Supplementary Information

A Copy Number Variation Morbidity Map of Developmental Delay

Gregory M. Cooper[1*#], Bradley P. Coe[1#], Santhosh Girirajan[1#], Jill A. Rosenfeld[2], Tiffany Vu[1], Carl Baker[1], Charles Williams[3], Heather Stalker[3], Rizwan Hamid[4], Vickie Hannig[4], Hoda Abdel-Hamid[5], Patricia Bader[6], Elizabeth McCracken[7], Dmitriy Niyazov[8], Kathleen Leppig[9], Heidi Thiese[9], Marybeth Hummel[10], Nora Alexander[10], Jerome Gorski[11], Jennifer Kussmann[11], Vandana Shashi[12], Krys Johnson[12],Catherine Rehder[13], Blake C. Ballif[2], Lisa G. Shaffer[2] and Evan E. Eichler[1,14]

[1]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA
[2]Signature Genomic Laboratories, LLC, Spokane, WA 99207, USA
[3]Department of Pediatrics, Division of Genetics, University of Florida, Gainesville, FL 32610, USA
[4]Vanderbilt University Medical Center, Nashville, TN 37232, USA
[5]Department of Pediatrics, Division of Child Neurology, University of Pittsburgh, Pittsburgh, PA 15201, USA
[6]Northeast Indiana Genetic Counseling Center, Ft. Wayne, IN 46845, USA
[7]Children's Hospital Pittsburgh, Pittsburgh, PA 15201, USA
[8]Ochsner Clinic, New Orleans, LA 70121, USA
[9]Group Health Cooperative, Seattle, WA 98112, USA
[10]West Virginia University, Morgantown, WV 26506, USA
[11]University of Missouri, Columbia, MO 65212, USA
[12]Departments of Pediatrics and Pathology, Duke University Medical Center, Durham, NC 27705, USA
[13]Clinical Molecular Diagnostic Laboratory, Duke University Health System, Durham, NC 27704, USA
[14]Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA
*Present address: HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA
[#]contributed equally to work

Contents

Nature Genetics: doi.10.1038/ng.909

Nature Genetics: doi.10.1038/ng.909

Supplementary Table 1. Phenotype by sample. (See Excel sheet)

Supplementary Table 2. Comorbid phenotypes and subclassifications.

Comorbidity	Neurological	Epilepsy	Autism Spectrum	ADD / ADHD	MCA	Craniofacial	Cardiovascular	Kidney	Hearing Loss	FTT
Neurological	8772									
Epilepsy	1002	1776								
Autism Spectrum	612	139	1379							
ADD/ADHD	102	15	29	247						
MCA	734	118	42	12	2247					
Craniofacial	2717	346	150	42	655	3898				
Cardiovascular	134	18	21	12	198	123	575			
Kidney	23	2	2	3	21	26	18	93		
Hearing Loss	70	9	8	2	15	26	7	3	121	
FTT	459	48	28	7	50	235	0	9	7	883

Supplementary Table 3. Size-wise validation of CNVs.

	>150 kbp to 400 kbp	>400 kbp to 1 Mbp	>1 Mbp	Chip
PARK2	4	11	2	High density custom targeted
EYS	1	3	5	High density custom targeted
AUTS2	7	0	0	High density custom targeted
16p12.1	0	33	5	High density custom targeted
16p11.2	0	77	0	High density custom targeted
17q21.31	0	3	0	High density custom targeted
Large CNVs	5	12	19	Hotspot 1
Large CNVs	0	5	5	High density custom targeted
15q25 CNVs	0	5	0	High density custom targeted
Subtelomeric CNVs	5	2	12	High density custom targeted
Large CNVs	0	45	93	FISH
Large CNVs	8	3	10	Hotspot 1
Large CNVs	0	38	7	Hotspot 1
Total	30	237	158	
Did not validate	3	1	19	
Validated	27	236	139	

Nature Genetics: doi.10.1038/ng.909

Supplementary Table 4. Description of control cohorts.

DataSet	DNA Source	Platform	Illumina Map Build	Calls	QC Passing Samples (hg18)	Ethnicity	Phenotypes
HGDP	Cell Line	HumanHap650Yv3_A	hg18	hg18 Native	984	Yoruba: 24, Yi: 10, Yakut: 23, Xibo: 8, Uygur: 10, Tuscan: 7, Tujia: 10, Tu: 9, Surui: 20, Sindhi: 23, She: 9, Sardinian: 27, San: 5, Russian: 24, Pima: 24, Pathan: 23, Papuan: 15, Palestinian: 49, Oroqen: 10, Orcadian: 14, Naxi: 8, Mozabite: 30, Mongola: 10, Miao: 9, Melanesian: 15, MbutiPygmy: 13, Maya: 21, Mandenka: 24, Makrani: 25, Lahu: 10, Karitiana: 22, Kalash: 24, Japanese: 27, Italian: 12, Hezhen: 9, Hazara: 21, Han-NChina: 10, Han: 33, French: 26, Druze: 45, Daur: 9, Dai: 6, Colombian: 13, Cambodian: 10, Burusho: 24, Brahui: 25, BiakaPygmy: 30, Bedouin: 46, Basque: 23, BantuSouthAfrica: 8, BantuKenya: 12, Balochi: 25, Adygei: 15,	Normal Individuals from Various Populations
NINDS (Coriell 550K)	Cell Line	Illumina 550K	hg18	hg18 Native	441	Caucasian	Neurologically normal individuals
NINDS (317K+240K)	Cell Line	Illumina317K+240K	hg17	hg17 -> 18 LiftOver	227	Caucasian	Neurologically normal individuals
PARC (CAP and PRINCE)	Cell Line	Illumina 317K	hg17	hg17 -> 18 LiftOver	936	European	Middle-age (40–70 years) individuals of European descent living in the United States with moderately high levels of total cholesterol
London (Asthmatic Parents)	Blood	Illumina 550K	hg18	hg18 Native	760	Mexican	Parents of asthmatic children (Mexican)
PARC2 (CAP2)	Cell Line	Illumina 610K Quad	hg18	hg18 Native	232	European	Middle-age (40–70 years) individuals of European descent living in the United States with moderately high levels of total cholesterol
PARC2 (Prince2)	Cell Line	Illumina 610K Quad	hg18	hg18 Native	534	European	Middle-age (40–70 years) individuals of European descent living in the United States with moderately high levels of total cholesterol
FHCRC	Blood	Human610-Quadv1_B	hg18	hg18 Native	1430	Asian/Pacific Islander: 20, Black: 47, Hispanic: 28, White: 1334	Post-menopausal (50-79 years) Female controls for pancreatic cancer, colon cancer, and cases and controls for a hip

Nature Genetics: doi.10.1038/ng.909

							fracture study
inChianti	Blood	HumanHap550v3_A	hg18	hg18 Native	695	Caucasian	Population-based study of older persons living in the Chianti geographic area
WTCCC2(NBS)	Blood	Custom Illumina 1.2M	hg18	hg18 Native	2090	Caucasian	UK Blood Service Control Group (blood donors, age range 18–69 years)

Supplementary Table 5. Definition of genomic disorder CNVs used for Table 1.

Chr	Start	End	Deletion Syndrome	Duplication Syndrome	Definition
chr1	0.00	10.00	1p36 deletion syndrome (*GABRD*)*	None	maximum
chr1	144.00	144.34	TAR deletion (*HFE2*)	None	exact
chr1	145.04	145.86	1q21.1 deletion (*GJA5*)	1q21.1 duplication	exact
chr2	57.60	61.59	2p15-16.1 microdeletion syndrome (*VRK2*)*	None	critical
chr2	96.09	97.04	2q11.2 deletion (*LMAN2L*, *ARID5A*)	2q11.2 duplication	exact
chr2	100.06	107.81	2q11.2q13 deletion (*NCK2*, *FHL2*)	None	exact
chr2	110.18	110.34	2q13 deletion (*NPHP1*)	None	exact
chr2	239.37	242.12	2q37 deletion (*HDAC4*)*	None	critical
chr3	197.23	198.84	3q29 deletion (*DLG1*)	3q29 duplication	exact
chr4	1.84	1.98	Wolf-Hirschhorn deletion (*WHSC1*, *WHSC2*)*	None	critical
chr5	175.65	176.99	Sotos syndrome deletion (*NSD1*)	None	exact
chr6	92.10	104.80	6q16 deletion (*FOXP1*, *SIM1*)*	None	critical
chr7	66.12	71.91	WBS-prox deletion (*AUTS2*)	WBS-prox duplication	exact
chr7	72.38	73.78	Williams syndrome deletion (*ELN*, *GTF2I*)	WBS duplication	exact
chr7	74.80	76.50	WBS-distal deletion (*RHBDD2*, *HIP1*)	WBS-distal duplication	exact
chr8	8.13	11.93	8p23.1 deletion (*SOX7*, *CLDN23*)	None	exact
chr9	136.95	140.20	9q34 deletion (*EHMT1*)*	9q34 duplication	critical
chr10	81.95	88.79	10q23 deletion (*NRG3*, *GRID1*)	None	exact
chr11	43.94	46.02	Potocki-Shaffer syndrome (*EXT2*)*	None	critical
chr11	67.51	70.96	SHANK2 FGFs deletion	None	exact
chr12	63.36	66.93	12q14 microdeletion syndrome (*GRIP1*, *HMGA2*)*	None	critical
chr13	19.71	19.91	13q12 deletion (*CRYL1*)*	None	critical
chr15	20.35	20.64	15q11.2 deletion (*NIPA1*)	None	exact
chr15	22.37	26.10	Prader-Willi/Angelman	PWS duplication	exact
chr15	28.92	30.27	15q13.3 deletion (*CHRNA7*)	15q13.3 duplication	exact
chr15	70.70	72.20	15q24 BP0-BP1 deletion (*BBS4*, *NPTN*, *NEO1*)	None	exact
chr15	70.70	73.58	15q24 BP0-BP1 (*PML*)	None	exact
chr15	72.20	73.38	15q24 critical BP1-2 deletion (*CLK3*,	None	exact

Nature Genetics: doi.10.1038/ng.909

			CSK)		
chr15	73.38	73.58	15q24 small deletion (*SIN3A*)	None	exact
chr15	73.76	75.99	15q24 BP2-BP3 deletion (*FBXO22, TPSAN3*)	None	exact
chr15	80.98	82.53	15q25.2 deletion (*HOMER2, BNC1*)	None	exact
chr15	97.18	100.34	None	15q26 overgrowth syndrome	critical
chr16	3.72	3.80	Rubinstein-Taybi Syndrome*	None	critical
chr16	15.41	16.20	16p13.11 deletion (*MYH11*)	16p13.11 duplication	exact
chr16	21.26	29.35	16p11.2p12.1 deletion	None	exact
chr16	21.85	22.37	16p12.1 deletion (*EEF2K, CDR2*)	None	exact
chr16	28.68	29.02	16p11.2 distal deletion (*SH2B1*)	None	exact
chr16	29.56	30.11	16p11.2 deletion (*TBX6*)	16p11.2 duplication	exact
chr17	0.05	2.54	17p13.3 deletion (both *YWHAE* and *PAFAH1B1*)*	17p13.3 duplication (both *YWHAE* and *PAFAH1B1*)	critical
chr17	0.50	1.30	17p13.3 deletion (including *PAFAH1B1*)*	17p13.3 duplication (including *PAFAH1B1*)	critical
chr17	2.31	2.87	17p13.3 deletion (including *YWHAE*)*	17p13.3 duplication (including *YWHAE*)	critical
chr17	14.01	15.44	HNPP (*PMP22*)	CMT1A	exact
chr17	16.65	20.42	Smith-Magenis syndrome deletion	Potocki-Lupski syndrome duplication	critical
chr17	26.19	27.24	NF1 microdeletion syndrome	None	exact
chr17	31.89	33.28	RCAD (renal cysts and diabetes) (*TCF2*)	17q12 duplication	exact
chr17	41.06	41.54	17q21.31 deletion (*MAPT*)	17q21.31 duplication	exact
chr17	55.01	55.43	17q23 deletion (*TUBD1, TMEM49*)	None	exact
chr17	55.42	57.66	17q23.1q23.2 deletion (*TBX2, TBX4*)	None	exact
chr22	17.40	18.67	DiGeorge/VCFS deletion	22q11.2 duplication	critical
chr22	20.24	21.98	22q11.2 distal deletion (*BCR, MAPK1*)	22q11.2 distal duplication	exact
chr22	49.46	49.52	Phelan-McDermid syndrome deletion (*SHANK3*)*	None	Covers *SHANK3* and is >500 kbp

All coordinates are according to build36. The genes in parentheses are potential candidate genes and identifiers of the genomic locations. *Denotes rearrangements not mediated by segmental duplications; VCFS – velocardiofacial syndrome, WBS – Williams-Beuren syndrome, HNPP – hereditary neuropathy with liability to pressure palsies, CMT1A – Charcot-Marie-Tooth disease type 1A.

Nature Genetics: doi.10.1038/ng.909

Supplementary Table 6. Frequency of classical genomic disorders in our cohort compared to the frequency known in the literature

Genomic disorder	Cases (n=15767)	*Observed frequency in the general population	*Reported frequency in the general population from literature	Reference
Williams syndrome deletion	42	0.6 in 10,000	1 in 10,000	Pober, B, NEJM, 2008; 6. Strømme P, J Child Neurol, 2002
Prader-Willi/Angelman	16	1 in 50,000	1 in 25,000 for Prader-Willi syndrome and 1 in 12,000 to 1 in 20,000 for Angelman	Whittington JE, JMG, 2001; Vogels, EJHG, 2004; 4. Thomson AK, Disabil Rehabil, 2006
Smith-Magenis syndrome deletion	16	1 in 50,000	1 in 15,000 to 1 in 25,000	Greenberg, AJHG, 1991 and AJMG, 1996; Girirajan and Elsea, EJHG, 2008
DiGeorge/VCFS deletion	96	0.7 per 5,000	1 in 3,500 to 1 in 5,000	Shprintzen, RJ, Dev Disabil Res Rev, 2008

*Calculations were made based on a 2.3% frequency of ID phenotypes in the general population (Ropers, Ann Rev Hum Genet Genom, 2011).

