A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay

Santhosh Girirajan, Jill A Rosenfeld, Gregory M Cooper, Francesca Antonacci, Priscillia Siswara, Andy Itsara, Laura Vives, Tom Walsh, Shane E McCarthy, Carl Baker, Heather C Mefford, Jeffrey M Kidd, Sharon R Browning, Brian L Browning, Diane E Dickel, Deborah L Levy, Blake C Ballif, Kathryn Platky, Darren M Farber, Gordon C Gowans, Jessica J Wetherbee, Alexander Asamoah, David D Weaver, Paul R Mark, Jennifer Dickerson, Bhuwan P Garg, Sara A Ellingwood, Rosemarie Smith, Valerie C Banks, Wendy Smith, Marie T McDonald, Joe J Hoo, Beatrice N French, Cindy Hudson, John P Johnson, Jillian R Ozmore, John B Moeschler, Urvashi Surti, Luis F Escobar, Dima El-Khechen, Jerome L Gorski, Jennifer Kussmann, Bonnie Salbert, Yves Lacassie, Alisha Biser, Donna M McDonald-McGinn, Elaine H Zackai, Matthew A Deardorff, Tamim H Shaikh, Eric Haan, Kathryn L Friend, Marco Fichera, Corrado Romano, Jozef Gécz, Lynn E deLisi, Jonathan Sebat, Mary-Claire King, Lisa G Shaffer & Evan E Eichler

Nature Genetics: doi:10.1038/ng.534

Supplementary Figures, Tables, and Supplemental Note

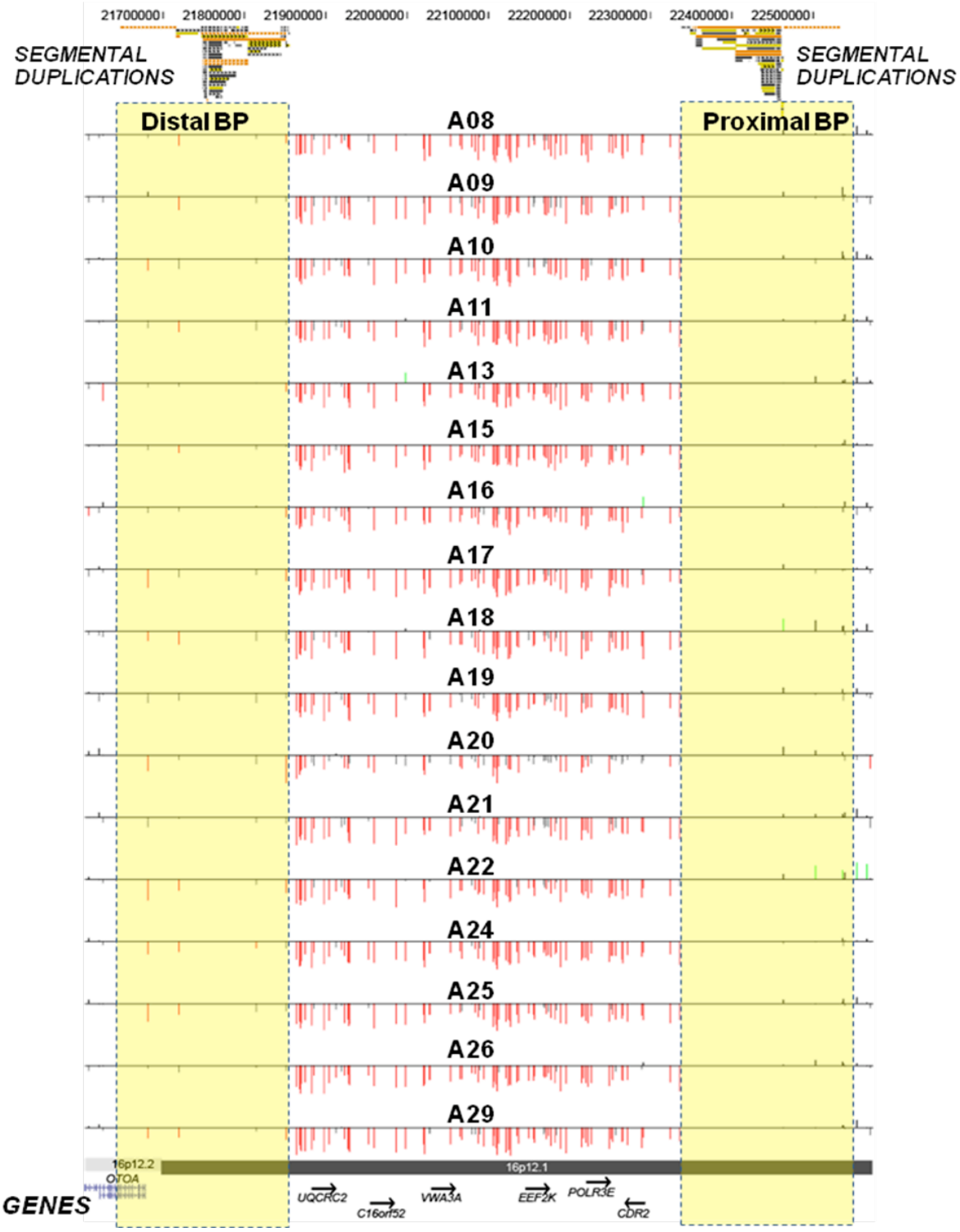
Contents

Supplementary Fig. 1. Representative 16p12.1 microdeletions in the replication set of cases. . 4
Supplementary Fig. 2. Second hits in replication set of 16p12.1 microdeletion. 5
Supplementary Fig. 3. Maternally inherited 16p12.1 microdeletion and a 15q11.2q13.1 duplication (*de novo* or paternally inherited) SG05 6
Supplementary Fig. 4. *De novo* 16p12.1 microdeletion in SG04. 7
Supplementary Table 1. Segmental duplications at 16p12.1 rearrangement breakpoints 8
Supplementary Table 2. Replication set of cases with 16p12.1 microdeletion and associated phenotypes 9
Supplementary Table 3. Clinical features identified in individuals with 16p12.1 microdeletion 10
Supplementary Table 4. Large CNV second hits in individuals with 16p12.1 deletions 14
Supplementary Table 5. Enrichment for 'second hit' CNVs among del 16p12.1 carriers 14
Supplementary Table 5a. Average size and genes in second hit CNVs in cases and controls 14
Supplementary Table 6. Inheritance of the 16p12.1 microdeletion and second hit 15
Supplementary Table 7. Parental ascertainment for 16p12.1 phenotypes 15
Supplementary Table 8. Frequency of clinical features in carrier parents 16
Supplementary Table 9. An excess of 16p12.1 del carrier parents manifesting phenotype 16
1. Supplementary Note: Copy-number variation (CNV) detection 17
1. CNV analysis in affected individuals 17
1.1. CNV analysis in cases with intellectual disabilities and developmental delay 17
Table S1.1. Detection rate of 16p12.1 microdeletion using BAC versus oligo arrays 17
1.2. CNV analysis in cases with schizophrenia 18
2. Controls for 16p12.1 analysis 18
Table S1.2. Composition of schizophrenia cases and controls 19
Table S1.3. Responses to the GAIN control questionnaire 19
3. Platforms and probe coverage for arrays used in cases and controls 23
Table S1.4. Platforms and probe coverage for arrays used for cases and controls 24
4. Validation of 16p12.1 microdeletions and the second hits 25
Table S1.5. High density 16p12.1 microarray design with control regions 25
Fig. S1.1. Validation by high density oligonucleotide arrays 26
2. Supplementary Note: CNV analysis in 480 patients with ID/DD from Italy and Australia 27
Fig. S2.1. Clinical features in 25514 28
Fig. S2.2. 16p12.1 microdeletion in Case #25514. 28
3. Supplementary Note: Comparison of 'two hit' phenotypes with second-hit-associated phenotypes. 29
Table S3.1. Phenotypes associated with homozygous 22q11.2 duplication 29
Table S3.2. Phenotypes associated with 22q13 deletion 30
Table S3.3. Phenotypes associated with 15q11.2q13.1 duplication 31
Table S3.4. Phenotypes associated with 5q15q23.3 duplication 32
4. Supplementary Note: Analysis of a multiplex family (LD1205) 33

Nature Genetics: doi:10.1038/ng.534

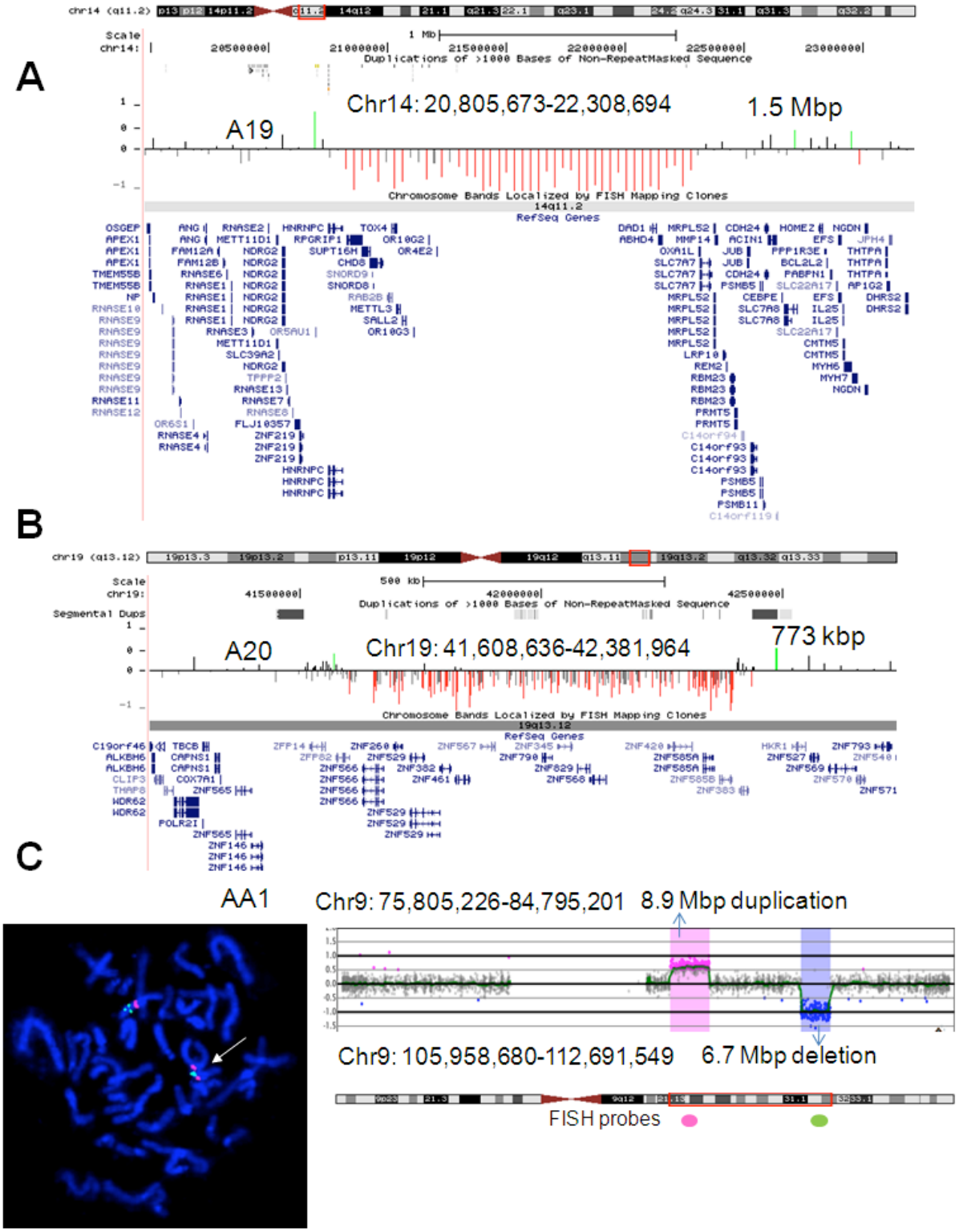
Fig. S4.1. A family with both schizophrenia and cognitive delay. 34
Fig. S4.2. 16p12.1 microdeletion in a multiplex family affected with schizophrenia. 36
5. Supplementary Note: Candidate genes within 16p12.1 microdeletion region..... 37
Table S5.1. RefSeq genes in the 16p12.1 microdeletion region 37
5.1. Candidate gene sequencing 37
Table S5.2. Primer sequences for the candidate gene PCR/sequencing 37
6. Supplementary Note: Haplotype analysis in patients and available parental DNA39
Fig. S6.1. IBD probability in unrelated probands 40
7. Supplementary Note: Analysis for two hits in 1q21.1 deletion patients 41
Table S7.1. Abnormalities in cases with 1q21.1 microdeletion and a second hit 41
Figure S7.1: Large CNVs in 1q21.1 microdeletion patients. 42
8. Supplementary Note: A two-hit model for 16p12.1 phenotypes. 43
9. Supplementary Note: References 44

Nature Genetics: doi:10.1038/ng.534



Supplementary Fig. 1. Representative 16p12.1 microdeletions in the replication set of cases. Probes with $\log_2$ ratios below a threshold of 1.5 standard deviations from the normalized mean $\log_2$ ratio denote deletions (red). The proximal and the distal breakpoints (BP) map to within the segmental duplications.

