

Copy-Number Variation of the Glucose Transporter Gene *SLC2A3* and Congenital Heart Defects in the 22q11.2 Deletion Syndrome

Elisabeth E. Mlynarski,¹ Molly B. Sheridan,¹ Michael Xie,² Tingwei Guo,³ Silvia E. Racedo,³ Donna M. McDonald-McGinn,^{1,4} Xiaowu Gai,⁵ Eva W.C. Chow,⁶ Jacob Vorstman,⁷ Ann Swillen,⁸ Koen Devriendt,⁸ Jeroen Breckpot,⁸ Maria Cristina Digilio,⁹ Bruno Marino,¹⁰ Bruno Dallapiccola,⁹ Nicole Philip,¹¹ Tony J. Simon,¹² Amy E. Roberts,¹³ Małgorzata Piotrowicz,¹⁴ Carrie E. Bearden,¹⁵ Stephan Eliez,¹⁶ Doron Gothelf,¹⁷ Karlene Coleman,¹⁸ Wendy R. Kates,¹⁹ Marcella Devoto,^{1,4,20,21} Elaine Zackai,^{1,4} Damian Heine-Suñer,²² Tamim H. Shaikh,²³ Anne S. Bassett,⁶ Elizabeth Goldmuntz,^{4,24} Bernice E. Morrow,³ Beverly S. Emanuel,^{1,4,*} and the International Chromosome 22q11.2 Consortium

The 22q11.2 deletion syndrome (22q11DS; velocardiofacial/DiGeorge syndrome; VCFS/DGS) is the most common microdeletion syndrome and the phenotypic presentation is highly variable. Approximately 65% of individuals with 22q11DS have a congenital heart defect (CHD), mostly of the conotruncal type, and/or an aortic arch defect. The etiology of this phenotypic variability is not currently known. We hypothesized that copy-number variants (CNVs) outside the 22q11.2 deleted region might increase the risk of being born with a CHD in this sensitized population. Genotyping with Affymetrix SNP Array 6.0 was performed on two groups of subjects with 22q11DS separated by time of ascertainment and processing. CNV analysis was completed on a total of 949 subjects (cohort 1, n = 562; cohort 2, n = 387), 603 with CHDs (cohort 1, n = 363; cohort 2, n = 240) and 346 with normal cardiac anatomy (cohort 1, n = 199; cohort 2, n = 147). Our analysis revealed that a duplication of *SLC2A3* was the most frequent CNV identified in the first cohort. It was present in 18 subjects with CHDs and 1 subject without ($p = 3.12 \times 10^{-3}$, two-tailed Fisher's exact test). In the second cohort, the *SLC2A3* duplication was also significantly enriched in subjects with CHDs ($p = 3.30 \times 10^{-2}$, two-tailed Fisher's exact test). The *SLC2A3* duplication was the most frequent CNV detected and the only significant finding in our combined analysis ($p = 2.68 \times 10^{-4}$, two-tailed Fisher's exact test), indicating that the *SLC2A3* duplication might serve as a genetic modifier of CHDs and/or aortic arch anomalies in individuals with 22q11DS.

Introduction

Congenital heart defects (CHDs) are the leading cause of birth defect-related deaths in newborns¹ and are estimated to occur in 0.5% to 1% of live births.² They can develop as an isolated abnormality or in conjunction with a syndromic condition. Approximately one third of CHDs result from malformations of the cardiac outflow tract and are collectively referred to as conotruncal heart defects (CTDs), examples of which include tetralogy of Fallot (TOF), pulmonary atresia with ventricular septal defect

(VSD), truncus arteriosus, and interrupted aortic arch type B.³ Both genetic and environmental etiologies of CTDs have been described.^{4–6} With respect to genetic etiologies, CTDs have been identified in individuals with single gene disorders, gain or loss of entire chromosomes, and submicroscopic unbalanced structural rearrangements or copy-number variants (CNVs). One of the most common CNVs associated with CTDs is the 22q11.2 deletion.^{7,8}

The 22q11DS (velocardiofacial syndrome; DiGeorge syndrome, VCFS/DGS [MIM: 192430, 188400]) is the most common microdeletion syndrome, affecting

¹Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; ²Center for Biomedical Informatics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; ³Department of Genetics, Albert Einstein College of Medicine, Bronx, NY 10461, USA; ⁴Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; ⁵Ocular Genomics Institute, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary and Harvard Medical School, Boston, MA 02114, USA; ⁶Clinical Genetics Research Program, Centre for Addiction and Mental Health and Department of Psychiatry, University of Toronto, Toronto, ON M5T 1R8, Canada; ⁷Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht 3584, the Netherlands; ⁸Center for Human Genetics, University of Leuven, Leuven 3000, Belgium; ⁹Medical Genetics, Bambino Gesù Hospital, 00165 Rome, Italy; ¹⁰Lorillard Spencer Cenci Foundation and Department of Pediatrics, La Sapienza University of Rome, Rome 00165, Italy; ¹¹Department of Medical Genetics, AP-HM and University of Mediterranean, Timone Children's Hospital, Marseille 13005, France; ¹²M.I.N.D. Institute & Department of Psychiatry and Behavioral Sciences, University of California, Sacramento, CA 95817, USA; ¹³Department of Cardiology and Division of Genetics, Boston Children's Hospital, Boston, MA 02115, USA; ¹⁴Department of Genetics, Polish Mother's Memorial Hospital - Research Institute, 93-338 Łódź, Poland; ¹⁵Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, CA 90095, USA; ¹⁶Office Médico-Pédagogique Research Unit, Department of Psychiatry, University of Geneva School of Medicine, 1211 Geneva 8, Switzerland; ¹⁷Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Tel Hashomer, and affiliated with the Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 52621, Israel; ¹⁸Children's Healthcare of Atlanta and Emory University School of Nursing, Atlanta, GA 30322, USA; ¹⁹Department of Psychiatry and Behavioral Sciences, and Program in Neuroscience, SUNY Upstate Medical University, Syracuse, NY 13210, USA; ²⁰Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; ²¹Department of Molecular Medicine, University of Rome La Sapienza, Rome 00185, Italy; ²²Genetics Department, Hospital Universitari Son Espases, Palma de Mallorca 07020, Spain; ²³Department of Pediatrics, University of Colorado, Denver, CO 80202, USA; ²⁴Division of Cardiology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

*Correspondence: beverly@mail.med.upenn.edu

<http://dx.doi.org/10.1016/j.ajhg.2015.03.007>. ©2015 by The American Society of Human Genetics. All rights reserved.

approximately 1 in 2,000–4,000 individuals.^{9,10} The vast majority of individuals with 22q11DS carry the typical 3 million base pair (3 Mb) deletion of one homolog of chromosome 22; nested, smaller interstitial 1.5–2 Mb 22q11.2 deletions are seen in <10% of individuals.¹¹ Both the typical 3 Mb deletion and most nested interstitial deletions occur between low copy repeats that punctuate the 22q11.2 region.¹² This deletion is usually de novo but can also be inherited.¹³ The 22q11DS phenotype is highly variable and includes CHDs, dysmorphic facial features, palatal anomalies, hypocalcemia, immunodeficiency, cognitive impairment, and various neuropsychiatric disorders. A variety of CHDs and/or aortic arch defects have been detected in approximately 65% of individuals with 22q11DS, the most prevalent of which are CTDs.^{14,15} The etiology of this cardiovascular phenotypic variability is not currently known, but it does not appear to correlate with sex, race, 22q11.2 deletion size, or parent of origin of the deletion.^{8,16,17}

The variable expressivity and reduced penetrance of CHDs in 22q11DS (including aortic arch anomalies) is probably influenced by genetic factors because individuals with 22q11DS and a CHD are more likely to have an unaffected relative with an isolated CHD than individuals with 22q11DS that have normal intracardiac and aortic arch anatomy.⁸ These findings are not explained by the inheritance of the non-deleted chromosome 22, suggesting that the variants that influence the development of CHD in these families lie outside of the 22q11.2 region.⁸ More than 40 genes are in the typically deleted region in 22q11DS. One of the strongest candidate genes for CHD on 22q11DS is *TBX1* (MIM: 602054), which encodes a T-box transcription factor.^{18–20} We previously sequenced coding exons of *TBX1* in this cohort and did not find evidence for mutation on the remaining allele.²¹ Therefore, we hypothesized that individuals with 22q11DS and CHDs have structural variants that affect their risk of being born with intracardiac and/or aortic arch malformations, possibly through epistatic interactions with the dosage-sensitive gene(s) in the 22q11.2 deleted region.

Our study is an investigation in search of CNV genetic modifiers involved in the variable 22q11DS cardiac phenotype and represents the largest genomic study of a micro-deletion syndrome performed to date. Genome-wide analysis of CNVs was performed on two separate cohorts of subjects with 22q11DS ascertained and processed in two different time periods: the first cohort consisted of 562 subjects (CHD, *n* = 363; no CHD, *n* = 199) and the second cohort comprised 387 subjects (CHD, *n* = 240; no CHD, *n* = 147) for a total of 949 subjects (CHD, *n* = 603; no CHD, *n* = 346). By analyzing 949 subjects with 22q11DS, we were able to identify a common CNV that was significantly enriched in 22q11DS-positive subjects with a CHD. This CNV and the gene it overlaps have not been previously reported in the CHD literature. The result supports the possibility that this is a genetic modifier of CHDs in individuals with 22q11DS.

Material and Methods

Subject Cohorts

Blood or saliva samples were obtained from subjects with 22q11DS, with their informed consent and in accordance with the ethical standards of the appropriate committees on human experimentation (Internal Review Board, 1999-201, Albert Einstein College of Medicine, NY; 07-005352_CR2 CHOP IRB). Two groups of subjects with 22q11DS were ascertained and processed at two distinct time points and were therefore treated as separate cohorts, referred to as cohort 1 and cohort 2. The recruitment goals for the two 22q11DS cohorts were a confirmed 22q11.2 deletion, self-reported as white of European descent, and (for familial cases) only one individual per family. Fluorescence in situ hybridization or multiplex ligation-dependent probe amplification (MLPA) testing was used to verify the 22q11.2 deletion in each subject and parents when available. Phenotypic information on intracardiac and aortic arch anatomy was obtained from echocardiograph and cardiology summary reports from the referring institutions; every subject enrolled in the study had an echocardiogram. The phenotypes of 227 of these subjects have been described in an earlier publication.²¹

A separate cohort of subjects with CTDs was recruited as part of a larger collaborative program (HD70454). These subjects tested negative for a 22q11.2 deletion and had no other recognizable genetic syndrome. A detailed description of the subject enrollment requirements, cardiac phenotypes, and array genotyping procedure for this non-deleted, CTD cohort has been published elsewhere.²²

Genome-wide SNP Array, Quality Control, and CNV Detection

Genomic DNA samples from subjects with 22q11DS were analyzed with the Affymetrix SNP Array 6.0 platform according to the manufacturer's instructions (Affymetrix) at the Genomics Core at Albert Einstein College of Medicine. Quality control values were calculated in Affymetrix Genotyping Console (Affymetrix) and any samples with contrast QC greater than 0.4 or mean absolute pairwise difference (MAPD) greater than 0.35 were excluded from further analysis. Only samples that had a typical 3 Mb 22q11.2 deletion or a proximal nested 22q11.2 deletion were included; all atypical deletions were excluded from the study. In addition, samples were removed if there was insufficient cardiac phenotype information about the subject or if the gender determined on the basis of X and Y chromosome SNP genotypes did not match their reported gender. SNP analysis was performed to exclude duplicate or related samples via estimation of identity by descent (IBD) with the PLINK software package.²³ Figure S1A depicts this initial quality assessment and sample elimination.

A custom copy number (CN) baseline reference was generated with 215 Affymetrix SNP 6.0 arrays from phenotypically normal, non-22q11DS control individuals that were ascertained concurrently with the 22q11DS cohorts. These reference arrays were run in the same facility, the Genomics Core at Albert Einstein College of Medicine, during the same time period and with an identical protocol as our experimental 22q11DS arrays in order to control for any batch variation. The reference arrays were subjected to and passed the same QC metrics, and equivalent ratio of male:female arrays were used to prevent gender bias (106 male, 109 female).

The CNV detection was performed on all three cohorts (both 22q11DS groups and the non-syndromic CTD group) via two methods: PennCNV²⁴ and CNV workshop.²⁵ The PennCNV-Affymetrix tool was used to extract the signal intensity data from the raw .cel files.²⁴ The canonical genotype clustering file used for CNV calling with PennCNV was generated from our custom CN baseline reference set.²⁴ The log₂ratios generated by PennCNV were used in CNV Workshop to produce CNV calls via circular binary segmentation.²⁵ The B allele frequency and log R ratio plots were visualized with the Affymetrix Chromosome Analysis Suite to support CNV calls. The final QC step occurred after CNV detection: any samples with elevated log₂ratio SD (>0.44) or a large number of detected CNVs (>300) were excluded from further analysis, as shown in Figure S1B. A total of 949 22q11DS samples, 562 in the first cohort and 387 in the second cohort, passed all of the QC metrics and were included in our analysis.

