A.; Wassilew, K.; Timmermann, B.; et al. Rare and private variations in neural crest, apoptosis and sarcomere genes define the polygenic background of isolated Tetralogy of Fallot. *Hum. Mol. Genet.* **2014**, *23*, 3115–3128. [[CrossRef](#)]
10. Grunert, M.; Dorn, C.; Cui, H.; Dunkel, I.; Schulz, K.; Schoenhals, S.; Sun, W.; Berger, F.; Chen, W.; Sperling, S.R. Comparative DNA methylation and gene expression analysis identifies novel genes for structural congenital heart diseases. *Cardiovasc. Res.* **2016**, *112*, 464–477. [[CrossRef](#)]
11. Yuan, W.; Xia, Y.; Bell, C.G.; Yet, I.; Ferreira, T.; Ward, K.J.; Gao, F.; Loomis, A.K.; Hyde, C.L.; Wu, H.; et al. An integrated epigenomic analysis for type 2 diabetes susceptibility loci in monozygotic twins. *Nat. Commun.* **2014**, *5*, 5719. [[CrossRef](#)]
12. Arora, M.; Reichenberg, A.; Willfors, C.; Austin, C.; Gennings, C.; Berggren, S.; Lichtenstein, P.; Anckarsäter, H.; Tammimies, K.; Bölte, S. Fetal and postnatal metal dysregulation in autism. *Nat. Commun.* **2017**, *8*, 15493. [[CrossRef](#)]
13. Castillo-Fernandez, J.E.; Spector, T.D.; Bell, J.T. Epigenetics of discordant monozygotic twins: Implications for disease. *Genome Med.* **2014**, *6*, 60. [[CrossRef](#)]
14. Lyu, G.; Zhang, C.; Ling, T.; Liu, R.; Zong, L.; Guan, Y.; Huang, X.; Sun, L.; Zhang, L.; Li, C.; et al. Genome and epigenome analysis of monozygotic twins discordant for congenital heart disease. *BMC Genom.* **2018**, *19*, 428. [[CrossRef](#)]
15. Andrews, S.R. FastQC. Available online: <https://github.com/s-andrews/FastQC> (accessed on 1 August 2020).
16. Okonechnikov, K.; Conesa, A.; García-Alcalde, F. Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **2016**, *32*, 292–294. [[CrossRef](#)]
17. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]

18. Poplin, R.; Chang, P.-C.; Alexander, D.; Schwartz, S.; Colthurst, T.; Ku, A.; Newburger, D.; Dijamco, J.; Nguyen, N.; Afshar, P.T.; et al. A universal SNP and small-indel variant caller using deep neural networks. *Nat. Biotechnol.* **2018**, *36*, 983–987. [[CrossRef](#)]
19. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.S.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genome Biol.* **2016**, *17*, 122. [[CrossRef](#)]
20. 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](#)]
21. Adzhubei, I.; Jordan, D.M.; Sunyaev, S.R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* **2013**, *7*, 7–20. [[CrossRef](#)]
22. Kumar, P.; Henikoff, S.; Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **2009**, *4*, 1073–1081. [[CrossRef](#)]
23. Schwarz, J.M.; Cooper, D.N.; Schuelke, M.; Seelow, D. MutationTaster2: Mutation prediction for the deep-sequencing age. *Nat. Methods* **2014**, *11*, 361–362. [[CrossRef](#)]
24. Kircher, M.; Witten, D.M.; Jain, P.; O’Roak, B.J.; Cooper, G.M.; Shendure, J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **2014**, *46*, 310–315. [[CrossRef](#)]
25. Chen, X.; Schulz-Trieglaff, O.; Shaw, R.; Barnes, B.; Schlesinger, F.; Källberg, M.; Cox, A.J.; Kruglyak, S.; Saunders, C.T. Manta: Rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **2016**, *32*, 1220–1222. [[CrossRef](#)]
26. Boeva, V.; Popova, T.; Bleakley, K.; Chiche, P.; Cappel, J.; Schleiermacher, G.; Janoueix-Lerosey, I.; Delattre, O.; Barillot, E. Control-FREEC: A tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* **2012**, *28*, 423–425. [[CrossRef](#)]