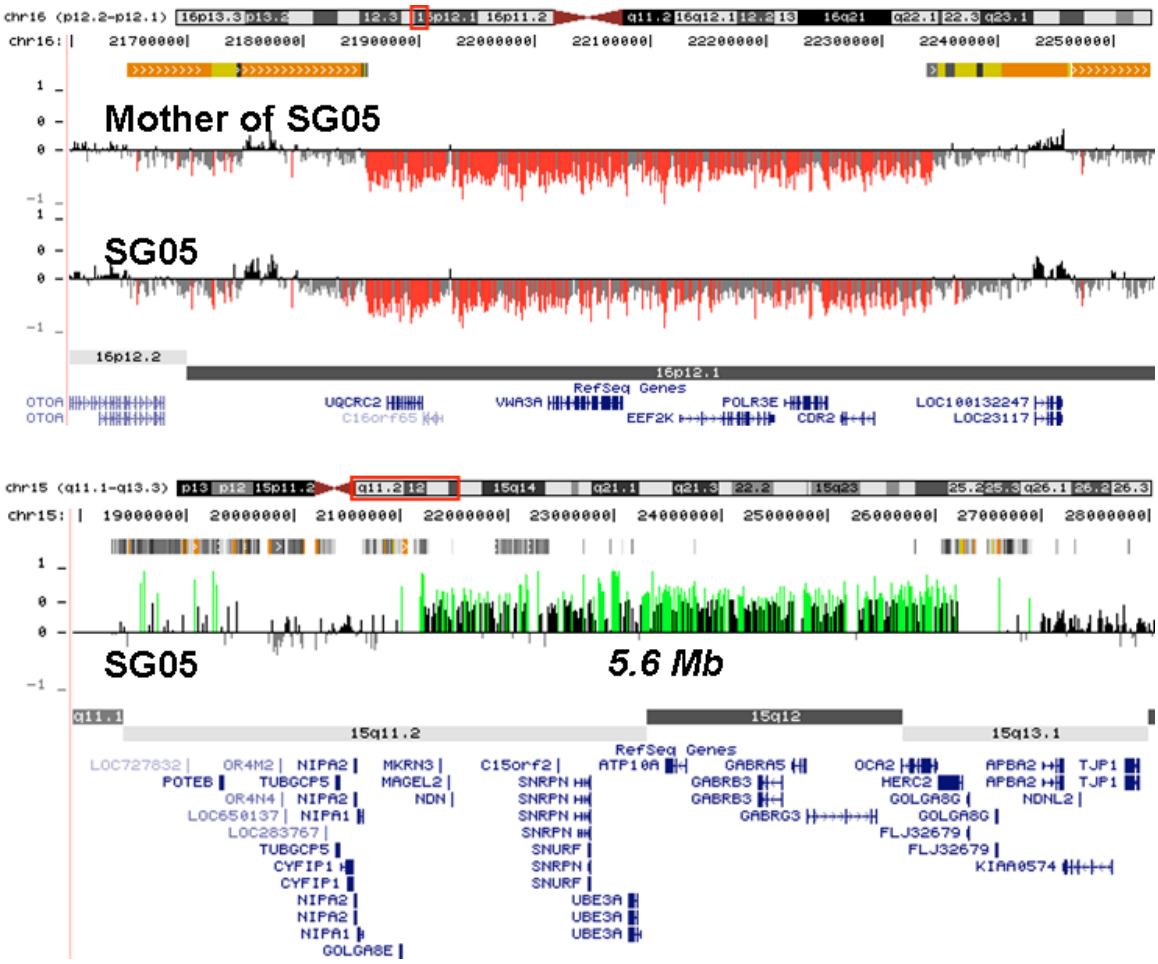
Nature Genetics: doi:10.1038/ng.534



Supplementary Fig. 2. Second hits in replication set of 16p12.1 microdeletion.

(A) and (B) show large deletions and (C) denotes a complex rearrangement on chr9. The FISH image shows deletion (green probe) and duplication (pink probe) on chr9 (arrow). Both the chr9 events were transmitted from the father who has a chr9 inversion; however, the 16p12.1 deletion was inherited from the mother.

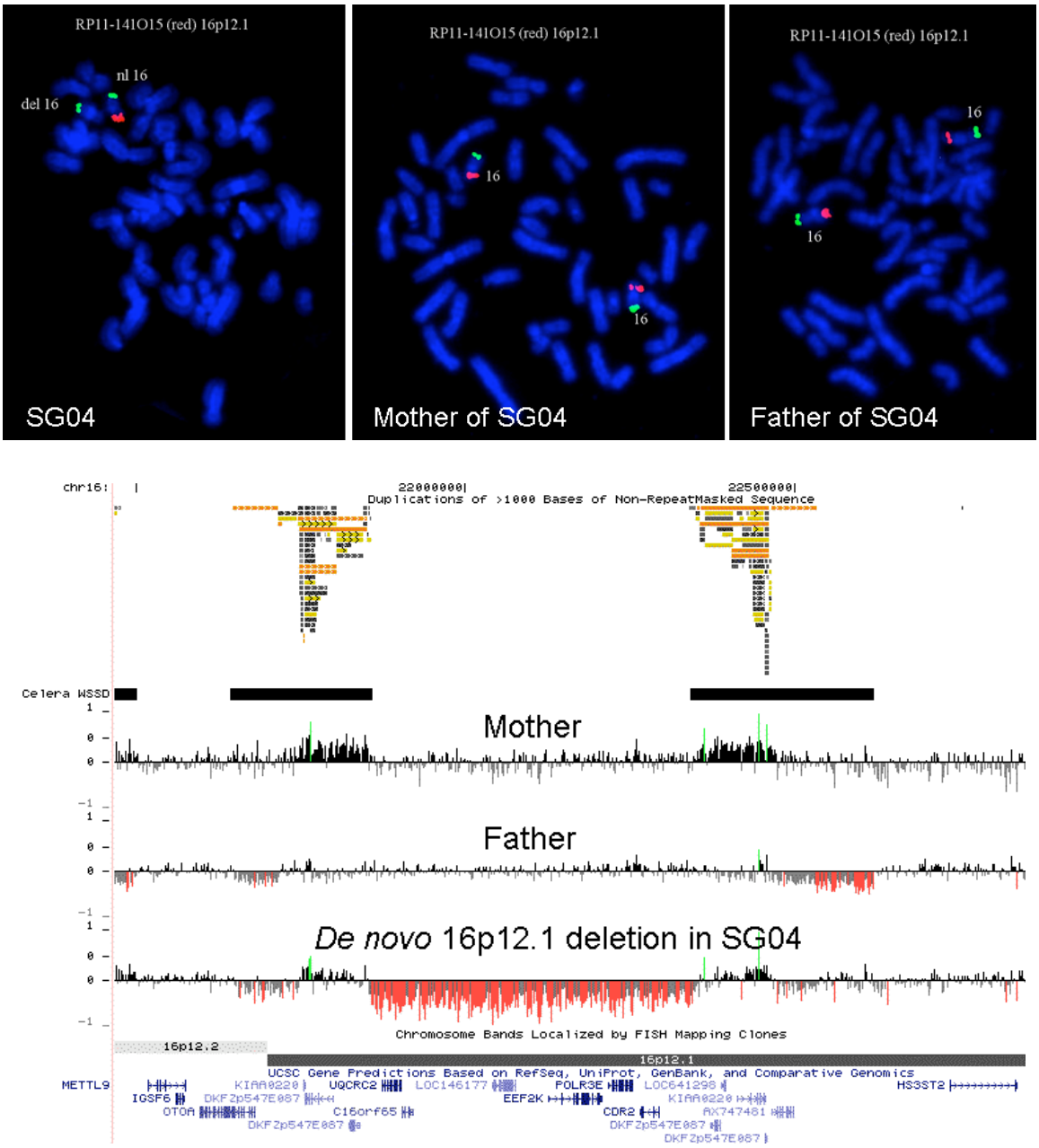
Nature Genetics: doi:10.1038/ng.534



Supplementary Fig. 3. Maternally inherited 16p12.1 microdeletion and a 15q11.2q13.1 duplication (*de novo* or paternally inherited) SG05.

High density array-CGH for SG05 and her mother is shown. Probes with $\log_2$ ratios above or below a threshold of 1.5 standard deviations from the normalized mean $\log_2$ ratio are colored green (duplication) or red (deletion), respectively. Also shown at the bottom panel is a 5.6-Mbp duplication of 15q11.2q13.1, associated with autism spectrum phenotypes. However, SG05 was not autistic at the time of evaluation.

Nature Genetics: doi:10.1038/ng.534



Supplementary Fig. 4. *De novo* 16p12.1 microdeletion in SG04.

(*Top panel*) Fluorescent *in situ* hybridization (FISH) analysis of metaphase chromosomal preparations from lymphoblast cells from SG04 and parents show the absence of a signal for the BAC probe (red, RP11-141O15) in the proband compared to the non-deleted parents who show two signals for the 16p12 probe. BAC probe RP11-7D23 from 16q subtelomere is labeled in green as a control. (*Bottom panel*) High density array-CGH for SG04 along with parental samples confirms the *de novo* 16p12.1 event.

Nature Genetics: doi:10.1038/ng.534

Supplementary Table 1. Segmental duplications at 16p12.1 rearrangement breakpoints

chrom	chromStart	chromEnd	Breakpoint	OtherStart	OtherEnd	size, bp	Hg18 orientation	Similarity
chr16	21647926	21716330	chr16	22464053	22532760	67986	same	99.51%
chr16	21716229	21720711	chr16	22351205	22355699	4456	same	99.44%
chr16	22351205	22355699	chr16			67986	-	99.9%
chr16	21748520	21851057	chr16	22357323	22460213	101892	inverted	99.75%
chr16	21790846	21794090	chr16	22460080	22463431	3202	inverted	98.77%

Nature Genetics: doi:10.1038/ng.534

Supplementary Table 2. Replication set of cases with 16p12.1 microdeletion and associated phenotypes

16p12.1 microdeletion				Second hit						
ID	**Gender**	**Clinical features**	**Inheritance**	**Chr**	**Start**	**End**	**Size**	**Event**	**Inheritance**	**Cytoband**
A8	M	Cardiac defects	Unknown	chrX	72819302	72990411	171,109	gain	Maternal	Xq13.2
A9	F	Developmental Delay	Maternal							
A10	F	Multiple Congenital Anomalies	Unknown							
A11	F	Developmental Delay, Seizure Disorder	Unknown							
A12	F	Seizures, developmental delay[1]	Unknown							
A13	M	Expressive language disorder	Unknown							
A14	M	Developmental Delay	Maternal							
A15	M	Developmental Delay, Hypotonia	Unknown							
AA1	M	Multiple congenital anomalies, reduction deformities of brain, profound mental retardation	Maternal	chr9	75,805,226	84,795,201	8,989,975	Gain	Paternal	9q21.13q21.32
				chr9	105,958,680	112,691,549	6,732,869	Loss	Paternal	9q31.1q31.3
A16	M	Developmental delay, lack of normal physiological development	Unknown							
A17	M	Undescended testicles	Unknown							
A18	M	Developmental delay	Maternal							
A19	M	Developmental delay	Unknown	chr3	5,411,510	6,088,506	676,996	Gain	Maternal	3p26.2p26.1
				chr14	20,805,673	22,308,694	1,503,021	Loss	Unknown	14q11.2
A20	M	Developmental delay	Paternal	chr19	41,608,636	42,381,964	773,328	Loss	Maternal	19q13.12
A21	F	Failure to thrive, short stature	Maternal	chr3	193,897,729	194,049,145	151,416	Loss	Unknown	3q29
A22	M	Microcephaly	Paternal							
A24	M	Failure to thrive	Maternal							
A25	M	Developmental delay[2]	Unknown							
A26	M	Congenital hydrocephalus	Maternal							
A28	M	Developmental delay, multiple congenital anomalies, dysmorphic features	Unknown							
A29	M	Delayed milestones	Unknown	chr4	94,053,559	94,374,372	320,813	Gain	Unknown	4q22.2
A30	M	Autistic disorder	Maternal	chr15	28,801,799	30,686,851	1,885,053	Loss	Maternal	15q13.2q13.3

[1]Has some features of Cornelia de Lange, but not convincing, is obese, and has developmental delay; [2]Has speech delays, delayed fine motor skills, mental retardation, headaches, recurrent ear infections, history of atrial displacement; Shaded rows represent cases with double hits >500 kbp.

Nature Genetics: doi:10.1038/ng.534

Supplementary Table 3. Clinical features identified in individuals with 16p12.1 microdeletion

Patient ID	Gender	Age at diagnosis	Craniofacial and skeletal features	Cognitive and neurological features	Other features	Family history	Inheritance
SG01	Female	7 years	Microcephaly	Profound psychomotor retardation, speech delay, hypotonia, West syndrome – seizure episodes, spastic quadriparesis, peripheral vision loss, delayed myelination, decreased white matter, and cerebral and cerebellar atrophy on MRI	NA	NA	Unknown
SG02	Male	15.8 years	Microcephaly, facial dysmorphology, growth retardation	Developmental delay, nonverbal, seizures, hypotonia, autistic, self-injurious behaviors	NA	NA	Unknown
SG03	Female	16 months	Facial dysmorphology, dysplastic ears, long eyelashes, periorbital fullness	Developmental delay, no speech at 2.25 years, hypotonia, periventricular changes on MRI	~2-Mbp terminal 22q deletion, strabismus	Mother with perceptual impairment; maternal half uncle with seizures; paternal great-grandmother with seizures	Maternal
SG04	Female	2 weeks	Telecanthus, depressed nasal bridge, bilateral clinodactyly	NA	Hypoplastic left heart with right-sided aortic arch, double-outlet right ventricle, mitral valve atresia, pulmonary stenosis, polysplenia, heterotaxy	NA	*De novo*
SG05	Female	3.9 years	Asymmetric face, small and deep-set eyes, prominent jaw, crowded teeth	Developmental delay, not autistic, speech delay, hypotonia, occasional staring spells	Dup(15q11.2q13.1), café-au-lait spot	Mother with depression, bipolar disorder, and post-traumatic stress disorder; two untested maternal half-sibs: one with ADHD and CHD; another with emotional problems, bifid uvula.	Maternal
SG06	Female	5.25 years	Microcephaly, triangular face, fleshy ears with small lobes, upslanting palpebral	Global developmental delay, no speech at 6 years, normal head CT, slowing	Still's murmur, 22q13ter del	Mother and father with learning disability; maternal grandmother	Maternal