CNV Analysis

A list of autosomal CNVs detected by greater than 10 contiguous probes for deletions and greater than 20 probes for duplications was generated with relevant annotations for analysis. Only CNVs detected by both PennCNV and CNV Workshop were included in the analysis, because these CNVs are less likely to be false positives due to variation in algorithms. CNV boundaries were determined by averaging the breakpoint locations predicted by CNV workshop and PennCNV. Any CNVs with a 50% or greater overlap with centromere, telomere, immunoglobulin regions, and/or segmental duplications were excluded. In addition, olfactory receptor genes were removed from further analysis. Finally, we merged CNVs separated by less than 10 kb to consider them as possible single contiguous events. CNV detection and analysis was performed with the GRCh36/hg18 build, and then CNV coordinates were converted to the GRCh37/hg19 build via the UCSC Genome Browser LiftOver tool. All genomic coordinates presented in figures and tables herein are based on the February 2009 Human Genome Build (GRCh37/hg19).

Common and rare CNVs were analyzed separately. We defined rare CNVs as those with a frequency of less than 0.1% of a previously published control population (dbVaR: nstd54);²⁶ all remaining CNVs were categorized as common. Here we focus on common CNVs to determine the impact of common variants as genetic modifiers of the 22q11DS cardiac phenotype.

In Silico Analysis of Gene Function

CNVs in subjects with 22q11DS that passed all selection criteria were annotated with the RefSeq gene set downloaded from the UCSC Table browser of the hg18 build. Gene Ontology (GO) annotations for each RefSeq gene were obtained from the Ensembl database and their Mammalian Phenotype Ontology (MPO) term annotations were retrieved from the Mouse Genome Informatics (MGI) database (July 2014 available version). Previously published analytical methods were employed to expand the annotation of the Gene Ontology and Mammalian Phenotype Ontology terms.²⁷ For each functional term (GO and MPO), we directly compared the frequency of occurrence between cases and controls via Fisher's exact test. Duplication and deletion events were evaluated separately after excluding genes deleted in the 22q11.2 region.

Statistical Analysis

The Fisher exact test or the Wilcoxon rank sum test was used for gene and CNV enrichment analyses. The Benjamin-Hochberg

false discovery rate procedure was adopted as our default multiple test correction method. Where appropriate, the permutation-based false discovery rate estimation was applied to correct multiple testing for functional analyses.²⁸

CNV Validation by qPCR

CNVs selected for validation were screened by real-time PCR. SYBR Green detection on an ABI SDS-7500 Fast Real-Time PCR system was used to quantify copy number (Applied Biosystems). Primer Express 3.0 software (ABI) and Primer 3 were used to design primers to amplify the region of interest; the specificity of each primer pair was tested with the UCSC In-Silico PCR tool. The average length of each amplicon was 62 base pairs (range: 51–97 base pairs). For each CNV, primers were designed to amplify two regions within the deleted/duplicated region and at least one set of primers was designed to amplify a flanking region with normal copy number. Each qPCR run included amplification of an endogenous control with known copy number (*RPPH1*). The sequences for all of the qPCR primers used in the CNV validation are listed in Table S5. Two DNA samples with normal copy number (one CEPH subject and one 22q11DS-positive subject with normal copy number at the particular CNV) were used as controls in each run. SYBR Green qPCR was also used to amplify the DNA from the available parents of subjects with 22q11DS to determine whether CNVs were inherited or de novo. 10 µl reactions were performed with 12.5 ng of DNA according to the manufacturer's recommended protocol.

Assessing *SLC2A3* Duplication Frequency in Control Populations

Publicly available databases were examined to determine the frequency of *SLC2A3* duplications in phenotypically normal individuals. Many of the studies submitted to these databases did not provide robust phenotypic information for control samples or a sample ID associated with each CNV. It is not possible to know how many samples are redundant or uniquely presented in a database without adequate sample ID information, which prevents a reliable calculation of the frequency of the *SLC2A3* duplication within the various databases. We were therefore restricted to previously published studies where sufficient sample ID information was provided and individuals were rigorously vetted to ensure a phenotypically "normal" control cohort (dbVaR: nstd21²⁹ and nstd54²⁶).

Mouse Embryo Dissections

Mouse embryos in the SW background were isolated in cold PBS at E9.5 and E10.5. Somite pairs were counted to define stages: 19–21 pairs of somites were defined as E9.5 and 30–32 pairs of somites were defined as E10.5.

Whole-Mount In Situ Hybridization

Digoxigenin-labeled RNA probe for *Slc2a3* was amplified by PCR from cDNA, with the primers 5'-TCCCCTCAGCTGCAGCCTA CTT-3' and 5'-TTGTTCAATCCCCCAGGGCCCT-3', forward and reverse, respectively. The forward primer contained the T3 polymerase priming sequence and the reverse primer contained the T7 polymerase priming sequence. The PCR products were purified with the PCR Purification Kit (QIAGEN), and antisense RNA was in vitro transcribed and labeled with T7 RNA polymerase (Roche) and the DIG RNA Labeling Mix (Roche), via the Digoxigenin Labeling Method. A sense RNA was generated with T3 RNA polymerase following the same procedure as the antisense RNA. The

Table 1. Frequency of CHDs in Subjects with 22q11DS

Type of CHD	First Cohort		Second Cohort	
	Total ^a	No. of Isolated ^b	Total ^a	No. of Isolated ^b
ASD	56	12	59	18
IAAB	41	3	25	0
PS	105	3	21	2
PTA	35	0	19	0
RAA	114	30	41	14
TOF	149	40	59	47
VSD ^c	118	15	81	47

Abbreviations are as follows: ASD, atrial septal defect; IAAB, interrupted aortic arch type B; PS, pulmonary atresia stenosis; PTA, persistent truncus arteriosus; RAA, right aortic arch; TOF, tetralogy of Fallot; VSD, ventricular septal defect.

^aTotal number of subjects with each type of CHD. Subjects with >1 heart defects are included in multiple categories.

^bNumber of subjects with isolated CHDs.

^cThis category does not include VSDs found in association with TOF and PTA.

sense RNA served as control for the specificity of the expression pattern observed with the antisense RNA. Digoxigenin-labeled RNA probes were purified with mini Quick Spin RNA Columns (Roche). Whole-mount in situ hybridization was performed as previously described.³⁰

Results

CNV Analysis in Two Consecutively Ascertained Cohorts of Subjects with 22q11DS

CNV analysis was performed with PennCNV and CNV workshop on 949 Affymetrix SNP6.0 arrays from two cohorts of subjects with 22q11DS, cohort 1 and cohort 2. The two cohorts were ascertained over different but consecutive time periods. When combined there were 603 individuals (cohort 1, $n = 363$; cohort 2, $n = 240$) with intracardiac defects and/or aortic arch defects and 346 (cohort 1, $n = 199$; cohort 2, $n = 147$) that had normal heart and aortic arch anatomy. The specific defects observed in these subjects are listed in Table 1. Each subject had a 22q11.2 deletion that included *TBX1* (Figure 1), a strong candidate gene for the physical defects of the syndrome including CHDs.^{18–20} Although parental DNA was not available to test from all subjects, the majority of the 22q11.2 deletions in the study cohort were de novo events. Analysis of the 22q11.2 deletion sizes (Table S1) revealed that 93.5% of subjects ($n = 564$) with a CHD carried the typical 3 Mb (LCR-A to LCR-D) 22q11.2 deletion^{12,31} compared to 94.8% of subjects ($n = 328$) with a normal heart ($p = 0.48$, two-tailed Fisher's exact test). Figure 1 illustrates the various 22q11.2 deletions and Table S1 contains the deletion size data broken down by cohort in addition to the overall frequencies. These findings indicate that the size of the 22q11.2 deletion does not play a role in the development of congenital cardiac defects.

A total of 13,518 CNVs outside of the 22q11.2 region were detected by both algorithms in the 949 subjects

(Table 2). These CNVs are unlikely to be false positives because they were identified by two methods and passed the probe cutoffs that were chosen based on extensive validation of such CNVs in the past. There was no significant difference in the number of CNVs detected (14.29 ± 5.34 versus 13.88 ± 5.01 , $p = 0.37$) or the average size of CNVs (607.72 ± 484.05 kb versus 575.48 ± 437.78 kb, $p = 0.22$) between subjects with CHDs including aortic arch anomalies and those with normal cardiovascular anatomy in either cohort or when the cohorts were combined (Table 2). CNVs affecting genes implicated in cardiac development and/or overrepresented in case or control subjects were chosen for qPCR validation (number of unique CNVs = 23; Table S2). The validation rate for CNVs that were identified by both algorithms (CNV Workshop and PennCNV, $n = 70$) was 100%, but was lower for CNVs predicted by only one algorithm (validation rate for CNVs identified only by CNV workshop = 35.5%; validation rate for CNVs identified only by PennCNV = 71.4%).

Identification and Analysis of the *SLC2A3* Duplication

In the analysis of the first cohort, only a single CNV showed a statistically significant difference in frequency between 22q11DS-positive subjects with and without CHDs. A common duplication of chromosome 12p13.31 was detected in 18 (5.0%) 22q11DS-positive subjects with a CHD and in 1 (0.5%) 22q11DS-positive subject with a normal heart ($p = 3.12 \times 10^{-3}$, two-tailed Fisher's exact test; Figure 2A). This duplication encompasses the entirety of *SLC2A3* (solute carrier family 2, member 3 [MIM: 138170]), the pseudogene *NANOGP1*, and part of *SLC2A14* (solute carrier family 2, member 14 [MIM: 611039]). *SLC2A14* is expressed only in the testes³² and the pseudogene *NANOGP1* is transcribed but not translated and therefore neither is relevant. Thus, this CNV at 12p13.31 is referred to as a duplication of *SLC2A3* hereafter.

The second cohort was then examined and an additional 20 *SLC2A3* duplications were identified: 17 (7.1%) subjects had a CHD and 3 (2.0%) did not. The *SLC2A3* duplication was also significantly enriched in individuals with CHDs within the second cohort ($p = 3.30 \times 10^{-2}$, two-tailed Fisher's exact test; Figure 2B). In total, 39 individuals with a duplication of *SLC2A3* were identified: 35 in 22q11DS-positive subjects with CHD (5.8%) and in 4 (1.1%) of the 22q11DS-positive subjects with normal hearts ($p = 2.68 \times 10^{-4}$, two-tailed Fisher's exact test; Figure 2). A duplication of the entire *SLC2A3* gene has been previously observed in 45/2,026 (2.2%) of healthy individuals.²⁹ Thus, it appears that the overall frequency of this duplication in our entire cohort with 22q11DS (4.1%) is substantially greater than that seen in healthy controls.

The *SLC2A3* duplication was validated with qPCR in all 37 individuals for whom DNA samples were still available (Figure 2C shows representative qPCR data). We

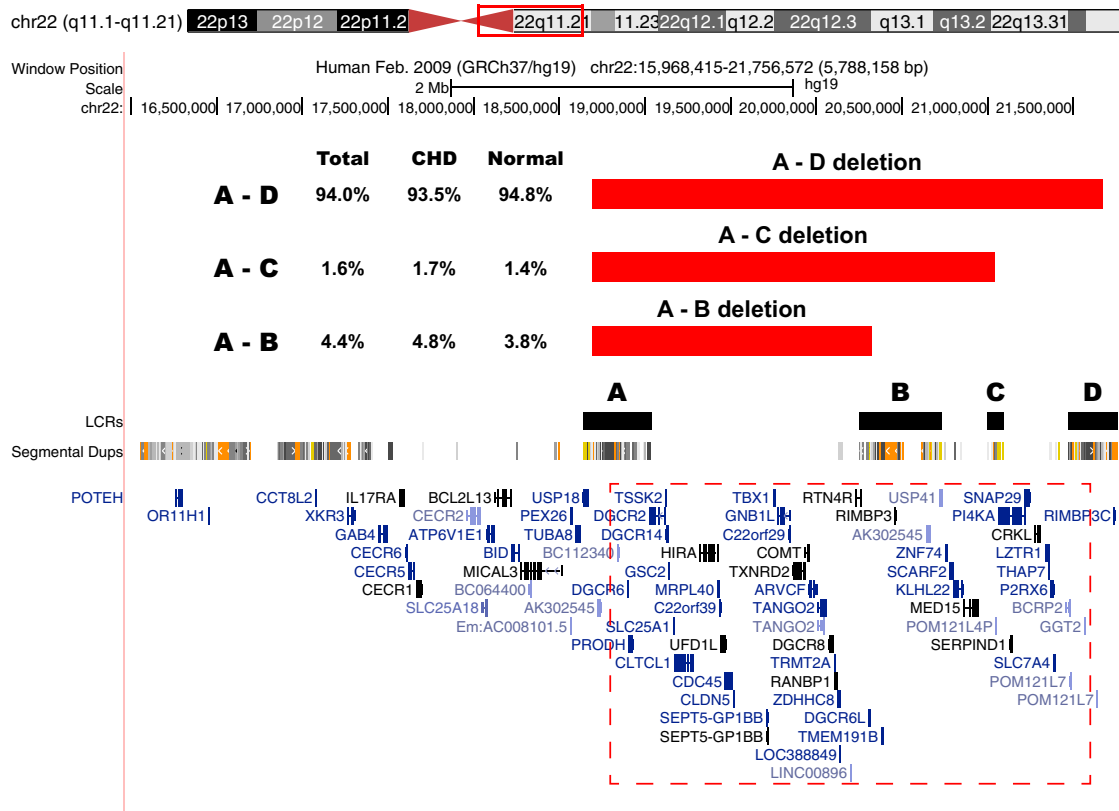


Figure 1. Chromosome 22 Deletion Sizes

UCSC Genome Browser view of the 22q11.2 deletion sizes in the 949 subjects with 22q11DS from both cohorts. The typical 22q11.2 deletions are mediated by low copy repeats (LCRs); the deletions start at LCR A and end within LCR B, C, or D. The LCRs are shown in black directly above the segmental duplication track. Table S1 contains the exact frequency and distribution for each of the typical deletions. Individuals with atypical 22q11.2 deletions were excluded from analysis. The red hatched box contains the genes that are typically deleted.