27. Krueger, F. Trim Galore. Available online: <https://github.com/FelixKrueger/TrimGalore> (accessed on 1 August 2020).
28. Krueger, F.; Andrews, S.R. Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **2011**, *27*, 1571–1572. [[CrossRef](#)]
29. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. 1000 Genome Project Data Processing Subgroup the Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
30. Devon, Ryan MethylDackel. Available online: <https://github.com/dpryan79/MethylDackel> (accessed on 1 August 2020).
31. Park, Y.; Figueroa, M.E.; Rozek, L.S.; Sartor, M.A. MethylSig: A whole genome DNA methylation analysis pipeline. *Bioinformatics* **2014**, *30*, 2414–2422. [[CrossRef](#)]
32. 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.; et al. Large-scale discovery of enhancers from human heart tissue. *Nat. Genet.* **2011**, *44*, 89–93. [[CrossRef](#)]
33. Kaynak, B.; von Heydebreck, A.; Mebus, S.; Seelow, D.; Hennig, S.; Vogel, J.; Sperling, H.-P.; Pregla, R.; Alexi-Meskishvili, V.; Hetzer, R.; et al. Genome-wide array analysis of normal and malformed human hearts. *Circulation* **2003**, *107*, 2467–2474. [[CrossRef](#)]
34. Toenjes, M.; Schueler, M.; Hammer, S.; Pape, U.J.; Fischer, J.J.; Berger, F.; Vingron, M.; Sperling, S. Prediction of cardiac transcription networks based on molecular data and complex clinical phenotypes. *Mol. Biosyst.* **2008**, *4*, 589–598. [[CrossRef](#)]
35. Ricci, M.; Xu, Y.; Hammond, H.L.; Willoughby, D.A.; Nathanson, L.; Rodriguez, M.M.; Vatta, M.; Lipshultz, S.E.; Lincoln, J. Myocardial alternative RNA splicing and gene expression profiling in early stage hypoplastic left heart syndrome. *PLoS ONE* **2012**, *7*, e29784. [[CrossRef](#)] [[PubMed](#)]
36. Schlesinger, J.; Tönjes, M.; Schueler, M.; Zhang, Q.; Dunkel, I.; Sperling, S.R. Evaluation of the LightCycler 1536 Instrument for high-throughput quantitative real-time PCR. *Methods* **2010**, *50*, S19–S22. [[CrossRef](#)]
37. Rodemoyer, A.; Kibiryeva, N.; Bair, A.; Marshall, J.; O’Brien, J.E.; Bittel, D.C. A tissue-specific gene expression template portrays heart development and pathology. *Hum. Genom.* **2014**, *8*, 6. [[CrossRef](#)]
38. Sheng, W.; Wang, H.; Ma, X.; Qian, Y.; Zhang, P.; Wu, Y.; Zheng, F.; Chen, L.; Huang, G.; Ma, D. LINE-1 methylation status and its association with tetralogy of fallot in infants. *BMC Med. Genom.* **2012**, *5*, 20. [[CrossRef](#)]

39. Sheng, W.; Qian, Y.; Zhang, P.; Wu, Y.; Wang, H.; Ma, X.; Chen, L.; Ma, D.; Huang, G. Association of promoter methylation statuses of congenital heart defect candidate genes with Tetralogy of Fallot. *J. Transl. Med.* **2014**, *12*, 31. [[CrossRef](#)]
40. Chowdhury, S.; Erickson, S.W.; MacLeod, S.L.; Cleves, M.A.; Hu, P.; Karim, M.A.; Hobbs, C.A. Maternal genome-wide DNA methylation patterns and congenital heart defects. *PLoS ONE* **2011**, *6*, e16506. [[CrossRef](#)]
