Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Mefford HC, Sharp AJ, Baker C, et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. N Engl J Med 2008;359. DOI: 10.1056/NEJMoa0805384.

SUPPLEMENTARY METHODS

Oligonucleotide array CGH

Rearrangements of 1q21.1 were analyzed utilizing custom oligonucleotide arrays (NimbleGen Systems, Madison, WI). The majority of cases (deletion cases 1-3, 6-15 and duplication cases 1-4) were analyzed on a custom array consisting of 166,000 isothermal probes (length 45-75 bp; mean probe density, 1 probe/130 bp) covering a number of chromosomal regions, including 27,958 probes covering a 4.5-Mb region of chromosome 1q21 (Build 35, chr11:142,000,000-146,500,000). Deletion cases 4, 5, and 16 and duplication cases 5 and 7 were analyzed using a second custom array consisting of 72,000 isothermal probes with 766 probes covering the same region of 1q21 (mean probe density 1 probe/5875 bp). Deletion cases 17-19 and duplication case 8 were analyzed using a third custom array consisting of 130,000 isothermal probes with 1155 probes in the same region based on Build 36 coordinates (chr1:143,500,000-148,000,000). Hybridizations were performed as described previously[1] and utilized a single normal male as a reference (GM15724, Coriell, Camden, NJ). Parent of origin was determined by analysis of microsatellite markers within the minimal deletion region. Insufficient DNA was available for additional oligo array CGH (after screening array CGH as in Table 3) for deletion case 21 and duplication case 6.

Quantitative PCR

Brilliant SYBR Green QPCR Master Mix (Stratagene) was used for quantitative PCR assays in 10µl reactions containing 2.5ng of template DNA from patient and control samples. Gene specific or control primers were used at a final concentration of 200nM with PCR conditions of 95°C denaturation for 10 mins followed by 40 cycles at 95°C for 30 secs, 60°C for 1 min and 72°C for 1 min followed by 1 min at 95°C, 30 secs at 55°C and a final step of 30 secs at 95°C to obtain a dissociation curve. Reactions were carried out in triplicate, and a passive internal reference dye (ROX) was used to correct any differences in reaction volume between wells. Results were analysed using MX3000P comparative quantitation software (Stratagene). Relative quantities of each gene were normalized by comparing to amplification levels of the control CFTR gene in each

sample using identical conditions. PCR amplification efficiencies for each primer pair were calculated from the slope of a standard curve obtained using a serial dilution of template DNA, where PCR Efficiency=$10^{(-1/\text{slope})-1}$.

TaqMan quantitative PCR

Patients from the USA (n=1040) were assayed for copy number of the 1q21.1 region using two *TaqMan* Gene Copy Number Assays. Primers and probes were designed from genomic sequence (Build 36) using Applied Biosystems proprietary software. Each assay was run as a duplex *TaqMan* real-time PCR reaction, utilizing a FAM dye-based assay targeted to 1q21.1 and a VIC dye-based assay for the reference gene, RNase P (PN 4316844 from Applied Biosystems, Foster City, CA). Each PCR assay was performed in quadruplicate and comprised 10 ng gDNA, 1xTaqMan probe/primer mix in 1xTaqMan Universal Master Mix in a 10 µl reaction amplified using an Applied Biosystems 7900HT SDS instrument for 2 mins at 50°C, 10 mins at 95°C, followed by 40 cycles of 15 secs at 95°C and 60 secs at 60°C. Real-time data were collected by the SDS 2.3 software. The method involves relative quantification of the test sequence versus a reference gene known to be two copies for diploid genome. Relative quantity is determined by the ΔΔCt [(FAM Ct- VIC Ct) $_{\text{sample}}$ - (FAM Ct – VIC Ct) $_{\text{calibrator}}$] method, where a reference sample or calibrator known to have two copies of the test sequence is used as the basis for comparative results. Gene copy number is 2× the relative quantity. The two regions assayed were (1) chr1:145,460,047-145,460,130 and (2) chr1:145,679,984-145,680,077 (Build 36 coordinates).

DNA methylation studies

Methylated DNA immunoprecipitation on patient C66 and her mother was performed as described previously.[1] Briefly, 10µg of genomic DNA from patient C66 and her mother were sonicated to a mean fragment size of ~500bp and immunoprecipitated using 5µg monoclonal mouse anti-5-methyl cytidine (Diagenode, Liege, Belgium). Immunoprecipitated DNA was bound to Protein A agarose beads (Invitrogen, Basel, Switzerland), washed and purified by phenol:chloroform extraction. Immunoprecipitated DNA and input (sonicated) DNA were labeled by random-priming with Cy3/Cy5

nonamers,[2] purified and hybridized to the same custom oligonucleotide array (mean density, 1 probe/130 bp) as described above for array CGH (NimbleGen Systems, Madison, WI).