investigated whether the *SLC2A3* duplication was de novo or inherited in the 13 probands with parental DNA available for qPCR analysis (11 case subjects with CHDs and 2 controls without). As expected for a common variant, 100% of the individuals tested had inherited the *SLC2A3* CNV from an unaffected parent that did not have a 22q11.2 deletion or a CHD (9 maternal; 3 paternal; 1 case where mother and father both carried the *SLC2A3* duplication). A deletion of the *SLC2A3* CNV was identified in 9 subjects with a CHD (1.4%) and 4 subjects with a normal heart (1.1%) ($p = 0.78$, two tailed Fisher's exact test). This indicates that only a duplication of *SLC2A3* and not the hemizygous deletion is associated with CHDs in subjects with 22q11DS.

The CNV analysis of the non-deleted CHD cohort (627 subjects with CHDs and 2,980 normal controls of European descent) was performed as previously described.²⁷ The detected CNVs were examined to determine whether *SLC2A3* duplications were associated with non-syndromic CHDs. A duplication of *SLC2A3* was identified in 19 of the 627 individuals with a CHD (3.0%) and in 75 of the 2,980 control subjects with no reported CHD (2.5%). The *SLC2A3* duplication was not enriched among individuals with CHDs in this non-syndromic cohort ($p = 0.49$, two

tailed Fisher's exact test). Deletions of *SLC2A3* were identified in 3 individuals with a non-syndromic CHD and 22 normal controls ($p = 0.60$, two tailed Fisher's exact test), which is a similar distribution as was seen in the 22q11DS cohorts.

Slc2a3 RNA In Situ Hybridization of Mouse Embryos

Whole-mount in situ hybridization was performed to determine whether *Slc2a3* is expressed in the murine pharyngeal apparatus and/or heart during development. Cells from the pharyngeal apparatus migrate into the cardiac outflow tract to form the conotruncal region. *Slc2a3* was expressed in the brain, pharyngeal arches, and outflow tract but less so in the heart and placenta at embryonic days 9.5 and 10.5 (Figure 3). It was also expressed in the pancreatic bud at embryonic day 10.5.

In Silico Analysis of CNV Function

To determine relevance of the common CNVs observed in subjects with 22q11DS to cardiac development, we used phenotype data from Gene Ontology and Mouse Genome Informatics Resource as previously described.²⁸ Gene Ontology (GO) analysis was performed in order to examine the annotated biological processes, cellular components,

Table 2. Total Number and Length of All Autosomal CNVs^a in 22q11DS-Positive Subjects with and without CHDs

	First Cohort		Second Cohort		Cohorts Combined	
	CHD (n = 363)	No CHD (n = 199)	CHD (n = 240)	No CHD (n = 147)	CHD (n = 603)	No CHD (n = 346)
No. of male subjects (Percent male)	174 (47.8%)	107 (53.8%)	119 (49.6%)	68 (46.3%)	293 (48.5%)	175 (50.6%)
Mean no. of CNVs per subject ± SD	14.34 ± 5.49	13.90 ± 5.39	14.09 ± 4.57	13.86 ± 4.45	14.29 ± 5.34	19.25 ± 5.01
Mean CNV length (kb) per subject ± SD (kb)	627.67 ± 519.78	568.15 ± 397.87	577.55 ± 423.61	585.39 ± 487.87	607.72 ± 484.05	575.48 ± 437.78
No. with ≥ 1 CNV ≥ 500 kb	47	24	25	19	72	43
					p Value	p Value
					0.53 ^b	0.59 ^b
					0.83 ^c	0.37 ^c
					0.87 ^c	0.22 ^c
					0.51 ^b	0.84 ^b

^aExcluding the 22q11.2 deletion.

^bFisher's exact test.

^cWilcoxon rank-sum test.

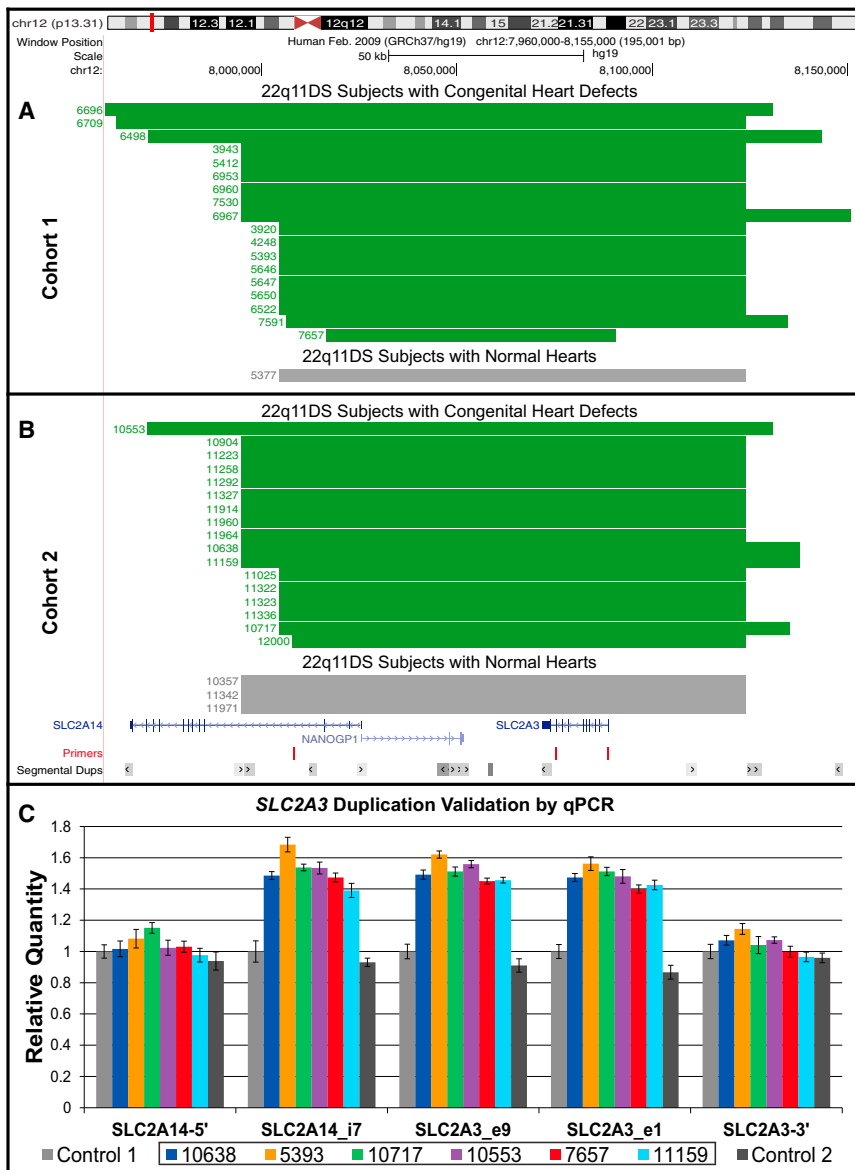
and/or molecular functions of genes impacted by CNVs in CHD case versus control subjects. Mammalian Phenotype (MP) analysis was performed to investigate the various phenotypes associated with genes impacted by CNVs in CHD case versus control subjects. The common and rare CNVs were evaluated separately for each analysis. The rare CNV analysis was less informative and will be reported elsewhere (E.E.M. and B.S.E., unpublished data). The results from the GO analysis are listed in Table S3. The only GO terms that were statistically significant after B-H/FDR adjustment were all related to *SLC2A3* ontologies and function. The results from the MP analyses are listed in Table S4. Several of the statistically significant MP terms enriched in 22q11DS-positive subjects with CHDs involved physiological defects and lethality such as “abnormal extraembryonic tissue physiology” and “complete embryonic lethality” (Table S4).

Discussion

The goal of our study was to determine whether structural genetic variants outside the 22q11.2 commonly deleted region explain the incomplete penetrance of CHDs in 22q11DS. Our analysis of subjects with 22q11DS, divided into two cohorts based upon time period of ascertainment and processing, revealed that the number of autosomal CNVs in those with CHDs compared to subjects with a normal heart was not significantly different in either cohort (Table 2). Furthermore, there was no correlation between the size of the 22q11.2 deletion and cardiac phenotype (Figure 1 and Table S1). However, one common CNV was found to be significantly associated with subjects with CHDs, suggesting that it might have an influence on cardiac development in the presence of a 22q11.2 deletion.

Duplication of *SLC2A3*

One common CNV, the duplication of *SLC2A3*, was significantly enriched in 22q11DS-positive subjects with a CHD. In total, the *SLC2A3* duplication was detected in 35 22q11DS-positive subjects with CHD (5.8%) and in 4 22q11DS-positive subjects with normal hearts (1.1%) ($p = 2.68 \times 10^{-4}$, two-tailed Fisher's exact test; Figure 2). This is the first study to report an association between a gain of *SLC2A3* and CHD. The *SLC2A3* duplication does not correlate with a specific type of heart defect in our 22q11DS-positive population, as shown in Table 3. This CNV is present in 2.2% of healthy individuals (dbVar: nstd21)²⁹ and was inherited from a parent in all 13 of the subjects with 22q11DS with available parental DNA. Cooper et al. identified a *SLC2A3* duplication in 1/575 subjects with cardiovascular disease (0.17%) and in 143/8,329 controls (1.7%) (dbVar: nstd54).²⁶ Collectively, these data suggest that the *SLC2A3* duplication is benign unless it is inherited in combination with the 22q11.2 deletion. This intriguing finding seems to be even more compelling because it was initially found in excess in the



first cohort and then it was “replicated” in a separate, consecutively ascertained group of subjects with 22q11DS (i.e., in cohort 2).

SLC2A3, formerly known as *GLUT3*, encodes a facilitated glucose transporter. *SLC2A3* was first isolated from neurons and was originally considered to be a neuronal-specific glucose transporter, but studies have shown that *SLC2A3* is expressed in a variety of human tissues.³³ *SLC2A3/GLUT3* is important in tissues with heightened energy demands and high metabolic rates because *GLUT3* has the highest glucose affinity and greatest transport capacity in the GLUT protein family; expression is most abundant in the brain, high in adult cardiac myocytes, liver, placenta, and at a barely detectable level in kidney.^{33,34} Here we have shown that murine *Slc2a3* is expressed in the pharyngeal apparatus and cardiac outflow tract at embryonic days 9.5 and 10.5 during murine cardiac morphogenesis (Figure 3). Previous work done in rats

showed that *GLUT3* is the prevailing glucose transporter in cardiomyoblasts and therefore it has been suggested that *GLUT3* has a predominant role during cardiac development.³⁵ Grover-McKay et al. determined that *SLC2A3* is expressed during human heart development because *SLC2A3* protein was present in the fetal myocardium at 10 weeks, increased protein levels were detected at 15 weeks, and the levels then decreased at 20 weeks of gestation.³⁶ Recent work has shown that there are dynamic expression changes in fetal myocardium during development: *SLC2A3* transcripts were detected at 15.9-fold higher levels than in newborn infants.³⁷ Together these observations indicate that *SLC2A3* might be involved in cardiac development, but the specific function of *SLC2A3* in the heart during embryogenesis has not been delineated.

