

# Rare Variants in *NR2F2* Cause Congenital Heart Defects in Humans

Saeed Al Turki,<sup>1,2,22</sup> Ashok K. Manickaraj,<sup>3,22</sup> Catherine L. Mercer,<sup>4,22</sup> Sebastian S. Gerety,<sup>1,22</sup> Marc-Phillip Hitz,<sup>1</sup> Sarah Lindsay,<sup>1</sup> Lisa C.A. D'Alessandro,<sup>3</sup> G. Jawahar Swaminathan,<sup>1</sup> Jamie Bentham,<sup>5</sup> Anne-Karin Arndt,<sup>6,7</sup> Jacoba Low,<sup>8,9</sup> Jeroen Breckpot,<sup>8</sup> Marc Gewillig,<sup>9</sup> Bernard Thienpont,<sup>8</sup> Hashim Abdul-Khaliq,<sup>10,11</sup> Christine Harnack,<sup>12</sup> Kirstin Hoff,<sup>7,13</sup> Hans-Heiner Kramer,<sup>7,11</sup> Stephan Schubert,<sup>11,14</sup> Reiner Siebert,<sup>13</sup> Okan Toka,<sup>11,15</sup> Catherine Cosgrove,<sup>16</sup> Hugh Watkins,<sup>16</sup> Anneke M. Lucassen,<sup>4</sup> Ita M. O'Kelly,<sup>4</sup> Anthony P. Salmon,<sup>4</sup> Frances A. Bu'Lock,<sup>17</sup> Javier Granados-Riveron,<sup>18</sup> Kerry Setchfield,<sup>18</sup> Chris Thornborough,<sup>17</sup> J. David Brook,<sup>18</sup> Barbara Mulder,<sup>19</sup> Sabine Klaassen,<sup>11,12,20</sup> Shoumo Bhattacharya,<sup>16</sup> Koen Devriendt,<sup>8</sup> David F. FitzPatrick,<sup>21</sup> UK10K Consortium, David I. Wilson,<sup>4,23</sup> Seema Mital,<sup>3,23,\*</sup> and Matthew E. Hurles<sup>1,23,\*</sup>

Congenital heart defects (CHDs) are the most common birth defect worldwide and are a leading cause of neonatal mortality. Nonsyndromic atrioventricular septal defects (AVSDs) are an important subtype of CHDs for which the genetic architecture is poorly understood. We performed exome sequencing in 13 parent-offspring trios and 112 unrelated individuals with nonsyndromic AVSDs and identified five rare missense variants (two of which arose de novo) in the highly conserved gene *NR2F2*, a very significant enrichment ( $p = 7.7 \times 10^{-7}$ ) compared to 5,194 control subjects. We identified three additional CHD-affected families with other variants in *NR2F2* including a de novo balanced chromosomal translocation, a de novo substitution disrupting a splice donor site, and a 3 bp duplication that cosegregated in a multiplex family. *NR2F2* encodes a pleiotropic developmental transcription factor, and decreased dosage of *NR2F2* in mice has been shown to result in abnormal development of atrioventricular septa. Via luciferase assays, we showed that all six coding sequence variants observed in individuals significantly alter the activity of *NR2F2* on target promoters.

## Introduction

Fewer than 20% of congenital heart defects (CHDs) can be attributed to large structural chromosomal variants or single-gene mutations causing monogenic syndromes.<sup>1</sup> The majority of CHDs are nonsyndromic (individuals without extracardiac phenotypes) and are of unknown etiology.<sup>2</sup> Mouse knockout studies have identified more than 300 genes in which (typically homozygous) loss-of-function mutations are sufficient to cause CHDs, and, given that only a minority of genes have been knocked out in mice thus far, hundreds more genes essential for cardiac development remain to be identified.<sup>3</sup>

Atrioventricular septal defects (AVSDs [MIM 606215]) cover a spectrum of CHDs characterized by a common atrioventricular junction coexisting with deficient atrioventricular septation. AVSDs represent 4%–5% of all CHDs and their prevalence ranges from 0.3 to 0.4 per 1,000 live births.<sup>4,5</sup> However, their prevalence is much higher in utero based on large fetal echocardiographic series where they were found to account for 18% of CHD-affected individuals.<sup>6</sup> The discrepancy in the prevalence may be attributed to the fact that many of the AVSD-affected fetuses will not survive until birth either because they die prematurely or because of elective termination. Postnatally, certain individual groups have a higher AVSD prevalence such as Down syndrome (DS [MIM 190685]) where 44% of

<sup>1</sup>Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK; <sup>2</sup>Department of Pathology, King Abdulaziz Medical City, P.O. Box 22490, Riyadh 11426, Saudi Arabia; <sup>3</sup>Division of Cardiology, Department of Pediatrics, Hospital for Sick Children, University of Toronto, Toronto, ON M5G 1X8, Canada; <sup>4</sup>Human Development and Health Academic Unit, Faculty of Medicine, University of Southampton, Southampton General Hospital, Southampton SO16 6YD, UK; <sup>5</sup>Department of Cardiology, Boston Children's Hospital, Harvard Medical School, 300 Longwood Avenue, Boston, MA 02459, USA; <sup>6</sup>Cardiovascular Division, Brigham and Women's Hospital, Harvard Medical School, and Harvard Stem Cell Institute, Boston, MA 02115, USA; <sup>7</sup>Department of Congenital Heart Disease and Pediatric Cardiology, University Hospital Schleswig-Holstein, Campus Kiel, 24105 Kiel, Germany; <sup>8</sup>Centre for Human Genetics, Katholieke Universiteit Leuven, 3000 Leuven, Belgium; <sup>9</sup>Pediatric Cardiology Unit, University Hospital Leuven, 3000 Leuven, Belgium; <sup>10</sup>Department of Pediatric Cardiology, Saarland University Hospital, 66421 Homburg, Germany; <sup>11</sup>Competence Network for Congenital Heart Defects; <sup>12</sup>Experimental and Clinical Research Center (ECRC), Charité Medical Faculty and Max-Delbrück-Center for Molecular Medicine, 13125 Berlin, Germany; <sup>13</sup>Institute of Human Genetics, Christian-Albrechts University Kiel & University Hospital Schleswig-Holstein, Campus Kiel, 24105 Kiel, Germany; <sup>14</sup>Department of Congenital Heart Disease and Pediatric Cardiology, Deutsches Herzzentrum Berlin, 13353 Berlin, Germany; <sup>15</sup>Department of Pediatric Cardiology, Children's Hospital, Friedrich-Alexander University, 91054 Erlangen, Germany; <sup>16</sup>Radcliffe Department of Medicine & Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK; <sup>17</sup>East Midlands Congenital Heart Centre, University Hospitals of Leicester NHS Trust, Leicester LE3 9QP, UK; <sup>18</sup>School of Life Sciences, University of Nottingham, Nottingham NG7 2UH, UK; <sup>19</sup>Heart Center, Academic Medical Center, 1105AZ Amsterdam, the Netherlands; <sup>20</sup>Department of Pediatric Cardiology, Charité University Medicine Berlin, 13353 Berlin, Germany; <sup>21</sup>MRC Human Genetics Unit, Institute of Genetic and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK

<sup>22</sup>These authors contributed equally to this work

<sup>23</sup>These authors contributed equally to this work

\*Correspondence: [seema.mital@sickkids.ca](mailto:seema.mital@sickkids.ca) (S.M.), [meh@sanger.ac.uk](mailto:meh@sanger.ac.uk) (M.E.H.)

<http://dx.doi.org/10.1016/j.ajhg.2014.03.007>. ©2014 The Authors

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

DS-affected individuals have CHDs, of which 39% are AVSDs.<sup>7</sup> The presence of three copies of chromosome 21 increases the risk of AVSDs ~2,000-fold,<sup>8</sup> but this is not sufficient to explain why half of those with DS do not exhibit either AVSDs or other CHDs. Many hypotheses have been proposed to explain this, for example that rare missense variants in VEGF-A pathway genes (on chromosome 21) increase the risk of AVSDs in DS.<sup>9</sup> AVSDs have also been observed in several other multisystem genetic syndromes that frequently result in CHDs. However, AVSDs account only for a minor fraction of CHD cases in these syndromic individuals. In nonsyndromic AVSDs, a small number of genes, including *CRELD1* (MIM 607170),<sup>10,11</sup> *ALK2* (MIM 102576),<sup>12</sup> *TBX5* (MIM 601620),<sup>13</sup> and *GATA4* (MIM 600576),<sup>14,15</sup> have been implicated, but either the genetic evidence for a pathogenic role in nonsyndromic AVSDs is weak or AVSDs are a much less frequent consequence than other CHD subtypes.

A recent exome-sequencing study suggested that de novo mutations collectively contribute to the underlying cause in 10% of a diverse collection of syndromic and nonsyndromic CHD-affected individuals.<sup>16</sup> Here we adopted a more targeted strategy, focusing on a specific subtype of nonsyndromic CHDs, AVSDs, initially in 13 parent-offspring trios and a larger cohort of 112 unrelated individuals.