Nature Genetics: doi:10.1038/ng.534

			fissures, hypoplastic alae nasi, overlapping toes, jaw angles to the left, poor growth	on EEG, hypertonic, behavior issues		with seizures; full sister with ADHD, OD, and LD, maternal half sister with febrile seizures	
SG07	Male	3.5 years	Microcephaly, coronal craniosynostosis, facial asymmetry, hypertelorism, telecanthus, epicanthus, inverted nipples, 2-3 toe syndactyly	Developmental delay, speech delay	Congenital heart defect - VSD, PFO, congenital absence of anterior portion of pericardium, chordee	Mother with learning disability and seizures; normocephalic younger sib with possible craniosynostosis but no 16p12.1 deletion	Maternal
SG08	Male	23 months	Microcephaly, short stature	Gross and fine motor delay, speech delay, hypotonia, seizure-like episodes with normal EEG, aggressive behaviors and ADHD	Strabismus, bilateral coloboma of iris	Mother with cognitive delay	Maternal
SG09	Female	4.5 years	Microcephaly, growth retardation, scoliosis, facial dysmorphology	Global developmental delay, nonverbal, abnormal stereotypic hand movements	Congenital heart defect – VSD and bicuspid aortic valve, tethered cord, neurogenic bladder	Mother with learning disability	Maternal
SG10	Female	16 months	Facial dysmorphology, downslanted palpebral fissures, bifid uvula, growth retardation, mild club foot bilaterally	Developmental delay, speech delay, decreased white matter on MRI, muscle weakness, ptosis	*De novo* del(5q15q23.3), deep sacral dimple, ectopic pupils	Mother with concerns of bipolar symptoms and family history of bipolar disorder	Maternal
SG11	Male	1 week	Flat nasal bridge, thin upper lip, prominent ears, bow legs, normal head size, length <3rd %ile	No delay at 8 months	Hypoplastic left heart diagnosed prenatally, small kidneys with intrarenal dilatation, shallow sacral dimple	Father reportedly in good health; mother has Blount syndrome	Paternal
SG12	Male	7.75 years	Facial dysmorphology	Developmental delay, seizures	Homozygous dup(22q11.21)	Mother and father with learning disability and dup(22q11.21); brother with learning disability, a cystic hygroma in the chest, homozygous dup(22q11.21), and no 16p12.1 deletion	Paternal
SG13	Male	13.5 years	Facial dysmorphology suggests a faint "gestalt" of trisomy 21, curved 5th	Learning disability, delayed fine motor skills and language, attention	Bilateral single palmar crease, asthma	Father with learning disability	Unknown

Nature Genetics: doi:10.1038/ng.534

			finger, pegged lateral incisor	deficit disorder			
SGA1	Male	11.5 years	Short stature, broad uvula, high-arched palate, alopecia, otherwise nondysmorphic	Developmental delay	Strabismus, hypospadias	NA	Unknown
SGA2	Male	10.5 years	Tall stature, large head (90th %ile), upslanted and deep-set eyes, retrognathia, short neck	Developmental delay, borderline intellectual function, psychiatric and behavioral issues, speech delay	Café-au-lait spots, dup(2p13p12)	Mother with bipolar disorder and developmental delay	Unknown
SGA3	Male	2.5 years	Hypertelorism, telecanthus, bifrontal narrowing, long eyelashes, inverted epicanthal folds, low-set ears, flat nasal root, normocephalic, dental carries	Developmental delay, speech delay, generalized hypotonia, febrile seizures, normal EEG and brain MRI, bilateral ptosis, myopathic mouth (holds open and drools)	Unilateral inguinal hernia, incomplete retinal vascularization (extreme prematurity), unilateral mild middle ear dysfunction	Father with childhood febrile seizures; mother with mild learning disability	Paternal
SGA5	Male	18 months	Deep-set eyes with epicanthus, short columella, wide mouth, prominent ears, OFC at 5th %ile, inverted nipples	Speech delay, global developmental delay, unilateral sensorineural hearing loss	NA	Father with childhood seizures; mother with learning disability, scoliosis, depression and anxiety; maternal grandfather with schizophrenia	Maternal
SGA6	Male	2 weeks	Broad nasal bridge, upturned nose, slight retrognathia, narrow palpebral fissures, short neck, broad nipples, low set posteriorly rotated ears, clinical diagnosis of cardiofacioucutaneous syndrome with mutation in *BRAF*	Complete agenesis of corpus callosum by brain MRI, myoclonus, normal EEG, unilateral hearing loss, epilepsy	Maternally inherited dup(14q32.13-q32.2), prenatal history of polyhydramnios and pleural effusions, congenital cardiac defect (PDA, PFO, pulmonary artery stenosis), bilateral hydronephrosis, thrombocytopenia, Hirschsprung disease	Maternal half-brother with learning disability; maternal grandfather reportedly with bipolar disorder	Mother does not have deletion, father unavailable for testing
SGA7	Male	2 weeks	Unilateral (left) cleft lip and palate, low-lying conus medullaris	Normal head ultrasound	Hypoplastic left heart, sacral dimple with tuft of hair	Brother with pyloric stenosis (untested); paternal family history of spina bifida	Unknown

Nature Genetics: doi:10.1038/ng.534

SGA8	Male	1.5 weeks	Mild dysmorphology, wide nose, dysplastic ear lobes, short neck, minor syndactyly of toes 2-3-4 with dorsal crease	Hypertonia/joint contractures of knees and fingers	Dextrocardia or mesocardia, total anomalous pulmonary venous connection, hypoplastic left heart, Shone's variant, hypoplastic aortic valve, and left aortic arch, maternally inherited duplication involving the *XIST* gene (chrX:72819302-72990411, hg18 coordinates)	NA	Unknown
SGA9	Female	23 months	Mild dysmorphology, poor growth	Pervasive developmental delay, speech delay, lost words, echolalia, very little spontaneous speech, hand flapping, hyperkinesis, poor attention, significant hypotonia, early gross motor milestones on time	Horseshoe kidney, lipoma on right labia, gastroesophogeal reflux, 2-vessel umbilical cord	Mother reportedly in good health; maternal uncles with developmental delays and social anxieties (untested)	Maternal
SGA28	Male	17 years	Microcephaly, bitemporal narrowing, overfolded helices, malocclusion of teeth with mixed underbite, absence of canines and duplication of lateral incisors of both the upper and lower dentition, lip pits in lower lip, aberrantly placed submandibular gland, cleft soft palate with absence of uvula, limited extension of elbows, bilateral fifth digit hypoplasia on feet	Impaired cognitive abilities, performs at the 2^{nd} grade level, psychiatric issues requiring hospitalization, aggression	Bilateral sensorineural hearing loss, clinical diagnosis of Van der Woode syndrome, prenatal exposure to alcohol and cocaine, possible history of abuse	Possible consanguinity, father has a history of violence, dyslexia and mild LD, mother with short stature, neither parent completed high school	Unknown
25514	Male	5 years	Plagiocephaly and brachycephaly, growth retardation (height 10^{th}-25^{th} centile; HC 2^{nd}-50^{th} centile), craniofacial dysmorphology	Developmental and speech delay, non-autistic, behavioral issues, sleep disturbance,	NA	Father had significant learning difficulties, concerns of speech and language development in a 21-month sister	Maternal

NA=information not available; MRI=magnetic resonance imaging; ADHD=attention deficit hyperactivity disorder; CHD=congenital heart defect; EEG=electroencephalogram; CT=computed tomography; OD=obsessive disorder; LD=learning disability; VSD=ventricular septal defect; PFO=patent foramen ovale; OFC=occipto-frontal circumference; PDA=patent ductus arteriosus

Nature Genetics: doi:10.1038/ng.534

Supplementary Table 4. Large CNV second hits in individuals with 16p12.1 deletions

	Case ID	Chrom	Start	End	Size (bp)	CNV	Comments	Inheritance
1	SG03	Chr22	47,457,855	49,549,855	2,092,000	Loss	22q13	Unknown
2	SG05	Chr15	21,000,000	26,609,706	5,609,706	Gain	15q11.2-q13.1	Unknown
3	SG10	Chr5	94,407,129	129,832,559	35,425,430	Loss	5q15-q23.3	*De novo*
4	SG12	Chr22	17,258,339	19,786,713	2,528,374	Gain	22q11.21	Both parents
5	SGA2	Chr2	75,333,401	76,640,253	1,306,852	Gain	2p13p12	Unknown
6	SGA6*	Chr14	95,166,172	96,014,653	848,481	Gain	14q32.13-q32.2	Maternal
7	SGAA1	Chr9	75,805,226	84,795,201	8,989,975	Gain	9q21.13q21.32	Paternal[1]
	SGAA1	Chr9	105,958,680	112,691,549	6,732,869	Loss	9q31.1q31.3	Paternal[1]
8	SGA19	Chr3	5,411,510	6,088,506	676,996	Gain	3p26.2p26.1	Maternal
	SGA19	Chr14	20,805,673	22,308,694	1,503,021	Loss	14q11.2	Unknown
9	SGA20	Chr19	41,608,636	42,381,964	773,328	Loss	19q13.12	Maternal
10	SGA30	Chr15	28,801,799	30,686,851	1,885,053	Loss	15q13.2q13.3	Maternal

*Also has a mutation (F468S) in the *BRAF* gene consistent with cardiofaciocutaneous syndrome; [1]Father has a balanced inversion on chr9

Supplementary Table 5. Enrichment for 'second hit' CNVs among del 16p12.1 carriers

We originally conditioned on the presence of both deletions and duplications (Table 3). We also reanalyzed the controls conditioning for the presence of a >500-kbp deletion only and then counted the number of observations with two hits. This analysis also suggests that there is enrichment of double hits in 16p12.1 microdeletion cases (p=0.0014, OR= 5.5).

	Two large CNVs	Total	Percent	Significance	OR
Discovery set cases	6	20	30%		
Discovery set controls[1]	5	68	7.3%	p=0.014	5.2
All discovery set controls[2]	9	2493	0.36%		
Replication set cases	4	22	18.2%		
Replication controls[1]	2	52	3.8%	p=0.06	5.4
All replication set controls[2]	12	2792	0.25%		
Combined cases	10	42	23.8%		
Combined controls[1]	7	120	5.8%	p=0.0014	5.5
All controls[2]	21	5285	0.39%		

[1]Controls conditioned to have at least one large deletion (>500 kbp) and finding of a second hit; [2]Controls having two CNVs >500 kbp.

Supplementary Table 5a. Average size and genes in second hit CNVs in cases and controls

	Number of second hits	Average size of the second hit (bp)	Average number of overlapping genes per individual
Cases (n=42)	10	6,837,152	42
Controls (n=120)	7	911,836	4

Overall, using the above analysis, we find that the average size of second hits in cases is 7.5-fold higher than in controls; there is also a 10-fold enrichment of RefSeq genes within the second hit CNVs in cases compared to that in controls.

Nature Genetics: doi:10.1038/ng.534

Supplementary Table 6. Inheritance of the 16p12.1 microdeletion and second hit

	Case ID	16p12.1 deletion inheritance	Second hit inheritance
1	SG03	Maternal	Paternal[1] or *de novo*
2	SG05	Maternal	Paternal[1] or *de novo*
3	SG10	Maternal	*De novo*
4	SG12	Paternal	Both parents
5	SGA2	Unknown	Unknown
6	SGA6	Paternal[1] or *de novo*	Maternal
7	SGAA1	Maternal	Paternal[2]
	SGAA1		Paternal[2]
8	SGA19	Paternal[1] or *de novo*	Maternal
	SGA19		Paternal[1] or *de novo*
9	SGA20	Paternal	Maternal
10	SGA30	Maternal	Maternal

[1]In these cases the father was not tested; [2]father has a balanced inversion on chr9; Shaded rows indicate inherited second hits

Supplementary Table 7. Parental ascertainment for 16p12.1 phenotypes

Parental phenotypic information was obtained based on family history information gathered during the clinic visit, prior to the genetic testing. In some cases, more detailed inquiries were made about phenotypes in parents found to carry the microdeletion.

Patient	**Ascertainment of parental phenotypes**
SG03	Parental information was obtained during the first visit before cytogenetic/CNV testing
SG04	Parental information was obtained during the first visit. No abnormal phenotypes were detected in the parents
SG05	Parental phenotype info collected during the initial consult, when array-CGH was ordered.
SG06	Some phenotypes were observed in the mother during the first visit but more clinical information was obtained after the child was tested
SG07	Phenotypic information was obtained before the cytogenetic/CNV testing
SG08	Parental phenotype information was obtained after the aCGH.
SG09	Minimal information obtained before aCGH, including maternal special education. More details about parental educational history were obtained after aCGH.
SG10	Phenotypic information was obtained after the diagnosis of mother as a carrier of the 16p12.1 microdeletion
SG11	Before the testing, during prenatal counseling
SG12	Both the parents' history of disabilities were known before the cytogenetic evaluation
SGA3	Clinical data on parents obtained before the child was tested.
SGA5	Parental phenotype information was taken at the time of the initial visit.
SGA9	Parental phenotype information was taken at the time of the initial visit.
25514	Family history was documented during the initial visit before the array-CGH evaluation

Nature Genetics: doi:10.1038/ng.534

Supplementary Table 8. Frequency of clinical features in carrier parents

	Case ID	Inheritance	Learning disabilities	Depression/bipolar disorder	Seizures	Other features
1	SG03	maternal	+	-	-	
2	SG05	maternal	-	+	-	
3	SG06	maternal	+	-	-	
4	SG07	maternal	+	-	+	
5	SG08	maternal	+	-	-	
6	SG09	maternal	+	-	-	
7	SG10	maternal	-	+	-	
8	SG11	paternal	-	-	-	Good health
9	SG12	paternal	+	-	-	
10	SGA3	paternal	-	-	+	
11	SGA5	maternal	+	+	-	Anxiety, scoliosis
12	SGA9	maternal	-	-	-	Good health
13	25514	maternal	-	-	N	Good health
	Total	**3 pat/10 mat**	**7/13**	**3/13**	**2/12**	

Supplementary Table 9. An excess of 16p12.1 del carrier parents manifesting phenotype

	Parents with phenotype	Parents with no phenotype
Carrier parents	10 parents (SG05d, SG12m, SG10m, SG06m, SG07m, SG08m, SG09m, SGA3d, SGA5m, SG03m)	3 parents (SG11d, SGA9m, 25514m)
Non-carrier parents	5 parents (SG12m, SG06d, SGA3m, SGA5d, 25514d)	9 parents (SG03d, SG11m, SG08d, SG07d, SG04m, SG04d, SG10d, SGA6d, SGA9d)

Parents not tested were not included in this calculation; d, dad; m, mom.
Fisher's exact test was performed to identify if there was a significant excess of carrier parents manifesting the phenotype compared to non-carrier parents; p=0.037; odds ratio = 6.0

Nature Genetics: doi:10.1038/ng.534

1. Supplementary Note: Copy-number variation (CNV) detection

1. CNV analysis in affected individuals

1.1. CNV analysis in cases with intellectual disabilities and developmental delay

Discovery cohort: A total of 11,393 cases with presenting complaints of intellectual disability, developmental delay and/or congenital anomalies were submitted to Signature Genomic Laboratories for analysis of copy-number variation (CNV) for the year spanning Nov 2007 – Dec 2008. These samples were ascertained from geneticists, pediatricians, and neurologists from several centers. We utilized these 11,393 samples along with 480 samples from Italy and Australia, totaling to 11,873 samples, with indications of developmental delay and congenital malformation for our initial discovery cohort.