Supplementary Table 7. Novel pathogenic genomic regions identified by sliding window analysis

NR set	Chrom	Start	End	Length	Type	Sig.pval	HotSpot	Telomere	Notes	Description	
1	chr1	0.7	9.9	9.2	del	2.21E-14	No	Yes	1tel	1p36	
1	chr1	0.85	3.65	2.8	dup	1.13E-03	No	Yes	1tel	1p36	
2	chr1	144.05	146.35	2.3	del	3.93E-10	Yes	No	1q21	1q21.1	
2	chr1	144.05	148.05	4	dup	2.31E-07	Yes	No	1q21	1q21.1	
3	chr1	170	170.6	0.6	del	7.85E-02	No	No	VUS	1q24.3	*
4	chr1	240.5	244.8	4.3	del	2.20E-02	No	Yes	1tel	1q44	
5	chr1	243.1	246.9	3.8	dup	4.54E-05	No	Yes	1tel	1q44	
6	chr10	2.6	3.2	0.6	dup	1.53E-03	No	No	VUS	10p15.3	*
7	chr10	46.35	48.05	1.7	del	1.35E-04	Yes	No	10q11.22	10q11.2	
7	chr10	46.35	48.05	1.7	dup	3.36E-05	Yes	No	10q11.22	10q11.2	
8	chr10	81.6	88.9	7.3	del	1.44E-02	Yes	No	10q23.1	10q23.1	*
9	chr10	127.75	135.25	7.5	del	3.53E-07	No	Yes	10tel	10q26.2	
10	chr11	0.3	3.4	3.1	del	1.35E-04	No	Yes	11tel	11p15.5	
11	chr11	127.55	134.35	6.8	del	2.63E-03	No	Yes	11tel	11q24.3	
12	chr12	0.1	3.5	3.4	dup	2.20E-02	No	Yes	12tel	12p13.3	
12	chr12	0.15	0.65	0.5	del	7.85E-02	No	Yes	12tel	12p13.3	
13	chr12	8.05	8.25	0.2	dup	7.85E-02	No	No	VUS	12p13.31	*
14	chr13	18.3	19.2	0.9	dup	5.14E-02	No	Yes	13tel	13q34	
15	chr13	112.65	114.05	1.4	del	7.85E-02	No	Yes	13tel	13p13	

Nature Genetics: doi.10.1038/ng.909

16	chr14	35.5	36.3	0.8	del	5.14E-02	No	No	VUS	14q13.2	
17	chr14	103.55	105.45	1.9	del	1.13E-03	No	Yes	14tel	14q32.33	
18	chr15	20.2	26.3	6.1	del	4.10E-06	Yes	No	PW_AS	PWS region	
18	chr15	20.2	29.7	9.5	dup	7.71E-09	Yes	No	PW_AS	PWS	
19	chr15	28.65	30.75	2.1	del	1.41E-09	Yes	No	15q13	15q13	
20	chr15	82.9	83.6	0.7	del	7.85E-02	Yes	No	15q25	15q25	*
21	chr15	98.2	100.2	2	del	4.52E-02	No	Yes	15tel	15q26.3	
21	chr15	98.85	99.55	0.7	dup	3.36E-02	No	Yes	15tel	15q26.3	
22	chr16	0.65	2.45	1.8	dup	5.14E-02	No	Yes	16tel	16p13.3	
22	chr16	0.1	5.1	5	del	7.71E-09	No	Yes	16tel	16p13.3	
23	chr16	3.65	5.15	1.5	dup	5.14E-02	No	Yes	16tel	16p13.3	
24	chr16	14.8	16.8	2	dup	2.09E-06	Yes	No	16p13	16p13.1	
24	chr16	14.8	18.2	3.4	del	5.87E-04	Yes	No	16p13-p12	16p13.11	
25	chr16	21.8	22.4	0.6	del	9.36E-07	Yes	No	16p12	16p12.1	
26	chr16	28.35	30.25	1.9	del	1.08E-09	Yes	No	16p11.2	16p11.2	
26	chr16	28.35	30.25	1.9	dup	9.52E-05	Yes	No	16p11.2	16p11.2	
27	chr16	82.35	88.75	6.4	del	3.15E-04	No	Yes	16tel	16p13.3	
27	chr16	86.3	88.7	2.4	dup	2.20E-02	No	Yes	16tel	16q23.3	
28	chr17	0.1	4.1	4	del	9.41E-03	No	Yes	17tel	17p13.3	
28	chr17	0.65	1.55	0.9	dup	2.20E-02	No	Yes	17tel	17p13.3	
29	chr17	16.6	22.5	5.9	dup	1.44E-02	Yes	No	17p11.2	17p11.2	
29	chr17	16.65	20.25	3.6	del	7.37E-04	Yes	No	17p11.2	17p11.2	
30	chr17	26.05	26.85	0.8	del	7.85E-02	Yes	No	17q11.2	17q11.2	
31	chr17	31.8	33.4	1.6	del	3.54E-02	Yes	No	17q12	17q12	
32	chr17	31.85	33.35	1.5	dup	3.61E-02	Yes	No	17q12	17q12	
33	chr17	41	41.5	0.5	del	8.83E-05	Yes	No	17q21	17q21.31	
34	chr17	69.6	78.7	9.1	del	5.04E-09	No	Yes	17tel	17q25	
35	chr18	0.1	5.3	5.2	del	5.07E-03	No	Yes	18tel	18p11.32	
36	chr18	6.75	7.35	0.6	dup	5.14E-02	No	Yes	18tel	18p11.31	
37	chr18	69.1	76	6.9	del	6.16E-03	No	Yes	18tel	18q22.1q23	
38	chr19	0.15	5.85	5.7	dup	2.95E-06	No	Yes	19tel	19p13.3	
38	chr19	0.15	8.65	8.5	del	2.63E-03	No	Yes	19tel	19p13.3	
39	chr19	59.55	63.75	4.2	dup	1.44E-02	No	Yes	19tel	19p13.42	
40	chr2	0.1	1.7	1.6	del	5.14E-02	No	Yes	2tel	2p25.3	
41	chr2	3.25	3.45	0.2	dup	5.14E-02	No	No	VUS	2p25.3	*
42	chr2	45.2	45.9	0.7	dup	2.20E-02	No	No	VUS	2p21	*
43	chr2	111.05	112.95	1.9	del	6.16E-03	Yes	No	2q13	2q13	*
43	chr2	111.05	112.85	1.8	dup	3.36E-02	Yes	No	2q13	2q13	*
44	chr2	165.4	166.1	0.7	del	5.14E-02	No	No	VUS	2q24.3	*
45	chr2	235.4	242.8	7.4	del	1.62E-05	No	Yes	2tel	2q37	

Nature Genetics: doi.10.1038/ng.909

46	chr20	0.1	1.1	1	del	5.14E-02	No	Yes	20tel	20p13	
47	chr20	59.7	62.3	2.6	del	2.20E-02	No	Yes	20tel	20q13.33	
47	chr21	19.95	20.25	0.3	del	5.14E-02	No	No	VUS	21q21.1	*
48	chr21	41.4	46.8	5.4	del	4.03E-03	No	Yes	21tel	21q22.3	
49	chr22	15.85	23.35	7.5	dup	7.09E-07	Yes	No	22q11	DGS region	
50	chr22	17.2	20.6	3.4	del	5.25E-19	Yes	No	22q11	DGS	
51	chr22	21.7	22.1	0.4	del	1.07E-02	Yes	No	22q11	DGS region	
52	chr22	23.7	24.5	0.8	del	7.18E-11	Yes	No	22q11	Distal 22q	
53	chr22	42.6	49.6	7	del	7.71E-09	No	Yes	22tel	22q13	
53	chr22	45.4	49.6	4.2	dup	9.41E-03	No	Yes	22tel	22q13	
54	chr3	0.1	0.9	0.8	del	6.16E-03	No	Yes	3tel	3p26.1	
55	chr3	0.8	1.4	0.6	dup	4.79E-02	No	Yes	3tel	3p26.1	
56	chr3	2.1	9.8	7.7	del	3.36E-02	No	Yes	3tel	3p26.1	
57	chr3	197.2	198.9	1.7	del	5.14E-02	Yes	No	3q29	3q29	
58	chr4	0.1	7.1	7	del	2.06E-04	No	Yes	4tel	4p16.3	
58	chr4	0.35	3.95	3.6	dup	5.14E-02	No	Yes	4tel	4p16.3	
59	chr4	9.45	10.45	1	dup	3.36E-02	No	No	VUS	4p16.1	*
60	chr4	81.95	83.35	1.4	del	3.36E-02	No	No	VUS	4q21.21q21.22	*
61	chr4	184.25	184.75	0.5	del	7.85E-02	No	Yes	4tel	4q35.1	
62	chr4	187.5	188.2	0.7	dup	2.20E-02	No	Yes	4tel	4q35.2	
63	chr5	0.1	1.5	1.4	del	3.36E-02	No	Yes	5tel	5p15.33	
64	chr5	3.75	4.45	0.7	del	7.85E-02	No	Yes	5tel	5p15.33	
65	chr5	175.45	177.45	2	del	2.20E-02	No	Yes	5tel	5q35.2	
66	chr5	180.05	180.75	0.7	del	7.37E-04	No	Yes	5tel	5q35.3	
67	chr6	0.1	5.8	5.7	del	1.44E-02	No	Yes	6tel	6p25.3	
68	chr6	20.85	21.25	0.4	del	7.85E-02	No	No	VUS	6p22.3	*
69	chr6	165.25	170.75	5.5	del	1.72E-03	No	Yes	6tel	6q27	
70	chr7	0.15	2.65	2.5	del	2.20E-02	No	Yes	7tel	7p22.1	
70	chr7	0.1	3.8	3.7	dup	6.16E-03	No	Yes	7tel	7p22.1	
71	chr7	5.7	6.2	0.5	dup	7.85E-02	No	Yes	7tel	7p22.1	
72	chr7	72.3	73.9	1.6	del	7.71E-09	Yes	No	WBS	WBS	
72	chr7	72.3	73.9	1.6	dup	4.82E-04	Yes	No	WBS	WBS dup	
73	chr8	0.15	11.95	11.8	del	1.13E-03	No	Yes	8tel	8p23.3	
73	chr8	0.35	1.05	0.7	dup	3.36E-02	No	Yes	8tel	8p23.2	
74	chr8	2.35	4.35	2	dup	2.20E-02	No	No	VUS	8p23.2	
75	chr8	8.25	11.95	3.7	dup	9.41E-03	Yes	No	8p23.1	8p23.1 deletion	
76	chr8	53.45	54.05	0.6	dup	5.14E-02	No	No	VUS	8q11.23	*
77	chr8	143.25	145.95	2.7	del	7.37E-04	No	Yes	8tel	8q24.3	
78	chr9	0.15	6.75	6.6	del	3.36E-02	No	Yes	9tel	9p24.3	
79	chr9	0.15	1.65	1.5	dup	3.36E-02	No	Yes	9tel	9p24.3	

Nature Genetics: doi.10.1038/ng.909

80	chr9	130.1	140.2	10.1	del	2.85E-13	No	Yes	9tel	9q34

*Novel loci

[a] The 14 novel loci presented in the main text (Table 2) were identified as a manually-curated (both at the locus and the individual CNV level) subset of the 80 loci listed above after removing loci of known pathogenicity, telomere-associated CNVs, and sites that appeared on manual inspection to be potentially spurious such as those with high assembly gap or segmental duplication content (e.g. 10p11.22). Note also that one novel locus (10p15 duplications) passed the above filters but was subsequently determined to be confounded with population-stratification between cases and controls (see Supplementary Note, below) and thus eliminated.

Supplementary Table 8.

Phenotypic and inheritance status for candidate loci.

Region	Source	Case	Start (bp)	End (bp)	Size (Mbp)	State	Phenotype	Inheritance	Ethnicity	Sibling
10p15.3	Signature	9895253	115,543	4,813,203	4.70	Gain	Cleft palate, der(7)t(7;10)(q34;p15)pat	Paternal (Unbalanced Translocation)	Caucasian	
10p15.3	Signature	9887090	128,680	6,841,656	6.71	Gain	Developmental Delay, Multiple Congenital Anomalies	Unknown	Other	
10p15.3	Signature	9883636	1,978,248	3,081,215	1.10	Gain	Developmental Delay, Mental Retardation, Speech and Language Delay	Unknown	Caucasian	
10p15.3	Signature	9901031	2,639,936	3,167,993	0.53	Gain	Developmental Delay	Unknown	Unknown	
10p15.3	Signature	9901027	2,639,936	3,167,993	0.53	Gain	Developmental Delay	Unknown	Unknown	
10p15.3	Signature	9903408	2,685,391	3,173,015	0.49	Gain	Developmental Delay, Learning Disibility	Unknown	Unknown	
10p15.3	Signature	9890718	2,716,992	3,196,525	0.48	Gain	Multiple fetal anomalies, Abnormal ultrasound, Normal karyotype	Unknown	Caucasian	
10p15.3	Signature	9897878	2,518,261	3,140,309	0.62	Gain	Developmental Delay	Unknown	Caucasian	
10p15.3	Signature	9885237	2,663,957	3,129,478	0.47	Gain	Developmental Delay, Hypogonadism	Unknown	Other	
10p15.3	Signature	9909907	2,685,191	3,114,260	0.43	Gain	Failure to Thrive	Unknown	Unknown	
10p15.3	Signature	9908718	2,685,191	3,114,260	0.43	Gain	Autism	Unknown	Caucasian	
10p15.3	Signature	9908524	2,685,191	3,114,260	0.43	Gain		Unknown	African-American	
10p15.3	Signature	9907464	2,685,191	3,114,260	0.43	Gain	Background rentinopathy, Congenital mitral insufficiency, Microcephalus, delayed milestones, Nystagmus	Unknown	African-American	
10p15.3	Signature	9906584	2,685,191	3,114,260	0.43	Gain	Delayed milestones, Failure to Thrive, Hypotonia, undescended testis	Unknown	Other	
10p15.3	Signature	9898606	2,685,191	3,104,448	0.42	Gain	Organic Encephalopathy	Unknown	African-	9898618

Nature Genetics: doi.10.1038/ng.909

									American	
10p15.3	Signature	9900428	2,685,391	3,125,543	0.44	Gain	Developmental Delay	Unknown	African-American	
10p15.3	Signature	9909307	2,685,391	3,114,121	0.43	Gain	Obstructive sleep apnea	Unknown	African-American	
10p15.3	Signature	9904855	2,685,391	3,114,121	0.43	Gain	multiple cardiac defects	Unknown	African-American	
10p15.3	Signature	9902650	2,685,391	3,114,121	0.43	Gain	FAS Poor Behavior, Learning Problems	Unknown	African-American	
10p15.3	Signature	9899894	2,685,391	3,114,121	0.43	Gain	Right midshaft femur fraction failure	Unknown	African-American	
10p15.3	Signature	9898648	2,685,391	3,114,121	0.43	Gain	Microcephaly	Unknown	African-American	
10p15.3	Signature	9898023	2,685,391	3,114,121	0.43	Gain	Dysmorphic Features, FTT, MR, Renal Agensis	Unknown	Unknown	
10p15.3	Signature	9895666	2,685,391	3,114,121	0.43	Gain	Hypoplastic left heart syndrome	Unknown	African-American	
10p15.3	Signature	9894429	2,685,391	3,114,121	0.43	Gain	Developmental Delay, Dysmorphic Features, Autistic Disorder	Unknown	African-American	
10p15.3	Signature	9894153	2,685,391	3,114,121	0.43	Gain	ADHD, Asthma, unspecified, Behavior and anger issues	Unknown	African-American	
10p15.3	Signature	9893919	2,685,391	3,114,121	0.43	Gain	Developmental Delay, Dysmorphic Features, Craniosynostosis	Unknown	African-American	
10p15.3	Signature	9890947	2,685,391	3,114,121	0.43	Gain	Congenital anomaly, Delayed Milestones	Unknown	African-American	
10p15.3	Signature	9890733	2,685,391	3,114,121	0.43	Gain	Lack of normal development	Unknown	Unknown	
10p15.3	Signature	9898618	2,685,391	3,104,448	0.42	Gain	ADHD, Organic Encpephalopathy	Unknown	African-American	9898606
10p15.3	Signature	9892737	2,686,834	3,116,378	0.43	Gain	Myasthenia gravis with (acute) exacerbation	Unknown	African-American	
10p15.3	Signature	9888112	2,686,834	3,116,378	0.43	Gain	Spasticity	Unknown	Unknown	
10p15.3	Signature	9885951	2,686,834	3,116,378	0.43	Gain	Congenital heart defect	Unknown	African-American	
10p15.3	Signature	9885908	2,686,834	3,116,378	0.43	Gain	Other specific developmental learning difficulties	Unknown	Caucasian	
10p15.3	Signature	9884622	2,686,834	3,116,378	0.43	Gain	Dysmorphic Features	Unknown	African-American	
10p15.3	Signature	9881122	2,686,834	3,116,378	0.43	Gain	Developmental Delay, Multiple Congenital Anomalies	Unknown	African-American	
2q13	Signature	9882799	108,874,790	114,926,862	6.05	Gain	Developmental Delay	Unknown	Caucasian	
2q13	Signature	9894411	111,114,738	112,817,963	1.70	Loss	Developmental Delay	Paternal	Caucasian	
2q13	Signature	9893921	111,114,738	112,817,963	1.70	Loss	Developmental Delay, Dysmorphic Features	Unknown	Caucasian	
2q13	Signature	9892734	111,114,738	112,817,963	1.70	Loss	Anomalies of skull and face bones and ears	Maternal	Caucasian	
2q13	Signature	9887975	111,114,738	112,817,963	1.70	Loss	Developmental Delay, Hypotonia	Unknown	Caucasian	