The importance of *SLC2A3/GLUT3* during development has been well documented. *SLC2A3* is the main glucose transporter responsible for transplacental transport of maternal glucose, thereby controlling the rate at which glucose is delivered to the fetus.³⁸ *SLC2A3* expression adaptively responds to glucose demands of fetal growth during normal development. Aberrant levels of *SLC2A3* have been linked to intrauterine growth retardation and pregnancy loss.^{39–41} Animal models have shown that

Table 3. Cardiac Phenotype for 22q11DS-Positive Subjects with the SLC2A3 Duplication

ID	Gender	Inherited	Intracardiac Phenotype	Aortic Arch Phenotype
Cohort 1				
6960	F	paternal	TOF, PDA	RAA, PS
5412	M	ND	TOF	PS
5393	F	ND	TOF	RAA
3920	M	ND	TOF	aberrant RSCA
3943	F	ND	TOF	RAA, aberrant LSCA
7657	F	ND	TOF	normal
5647	F	ND	TOF	normal
6522	M	ND	VSD, ASD, bicuspid aortic valve, PDA	IAAB
6953	M	ND	VSD, PFO	aberrant RSCA, IAAB
6498	F	maternal	VSD, ASD	aberrant RSCA, TGA, PS
7591	M	ND	VSD, ASD	aberrant RSCA
5646	F	ND	VSD, bicuspid aortic valve	RAA, PS
6709	M	maternal	VSD	RAA, aberrant LSCA
5650	F	ND	ASD	normal
6696	F	ND	PTA	aberrant RSCA
4248	F	maternal	normal	aberrant RSCA
6967	F	ND	normal	aberrant LSCA
7530	F	maternal	normal	aberrant RSCA
5377	M	ND	normal	normal
Cohort 2				
10638	M	ND	VSD, ASD, PDA	coarctation
11025	F	ND	VSD, ASD	normal
11336	M	ND	VSD, ASD	normal
10553	M	paternal	VSD	IAAB
11964	F	both ^a	VSD	IAAB
10717	F	maternal	VSD	normal
11223	F	ND	VSD	normal
11914	M	ND	VSD	normal
11292	F	ND	VSD	normal
11322	M	ND	VSD	PS
10904	M	ND	ASD, PDA	normal
11327	F	ND	ASD	normal
11258	M	maternal	PDA	RAA, aberrant LSCA
11960	F	paternal	bicuspid aortic valve	aberrant RSCA

Table 3. Continued

ID	Gender	Inherited	Intracardiac Phenotype	Aortic Arch Phenotype
11323	F	ND	unspecified congenital heart defect ^b	normal
11159	F	maternal	normal	RAA, aberrant LSCA, vascular ring
12000	F	ND	normal	aberrant LSCA
10357	F	ND	normal	normal
11342	M	maternal	normal	normal
11971	F	maternal	normal	normal

Abbreviations are as follows: LSCA, left subclavian artery; RSCA, right subclavian artery; PDA, patent ductus arteriosus; PS, pulmonary atresia stenosis; PTA, persistent truncus arteriosus; RAA, right aortic arch; TGA, transposition of the great arteries; ND, DNA not available for testing.

^aMother and father both carry the *SLC2A3* duplication.

^bRequired surgical intervention 3 months after birth.

SLC2A3 alterations can cause extremely deleterious developmental defects. The homozygous null deletion of *Slc2a3* is embryonic lethal in mice and the heterozygous deletion is not lethal but results in intrauterine growth retardation.^{39,42} Knockdown of the *SLC2A3* zebrafish ortholog, *slc2a3a*, increased apoptosis and was embryonic lethal.⁴³ Thus, although animal models have clearly demonstrated that loss of *SLC2A3* can cause significant defects, currently there are no reports in the literature about the developmental effect of a *SLC2A3* duplication or overexpression in animal models. However, it has been shown in a variety of human cell types that a duplication of *SLC2A3* results in significantly increased expression and protein levels.^{44,45}

There are two recent reports of *SLC2A3* CNVs in association with human disease phenotypes.^{45,46} Both duplications and deletions of *SLC2A3* were identified in a genetic study of rheumatoid arthritis (MIM: 180300). The deletions were deemed protective against rheumatoid arthritis whereas duplications of *SLC2A3* had no effect.⁴⁶ *SLC2A3* duplications and deletions were also detected in subjects with Huntington disease (MIM: 143100); the duplication correlated with delayed age of onset.⁴⁵ It is important to note that the individuals in the Huntington disease and rheumatoid arthritis studies do not carry the 22q11.2 deletion. In the absence of the 22q11.2 deletion, individuals with a duplication of *SLC2A3* do not present with congenital heart defects, indicating that both mutations might be required for the manifestation of a CHD.

Gains overlapping *SLC2A3* have not been reported in studies of non-syndromic CHDs although they have been detected at the same frequency as the general population.⁴⁷ In our analysis of a different non-syndromic cohort, the *SLC2A3* duplication showed no enrichment ($p = 0.49$, two-tailed Fisher's exact test) (unpublished data). These results are consistent with the fact that, in

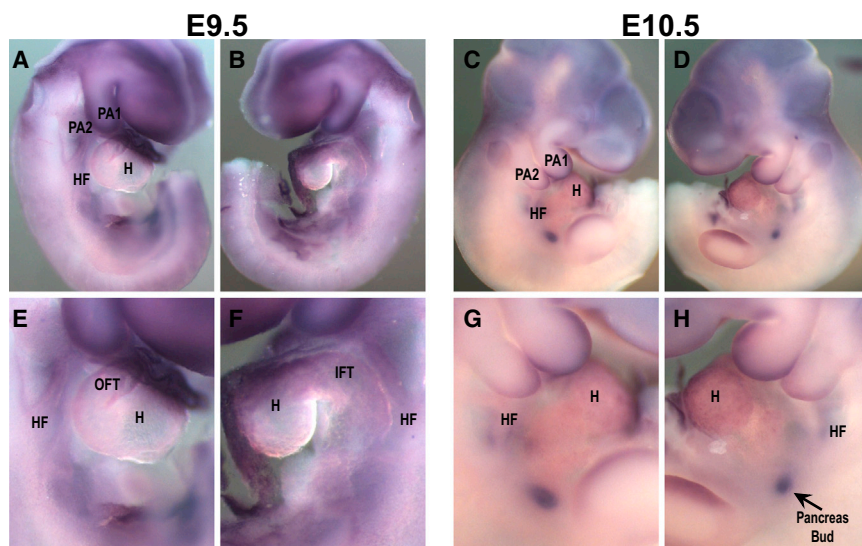


Figure 3. *Slc2a3* RNA In Situ Hybridization

(A–D) Whole-mount in situ hybridization on wild-type mouse embryos showing the expression pattern of *Slc2a3* (*Glut3*) at embryonic stages E9.5 (A and B) and E10.5 (C and D). Panels (A) and (C) show the right side of the embryos, and panels (B) and (D) show the left side. *Slc2a3* is expressed in the brain, pharyngeal arches 1 and 2 (PA1 and PA2), in the heart field (HF), and in the heart (H) at both embryonic stages.

(E–H) Higher magnification of the central part of the embryos shown in (A)–(D). *Slc2a3* is also expressed in the placenta (not shown) and in the pancreas bud at E10.5 (H, white arrow).

Abbreviations are as follows: OFT, outflow tract; IFT, inflow tract.

the 22q11.2-deleted cohorts, the *SLC2A3* duplication is inherited from unaffected parents. Thus, it appears likely that there might be an epistatic interaction between the *SLC2A3* duplication and dosage-sensitive gene(s) in the 22q11.2-deleted region that increase the likelihood of a structural cardiac defect. Therefore, *SLC2A3* is probably a modifier of the 22q11DS cardiac phenotype that exemplifies a “two-hit” model.

Genetic modifiers and the two-hit model are not unique to 22q11DS, but thus far the two-hit model has been used to explain only the variable expressivity of CNVs associated with neurodevelopmental phenotypes.⁴⁸ The *SLC2A3* duplication might be a genetic modifier of the 22q11DS cardiac phenotype. However, it does not completely explain the etiology of heart defects in the 22q11DS-positive population because the CNV was seen in only 5.8% of subjects with a CHD and also in 4 individuals with a purportedly “normal” heart. We classified these 4 subjects as controls because intracardiac defects were not noted in their echocardiogram report. Although these individuals have a normal left arch, it is quite possible they have an aberrant right subclavian that was not reported because very few look for this type of abnormality or because the imaging was insufficient to detect it. Alternatively, the combined effect of the *SLC2A3* duplication with the 22q11.2 deletion increases the risk of having a CHD but might not be sufficient to cause a structural defect, and perhaps something additional in the genetic background and/or possibly exposure to epigenetic, environmental, or maternal factors in utero might be necessary for the manifestation of a CHD.

Previously Identified CNVs Associated with Non-syndromic CHDs

A number of recent studies have examined the prevalence of CNVs in non-syndromic individuals with CHDs. These studies have illustrated that rare CNVs can play a role in the pathogenesis of non-syndromic CHDs.^{47,49–53} Both

22q11DS cohorts were assessed for the CNVs reported in these non-syndromic studies. With the exception of the 22q11.2 deletion reported by Greenway et al.,⁵² none of the previously identified CNVs were detected or significantly associated with CHDs in our 22q11DS cohorts. This is not surprising because the non-syndromic CHD studies focused predominantly on rare CNVs, which are mechanistically distinct from the goal of this study. Rather than looking for causal primary lesions in the form of rare CNVs, we started with the 22q11.2 deletion and investigated the genomes of deleted individuals for CNVs that act as possible genetic modifiers.

Conclusion

This study sheds new insight onto the idea that copy-number changes at two different loci might cause or affect penetrance of a structural birth defect during development.⁴⁸ Individuals with 22q11DS carry an initial genetic lesion that significantly elevates their risk of developing CHDs with variable expressivity. The 22q11.2 deletion was the only CNV present in every 22q11DS-positive individual with a CHD; however, one common CNV was significantly enriched and associated with CHDs in our two cohorts of subjects with 22q11DS. It suggests that variability of CHD phenotype in 22q11DS might be in part due to a duplication of *SLC2A3*. This finding supports a possible “two-hit” model where CNVs outside of the deleted region might explain the incomplete penetrance of the 22q11.2 deletion as a facilitator of congenital heart defects in a subset of subjects with 22q11DS. In the future, functional validation studies with a model system will help elucidate the role of *SLC2A3* duplication as a cardiac modifier.

Supplemental Data

Supplemental Data include one figure and five tables and can be found with this article online at <http://dx.doi.org/10.1016/j.ajhg.2015.03.007>.

Acknowledgments

Thanks to Colleen Franconi, Meghan McNamara, April Hacker, Tracy Busse, Mark Bowser, Anne Marie Higgins, Chad Haldeman-Englert, Petra Warner, Jenna Walck, and Danica Kuncio for technical assistance. Special thanks go to the subjects and their families for their willing participation. We also thank the Molecular Cytogenetics and Genomics Cores at Einstein for preparing DNA and genotyping with Affymetrix 6.0 arrays. This work was funded by NIH (MH87636, HL84410, HD070454, and HD026979) and in part by NIH/NCATS (National Center for Advancing Translational Sciences), Grant UL1TR000003.

Received: December 4, 2014

Accepted: March 11, 2015

Published: April 16, 2015

Web Resources

The URLs for data presented herein are as follows:

dbVar, <http://www.ncbi.nlm.nih.gov/dbvar/>
Gene Ontology Consortium, <http://geneontology.org/>
Mouse Genome Informatics, <http://www.informatics.jax.org/>
OMIM, <http://www.omim.org/>
PLINK, <http://pngu.mgh.harvard.edu/~purcell/plink/>
Primer3, <http://bioinfo.ut.ee/primer3>
UCSC Genome Browser, <http://genome.ucsc.edu>
UCSC In-Silico PCR, <http://genome.ucsc.edu/cgi-bin/hgPcr?command=start>