Replication cohort: A total of 9,254 DNA samples were utilized for this set. These samples were submitted to Signature Genomics Laboratories for CNV analysis during 2009. This set is independent of the discovery set and no overlap is noted between any cases. All these samples were ascertained independent of the first discovery set for major indications of congenital neurodevelopmental disease.

CNV detection: Microarray-based comparative genomic hybridization (CGH) was performed with a whole-genome bacterial artificial chromosome (BAC) microarray chip (SignatureChipWG®) and/or an oligo-based (SignatureChipOS®) chip (either 105K custom-designed by Agilent Technologies or 135K custom-designed by Roche NimbleGen). The whole-genome SignatureChipWG® has three BAC clones (RP11-1152N5, RP11-141O15, and RP11-21M24), and the SignatureChipOS® has 42 oligo probes (105K) or 74 oligo probes (135K) spanning the 16p12.1 region (chr16: 21,857,845-22,336,067). Microarray hybridizations were performed as described previously [1,2]. Table S1.1 shows a breakdown of samples (cases) analyzed using BAC arrays versus the oligo arrays. There was no difference in the detection rate with the BAC arrays compared to oligo-based arrays (p=0.12, Fisher's exact test).

Table S1.1. Detection rate of 16p12.1 microdeletion using BAC versus oligo arrays

Platform	Discovery set	Replication set	Platform-specific total
BAC arrays	12/7007	2/2123	14/9130
Oligo arrays	7/4386	20/7131	27/11,517
Total	19/11,393	22/9254	

Nature Genetics: doi:10.1038/ng.534

1.2. CNV analysis in cases with schizophrenia

DNA samples from schizophrenia cases (n=416) were analyzed on the NimbleGen HD2 array-CGH platform at Cold Spring Harbor Laboratories and a further 2,645 DNA samples with schizophrenia were analyzed on the Affymetrix 6.0 platform by the Genetic Association Information Network (GAIN) project for the study of schizophrenia (phs000021.v2.p1). In addition, we included 96 individuals affected with schizophrenia or schizo-affective disorders from 26 multiplex families. These families were interviewed, diagnosed, and sampled as previously described [3,4].

2. Controls for 16p12.1 analysis

Discovery cohort: The control copy-number variation data (n=8540) consisted of six sets: set 1 comprised 671 individuals of European descent with no family history or first-degree relative with amyotrophic lateral sclerosis, ataxia, autism, brain aneurysm, dystonia, Parkinson's disease, or schizophrenia; set 2 comprised 936 middle-aged (40–70 years) individuals of European descent living in the United States tested for statin use and cholesterol levels; set 3 consisted of 886 individuals from the Human Genome Diversity Panel (HGDP); set 4 data (n=3,181) were obtained from control individuals used for a large study of schizophrenia [5]; set 5 is composed of 446 schizophrenia control samples; and set 6 consisted of data obtained from 2,420 GAIN controls utilized for a genome-wide association study of schizophrenia. The GAIN cohort was collected to represent unrelated cases or controls. Their health and ethnicity were evaluated through a web-based questionnaire and ascertained by source of sample, geographic representativeness, comparability of cases and controls (Supplementary Table 3), and comprehensiveness of trait and phenotypic definitions [6]. The University of Washington Committee on Research Involving Human Subjects and the Institutional Review Board approved this study. A breakdown of the ethnic origin of each schizophrenia case and control samples is shown in Table S1.2.

Two-hit frequency: For comparing the double-hit frequency 2,493 controls were derived from the following (1) 671 individuals of European descent with no family history or first-degree relative with amyotrophic lateral sclerosis, ataxia, autism, brain aneurysm, dystonia, Parkinson's disease, or schizophrenia; (2) 936 middle-aged (40–70 years) individuals of European descent living in the United States tested for statin response and cholesterol levels; (3) 886 individuals from the Human Genome Diversity Panel (HGDP) [7].

Nature Genetics: doi:10.1038/ng.534

Table S1.2. Composition of schizophrenia cases and controls

Source	CSHL		GAIN	
Platform	NimbleGen HD2		Affymetrix 6.0	
	Cases	Controls	Cases	Controls
Caucasian	303	396	1404	1442
African American	27	35	1241	978
East Asian	5	7	0	0
Hispanic	6	8	0	0
Unknown	75	0	0	0
Total	416	446	2645	2420

Replication cohort: We performed a replication study utilizing the control data from Welcome Trust Case Control Consortium (WTCCC) reported by Kirov and colleagues [8] and published control data from Shaikh and colleagues [9]. Shaikh controls (n=2026) consist of children and their parents who visited the Children's Hospital of Philadelphia and its satellite sites. Individuals with developmental delays or neurological disorders were excluded. However, most subjects were too young to be diagnosed or excluded for psychiatric phenotypes. Adults were only screened for autism or any behavioral disorders. The WTCCC control DNAs (n=2,792) were generated from blood donors or general population sampling and were not screened for any phenotypes [10]. Moreover, these controls were ascertained for genome-wide association studies involving common variants and a proportion of these controls are likely to have disease of interest [10]. Thus, it is likely that clinical features such as depression, seizures or bipolar disorders are present in association with 16p12.1 and that these controls represent the general population as opposed to a true "unaffected" control set. The comparisons in the replication set are thus conservative. Two controls from the Shaikh control set and three controls from Kirov set were identified to carry a 16p12.1 microdeletion. However, utilizing the identifiers from Kirov sample data, we find that none of the three WTCCC controls with the microdeletion carry a large CNV second hit.

Further investigation of phenotypes in two control individuals from the GAIN cohort revealed a diagnosis of major depressive disorder (MDD) in one of the controls (26140) (Table S1.3).

Table S1.3. Responses to the GAIN control questionnaire

Variable accession	Variable name	Variable description	Sample1	Sample2
			18125	**26140***

Nature Genetics: doi:10.1038/ng.534

phv00020270.v1.p1	A1A	Have you ever had a time in your life when you felt sad, blue, or depressed for two weeks or more in a row?	yes	yes
phv00020271.v1.p1	A1B	Have you ever had a time in your life lasting two weeks or more when you lost interest in most things like hobbies, work, or activities that usually give you pleasure?	yes	no
phv00020272.v1.p1	A1C	How much of the day did these feelings usually last?	about half of the day	most of the day
phv00020273.v1.p1	A1D	Please think of the two-week period in your life when your feelings of depression or loss of interest were worst: did you feel this way? (every day, almost every day, less often)	not asked	everyday
phv00020274.v1.p1	A2	Did you feel more tired out or low on energy than is usual for you?	not asked	yes
phv00020275.v1.p1	A3	Did you gain or lose weight without trying, or did you stay about the same weight?	not asked	lost
phv00020276.v1.p1	A3AA	About how much weight did you gain?	not asked	not asked
phv00020277.v1.p1	A3AB	About how much weight did you lose?	not asked	~20lbs
phv00020278.v1.p1	A3ACGAIN	(Gained): about how much weight did you gain and lose?	not asked	not asked
phv00020279.v1.p1	A3ACLOST	(Lost): about how much weight did you gain and lose?	not asked	not asked
phv00020280.v1.p1	A4	Did you have more trouble falling asleep than you usually do?	not asked	yes
phv00020281.v1.p1	A4A	Did you have more trouble falling asleep than you usually do? If yes: how often did that happen?	not asked	nearly every night
phv00020282.v1.p1	A5	Did you have a lot more trouble concentrating than usual?	not asked	yes
phv00020283.v1.p1	A6	People sometimes feel down on themselves, no good, or worthless. Did you feel this way?	not asked	yes
phv00020284.v1.p1	A7	Did you think a lot about death - either your own, someone else's, or death in general?	not asked	yes

Nature Genetics: doi:10.1038/ng.534

phv00020285.v1.p1	COUNT	Count of 'yes' answers to A1B (Have you ever had a time in your life lasting two weeks or more when you lost interest in most things like hobbies, work, or activities that usually give you pleasure?), A2 (Did you feel more tired out or low on energy than is usual for you?), A3 (Did you gain or lose weight without trying, or did you stay about the same weight?), A4 (Did you have more trouble falling asleep than you usually do?), A5 (Did you have a lot more trouble concentrating than usual?), A6 (People sometimes feel down on themselves, no good, or worthless. Did you feel this way?), and A7 (Did you think a lot about death - either your own, someone else's, or death in general?) question in questionnaire	-2	5
phv00020286.v1.p1	A8	About how many weeks altogether did you feel this way? Count the weeks before, during and after the worst two weeks. The total period of depression/loss of interest was:	not asked	12
phv00020287.v1.p1	A8A	How many periods like this did you have in your life, lasting two or more weeks?	not asked	3
phv00020288.v1.p1	A8B	About how old were you the FIRST time you had a period of two weeks like this? (Whether or not you received any help for it?) Years of age when you first felt this way	not asked	26
phv00020289.v1.p1	A8C	About how old were you the LAST time you had a period of two weeks like this? (Whether or not you received any help for it?) Years of age when the most recent episode happened	not asked	38
phv00020290.v1.p1	A8D	Did you ever tell a professional about these problems (medical doctor, psychologist, social worker, counselor, nurse, clergy, or other helping professional)?	not asked	no

Nature Genetics: doi:10.1038/ng.534

phv00020291.v1.p1	A8E	Did you take medication or use drugs or alcohol more than once for these problems?	not asked	yes
phv00020292.v1.p1	A8F	How much did these problems (feeling depression or loss of interest) interfere with your life or activities? (A lot, some, a little, not at all)	not asked	some

*The 'control' individual 26140 fits into a diagnosis of major depressive disorder (MDD).

Nature Genetics: doi:10.1038/ng.534

3. Platforms and probe coverage for arrays used in cases and controls

Numerous studies have shown that different platforms show comparable sensitivity and specificity for large CNVs (typically >200 kbp in size)[11,12]. Based on our own studies and those of Redon and colleagues[13], two or more unique BACs (without segmental duplications) can be used to accurately detect microdeletions and microduplications although the boundaries are not precisely identified. In this study, we note that all events that were detected by BAC-based microarrays were subsequently confirmed by high-density oligonucleotide microarrays providing near-breakpoint resolution of the boundaries. We were able to validate all (37/37, 100%) samples with the 16p12.1 microdeletion using an orthogonal platform (see next section). Nevertheless, because of the larger number of platforms under consideration, we compared detection rates in cases and controls as a function of the platform. All control samples were detected using oligonucleotide microarrays. The 16p12.1 microdeletion is >500 kbp in size and should be easily detectable using currently available array platforms. Our previous analysis (HMMSeq) [7] and those of other groups suggests ten probes are typically sufficient to detect such events. The different oligonucleotide different array platforms that were used to generate the control data have adequate coverage throughout the genome and for the 16p12.1 region. A list of probe coverage for each of the platforms used in cases and control CNV detection is shown in Table S1.4. Note the study of Kirov and colleagues had the fewest number of probes (n=26) and yet identified the greatest number (n=3) of 16p12.1 microdeletions.

Nature Genetics: doi:10.1038/ng.534

Table S1.4. Platforms and probe coverage for arrays used for cases and controls

Screening set cases	n	Evaluation platform	Probe coverage	Deletions
Meta-analysis (ISC cases)	3,391	Affymetrix 5.0/6.0	141/251 probes	4
Meta-analysis (Marshall et al)	427	Affymetrix GeneChip 500K	26 probes	1
Discovery set cases				
Cases with LD/DD	7,007	SignatureChipWG BAC arrays	3 BACs	12
	4,386	105K Agilent oligo arrays	42 probes	7
	480	Nimblegen 12 x 135 K	200 probes	1
Schizophrenia cases	3,061	NimblegenHD2/Affymetrix 6.0/ROMA	368/251/18 probes	3
Discovery set controls				
Scz controls	2,866	NimblegenHD2/Affymetrix 6.0/ROMA	368/251/18 probes	2
Itsara et al (PARC)	936	Illumina Human Hap300	30 probes	0
Itsara et al (NINDS)	671	Illumina Human Hap550	47 probes	0
Itsara et al (HGDP)	886	Illumina Human Hap650Y	55 probes	0
Itsara et al (ISC controls)	3,181	Affymetrix 5.0/6.0	141/251 probes	0
Replication set cases				
Cases with LD/DD	2,123	SignatureChipWG BAC arrays	3 BACs	2
	7,131	105K Agilent/135K Nimblegen oligo arrays	42/74 probes	20
Replication set controls				
Shaikh controls	2,026	Illumina Human Hap550	47 probes	2
Kirov controls	2,792	Affymetrix GeneChip 500K	26 probes	3
PARC	782	Illumina Quad 610	48 probes	1
Inchianti samples	699	Illumina Human Hap550	47 probes	0

Note: screening set indicates the initial study by Itsara et al, 2009; Discovery set denotes the first set used and replication set consists of independent set of cases and controls.

Nature Genetics: doi:10.1038/ng.534

4. Validation of 16p12.1 microdeletions and the second hits

Two types of validations were performed: (1) events identified by oligo/BAC arrays in ID/DD cohorts were confirmed by fluorescent *in situ* hybridization (FISH) (2) further validation was performed to delineate the breakpoints (for both 16p12.1 events and the second hits) using custom targeted array-CGH.

(1) All 16p12.1 microdeletions, with the exception of a single case where only extracted DNA was available, from the ID/DD cohorts were confirmed by fluorescent *in situ* hybridization (FISH) analysis on metaphase chromosomal preparations using one or more BAC clones spanning the region, as described previously [14].

(2) To refine the breakpoints and for validation of the 16p12.1 deletions identified by whole-genome BAC/oligo arrays or ROMA/NimbleGen HD2/Affymetrix 6.0 platforms, a custom high density oligonucleotide array (NimbleGen Systems) was used (Fig. S1.1). The high density array consisted of 385,000 isothermal probes (45-75 bp) including a 4-Mb 16p12.1 region (80,000 probes, mean density 1 probe every 50 bp) and three or more known microdeletion regions as controls (305,000 probes, mean density 1 probe every 100 bp) (Table S1.5). We were able to validate all (37/37, 100%) available DNA samples with the 16p12.1 microdeletion using an orthogonal platform (Supplementary Fig. 1, and Fig. 1).