In the current study we report an enrichment of likely causal variants in *NR2F2* in families with isolated AVSDs as well as other isolated CHD phenotypes including coarctation of aorta (CoA [MIM 120000]). Two published mouse models indicate an important role for *Nr2f2* in the development of the heart, displaying atrioventricular septal and valvular defects.<sup>17,18</sup> We also demonstrate the expression of *NR2F2* in the developing human fetal heart including the atria, coronary vessels, and aorta. In addition, the results from luciferase assays showed that all six coding sequence variants observed in cases significantly alter the activity of *NR2F2* on target promoters. Taken together, these data support our hypothesis that rare and private variants in *NR2F2* probably contribute to AVSDs and other CHDs during human development.

## Subjects and Methods

### Subjects

Individuals were enrolled prospectively in an Ontario province-wide Biobank registry and GO-CHD (Oxford). Informed consent was obtained from parents or legal guardians. The replication cohort included 245 individuals from different centers. These included 120 individuals from the CONCOR registry and DNA-bank, a joint registry of the Dutch Heart Foundation and the Inter-university Cardiology Institute Netherlands (ICIN) of adults of European ancestry with congenital heart disease and 18 individuals from the National Register for Congenital Heart Defects, Germany (for details, see [Tables S1A–S1C](#) available online). Other smaller sample sets were collected by the same criteria from different centers. The local ethics committees of each of the centers that recruited the participating individuals approved this study.

### Exome Sequencing

Samples were sequenced at the Wellcome Trust Sanger Institute. Genomic DNA from venous blood or saliva was obtained and captured by SureSelect Target Enrichment V3 (Agilent) and sequenced (HiSeq Illumina 75 bp pair-end reads). Reads were mapped to the reference genome via BWA.<sup>19</sup> Single-nucleotide variants were called by SAMtools<sup>20</sup> and GATK<sup>21</sup> and indels were called by SAMtools and Dindel.<sup>22</sup> Variants were annotated for allele frequency by 1000 Genomes (June 2012 release), NHLBI-ESP (6503) project, and UK10K cohorts. Variant Effect Predictor<sup>23</sup> was used to annotate the impact on the protein structure and GERP for nucleotide conservation scores.<sup>24</sup>

### Identification and Confirmation of De Novo Mutations

We used DenovoGear<sup>25</sup> to detect de novo variants from BCF files generated by SAMtools mpileup and BCFtools.<sup>20</sup> To minimize the false positive rate, we excluded variants in tandem repeat or segmental duplication regions (UCSC genome tables) or common variants with allele frequency of >1% in 1000 Genomes, NHLBI-ESP project, and UK10K. We also filtered out variants with >10% of alternative reads supporting the alternative allele in at least one parent and manually checked the sequencing context for all coding variants via Integrative Genomics Viewer (IGV). All coding variants (silent, splice site, missense, frameshift, or stop gain and stop lost) underwent validation by capillary sequencing in the child and both parents ([Table S2](#)).

### Burden Test of Rare and De Novo Missense Mutations in *NR2F2*

Before performing the burden analysis, we removed related samples based on the reported pedigree and by additionally checking for relatedness between samples via SNPRelate R package.<sup>26</sup> To avoid the confounding effect of population stratification, we compared the case cohort and UK10K control cohort to four HapMap populations (CEU, YRI, CHB, and JPT)<sup>27</sup> via principal component analysis and removed non-European samples. A rare variant is defined as a variant with an allele frequency of <1% in the 1000 Genomes Project data. The exome-sequenced controls comprise 894 UK10K samples with autism or schizophrenia and 4,300 European Americans in the NHLBI-ESP project. We generated a 2 by 2 table for each gene for the number of reference and alternative alleles in cases and controls and assessed significance by the Fisher's exact test ([Table S3](#)). The mutation rate of missense variants in *NR2F2* was estimated (for de novo burden analyses) based on the length of the *NR2F2* coding region (1,245 bp), an average single-nucleotide mutation rate in coding regions of  $1.5 \times 10^{-8}$  per base per generation, and the expected proportion of de novo mutations that are missense.<sup>28</sup>

### *NR2F2* Expression Plasmids and Luciferase Constructs

To generate expression plasmids for *NR2F2* (MIM 107773) and its variants, the human wild-type *NR2F2* (RefSeq accession numbers NM\_021005.3, NP\_066285.1) coding sequence was PCR amplified from a full-length EST (GenBank accession number BC042897) and cloned by Gibson assembly (New England Biolabs) into a CMV-driven pCS2-Cherry plasmid. We recreated the mutant forms of *NR2F2* c.222\_224dup (p.Gln75dup), c.509A>T (p.Asp170Val), c.614A>T (p.Asn205Ile), c.753G>C (p.Glu251Asp), c.1022C>A (p.Ser341Tyr), and c.1234G>T (p.Ala412Ser) by amplifying two PCR fragments introducing

each mutation, and cloned these as above. All nucleotide changes above relate to RefSeq NM\_021005.3. These expression constructs produce fusion proteins with fluorescent cherry domain<sup>29</sup> in order to monitor expression and localization. To create the *NGFI-A* (MIM 128990) and *APOB* (MIM 107730) promoter-driven Luciferase plasmids, we cloned synthetic DNA fragments for the rat *NGFI-A* upstream genomic region from –389 to +43<sup>30</sup> and the human *APOB* upstream region from –139 to +121<sup>31</sup> into a promoterless pGL3 Luciferase plasmid (Promega) by Gibson assembly (New England Biolabs).

### Luciferase Assays

HEK293T and HEPG2 cells were plated in 96-well plates and transfected with 30 ng of either *NGFI-A* or *APOB* luciferase plasmids, 0.75 ng of RL-TK renilla plasmid (Promega), and either 30 ng of *NR2F2* expression plasmid (wild-type or variants) or 30 ng of Cherry plasmid as a control. Two days after transfection, the cells were lysed and assayed for luciferase activity by the Dual-Luciferase Reporter Assay System, according to the manufacturer's instructions (Promega). Each transfection was done in replicates (minimum three times) and the experiments were repeated three to four times. Luciferase readings were first normalized to the transfection control (renilla plasmid). Relative Response Ratios (Promega) were calculated based on negative and positive controls (cherry and *NR2F2* plasmid transfections), and outliers across all experiments were identified by a median absolute deviation ratio >3. A *t* test was performed to identify significant differences between variants and between promoters.

### Mapping Breakpoint Sequence

Flow-sorted derivative chromosomes 14 and 15 were used as template to map the breakpoint via methods previously described.<sup>32</sup> The derivative 14 breakpoint was identified with the combination of forward primer 5'-TGGGTGACACAGCAAGACTG-3' (chr 14) and reverse primer 5'-GGGGAGGAAAGGAGACACTC-3' (chr 15), which amplified a product of 431 bases that was capillary sequenced.

### Immunohistochemistry

Immunolocalization of proteins in human fetal heart tissue was carried out via protocols previously reported.<sup>33–35</sup> Fetal tissue was obtained with informed consent and according to the protocol ethically approved by Southampton and South West Hants LREC. Slides were incubated with primary antibodies (anti-rabbit raised to *NR2F2*, 1 in 400 [Abcam]; anti-mouse to CD34, 1 in 200 [Novocastra]; Troponin C 1 in 200 [Novocastra]; and SMA, 1 in 100 [Novocastra]). Secondary antibodies were applied (FITC-conjugated anti-rabbit Ig [Sigma, 1 in 200] and Alexa-594 conjugated anti-mouse Ig [Life Technologies, 1 in 200]). Slides were further washed in PBS before dehydrating and mounting sections in Vectashield (Vector Laboratories) with DAPI nuclear counterstain. Visualization and image capture of sections was performed with a Zeiss Axioplan fluorescence microscope and software (Carl Zeiss).

### Whole-Mount In Situ Hybridization

Primers including T3 and T7 promoter sequences were designed against the 3' UTR of *Nr2f2* (T3-Forward 5'-AATTAACCTCAC TAAAGGAGCCAAGGAATGTGTCCAAG-3' and T7-Reverse 5'-TAA TACGACTCACTATAGGGAGAACTCACAGGGGCTCAG-3'). PCR products were generated with murine DNA from outbred albino mouse strain CD-1. Sense (T3) and antisense (T7) riboprobes

were made by in vitro transcription with T3 or T7 polymerase (Roche) and with the PCR products as template. Riboprobes were labeled with DIG RNA Labeling Mix (Roche). Whole-mount in situ hybridization with 10.5 dpc mouse embryos was conducted with protocols previously reported.<sup>36,37</sup> Embryos were mounted in 1% agarose and imaged with optical projection tomography (OPT) described previously<sup>38</sup> by a Biotronics OPT Scanner 3001 (Biotronics). Data were processed with Biotronics proprietary software (Biotronics, MRC Technology) and images analyzed by Drishti software.