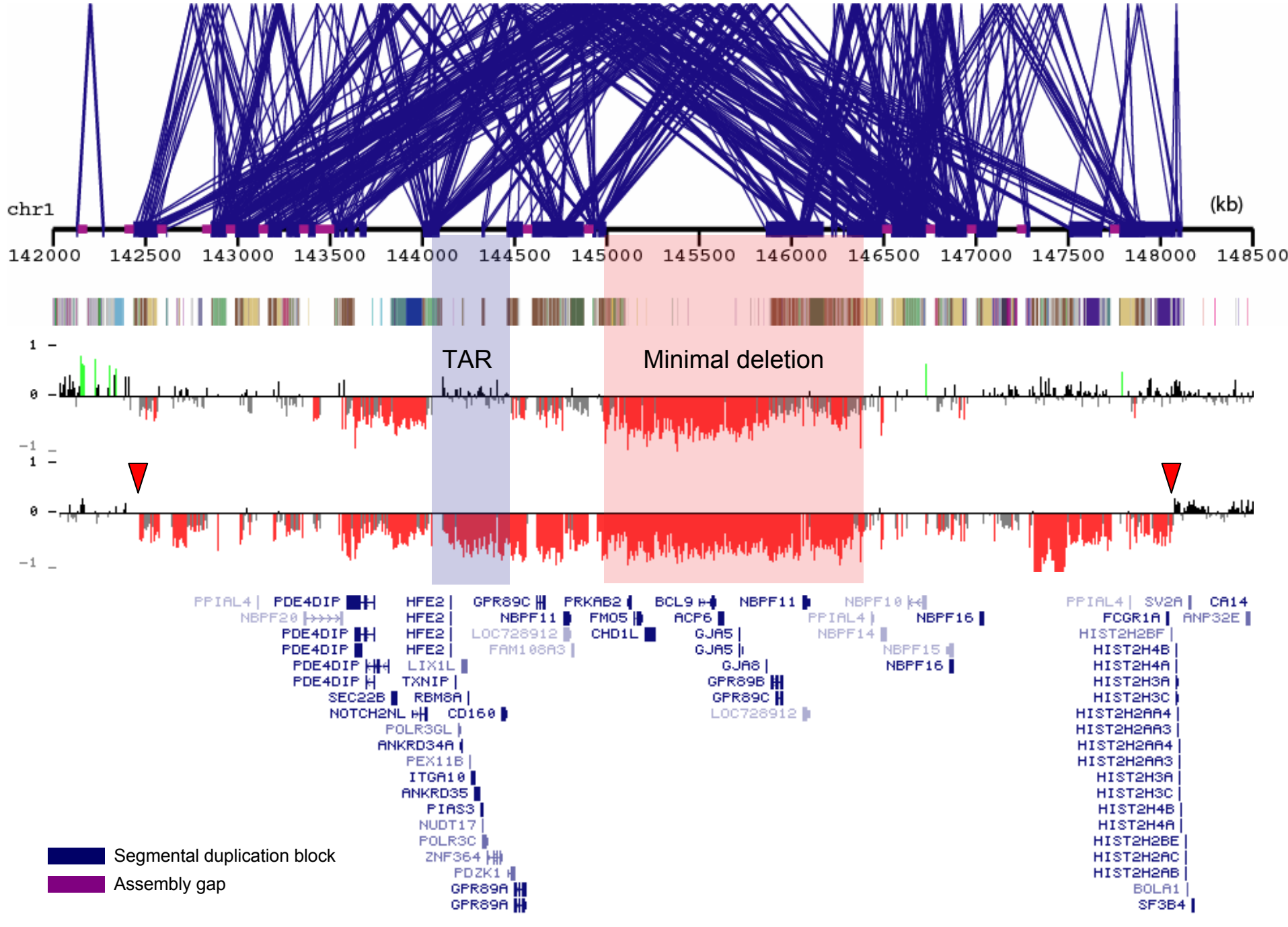
Fluorescence *in situ* hybridization

Metaphase spreads were obtained from lymphoblast immortalized cell lines. FISH experiments were performed using the fosmid ABC14_50192300_M1, directly labeled by nick-translation with Cy3-dUTP (Perkin-Elmer) as previously described[3] with minor modifications. Briefly: 300 ng of labeled probe was hybridized at 37°C in 2 SSC, 50% (v/v) formamide, 10% (w/v) dextran sulphate, with 5µg COT1 DNA (Roche) and 3µg sonicated salmon sperm DNA in a volume of 10µL. Post-hybridization washing was at 60°C in 0.1 SSC (three times, high stringency). Nuclei were simultaneously DAPI stained. Digital images were obtained using a Leica DMRXA2 epifluorescence microscope equipped with a cooled CCD camera (Princeton Instruments). DAPI and Cy3 fluorescence signals, detected with specific filters, were recorded separately as gray scale images. Pseudocoloring and merging of images were performed using Adobe Photoshop software.

Sequence analysis

The coding exon and upstream region of *GJA8* and of *GJA5* were sequenced using standard BigDye terminator chemistry. Primer sequences are available upon request.