Table S1.5. High density 16p12.1 microarray design with control regions

Region	chr	start	stop	size	density	#probes
16p12	chr16	20000000	24000000	4,000,000	50	80000
15q25 distal	chr15	82400000	84000000	1,600,000	100	16000
15q25 prox	chr15	80000000	82400000	2,400,000	100	24000
15q13	chr15	25500000	31000000	5,500,000	100	55000
17q23	chr17	54500000	58000000	3,500,000	100	35000
7p15	chr7	29000000	35200000	6,200,000	100	62000
9q22	chr9	89600000	90000000	400,000	100	4000
19q13	chr19	41400000	42600000	1,200,000	100	12000
19q13	chr19	53000000	55500000	2,500,000	100	25000
15q24	chr15	70000000	76500000	6,500,000	100	65000
			Total	33,800,000		378,000

Nature Genetics: doi:10.1038/ng.534

For validation of the second hits, targeted arrays comprising of 135,000 probes on genomic regions flanked by segmental duplications or "hotspots" (mean density 1 probe every 2500 bp) and a genomic backbone (mean density 1 probe every 36,000 bp) were also utilized (Girirajan and Eichler, unpublished). All microarray hybridization experiments were performed as described previously [15], using a single unaffected male (GM15724 [Coriell]) as reference. All validations for the replication set of cases were performed using the 'hotspot' chip where the density of probes in the 16p12.1 region is one probe every ~2.5 kbp totaling to about 200 probes in the region (Supplementary Fig. 1).

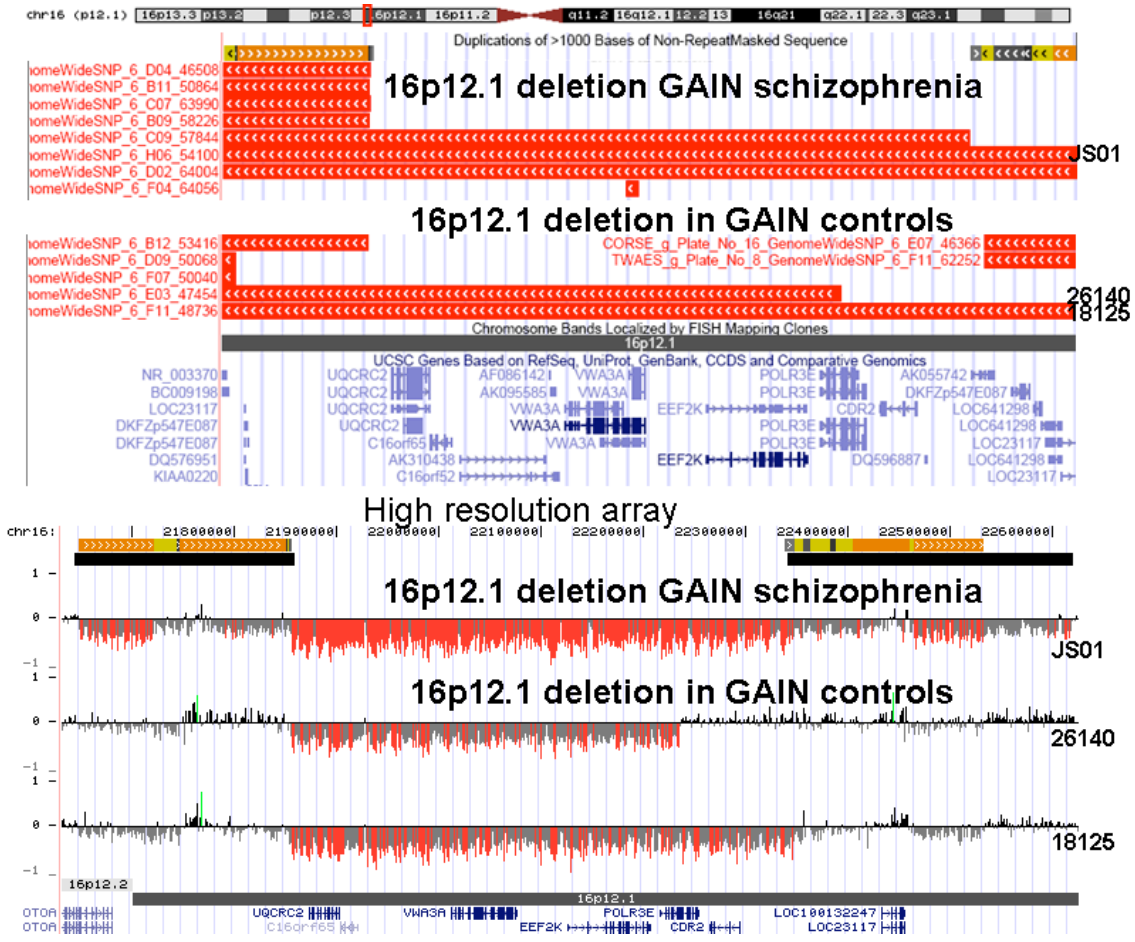


Fig. S1.1. Validation by high density oligonucleotide arrays. The upper panel shows the location and the extent of 16p12.1 microdeletion in schizophrenia cases and controls assayed by HD2 and Affymetrix 6.0 arrays. The lower panel shows the validation of a subset of deletion cases (JS01) and controls (26140, 18125) using NimbleGen high resolution arrays. Note that the breakpoint in the control sample 26140 maps outside of the segmental duplication.

Nature Genetics: doi:10.1038/ng.534

2. Supplementary Note: CNV analysis in 480 patients with ID/DD from Italy and Australia

We analyzed 480 individuals (458 individuals from Italy and 22 individuals from Australia) with indications of intellectual disabilities and multiple congenital anomalies using the hotspot array. We identified one individual (25514) from the Australian cohort with the 16p12.1 microdeletion (Fig. S2.1, Fig. S2.2). The clinical features of this case are as follows: 25514 has fine motor and speech/language delay, marked plagiocephaly and mild dysmorphic facial features.

Pregnancy was complicated by thrombocytopenia, shingles in the 3rd trimester and reduced amniotic fluid volume. Labour was induced at 39 weeks and he was born by normal vaginal delivery with birth weight 3325 g (25^{th}-50^{th} centile), birth length 50 cm (10^{th}-50^{th} centile and birth head circumference 34.5 cm (10^{th}-50^{th} centile) (Supplementary Fig. 3). He was well at birth, fed well and thrived. There were concerns about his head shape from 6 months of age. Skull X-ray at 4½ years of age and CT skull at 5 years of age showed plagiocephaly with all sutures patent. Development was delayed. He sat at 9 months and walked at 24 months. Concerns were raised about his speech/language and fine motor development when he started kindergarten at 4 years of age. He receives speech pathology support. Formal developmental assessment has not yet been performed. Hearing and eyesight have been assessed and are normal. Behaviourally, he is described as a happy boy who is not a good listener and has difficulties sitting still, is distractible and tends to walk on tip toes. There are no autistic features. He was not a good sleeper, waking at least once each night, until he started school at 5 years. He has one sibling, a 21 month old sister. She was born prematurely (32 weeks) and there are concerns about her speech/language development. Her development has not yet been assessed formally. Mother completed Year 11 but left school early in Year 12. She worked as a waitress and receptionist prior to having her children. Father had significant learning difficulties, received extra help in Year 8, repeated Year 8 and left school during Year 9. He worked as a painter/decorator for some time after leaving school and, for the past 8 years, as a factory worker. Father's father had learning difficulties at school. There is no other family history of learning difficulties/intellectual disability or psychiatric illness on either side of the family. On examination, height was 104 cm (10^{th}-25^{th} centile), weight was 15.8 kg (10^{th} centile) and head circumference was 48.4 cm (2^{nd}-50^{th} centile). There was pronounced plagiocephaly and brachycephaly with flattening of the left occipitoparietal region. He had an anterior cowlick. There were a broad, relatively high forehead with frontal bossing, relatively high nasal bridge, relatively widely set eyes and a short philtrum. The palate and dentition were normal. He had joint hypermobility with Beighton score 6-7/9 associated with hypotonia. Hands and feet were normal apart from small finger nails. No abnormality was detected on examination of the cardiovascular system, abdomen or genitalia. Past investigations have included CT brain, chromosomes, subtelomere MLPA, molecular testing for fragile X syndrome, creatine kinase, thyroid function tests, urine amino and organic acids, calcium/phosphate/vitamin D, iron studies, full blood examination and urea/electrolytes.

Nature Genetics: doi:10.1038/ng.534

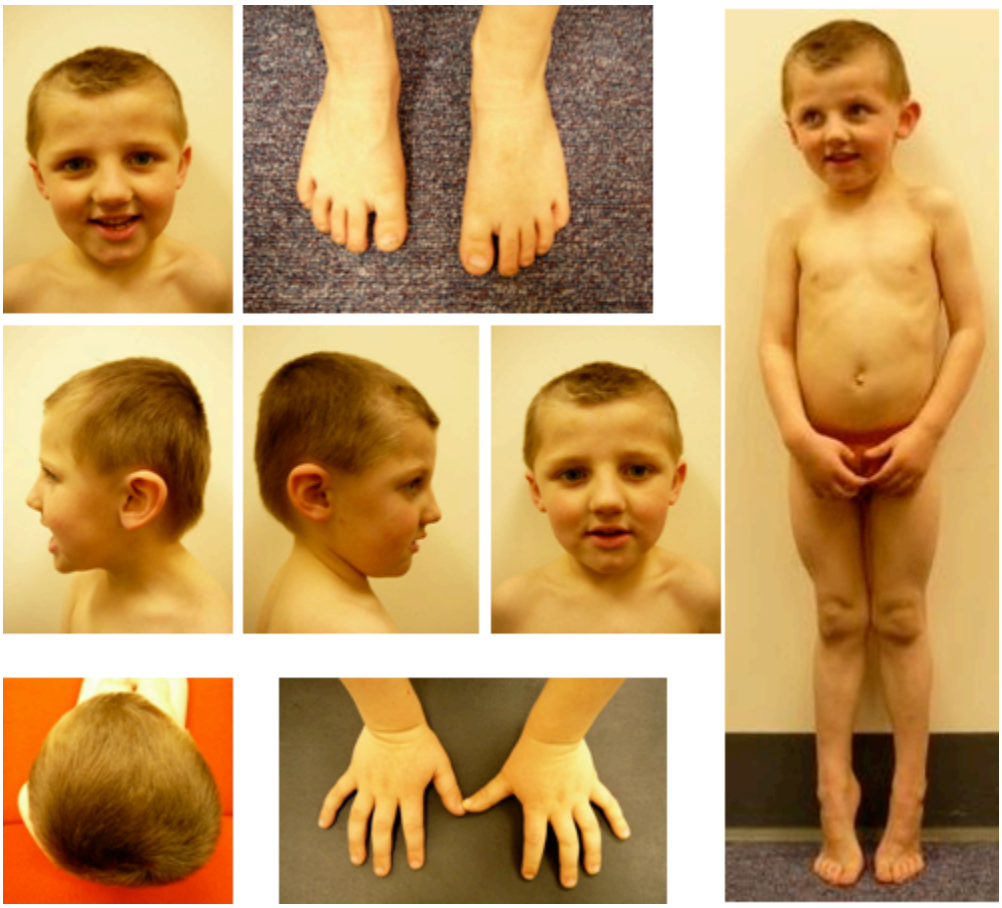


Fig. S2.1. Clinical features in 25514. Photographs of 25514 depicts craniofacial and skeletal features observed. The child is 2 years old.

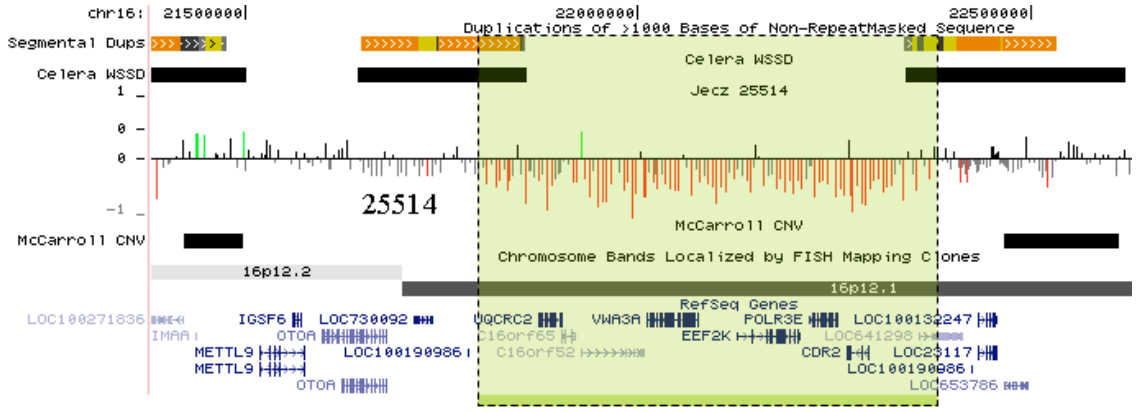


Fig. S2.2. 16p12.1 microdeletion in Case #25514.

Nature Genetics: doi:10.1038/ng.534

3. Supplementary Note: Comparison of 'two hit' phenotypes with second-hit-associated phenotypes

We compared available phenotypes in cases with 16p12.1 microdeletion and a second hit to those from (1) published reports of phenotypes associated with second hit, (2) with DECIPHER [16] data entries for unpublished cases with the second hit, or (3) ECARUCA [17] entries when no DECIPHER or published reports were available. We were able to obtain phenotypic information on four known CNVs from published/DECIPHER sources.

Notes: Table S3.2. About 94% of patients with Phelan-McDermid syndrome (see Table S3.2) meet with the criteria of autism spectrum disorders [18]. N denotes phenotypic data not available. Table S3.3. Hogart et al. (2008) [19] review on chromosome 15q11.2q13 abnormalities; Cook et al. (1997) [20] noted that parental origin for duplications of 15q11-q13; paternal inheritance led to a normal phenotype, whereas maternal inheritance led to autism or atypical autism. Clayton-Smith et al. (1993) [21] reported mental retardation, ataxia, and personality disorders in individuals with 15q11.2q13 duplication. Thomas et al. (2003) [22] reported autism spectrum phenotypes and milder developmental phenotypes in individuals with maternally inherited duplication. N denotes phenotypic data not available.

Table S3.1. Phenotypes associated with homozygous 22q11.2 duplication

	Patient SG12, 16p12 deletion	Patient 4 from Yobb et al*	DECIPHER
Second hit	homozygous dup22q11.2	homozygous dup22q11.2	homozygous dup22q11.2
Age at diagnosis	7.75 years		N
Developmental delay	+	-	present
Learning disabilities	N	+	
Speech delay	N	+	nasal speech
Dysmorphic features	+	+, left ear pit, bulbous nasal tip, broads hands with square tipped fingers and fetal finger pads	telecanthus
Congenital cardiac defect	N	-	N
Hypoplastic heart	N	-	N
Seizures	+	-	N
Psychiatric/behavioral abnormalities	N	+, social immaturity	N
Hearing loss	N	+, due to chronic otitis media	N
Family history of any health/psychiatric problems	Both parents with learning disabilities and dup22q11.2	+, father has learning disability carries 22q11.2 dup; mother has 22q11.2 dup but has only hand abnormalities	

*[23]. N denotes phenotypic data not available.