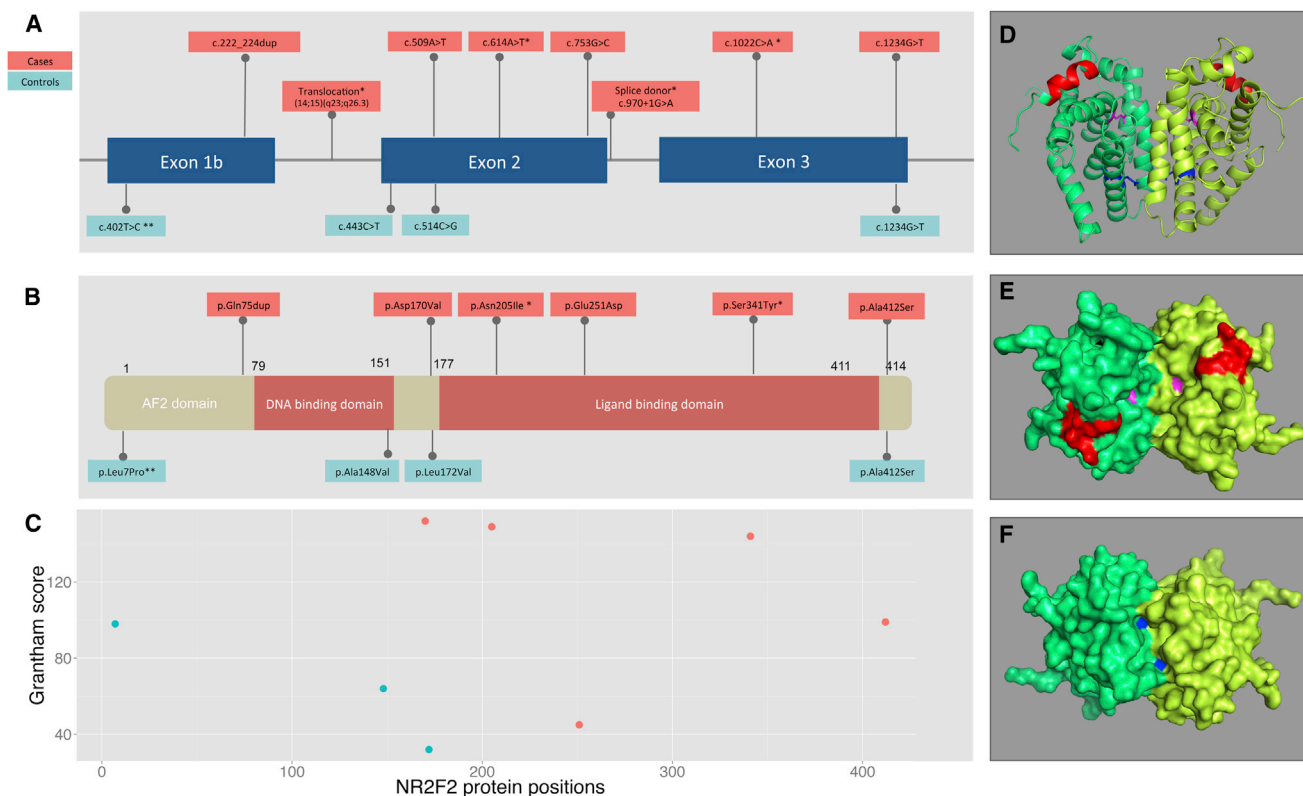
## Results

### De Novo Mutations Identified in Nonsyndromic AVSD-Affected Parent-Offspring Trios

By exome sequencing, we identified and subsequently validated 13 de novo coding mutations in the 13 trios: nine missense and four synonymous variants (Subjects and Methods; Table S2 and Figure S1). Two of the genes with missense mutations are known to be expressed in heart tissue (*ZMYND8* and *NR2F2*),<sup>18,39</sup> of which only *NR2F2* has a mouse knockout with a cardiac phenotype (Table S4). The numbers of missense de novo variants are higher than the silent variants but the burden of de novo missense variants is not statistically significant (exact binomial test, *p* = 0.69, Figure S1) compared with the expected proportion of de novo missense mutations proposed previously.<sup>28</sup>

### Burden of Rare Missense Variants Analysis

We then tested each of the nine genes identified above as harboring de novo functional mutations for a burden of rare coding mutations in all 125 exome-sequenced unrelated AVSD-affected individuals (13 affected children from the trios and 112 unrelated AVSD-affected individuals) compared to 5,194 population-matched control subjects. We found *NR2F2* to be the only gene, of the nine genes with de novo variants identified in the original trios, with a significant enrichment of rare missense variants (Fisher's exact *p* =  $7.7 \times 10^{-7}$ , odds ratio = 54.1) (see Subjects and Methods; Table S3 and Figure S2). This analysis detected four additional rare missense mutations in AVSD-affected individuals and four rare missense mutations in control subjects (Figures 1 and 2C–2F). Only one of the missense variants in affected individuals (c.1234G>T [RefSeq NM\_021005.3]; p.Ala412Ser [RefSeq NP\_066285.1]) has previously been observed in population data, in a single individual, in the 4,300 European American exomes from the NHLBI-ESP project. Using parental samples where available, we showed that in addition to the de novo mutation c.1022C>A (p.Ser341Tyr) identified initially, the variant c.614A>T (p.Asn205Ile) also arose de novo, whereas two of the other three missense variants observed in affected individuals (c.753G>C [p.Glu251Asp] and c.1234G>T [p.Ala412Ser]) were inherited from an apparently healthy parent (Figures 1A and 1B and Table 1), suggesting either incomplete



**Figure 1. Structure of *NR2F2* and the Encoded Protein**

(A) *NR2F2* has three coding exons and four transcripts (see Figure S3C). The transcript that generates the full-length protein (RefSeq NM\_021005) is shown here annotated with functional variants in cases (red) and controls (blue).

(B) Similar to other nuclear receptors, *NR2F2* has three main domains: a ligand-binding (LBD), DNA-binding (DBD), and an activation binding motif (AF2). Three mutations in cases are located in the ligand-binding domain (LBD).

Asterisk (\*) denotes de novo variant.

(C) The Grantham score for the missense mutations.

(D) Two missense variants mapped onto the partial crystal structure for the *NR2F2* ligand-binding domain.<sup>42</sup>

(E) c.753G>C (RefSeq NM\_021005.3); p.Glu251Asp (RefSeq NP\_066285.1) (purple) falls in the ligand-binding groove of the dimer, which in the repressed conformation is occupied by helix AF2 (red), and thus this variant is likely to perturb ligand binding.

(F) c.1022C>A (RefSeq NM\_021005.3); p.Ser341Tyr (RefSeq NP\_066285.1) (blue) is likely to destabilize helix A10 through steric hindrance and thus decrease the stability of *NR2F2* homodimerization (see Figure S5).

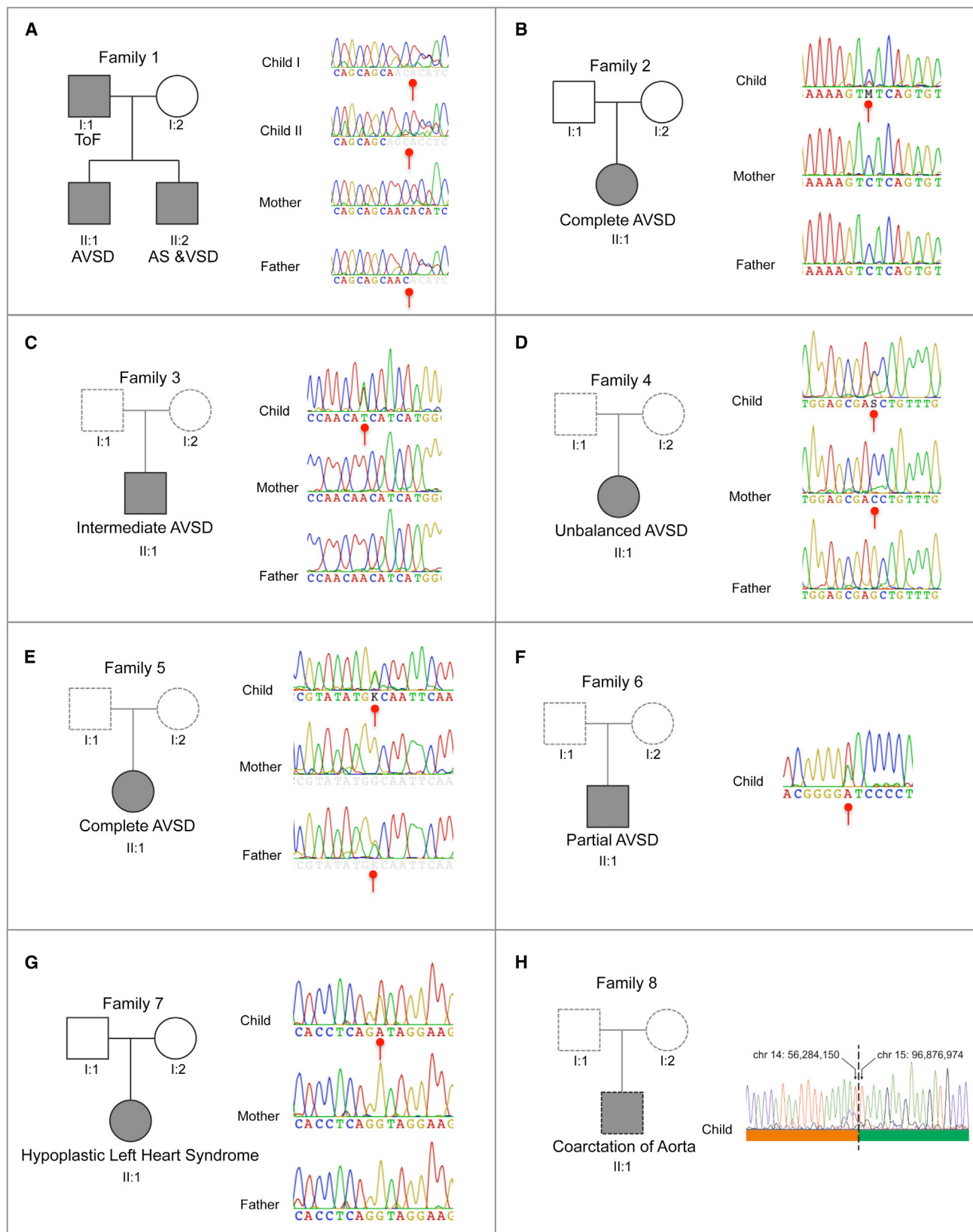
penetrance or that these are rare and benign variants. However, the high odds ratio for rare missense variants in this gene argues that it is unlikely that both of these variants inherited from unaffected parents are etiologically irrelevant. p.Ala412Ser is least likely to be disease causing because it is inherited from an unaffected parent and observed in a control individual not known to have CHDs. Moreover, the amino acid changes observed in cases appear to be more disruptive than those observed in controls, as measured by the Grantham score, but with so few variants observed in control subjects, this trend is not statistically significant (Figure 1C). We also screened the three coding exons of the major transcript of *NR2F2* in an additional 245 AVSD-affected individuals via capillary sequencing but observed no rare functional variants.