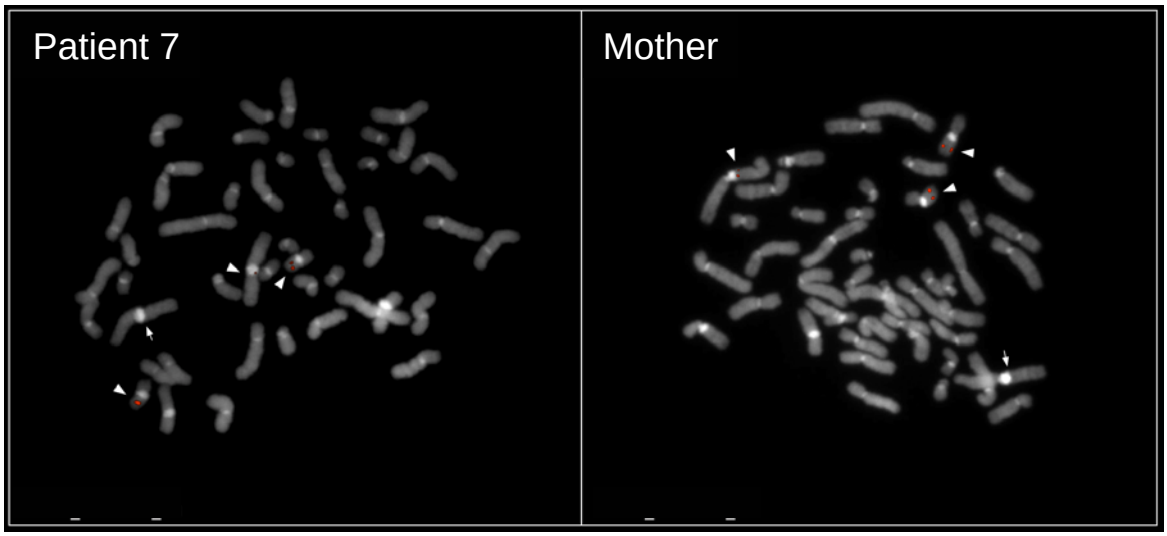
Supplementary Figure 1



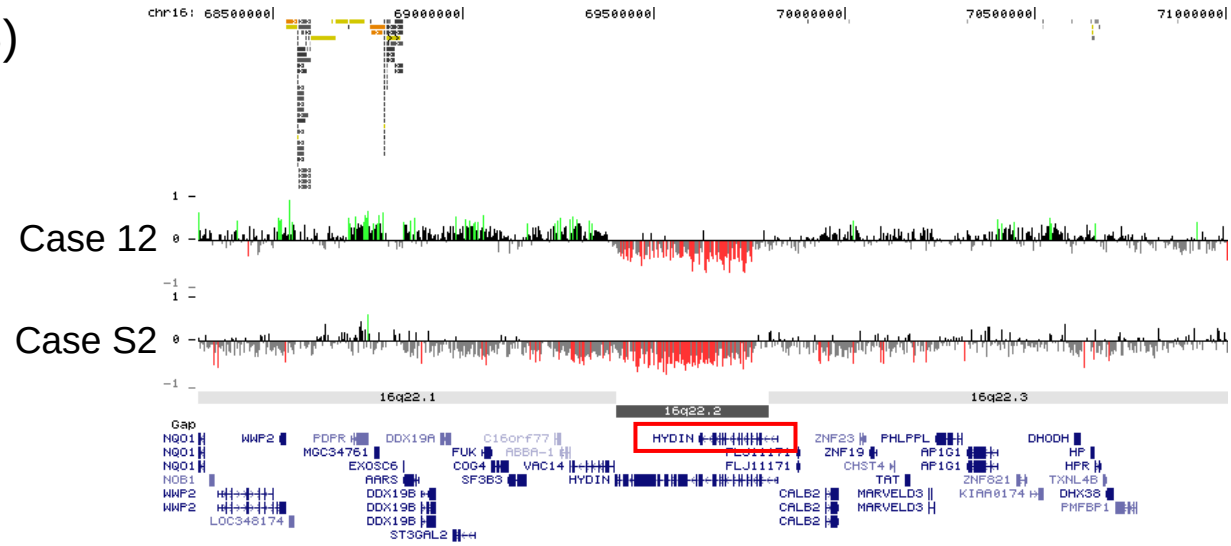
Supplementary Figure 1: Duplication architecture of 1q21 breakpoint regions and large atypical deletion. Build 36, chr1: 142,000,000-148,500,000: paralogy between large (≥10 kb), highly identical (≥95%) intrachromosomal segmental duplications (*blue bars*) are shown as pairwise alignments (*blue lines*). Sequence assembly gaps are shown as purple bars. The underlying duplicon structure is shown, with blocks of identical color representing those that share the same evolutionary origin[4]. Oligo array CGH data (NimbleGen HD2 whole-genome oligo array) for deletion cases 10 and 12 are shown below. The breakpoints of the large deletion in case 12 are indicated by inverted red triangles. Pink shading indicates minimal rearrangement region and blue shading represents the TAR locus as in Figure 1. RefSeq genes are also shown. Note that case 10 appears to have a second deletion proximal to the TAR locus in addition to the minimal 1q21.1 deletion; this is a region of known copy number variation. An alternative explanation is these two apparently separate events actually represent one event, because the structure of 1q21.1 in this individual includes inversion(s), resulting in a structure different than that represented by the reference sequence.

Supplementary Figure 2

A)