Nature Genetics: doi.10.1038/ng.909

2q13	Signature	9887950	111,114,738	112,817,963	1.70	Loss	Congenital anomalies of diaphragm, Double outlet right ventricle, Fetal growth retardation,history of sudden cardiac arrest	*de novo*	Unknown
2q13	Signature	9887419	111,114,738	112,817,963	1.70	Gain	autism, dysmorphic features	Unknown	Other
2q13	Signature	9886256	111,114,738	112,817,963	1.70	Loss	Persistent open anterior fontanelle	Maternal	Caucasian
2q13	Signature	9886105	111,114,738	112,817,963	1.70	Gain	Developmental Delay	Unknown	Unknown
2q13	Signature	9885572	111,114,738	112,817,963	1.70	Loss	Dysmorphic Features, Mental retardation	Unknown	Caucasian
2q13	Signature	9883850	111,114,738	112,817,963	1.70	Loss	Developmental Delay, Seizure Disorder, Lack of normal physiological development, unspecified	Unknown	Caucasian
2q13	Signature	9883704	111,114,738	112,817,963	1.70	Loss	Hypotonia, Not tolerating breast milk or milk formula, vomiting	Unknown	Caucasian
2q13	Signature	9908137	111,131,409	112,771,623	1.64	Loss	Frontal Lobe Syndrome	Maternal	Caucasian
2q13	Signature	9907637	111,131,409	112,771,623	1.64	Gain	Developmental Delay, Dysmorphic Features	Paternal	Caucasian
2q13	Signature	9901189	111,131,609	112,771,484	1.64	Gain	Developmental Delay, Dysmorphic Features	Maternal	Caucasian
2q13	Signature	9899071	111,131,609	112,771,484	1.64	Loss	Developmental Delay, Heart Defect	Unknown	African-American
2q13	Signature	9893017	111,131,609	112,771,484	1.64	Loss	Pervasive developmental disorders	*de novo*	Caucasian
2q13	Signature	9890489	111,131,609	112,771,484	1.64	Gain	Developmental Delay	Unknown	Caucasian
2q13	Signature	9894264	110,183,373	111,085,361	0.90	Gain	NICHD Prenatal Study	Unknown	Unknown
2q13	Signature	9900304	110,199,003	111,141,741	0.94	Gain	Developmental Delay, Dysmorphic Features-severe	Unknown	Other
2q13	Signature	9906602	110,199,003	111,131,670	0.93	Gain	Congenital anomaly	Unknown	Caucasian
2q13	Decipher	254017	111,085,360	112,782,191	1.70	Gain	Cleft palate, Microcephaly, ORAL REGION, Thin	Unknown	Unknown
2q13	Decipher	250127	111,142,809	112,762,656	1.62	Gain	Hydrocephalus/large ventricles, non-specific, Hypoglycaemia, Mental retardation/developmental delay, Microcephaly	Inherited	Unknown

Nature Genetics: doi.10.1038/ng.909

2q13	Decipher	251377	111,142,808	112,762,656	1.62	Loss	Advanced bone age/large epiphyses, Broad hands, Depressed/flat nasal bridge, Dolichocephaly/scaphocephaly, Macrocephaly, Mental retardation/developmental delay, Microstomia, Palpebral fissures slant up, Philtrum, general abnormalities, Prominent forehead/frontal bossing, Pyloric stenosis, Scoliosis, Small mandible/micrognathia, Small/short nose, Spasticity/brisk reflexes/Babinski, Wide feet	Unknown	Unknown	
10q23.1	Signature	9893676	81,674,576	88,837,361	7.16	Loss	Autistic Disorder, Microcephaly	*de novo*	Other	
10q23.1	Signature	9887883	81,674,576	88,837,361	7.16	Loss	Developmental Delay, Dysmorphic Features	Unknown	Caucasian	
10q23.1	Signature	9883729	81,674,576	88,837,361	7.16	Loss	Dysmorphic Features, Multiple Congenital Anomalies, Family history of other condition Clinical diagnosis of weaver syndrome	Unknown	Caucasian	
10q23.1	Signature	9882409	81,674,576	88,837,361	7.16	Loss	Developmental Delay, Known 10q deletion	Unknown	Caucasian	
10q23.1	Signature	9909469	81,682,644	88,931,994	7.25	Loss	Developmental Delay, Dysmorphic Features, Joint Hypermobility, Mild Hypotonia, del(10)(q23.1q23.2)dn	*de novo*	Caucasian	
10q23.1	Signature	9905075	81,682,844	88,931,855	7.25	Loss	Dysmorphic Features	Unknown	Caucasian	
10q23.1	Signature	9903971	81,682,844	88,931,855	7.25	Loss	Developmental Delay, Macrocephaly, Failure to Thrive	*de novo*	Caucasian	
10q23.1	Signature	9885281	80,739,721	83,788,125	3.05	Loss	Developmental delay, History of congenital heart disease, Speech delay	Unknown	Caucasian	9884711
10q23.1	Signature	9884711	80,739,721	83,788,125	3.05	Loss	Sibling with 2q11.2 Deletion, 10q22.3q23.1 Deletion	Unknown	Caucasian	9885281
10q23.1	Signature	9909892	84,765,247	85,211,522	0.45	Loss	Developmental Delay	Paternal	Other	
10q23.1	Signature	9893611	83,955,196	84,423,692	0.47	Loss	Developmental Delay, Anomalies of skull and face bones	Maternal	Unknown	9884218, 9884217
10q23.1	Signature	9884218	83,955,196	84,423,692	0.47	Loss	Half sister and mother with 10q23.1 duplication, 18p duplications	Maternal	Unknown	9893611, 9884217
10q23.1	Signature	9884217	83,955,196	84,423,692	0.47	Loss	Mother and half sister with 10q23.1 deletion	Maternal	Unknown	9893611, 9884218

Nature Genetics: doi.10.1038/ng.909

10q23.1	Decipher	250977	81,482,307	89,136,979	7.65	Loss	Club foot, varus, Dyspraxia/apraxia including gait apraxia, Hyperactivity, Hypertelorism, Mental retardation/developmental delay, Metatarsus varus, Speech defect/dysarthria, Speech delay	Unknown	Unknown	
2p21	Signature	9890394	45,248,672	45,810,225	0.56	Gain	Absent Corpus Callosum	Unknown	Caucasian	
2p21	Signature	9902819	45,272,711	45,815,087	0.54	Gain	Exotropia	Unknown	Caucasian	
2p21	Signature	9901154	45,272,711	45,815,087	0.54	Gain	AO Dilation	Unknown	Unknown	
2p21	Signature	9897440	45,272,711	45,815,087	0.54	Gain	Developmental Delay, Failure to Thrive, Short Stature, Supravalcular Pulmonary Artery Stenosis	Unknown	Caucasian	
2p21	Signature	9896183	45,272,711	45,815,087	0.54	Gain	Encephalopathy	Unknown	Caucasian	
2p21	Signature	9889969	45,272,711	45,815,087	0.54	Gain	Attention deficit disorder of childhood with hyperactivity	Unknown	Caucasian	
2p21	Signature	9886416	45,277,629	45,810,225	0.53	Gain	Developmental Delay, Hydrocephalus	Unknown	Caucasian	
2p21	Signature	9885152	45,277,629	45,810,225	0.53	Gain	Developmental Delay, Agenesis of corpus callosum	Unknown	Caucasian	
2p21	Signature	9884923	45,277,629	45,810,225	0.53	Gain	Developmental Delay, Dysmorphic Features, Microcephaly	Unknown	Caucasian	
2p21	Decipher	251298	45,265,479	45,828,360	0.56	Gain	Coarctation/interrupted aorta, Frontal upsweep/cowlick, Mental retardation/developmental delay, Short stature, general abnormalities	Unknown	Unknown	
2p21	Decipher	252059	45,349,137	45,815,194	0.47	Gain	Coloboma involving optic nerve, Dystopia canthorum (telecanthus), Hypospadias, Long toes, Mental retardation/developmental delay, Microcephaly, Myopia, Strabismus, Tapering fingers	*de novo*	Unknown	
4p16.1	Signature	9893927	9,237,041	10,356,984	1.12	Gain	Ambiguous Genitalia	Unknown	Caucasian	
4p16.1	Signature	9898387	9,314,766	10,484,553	1.17	Gain	Loss of weight, Gastroesophageal reflux	Unknown	Caucasian	
4p16.1	Signature	9892775	9,414,321	10,320,088	0.91	Gain	Developmental Delay	Unknown	Caucasian	
4p16.1	Signature	9885183	9,414,321	10,219,850	0.81	Gain	Developmental Delay, Dysmorphic Features	Unknown	African-American	
4p16.1	Signature	9891544	9,423,818	10,283,626	0.86	Gain	Developmental Delay	Unknown	Caucasian	
4p16.1	Signature	9890606	9,505,269	10,548,999	1.04	Gain	Autistic Disorder	Unknown	Caucasian	9890605
4p16.1	Signature	9890605	9,505,269	10,548,999	1.04	Gain	Autistic Disorder	Unknown	Caucasian	9890606
4p16.1	Signature	9882209	9,512,664	10,320,088	0.81	Gain	Developmental Delay, Dysmorphic Features, Failure To Thrive	Unknown	Other	

Nature Genetics: doi.10.1038/ng.909

4p16.1	Decipher	250927	9,438,125	10,582,606	1.14	Gain	Mental retardation/developmental delay, Short attention span, Speech delay	Inherited	Unknown
4q21.21 q21.22	Signature	9898554	80,082,313	86,941,434	6.86	Loss	Developmental Delay	*de novo*	Caucasian
4q21.21 q21.22	Signature	9906434	80,882,789	87,801,982	6.92	Loss	Developmental Delay	*de novo*	Unknown
4q21.21 q21.22	Signature	9894220	81,869,369	83,889,140	2.02	Loss	Developmental Delay, Dysmorphic Features	*de novo*	Caucasua n
4q21.21 q21.22	Signature	9882016	82,021,578	86,946,838	4.93	Loss	Developmental Delay, Dysmorphic Features, Other conditions due to autosomal anomalies	Unknown	Caucasian
4q21.21 q21.22	Signature	9890651	82,632,273	84,004,580	1.37	Loss	Other encephalopathy	Unknown	Caucasian
4q21.21 q21.22	Signature	9892313	81,033,420	82,731,438	1.70	Loss	Multiple Brain Malformations	Unknown	African-American
4q21.21 q21.22	Signature	9890179	81,267,974	83,227,291	1.96	Loss	Developmental Delay, Dysmorphic Features, Multiple Congenital Anomalies	Unknown	Caucasian
4q21.21 q21.22	Signature	9882411	82,021,578	82,846,844	0.83	Loss	Congenital Disorder, Polymicrogyria, Hemiplegia, Unspecified congenital anomaly of brain spinal cord and nervous system	Unknown	Caucasian
4q21.21 q21.22	Decipher	248316	73,321,374	82,333,001	9.01	Loss	Feeding problems in infants, SEIZURES, general abnormalities, Short stature, general abnormalities	Unknown	Unknown
4q21.21 q21.22	Decipher	251258	78,304,421	84,254,989	5.95	Loss	Autism/autistic behaviour, Diplegia, Feeding problems in infants, Mental retardation/developmental delay Palpebral fissures slant down, Prominent forehead/frontal bossing, Self-mutilation, Short philtrum, Short stature, proportionate, Simple/absent philtrum, Thick fingers, Tonic/clonic (grand-mal), Wide palpebral fissures	Unknown	Unknown
4q21.21 q21.22	Decipher	749	80,383,883	82,332,981	1.95	Loss	Autism/autistic behaviour, Cryptorchid testes, Hypospadias, Mental retardation/developmental delay, Scoliosis, Short stature, general abnormalities, Strabismus, Tonic/clonic (grand-mal)	*de novo*	Unknown

Nature Genetics: doi.10.1038/ng.909

4q21.21 q21.22	Decipher	4665	82,229,075	86,710,919	4.48	Loss	Abnormal/absent metatarsals, Clinodactyly, Hypoplastic toes (including phalanges), Mental retardation/developmental delay, Short stature, general abnormalities, Small hands, Speech delay	*de novo*	Unknown
4q21.21 q21.22	Decipher	4539	82,301,439	85,425,757	3.12	Loss	Atrial septum defect, Cleft uvula, Depressed/flat nasal bridge, Down-turned corners of the mouth, Dystopia canthorum (telecanthus), Epicanthic folds, High birth weight (> 90th centile), Hypotonia (non-myopathic), Mental retardation/developmental delay, Open mouth appearance, Palpebral fissures slant up, Short philtrum, Small feet, Small hands, Speech delay, Submucous cleft palate, Urinary reflux	*de novo*	Unknown
2p25.3	Signature	9882501	3,478,943	3,912,501	0.43	Gain	ADHD, ODD, Asperger's, Short stature	*Unknown*	Unknown
2p25.3	Signature	9896126	56,096	4,882,972	4.83	Gain	Developmental Delay, Dysmorphic Features, 46,XX,der (13)t(2;13)(p25.1;q32)	Unknown	Other
2p25.3	Signature	9896073	658,705	7,577,621	6.92	Gain	Developmental Delay	Unknown	Caucasian
2p25.3	Signature	9906154	1,556,612	4,614,213	3.06	Gain	Abnormal Ultrasound	Unknown	Unknown
2p25.3	Signature	9882074	1,802,677	3,464,495	1.66	Gain	Sibling and Mother with 2p25.3 Duplication, Multiple Congenital Anomalies, ADHD	Maternal	Caucasian
2p25.3	Signature	9890175	3,342,484	4,882,972	1.54	Gain	Autistic Disorder	Unknown	Caucasian
2p25.3	Signature	9903386	3,304,456	3,789,869	0.49	Gain	Family history chromosome abnormality	Unknown	Unknown
2q24.3	Signature	9887654	164,946,224	168,513,416	3.57	Loss	Hypotonia, Developmental Delay	Unknown	Caucasian
2q24.3	Signature	9902525	165,485,317	171,030,878	5.55	Loss	Developmental Delay, Dysmorphic Features	Unknown	Caucasian
2q24.3	Signature	9908201	161,755,544	165,849,815	4.09	Loss	Developmental Delay	*de novo*	Other
2q24.3	Signature	9904266	161,919,306	166,011,752	4.09	Loss	Developmental Delay, Hypotonia, 46,XY,t(7;10)(q22;q26)	*de novo*	Caucasian
2q24.3	Signature	9904626	164,188,377	165,611,978	1.42	Loss	Developmental Delay, Myopathy	*de novo*	Unknown
2q24.3	Signature	9906141	165,118,051	167,362,094	2.24	Loss	Seizure Disorder	*de novo*	Caucasian
2q24.3	Signature	9882986	165,595,793	166,186,269	0.59	Loss	Developmental Delay	Unknown	Caucasian

Nature Genetics: doi.10.1038/ng.909

2q24.3	Decipher	254495	162,536,101	168,251,297	5.72	Loss	Cleft palate, Mental retardation/developmental delay, Tonic seizures	*de novo*	Unknown	
2q24.3	Decipher	249144	162,861,293	165,685,811	2.82	Loss	BEHAVIOURAL PROBLEMS, general abnormalities, Mandible, general abnormalities, Mental retardation/developmental delay, Ptosis of eyelids, Strabismus	*de novo*	Unknown	
2q24.3	Decipher	2859	165,061,403	168,143,091	3.08	Loss	1-[12] months/onset, Hypotonia (non-myopathic), Mental retardation/developmental delay, Microcephaly, Scoliosis, Strabismus, Tonic seizures	*de novo*	Unknown	
2q24.3	Decipher	249811	165,238,836	174,174,805	8.94	Loss	Face, general abnormalities, Low birthweight (< 3rd centile), SEIZURES, general abnormalities, Syndactyly 2-3 of toes	Unknown	Unknown	
21q21.1	Signature	9909341	13,600,026	20,175,986	6.58	Loss	Microcephaly, Hypoplastic nails, Micrognathia, Congenital contraction of fingers	Maternal	Caucasian	
21q21.1	Signature	9893857	17,002,975	20,246,507	3.24	Loss	Multiple Congenital Anomalies, Chromosome Abnormality	Unknown	Unknown	
21q21.1	Signature	9895119	17,117,282	20,824,736	3.71	Loss	Multiple Congenital Anomalies	Unknown	Unknown	
21q21.1	Signature	9890191	18,764,366	22,855,093	4.09	Loss	Developmental Delay	Unknown	Unknown	
21q21.1	Signature	9892738	19,441,488	20,059,094	0.62	Loss	Delayed Milestones	Paternal	African-American	
21q21.1	Signature	9895945	20,042,753	20,523,162	0.48	Loss	Developmental Delay, Congenital hip dysplasia	Unknown	Other	9895486
21q21.1	Signature	9895486	20,042,753	20,523,162	0.48	Loss	Hydrocephalus, Velopharyngeal insufficiency, Cleft palate, Family history, Cleft palate, Schizophrenia	Unknown	Other	9895945
21q21.1	Decipher	249224	14,320,039	20,690,911	6.37	Loss	Speech delay	*de novo*	Unknown	
8q11.23	Signature	9887107	53,483,464	54,080,009	0.60	Gain	Developmental Delay, Autistic Disorder	Unknown	Caucasian	
8q11.23	Signature	9892323	53,494,612	54,019,382	0.52	Gain	Suspected Wolf-Hirschhorn, Mental retardation	Unknown	Caucasian	
8q11.23	Signature	9891360	53,494,612	53,928,978	0.43	Gain	Moderate Mental Retardation	Unknown	Caucasian	
8q11.23	Signature	9884625	53,521,373	54,080,009	0.56	Gain	Pervasive development disorder	Unknown	Caucasian	
8q11.23	Signature	9883029	53,521,373	54,018,486	0.50	Gain	Developmental Delay, Dysmorphic Features, Multiple Congenital	Unknown	Caucasian	