References

1. Rosano, A., Botto, L.D., Botting, B., and Mastroiacovo, P. (2000). Infant mortality and congenital anomalies from 1950 to 1994: an international perspective. *J. Epidemiol. Community Health* *54*, 660–666.
2. van der Bom, T., Zomer, A.C., Zwinderman, A.H., Meijboom, F.J., Bouma, B.J., and Mulder, B.J. (2011). The changing epidemiology of congenital heart disease. *Nat. Rev. Cardiol.* *8*, 50–60.
3. Goldmuntz, E., Clark, B.J., Mitchell, L.E., Jawad, A.F., Cuneo, B.F., Reed, L., McDonald-McGinn, D., Chien, P., Feuer, J., Zackai, E.H., et al. (1998). Frequency of 22q11 deletions in patients with conotruncal defects. *J. Am. Coll. Cardiol.* *32*, 492–498.
4. 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.
5. Zhu, H., Kartiko, S., and Finnell, R.H. (2009). Importance of gene-environment interactions in the etiology of selected birth defects. *Clin. Genet.* *75*, 409–423.
6. Jenkins, K.J., Correa, A., Feinstein, J.A., Botto, L., Britt, A.E., Daniels, S.R., Elixson, M., Warnes, C.A., and Webb, C.L.; American Heart Association Council on Cardiovascular Disease in the Young (2007). Noninherited risk factors and congenital cardiovascular defects: current knowledge: a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* *115*, 2995–3014.
7. Lammer, E.J., Chak, J.S., Iovannisci, D.M., Schultz, K., Osoegawa, K., Yang, W., Carmichael, S.L., and Shaw, G.M. (2009). Chromosomal abnormalities among children born with conotruncal cardiac defects. *Birth Defects Res. A Clin. Mol. Teratol.* *85*, 30–35.
8. Swaby, J.A., Silversides, C.K., Bekeschus, S.C., Piran, S., Oechslin, E.N., Chow, E.W., and Bassett, A.S. (2011). Complex congenital heart disease in unaffected relatives of adults with 22q11.2 deletion syndrome. *Am. J. Cardiol.* *107*, 466–471.
9. Burn, J., and Goodship, J. (1996). Developmental genetics of the heart. *Curr. Opin. Genet. Dev.* *6*, 322–325.
10. Robin, N.H., and Shprintzen, R.J. (2005). Defining the clinical spectrum of deletion 22q11.2. *J. Pediatr.* *147*, 90–96.
11. Emanuel, B.S. (2008). Molecular mechanisms and diagnosis of chromosome 22q11.2 rearrangements. *Dev. Disabil. Res. Rev.* *14*, 11–18.
12. Shaikh, T.H., Kurahashi, H., Saitta, S.C., O'Hare, A.M., Hu, P., Roe, B.A., Driscoll, D.A., McDonald-McGinn, D.M., Zackai, E.H., Budarf, M.L., and Emanuel, B.S. (2000). Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum. Mol. Genet.* *9*, 489–501.
13. McDonald-McGinn, D.M., Tonnesen, M.K., Laufer-Cahana, A., Finucane, B., Driscoll, D.A., Emanuel, B.S., and Zackai, E.H. (2001). Phenotype of the 22q11.2 deletion in individuals identified through an affected relative: cast a wide FISHing net!. *Genet. Med.* *3*, 23–29.
14. McDonald-McGinn, D.M., LaRossa, D., Goldmuntz, E., Sullivan, K., Eicher, P., Gerdes, M., Moss, E., Wang, P., Solot, C., Schultz, P., et al. (1997). The 22q11.2 deletion: screening, diagnostic workup, and outcome of results; report on 181 patients. *Genet. Test.* *1*, 99–108.
15. Ryan, A.K., Goodship, J.A., Wilson, D.I., Philip, N., Levy, A., Seidel, H., Schuffenhauer, S., Oechsler, H., Belohradsky, B., Prieur, M., et al. (1997). Spectrum of clinical features associated with interstitial chromosome 22q11 deletions: a European collaborative study. *J. Med. Genet.* *34*, 798–804.
16. Sandrin-Garcia, P., Abramides, D.V., Martelli, L.R., Ramos, E.S., Richieri-Costa, A., and Passos, G.A. (2007). Typical phenotypic spectrum of velocardiofacial syndrome occurs independently of deletion size in chromosome 22q11.2. *Mol. Cell. Biochem.* *303*, 9–17.
17. Goldmuntz, E., Driscoll, D.A., Emanuel, B.S., McDonald-McGinn, D., Mei, M., Zackai, E., and Mitchell, L.E. (2009). Evaluation of potential modifiers of the cardiac phenotype in the 22q11.2 deletion syndrome. *Birth Defects Res. A Clin. Mol. Teratol.* *85*, 125–129.
18. Merscher, S., Funke, B., Epstein, J.A., Heyer, J., Puech, A., Lu, M.M., Xavier, R.J., Demay, M.B., Russell, R.G., Factor, S., et al. (2001). TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell* *104*, 619–629.
19. Jerome, L.A., and Papaioannou, V.E. (2001). DiGeorge syndrome phenotype in mice mutant for the T-box gene, *Tbx1*. *Nat. Genet.* *27*, 286–291.
20. Lindsay, E.A., Vitelli, F., Su, H., Morishima, M., Huynh, T., Pramparo, T., Jurecic, V., Ogunrinu, G., Sutherland, H.F.,

- Scambler, P.J., et al. (2001). Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* *410*, 97–101.
21. Guo, T., McDonald-McGinn, D., Blonska, A., Shanske, A., Bassett, A.S., Chow, E., Bowser, M., Sheridan, M., Beemer, F., Devriendt, K., et al.; International Chromosome 22q11.2 Consortium (2011). Genotype and cardiovascular phenotype correlations with TBX1 in 1,022 velo-cardio-facial/DiGeorge/22q11.2 deletion syndrome patients. *Hum. Mutat.* *32*, 1278–1289.
 22. Agopian, A.J., Mitchell, L.E., Glessner, J., Bhalla, A.D., Sewda, A., Hakonarson, H., and Goldmuntz, E. (2014). Genome-wide association study of maternal and inherited loci for conotruncal heart defects. *PLoS ONE* *9*, e96057.
 23. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* *81*, 559–575.
 24. Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., Hakonarson, H., and Bucan, M. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* *17*, 1665–1674.
 25. Gai, X., Perin, J.C., Murphy, K., O'Hara, R., D'arcy, M., Wenocur, A., Xie, H.M., Rappaport, E.F., Shaikh, T.H., and White, P.S. (2010). CNV Workshop: an integrated platform for high-throughput copy number variation discovery and clinical diagnostics. *BMC Bioinformatics* *11*, 74.
 26. Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V., et al. (2011). A copy number variation morbidity map of developmental delay. *Nat. Genet.* *43*, 838–846.
 27. White, P.S., Xie, H.M., Werner, P., Glessner, J., Latney, B., Hakonarson, H., and Goldmuntz, E. (2014). Analysis of chromosomal structural variation in patients with congenital left-sided cardiac lesions. *Birth Defects Res. A Clin. Mol. Teratol.* *100*, 951–964.
 28. Gai, X., Xie, H.M., Perin, J.C., Takahashi, N., Murphy, K., Wenocur, A.S., D'arcy, M., O'Hara, R.J., Goldmuntz, E., Grice, D.E., et al. (2012). Rare structural variation of synapse and neurotransmission genes in autism. *Mol. Psychiatry* *17*, 402–411.
 29. Shaikh, T.H., Gai, X., Perin, J.C., Glessner, J.T., Xie, H., Murphy, K., O'Hara, R., Casalunovo, T., Conlin, L.K., D'Arcy, M., et al. (2009). High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res.* *19*, 1682–1690.
 30. Nowotschin, S., Liao, J., Gage, P.J., Epstein, J.A., Campione, M., and Morrow, B.E. (2006). Tbx1 affects asymmetric cardiac morphogenesis by regulating Pitx2 in the secondary heart field. *Development* *133*, 1565–1573.
 31. Edelman, L., Pandita, R.K., and Morrow, B.E. (1999). Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. *Am. J. Hum. Genet.* *64*, 1076–1086.
 32. Wu, X., and Freeze, H.H. (2002). GLUT14, a duplicon of GLUT3, is specifically expressed in testis as alternative splice forms. *Genomics* *80*, 553–557.
 33. Simpson, I.A., Dwyer, D., Malide, D., Moley, K.H., Travis, A., and Vannucci, S.J. (2008). The facilitative glucose transporter GLUT3: 20 years of distinction. *Am. J. Physiol. Endocrinol. Metab.* *295*, E242–E253.
 34. Shepherd, P.R., Gould, G.W., Colville, C.A., McCoid, S.C., Gibbs, E.M., and Kahn, B.B. (1992). Distribution of GLUT3 glucose transporter protein in human tissues. *Biochem. Biophys. Res. Commun.* *188*, 149–154.
 35. Abel, E.D. (2004). Glucose transport in the heart. *Front. Biosci.* *9*, 201–215.
 36. Grover-McKay, M., Walsh, S.A., and Thompson, S.A. (1999). Glucose transporter 3 (GLUT3) protein is present in human myocardium. *Biochim. Biophys. Acta* *1416*, 145–154.
 37. Kong, B., Liu, Y.L., and Lü, X.D. (2008). Microarray-bioinformatics analysis of altered genomic expression profiles between human fetal and infant myocardium. *Chin. Med. J. (Engl.)* *121*, 1257–1264.
 38. Fowden, A.L., Forhead, A.J., Coan, P.M., and Burton, G.J. (2008). The placenta and intrauterine programming. *J. Neuroendocrinol.* *20*, 439–450.
 39. Ganguly, A., McKnight, R.A., Raychaudhuri, S., Shin, B.C., Ma, Z., Moley, K., and Devaskar, S.U. (2007). Glucose transporter isoform-3 mutations cause early pregnancy loss and fetal growth restriction. *Am. J. Physiol. Endocrinol. Metab.* *292*, E1241–E1255.
 40. Janzen, C., Lei, M.Y., Cho, J., Sullivan, P., Shin, B.C., and Devaskar, S.U. (2013). Placental glucose transporter 3 (GLUT3) is up-regulated in human pregnancies complicated by late-onset intrauterine growth restriction. *Placenta* *34*, 1072–1078.
 41. Lesage, J., Hahn, D., Léonhardt, M., Blondeau, B., Bréant, B., and Dupouy, J.P. (2002). Maternal undernutrition during late gestation-induced intrauterine growth restriction in the rat is associated with impaired placental GLUT3 expression, but does not correlate with endogenous corticosterone levels. *J. Endocrinol.* *174*, 37–43.
 42. Schmidt, S., Hommel, A., Gawlik, V., Augustin, R., Junicke, N., Florian, S., Richter, M., Walther, D.J., Montag, D., Joost, H.G., and Schürmann, A. (2009). Essential role of glucose transporter GLUT3 for post-implantation embryonic development. *J. Endocrinol.* *200*, 23–33.
 43. Carayannopoulos, M.O., Xiong, F., Jensen, P., Rios-Galdamez, Y., Huang, H., Lin, S., and Devaskar, S.U. (2014). GLUT3 gene expression is critical for embryonic growth, brain development and survival. *Mol. Genet. Metab.* *111*, 477–483.
 44. Laurent, L.C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J.V., Lee, S., Barrero, M.J., et al. (2011). Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell* *8*, 106–118.
 45. Vittori, A., Breda, C., Repici, M., Orth, M., Roos, R.A., Outeiro, T.F., Giorgini, F., and Hollox, E.J.; REGISTRY investigators of the European Huntington's Disease Network (2014). Copy-number variation of the neuronal glucose transporter gene SLC2A3 and age of onset in Huntington's disease. *Hum. Mol. Genet.* *23*, 3129–3137.
 46. Veal, C.D., Reekie, K.E., Lorentzen, J.C., Gregersen, P.K., Padyukov, L., and Brookes, A.J. (2014). A 129-kb deletion on chromosome 12 confers substantial protection against rheumatoid arthritis, implicating the gene SLC2A3. *Hum. Mutat.* *35*, 248–256.
 47. Soemedi, R., Wilson, I.J., Bentham, J., Darlay, R., Töpf, A., Zelenika, D., Cosgrove, C., Setchfield, K., Thornborough, C.,

- Granados-Riveron, J., et al. (2012). Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *Am. J. Hum. Genet.* *91*, 489–501.
48. Girirajan, S., Rosenfeld, J.A., Cooper, G.M., Antonacci, F., Siswara, P., Itsara, A., Vives, L., Walsh, T., McCarthy, S.E., Baker, C., et al. (2010). A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat. Genet.* *42*, 203–209.
49. Erdogan, F., Larsen, L.A., Zhang, L., Tümer, Z., Tommerup, N., Chen, W., Jacobsen, J.R., Schubert, M., Jurkatis, J., Tzschach, A., et al. (2008). High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. *J. Med. Genet.* *45*, 704–709.
50. Lalani, S.R., Shaw, C., Wang, X., Patel, A., Patterson, L.W., Kolodziejska, K., Szafranski, P., Ou, Z., Tian, Q., Kang, S.H., et al. (2013). Rare DNA copy number variants in cardiovascular malformations with extracardiac abnormalities. *Eur. J. Hum. Genet.* *21*, 173–181.
51. 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. (2012). Human gene copy number spectra analysis in congenital heart malformations. *Physiol. Genomics* *44*, 518–541.
52. 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. (2009). De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat. Genet.* *41*, 931–935.
53. 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. (2012). Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet.* *8*, e1002843.