Nature Genetics: doi:10.1038/ng.534

Table S3.2. Phenotypes associated with 22q13 deletion

	Patient SG03 16p12.1 microdeletion	22q13 deletion (Phelan et al, AJMG)*	DECIPHER 132	DECIPHER 865	DECIPHER 220	DECIPHER 2322
Second hit	2 Mbp 22q13 deletion (chr22:47457855-49549855)	22q13 deletion	22q13 deletion 1.22 Mbp deletion (chr22: 48318025-49534710)	22q13 deletion 3.29 Mbp deletion (chr22: 462-49521447)	22q13 deletion 1.71 Mbp deletion (chr 22: 47740349-49449200)	22q13 deletion 1.59 Mbp (chr22:47937237-49525130)
Birth weight/length/OFC	8 lb 3 oz (75th-90th percentile)	normal growth	N	N	N	N
Weight	13.2 kg (50-75th %ile)	normal growth	N	Obesity	N	N
Length	87.2 cm (25-50th %ile)	normal growth	N	N	N	N
OFC	mild brachycephaly; HC 46 cm (5-10th %ile)	dolicocephaly	N	N	N	N
Developmental delay	+	+	+	+	+	+
Speech delay	+	+	+	Low-pitched voice	N	+
Dysmorphic features	Minor dysmorphology, microcephaly, exotropia, periventricular abnormalities detected by MRI	Minor dysmorphology	N	Macrocephaly	small mandible, micrognathia, hyper-mobile/extensible fingers, adducted thumbs, syndactyly, camptodactyly	
Psychiatric/behavioral abnormalities1	none yet	autistic-like behaviors, stereotypic behaviors	Autism/autistic behavior	Autism/autistic behavior	N	autism/autistic behavior
Hypotonia	+	+	+	N	+	+
Other features	strabismus (repaired)	Common physical traits include long eye lashes, large or unusual ears, relatively large hands, dysplastic toenails, full brow, dolicocephaly, full cheeks, bulbous nose, and pointed chin. Behavior is autistic-like with decreased perception of pain and habitual chewing or mouthing		Has also a 2.87 Mbp, deletion 2q37.3q37.3	N	

Nature Genetics: doi:10.1038/ng.534

Table S3.3. Phenotypes associated with 15q11.2q13.1 duplication

	Patient SG05 with a 16p12.1 deletion	Literature review	DECIPHER 248714	DECIPHER 248550
Second hit	5 Mbp dup 15q11.2q13.1	dup15q11.2q13.1	4.9 Mbp dup15q11.2q13.1	5 Mbp dup15q11.2q13.1
Gestational age	full term	N	N	N
Birth weight/length/OFC	8 lb 2 oz (50th-75th percentile), 21.5 inches (90th-97th percentile), OFC unknown	N	N	N
Age at most recent visit	3 yr, 10 mo	N	N	N
Weight	36.5 lb (50th-75th percentile)	normal growth	N	Obesity
Length	40.75 in (50th-75th percentile)	normal growth	short stature	N
OFC	49.5 cm (0 to -1 SD)	normal growth	N	N
Developmental delay	+	-	+	+
Autism/autism spectrum disorder	NOT autistic	+	N	N
Speech delay	receiving speech therapy	+	N	N
Dysmorphic features	Asymmetric face, small & deep-set eyes, prominent jaw, asymmetric face, crowded teeth	-	N	lowset ears, long philtrum, cleft uvula, small mandible, micrognathia, everted lip, upturned nose
Seizures	occasional staring spells, no known seizure activity	-	+	N
Psychiatric/behavioral abnormalities	N	N	N	self mutilation, aggressive behavior
Hearing loss	PE tubes in past; no reliable audiology evaluation due to poor cooperation	-	N	N
Hypotonia	+	-	N	N
Other features	1 café-au-lait spot		Ataxia	

Nature Genetics: doi:10.1038/ng.534

Table S3.4. Phenotypes associated with 5q15q23.3 duplication

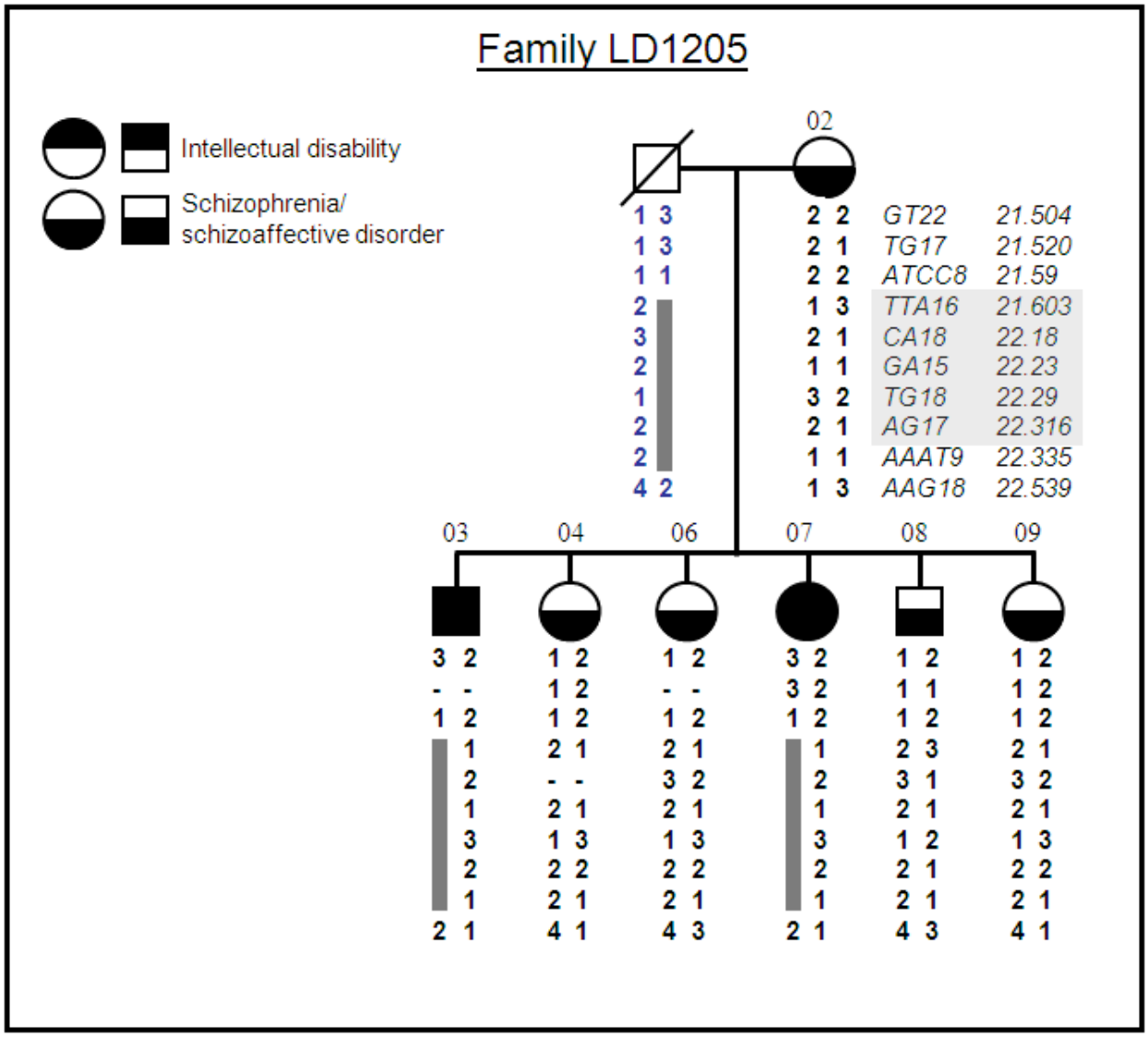
	Patient SG10 with 16p12.1 deletion	Lindgren et al, AJHG, 1992*	Lindgren et al, AJHG, 1992*	DECIPHER 797
Second hit	5q15q23.3 deletion	Del 5q15q21.3	5q15q23.2 microdeletion	5q15q23.2 deletion
Gestational age	38 weeks	38 weeks	N	
Birth weight/length/OFC	3129 g (25th-50th percentile)	3770 g	3900 g	
Age at most recent visit	16 months	13 y	30 y	28 y
Weight	7.9 kg (<3rd percentile)	25th-50th percentile	N	N
Length	72.3 cm (<3rd percentile)	25th-50th percentile	N	N
OFC	46.5 cm (50th percentile)	N	N	N
Developmental delay	+	+	+	N
Speech delay	+	+	N	N
Dysmorphic features	+, exotropia, ectopic pupils, flat nose, wide nasal bridge, philtrum Likert scale 5	+	+, small ears, feet showed metatarsus varsus deformity	joint laxity. Club foot
Congenital cardiac defect	-	-	+	N
Seizures	+	-	+	N
Psychiatric/behavioral abnormalities	decreased activity	-	-	N
Hearing loss	+	-	+	N
Hypotonia	+	-	+	N
Intestinal polyps	N	Single benign polyps	Multiple adenomatous polyps in the intestine	colonic tumors
Other features	Ptosis, hypertelorism, bifid uvula, broad nasal septum, sacral dimple, bridged simian crease of the left hand, clinodactyly of fifth toes bilaterally, bilateral club feet tapering fingers, camptodactyly of left fifth digit	Bilateral posterior pole cataracts and hypoplastic optic discs, hypertelorism, prominent forehead, bulbous nasal tip, prominent jaw (macrognathia), anteverted nostrils, wide philtum, prognathism, high-arched palate	Nystagmus and strabismus, polycystic ovaries	

* Lindgren et al, AJHG, 1992 [24]. N denotes phenotypic data not available.

Nature Genetics: doi:10.1038/ng.534

4. Supplementary Note: Analysis of a multiplex family (LD1205)

In order to evaluate the role of the 16p12.1 microdeletion in very severely affected families with neuropsychiatric disease, we screened probands from 23 multiplex families, each of which included at least three relatives with schizophrenia or schizoaffective disorder. Clinical characteristics of the families have been previously described [3,25]. A genome-wide screen for copy-number variation, initially using ROMA [26], revealed a deletion at 16p12.1 in patient LD1205.03 (Fig. S4.1). The deletion was confirmed by high density array-CGH (Fig. 1, Fig. S4.2), and further evaluation of the family identified the 16p12.1 microdeletion also in one of the proband's sisters (07), but not in the mother (02) or other siblings (04, 06, 08, 09). Genotypes at ten microsatellite markers in the critical 1-Mbp region of 16p12 demonstrated that the deletion was inherited from the deceased father. The phenotypes of members of family LD1205 are interesting in relation to their genotypes.



Nature Genetics: doi:10.1038/ng.534

Fig. S4.1. A family with both schizophrenia and cognitive delay.
Family LD1205 includes two siblings with both schizophrenia and intellectual disability (03 and 07) and four siblings with schizophrenia (04, 06, 08) or major depression (09), but without cognitive impairment. High-density array-CGH (see Methods) and genotypes at microsatellite markers from chromosome 16p12 indicate that siblings 03 and 07 inherited the deletion from their deceased father.

All six siblings in the study were diagnosed with schizophrenia or a major depressive psychosis. However, learning disabilities were diagnosed only in the siblings who harbor the deletion (03 and 07). The mother (02) was diagnosed with a psychotic disorder based on recurrent illness many years prior but was apparently symptom-free when interviewed at age 72 and had no cognitive impairment. No clinical information was available for the deceased father. The co-occurrence of cognitive defects with schizophrenia among siblings carrying the 16p12.1 microdeletion suggests that the deletion may exacerbate the psychosis phenotype, which is due to a different, and still unknown (but plausibly inherited from the mother), cause. The details of clinical diagnoses are given below.

This family lives in Costa Rica and was interviewed by one of us (L.D.) in 1998.
LD1205-02. The subject is the mother of the family, born in 1926 and interviewed at age 72. When she was age 28, she had two episodes, each lasting less than six months, with multiple delusions and auditory hallucinations. In the first episode, she talked nonsense, wandered the streets and was very paranoid. In the second episode, she was anxious, heard voices, and saw shadows. When interviewed at age 72, she was not taking medication and appeared symptom free and with no cognitive impairment. Her diagnosis is schizophreniform disorder.

LD1205-03. This subject is the proband of the family, born in 1949 and interviewed at age 49. He is single with no children and lives with his mother. He completed the 2nd grade of school and his records indicate mild mental retardation. He had a history of auditory hallucinations and multiple delusions, and presented with flat affect, formal thought disorder, and multiple depressive symptoms. He also has a history of alcohol abuse. His diagnosis is schizophrenia, chronic undifferentiated, with mental retardation and alcohol abuse.

LD1205-04. This subject lives with her husband and has two children but both children have been removed from her care. When interviewed, she was very dirty, laughing inappropriately, staring at the floor, and otherwise mute. She was not taking medication when seen. Medical records and her relatives indicate that she has had multiple delusions and hallucinations since about age 13. Her diagnosis is schizophrenia.

LD1205-06. This subject was born in 1955 and was interviewed at age 47. She lives with her husband and 3 of her 6 children. Eight years prior to interview she was hospitalized and medicated because “witches were after her.” She heard voices telling her to keep walking, because figures with masks were coming after her. She had visual hallucinations in which she saw the figures. She has had several subsequent hospitalizations. Her medical records describe schizophrenia and disorganized speech disturbance, beginning at age 17. There is no indication of cognitive impairment. She presents with flat affect, hallucinations, and paranoia, and reports periods of depression and suicidal thoughts. Her diagnosis is paranoid schizophrenia.

LD1205-07. This subject did not understand questions and gave irrelevant answers. She was dirty and unkempt. She was previously married, has 8 children and lives with her mother. Although her

Nature Genetics: doi:10.1038/ng.534

interview was unreliable for her history, the medical records were extensive and clear. She has been frequently hospitalized since age 22 (postpartum), and medical records indicate her onset of symptoms as age 19. The medical records indicate that she has delusions and hallucinations, hears voices, has disorganized speech, and is mentally retarded, all consistent with the interview. Her diagnosis is chronic schizophrenia with mild mental retardation.