Nature Genetics: doi.10.1038/ng.909

							Anomalies		
8q11.23	Signature	9885106	53,521,373	53,987,142	0.47	Gain	Seizures	Unknown	Caucasian
8q11.23	Signature	9890001	53,650,602	54,210,422	0.56	Gain	46,XX,t(4;16)(p16;p13.3) pat, Parental concern	Maternal	Other
8q11.23	Decipher	254185	53559679	54003317	0.44	Gain	Mental retardation/developmental delay	Inherited	Unknown
1q24.3	Signature	9901721	163,760,542	172,532,315	8.77	Loss	Delayed Milestones, Obsruction of nasolacrimal duct, Generlized anxiety disorder	Unknown	Caucasian
1q24.3	Signature	9906143	164,099,272	170,502,726	6.40	Loss	Developmental Delay, Dysmorphic Features	Unknown	Unknown
1q24.3	Signature	9898176	165,344,243	172,662,904	7.32	Loss	Conditions due to anomaly unspecified chromosome	Unknown	Caucasian
1q24.3	Signature	9910356	170,099,984	179,914,925	9.81	Loss	Developmental Delay, Dysmorphic Features, Multiple Congenital Anomalies	Unknown	Caucasian
1q24.3	Signature	9882995	168,643,158	171,658,282	3.02	Loss	Anomalies of diaphragm, Anomaly of heart, Anomalies of skull and face bones	Maternal	Caucasian
1q24.3	Signature	9903173	168,991,860	174,645,488	5.65	Loss	Developmental Delay, Dysmorphic Features, Multiple Congenital Anomalies	Unknown	Caucasian
1q24.3	Decipher	254308	166,314,077	172,150,939	5.84	Loss	Mental retardation/developmental delay	*de novo*	Unknown
1q24.3	Decipher	248515	168,899,278	174,853,889	5.95	Loss	Brachydactyly, Large nose, Short stature, general abnormalities, Speech delay, Tapering fingers	*de novo*	Unknown
12p13.31	Signature	9887300	60,861	9,407,180	9.35	Gain	Developmental Delay, Dysmorphic Features, Hypotonia	Unknown	Caucasian
12p13.31	Signature	9889868	60,861	8,623,949	8.56	Gain	Congenital Heart disease, dysmorphic features, add (22)(q13) ARSA deleted	Unknown	Caucasian
12p13.31	Signature	9895643	6,578,498	8,751,021	2.17	Loss	Developmental Delay	*de novo*	Caucasian
12p13.31	Signature	9899612	7,670,027	8,499,498	0.83	Gain	Developmental Delay	Unknown	Caucasian
12p13.31	Signature	9902132	8,136,631	8,586,879	0.45	Gain	Developmental Delay	Unknown	Caucasian
12p13.31	Signature	9891072	60,861	8,201,023	8.14	Gain	Developmental Delay, Dysmorphic Features, Seizure Disorder, Autosomal anomaly	Unknown	Caucasian
12p13.31	Signature	9905711	85,117	8,165,257	8.08	Gain	Developmental Delay, Dysmorphic Features	Unknown	Caucasian
15q25	Signature	9903883	82,553,367	89,111,260	6.56	Loss	Short Stature	Paternal	African-American

Nature Genetics: doi.10.1038/ng.909

15q25	Signature	9887066	82,944,098	83,484,862	0.54	Loss	Developmental Delay, Seizure Disorder	Unknown	African-American	
15q25	Signature	9885639	82,944,098	83,484,862	0.54	Loss	Developmental Delay	Unknown	Other	9883421
15q25	Signature	9883421	82,944,098	83,484,862	0.54	Loss	Autism, Brother with 15q25.2q25.3 deletion	Unknown	Other	9885639
15q25	Signature	9890732	82,950,693	83,498,403	0.55	Loss	Muscle weakness	Maternal	Unknown	
15q25	Signature	9890571	82,950,693	83,498,403	0.55	Loss	Developmental Delay, Delayed milestones	Unknown	Caucasian	
15q25	Decipher	248182	82,743,637	83,527,718	0.78	Loss	Tall stature, proportionate, Ulnar deviation of fingers	Unknown	Unknown	
6p22.3	Signature	9910583	14,957,866	21,529,824	6.57	Loss	Developmental Delay	Unknown	Unknown	
6p22.3	Signature	9882480	16,190,225	21,116,367	4.93	Loss	Mental retardation, Obesity, Joint laxity	Unknown	Unknown	
6p22.3	Signature	9900792	16,294,370	21,529,684	5.24	Loss	Developmental Delay	*de novo*	Other	
6p22.3	Signature	9901522	20,019,727	21,706,014	1.69	Loss	Developmental Delay, Dysmorphic Features	Unknown	African-American	
6p22.3	Signature	9890404	20,622,416	21,316,606	0.69	Loss	Developmental Delay, Dysmorphic Features	*de novo*	Caucasian	
6p22.3	Signature	9890894	20,941,545	22,499,254	1.56	Loss	Developmental Delay, Dysmorphic Features	Maternal	Unknown	

Supplementary Table 9. Clinical features of 15q25.2 microdeletions.

Individual ID	9885639	9887066	9890571	9890732
Gender	Male	Male	Male	Female
Age at evaluation	2.5y	12y	9y	15y
Stature	NA	NA	Short stature	Tall stature
Failure to thrive	NA	NA	+	-
Head size	NA	NA	25th-50th	97th
Developmental delay/learning disability	+	+	+	Mild motor delay
Speech delay	+	NA	+	-
Other neurologic features	PDD-NOS	Seizures	Arnold Chiari type I malformation	Hypotonia; congenital strabismus; headaches; normal brain MRI
Craniofacial features	NA	NA	Non dysmorphic	Non dysmorphic
Musculoskeletal features	NA	NA	Scoliosis	Congenital myopathy; slender fingers
Gastrointestinal features	NA	NA	Recurrent vomiting	Severe constipation
Congenital heart disease	NA	NA	+	-
Inheritance	Unknown	Unknown	Not maternal	Maternal
Family history	Brother with autism also has deletion	NA	LD in mother and maternal and paternal half sibs	Mother has migraines; father has dyslexia

LD = learning disability; NA = information not available

Nature Genetics: doi.10.1038/ng.909

Supplementary Table 10.

Frequency of SD-mediated HotSpot CNVs.

Chr	hotspotStart	hotspotEnd	Size	SD orientation	Classification	Deletions (Cases)	Gains (Cases)	Deletions (controls)	Gains (controls)
chr1	39762734	40000001	237,267	Inverted SD	Inactive	0	0	0	0
chr1	47188614	47365812	177,198	Inverted SD	Inactive	0	0	0	0
chr1	142513048	147888332	5,375,284	Direct SD	Active	4	6	1	0
chr1	142875105	144745644	1,870,539	Inverted SD	Inactive	0	0	0	0
chr1	143569171	144927245	1,358,074	Direct SD	Active	28	13	6	2
chr1	144544474	145861130	1,316,656	Inverted SD	Active	5	5	0	2
chr2	87471381	89905818	2,434,437	Direct SD	Active	2	0	3	2
chr2	95521267	97438085	1,916,818	Direct SD	Active	2	2	0	0
chr2	100087627	109875314	9,787,687	Inverted SD	Inactive	0	0	0	0
chr2	108505695	110692620	2,186,925	Direct SD	Active	1	1	0	0
chr2	108513314	112821215	4,307,901	Inverted SD	Inactive	0	0	0	0
chr2	110179629	110498651	319,022	Inverted SD	Active	119	78	32	31
chr2	110768306	112829142	2,060,836	Direct SD	Active	5	12	0	1
chr2	130595668	130924110	328,442	Direct SD	Active	0	0	5	4
chr2	130612254	131674647	1,062,393	Inverted SD	Inactive	0	0	0	0
chr2	130984375	131691306	706,931	Direct SD	Active	0	1	0	1
chr2	232912452	233014237	101,785	Direct SD	Active	7	2	5	3
chr3	10091035	11886214	1,795,179	Direct SD	Inactive	0	0	0	0
chr3	125183427	127172186	1,988,759	Inverted SD	Inactive	0	0	0	0
chr3	126972071	131200599	4,228,528	Inverted SD	Inactive	0	0	0	0
chr3	130595100	130688253	93,153	Direct SD	Active	0	0	1	0
chr3	196880220	198843765	1,963,545	Direct SD	Active	3	4	0	0
chr4	3980496	9341265	5,360,769	Inverted SD	Inactive	0	0	0	0
chr4	69731461	69904884	173,423	Inverted SD	Inactive	0	0	0	0
chr4	75542351	75671304	128,953	Direct SD	Inactive	0	0	0	0
chr4	119778181	120543675	765,494	Inverted SD	Inactive	0	0	0	0
chr5	312327	1622973	1,310,646	Inverted SD	Active	0	1	1	0
chr5	21558880	29473365	7,914,485	Inverted SD	Inactive	0	0	0	0
chr5	98798445	99694503	896,058	Inverted SD	Inactive	0	0	0	0
chr5	98868173	99433874	565,701	Inverted SD	Inactive	0	0	0	0
chr5	175454438	177400971	1,946,533	Direct SD	Active	0	7	0	0
chr6	167530391	167691351	160,960	Inverted SD	Active	0	0	2	0
chr7	29691962	30290871	598,909	Inverted SD	Inactive	0	0	0	0
chr7	29707202	32734087	3,026,885	Direct SD	Inactive	0	0	0	0
chr7	32687455	35178799	2,491,344	Direct SD	Inactive	0	0	0	0
chr7	35973959	45733009	9,759,050	Inverted SD	Inactive	0	0	0	0

Nature Genetics: doi.10.1038/ng.909

chr7	45818867	51421882	5,603,015	Inverted SD	Inactive	0	0	0	0
chr7	45825273	55771836	9,946,563	Inverted SD	Inactive	0	0	0	0
chr7	55798223	56407429	609,206	Direct SD	Active	1	0	0	0
chr7	64204271	64854836	650,565	Direct SD	Active	1	0	0	4
chr7	64272835	64623314	350,479	Inverted SD	Active	5	63	11	45
chr7	66164905	71867176	5,702,271	Inverted SD	Inactive	0	0	0	0
chr7	72135927	74963306	2,827,379	Direct SD	Active	15	41	0	0
chr7	72164317	76478774	4,314,457	Direct SD	Active	0	1	0	0
chr7	73973428	76478774	2,505,346	Direct SD	Active	0	2	0	0
chr7	149258505	153385317	4,126,812	Inverted SD	Active	1	1	0	0
chr8	2182892	2316618	133,726	Inverted SD	Inactive	0	0	0	0
chr8	6986986	12511916	5,524,930	Direct SD	Active	7	6	0	0
chr9	33394043	42846527	9,452,484	Inverted SD	Inactive	0	0	0	0
chr9	33599994	38525756	4,925,762	Inverted SD	Inactive	0	0	0	0
chr9	85682016	87551412	1,869,396	Inverted SD	Active	0	1	0	0
chr9	89747431	89915010	167,579	Inverted SD	Inactive	0	0	0	0
chr9	96162603	98699217	2,536,614	Inverted SD	Inactive	0	0	0	0
chr10	27628183	28282277	654,094	Inverted SD	Active	0	0	1	0
chr10	45517745	51518749	6,001,004	Direct SD	Active	2	5	0	1
chr10	47859358	48519351	659,993	Inverted SD	Inactive	0	0	0	0
chr10	81144969	89104220	7,959,251	Direct SD	Active	1	7	0	0
chr11	48623166	51233511	2,610,345	Inverted SD	Inactive	0	0	0	0
chr11	67351106	71153513	3,802,407	Inverted SD	Active	0	1	0	0
chr12	9366975	9452534	85,559	Inverted SD	Inactive	0	0	0	0
chr13	18912385	24407424	5,495,039	Direct SD	Inactive	0	0	0	0
chr13	23376878	23779362	402,484	Inverted SD	Inactive	0	0	0	0
chr13	24076608	24417171	340,563	Inverted SD	Inactive	0	0	0	0
chr15	18963711	30468167	11,504,456	Direct SD	Active	26	16	0	0
chr15	20300717	21117762	817,045	Inverted SD	Active	63	94	36	19
chr15	26774623	28479130	1,704,507	Direct SD	Active	1	2	1	0
chr15	26774623	30601869	3,827,246	Direct SD	Active	1	0	0	0
chr15	28741566	30486859	1,745,293	Inverted SD	Active	17	29	2	0
chr15	70750592	72145145	1,394,553	Inverted SD	Inactive	0	0	0	0
chr15	70750592	73354113	2,603,521	Direct SD	Active	1	4	0	0
chr15	72170064	73354113	1,184,049	Inverted SD	Inactive	0	0	0	0
chr15	73772788	75973923	2,201,135	Inverted SD	Inactive	0	0	0	0
chr15	80431047	83582321	3,151,274	Direct SD	Active	0	1	0	0
chr15	82707944	83527197	819,253	Direct SD	Active	1	5	0	0
chr16	11945037	21281806	9,336,769	Direct SD	Inactive	0	0	0	0
chr16	14730726	18640357	3,909,631	Direct SD	Active	18	8	2	0
chr16	14771038	16335655	1,564,617	Direct SD	Active	24	18	10	3

Nature Genetics: doi.10.1038/ng.909

chr16	21341194	30193879	8,852,685	Direct SD	Active	1	2	0	0
chr16	21419562	21746862	327,300	Direct SD	Active	31	25	8	18
chr16	21716330	22464053	747,723	Direct SD	Active	3	37	1	3
chr16	28282332	29404449	1,122,117	Direct SD	Active	1	0	0	0
chr16	28282332	30144419	1,862,087	Direct SD	Active	1	2	0	0
chr16	28390714	28956061	565,347	Inverted SD	Active	2	3	0	0
chr16	28971100	30144419	1,173,319	Direct SD	Active	28	64	2	3
chr16	68564256	73116829	4,552,573	Inverted SD	Inactive	0	0	0	0
chr17	2932052	3071880	139,828	Inverted SD	Active	0	1	0	0
chr17	14038639	15411628	1,372,989	Direct SD	Active	9	3	1	0
chr17	15411625	18667826	3,256,201	Direct SD	Active	1	1	0	0
chr17	15617519	20727080	5,109,561	Direct SD	Active	8	15	0	0
chr17	18692134	20168437	1,476,303	Direct SD	Inactive	0	0	0	0
chr17	31529840	33409941	1,880,101	Direct SD	Active	14	11	3	2
chr17	33724392	42907078	9,182,686	Inverted SD	Inactive	0	0	0	0
chr17	40987062	42445986	1,458,924	Inverted SD	Inactive	0	0	0	0
chr17	41060948	41650183	589,235	Inverted SD	Active	2	23	0	0
chr17	55032245	55407548	375,303	Inverted SD	Inactive	0	0	0	0
chr17	55376032	57709940	2,333,908	Direct SD	Inactive	0	0	0	0
chr18	10646102	12157145	1,511,043	Inverted SD	Inactive	0	0	0	0
chr19	22381717	22624454	242,737	Inverted SD	Inactive	0	0	0	0
chr19	41507164	42436859	929,695	Inverted SD	Active	0	1	0	0
chr19	46025012	46318700	293,688	Inverted SD	Inactive	0	0	0	0
chr19	53152183	55286266	2,134,083	Inverted SD	Inactive	0	0	0	0
chr20	45900945	46554106	653,161	Direct SD	Active	0	1	1	0
chr22	15418962	20940408	5,521,446	Inverted SD	Active	0	1	0	0
chr22	17067917	20039629	2,971,712	Direct SD	Active	22	92	4	0
chr22	17161131	23372017	6,210,886	Direct SD	Active	4	2	0	0
chr22	18748581	19868650	1,120,069	Direct SD	Active	6	7	6	0
chr22	18776142	19384612	608,470	Direct SD	Active	1	0	1	0
chr22	19404326	20149942	745,616	Direct SD	Active	9	3	1	1
chr22	19813332	21303846	1,490,514	Direct SD	Active	2	7	1	0
chr22	19814802	23372017	3,557,215	Direct SD	Active	4	5	0	0
chr22	21327581	22967972	1,640,391	Inverted SD	Inactive	0	0	0	0

Nature Genetics: doi.10.1038/ng.909

Supplementary Table 11. Clinical features associated with 17q21.31 microdeletion.