#### De Novo and Inherited *NR2F2* Mutations in Non-AVSD Congenital Heart Defect-Affected Families

There is considerable phenotypic heterogeneity in CHDs whereby the same genes can be associated with diverse

forms of CHDs in humans, e.g., *GATA4*, *NOTCH1* (MIM 190198), *NKX2-5* (MIM 600584), and *CITED2*. Almost 45% of the CHD candidate genes identified from mice knockouts have caused diverse cardiac phenotypes.<sup>3,40</sup> We therefore explored the frequency of *NR2F2* variants in other non-AVSD CHD cohorts available to us (including 293 families with exome-sequencing data). We identified three additional CHD-affected families with non-AVSD phenotypes with previously unidentified functional variants in *NR2F2*. In an individual with Tetralogy of Fallot (TOF [MIM 187500]), we detected 3 bp duplication (c.222\_224dup [p.Gln75dup]), which had been transmitted to two affected sons (one with AVSDs and the other with aortic stenosis and a ventricular septal defect) (Figure 2A). We also investigated a previously reported child with coarctation of the aorta with a de novo balanced chromosomal translocation 46,XY,t(14;15)(q23;q26.3).<sup>41</sup> By using flow-sorted derivative 14 and 15 chromosomes, we fine-mapped the translocation to the first intron of *NR2F2*. The breakpoint was predicted to truncate all





**Figure 2. Pedigree Charts and Capillary Sequencing Results of *NR2F2* Variants in Eight CHD-Affected Families**

Solid lines in pedigree charts indicate both whole-exome sequencing data and capillary sequencing are available; dashed lines indicate samples with *NR2F2* capillary sequencing data only. See Table 1 for details.

**Table 1. NR2F2 Sequence Alterations Identified in Individuals with AVSDs and Other Heart Structural Phenotypes**

Family	Subject	Sex	Phenotype	DNA Mutation <sup>a</sup>	Protein Change <sup>b</sup>	Variant Type	GERP++ <sup>c</sup>	De Novo or Inherited	Seen in Unrelated Control Subjects <sup>d</sup>
1	I:1	M	TOF	c.220_222dup	p.Gln75dup	in-frame duplication	–	ND	no
1	II:1	M	cAVSD	c.220_222dup	p.Gln75dup	in-frame duplication	–	inherited from affected father	no
1	II:2	M	AS and VSD	c.220_222dup	p.Gln75dup	in-frame duplication	–	inherited from affected father	no
2	II:1	F	cAVSD	c.1022C>A	p.Ser341Tyr	missense	5.15	de novo	no
3	II:1	M	iAVSD	c.614A>T	p.Asn205Ile	missense	5.05	de novo	no
4	II:1	F	ubAVSD	c.753G>C	p.Glu251Asp	missense	4.17	inherited from unaffected mother	no
5	II:1	F	cAVSD	c.1234G>T	p.Ala412Ser	missense	5.74	inherited from unaffected father	yes
6	II:1	M	pAVSD	c.509A>T	p.Asp170Val	missense	5.00	ND	no
7	II:1	F	HLHS	c.970+1G>A	–	splice donor	4.06	de novo	no
8	II:1	M	CoA	(14;15)(q23;q26.3)	–	balanced translocation	–	de novo	no

Abbreviations are as follows: AVSD, atrioventricular septal defect; pAVSD, partial AVSD; cAVSD, complete AVSD; ucAVSD, unbalanced complete AVSD; iAVSD, intermediate AVSD; TOF, tetralogy of Fallot; HLHS, hypoplastic left heart syndrome; AS, aortic stenosis; VSD, ventricular septal defect; CoA, coarctation of aorta; –, not applicable; ND, parent DNA was unavailable.

<sup>a</sup>Position on NR2F2 cDNA RefSeq NM\_021005.3.

<sup>b</sup>Position on NR2F2 protein product RefSeq NP\_066285.1.

<sup>c</sup>GERP++ are single-nucleotide conservation scores.

<sup>d</sup>Control subjects include 894 and 4,300 European samples from UK10K and NHLBI-ESP data sets, respectively.

annotated transcripts, thus probably generating a null allele (Figures 2H and S3). In the third family, a trio of two healthy parents of an affected child with hypoplastic left heart syndrome (HLHS [MIM 241550]), we identified a de novo splice site mutation (c.2359+1G>A [RefSeq NM\_021005.3]) that is likely to skip the third exon (Figure 2G). In summary, we identified eight CHD-affected families with different rare, functional, variants in NR2F2, four of which arose de novo, and one of which segregated with CHDs in a multiplex family (Table 1, Figures 1 and S4).

### Expression Pattern of NR2F2 in the Developing Mammalian Embryo

To explore the expression of NR2F2 in mammalian development, we used whole-mount in situ hybridization and optical projection tomography to map the pattern of *Nr2f2* mRNA expression in the developing mouse embryo (Figure 3). We observed *Nr2f2* mRNA expression in the atria of the heart, branchial arches, somites, and olfactory placode at 10.5 dpc. We also demonstrated that NR2F2 is expressed in several structures of the developing human fetal heart, including the atria, coronary vessels, and aorta (Figure 4).

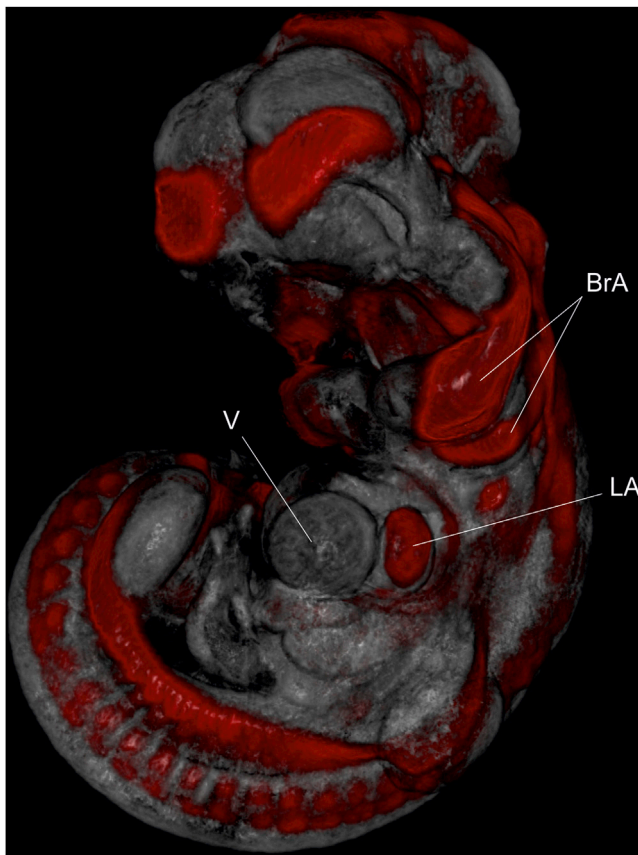
### Mapping Mutations on the Crystal Structure of the NR2F2 Ligand-Binding Domain

The missense variants seen in cases are distributed throughout NR2F2 protein, with three falling in the

ligand-binding domain (p.Asn205Ile, p.Glu251Asp, and p.Ser341Tyr), of which two can be mapped to a previously determined partial crystal structure for this domain<sup>42</sup> (Figures 1D–1F and S5). We analyzed the conformational constraints introduced on the local environment of the protein and attempted to minimize unacceptable and close contacts by using different rotamers of the mutated residue. None of the possible rotamers for the mutated residues could eliminate stereo-chemical clashes in the local environment, leading to the conclusion that these mutations could be accommodated only by a conformation change in the local fold, which in turn would disrupt dimerization (p.Ser341Tyr) or affect the ligand-binding properties of the protein (p.Asn205Ile).