The American Journal of Human Genetics

Supplemental Data

**Copy-Number Variation of the Glucose Transporter
Gene *SLC2A3* and Congenital Heart Defects
in the 22q11.2 Deletion Syndrome**

Elisabeth E. Mlynarski, Molly B. Sheridan, Michael Xie, Tingwei Guo, Silvia E. Racedo, Donna M. McDonald McGinn, Xiaowu Gai, Eva W.C. Chow, Jacob Vorstman, Ann Swillen, Koen Devriendt, Jeroen Breckpot, Maria Cristina Digilio, Bruno Marino, Bruno Dallapiccola, Nicole Philip, Tony J. Simon, Amy E. Roberts, Małgorzata Piotrowicz, Carrie E. Bearden, Stephan Eliez, Doron Gothelf, Karlene Coleman, Wendy R. Kates, Marcella Devoto, Elaine Zackai, Damian Heine-Suñer, Tamim H. Shaikh, Anne S. Bassett, Elizabeth Goldmuntz, Bernice E. Morrow, Beverly S. Emanuel and the International Chromosome 22q11.2 Consortium

Figure S1:

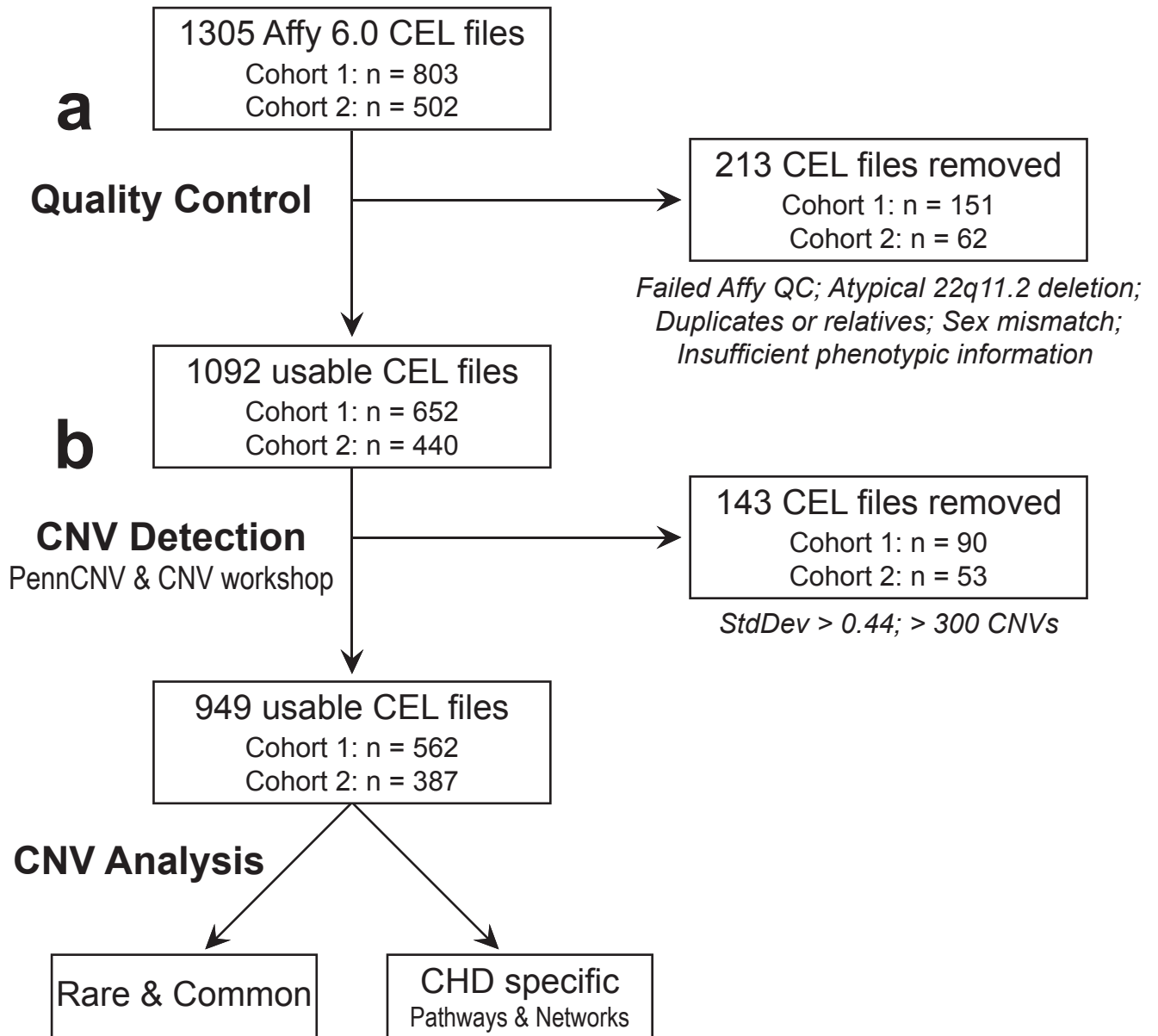


Figure S1: CNV detection and analysis. Flow chart representing our CNV detection and analysis and the number of .CEL files from each cohort included at each step. (a) initial assessment of .cel file quality. (b) CNV detection and quality control.

Table S1: 22q11.2 Deletion Sizes

	First Cohort ^a			Second Cohort ^b			Cohorts Combined ^c		
	Total (n = 562) # %	CHD (n = 363) # %	No CHD (n = 199) # %	Total (n = 387) # %	CHD (n = 240) # %	No CHD (n = 147) # %	Total (n = 949) # %	CHD (n = 603) # %	No CHD (n = 346) # %
A-D	534 95.0	343 94.5	191 96.0	358 92.5	221 92.1	137 93.2	892 94.0	564 93.5	328 94.8
A-C	7 1.3	5 1.4	2 1.0	8 2.1	5 2.1	3 2.0	15 1.6	10 1.7	5 1.4
A-B	21 3.7	15 4.1	6 3.0	21 5.4	14 5.8	7 4.8	42 4.4	29 4.8	13 3.8

^a The frequency of the 22q11.2 deletion sizes in first cohort subjects with CHDs, without CHDs and combined.

^b The frequency of the 22q11.2 deletion sizes in second cohort subjects with CHDs, without CHDs and combined.

^c Overall 22q11.2 deletion frequencies with the first and second cohorts combined.

Table S2: Validated CNVs (GRCh37/hg19 genome build)

Sample	Phenotype	Chr	Start (bp)	End (bp)	Size (bp)	Type	# of Probes	CNV Detection Programs	Inherited	Gene(s)
7462	CHD	1	228,807,528	228,957,150	149,622	Duplication	86	CNV Workshop & PennCNV	Not tested	RHOJ
5471	CHD	2	27,263,122	27,648,118	384,996	Duplication	158	CNV Workshop & PennCNV	Not tested	EMILIN1, TRIM54
6504	CHD	2	32,627,998	33,331,779	703,781	Duplication	453	CNV Workshop & PennCNV	Not tested	LTBP1
5373	CHD	2	32,630,549	33,331,779	701,230	Duplication	452	CNV Workshop & PennCNV	Not tested	LTBP1
6719	CHD	2	113,926,885	114,236,566	309,681	Duplication	150	CNV Workshop & PennCNV	Not tested	PAX8
4160	CHD	3	7,615,567	8,156,424	540,857	Duplication	503	CNV Workshop & PennCNV	Not tested	GRM7
6501	CHD	5	179,612,667	179,883,302	270,635	Duplication	175	CNV Workshop & PennCNV	Not tested	MAPK9
5387	No CHD	9	8,748,942	8,809,841	60,899	Deletion	61	CNV Workshop & PennCNV	Not tested	PTPRD
4212	No CHD	9	9,798,243	9,822,054	23,811	Deletion	57	CNV Workshop & PennCNV	Yes	PTPRD
4251	No CHD	9	9,847,865	9,966,099	118,234	Deletion	144	CNV Workshop & PennCNV	Yes	PTPRD
3935	CHD	9	80,063,217	80,086,725	23,508	Deletion	23	CNV Workshop & PennCNV	Yes	GNA14
6516	CHD	9	119,411,347	119,659,873	248,526	Deletion	267	CNV Workshop & PennCNV	Not tested	ASTN2;TRIM32
7542	No CHD	10	53,383,564	53,404,882	21,318	Deletion	18	CNV Workshop & PennCNV	Yes	PRKG1
5456	No CHD	10	79,250,577	79,370,519	119,942	Deletion	98	CNV Workshop & PennCNV	Not tested	KCNMA1
3920	CHD	12	8,004,426	8,123,238	118,812	Duplication	59	CNV Workshop & PennCNV	Not tested	SLC2A3
3943	CHD	12	7,994,787	8,123,900	129,113	Duplication	25	CNV Workshop & PennCNV	Not tested	SLC2A3
4248	CHD	12	8,004,426	8,130,958	126,532	Duplication	61	CNV Workshop & PennCNV	Yes	SLC2A3
5355	No CHD	12	8,004,425	8,123,900	119,475	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
5368	CHD	12	8,004,425	8,135,092	130,667	Duplication	64	CNV Workshop & PennCNV	Not tested	SLC2A3
5377	No CHD	12	8,004,425	8,123,900	119,475	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
5393	CHD	12	8,004,425	8,123,900	119,475	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
5646	CHD	12	8,004,425	8,123,900	119,475	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
5647	CHD	12	8,004,425	8,123,900	119,475	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
5650	CHD	12	8,004,425	8,123,900	119,475	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
6498	CHD	12	7,971,035	8,143,483	172,448	Duplication	69	CNV Workshop & PennCNV	Yes	SLC2A3
6522	CHD	12	8,004,426	8,123,901	119,475	Duplication	59	CNV Workshop & PennCNV	Not tested	SLC2A3
6696	CHD	12	7,958,406	8,130,958	172,552	Duplication	72	CNV Workshop & PennCNV	Not tested	SLC2A3
6708	CHD	12	8,004,426	8,123,901	119,475	Deletion	60	CNV Workshop & PennCNV	Yes	SLC2A3
6709	CHD	12	7,962,762	8,123,901	161,139	Duplication	69	CNV Workshop & PennCNV	Yes	SLC2A3
6953	CHD	12	7,994,787	8,123,900	129,113	Duplication	61	CNV Workshop & PennCNV	Yes	SLC2A3
6960	CHD	12	7,994,787	8,123,900	129,113	Duplication	56	CNV Workshop & PennCNV	Yes	SLC2A3
6967	CHD	12	7,994,787	8,150,607	155,820	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
7530	CHD	12	7,994,787	8,123,900	129,113	Duplication	59	CNV Workshop & PennCNV	Yes	SLC2A3
7591	CHD	12	8,006,507	8,134,669	128,162	Duplication	61	CNV Workshop & PennCNV	Not tested	SLC2A3
7657	CHD	12	8,016,749	8,090,715	73,966	Duplication	26	CNV Workshop & PennCNV	Not tested	SLC2A3

10357	No CHD	12	7,994,787	8,123,900	129,113	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
10553	CHD	12	7,970,721	8,130,958	160,237	Duplication	65	CNV Workshop & PennCNV	Yes	SLC2A3
10638	CHD	12	7,994,787	8,137,722	142,935	Duplication	65	CNV Workshop & PennCNV	Not tested	SLC2A3
10717	CHD	12	8,004,425	8,135,092	130,667	Duplication	63	CNV Workshop & PennCNV	Yes	SLC2A3
10904	CHD	12	7,994,787	8,123,900	129,113	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
11025	CHD	12	8,004,425	8,123,900	119,475	Duplication	59	PennCNV	Not tested	SLC2A3
11159	CHD	12	7,994,787	8,137,722	142,935	Duplication	65	CNV Workshop & PennCNV	Yes	SLC2A3
11223	CHD	12	7,994,787	8,123,900	129,113	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
11258	CHD	12	7,994,787	8,123,900	129,113	Duplication	60	CNV Workshop & PennCNV	Yes	SLC2A3
11292	CHD	12	7,994,787	8,123,900	129,113	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
11316	No CHD	12	8,004,425	8,123,900	119,475	Duplication	59	PennCNV	Yes	SLC2A3
11322	CHD	12	8,004,425	8,123,900	119,475	Duplication	59	CNV Workshop & PennCNV	Not tested	SLC2A3
11323	CHD	12	8,004,425	8,123,900	119,475	Duplication	59	CNV Workshop & PennCNV	Not tested	SLC2A3
11327	CHD	12	7,994,787	8,123,900	129,113	Duplication	60	CNV Workshop & PennCNV	Not tested	SLC2A3
11336	CHD	12	8,004,425	8,123,900	119,475	Duplication	59	PennCNV	Not tested	SLC2A3
11342	No CHD	12	7,994,787	8,123,900	129,113	Duplication	60	CNV Workshop & PennCNV	Yes	SLC2A3
11960	CHD	12	7,994,787	8,123,900	129,113	Duplication	60	CNV Workshop & PennCNV	Yes	SLC2A3
11964	CHD	12	7,994,787	8,123,900	129,113	Duplication	60	CNV Workshop & PennCNV	Yes	SLC2A3
11971	No CHD	12	7,994,787	8,123,900	129,113	Duplication	60	PennCNV	Yes	SLC2A3
12000	CHD	12	8,007,799	8,123,900	116,101	Duplication	57	CNV Workshop & PennCNV	Not tested	SLC2A3
5402	CHD	13	110,593,018	111,185,674	592,656	Duplication	570	CNV Workshop & PennCNV	Not tested	COL4A1
6502	CHD	14	64,731,581	65,007,547	275,966	Duplication	184	CNV Workshop & PennCNV	Not tested	MTHFD1
5496	No CHD	14	73,018,692	73,188,701	170,009	Duplication	125	CNV Workshop & PennCNV	Not tested	DPF3
5491	CHD	15	22,673,400	23,478,050	804,650	Duplication	322	CNV Workshop & PennCNV	Not tested	15q11.2
6948	CHD	15	22,681,251	23,487,547	806,296	Duplication	323	CNV Workshop & PennCNV	Yes	15q11.2
6696	CHD	15	22,751,681	23,282,812	531,131	Deletion	308	CNV Workshop & PennCNV	Yes	15q11.2
4194	CHD	16	14,835,068	16,396,512	1,561,444	Deletion	679	CNV Workshop & PennCNV	Yes	MYH11
6501	CHD	16	82,714,047	82,765,607	51,560	Deletion	47	CNV Workshop & PennCNV	Not tested	CDH13
6700	CHD	16	83,308,001	83,353,031	45,030	Duplication	59	CNV Workshop & PennCNV	Yes	CDH13
3935	CHD	17	34,816,256	36,295,000	1,478,744	Duplication	996	CNV Workshop & PennCNV	Yes	17q12
5482	CHD	17	72,629,123	73,413,449	784,326	Duplication	382	CNV Workshop & PennCNV	Not tested	GRB2
4178	CHD	18	23,634,475	28,940,034	5,305,559	Deletion	3360	CNV Workshop & PennCNV	Not tested	18q11.2
4165	CHD	21	18,660,957	19,220,122	559,165	Duplication	374	CNV Workshop & PennCNV	Not tested	CXADR
6963	CHD	21	18,736,782	19,124,700	387,918	Duplication	295	CNV Workshop & PennCNV	Not tested	CXADR
5442	CHD	21	35,722,995	35,903,830	180,835	Duplication	109	CNV Workshop & PennCNV	Yes	RCAN1