Clinical features	Previous studies	Percentage	9888884
Low birth weight	6/22	27.3%	+
Microcephaly	1/22	4.5%	+
Short stature	4/22	18.2%	+
Developmental delay	22/22	100.0%	+
Hypotonia	21/22	95.5%	+
Epilepsy	11/22	50.0%	-
Facial features			
Abnormal hair color/texture	13/22	59.1%	+
Broad forehead	15/22	68.2%	-
Long face	14/19	73.7%	+
Hypermetropia	8/22	36.4%	-
Nasal speech	6/12	50.0%	-
Pale irides	10/22	45.5%	-
Strabismus	10/22	45.5%	+
Upslanting palpebral fissures	15/22	68.2%	+
Blepharophimosis	8/22	36.4%	-
Ptosis	11/22	50.0%	-
Epicanthal folds	15/22	68.2%	+
Tubular or pear shaped nose	18/22	81.8%	-
Bulbous nasal tip	20/21	95.2%	+
Prominent ears	13/22	59.1%	+
Narrow/high-arched palate	11/22	50.0%	-
Cleft palate	2/22	9.1%	-
Broad chin	9/21	42.9%	-
Extremities			
Narrow hands	5/22	22.7%	+
Slender/long fingers	5/18	27.8%	+
Hypoplastic hand muscles	5/17	29.4%	+
Dislocation of hip(s)	6/22	27.3%	+
Slender lower limbs	7/17	41.2%	-
Positional feet deformity	6/22	27.3%	-
Other features			
Heart defects	6/22	27.3%	+
Kidney and urologic anomalies	7/22	31.8%	-
Cryptorchidism	7/9	77.8%	-
Scoliosis/kyphosis	8/22	36.4%	-
Pectus excavatum	5/22	22.7%	-
Ventriculomegaly	6/16	37.5%	-
Friendly behavior	16/18	88.9%	+

Nature Genetics: doi.10.1038/ng.909

Clinical features comparing cases with typical (SD-mediated) 17q21.31 deletion (Koolen et al, JMG) [31]versus the atypical deletion are shown above. One five-year-old female was identified to have an atypical, partial deletion of the 17q21.31 commonly deleted region (chr17:41,356,798-41,631,306, UCSC hg18 coordinates; Figure 3). This patient narrows the critical region for the 17q21.31 microdeletion syndrome and narrows the search for the critical gene(s), as she presented with the features of the syndrome. She had the common features of the syndrome, including distinctive dysmorphic features with a bulbous nasal tip, upslanting and almond-shaped palpebral fissures, long face, strabismus, epicanthal folds, and prominent ears; developmental delay with limited speech; hypotonia in infancy; and a friendly disposition. She had additional features that have been seen in some individuals with the microdeletion syndrome, including low birth weight, short stature, microcephaly, long fingers, and heart defects (ASD vs. PFO and thickening of the mitral and aortic valves). Additionally, she presented with postaxial polysyndactyly, neonatal cholestasis, resolved leucopenia, dry skin with some hyperpigmented lesions, and an anteriorly split tongue.

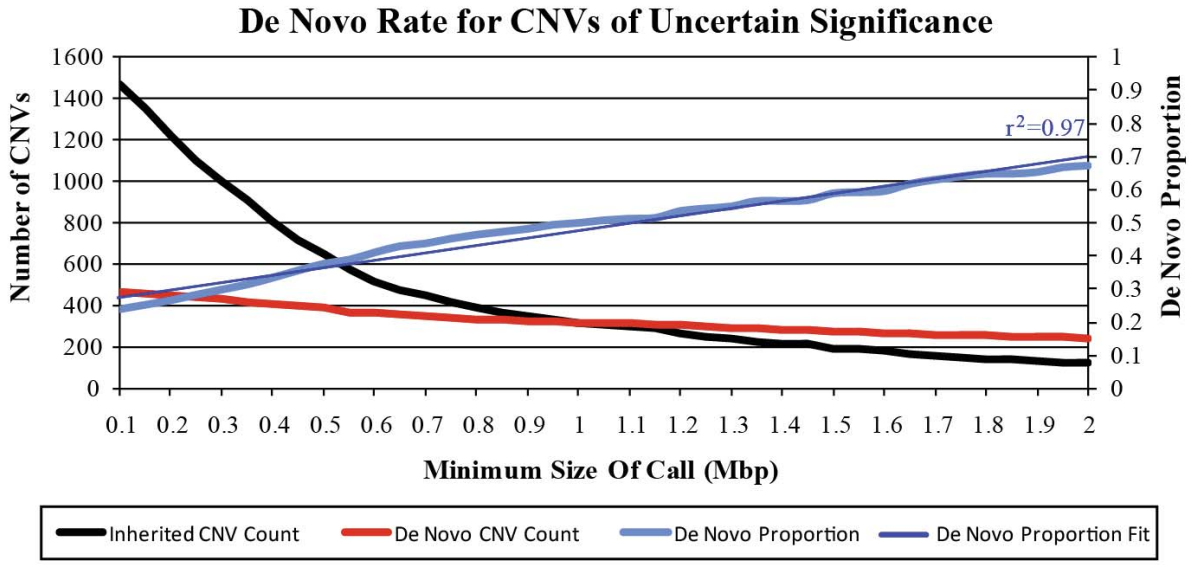
Supplementary Table 12. Gene level statistics. (See Excel sheet)

Supplementary Table 13. Control CNV burden by gene. (See Excel sheet)

Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 1. *De novo* rates for CNVs of uncertain significance.

We examined 2,058 CNVs of uncertain clinical significance for inheritance status (excluding events due to unbalanced translocations as there is a possibility that a parent carries a balanced event). Of these CNVs, 486 were *de novo* and 1,572 were inherited. As expected, *de novo* events tended to be larger (median size 2 Mbp) compared to inherited events (median size 400 kbp). Additionally, the proportion of *de novo* CNVs is strongly correlated with CNV size (R2 = 0.97) and *de novo* events predominate beyond 1 Mbp.



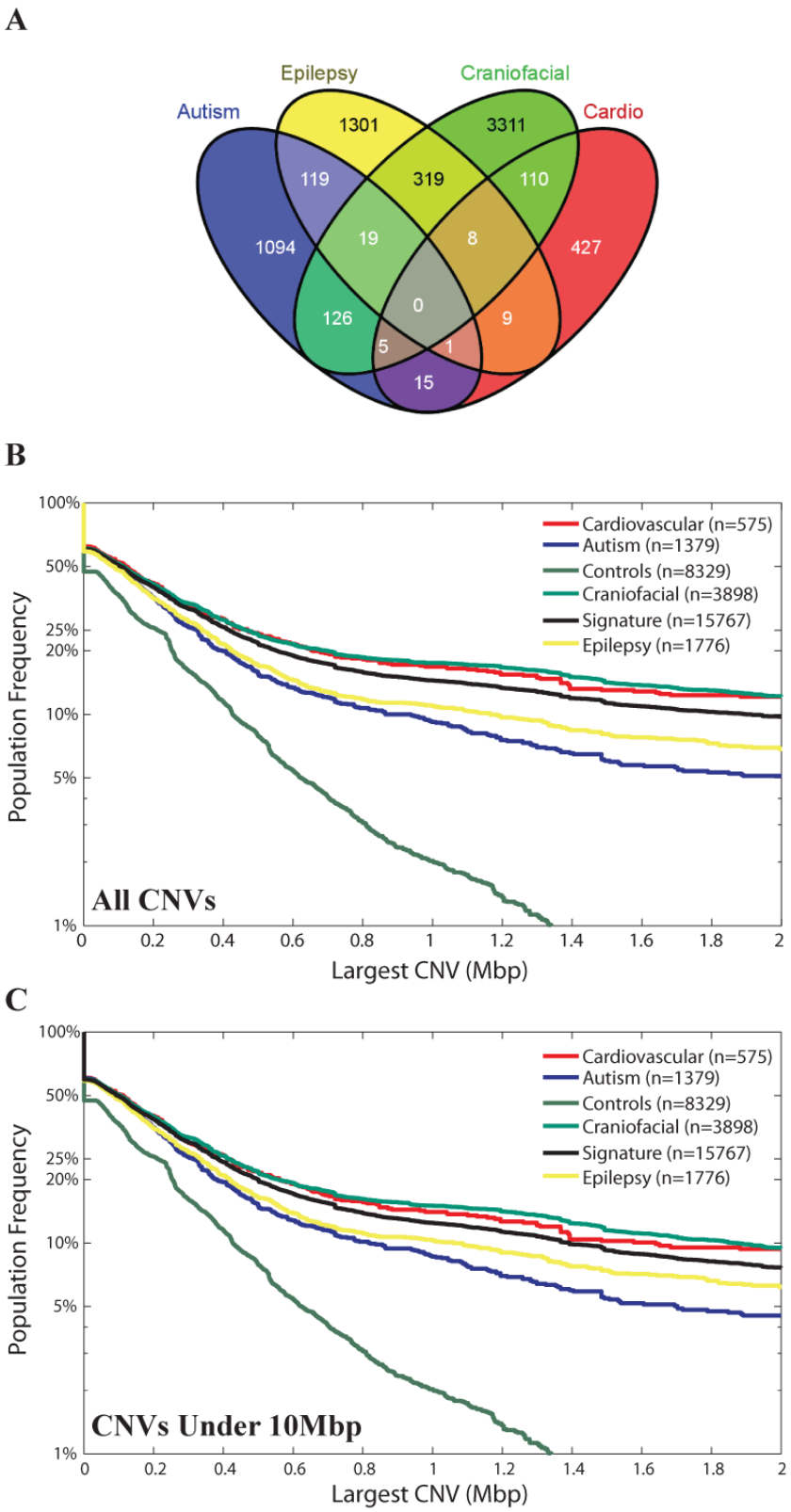
Event Type	Count	Count over 400kbp	Count over 500 kbp	Count over 750 kbp	Count over 1 Mbp	Count over 1.25 Mbp	Count over 1.5 Mbp
De Novo	486	408	392	344	318	299	271
Inherited	1572	808	647	416	319	252	190

Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 2. Phenotype breakdown and CNV burden analysis. (A) Breakdown of phenotype definitions in the Signature Genomics set. **(B)** The population distribution of large CNVs in each subtype of the cases and controls is displayed as a survivor function with the proportion of samples carrying a CNV of a given size displayed as a curve. An increased CNV burden is immediately apparent in the cardiovascular and craniofacial phenotypes, particularly compared to the less severe autism and epilepsy cases. This increased burden is also apparent after filtering large (>10 Mbp) calls **(C)** from our dataset, implying the increased CNV burden is not purely resultant from karyotypically visible abnormalities.

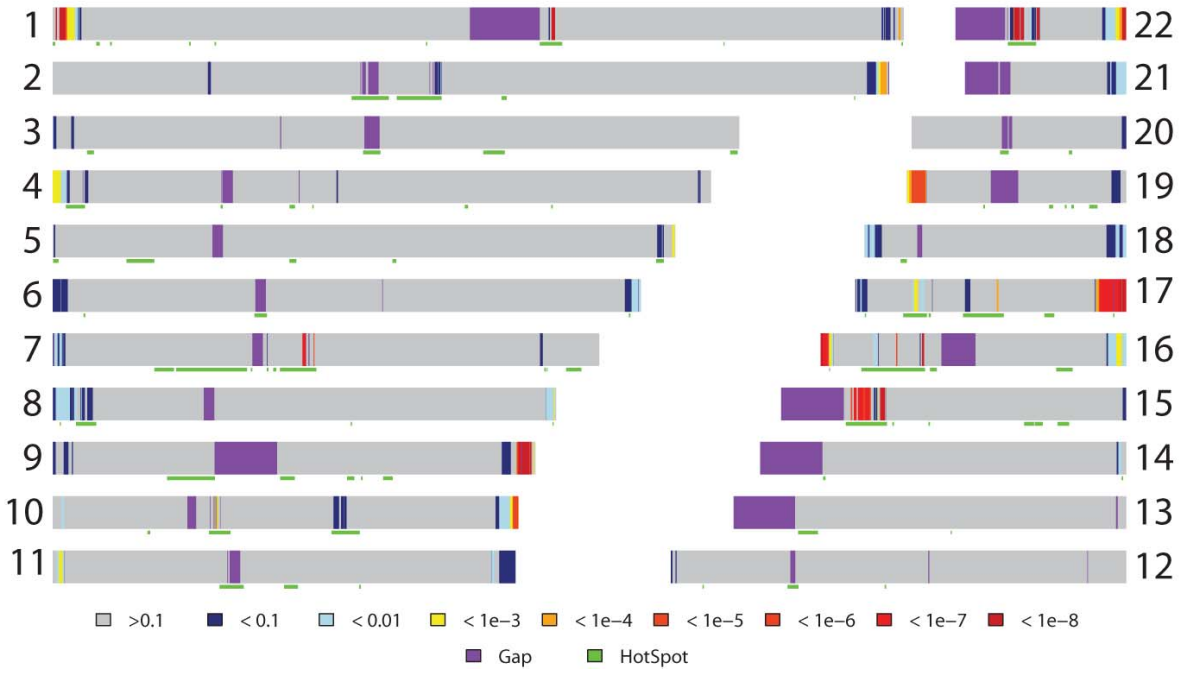
Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 2



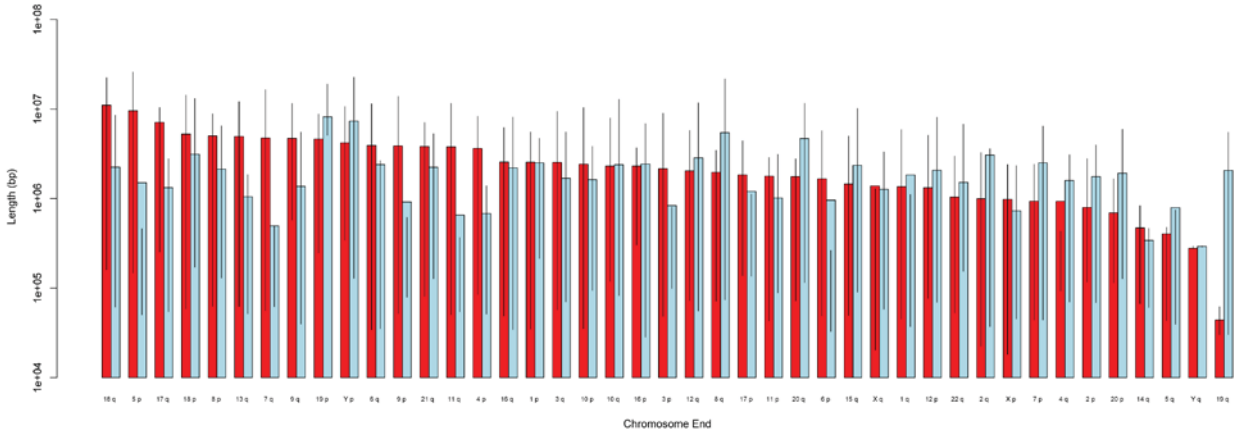
Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 3. Genome-wide heatmap of affected individual CNV enrichment p-values. For each indicated chromosome (outer edges), colored rectangles are plotted for consecutive 200 kbp windows according to the color scale indicated at bottom, with the p-value computed using a one-tailed Fisher's exact test comparing deletion or duplication counts in affected and unaffected individuals; note that the smaller of the deletion or duplication p-value was used. Regions of gaps in the hg18 assembly are colored purple, while duplication-mediated CNV hotspots are indicated with green lines underneath each chromosome.