### Functional Impact of NR2F2 Variants on Transcriptional Activity

Despite the availability of computational methods predicting the effect of missense variants on protein function, interpreting the significance of these mutations in human disease is notoriously difficult. We therefore sought to test the consequence of the identified NR2F2 variants in a functional assay. NR2F2 is a transcriptional regulator, with both activating and repressive effects on target gene expression.<sup>43</sup> A number of NR2F2-responsive genomic elements have been identified that, when placed upstream of a reporter gene, can quantitate transcriptional regulator function of NR2F2 variants.<sup>30,31,42</sup> By using the most widely employed element, the promoter region of



**Figure 3. *Nr2f2* Expression in the Developing Mouse Embryo**  
*Nr2f2* mRNA expression (red) is detected in the atria of the heart, branchial arches, somites, and olfactory placode at 10.5 dpc by whole-mount in situ hybridization.

*NGFI-A*,<sup>30</sup> to drive a luciferase reporter in HEK293 cells, we compared its level of activation by wild-type NR2F2 with that of the case-derived variants. We observed robust luciferase activation by wild-type NR2F2 and equivalent levels of activity from variants p.Asp170Val and p.Ala412Ser. However, two variants (p.Glu251Asp and p.Ser341Tyr) show a significantly lower activity in this assay (20%–24% reduction,  $p < 0.01$ ), whereas variants p.Gln70dup and p.Asn205Ile have an increased activity (13%–15% increase,  $p < 0.03$ ) (Figure 5).

Because the function of nuclear receptors involves a complex interaction with other transcriptional coregulators, we hypothesized that the consequence of NR2F2 variants might be promoter context dependent. We therefore performed the luciferase assay on an alternative promoter fragment from the *APOB* that has previously been shown to be bound by NR2F2 and used for structure-function studies.<sup>31</sup> In agreement with our prediction, the activities of the variants on the *APOB* promoter in HEK293 cells were significantly different from those using the *NGFI-A* promoter (Figure 5). Variants p.Asp170Val, p.Asn205Ile, p.Glu251Asp, and p.Ser341Tyr all show strong reductions in transcriptional activity compared to wild-type NR2F2 (26%–52% reduction,  $p < 0.001$ ), and

p.Ala412Ser now has significantly higher activity (12.9% increase,  $p = 0.006$ ). Strikingly, variant p.Asn205Ile reduces the activity of NR2F2 on the *APOB* promoter while increasing it on the *NGFI-A* promoter (down 26% versus up 15%,  $p = 0.0003$ ).

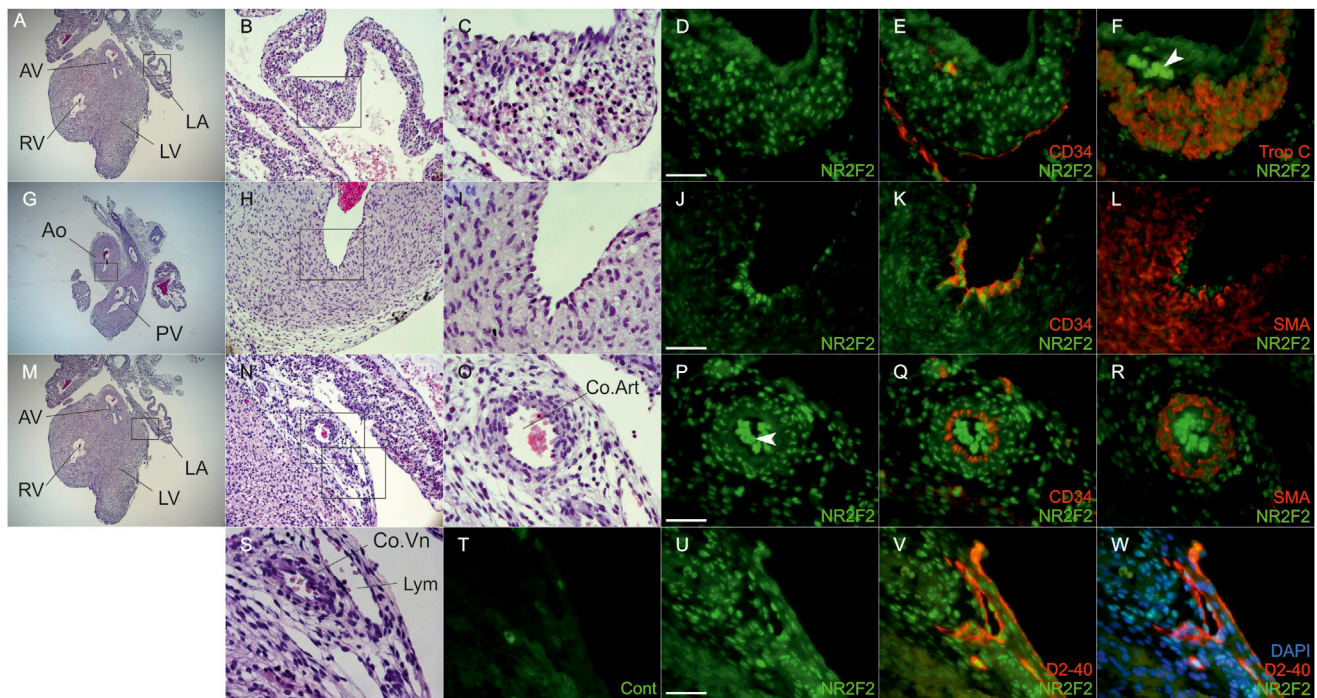
Finally, we asked whether the known repressor function of NR2F2 was affected by any of the identified variants. In HEPG2 cells, NR2F2 represses the *APOB* promoter, whose basal activity is high in this cellular context.<sup>31</sup> When we performed the luciferase assay in HEPG2 cells, we found that the expected repressive activity of NR2F2 is not affected by any of the variants observed in individuals with CHDs.

## Discussion

By using exome data from a combined study design of affected parent-offspring trios and index cases, we were able to identify 2 out of 370 affected individuals (125 with exome, 245 with capillary sequencing) with de novo missense variants in *NR2F2* (an observation which, given the mutation rate of *NR2F2*, has a Poisson  $p$  value of  $p = 4.8 \times 10^{-5}$ ) and another 3 affected individuals with rare missense variants that represent a very significant enrichment compared to 5,194 control subjects ( $p = 7.7 \times 10^{-7}$ , Fisher exact test of case and control subjects). Moreover, we identified three additional CHD families with other variants in *NR2F2* including a de novo balanced chromosomal translocation in an individual with coarctation of the aorta, a 3 bp duplication that cosegregated in a multiplex family of a father with tetralogy of Fallot and two sons (one with AVSDs and the other with aortic stenosis and ventricular septal defect), and a de novo substitution disrupting a splice donor site in a hypoplastic left heart syndrome individual. Thus we observed three functional de novo substitutions in *NR2F2* across all 663 CHD-affected individuals.

NR2F2 belongs to a small family of the steroid/thyroid hormone receptor nuclear superfamily of transcription factors that includes two related but distinct genes: *NR2F1* (or *COUP-TF1*) and *NR2F2* (or *COUP-TFII*). Both genes are involved in many cellular and developmental processes. Whereas *NR2F1* is mainly involved in neural development, *NR2F2* is expressed and involved in the organogenesis of the stomach, limbs, skeletal muscles, and heart.<sup>43</sup> The *Nr2f2* mouse null model leads to embryonic lethality with severe hemorrhage and failure of the atria and sinus venosus to develop past the primitive tube stage.<sup>17</sup> A recent hypomorphic *Nr2f2* mouse mutant exhibits a spectrum of cardiac defects including left atrial hypoplasia, ventricular hypoplasia, and atrioventricular septal defects resulting from the disruption of endocardial cushion development in a dosage-sensitive fashion. The latter is partially driven by defective endothelial-mesenchymal transformation and hypocellularity of the atrioventricular canal.<sup>18</sup> These mouse models and our





**Figure 4. NR2F2 Localization in the Developing Human Heart**

Immunofluorescent analysis of NR2F2 in fixed human fetal heart via anti-NR2F2 (D–F, J–L, P–R, U–W, green) and colabelling (red) with CD34 (E, K, Q), troponin C (F), SMA (L, R), D2-40, and DAPI (W, blue). Haematoxylin and eosin staining (A–C, G–I, M–O, S). An additional autofluorescence artifact was detected (arrowhead F, P–R) from hemaglobin within erythrocytes. Negative control for NR2F2 shown in (T). The fields shown in (C), (I), (O), and (S) are from haematoxylin and eosin-stained serial sections adjacent to the fields shown in (D)–(F), (J)–(L), (P)–(R), and (T)–(W), respectively. The boxed areas in haematoxylin and eosin-stained fields represent the area shown in higher magnification in the adjacent field to the right.

Abbreviations are as follows: LA, left atrium; Ao, aorta; Co.Art, coronary artery; Co.Vn, coronary vein; Lym, lymphatic vessel. Scale bars represent 100  $\mu$ m.