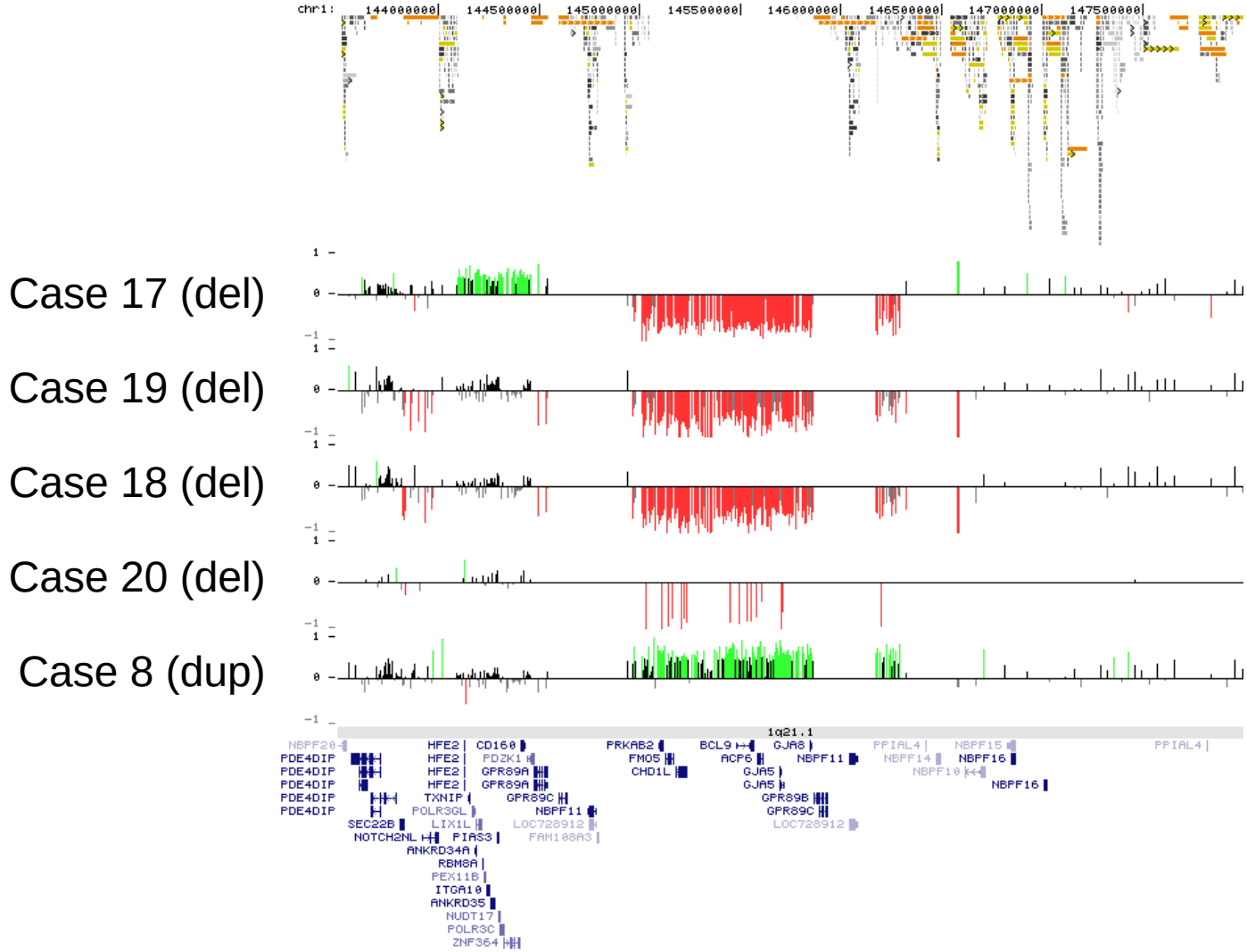


B)



Supplementary Figure 2: A) FISH study in patient 7 and her mother, who each have a 1q21.1 deletion, using a chromosome-16 fosmid (ABC14_50192300_M1) that maps to the *HYDIN* locus. Both individuals have a positive signal on both homologs of chromosome 16 and one homolog of chromosome 1, suggesting that a *HYDIN* homolog maps to the deleted region of 1q21.1 in these individuals. B) Oligonucleotide array CGH data (NimbleGen HD2, 2.1-million oligo whole-genome chip) for two patients with 1q21.1 deletions; for each patient, there is also an apparent deletion on chromosome 16q22 encompassing most of the *HYDIN* gene. FISH validation experiments in three additional 1q21.1 deletion patients who exhibit similar 16q22 “deletions” failed to show deletion of 16q22 (data not shown) since the signal actually represents the deleted copy of *HYDIN2* on 1q21.1.

Supplementary Figure 3



Supplementary Figure 3: Array CGH results for deletion cases 17-20 and duplication case 8 (Build 36, chr1:143,500,000-148,000,000). Data for deletion cases 17-19 and duplication case 8 are from custom NimbleGen arrays (see Supplementary Methods). Data shown for case 20 is from the original screening array (Agilent 44k; see Table 3). Results for deletion case 21 and duplication case 6 are not shown as there was insufficient DNA for additional assays (aside from screening array CGH described in Table 1) and the screening array data is not available in this presentation format.

Supplementary Table 1A. Details on Enrollment and Results of Screening of the 5218 Study Patients, According to Series.*

Series	Primary Center	Source of Referrals	Ascertainment and Inclusion Criteria	Exclusion Criteria	Screening Method†	Detected Cases of Deletion (N=25)
1 (459 patients)	Wellcome Trust Center for Human Genetics, Oxford, United Kingdom	United Kingdom (50% of referrals from clinical geneticists, 35% from community-learning disability teams, and 15% from other physicians)	IMR±MCA (all had mild-to-severe MR with or without other features, normal karyotype and subtelomeric studies, and normal fragile X studies [when indicated])	Normal intellect and abnormal karyotype, abnormal subtelomeric studies, or fragile X syndrome or other known causes of MR	Targeted BAC array	Patient 1
2 (104 patients)	Case Western Reserve University, Cleveland	Case Western Reserve University Hospital and University Hospitals of Cleveland	IMR (all had normal karyotype and normal fragile X studies, when indicated)	Normal intellect and abnormal karyotype, abnormal subtelomeric studies, or fragile X syndrome or other known causes of MR	Targeted BAC array	Patient 14
3 (84 patients)	Vanderbilt University, Nashville	Tufts Medical Center and Vanderbilt University	Autism (probands were recruited from families with two or more family members with autism, as determined with ADI-R and ADOS)	Fragile X syndrome, chromosomal abnormality, IQ <35, and significant dysmorphic features	Targeted BAC array	—
4 (228 patients)	Istituto di Ricovero e Cura a Corattere Scientifico (IRCCS) Associazione Oasi Maria Santissima, Troina, Italy	Clinical geneticists in Sicily (95% of referrals) and mainland Italy (5% of referrals)	MR+DD±mild dysmorphic features (30% of patients) or MR+DD±major dysmorphic features or anomalies (70% of patients) (25% also had autistic features or another neuropsychiatric disorder; all patients had mild-to-severe MR [as defined by DSM-IV-TR criteria], normal karyotype, and normal subtelomeric screening by FISH or MLPA)	Normal intellect and abnormal karyotype, abnormal subtelomeric studies, or fragile X syndrome or other known cause of MR or DD	Agilent 44k oligonucleotide array (61 patients), Agilent 105k oligonucleotide array (60 patients), and Agilent 244k oligonucleotide array (107 patients)	Patients 4, 5