LD1205-08. This subject was born in 1970 and interviewed at age 28. He lives with his mother and is unemployed. His behavior was very aggressive, particularly when drinking. He was first hospitalized at age 13 and at least three times subsequently. His medical records indicate paranoid delusions and auditory hallucinations. There was no indication of cognitive dysfunction. He has attempted suicide by hanging. At the interview, he did not speak, had only negative symptoms, and was unkempt and very dirty. His diagnosis is schizophrenia with alcohol abuse.

LC1205-09. This subject was born in 1960 and interviewed at age 38. She lives with her mother and works as a housekeeper outside the home. Her first hospitalization was at age 34 following an attempted suicide. She has had five suicide attempts, by multiple methods, and multiple hospitalizations. Her symptoms include depression and paranoid thoughts. No cognitive disturbance was mentioned or apparent. Her diagnosis is recurrent major depression with psychotic features.

No clinical information is available for the deceased father. A third brother is deceased; no information is available for him. Three additional sisters are not affected, do not live at home, and were not included in the study.

Nature Genetics: doi:10.1038/ng.534

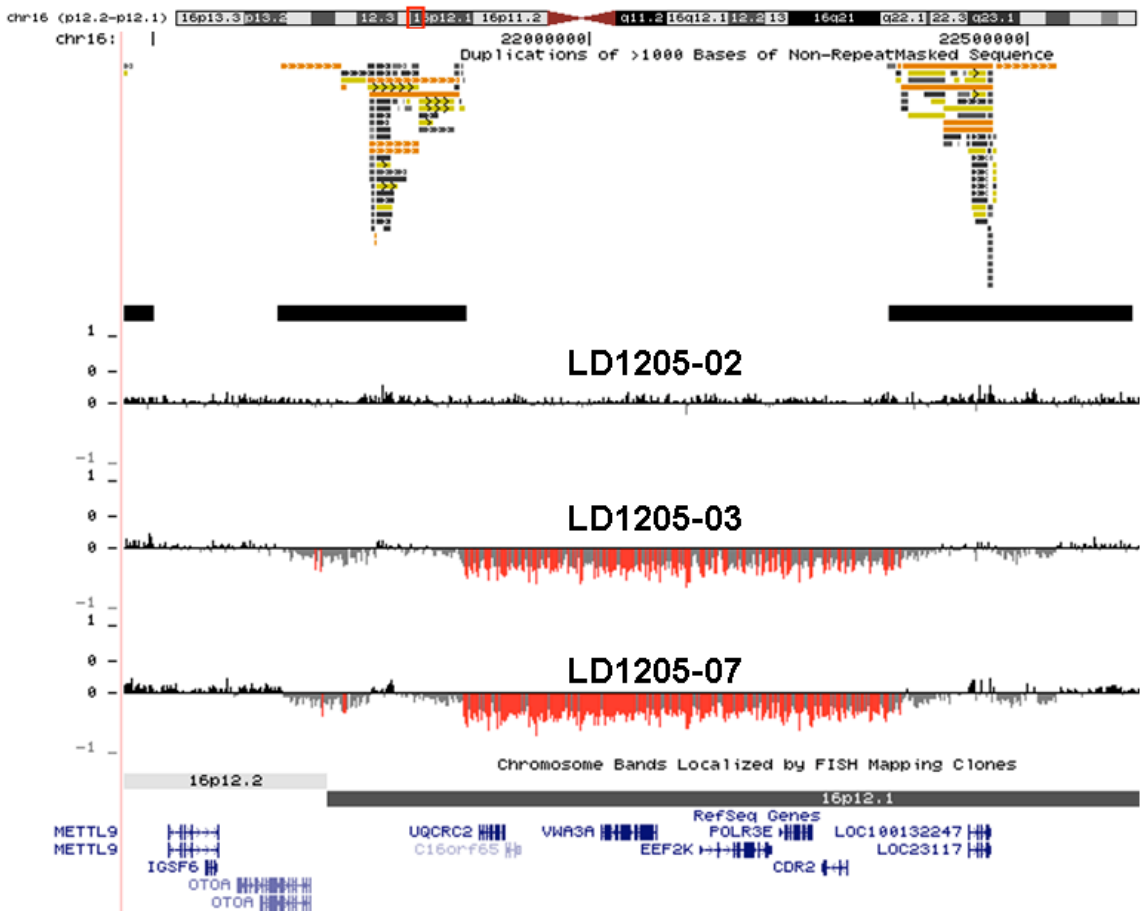


Fig. S4.2. 16p12.1 microdeletion in a multiplex family affected with schizophrenia. The figure shows high resolution microarray analysis on two siblings and the mother from family LD1205.

Nature Genetics: doi:10.1038/ng.534

5. Supplementary Note: Candidate genes within 16p12.1 microdeletion region

Six RefSeq genes are annotated within the 16p12.1 microdeletion region.

Table S5.1. RefSeq genes in the 16p12.1 microdeletion region

Gene ID	Description	Associated phenotype/function*
CDR2	Cerebellar degeneration-related autoantigen 2	Paraneoplastic carcinomas of lung, ovary, breast, and lymph nodes (Hodgkin disease), antibodies (anti-Yo) to CDR2 is a cancer-biomarker
EEF2K	Eukaryotic elongation factor 2 kinase	Calmodulin-related signaling pathway, cytoplasmic protein catalyzes the movement of ribosome along mRNA during translation. *EEF2K* regulates the activity of eukaryotic elongation factor-2, *EEF2*, by phosphorylation. *EEF2* is essential for cardiac myocyte elongation by mediating the translocation step of peptide-chain elongation. *EEF2K* is, in turn, regulated by several factors both covalently and by allosteric stimulation
VWA3A	von Willebrand factor A domain containing 3A	Unknown function
UQCRC2	Ubiquinol-cytochrome C reductase core protein 2	Peripheral membrane protein, involved in the assembly of mitochondrial respiratory chain complex, oxidative-phosphorylation
POLR3E	Polymerase RNA (DNA-directed) III, polypeptide E	Catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates
C16Orf65	Hypothetical protein LOC255762	Unknown

*from Online Mendelian Inheritance in Man.

5.1. Candidate gene sequencing

Three candidate genes were selected for sequencing (*CDR2*, *EEF2K*, and *UQCRC2*). These genes were selected based on their known functions (from OMIM search) or by published reports. DNA samples from four cases with 16p12.1 microdeletion (SG04, SG07, SG11, LD1205-03) and two non-deleted controls were utilized. Sequencing of coding exons did not reveal any mutation.

Table S5.2. Primer sequences for the candidate gene PCR/sequencing

Primer Name	Forward Primer	Reverse Primer
UQCRC2_1	gaaaaggggtgcatactgga	caccttgtcgacacactgct
UQCRC2_2	ggtggttattgctccttcca	gcaaatgggtggttcctaaa
UQCRC2_3	catgccttactctccttggtaa	gcatttccaaacacttgagga
UQCRC2_4	ggcaaacttggcattgaaaa	aagaaatcctcaacctccttga
UQCRC2_5	ctttccaaaggaattggttga	agctgtaatgctctcaatgtgaa
UQCRC2_6	tggattattgttcataaactgctca	tagaatcaaatgtgatctggaaatg
UQCRC2_7	tgaatgccagaagatgacca	tgaaaaagctagcacgctga
UQCRC2_8	tcccagggtaattcacctca	tgttggattctgctttccaa
UQCRC2_9	agtggcattatggcaggaag	atacaacacccaagccatcc
UQCRC2_10	ccgagctgcagaaaagactc	tcaagcctggctatctcttca

Nature Genetics: doi:10.1038/ng.534

UQCRC2_11	ccgtgaaggtccaagttttt	tgcaacttgtccaaccagaa
UQCRC2_12	gcatggacactgtgctgttt	tcccaaagtgctgggattag
UQCRC2_13	tgacttctgccttttattgga	tgcatccttctgaaattgga
CDR2_1	cgctgcccctagaagacc	ccgccacaaaaagcaact
CDR2_2	aaattgttgatgtgagatcatttg	ggcataacagcaacacatcg
CDR2_3	tttgtgtcacctggtcctga	tccccctagtccttccttgt
CDR2_4	aagcagcagtggcttctgtc	cagtccctgctgtaactcca
CDR2_5a	caccagtcttcctctcgtca	CTCCTGATGCAGGTCTCCTC
CDR2_5b	GCTGGTGCCAGACTCTCTGT	ggagagagaggcgataggc
EEF2K_01	ttgtaggaccttcgcctctg	ccagggtttggaaatacgaa
EEF2K_02	tgcctgtgagtgcagatagg	tgtgagaaccacgagaagga
EEF2K_03	GGAGAGTGGGTTGGTTCTGA	AACAAAGAATCCCCCGAAGT
EEF2K_04	aaggcctctatggctttggt	ctgttcttcctgagcccatc
EEF2K_05	ctgtcccagctctttgaagg	ATTCACCATGGCCCTCTTTC
EEF2K_06	attctgggcatcatcaggag	agaagctgtcatggggtgac
EEF2K_07	aagagaaagcccctcagcat	tatggccagggcaggtataa
EEF2K_08	cactgtacgcgctttttcaa	ttccagccctggtattcatc
EEF2K_09	ccctttcttggaggccttag	ctgtcccctccaatcacagt
EEF2K_10	tttagcctccaggtgctgat	tggaggctgacattcatctg
EEF2K_11	atgggacagaggggagaaag	catgaccggcctgaaatact
EEF2K_12	gcagagaagggttccagtgt	ccctgcagaggttcaaagag
EEF2K_13	acatcgtgagggttcactcc	tggacctgccacacagacta
EEF2K_14	aagtccattttggggaaggt	aaccacagggtgaccaagtc
EEF2K_15	gccccctcctatctcacttc	aataagttgcccatggttgc
EEF2K_16	acctccacctttccctctgt	acccccctacaaagccgtaat
EEF2K_17	agggcacacggagtagtttg	gtagctgccaaggagcaaag

Primers in small case map within introns and upper case primers map within exons.

Nature Genetics: doi:10.1038/ng.534

6. Supplementary Note: Haplotype analysis in patients and available parental DNA

We genotyped 32 individuals comprised of 17 probands, four available parents, one affected sibling, and 11 controls to determine:

1. If there is one specific sequence haplotype that predisposes to this microdeletion
2. If there is a specific haplotype background (founder effect) that harbors this event (since 96% of the probands inherited this microdeletion).

SNP-based genotyping analysis was performed using Illumina 550K chip and genotyping calls were analyzed for chromosome 16p as described below. The genotypic calls were phased and identity-by-descent (IBD) detection was performed for cryptic relatedness. Phasing of the genotypic calls was performed using Beagle [27], using Mendelian parent-offspring information when available [28]. We used Illumina 550K data from the 1958 British Birth Cohort (1958BC) [29] as a "reference panel" in our analyses since the use of additional individuals improves estimation of haplotype frequencies. We directly examined the phased haplotypes, and also clustered the haplotypes using Beagle [27]. Investigation of the haplotypes and haplotype clusters suggested that the deletions occurred on different haplotypes, however, it is it is difficult to be certain since phasing of haplotypes from unrelated individuals has limited accuracy.

We used an extension for diploid data of an IBD detection algorithm for haploid data previously described [30]. To do this analysis, we estimated IBD probabilities between all pairs of unrelated individuals in the genotyping data, excluding parents and one of the full sibs, and between the known parent-child pairs. We took out the SNPs in the region covered by the deletion, and put in a bi-allelic marker for the deletion (alleles were deletion and non-deletion). We calculated posterior probabilities of IBD for each pair at each position. Strong evidence for IBD was observed for the parent-child pairs and for two unrelated probands SG06 and SG12. This latter pair was estimated to share a haplotype IBD for approximately 3 cM ending at or shortly after the deletion. The plot below shows the estimated IBD probabilities for the two unrelated probands with evidence of IBD.

For comparison, we tested 10,000 randomly selected pairs of individuals from the 1958BC. We tested these for IBD using the same set of markers and linkage disequilibrium model as for analysis of the genotyping data described above. We found 4 pairs out of 10,000 with evidence of IBD in the 1958BC data, with IBD lengths 1.5, 2, 5 and 5 cM respectively. The finding of one pair with IBD out of 136 pairs of unrelated probands tested compared to 4 out of 10,000 for the 1958BC data has marginal statistical significance (p=0.065; Fisher's exact test).

Nature Genetics: doi:10.1038/ng.534

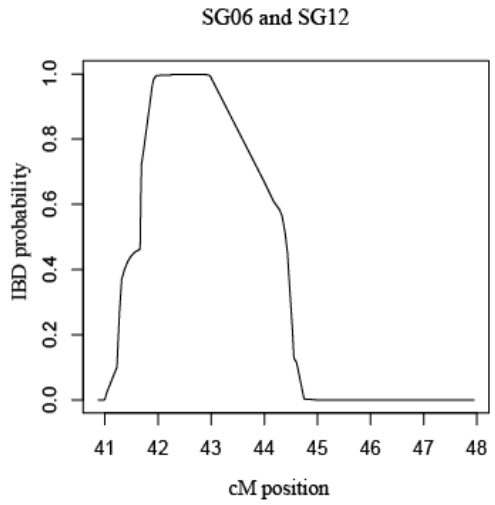


Fig. S6.1. IBD probability in unrelated probands. The cM positions are taken from the HapMap phase 2 build35 estimated recombination map [31]. The deletion is located at 43.0 cM; there are no SNPs between the deletion and 44.0 cM in the data used to estimate IBD probabilities.

Our IBD detection method estimates the haplotype that is shared IBD, along with a probability that the shared allele at each marker is correctly identified. When we build our haplotype frequency model with a large cohort (in this case the 1958BC), the probability that the shared allele is correctly identified is typically high (>0.995) in the interior of the IBD region, and consequently, the estimated haplotypes are generally highly accurate. Using IBD, we estimated the transmitted deletion haplotypes from two genotyped proband-parent pairs (SG04 and mother, and SG05 and SG14), the shared deletion haplotype between a proband and affected sibling (MC157 and MC160), and the deletion haplotype shared IBD between probands SG06 and SG12. All 4 haplotypes differed in the immediate vicinity of the deletion which suggests that the microdeletion is caused by recent recurrent events.

The 1958 British Birth Cohort data genotyped on the Illumina 550K chip that we used was generated by the Wellcome Trust Sanger Institute in collaboration with the 1958 British Birth Cohort, but is being distributed as part of the Wellcome Trust Case Control Consortium (WTCCC). A full list of the investigators who contributed to the generation of the WTCCC data is available from www.wtccc.org.uk. Funding for the Wellcome Trust Case Control Consortium project was provided by the Wellcome Trust under award 076113.