41. Sheng, W.; Qian, Y.; Wang, H.; Ma, X.; Zhang, P.; Diao, L.; An, Q.; Chen, L.; Ma, D.; Huang, G. DNA methylation status of NKX2-5, GATA4 and HAND1 in patients with tetralogy of fallot. *BMC Med. Genom.* **2013**, *6*, 46. [[CrossRef](#)]
42. Bansal, V.; Dorn, C.; Grunert, M.; Klaassen, S.; Hetzer, R.; Berger, F.; Sperling, S.R. Outlier-based identification of copy number variations using targeted resequencing in a small cohort of patients with Tetralogy of Fallot. *PLoS ONE* **2014**, *9*, e85375. [[CrossRef](#)]
43. Warburton, D.; Ronemus, M.; Kline, J.; Jobanputra, V.; Williams, I.; Anyane-Yeboah, K.; Chung, W.; Yu, L.; Wong, N.; Awad, D.; et al. The contribution of de novo and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. *Hum. Genet.* **2014**, *133*, 11–27. [[CrossRef](#)]
44. Greenway, S.C.; Pereira, A.C.; Lin, J.C.; DePalma, S.R.; Israel, S.J.; Mesquita, S.M.; Ergul, E.; Conta, J.H.; Korn, J.M.; McCarroll, S.A.; et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat. Genet.* **2009**, *41*, 931–935. [[CrossRef](#)]
45. Tomita-Mitchell, A.; Mahnke, D.K.; Struble, C.A.; Tuffnell, M.E.; Stamm, K.D.; Hidestrand, M.; Harris, S.E.; Goetsch, M.A.; Simpson, P.M.; Bick, D.P.; et al. Human gene copy number spectra analysis in congenital heart malformations. *Physiol. Genom.* **2012**, *44*, 518–541. [[CrossRef](#)] [[PubMed](#)]
46. Payne, A.R.; Chang, S.-W.; Koenig, S.N.; Zinn, A.R.; Garg, V. Submicroscopic chromosomal copy number variations identified in children with hypoplastic left heart syndrome. *Pediatr. Cardiol.* **2012**, *33*, 757–763. [[CrossRef](#)]
47. de Souza, K.R.; Mergener, R.; Huber, J.; Campos Pellanda, L.; Riegel, M. Cytogenomic Evaluation of Subjects with Syndromic and Nonsyndromic Conotruncal Heart Defects. *Biomed. Res. Int.* **2015**, *2015*, 401941. [[CrossRef](#)]
48. Geng, J.; Picker, J.; Zheng, Z.; Zhang, X.; Wang, J.; Hisama, F.; Brown, D.W.; Mullen, M.P.; Harris, D.; Stoler, J.; et al. Chromosome microarray testing for patients with congenital heart defects reveals novel disease causing loci and high diagnostic yield. *BMC Genom.* **2014**, *15*, 1127. [[CrossRef](#)] [[PubMed](#)]
49. Bittel, D.C.; Zhou, X.-G.; Kibiriyeva, N.; Fiedler, S.; O'Brien, J.E.; Marshall, J.; Yu, S.; Liu, H.-Y. Ultra high-resolution gene centric genomic structural analysis of a non-syndromic congenital heart defect, Tetralogy of Fallot. *PLoS ONE* **2014**, *9*, e87472. [[CrossRef](#)] [[PubMed](#)]
50. Aguayo-Gómez, A.; Arteaga-Vázquez, J.; Svyryd, Y.; Calderón-Colmenero, J.; Zamora-González, C.; Vargas-Alarcón, G.; Mutchinick, O.M. Identification of Copy Number Variations in Isolated Tetralogy of Fallot. *Pediatr. Cardiol.* **2015**, *36*, 1642–1646. [[CrossRef](#)]
51. Hitz, M.-P.; Lemieux-Perreault, L.-P.; Marshall, C.; Feroz-Zada, Y.; Davies, R.; Yang, S.W.; Lionel, A.C.; D'Amours, G.; Lemyre, E.; Cullum, R.; et al. Rare copy number variants contribute to congenital left-sided heart disease. *PLoS Genet.* **2012**, *8*, e1002903. [[CrossRef](#)] [[PubMed](#)]