5 (474 patients)	Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands	Clinical geneticists in southeast Netherlands (90% of referrals) and Ireland and United Kingdom (10% of referrals)	Isolated MR (30% of patients) or MR±MCA (70% of patients) (all patients had normal chromosomal G bands after trypsin and Giemsa staining and normal subtelomeric screening by MLPA)	Normal intellect, abnormal karyotype, abnormal subtelomeric studies, and fragile X syndrome or other known causes of MR	32K BAC array	Patient S3
6 (1180 patients)	Centre for Human Genetics, Catholic University, Leuven, Belgium	Clinical geneticists in Belgium	IMR±dysmorphic features (95% of patients) or dysmorphic features + additional anomalies with normal intellect (5% of patients)	Recognizable syndrome and known karyotype abnormality or known cause of MR	1-Mb BAC array	Patients 3, 6, 13
7 (367 patients)	Wessex Regional Genetics Laboratory, Salisbury, United Kingdom	Clinical geneticists in southern England	IMR or DD + MCA (85% of patients) or MCA without MR or in those too young for evaluation for MR (15% of patients) (all patients had normal karyotype and normal fragile X studies [when indicated])	Known karyotype abnormality and fragile X syndrome or other known cause of MR	Custom 44k Agilent oligonucleotide array	Patients 8, 9, 11
8 (187 patients)	University of Manchester, Manchester, United Kingdom	Clinical geneticists in northern England	IMR + other features suggestive of chromosomal abnormality (all patients had mild to severe MR, normal karyotype, and normal subtelomeric studies)	Normal intellect and abnormal karyotype, other known cause of MR, or judgment by a panel of clinical geneticists of probable cause other than chromosomal	BlueGnome Cytochip BAC array or Affymetrix 250k Nsp SNP array	Patient 10
9 (180 patients)	University of Cambridge, Cambridge, United Kingdom	Clinical geneticists in East Anglia, United Kingdom	IMR + congenital anomalies or dysmorphic features (all patients had apparently normal karyotype)	Normal intellect and abnormal karyotype or other known cause of MR	Affymetrix 250k Nsp SNP array	Patients 2, 15, S1, S2
10 (302 patients)	Instituto G. Gaslini, Genoa, Italy, and Università Federico II, Naples, Italy	Neurologists (50% of referrals), clinical geneticists (35% of referrals) and other physicians (15% of referrals) in Italy	IMR±MCA with normal karyotype and normal subtelomeric studies	Normal intellect and abnormal karyotype, abnormal subtelomeric studies, or other known cause of MR	Agilent 44k oligonucleotide array	Patient 7

Supplementary Table 1A. (Continued.)

Series	Primary Center	Source of Referrals	Ascertainment and Inclusion Criteria	Exclusion Criteria	Screening Method†	Detected Cases of Deletion (N=25)
11 (212 patients)	Geneva University Hospital and Inselspital, Bern, Switzerland	Clinical geneticists in Switzerland	IMR + MCA (90% of patients), isolated MR (5% of patients), or autistic spectrum disorder (5% of patients) (all had normal karyotype and normal fragile X studies [when indicated])	Abnormal karyotype, fragile X syndrome, or other known cause of MR	Agilent 244k oligonucleotide array	Patient 16
12 (57 patients)	Irish Autism Genetics Study, Trinity College, Dublin	Clinical geneticists and other physicians in Ireland	Autism (diagnosis by ADI-R and ADOS criteria, IQ >35, or mental age score >18 mo)	Abnormal karyotype, fragile X syndrome, extreme prematurity, infantile rubella, and exposure in utero to medications known to cause autism or other known medical cause of autism	ROMA	—
13 (1040 patients)	Greenwood Genetic Center, Greenwood, SC	South Carolina Department of Disabilities and Special Needs	IMR with normal karyotype and normal fragile X studies (when indicated)	Known cause of MR including fragile X syndrome, abnormal karyotype, and known metabolic disorder	TaqMan quantitative PCR	Patients 12, S4
14 (25 patients)	Pompeu Fabra University, Barcelona	Clinical geneticists in Sabadell, Spain	IMR±MCA (features suggestive of 22q11 deletion syndrome but normal karyotype, normal FISH results for deletion in 22q11, normal subtelomeric studies, and normal fragile X studies)	Abnormal karyotype, presence of del 22q11, CGG-repeat expansions in the fragile X mental retardation 1 gene (fragile X syndrome), and subtelomeric alteration	Targeted BAC array	Patient 17
15 (319 patients)	Centre for Medical Genetics, Ghent University Hospital, Ghent, Belgium	Clinical geneticists in Belgium	IMR±MCA (majority of patients with dysmorphic features)	Abnormal karyotype, fragile X studies, and other known causes of MR	Agilent 44k oligonucleotide array	Patients 18–21

* ADI-R denotes Autism Diagnostic Interview–Revised; ADOS Autism Diagnostic Observation Schedule; DD developmental delay; DSM-IV-TR *Diagnostic and Statistical Manual of Mental Disorders,* fourth edition, text revision; FISH fluorescence in situ hybridization; IMR idiopathic mental retardation; MCA multiple congenital abnormalities; MLPA multiplex ligation-dependent probe amplification; MR mental retardation; PCR polymerase chain reaction; and SNP single-nucleotide polymorphism.

† Screening with the use of the targeted bacterial artificial chromosome (BAC) array as used in series 1, 2, and 3 is described by Sharp et al.[18] and as used in series 14 is described by Cuscó et al.[19] Screening with the use of the 32K BAC array as used in series 5 is described by de Vries et al.[2] Screening with the use of a custom 44k Agilent oligonucleotide array as used in series 7 is described by Barber et al.[1] Screening with the use of representational oligonucleotide microarray analysis (ROMA) as used in series 12 is described by Sebat et al.[20]

Supplementary Table 1B: Phenotypic features of 1q21.1 deletion (n=3) and duplication (n=3) patients from Dutch series*									
Case	**Inheritance**	**Origin**	**Cognitive**	**Growth**	**Facial features**	**Skeletal**	**Heart**	**Eyes**	**Neurologic**
1 (del)	De novo	Unk	MR	Postnatal growth parameters normal	Microcephaly, narrow eyes, periorbital fullness	Normal	Patent ductus arteriosus (surgically repaired)	Normal	No epilepsy; broad-based gait
2 (del)	De novo	Unk	Expressive language delay, language dyspraxia, mild motor delay	Postnatal growth parameters low normal range	Prominent forehead, triangular face, broad nose and mouth	Polydactyly	Normal	Normal	No epilepsy; hypotonia; delayed myelination
3 (del)	Unk	Unk	MR; self-harming behaviors	Postnatal growth parameters low normal range	Straight nose	Normal	Normal	Normal	No epilepsy
4 (dup)	De novo	Unk	IQ within normal range	Mild intrauterine growth retardation; OFC normal at 2 years	Normal	Normal	Perimembranous VSD	Normal	No epilepsy Other: also has hypospadias
5 (dup)	Inh, father normal	P	MR	Normal	Normal	Normal	Normal	Normal	Idiopathic generalized epilepsy
6 (dup)	Unk	Unk	MR	Normal	Normal	Normal	Normal	Normal	Idiopathic generalized epilepsy; hypoplasia of the corpus callosum and cerebellar vermis

* Additional patients analyzed at University Medical Center, Utrecht, The Netherlands included 788 patients with MR/MCA analyzed using the Cytochip BAC array (BlueGnome, Cambridge, UK) and 1-Mb BAC array[5]. Inclusion criteria were: DeVries score >3 (ref. 2)[6], no known karyotype abnormality. Exclusion criteria were autism spectrum disorder, known karyotype abnormality, Fragile X or other known cause of MR.