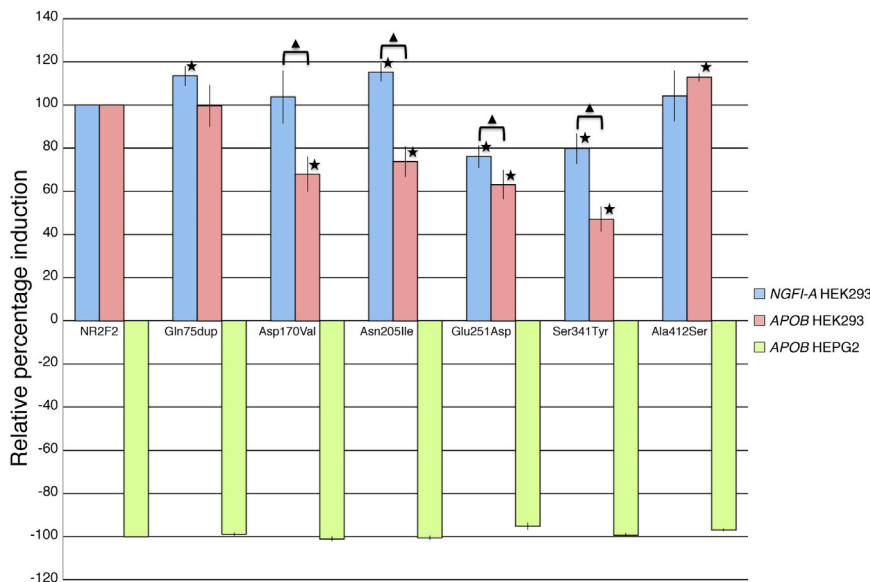
expression data strongly support a role for *NR2F2* in several different cardiac developmental processes, including endocardial cushion development, and specifically that cardiac development is likely to be sensitive to the dosage of functional *NR2F2*.

In humans, previous case reports of 15q terminal deletions that, in addition to *NR2F2*, encompass several genes and regulatory regions, have suggested *NR2F2* as a candidate gene for both CHDs<sup>44</sup> and congenital diaphragmatic hernia (CDH).<sup>45,46</sup> A role in CHDs was proposed on the basis that *NR2F2* falls within a critical interval deleted in the subset of individuals that have CHDs in addition to the syndromic features typically associated with these deletions.<sup>44</sup> A role for *NR2F2* in CDH is supported by the tissue-specific ablation of *Nr2f2* in mice, which results in malformation of the diaphragm.<sup>47</sup> However, *NR2F2* resequencing studies failed to reveal any pathological mutations in CDH cases,<sup>48,49</sup> which led the authors to hypothesize that variants in the noncoding region surrounding *NR2F2* may contribute to the development of isolated CHDs.<sup>50</sup> Additionally, the conditional mouse model cannot distinguish between the importance of coding and noncoding sequence, because excision of the allele removes the entire coding interval, including 4.4 kb of noncoding sequence. Consistent with their hypoth-

esis of a role for noncoding variation in *CDH*, none of the individuals in our study with *NR2F2* missense or loss-of-function sequence variants manifested CDH.

To assess the possibility of a correlation between the severity of the *NR2F2* variant and the resulting CHD phenotype, we collated the cardiac phenotypes associated with 11 published whole-gene deletions of *NR2F2* and combined these with the phenotypes of individuals with *NR2F2* described in our study. We observed a highly significant phenotypic difference between the 13 individuals with loss-of-function variants and the 8 individuals with missense variants: 9 of the individuals with loss-of-function variants had left ventricular outflow tract obstruction (LVOTO) but none had AVSDs, whereas 6 out of 8 individuals with missense variants had AVSDs and only one had LVOTO ( $p = 0.0009$ , Fisher's exact test). In addition, 8 out of 11 individuals with *NR2F2* deletions also had either an atrial septal defect or a ventricular septal defect (Tables S5–S7 and Figure S6). This emerging genotype-phenotype correlation in humans parallels the mouse studies that showed that complete loss-of-function resulted in more complex cardiac phenotypes, whereas hypomorphic variants resulted in more specific deficits in the development of the endocardial cushions.





**Figure 5. NR2F2 Variants in AVSD-Affected Probands Affect Transcriptional Activity**

An NR2F2-responsive luciferase reporter driven by the *NGFI-A* or *APOB* upstream region was cotransfected with wild-type NR2F2, or identified coding variants (p.Gln75dup, p.Asp170Val, p.Asn205Ile, p.Glu251Asp, p.Ser341Tyr, and p.Alc412Ser) into HEK293T (*NGFI-A* and *APOB*) and HEPG2 (*APOB*) cells (see [Subjects and Methods](#) for details). Bar chart values are activity relative to wild-type NR2F2 (mean percentage  $\pm$  SD). Repression of the *APOB* promoter in HEPG2 cells is shown as negative values to illustrate the direction of change from negative control. In HEK293 cells, all variants show significant difference from wild-type on one or both promoters. The p.Asn205Ile variant shows the reverse direction of change depending on which promoter was used. In HEPG2 cells, all variants retain wild-type levels of repressive activity. Asterisk indicates significant change from wild-type activity. Triangle indicates significant difference between promoters.

Our in vitro experimental data indicate that all six *NR2F2* missense variants we identified have a measurable impact on the transcriptional activator function of NR2F2 in at least one of two assays. In contrast, the repressor function of NR2F2 appears intact. That individual mutations have promoter-specific effects on gene function probably reflects the complexity of the protein-protein interactions NR2F2 engages in depending on tissue, stage, and genomic context. The diversity of both human and mouse cardiac phenotypes associated with NR2F2 variation suggests that NR2F2 plays a critical role in several temporally and spatially distinct cardiac developmental processes. Moreover, the human and mouse genetic data suggest that the development of endocardial cushions is more sensitive to dosage of functional NR2F2 than other cardiac developmental processes. Indeed, the nonsyndromic CHD presentation of individuals with *NR2F2* variants, despite its broader embryonic expression, suggest, more generally, that the heart is more sensitive to NR2F2 dosage than other organs. It will be necessary to identify the etiologically relevant NR2F2-target promoter(s) and cell type(s) to understand the specific molecular mechanisms by which these variants perturb cardiac developmental networks.

To place our observations of *NR2F2* in the context of other genes harboring variants that cause CHDs, it is important to distinguish between genes with compelling genetic evidence for a role in CHDs versus those with much weaker evidence, often in the form of small numbers of rare missense variants of unknown inheritance observed in a small fraction of individuals with CHDs. The genes with the most robust associations to CHDs are typically seen in the context of multisystem syndromes that include

CHDs as a component phenotype (e.g., *TBX5*, *GATA6*, *EVC2*). There are relatively few genes robustly associated with nonsyndromic CHDs and none clearly associated with nonsyndromic AVSDs in particular. Further, these genes are often associated with a wide range of CHD phenotypes, albeit with an appreciable bias toward some CHD subtypes. Examples include a bias toward right-sided heart defects with pathogenic variants in *JAG1*,<sup>51</sup> a bias toward transposition of the great arteries and conotruncal heart defects for *NODAL*, and a bias toward atrial septal defects for *GATA4*.<sup>52</sup> This reflects the differential sensitivity of different genes to temporally and spatially distinct cardiac developmental processes. Our genotype-phenotype observations with *NR2F2* missense and loss-of-function variants fit with this view, with a bias toward AVSDs and LVOTO, respectively.

Most of the robustly CHD-associated genes, especially those that act in a dominant fashion, encode proteins involved in extracellular or intracellular signaling (e.g., *JAG1*, *NOTCH1*, *PTPN11*), developmental transcription factors (e.g., *TBX1*, *TBX5*, *NKX2-5*, *GATA4*, *GATA6*, *ZIC3*), chromatin remodeling factors (e.g., *MLL2*, *CREBBP*, *CHD7*), or structural proteins (e.g., *MYH6*, *ELN*), and most of these dominant genes operate through variants causing complete or partial loss-of-function that result in altered dosage of functional protein. As a developmental transcriptional factor, *NR2F2* probably operates by a similar, dosage-sensitive mechanism to other known CHD-associated genes.

In addition to the direct role of *NR2F2* mutations in causing CHDs, given its dosage sensitivity, *NR2F2* may potentially also act as an environmentally responsive factor by mediating the effect of known nongenetic CHD risk factors such as high glucose<sup>53</sup> and retinoic acid

levels.<sup>54</sup> Insulin and glucose levels are known to negatively control *NR2F2* expression via the forkhead box protein O1 (Foxo1) pathway in hepatocyte and pancreatic cells.<sup>55</sup> Furthermore, *NR2F2* has been shown to play a critical role in retinoic acid signaling during development.<sup>56</sup> Further investigation is needed to determine how glucose and retinoic acid levels may alter *NR2F2* expression in the developing heart.

In summary, our findings add *NR2F2* to the short list of dosage-sensitive regulators such as *TBX5*, *TBX1*, *NKX2-5*, and *GATA4* that have been shown, when mutated, to interfere with normal heart development and that lead to CHDs in both mice and humans. By virtue of their dosage sensitivity, these master regulators potentially play a key role in integrating genetic and environmental risk factors for abnormal cardiac development.