52. Erdogan, F.; Larsen, L.A.; Zhang, L.; Tümer, Z.; Tommerup, N.; Chen, W.; Jacobsen, J.R.; Schubert, M.; Jurkatis, J.; Tzschach, A.; et al. High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. *J. Med. Genet.* **2008**, *45*, 704–709. [[CrossRef](#)]
53. Soemedi, R.; Wilson, I.J.; Bentham, J.; Darlay, R.; Töpf, A.; Zelenika, D.; Cosgrove, C.; Setchfield, K.; Thornborough, C.; Granados-Riveron, J.; et al. Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *Am. J. Hum. Genet.* **2012**, *91*, 489–501. [[CrossRef](#)]
54. Wilson, K.D.; Shen, P.; Fung, E.; Karakikes, I.; Zhang, A.; InanlooRahatloo, K.; Odegaard, J.; Sallam, K.; Davis, R.W.; Lui, G.K.; et al. A rapid, high-quality, cost-effective, comprehensive and expandable targeted next-generation sequencing assay for inherited heart diseases. *Circ. Res.* **2015**, *117*, 603–611. [[CrossRef](#)]
55. Silversides, C.K.; Lionel, A.C.; Costain, G.; Merico, D.; Migita, O.; Liu, B.; Yuen, T.; Rickaby, J.; Thiruvahindrapuram, B.; Marshall, C.R.; et al. Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet.* **2012**, *8*, e1002843. [[CrossRef](#)]
56. Richards, A.A.; Santos, L.J.; Nichols, H.A.; Crider, B.P.; Elder, F.F.; Hauser, N.S.; Zinn, A.R.; Garg, V. Cryptic chromosomal abnormalities identified in children with congenital heart disease. *Pediatr. Res.* **2008**, *64*, 358–363. [[CrossRef](#)]

57. O'Brien, J.E.; Kibiryeveva, N.; Zhou, X.-G.; Marshall, J.A.; Lofland, G.K.; Artman, M.; Chen, J.; Bittel, D.C. Noncoding RNA expression in myocardium from infants with tetralogy of Fallot. *Circ. Cardiovasc. Genet.* **2012**, *5*, 279–286. [[CrossRef](#)]
58. Sucharov, C.C.; Sucharov, J.; Karimpour-Fard, A.; Nunley, K.; Stauffer, B.L.; Miyamoto, S.D. Micro-RNA expression in hypoplastic left heart syndrome. *J. Card. Fail.* **2015**, *21*, 83–88. [[CrossRef](#)]
59. Grunert, M.; Appelt, S.; Dunkel, I.; Berger, F.; Sperling, S.R. Altered microRNA and target gene expression related to Tetralogy of Fallot. *Sci. Rep.* **2019**, *9*, 19063. [[CrossRef](#)]
60. Grunert, M.; Appelt, S.; Schönhals, S.; Mika, K.; Cui, H.; Cooper, A.; Cyganek, L.; Guan, K.; Sperling, S.R. Induced pluripotent stem cells of patients with Tetralogy of Fallot reveal transcriptional alterations in cardiomyocyte differentiation. *Sci. Rep.* **2020**, *10*, 10921. [[CrossRef](#)]
61. Wu, B.; Baldwin, H.S.; Zhou, B. Nfatc1 directs the endocardial progenitor cells to make heart valve primordium. *Trends Cardiovasc. Med.* **2013**, *23*, 294–300. [[CrossRef](#)]
62. Kirk, E.P.; Sunde, M.; Costa, M.W.; Rankin, S.A.; Wolstein, O.; Castro, M.L.; Butler, T.L.; Hyun, C.; Guo, G.; Otway, R.; et al. Mutations in cardiac T-box factor gene TBX20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. *Am. J. Hum. Genet.* **2007**, *81*, 280–291. [[CrossRef](#)]
63. Grunert, M.; Dorn, C.; Rickert-Sperling, S. Cardiac Transcription Factors and Regulatory Networks. In *Congenital Heart Diseases: The Broken Heart*; Rickert-Sperling, S., Kelly, R.G., Driscoll, D.J., Eds.; Springer: Vienna, Austria, 2016; pp. 139–152. ISBN 978-3-7091-1882-5.