Inh, inherited; P paternal; Unk, unknown; MR, mental retardation; Ht, height; Wt, weight; OFC, occipitofrontal circumference

Supplementary Table 2: Evaluation of copy number variation in unaffected individuals					
STUDY	**n**	**Evaluation method**	**Coverage of minimal deletion region**	**Deletions detected**	**Duplications detected**
Itsara *et al* (in preparation)*	2063	Illumina Infinium arrays (HumanHap 300 or HumanHap 550)	99 probes (HumanHap 300) 212 probes (HumanHap 550)	0	0
Cohort viii: Manchester, UK	300	SYBR Green quantitative PCR (see Methods)	5 loci within region	0	0
Locke et al (2006), Redon et al, (2006)[7, 8]	209	Targeted BAC array CGH, 32K WGTP BAC array CGH and Affymetrix 5.0	13 clones (targeted BAC); 6 clones (WGTP); 348 probes (Affy 5.0)	0	0
Iafrate et al (2004)[9]	55	1-Mb BAC array CGH	2 clones	0	0
Sebat et al (2004)[10]	20	ROMA	37 probes (1 probe / 35kb)	0	0
Sharp et al (2005)[11]	47	Targeted BAC array CGH	13 clones	0	0
Simon-Sanchez et al (2007)[12]	276	HumanHap 300 and Infinium Human-I	99 probes (HumanHap 300)	0	0
Zogopolous et al (2007)[13]	1190	Affymetrix GeneChip 100K and 500K	50 probes (100K) 221 probes (500K)	0	1**
Pinto et al (2007)[14]	506	Affymetrix 500K SNP array	221 probes	0	0
de Stahl et al (2008)[15]	71	32K BAC array	6 clones	0	0
TOTAL	**4737**			**0**	**1**

* 22% of the individuals in this series are a subset of the Human Genome Diversity Panel (HGDP) reported by Jakobssen et al (ref. 12)[16]; 35% are adult neurologic controls from the NINDS collection screened for neurologic disorders including autism, bipolar disorder and schizophrenia.[12, 17] The remaining 43% individuals are adult self-reported Caucasians from the United States who are enrolled in the PARC study, which aims to identify genetic contributors to the variable efficacy of statin drugs on cardiovascular disease risk (http://www.pharmgkb.org/network/members/parc.jsp#team).

** 4 duplications were reported in this study; however only 1 was validated (L. Feuk, personal communication).

Supplementary Table 3. Phenotypic Features of Probands with a Duplication in Chromosome 1q21.1.*

Patient†	Inheritance	Parental Origin	Cognitive Features	Growth Features	Facial Features	Skeletal Features	Features of the Heart	Features of the Eyes	Neurologic Features
1	Inherited (normal parent and affected brother carry 1q21.1 duplication)	P	Autism; no other data available	Unknown	No notable dysmorphic features	Unknown	Unknown	Unknown	No known seizures
2	Unknown	Unknown	Moderate MR	Height, <3rd percentile for age; OFC, 85th percentile for age	Frontal bossing, downslanting palpebral fissures, epicanthic folds, hypertelorism, high palate, bifid uvula, low ears	Normal	Normal	Rotational nystagmus	Normal
3	De novo	Unknown	MR	Macrocephaly	Frontal bossing, downslanting palpebral fissures, low set ears	Normal	Normal	Normal	Normal
4	Unknown	Unknown	DD, mild learning disability, challenging behavior	Height, <3rd percentile for age; weight, 50th percentile for age; OFC, 90th percentile for age	Small low-set ears, upturned lobes; small mouth and chin; high palate; broad nasal bridge; telecanthus	Joint laxity	Small ventricular septal defect, closed spontaneously (no intervention)	Normal	Small sacral dimple, mild hypotonia
5	Inherited (normal father)	P	Autism, learning disabilities	Height and weight, 5th–10th percentile for age	Normal	Hemivertebra L4	Univentricular heart	Papillary atrophy‡	Seizure disorder
6	Unknown	Unknown	Moderate MR, autistic features, severe speech delay	Height, 75th percentile for age; gross obesity; OFC, 97th percentile for age	Everted lower lip	Normal	Normal	Normal	Normal findings on head CT
7	De novo	Unknown	Autism, borderline mild MR	Height, 90th percentile for age; OFC, >90th percentile for age	Normal	Normal	Normal	Normal	Normal
8	Unknown	Unknown	Psychomotor retardation	Weight, 90th percentile for age; OFC, 97th percentile for age	Hypertelorism	Normal	Normal	Normal	Normal findings on brain MRI

* CT denotes computed tomography, DD developmental delay, M maternal, MR mental retardation, MRI magnetic resonance imaging, OFC occipitofrontal circumference, and P paternal.
† Patient 1 is from series 3, Patients 2 and 3, series 5; Patient 4, series 8; Patients 5 and 6, series 6; Patient 7, series 12; and Patient 8, series 15.
‡ Patient 5 also had a choroid plexus carcinoma for which he had surgery, radiotherapy, and chemotherapy. Papillary atrophy is probably secondary to repeated episodes of increased intracranial hypertension related to the tumor.