### Supplemental Data

Supplemental Data include six figures and seven tables and can be found with this article online at <http://www.cell.com/ajhg/>.

### Consortia

Members of the UK10K Rare Diseases Cohorts Working Group are Matthew Hurles (cochair), David R. FitzPatrick (cochair), Saeed Al-Turki, Carl Anderson, Inês Barroso, Philip Beales, Jamie Bentham, Shoumo Bhattacharya, Keren Carss, Krishna Chatterjee, Sebhatten Cirak, Catherine Cosgrove, Allan Daly, Jamie Floyd, Chris Franklin, Marta Futema, Steve Humphries, Shane McCarthy, Hannah Mitchison, Francesco Muntoni, Alexandros Onoufriadis, Victoria Parker, Felicity Payne, Vincent Plagnol, Lucy Raymond, David Savage, Peter Scambler, Miriam Schmidts, Robert Semple, Eva Serra, Jim Stalker, Margriet van Kogelenberg, Parthiban Vijayarangakannan, Klaudia Walter, and Gretta Wood.

### Acknowledgments

The authors would like to thank the individuals and their families for their support and participation and Don Conrad for the DeNovoGear software. This study was supported by funding from the Wellcome Trust (grant number WT098051), an MRC training fellowship (to C.L.M.), Little Hearts Matter and the Competence Network for Congenital Heart Defects/National Register for Congenital Heart Defects (Germany) funded by the Federal Ministry of Education and Research (BMBF), Support Code FKZ 01GI0601, the DZHK (German Centre for Cardiovascular Research), and a Heart and Stroke Foundation of Canada research fellowship (to A.K.M.). This study makes use of data generated by the UK10K Consortium, derived from samples from TwinsUK and ALSPAC. A full list of the investigators who contributed to the generation of the data is available from <http://www.UK10K.org>. Funding for UK10K was provided by the Wellcome Trust under award WT091310. The authors would like to thank the NHLBI GO Exome Sequencing Project and its ongoing studies, which produced and provided exome variant calls for comparison: the Lung GO Sequencing Project (HL-102923), the WHI Sequencing Project (HL-102924), the Broad GO Sequencing Project (HL-102925), the Seattle GO Sequencing Project (HL-102926), and the Heart GO Sequencing Project (HL-103010).

Received: October 24, 2013

Accepted: March 12, 2014

Published: April 3, 2014

### Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, <http://browser.1000genomes.org>

Drishti software, <http://anusf.anu.edu.au/Vizlab/drishti/index.shtml>

European Genome-phenome Archive (EGA), <https://www.ebi.ac.uk/ega>

Online Mendelian Inheritance in Man (OMIM), <http://www.omim.org/>

RefSeq, <http://www.ncbi.nlm.nih.gov/RefSeq>

### Accession Numbers

The EGA accession numbers for exome-sequencing data reported in this paper are EGAS00001000125, EGAS00001000317, and EGAS00001000185.

### References

1. Blue, G.M., Kirk, E.P., Sholler, G.F., Harvey, R.P., and Winlaw, D.S. (2012). Congenital heart disease: current knowledge about causes and inheritance. *Med. J. Aust.* 197, 155–159.
2. Pierpont, M.E., Basson, C.T., Benson, D.W., Jr., Gelb, B.D., Giglia, T.M., Goldmuntz, E., McGee, G., Sable, C.A., Srivastava, D., and Webb, C.L.; American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young (2007). Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* 115, 3015–3038.
3. Bentham, J., and Bhattacharya, S. (2008). Genetic mechanisms controlling cardiovascular development. *Ann. N. Y. Acad. Sci.* 1123, 10–19.
4. Reller, M.D., Strickland, M.J., Riehle-Colarusso, T., Mahle, W.T., and Correa, A. (2008). Prevalence of congenital heart defects in metropolitan Atlanta, 1998–2005. *J. Pediatr.* 153, 807–813.
5. Hoffman, J.I. (1995). Incidence of congenital heart disease: I. Postnatal incidence. *Pediatr. Cardiol.* 16, 103–113.
6. Allan, L.D., Sharland, G.K., Milburn, A., Lockhart, S.M., Groves, A.M., Anderson, R.H., Cook, A.C., and Fagg, N.L. (1994). Prospective diagnosis of 1,006 consecutive cases of congenital heart disease in the fetus. *J. Am. Coll. Cardiol.* 23, 1452–1458.
7. Freeman, S.B., Bean, L.H., Allen, E.G., Tinker, S.W., Locke, A.E., Druschel, C., Hobbs, C.A., Romitti, P.A., Royle, M.H., Torfs, C.P., et al. (2008). Ethnicity, sex, and the incidence of congenital heart defects: a report from the National Down Syndrome Project. *Genet. Med.* 10, 173–180.
8. Ferencz, C., Boughman, J.A., Neill, C.A., Brenner, J.I., and Perry, L.W.; Baltimore-Washington Infant Study Group (1989). Congenital cardiovascular malformations: questions on inheritance. *J. Am. Coll. Cardiol.* 14, 756–763.

9. Ackerman, C., Locke, A.E., Feingold, E., Reshey, B., Espana, K., Thusberg, J., Mooney, S., Bean, L.J., Dooley, K.J., Cua, C.L., et al. (2012). An excess of deleterious variants in VEGF-A pathway genes in Down-syndrome-associated atrioventricular septal defects. *Am. J. Hum. Genet.* 91, 646–659.
10. Robinson, S.W., Morris, C.D., Goldmuntz, E., Reller, M.D., Jones, M.A., Steiner, R.D., and Maslen, C.L. (2003). Missense mutations in CRELD1 are associated with cardiac atrioventricular septal defects. *Am. J. Hum. Genet.* 72, 1047–1052.
11. Zatyka, M., Priestley, M., Ladusans, E.J., Fryer, A.E., Mason, J., Latif, F., and Maher, E.R. (2005). Analysis of CRELD1 as a candidate 3p25 atrioventricular septal defect locus (AVSD2). *Clin. Genet.* 67, 526–528.
12. Karkera, J.D., Lee, J.S., Roessler, E., Banerjee-Basu, S., Ouspenskaya, M.V., Mez, J., Goldmuntz, E., Bowers, P., Towbin, J., Belmont, J.W., et al. (2007). Loss-of-function mutations in growth differentiation factor-1 (GDF1) are associated with congenital heart defects in humans. *Am. J. Hum. Genet.* 81, 987–994.
13. Basson, C.T., Huang, T., Lin, R.C., Bachinsky, D.R., Weremowicz, S., Vaglio, A., Bruzzone, R., Quadrelli, R., Lerone, M., Romeo, G., et al. (1999). Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. *Proc. Natl. Acad. Sci. USA* 96, 2919–2924.
14. Rajagopal, S.K., Ma, Q., Obler, D., Shen, J., Manichaikul, A., Tomita-Mitchell, A., Boardman, K., Briggs, C., Garg, V., Srivastava, D., et al. (2007). Spectrum of heart disease associated with murine and human GATA4 mutation. *J. Mol. Cell. Cardiol.* 43, 677–685.
15. Garg, V., Kathiriyi, I.S., Barnes, R., Schluterman, M.K., King, I.N., Butler, C.A., Rothrock, C.R., Eapen, R.S., Hirayama-Yamada, K., Joo, K., et al. (2003). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* 424, 443–447.
16. 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. (2013). De novo mutations in histone-modifying genes in congenital heart disease. *Nature* 498, 220–223.
17. Pereira, F.A., Qiu, Y., Zhou, G., Tsai, M.J., and Tsai, S.Y. (1999). The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev.* 13, 1037–1049.
18. Lin, F.J., You, L.R., Yu, C.T., Hsu, W.H., Tsai, M.J., and Tsai, S.Y. (2012). Endocardial cushion morphogenesis and coronary vessel development require chicken ovalbumin upstream promoter-transcription factor II. *Arterioscler. Thromb. Vasc. Biol.* 32, e135–e146.
19. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760.
20. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
21. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303.
22. Albers, C.A., Lunter, G., MacArthur, D.G., McVean, G., Ouwehand, W.H., and Durbin, R. (2011). Dindel: accurate indel calls from short-read data. *Genome Res.* 21, 961–973.
23. McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* 26, 2069–2070.
24. Cooper, G.M., Stone, E.A., Asimenos, G., Green, E.D., Batzoglu, S., and Sidow, A.; NISC Comparative Sequencing Program (2005). Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res.* 15, 901–913.
25. Conrad, D.F., Keebler, J.E., DePristo, M.A., Lindsay, S.J., Zhang, Y., Casals, F., Idaghdour, Y., Hartl, C.L., Torroja, C., Garimella, K.V., et al.; 1000 Genomes Project (2011). Variation in genome-wide mutation rates within and between human families. *Nat. Genet.* 43, 712–714.
26. Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C., and Weir, B.S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28, 3326–3328.
27. International HapMap Consortium (2003). The International HapMap Project. *Nature* 426, 789–796.
28. Kryukov, G.V., Pennacchio, L.A., and Sunyaev, S.R. (2007). Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *Am. J. Hum. Genet.* 80, 727–739.
29. Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* 22, 1567–1572.
30. Pipaón, C., Tsai, S.Y., and Tsai, M.J. (1999). COUP-TF upregulates NGFI-A gene expression through an Sp1 binding site. *Mol. Cell. Biol.* 19, 2734–2745.
31. Achatz, G., Hölzl, B., Speckmayer, R., Hauser, C., Sandhofer, F., and Paulweber, B. (1997). Functional domains of the human orphan receptor ARP-1/COUP-TFII involved in active repression and transrepression. *Mol. Cell. Biol.* 17, 4914–4932.
32. Hearn, T., Renforth, G.L., Spalluto, C., Hanley, N.A., Piper, K., Brickwood, S., White, C., Connolly, V., Taylor, J.F., Russell-Eggitt, I., et al. (2002). Mutation of ALMS1, a large gene with a tandem repeat encoding 47 amino acids, causes Alström syndrome. *Nat. Genet.* 31, 79–83.
33. Piper, K., Brickwood, S., Turnpenny, L.W., Cameron, I.T., Ball, S.G., Wilson, D.I., and Hanley, N.A. (2004). Beta cell differentiation during early human pancreas development. *J. Endocrinol.* 181, 11–23.
34. Hearn, T., Spalluto, C., Phillips, V.J., Renforth, G.L., Copin, N., Hanley, N.A., and Wilson, D.I. (2005). Subcellular localization of ALMS1 supports involvement of centrosome and basal body dysfunction in the pathogenesis of obesity, insulin resistance, and type 2 diabetes. *Diabetes* 54, 1581–1587.
35. Gill, H.K., Parsons, S.R., Spalluto, C., Davies, A.F., Knorz, V.J., Burlinson, C.E., Ng, B.L., Carter, N.P., Ogilvie, C.M., Wilson, D.I., and Roberts, R.G. (2009). Separation of the PROX1 gene from upstream conserved elements in a complex inversion/translocation patient with hypoplastic left heart. *Eur. J. Hum. Genet.* 17, 1423–1431.
36. Harewood, L., Liu, M., Keeling, J., Howatson, A., Whiteford, M., Branney, P., Evans, M., Fantes, J., and Fitzpatrick, D.R. (2010). Bilateral renal agenesis/hypoplasia/dysplasia (BRAHD): postmortem analysis of 45 cases with breakpoint mapping of two de novo translocations. *PLoS ONE* 5, e12375.
37. Rainger, J., van Beusekom, E., Ramsay, J.K., McKie, L., Al-Gazali, L., Pallotta, R., Saponari, A., Branney, P., Fisher,