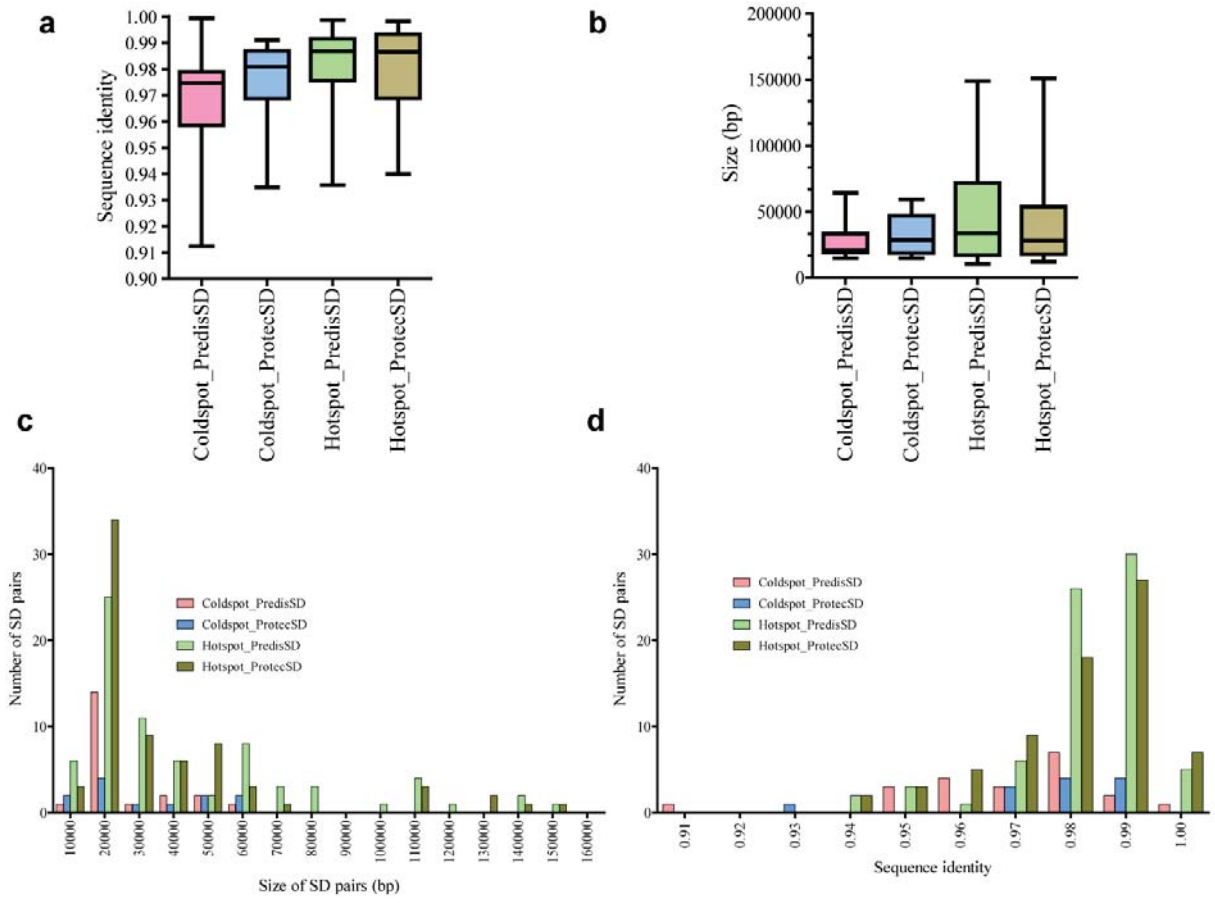
Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 4. Size distributions for subtelomeric (<1.5 Mbp from a chromosome end) deletions (red) and duplications (light blue), excluding all variants >30 Mbp. For each chromosome end of hg18 (X-axis), the bar height (Y-axis, log-scaled) shows the mean length, while the thin black line spans the 10th to 90th percentile of the size distribution. Chromosome ends are sorted by mean deletion length. Note that in some cases the mean lies outside the 10th to 90th percentile range (e.g. duplications on 5p).



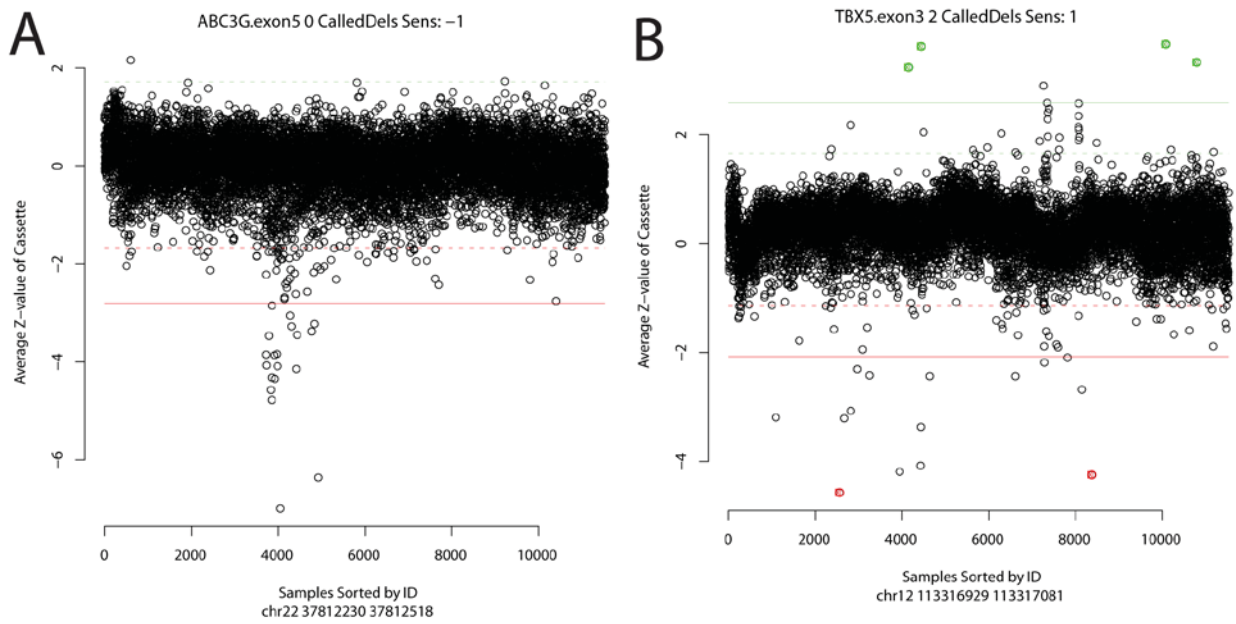
Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 5. Segmental duplication architecture and predisposition to rearrangements. Segmental duplications (SDs) predispose to rearrangements if they are in direct orientation and do not predispose to rearrangements if they are in opposing orientations. **(A)** Sequence identity of interspersed (distance >50 kbp to 5 Mbp) SDs flanking active and "cold" hotspot regions are shown. Active hotspots are defined as those with CNV events and cold hotspots are refractory for CNV rearrangements. (**B**) SD sizes for predisposing and protective SDs are shown. (**C**) Distribution of the number of SD pairs versus the size is shown for all categories of hotspot regions. (**D**) Distribution of the number of SD pairs as a function of sequence identity is shown for all the hotspot regions.



Nature Genetics: doi.10.1038/ng.909

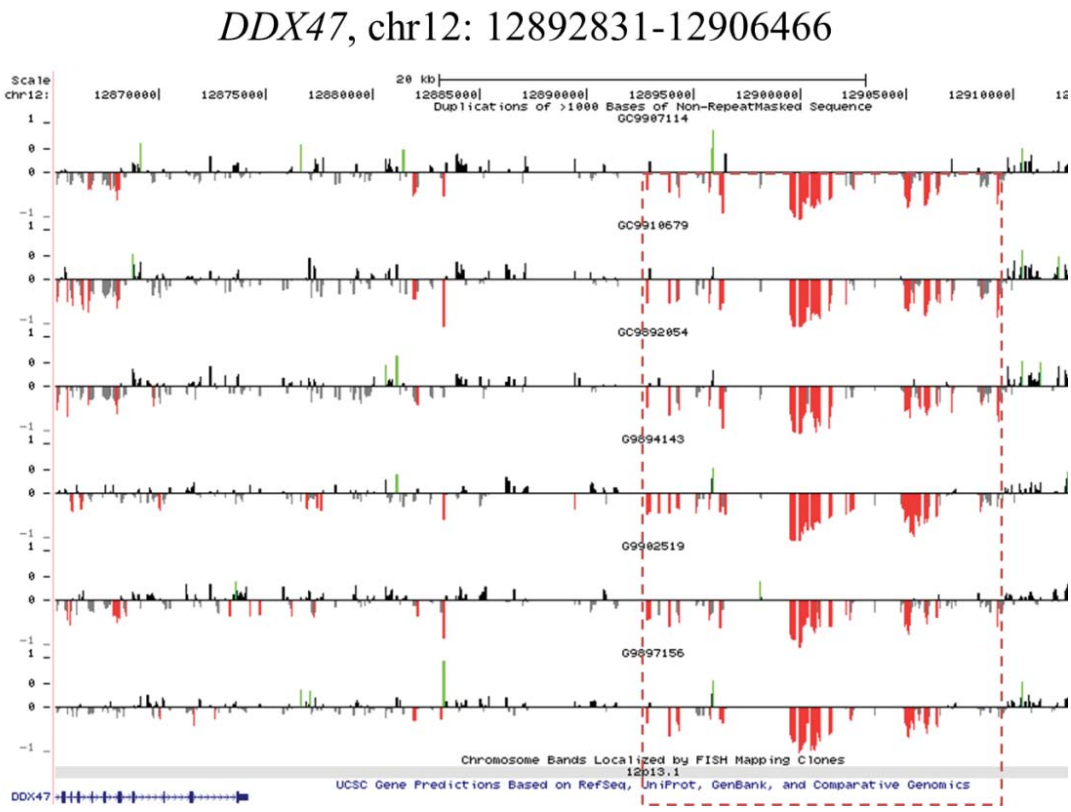
Supplementary Figure 6. Examples of "bad" (A) and "good" (B) exon "cassettes". For both panels, the average cassette score (Y-axis) is shown for each sample, plotted in sorted sample identity order (X-axis). Solid red and green lines correspond to 5 standard deviations below and above, respectively, the global mean, while dashed lines indicate 3 standard deviations. These lines are only illustrative, as thresholds for identifying outliers were computed for each array type independently. Samples that carry a larger, "called" deletion that spans the cassette are colored red, and duplications are colored green. The fraction of samples that would be correctly classified as deleted using only this cassette is shown in the plot title ("Sens"; set to -1 when no large deletions are present). In panel A, all the outliers cluster together suggestive of a batch artifact and there is not a clear visual separation between the outliers and bulk of the samples. Conversely, in panel B (corresponding to subsequently validated predictions shown in Table 3), the outliers are spread across array type and have clear visual separation from the bulk distribution.



Nature Genetics: doi.10.1038/ng.909

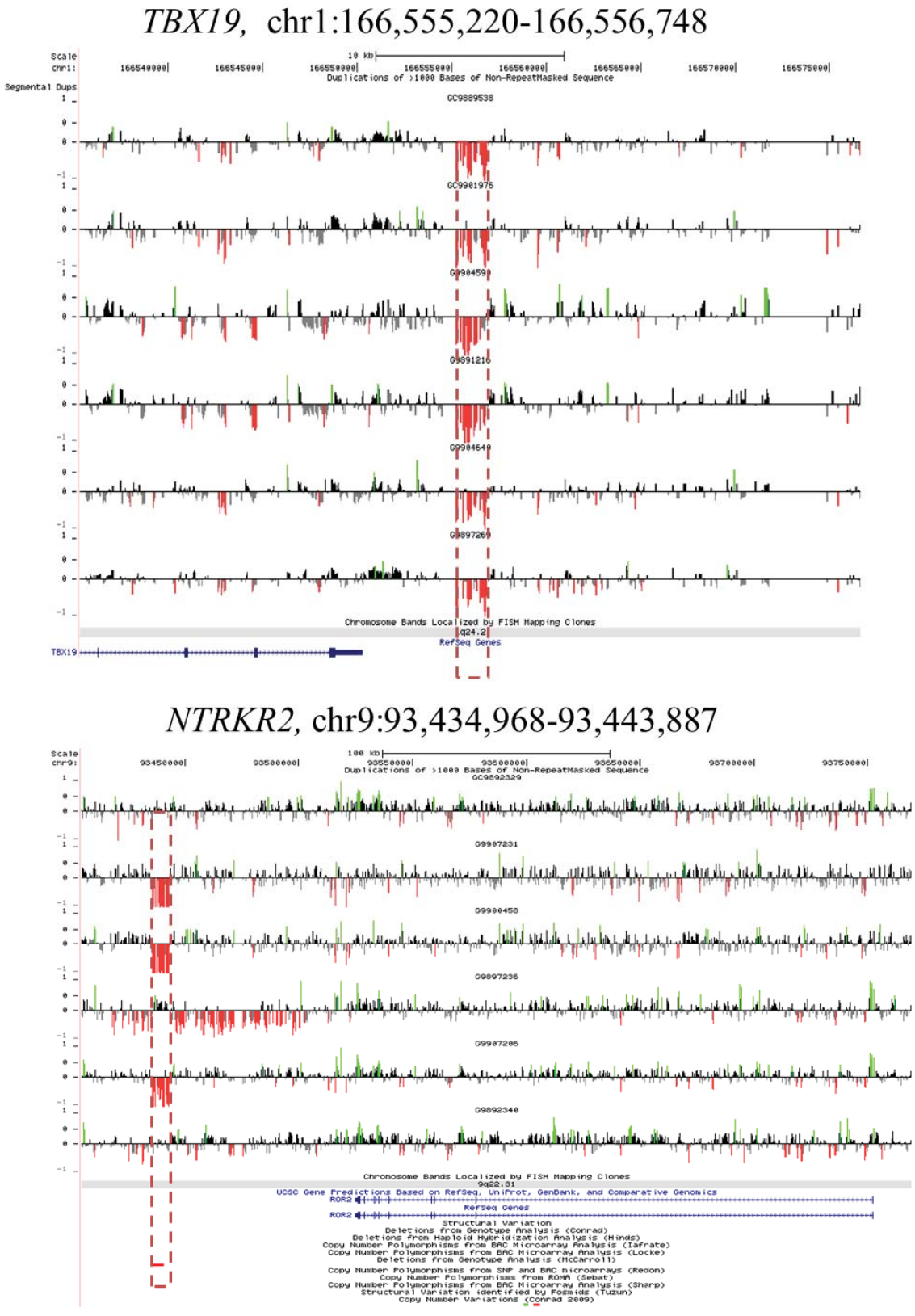
Supplementary Figure 7. ArrayCGH validation for 10 CCDS gene deletions.

Please note that these deletions were based on UCSC CCDS gene predictions, not RefSeq genes.



Nature Genetics: doi.10.1038/ng.909

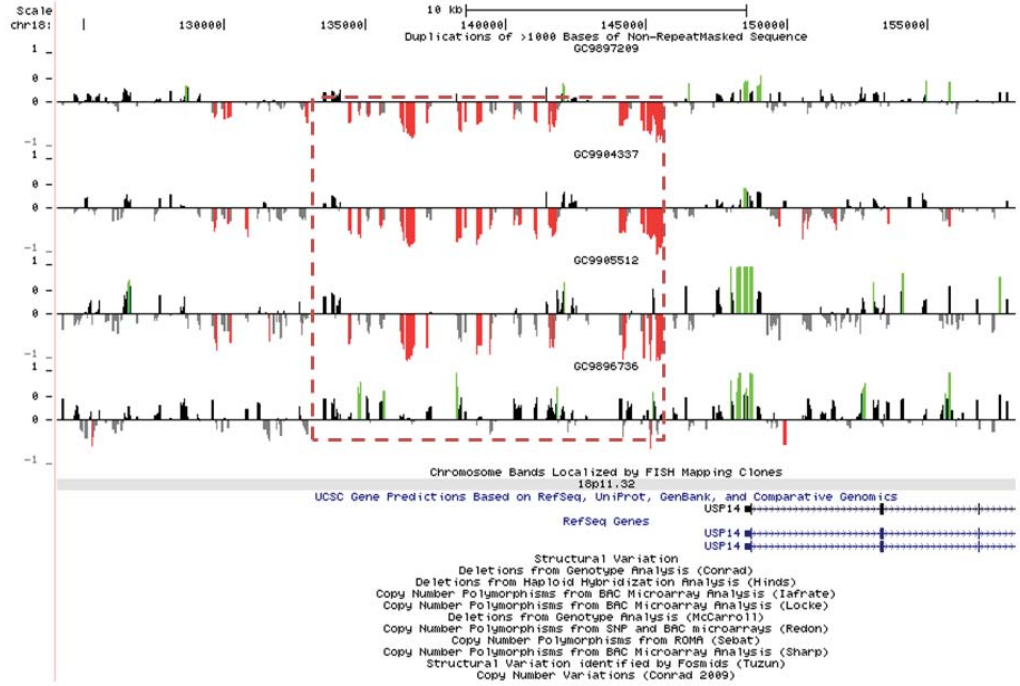
Supplementary Figure 7 continued.