Nature Genetics: doi:10.1038/ng.534

7. Supplementary Note: Analysis for two hits in 1q21.1 deletion patients

As a preliminary test, we began by reanalyzing the frequency and effect of two hits in cases with 1q21.1 microdeletion which we reported recently (Mefford et al, 2008) [32]. Among 16% (4/25) of the 1q21.1 microdeletion, a second hit was identified among children with generally more severe phenotypes. These cases with two hits included patients with severe neurological deficit and craniofacial dysmorphology (S1), overt craniofacial features (S2) and a case with severe schizophrenia without cognitive impairment (S5) (Table S7.1). The frequency of the double hit among 1q21.1 patients (16%) is 40-fold enriched when compared to normal controls (0.39%) and is also significantly enriched when conditioning on controls for the presence of a large CNV first hit (*p=0.03*). The second hit CNVs observed in 1q21.1 microdeletion patients are shown in Fig. S7.1 (Mefford and Eichler).

Table S7.1. Abnormalities in cases with 1q21.1 microdeletion and a second hit

ID	Cognitive	Neurological	Craniofacial	Skeletal	Cardiac	Psychiatric
S1	Not assessed	hypotonia	Short stub nose	Severe growth retardation	Small persistent foramen ovale	Absent
S2	Early delays	hypotonia	Asymmetry of cranial vault, facial dysmorphism	Bilateral clinodactyly	Not assessed	Absent
S3	Severe MR, motor delay	Ataxia, hypotonia	Palatoschisis, micrognathia, high, broad nasal bridge, dysplastic low-set ears	Bilateral clinodactyly	Atrial septal defect	Absent
S5 *	Absent	Absent	Not assessed	Absent	Not assessed	Schizophrenia, severe paranoia, delusions, and catatonia

*S5 carries a 671-kbp 2p16.3 duplication in addition to the 1q21.1 deletion (reported by Walsh et al, Science, 2008); Shaded boxes indicate severe phenotypes and/or and variable features among individuals with the 1q21.1 deletion (Mefford et al, NEJM, 2008).

Nature Genetics: doi:10.1038/ng.534

Second hits in 1q21.1 microdeletion cases

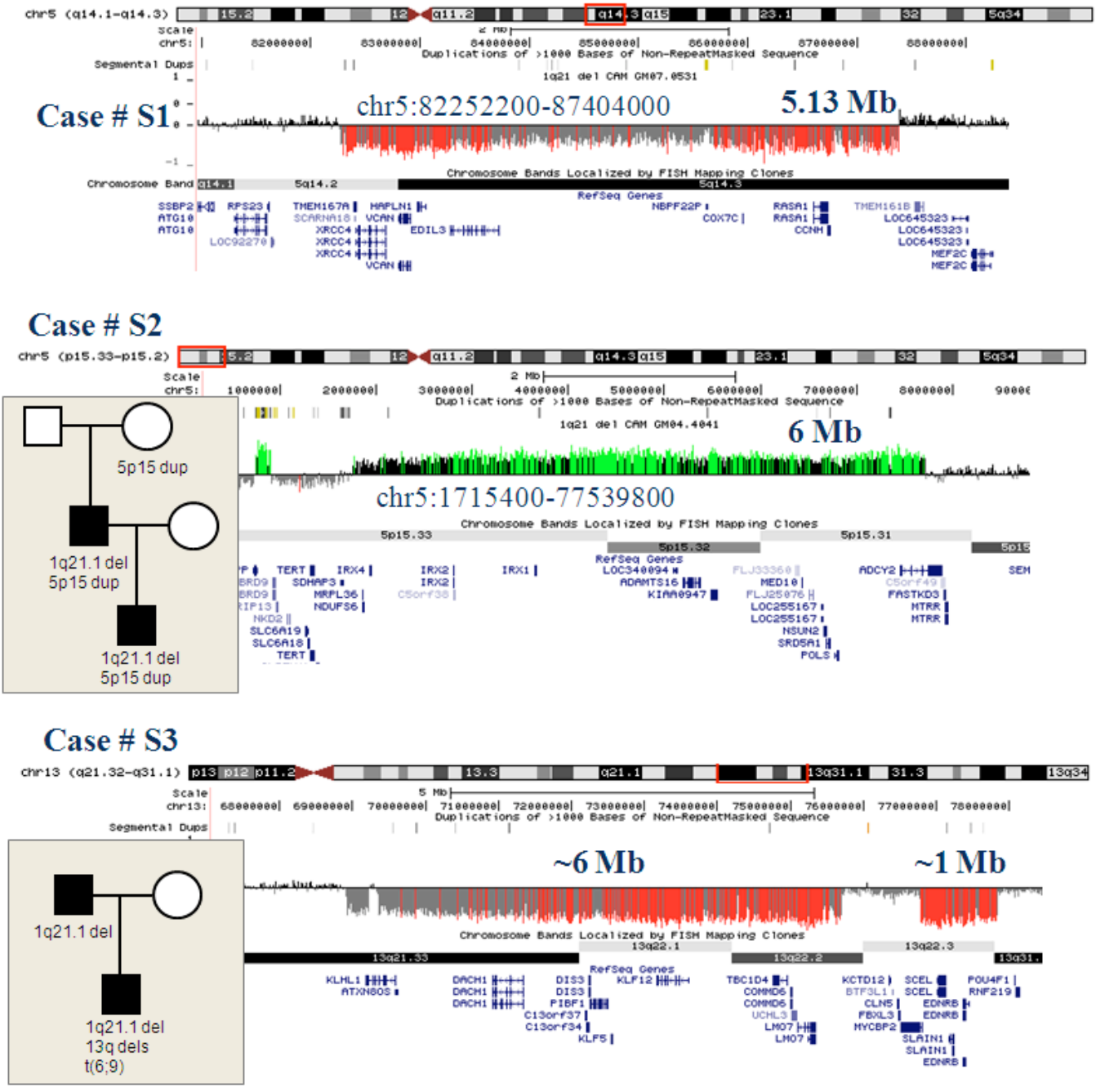


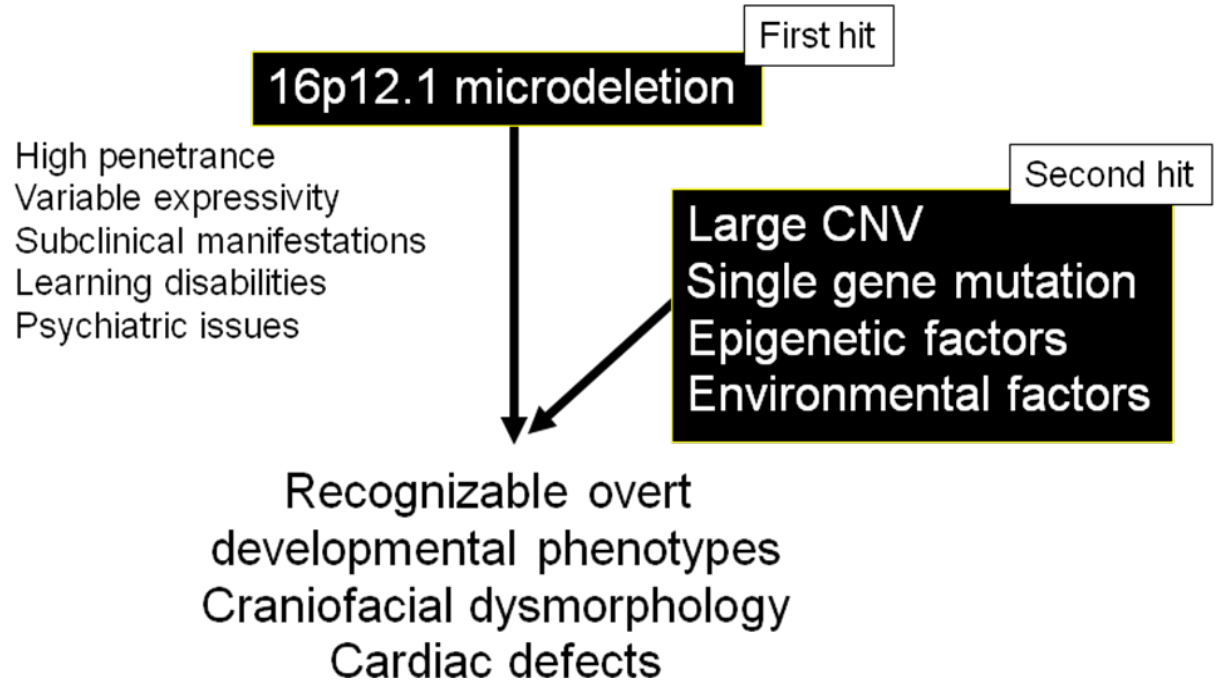
Figure S7.1: Large CNVs in 1q21.1 microdeletion patients.

The 1q21.1 microdeletion in case#1 arose *de novo*, while case #S2 inherited both the deletion and the 5p15 duplication from the affected father and case #S3 carries two deletions on chr13q. Note, case #S2 also carries a der(8)t(8;21)(q24.3;q22.3).

Nature Genetics: doi:10.1038/ng.534

8. Supplementary Note: A two-hit model for 16p12.1 phenotypes

Our analysis suggests that deletion of 16p12.1 is an independent risk factor for intellectual disability and developmental delay. Further whole-genome analysis shows that the 16p12.1 deletion acts in concert with other factors to alter neurological phenotypes. We propose a "two-hit" model, wherein a secondary insult is necessary during development to result in a more severe clinical manifestation as pediatric disease. The second hit could potentially be another CNV, mutation in a phenotypically-related gene, or an environmental event influencing the phenotype.



Nature Genetics: doi:10.1038/ng.534

9. Supplementary Note: References

1. Ballif, B.C. et al. Expanding the clinical phenotype of the 3q29 microdeletion syndrome and characterization of the reciprocal microduplication. *Mol Cytogenet* **1**, 8 (2008).
2. Ballif, B.C. et al. Identification of a previously unrecognized microdeletion syndrome of 16q11.2q12.2. *Clin Genet* **74**, 469-75 (2008).
3. DeLisi, L.E. et al. Clinical characteristics of schizophrenia in multiply affected Spanish origin families from Costa Rica. *Psychiatr Genet* **11**, 145-52 (2001).
4. DeLisi, L.E. et al. Genome-wide scan for linkage to schizophrenia in a Spanish-origin cohort from Costa Rica. *Am J Med Genet* **114**, 497-508 (2002).
5. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237-41 (2008).
6. Manolio, T.A. et al. New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* **39**, 1045-51 (2007).
7. Itsara, A. et al. Population analysis of large copy number variants and hotspots of human genetic disease. *Am J Hum Genet* **84**, 148-61 (2009).
8. Kirov, G. et al. Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum Mol Genet* **18**, 1497-503 (2009).
9. Shaikh, T.H. et al. High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res* **19**, 1682-90 (2009).
10. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).
11. Cooper, G.M., Zerr, T., Kidd, J.M., Eichler, E.E. & Nickerson, D.A. Systematic assessment of copy number variant detection via genome-wide SNP genotyping. *Nat Genet* **40**, 1199-203 (2008).
12. McCarroll, S.A. et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* **40**, 1166-74 (2008).
13. Redon, R. et al. Global variation in copy number in the human genome. *Nature* **444**, 444-54 (2006).
14. Coppinger, J. et al. Identification of familial and de novo microduplications of 22q11.21-q11.23 distal to the 22q11.21 microdeletion syndrome region. *Hum Mol Genet* **18**, 1377-83 (2009).
15. Selzer, R.R. et al. Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. *Genes Chromosomes Cancer* **44**, 305-19 (2005).
16. Firth, H.V. et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet* **84**, 524-33 (2009).
17. Feenstra, I. et al. European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations (ECARUCA); an online database for rare chromosome abnormalities. *Eur J Med Genet* **49**, 279-91 (2006).
18. Bill, B.R. & Geschwind, D.H. Genetic advances in autism: heterogeneity and convergence on shared pathways. *Curr Opin Genet Dev* **19**, 271-8 (2009).
19. Hogart, A., Wu, D., Lasalle, J.M. & Schanen, N.C. The comorbidity of autism with the genomic disorders of chromosome 15q11.2-q13. *Neurobiol Dis* (2008).
20. Cook, E.H., Jr. et al. Autism or atypical autism in maternally but not paternally derived proximal 15q duplication. *Am J Hum Genet* **60**, 928-34 (1997).

Nature Genetics: doi:10.1038/ng.534

21. Clayton-Smith, J., Webb, T., Cheng, X.J., Pembrey, M.E. & Malcolm, S. Duplication of chromosome 15 in the region 15q11-13 in a patient with developmental delay and ataxia with similarities to Angelman syndrome. *J Med Genet* **30**, 529-31 (1993).
22. Thomas, J.A. et al. Genetic and clinical characterization of patients with an interstitial duplication 15q11-q13, emphasizing behavioral phenotype and response to treatment. *Am J Med Genet A* **119A**, 111-20 (2003).
23. Yobb, T.M. et al. Microduplication and triplication of 22q11.2: a highly variable syndrome. *Am J Hum Genet* **76**, 865-76 (2005).
24. Lindgren, V., Bryke, C.R., Ozcelik, T., Yang-Feng, T.L. & Francke, U. Phenotypic, cytogenetic, and molecular studies of three patients with constitutional deletions of chromosome 5 in the region of the gene for familial adenomatous polyposis. *Am J Hum Genet* **50**, 988-97 (1992).
25. Shaw, S.H. et al. A genome-wide search for schizophrenia susceptibility genes. *Am J Med Genet* **81**, 364-76 (1998).
26. Sebat, J. et al. Large-scale copy number polymorphism in the human genome. *Science* **305**, 525-8 (2004).
27. Browning, S.R. & Browning, B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* **81**, 1084-97 (2007).
28. Browning, B.L. & Browning, S.R. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet* **84**, 210-23 (2009).
29. Power, C. & Elliott, J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol* **35**, 34-41 (2006).
30. Browning, S.R. Estimation of pairwise identity by descent from dense genetic marker data in a population sample of haplotypes. *Genetics* **178**, 2123-32 (2008).
31. Frazer, K.A. et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851-61 (2007).
32. Mefford, H.C. et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Engl J Med* **359**, 1685-99 (2008).

Nature Genetics: doi:10.1038/ng.534