64. Bocconi, P.; MacGrogan, D.; Scandura, J.M.; Nimer, S.D. The human L(3)MBT polycomb group protein is a transcriptional repressor and interacts physically and functionally with TEL (ETV6). *J. Biol. Chem.* **2003**, *278*, 15412–15420. [[CrossRef](#)]
65. Trojer, P.; Li, G.; Sims, R.J.; Vaquero, A.; Kalakonda, N.; Bocconi, P.; Lee, D.; Erdjument-Bromage, H.; Tempst, P.; Nimer, S.D.; et al. L3MBTL1, a histone-methylation-dependent chromatin lock. *Cell* **2007**, *129*, 915–928. [[CrossRef](#)]
66. Ziller, M.J.; Gu, H.; Müller, F.; Donaghey, J.; Tsai, L.T.-Y.; Kohlbacher, O.; De Jager, P.L.; Rosen, E.D.; Bennett, D.A.; Bernstein, B.E.; et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature* **2013**, *500*, 477–481. [[CrossRef](#)]
67. Chen, P.-Y.; Feng, S.; Joo, J.W.J.; Jacobsen, S.E.; Pellegrini, M. A comparative analysis of DNA methylation across human embryonic stem cell lines. *Genome Biol.* **2011**, *12*, R62. [[CrossRef](#)]
68. Gardiner-Garden, M.; Frommer, M. CpG islands in vertebrate genomes. *J. Mol. Biol.* **1987**, *196*, 261–282. [[CrossRef](#)]
69. Deaton, A.M.; Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **2011**, *25*, 1010–1022. [[CrossRef](#)]
70. Jones, P.A. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **2012**, *13*, 484–492. [[CrossRef](#)]
71. Yehya, A.; Souki, R.; Bitar, F.; Nemer, G. Differential duplication of an intronic region in the NFATC1 gene in patients with congenital heart disease. *Genome* **2006**, *49*, 1092–1098. [[CrossRef](#)]
72. Abdul-Sater, Z.; Yehya, A.; Beresian, J.; Salem, E.; Kamar, A.; Baydoun, S.; Shabbani, K.; Soubra, A.; Bitar, F.; Nemer, G. Two heterozygous mutations in NFATC1 in a patient with Tricuspid Atresia. *PLoS ONE* **2012**, *7*, e49532. [[CrossRef](#)]
73. Takeuchi, J.K.; Mileikowska, M.; Koshiba-Takeuchi, K.; Heidt, A.B.; Mori, A.D.; Arruda, E.P.; Gertsenstein, M.; Georges, R.; Davidson, L.; Mo, R.; et al. Tbx20 dose-dependently regulates transcription factor networks required for mouse heart and motoneuron development. *Development* **2005**, *132*, 2463–2474. [[CrossRef](#)]
74. Irizarry, R.A.; Ladd-Acosta, C.; Wen, B.; Wu, Z.; Montano, C.; Onyango, P.; Cui, H.; Gabo, K.; Rongione, M.; Webster, M.; et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.* **2009**, *41*, 178–186. [[CrossRef](#)]
75. Xu, Y.; Li, T.; Pu, T.; Cao, R.; Long, F.; Chen, S.; Sun, K.; Xu, R. Copy Number Variants and Exome Sequencing Analysis in Six Pairs of Chinese Monozygotic Twins Discordant for Congenital Heart Disease. *Twin Res. Hum. Genet.* **2017**, *20*, 521–532. [[CrossRef](#)]
76. Breckpot, J.; Thienpont, B.; Gewillig, M.; Allegaert, K.; Vermeesch, J.R.; Devriendt, K. Differences in Copy Number Variation between Discordant Monozygotic Twins as a Model for Exploring Chromosomal Mosaicism in Congenital Heart Defects. *Mol. Syndromol.* **2012**, *2*, 81–87. [[CrossRef](#)]

77. Chaivasap, P.; Kulawonganunchai, S.; Srichomthong, C.; Tongshima, S.; Suphapeetiporn, K.; Shotelersuk, V. Whole genome and exome sequencing of monozygotic twins with trisomy 21, discordant for a congenital heart defect and epilepsy. *PLoS ONE* **2014**, *9*, e100191. [[CrossRef](#)]
78. Hui, D.S.; Bonow, R.O.; Stolker, J.M.; Braddock, S.R.; Lee, R. Discordant Aortic Valve Morphology in Monozygotic Twins: A Clinical Case Series. *JAMA Cardiol.* **2016**, *1*, 1043–1047. [[CrossRef](#)]