Supplementary Table 4: Phenotypic features for patients with 1q21.1 rearrangements and additional chromosomal abnormalities										
Case	**Inheritance**	**Origin**	**Additional imbalances**	**Cognitive**	**Growth**	**Facial features**	**Skeletal**	**Heart**	**Eyes**	**Neurologic**
S1 (del)	De novo	Unk	5q14 deletion (5.3 Mb; de novo)	N.I.	Wt P0.4-2 OFC <P0.4	Hypotelorism, short stub nose	Bilateral post-axial skin tags	Small persistent patent foramen ovale (asymptomatic)	N.I.	Hypotonia
S2 (del)	Inherited	P	5p15.3 duplication (6.1 Mb; inherited from father) der(8)t(8;21)(q24.3;q22.3)	Early delays (sat >12 mo, walked at 20 mo); now progressing	At 2 yrs: Wt P9 Ht P20 OFC <P0.4	Asymmetry of cranial vault, downslanting palpebral fissures, tubular nose	Bilateral 5th finger clinodactyly	N.I.	Strabismus	Hypotonia
S3 (del)	Inherited	P	t(6,9) 13q21-q31 del (6.6 Mb) and 13q31 del (1 Mb)	MR, motor delays	Pre- and post-natal growth retardation; microcephaly	Palatoschisis, ptosis, long eyelashes, high/broad nasal bridge, dysplastic and low set ears micrognathia	Bilateral 5th finger clinodactyly	Atrial septal defect, no intervention required	Strabismus	Ataxia, hypotonia
S4 (del)	Unk; half-siblings and maternal uncles with MR	Unk	*DLG3* mutation	Moderate MR (IQ 41)	Microcephaly	Sloping forehead, L ptosis, wide-spaced teeth	Normal	Murmur, normal echocardiogram	Normal	No seizures
S5* (del)	Unk	Unk	2p16.3 duplication chr2:48,648,754-49,319,683 (671 kb)	Likely normal; achieved GED; honorably discharged from military service	Mildly obese	N.I.	N.I.	N.I.	N.I.	Schizophrenia, onset age 25; severe paranoia, delusions, catatonia
S6 (dup)	De novo	Unk	1q21-q23 duplication (9.4 Mb)	MR, severe with little progress	Normal but poor feeding	Myopathic facies, prominent lower lip	Normal	Normal	Myopia and strabismus	Deafness, delayed milestones

Del, deletion; dup, duplication; Unk, unknown; M, maternal; P, paternal; MR, mental retardation; Ht, height; Wt, weight; OFC, occipitofrontal circumference; N.I., no information; *previously reported by Walsh and colleagues[18]

Supplementary Table 5: All pairwise alignments between breakpoint regions

(Pairwise alignments with longest stretch of highest identity for each pair in bold)

A) Paired duplications between BP1 and BP2

BP1 start	BP1 stop	BP2 start	BP2 stop	Size (kb)	Identity	orientation
142,857,141	**142,927,783**	**144,010,043**	**144,080,196**	**70.6**	**95.52%**	**+**
143,522,082	143,540,019	144,010,043	144,027,990	17.9	98.51%	+
143,534,672	143,552,732	144,074,521	144,092,620	18.1	96.07%	+

B) Paired duplications between BP1 and BP3

BP1 start	BP1 stop	BP3 start	BP3 stop	Size (kb)	Identity	orientation
142,857,141	142,875,105	144,763,589	144,745,644	18.0	97.52%	-
143,308,712	143,333,770	144,786,645	144,761,595	25.1	97.49%	-
143,522,082	143,545,357	144,763,589	144,740,310	23.3	97.98%	-
143,559,134	143,575,976	144,740,310	144,723,753	16.8	96.28%	-
143,594,887	**143,627,819**	**144,963,196**	**144,995,604**	**32.9**	**95.72%**	**+**

C) Paired duplications between BP1 and BP4

BP1 start	BP1 stop	BP4 start	BP4 stop	Size (kb)	Identity	orientation
142,436,049	**142,562,525**	**146,298,638**	**146,425,103**	**126.5**	**99.59%**	**+**
142,535,781	142,562,525	146,542,663	146,572,240	26.7	99.09%	+
142,542,168	142,562,525	146,927,802	146,907,552	20.4	98.94%	-
142,542,168	142,562,525	147,089,711	147,069,563	20.4	99.03%	-
142,857,141	142,875,105	146,060,642	146,042,692	18.0	97.52%	-
142,858,092	142,875,057	147,006,246	147,023,138	17.0	98.44%	+
142,858,092	142,875,058	146,844,161	146,861,095	17.0	98.47%	+
142,859,972	142,881,664	146,492,662	146,471,017	21.7	95.60%	-
142,985,839	143,002,546	147,523,840	147,505,292	16.7	96.49%	-
142,985,839	143,006,944	146,369,533	146,348,380	21.1	97.39%	-
142,985,839	143,113,101	147,909,455	147,782,119	127.3	99.68%	-
143,015,785	143,031,533	146,348,462	146,332,665	15.7	97.94%	-
143,032,741	143,113,101	146,379,171	146,459,532	80.4	99.61%	+
143,051,940	143,113,038	146,542,663	146,606,116	61.1	98.79%	+
143,058,333	143,113,038	146,927,802	146,873,659	54.7	98.71%	-
143,058,333	143,113,038	147,089,711	147,035,687	54.7	98.75%	-
143,163,102	143,182,782	147,526,837	147,505,292	19.7	96.86%	-
143,163,102	143,187,183	146,372,523	146,348,380	24.1	97.17%	-
143,163,102	143,238,033	147,912,434	147,837,397	74.9	99.43%	-
143,196,013	143,211,780	146,348,462	146,332,665	15.8	97.92%	-
143,212,991	143,238,033	146,379,171	146,404,289	25.0	99.08%	+
143,263,101	143,331,520	146,777,983	146,849,231	68.4	98.49%	+
143,281,711	143,292,847	146,286,961	146,278,091	11.1	95.37%	-
143,308,712	143,333,770	146,083,748	146,058,648	25.1	97.51%	-
143,310,744	143,333,770	146,727,982	146,706,108	23.0	98.74%	-
143,321,088	143,331,520	147,000,772	147,011,318	10.4	95.64%	+
143,522,082	143,540,019	146,708,099	146,690,133	17.9	98.21%	-
143,522,082	143,545,357	146,060,642	146,037,358	23.3	97.96%	-
143,523,027	143,542,193	146,844,161	146,863,310	19.2	98.38%	+
143,523,027	143,542,193	147,006,246	147,025,354	19.2	98.40%	+
143,524,924	143,540,019	146,492,662	146,477,559	15.1	98.53%	-
143,559,134	143,575,976	146,037,358	146,020,797	16.8	96.27%	-