Table S3: Gene Ontology - Common CNVs

A: All CNV types (17 statistically significant terms)

GO ID	GO Term Description	GO Term Type	CHD (#)	No CHD (#)	p Value	padjust_BH	FDR
GO:0005996	monosaccharide metabolic process	biological process	50	7	2.65E-05	0.032536	0.0540
GO:0005975	carbohydrate metabolic process	biological process	78	17	2.83E-05	0.032536	0.0275
GO:0044723	single-organism carbohydrate metabolic process	biological process	75	17	6.80E-05	0.039545	0.0507
GO:0007275	multicellular organismal development	biological process	45	7	0.000149	0.039545	0.0488
GO:0005355	glucose transmembrane transporter activity	molecular function	43	7	0.000292	0.039545	0.0166
GO:0006766	vitamin metabolic process	biological process	43	7	0.000292	0.039545	0.0166
GO:0006767	water-soluble vitamin metabolic process	biological process	43	7	0.000292	0.039545	0.0166
GO:0008643	carbohydrate transport	biological process	43	7	0.000292	0.039545	0.0166
GO:0008645	hexose transport	biological process	43	7	0.000292	0.039545	0.0166
GO:0015144	carbohydrate transmembrane transporter activity	molecular function	43	7	0.000292	0.039545	0.0166
GO:0015145	monosaccharide transmembrane transporter activity	molecular function	43	7	0.000292	0.039545	0.0166
GO:0015149	hexose transmembrane transporter activity	molecular function	43	7	0.000292	0.039545	0.0166
GO:0015749	monosaccharide transport	biological process	43	7	0.000292	0.039545	0.0166
GO:0015758	glucose transport	biological process	43	7	0.000292	0.039545	0.0166
GO:0019852	L-ascorbic acid metabolic process	biological process	43	7	0.000292	0.039545	0.0166
GO:0051119	sugar transmembrane transporter activity	molecular function	43	7	0.000292	0.039545	0.0166
GO:1901476	carbohydrate transporter activity	molecular function	43	7	0.000292	0.039545	0.0166

B: Duplications (28 statistically significant terms)

GO ID	GO Term Description	GO Term Type	CHD (#)	No CHD (#)	p Value	padjust_BH	FDR
GO:0044723	single-organism carbohydrate metabolic process	biological process	41	3	4.13E-06	0.003262	0
GO:0005975	carbohydrate metabolic process	biological process	44	4	5.68E-06	0.003262	0
GO:0005996	monosaccharide metabolic process	biological process	40	3	6.21E-06	0.003262	0
GO:0055085	transmembrane transport	biological process	36	3	3.11E-05	0.004497	0.0005
GO:0044281	small molecule metabolic process	biological process	46	6	3.55E-05	0.004497	0.0004
GO:0005355	glucose transmembrane transporter activity	molecular function	34	3	6.85E-05	0.004497	0.0016
GO:0006766	vitamin metabolic process	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0006767	water-soluble vitamin metabolic process	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0007275	multicellular organismal development	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0007276	gamete generation	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0007283	spermatogenesis	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0008643	carbohydrate transport	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0008645	hexose transport	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0015144	carbohydrate transmembrane transporter activity	molecular function	34	3	6.85E-05	0.004497	0.0016

GO:0015145	monosaccharide transmembrane transporter activity	molecular function	34	3	6.85E-05	0.004497	0.0016
GO:0015149	hexose transmembrane transporter activity	molecular function	34	3	6.85E-05	0.004497	0.0016
GO:0015749	monosaccharide transport	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0015758	glucose transport	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0019752	carboxylic acid metabolic process	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0019852	L-ascorbic acid metabolic process	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0030154	cell differentiation	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0048232	male gamete generation	biological process	34	3	6.85E-05	0.004497	0.0016
GO:0051119	sugar transmembrane transporter activity	molecular function	34	3	6.85E-05	0.004497	0.0016
GO:1901476	carbohydrate transporter activity	molecular function	34	3	6.85E-05	0.004497	0.0016
GO:0006082	organic acid metabolic process	biological process	36	4	0.000123	0.007459	0.0020
GO:0043436	oxoacid metabolic process	biological process	36	4	0.000123	0.007459	0.0020
GO:0048869	cellular developmental process	biological process	35	4	0.000178	0.010416	0.0025
GO:0044710	single-organism metabolic process	biological process	80	22	0.000483	0.027198	0.0071

C: Deletions (none are significant - top 5 are shown)

GO ID	GO Term Description	GO Term Type	CHD (#)	No CHD (#)	p Value	padjust_BH	FDR
GO:0031424	keratinization	biological process	274	136	0.038366	0.921843	1
GO:0044707	single-multicellular organism process	biological process	346	178	0.044478	0.921843	1
GO:0016021	integral component of membrane	cellular component	277	141	0.069179	0.921843	1
GO:0044430	cytoskeletal part	cellular component	86	39	0.112198	0.921843	1
GO:0044446	intracellular organelle part	cellular component	164	81	0.113582	0.921843	1

p Value = two-tailed Fisher's exact test

padjust_BH = B-H adjusted p Value

FDR = false discovery rate

Table S4: Mammalian Phenotype - Common CNVs

A: All CNV types (none are significant - top 5 are shown)

MP ID	MP Term Description	CHD (#)	No CHD (#)	p Value	padjust_BH	FDR
MP:0011091	complete prenatal lethality	47	8	0.000205	0.086782	0.1470
MP:0001488	increased startle reflex	43	7	0.000292	0.086782	0.0273
MP:0004264	abnormal extraembryonic tissue physiology	43	7	0.000292	0.086782	0.0273
MP:0004265	abnormal placental transport	43	7	0.000292	0.086782	0.0273
MP:0010038	abnormal placenta physiology	43	7	0.000292	0.086782	0.0273

B: Duplications (14 statistically significant terms)

MP ID	MP Term Description	CHD (#)	No CHD (#)	p Value	padjust_BH	FDR
MP:0001488	increased startle reflex	34	3	0.000068	0.011375	0.0013
MP:0003088	abnormal prepulse inhibition	34	3	0.000068	0.011375	0.0013
MP:0004264	abnormal extraembryonic tissue physiology	34	3	0.000068	0.011375	0.0013
MP:0004265	abnormal placental transport	34	3	0.000068	0.011375	0.0013
MP:0009142	decreased prepulse inhibition	34	3	0.000068	0.011375	0.0013
MP:0010038	abnormal placenta physiology	34	3	0.000068	0.011375	0.0013
MP:0011095	complete embryonic lethality between implantation and placentation	34	3	0.000068	0.011375	0.0013
MP:0011091	complete prenatal lethality	37	4	0.000085	0.012442	0.0011
MP:0009850	embryonic lethality between implantation and placentation	36	4	0.000123	0.014403	0.0013
MP:0011090	partial perinatal lethality	36	4	0.000123	0.014403	0.0013
MP:0001363	increased anxiety-related response	34	4	0.000258	0.02158	0.0016
MP:0001486	abnormal startle reflex	34	4	0.000258	0.02158	0.0016
MP:0001961	abnormal reflex	34	4	0.000258	0.02158	0.0016
MP:0004994	abnormal brain wave pattern	34	4	0.000258	0.02158	0.0016

C: Deletions (none are significant - top 5 are shown)

MP ID	MP Term Description	CHD (#)	No CHD (#)	p Value	padjust_BH	FDR
MP:0000278	abnormal myocardial fiber morphology	101	46	0.092162	0.967473	1
MP:0004087	abnormal muscle fiber morphology	101	46	0.092162	0.967473	1
MP:0004937	dilated heart	101	46	0.092162	0.967473	1
MP:0002972	abnormal cardiac muscle contractility	107	50	0.109976	0.967473	1
MP:0000274	enlarged heart	107	50	0.109976	0.967473	1

p Value = two-tailed Fisher's exact test

padjust_BH = B-H adjusted p Value

FDR = false discovery rate

Table S5: qPCR primers used in CNV validation

Primer Name	Primer Sequence 5' -> 3'
RPPH1F (control)	GCGGACGGAAGCTCATCAG
RPPH1R (control)	TCAGACCTTCCCAAGGGACAT
ASTN2e10F	ACTCACCCGCAGAACATGAAG
ASTN2e10R	TGCCCTGCCCTACGAT
ASTN2e12F	CCTGTGGAGAGGAGCAGAGAGA
ASTN2e12R	GCCCTGTAGTCCCTCAGCAGTCT
ASTN2e16F	CCTGAAATGAGCTGGTCATCTG
ASTN2e16R	CATCACCTACCTCTCAGGTTTGC
ASTN2i9F	AGGAGGAGCCCAGGATTAGTG
ASTN2i9R	AAGAACATCTCACATTTTGGGAGAA
ASTN1i16F	TGTTCAAAAACAACAGCACAAAGGT
ASTN1i16R	GATGATGCCTTCTGCCTAAGCT
CDH13e1F	TCCCACGAAAATATGCTCAGT
CDH13e1R	CCGGCGGCGTTTTCA
CDH13i1aF	GTGGGCACCACTTGAGACAGT
CDH13i1aR	GGTTTACCGGCCAGACACA
CDH13i1bF	ACGTGAGCCACCTGTGAA
CDH13i1bR	AGCTTCTGGTGCTGGACAT
CDH13e2F	GAAAGTTAGCAGGGCAAACACA
CDH13e2R	CAGCACCTGAGAAAACCAAAGC
CDH13e5F	CACTGGAAAGGGAGTGGATCA
CDH13e5R	CGCTCCCTGTGTTCTCATTGA
CDH13i5aF	AAAAGAAGAAAGGCAGGATAATCAAT
CDH13i5aR	CCATCCAAATGCAAGGTCATAA
CDH13i5bF	GGGATCCCAGCGAGCATAAC
CDH13i5bR	GGTGGTTGGAGAGGACTTATTCC
CDH13e6F	TGCCTCTGGAAGTCATTGTGAT
CDH13e6R	GGGCCTTCCCAGAAAGATC
COL4A1+205kbF	AGGGCCACAGAGACTTGGAA
COL4A1+205kbR	TGTCCTGACCCCATCTATCA
COL4A1+212kbF	CATCACCTGGAAAAAAGGTCTGA
COL4A1+212kbR	GCCTGGGCTCACTCTCTTCA
COL4A1-222kbF	CCCTGGTGGTGAGGGAAAC
COL4A1-222kbR	AGTGTGGCTGATCAGACATTGG
COL4A1-227kbF	GAGGCCGAGCCCTACA
COL4A1-227kbR	TTGACCGGCACACTCGTACA
COL4A1e49F	CAGCCAGTACGAGTAGTCATTTCG
COL4A1e49R	GCCCTTCTGTTCTGCAATATT
COL4A1e50F	TGGCGTCTCCCAGACA
COL4A1e50R	GGTGGTCTCGCTGTGGAT
COL4A1i49F	CAAAGCAGCAGCTTCAAGTACA
COL4A1i49R	ACACATCGAAGCTCAGCATGA
CXADR-275kbF	GTGGATTGAGCTACCCTGCAA
CXADR-275kbR	CCAGACACCTCGGCTAGTCTTT
CXADR-148kbF	GGCTTAAAATGGTGGCTTTATCC
CXADR-148kbR	GAATAGAATGAAGGAAATTAATGGTTAAGG
CXADRRe6F	GCTTCAAATAAAGCTGGACTAATTGC
CXADRRe6R	TGATAAGACCAATGAGCGCTAGAG
CXADR+274kbF	TCTGACCCCAAGGACCTCTA
CXADR+274kbR	TGCCTTATGCTTTTCCAATTAATC
CXADR+292kbF	GATGCTCTTCCCCTTTGCTT
CXADR+292kbR	GGTAAATCTTTAGATGAGCCTGTTTG
DEFB-347kbF	CTCCTCTGCTCGTGCCTTTC
DEFB-347kbR	GGTGTTCAAGGTCCTGCTCTAG
DEFB-335kbF	AGCCACCATCACTGTCAATCC
DEFB-335kbR	CCCCAGAGCCCAATTCTC
DEFBe1F	TGACGGCCGAGGTGAGA