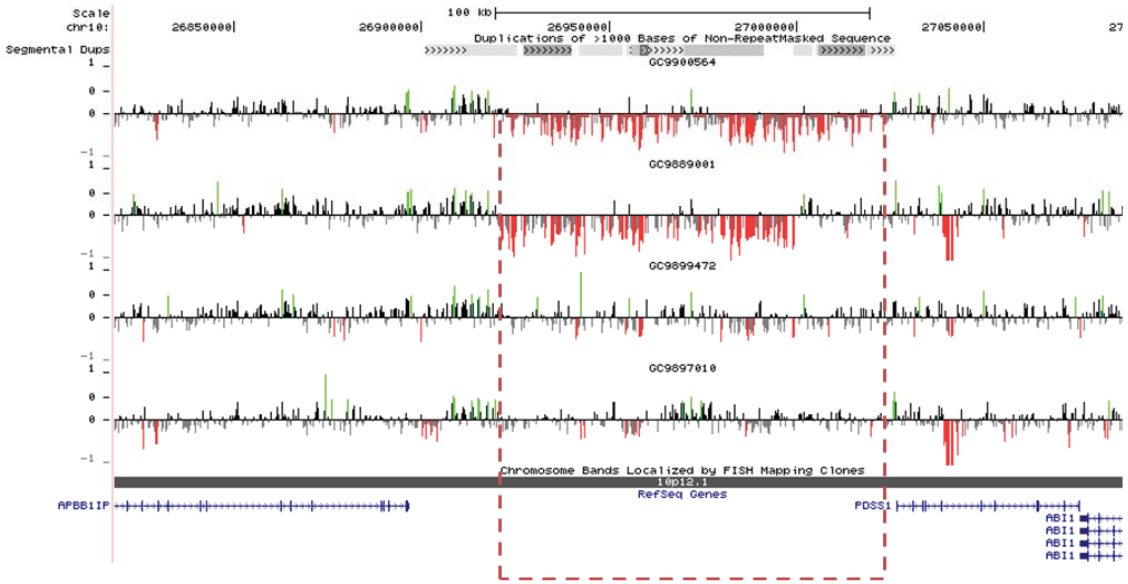
Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 7 continued.

USP14, chr18:134300-145900



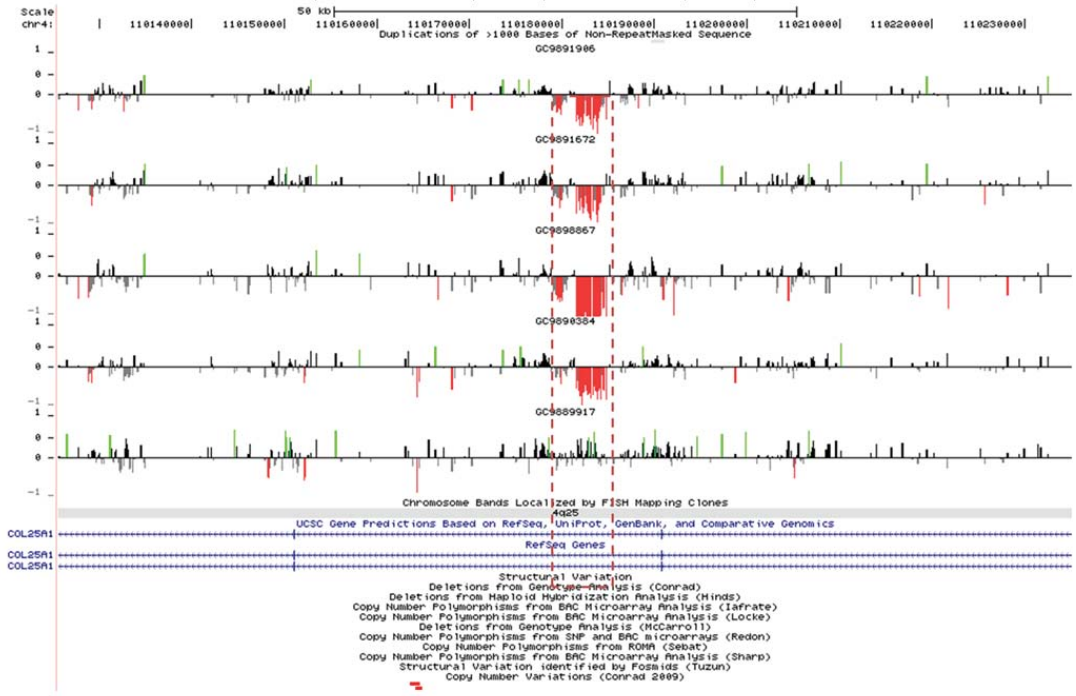
APBB1IP, chr10 2691900 27020000



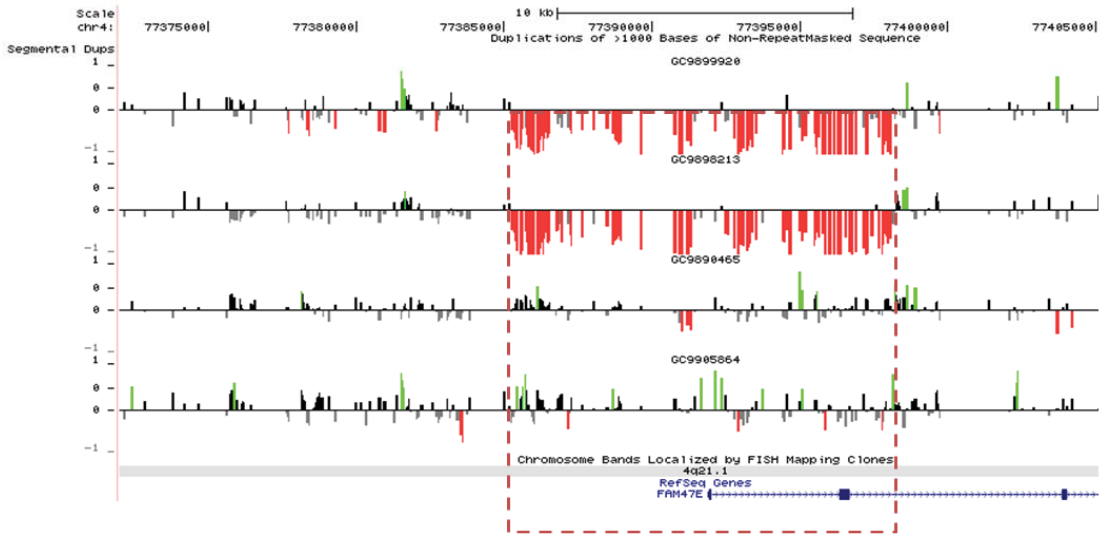
Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 7 continued.

COL25A1, chr4:110,178,588-110,184,801



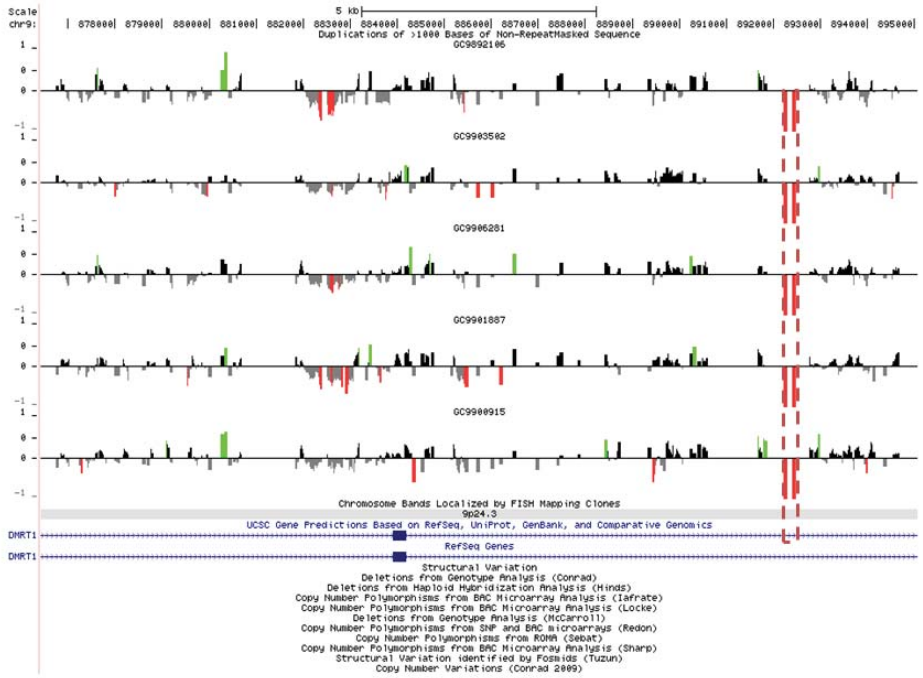
FAM47E, chr4:77,385,201-77,398,400



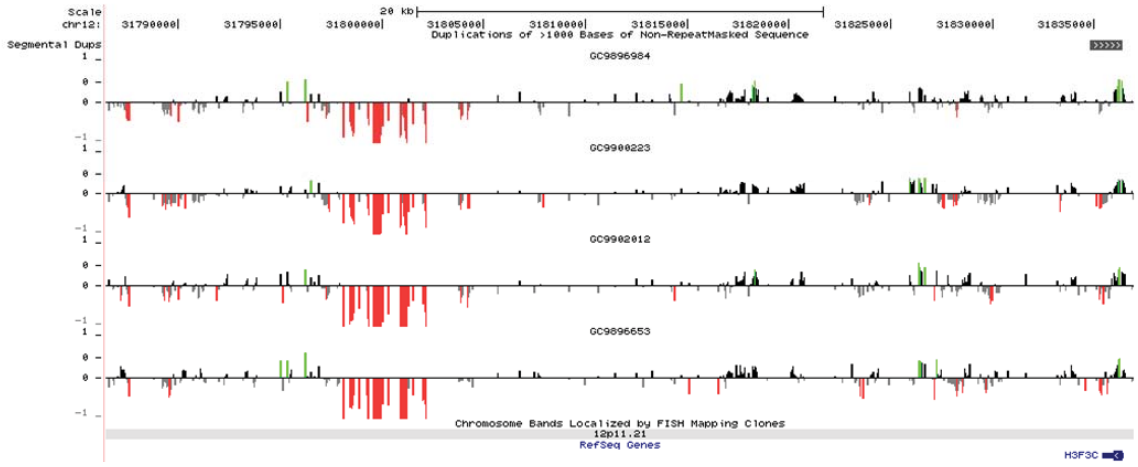
Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 7 continued.

DMRT1, chr9:892,036-892,565



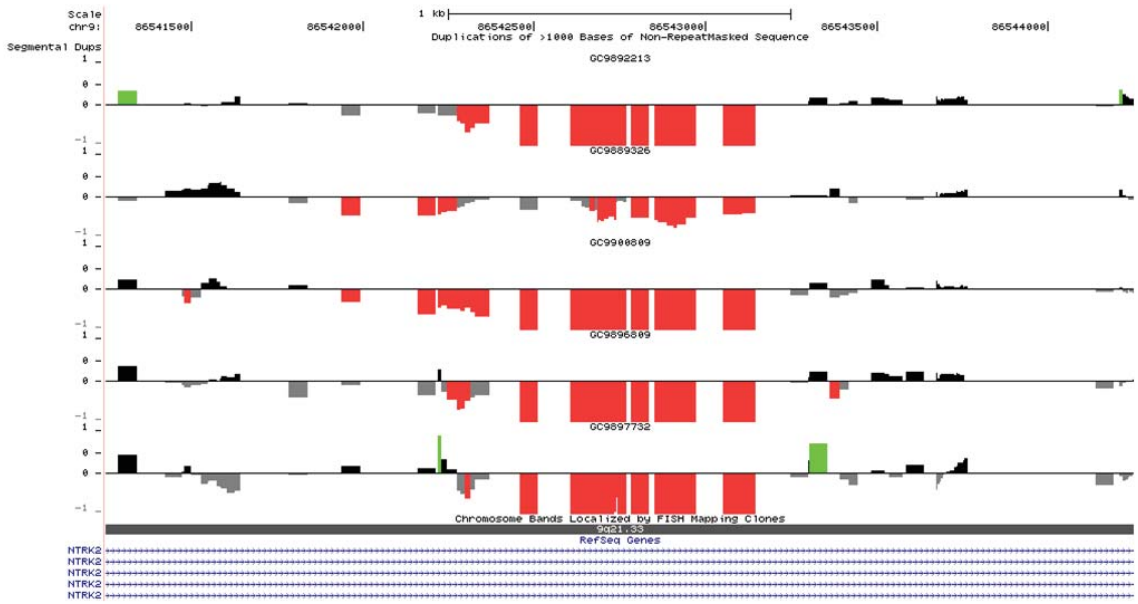
H3F3C, chr12:31,835,114-31,835,838



Nature Genetics: doi.10.1038/ng.909

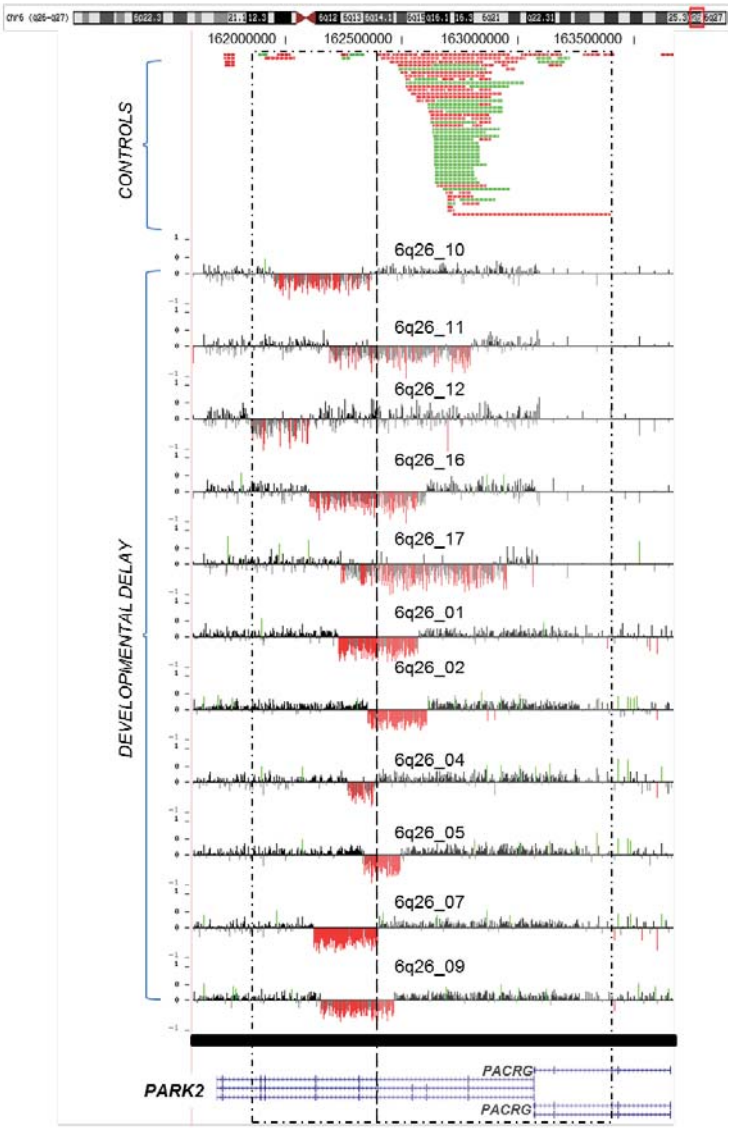
Supplementary Figure 7 continued.

NTRK2, chr9:86,542,251-86,543,250



Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 8. Array CGH data shows deletions within *PARK2*. Shown are array-CGH results from a custom array (see Supplementary Note below) for the genomic region including *PARK2*. The log2-ratio values for each probe are shown along the X-axes, in hg18 coordinates (indicated at top). Strongly negative probes are colored red and strongly positive probes are colored green. There is a potential 3' bias of case relative to control deletions (vertical dashed lines), although this difference is not statistically significant. Note also that these deletion carriers were identified during a validation and confirmation effort of prediction deletions that are separate from the exon-specific deletion analysis summarized in Table 3 of the main text.