D) Paired duplications between BP2 and BP3						
BP2 start	**BP2 stop**	**BP3 start**	**BP3 stop**	**Size (kb)**	**Identity**	**orientation**
144,004,163	144,027,990	144,769,399	144,745,644	23.8	97.84%	-
144,073,367	**144,092,620**	**144,926,035**	**144,945,324**	**19.3**	**95.08%**	**+**

E) Paired duplications between BP2 and BP4						
BP2 start	**BP2 stop**	**BP4 start**	**BP4 stop**	**Size (kb)**	**Identity**	**orientation**
144,004,163	144,027,990	146,066,503	146,042,692	23.8	97.81%	-
144,004,163	144,079,919	146,713,980	146,637,664	75.8	98.19%	-
144,011,009	144,027,942	147,006,246	147,023,138	16.9	98.41%	+
144,011,009	**144,027,943**	**146,844,161**	**146,861,095**	**16.9**	**98.42%**	**+**
144,012,902	144,034,531	146,492,662	146,471,013	21.6	95.65%	-
144,065,165	144,076,718	146,482,897	146,471,341	11.6	97.86%	-

F) Paired duplications between BP3 and BP4						
BP3 start	**BP3 stop**	**BP4 start**	**BP4 stop**	**Size (kb)**	**Identity**	**orientation**
144,322,642	144,337,061	146,017,954	146,003,508	14.4	99.23%	-
144,459,423	144,544,474	145,948,719	145,861,130	85.1	99.12%	-
144,594,476	**144,876,007**	**145,891,442**	**146,173,101**	**281.5**	**99.94%**	**+**
144,743,482	144,762,642	146,863,310	146,844,161	19.2	97.94%	-
144,743,482	144,762,642	147,025,354	147,006,246	19.2	97.99%	-
144,745,644	144,760,751	146,477,559	146,492,662	15.1	97.90%	+
144,745,644	144,784,610	146,690,133	146,727,982	39.0	97.85%	+
144,763,840	144,774,233	147,011,318	147,000,772	10.4	95.80%	-
144,763,840	144,786,645	146,849,231	146,826,314	22.8	96.54%	-
144,927,245	144,937,953	146,048,059	146,037,358	10.7	98.49%	-
144,953,161	144,963,193	146,037,358	146,027,332	10.0	98.12%	-

Supplementary Table 6: Sequencing results							
			GJA5				*GJA8*
			rs1692137	rs791286	rs35594137	rs1692141	rs3766503
Patient	**CHD**	**Cataracts**	**upstream of tss (-32bp)**	**Exon 1 (5' UTR)**	**Intron1**	**3' UTR**	**Exon 2 (synonymous)**
Case 2*	+	+	G	C	A	A	G
Case 3	--	--	G	C	A	A	G
Case 6	--	--	G	C	G	G	G
Case 7	+	--	G	C	G	G	G
Case 8	--	--	G	C	G	G	G
Case 9	--	+	G	C	G	N.D.	G
Case 10	--	--	G	C	G	G	G
Case 11	--	--	G	C	G	G	A
Case 12	--	--	G	C	G	G	G
Case 15	--	--	G	C	G	G	G
Case 16	+	+	G	C	A	A	G

* Case 2 also has an A>C change 12 bp downstream of the stop codon in *GJA8*
All other changes were at sites of known SNPs, as indicated in table.
CHD, congenital heart disease

References

1. Weber M, Davies JJ, Wittig D, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 2005;37(8):853-62.
2. Selzer RR, Richmond TA, Pofahl NJ, et al. Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. Genes Chromosomes Cancer 2005;44(3):305-19.
3. Lichter P, Tang CJ, Call K, et al. High-resolution mapping of human chromosome 11 by in situ hybridization with cosmid clones. Science 1990;247(4938):64-9.
4. Jiang Z, Tang H, Ventura M, et al. Ancestral reconstruction of segmental duplications reveals punctuated cores of human genome evolution. Nat Genet 2007;39(11):1361-8.
5. Vissers LE, de Vries BB, Osoegawa K, et al. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. Am J Hum Genet 2003;73(6):1261-70.
6. de Vries BB, White SM, Knight SJ, et al. Clinical studies on submicroscopic subtelomeric rearrangements: a checklist. J Med Genet 2001;38(3):145-50.
7. Locke DP, Sharp AJ, McCarroll SA, et al. Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. Am J Hum Genet 2006;79(2):275-90.
8. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature 2006;444(7118):444-54.
9. Iafrate AJ, Feuk L, Rivera MN, et al. Detection of large-scale variation in the human genome. Nat Genet 2004;36(9):949-51.
10. Sebat J, Lakshmi B, Troge J, et al. Large-scale copy number polymorphism in the human genome. Science 2004;305(5683):525-8.
11. Sharp AJ, Locke DP, McGrath SD, et al. Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 2005;77(1):78-88.
12. Simon-Sanchez J, Scholz S, Fung HC, et al. Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. Hum Mol Genet 2007;16(1):1-14.
13. Zogopoulos G, Ha KC, Naqib F, et al. Germ-line DNA copy number variation frequencies in a large North American population. Hum Genet 2007;122(3-4):345-53.
14. Pinto D, Marshall C, Feuk L, Scherer SW. Copy-number variation in control population cohorts. Hum Mol Genet 2007;16 Spec No. 2:R168-73.
15. de Stahl TD, Sandgren J, Piotrowski A, et al. Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array. Hum Mutat 2008;29(3):398-408.
16. Jakobsson M, Scholz SW, Scheet P, et al. Genotype, haplotype and copy-number variation in worldwide human populations. Nature 2008;451(7181):998-1003.
17. Schymick JC, Scholz SW, Fung HC, et al. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. Lancet Neurol 2007;6(4):322-8.
18. Walsh T, McClellan JM, McCarthy SE, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 2008;320(5875):539-43.