Primer Name	Primer Sequence 5' -> 3'
DEFBe1R	CCCCGTCGGTATAAGGATGA
DEFB+6kbF	TGCCTTGGTGGCCTCTAGTG
DEFB+6kbR	AACACGCTTGCAGCTCTGACT
DEFB+38kbF	CCCATGTACTCCCTGCAGTATTT
DEFB+38kbR	TTCCGCTCGTCCGTCAA
GNA14i2aF	GAGCATCCTCCAGGGAAAG
GNA14i2aR	CCCCTAAATTTTGTGCCCTAGA
GNA14i2bF	GGAGAATGCTCACGGGTACAG
GNA14i2bR	CCCCACCCAGAGGTTTCAGT
GNA14i2cF	CATTCATCAGGCATGCGTGTA
GNA14i2cR	CTCGAATAATGCCTGGCAAAG
GNA14i2dF	CGAGCACTAAGACCTCCTCGTT
GNA14i2dR	TCCTGTCTGGTGTACCTGCACTT
GRB2-19kbF	CTCTCCTGAGCATACCTGTATTGC
GRB2-19kbR	GGTCGTAATCTACAAAGCGTTTTG
GRB2-231F	TCCCCCGTCTAATCACTTTATTTT
GRB2-231R	GCCCCTACCCATCCAAGT
GRB2e3F	GAAGCCGCTTTTTCCATTAAGC
GRB2e3R	CAGGTTTTGAACGAAGAATGTGAT
GRB2+673kbF	GGGCTGACCATGCTACAGAGA
GRB2+673kbR	GGGAGGCACCAAAGCAGAA
GRB2+695kbF	CTGCTAGAATTGCCCTTTCCA
GRB2+695kbR	GCGCAGGTCCTCTGAAAA
GRM7i7F	GCAAAGATCAGAAAGGATGATTCA
GRM7i7R	CTTGCCAGGAGCAAACAGAT
GRM7e8F	GCCTTCATTCCAATTTTTTTTGG
GRM7e8R	TGATGTGCATTTTCACTTACCTTTTC
GRM7e14F	CCAAAGGAGTGTGAAACTCAA
GRM7e14R	GGCTCTCAGTGGTCCATTCTAAA
GRM7+370kbF	TCAAATCTGGCTTCCAGGAACT
GRM7+370kbR	GAGGAAGATCTAGCCAATGGAAGA
GRM7+373kbF	TGGAGGAGGGAGAAACCATTT
GRM7+373kbR	GAGAATCACCTCCTTTGGATGTCT
KCNMA1e1F	AGAGGCCCCCGAAGAAAGT
KCNMA1e1R	GCATGTGGTGGGCTTTTCC
KCNMA1i1aF	GGGTGTCATTAGGGTCTGTGGAT
KCNMA1i1aR	GGGTATTTGAACCGGGTGTATT
KCNMA1i1bF	GACAGGCATGGTGCTTCCA
KCNMA1i1bR	GCGAGGCTTCTCCAAGTGT
KCNMA1i1cF	CAGGAAGAGCCACCAAGAGTTT
KCNMA1i1cR	AAGGTGCCTGTGGGCTATTAG
LTBP1e2F	CAGCGATCTTGCTTTGTTTCAG
LTBP1e2R	GGCCTTACTCCAGCCATGAC
LTBP1e4F	GACCCAGGAATACGTGCTCAA
LTBP1e4R	GTGGACTGCTCTCCTGAAATCC
LTBP1e3F	CCATGTCAGAATGGAGGGATGT
LTBP1e3R	TGGTCCCTGGTTTACACACACA
LTBP1i3F	CCCTCGGCTTTGAACCAA
LTBP1i3R	AAGGATTCAGCAACTTGAAGGT
LTBP1+550kbF	TGCAGCCTGCCTCATCTG
LTBP1+550kbR	CAACAAACGAGGCATACACTACAAG
MAPK9+56kbF	GGGCGTCAGGTATGCAAAC
MAPK9+56kbR	TTCTCAGGGTTCCTTTCAGTATT
MAPK9-41kbF	CAGGAGGAATCCACAGTGTGTCT
MAPK9-41kbR	CGCCACCCCAACCCTATT
MAPK9e4F	GCTCCATGTGAATAACCTGACATAA
MAPK9e4R	CCTTTTCTCCCTCAGGTATTTGG

Primer Name	Primer Sequence 5' -> 3'
MAPK9-161kbF	TTGACGGCACTAGAAAGCAAAG
MAPK9-161kbR	GCGCGGGCTTCCAAGT
MAPK9-180kbF	GAGTCGGTGGTGAAGCACAA
MAPK9-180kbR	CATGAGAGAGGAAGGCCTGAGA
MTHFD1e1F	GGGTTGGGTTGTCCTGCTT
MTHFD1e1R	CGATGGACACCACCAATATCG
MTHFD1e9F	GTTTTGATTTCTCCCCACTTG
MTHFD1e9R	CACCCACAACTTTTCTCCATT
MTHFD1e27F	TGCATGTCTGTTTACTTTAGTGACGTT
MTHFD1e27R	TGCAACACCAAGATGGCAAA
MTHFD1+85kbF	CAATTTATCAGTGCTGGTGTCA
MTHFD1+85kbR	CATGAAACACCTTAAATGCTCAT
MYH11e1F	CCCAGCCTTCCCAACTC
MYH11e1R	TCACAATGTCGTTGGCTTTTCT
MYH11e42F	ATAACTCTACGCTCCAGACCTTCT
MYH11e42R	AACCCACAGGCGAGGAAAC
MYH11e2F	TGGTCCCTGTGGAATAAGG
MYH11e2R	CTTCCGGGCTTGGTTGTG
PAX8e13F	CGGGAAAGAAAGGCGAGTTT
PAX8e13R	GGCTTGGGCTGTGCTTTG
PAX8e2F	CAGCCCTCCATGGCCTAAG
PAX8e2R	CAAACCTCTACCTGACCCTGACA
PAX8e1F	CTGGGCCCGGTGTCTCT
PAX8e1R	GGTGATGCCGGGTGGAT
PAX8-50kbF	TGCAAATCATACCCGCTTGA
PAX8-50kbR	AATGCTCATTAGCCCTGATTGG
PRKG1i3aF	CCTGGGAGGGTGGAGTATTACA
PRKG1i3aR	CACAAAGGAAAACACCACACACA
PRKG1i3bF	CCTATGCGGGAGTTTGAACCTG
PRKG1i3bR	AAAAGAGCCCTGAACACTAAAAATTT
PRKG1i3cF	AAATGGCTGTACAATGGTACAA
PRKG1i3cR	CGTTTCCAGACAGCGTTTCCAG
PRKG1i3dF	GCTGCCAGTGGGTA AAAAGG
PRKG1i3dR	AATCACGGAAGTCACAAAGCATT
PTPRDe11F	CAGCGAGTCTGTCCGATCTG
PTPRDe11R	CCCCTTTCCAGGCTCTGA
PTPRDi10aF	TGCCAATTTCTTGGCTTT
PTPRDi10aR	AGGAGGGGAGAAGGCCCTACA
PTPRDi10bF	AATGGTGTGAGAGGTGAACAACA
PTPRDi10bR	TTGTGGTGGAGGGGAGCAAAT
PTPRDi10cF	TCTCATTTGCGTTCTTGAAAAA
PTPRDi10cR	GCTGACCGTTTCTTTGTACATGAA
RCAN1e1F	GCATCCTGTTTGGACAGCAA
RCAN1e1R	CCTGTCTGCCTGCAAGCAT
RCAN1e4F	TCGCTGCGTGCAATTCAT
RCAN1e4R	CTGAGTGACCCTGCGATTATTTT
RCAN1-58kbF	GGATTTTCTCTGGTATTGGAA
RCAN1-58kbR	TGCTGAGTTCACACAGTTGCTTAA
RCAN1-170kbF	GGTGGTGGCTGGCTCTGT
RCAN1-170kbR	TGTCCTTCAATTGGACTTTTTGC
RCAN1-163kbF	GAACCAGGCACAGAGAAAGTAGAAG
RCAN1-163kbR	GGGCATTGAGAATCTCTGGTAATT
RCAN1i2F	TGGGCTCAGGAGTGATTTCTCT
RCAN1i2R	TCCCTGTGGCCAAGCAA
RHOU+70kbF	CTGGGAAAAGTGGACGTTGGT
RHOU+70kbR	AGCAAGTATGAAACCCCTCCTAAA
RHOU+80kbF	AGCAGCCTTGCCTATGCA

Primer Name	Primer Sequence 5' -> 3'
RHOU+80kbR	GCTGATTCTGGGAACCAGTTATG
RHOU-62kbF	AGGTAGCAGCAGCCCGATT
RHOU-62kbR	AGTAAAGAATAAGCTGTGGCAGTTCA
RHOU-82kbF	AGGCCTTTTGGTGCTTACCA
RHOU-82kbR	TCTGCCTCTCTCTATTTCAGTCA
RHOUe2F	CGTGAGACTCCAACCTCTGTGACA
RHOUe2R	CCCAGCACTGAGCTGTAACG
SLC2A14-5'F	CATGGCCTTCAGAAGAAGTTAAGAT
SLC2A14-5'R	ACACTGGTTGCTGTAGCAGAAACTA
SLC2A14i7F	AGCATCATCTGTCCAGTCTTCATC
SLC2A14i7R	GCCTAATGTCCAGAGTGGGAAT
SLC2A3e9F	GGCCACAATAAACCCAGGGAAT
SLC2A3e9R	GCTATCTTGGTCTTTGTAGCCTTCTT
SLC2A3e1F	TTTTTCAGCCAACAAAACCTTCA
SLC2A3e1R	TCCTGAGGACGTGGAGAAAACCT
SLC2A3-3'F	CACAACAGAGGCAAGGGTACATAT
SLC2A3-3'R	AAGAAGGAGCAAATGCCAAGTT
TRIM54+102kbF	CCCCTCTCCAGTGTTTTCA
TRIM54+102kbR	GGGTTTGCAGGAGTCTGAGAGT
TRIM54+126kbF	TCTCAGGTGAATCAACGGAATG
TRIM54+126kbR	CCTCTGTATCCATGGCCAGGTA
TRIM54-198kbF	TCTCACAGGTCCTCAAGGTGAA
TRIM54-198kbR	TCACTTCTCCCTCCAGAATCAAA
TRIM54-242kbF	TGAGGGTTGTGGGTCTAAGCTT
TRIM54-242kbR	TGAGTATGCACGTGAGTAGCTGTG
TRIM54-255kbF	GCCATGTGTCCCTCAACTGA
TRIM54-255kbR	GGTGCATGGGAGGTGGAA
TRIM54e4F	TTCTTAGGACCCTTCATGCTTAAAG
TRIM54e4R	GGAGTGCAGCGGCCTAGAG
15q11.2aF	GTGTGCTGGAGAAAGGATGCT
15q11.2aR	CCATCAGGCAGGCAAATAATC
15q11.2bF	CCTGGCTGTGGAAGGTGACT
15q11.2bR	TGTCAGTGACCAAGCCATCTG
15q11.2cF	TCCGCCATGTAAATTACGAGTTC
15q11.2cR	GATGTGGCCCCGACTCT
15q11.2dF	GCAACAAGGCCTGCAGGTAT
15q11.2dR	TGTGTCTCCAAATGGGCAGTT
17q12aF	CATGGCGTCGCTCAAATGTA
17q12aR	TGGGCTTCTCCAAGCAGATC
17q12bF	CCCCCTGCAGCAGCAA
17q12bR	GAGCTAAATTTCCGCAGCAATAA
17q12cR	CCCATCTTCCAGCCCTGAA
17q12cR	TCCTCGCACTCGCCACTAG
17q12dF	GGGCTTGGTTGGTGGGTAA
17q12dR	GAGGGCCAGGGATCATCAC
18q11aF	GGTTGTCCCTGACTCATTGTCA
18q11aR	ACCATTCTGTGCCATCATCACA
18q11.2bF	CACATGCGCTCCACAAGTG
18q11.2bR	GCAGACACACAAATTCAAATATTACA
18q11cF	CCGCTGTGACTGCTTTCCA
18q11cR	AGAACATCATGGTGGCTTTCAA
18q11dF	TGGATAGGCTCATCAGGATCAA
18q11dR	TGCAAACCCACCATGTCTATC
18q11eF	CCATGTGCCTGTCTCTT
18q11eR	CCTGCCAAAGCCCTTTAG
18q11fF	GCAGAATTCGGAAGTGAAGAAGAC
18q11fR	TTGCACGCAAATCCTGTGA