79. Zhang, R.; Thiele, H.; Bartmann, P.; Hilger, A.C.; Berg, C.; Herberg, U.; Klingmüller, D.; Nürnberg, P.; Ludwig, M.; Reutter, H. Whole-Exome Sequencing in Nine Monozygotic Discordant Twins. *Twin Res. Hum. Genet.* **2016**, *19*, 60–65. [[CrossRef](#)]
80. Erlich, Y. Blood ties: Chimerism can mask twin discordance in high-throughput sequencing. *Twin Res. Hum. Genet.* **2011**, *14*, 137–143. [[CrossRef](#)]
81. Zaidi, S.; Choi, M.; Wakimoto, H.; Ma, L.; Jiang, J.; Overton, J.D.; Romano-Adesman, A.; Bjornson, R.D.; Breitbart, R.E.; Brown, K.K.; et al. De novo mutations in histone-modifying genes in congenital heart disease. *Nature* **2013**, *498*, 220–223. [[CrossRef](#)]
82. Lindinger, A.; Schwedler, G.; Hense, H.-W. Prevalence of congenital heart defects in newborns in Germany: Results of the first registration year of the PAN Study (July 2006 to June 2007). *Klin. Padiatr.* **2010**, *222*, 321–326. [[CrossRef](#)]
83. Cosselman, K.E.; Navas-Acien, A.; Kaufman, J.D. Environmental factors in cardiovascular disease. *Nat. Rev. Cardiol.* **2015**, *12*, 627–642. [[CrossRef](#)]
84. Porter, G.A. Environmental Signals. In *Congenital Heart Diseases: The Broken Heart*; Rickert-Sperling, S., Kelly, R.G., Driscoll, D.J., Eds.; Springer: Vienna, Austria, 2016; pp. 223–235. ISBN 978-3-7091-1882-5.
85. Abu-Halima, M.; Weidinger, J.; Poryo, M.; Henn, D.; Keller, A.; Meese, E.; Abdul-Khaliq, H. Micro-RNA signatures in monozygotic twins discordant for congenital heart defects. *PLoS ONE* **2019**, *14*, e0226164. [[CrossRef](#)]
86. Czyz, W.; Morahan, J.M.; Ebers, G.C.; Ramagopalan, S.V. Genetic, environmental and stochastic factors in monozygotic twin discordance with a focus on epigenetic differences. *BMC Med.* **2012**, *10*, 93. [[CrossRef](#)]
87. Litviňuková, M.; Talavera-López, C.; Maatz, H.; Reichart, D.; Worth, C.L.; Lindberg, E.L.; Kanda, M.; Polanski, K.; Heinig, M.; Lee, M.; et al. Cells of the adult human heart. *Nature* **2020**. [[CrossRef](#)]
88. Karatza, A.A.; Wolfenden, J.L.; Taylor, M.J.O.; Wee, L.; Fisk, N.M.; Gardiner, H.M. Influence of twin-twin transfusion syndrome on fetal cardiovascular structure and function: Prospective case-control study of 136 monochorionic twin pregnancies. *Heart* **2002**, *88*, 271–277. [[CrossRef](#)]
89. Marceau, K.; McMaster, M.T.B.; Smith, T.F.; Daams, J.G.; van Beijsterveldt, C.E.M.; Boomsma, D.I.; Knopik, V.S. The Prenatal Environment in Twin Studies: A Review on Chorionicity. *Behav. Genet.* **2016**, *46*, 286–303. [[CrossRef](#)]
90. Wong, H.S.; Kidd, A.; Zuccollo, J.; Parker, S.; Richardson, V.; Tait, J.; Pringle, K.C. A case of amyoplasia in a monochorionic twin pregnancy: A sequela from twin-twin transfusion syndrome? *Fetal Diagn. Ther.* **2009**, *25*, 31–35. [[CrossRef](#)] [[PubMed](#)]
91. Norris, A.W.; Wang, C.; Yao, J.; Walsh, S.A.; Sawatzke, A.B.; Hu, S.; Sunderland, J.J.; Segar, J.L.; Ponto, L.L.B. Effect of insulin and dexamethasone on fetal assimilation of maternal glucose. *Endocrinology* **2011**, *152*, 255–262. [[CrossRef](#)]
92. Mahle, W.T. What we can learn from twins: Congenital heart disease in the danish twin registry. *Circulation* **2013**, *128*, 1173–1174. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).