- M., Morrison, H., et al. (2011). Loss of the BMP antagonist, SMOC-1, causes Ophthalmo-acromelic (Waardenburg Anophthalmia) syndrome in humans and mice. *PLoS Genet.* 7, e1002114.
38. Sharpe, J., Ahlgren, U., Perry, P., Hill, B., Ross, A., Hecksher-Sørensen, J., Baldock, R., and Davidson, D. (2002). Optical projection tomography as a tool for 3D microscopy and gene expression studies. *Science* 296, 541–545.
39. Fossey, S.C., Kuroda, S., Price, J.A., Pendleton, J.K., Freedman, B.I., and Bowden, D.W. (2000). Identification and characterization of PRKCBP1, a candidate RACK-like protein. *Mamm. Genome* 11, 919–925.
40. Winston, J.B., Erlich, J.M., Green, C.A., Aluko, A., Kaiser, K.A., Takematsu, M., Barlow, R.S., Sureka, A.O., LaPage, M.J., Janss, L.L., and Jay, P.Y. (2010). Heterogeneity of genetic modifiers ensures normal cardiac development. *Circulation* 121, 1313–1321.
41. Baptista, J., Mercer, C., Prigmore, E., Gribble, S.M., Carter, N.P., Maloney, V., Thomas, N.S., Jacobs, P.A., and Crolla, J.A. (2008). Breakpoint mapping and array CGH in translocations: comparison of a phenotypically normal and an abnormal cohort. *Am. J. Hum. Genet.* 82, 927–936.
42. Kruse, S.W., Suino-Powell, K., Zhou, X.E., Kretschman, J.E., Reynolds, R., Vornrhein, C., Xu, Y., Wang, L., Tsai, S.Y., Tsai, M.J., and Xu, H.E. (2008). Identification of COUP-TFII orphan nuclear receptor as a retinoic acid-activated receptor. *PLoS Biol.* 6, e227.
43. Lin, F.J., Qin, J., Tang, K., Tsai, S.Y., and Tsai, M.J. (2011). Coup d'état: an orphan takes control. *Endocr. Rev.* 32, 404–421.
44. Nakamura, E., Makita, Y., Okamoto, T., Nagaya, K., Hayashi, T., Sugimoto, M., Manabe, H., Taketazu, G., Kajino, H., and Fujieda, K. (2011). 5.78 Mb terminal deletion of chromosome 15q in a girl, evaluation of NR2F2 as candidate gene for congenital heart defects. *Eur. J. Med. Genet.* 54, 354–356.
45. Brady, P.D., DeKoninck, P., Fryns, J.P., Devriendt, K., Deprest, J.A., and Vermeesch, J.R. (2013). Identification of dosage-sensitive genes in fetuses referred with severe isolated congenital diaphragmatic hernia. *Prenat. Diagn.* 33, 1283–1292.
46. Klaassens, M., van Dooren, M., Eussen, H.J., Douben, H., den Dekker, A.T., Lee, C., Donahoe, P.K., Galjaard, R.J., Goemaere, N., de Krijger, R.R., et al. (2005). Congenital diaphragmatic hernia and chromosome 15q26: determination of a candidate region by use of fluorescent in situ hybridization and array-based comparative genomic hybridization. *Am. J. Hum. Genet.* 76, 877–882.
47. You, L.R., Takamoto, N., Yu, C.T., Tanaka, T., Kodama, T., Demayo, F.J., Tsai, S.Y., and Tsai, M.J. (2005). Mouse lacking COUP-TFII as an animal model of Bochdalek-type congenital diaphragmatic hernia. *Proc. Natl. Acad. Sci. USA* 102, 16351–16356.
48. Scott, D.A., Klaassens, M., Holder, A.M., Lally, K.P., Fernandes, C.J., Galjaard, R.J., Tibboel, D., de Klein, A., and Lee, B. (2007). Genome-wide oligonucleotide-based array comparative genome hybridization analysis of non-isolated congenital diaphragmatic hernia. *Hum. Mol. Genet.* 16, 424–430.
49. Slavotinek, A.M., Moshrefi, A., Davis, R., Leeth, E., Schaeffer, G.B., Burchard, G.E., Shaw, G.M., James, B., Ptacek, L., and Pennacchio, L.A. (2006). Array comparative genomic hybridization in patients with congenital diaphragmatic hernia: mapping of four CDH-critical regions and sequencing of candidate genes at 15q26.1-15q26.2. *Eur. J. Hum. Genet.* 14, 999–1008.
50. Arrington, C.B., Bleyl, S.B., Matsunami, N., Bowles, N.E., Leppert, T.I., Demarest, B.L., Osborne, K., Yoder, B.A., Byrne, J.L., Schiffman, J.D., et al. (2012). A family-based paradigm to identify candidate chromosomal regions for isolated congenital diaphragmatic hernia. *Am. J. Med. Genet. A.* 158A, 3137–3147.
51. McElhinney, D.B., Krantz, I.D., Bason, L., Piccoli, D.A., Emerick, K.M., Spinner, N.B., and Goldmuntz, E. (2002). Analysis of cardiovascular phenotype and genotype-phenotype correlation in individuals with a JAG1 mutation and/or Alagille syndrome. *Circulation* 106, 2567–2574.
52. Barriot, R., Breckpot, J., Thienpont, B., Brohée, S., Van Vooren, S., Coessens, B., Tranchevent, L.C., Van Loo, P., Gewillig, M., Devriendt, K., and Moreau, Y. (2010). Collaboratively charting the gene-to-phenotype network of human congenital heart defects. *Genome Med* 2, 16.
53. 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., and Reece, E.A. (2008). Diabetes mellitus and birth defects. *Am. J. Obstet. Gynecol.* 199, e1–e9.
54. Botto, L.D., Loffredo, C., Scanlon, K.S., Ferencz, C., Khoury, M.J., David Wilson, P., and Correa, A. (2001). Vitamin A and cardiac outflow tract defects. *Epidemiology* 12, 491–496.
55. Perilhou, A., Tourrel-Cuzin, C., Kharroubi, I., Henique, C., Fauveau, V., Kitamura, T., Magnan, C., Postic, C., Prip-Buus, C., and Vasseur-Cognet, M. (2008). The transcription factor COUP-TFII is negatively regulated by insulin and glucose via Foxo1- and ChREBP-controlled pathways. *Mol. Cell. Biol.* 28, 6568–6579.
56. Vilhais-Neto, G.C., Maruhashi, M., Smith, K.T., Vasseur-Cognet, M., Peterson, A.S., Workman, J.L., and Pourquié, O. (2010). Rere controls retinoic acid signalling and somite bilateral symmetry. *Nature* 463, 953–957.