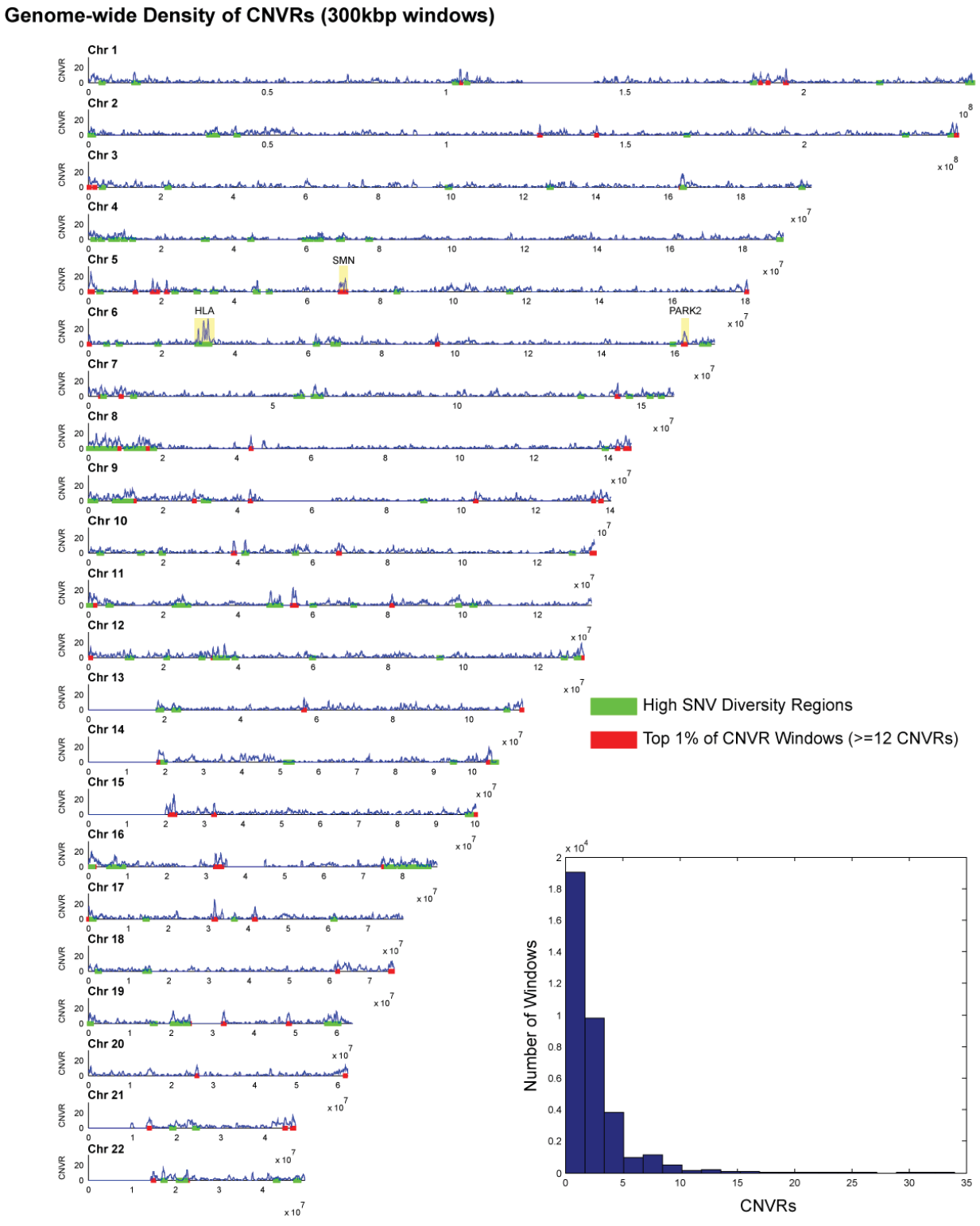


Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 9. Genome-wide density of CNVRs. The blue line represents the total unique CNVR count for each 300 kbp window across the genome. Green highlights indicate regions of high single-nucleotide variant (SNV) diversity while red highlights represent the top 1% of CNVR windows (green supersedes red). Highlighted regions of particularly high CNVR density include the HLA locus on chromosome 6p that corresponds to a region of high SNV diversity, and the SMN locus on chromosome 5 that does not demonstrate significant SNV diversity. Strikingly, the high CNVR density regions include several disease-associated genes including *PARK2* (highlighted).

Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 9

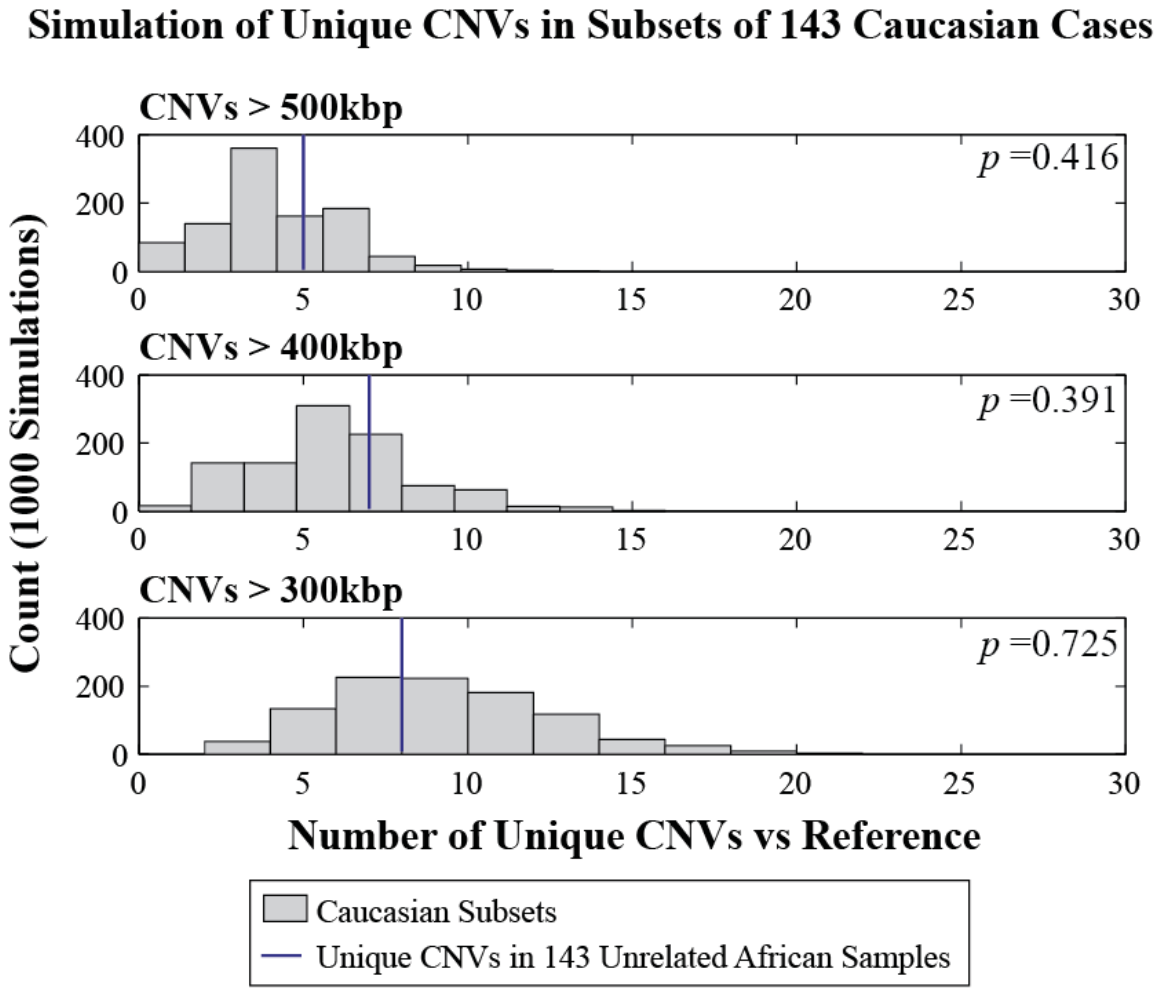


Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 10. Unique large CNV rates among ethnic groups.

Due to the slight enrichment of African/African American Individuals in our case data (also see Supplementary Figure 10), we tested for normal differences in large CNV loci between control populations.

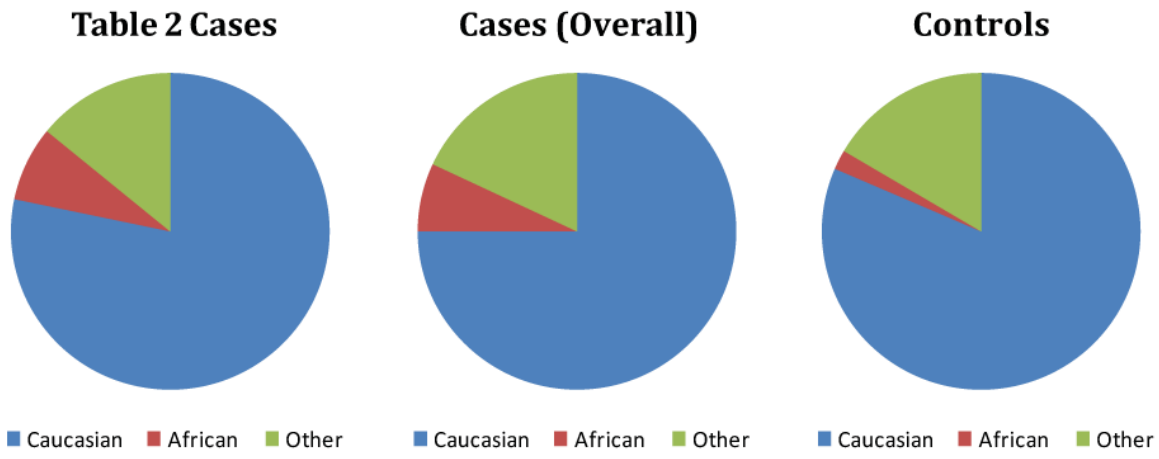
To simulate the expected number of unique CNVs in a given subpopulation, we first selected 3,381 random Caucasian controls as a reference group, and selected 1,000 random sets of 143 Caucasian individuals from an additional pool of 3,380 Caucasian controls. We then enumerated the number of unique CNVs (defined as not having a 50% reciprocal overlapping CNV in the reference group) found in each set of 143 samples. Finally we calculated the number of unique CNVs detect in our 143 unrelated African controls and compared this value to the simulation data. We noted no significant enrichment of unique CNVs in the 143 African samples at 300, 400, and 500 kbp, suggesting that population stratification is unlikely to be a major confounding factor in our identification of large pathogenic CNVs.



Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 11. Ethnic Stratification in Table 2 Regions, Cases and Controls.

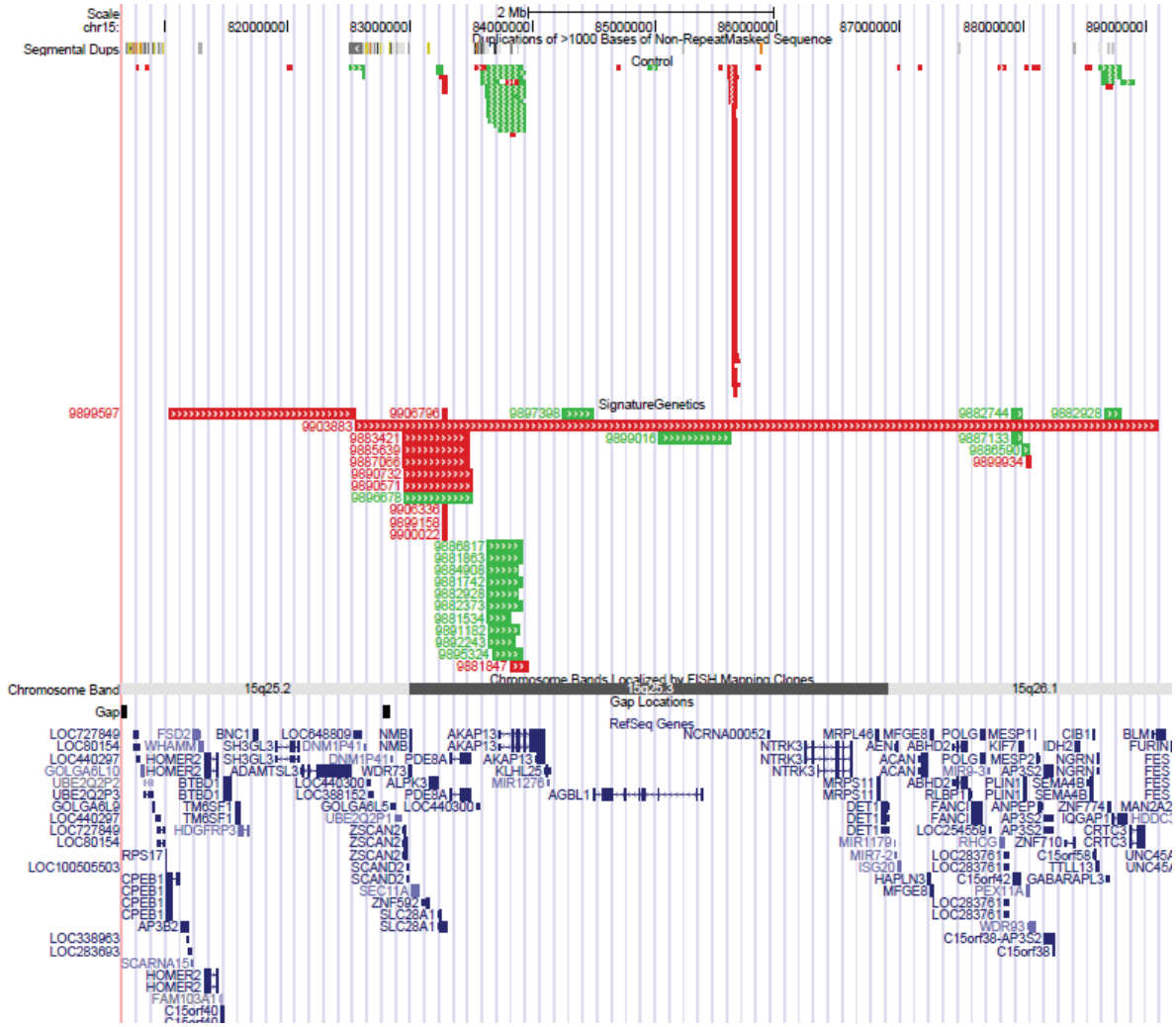
Indicated are the distributions of ethnicities in the Control and Case cohorts.



Nature Genetics: doi.10.1038/ng.909

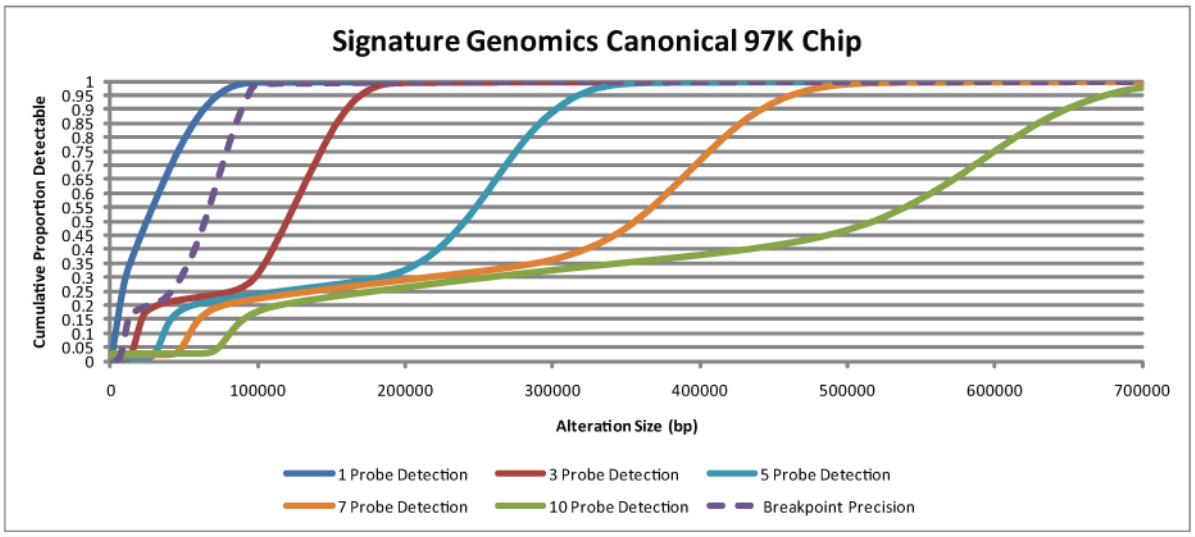
Supplementary Figure 12. Deletions on chromosome 15q and 15q25.2q25.3.

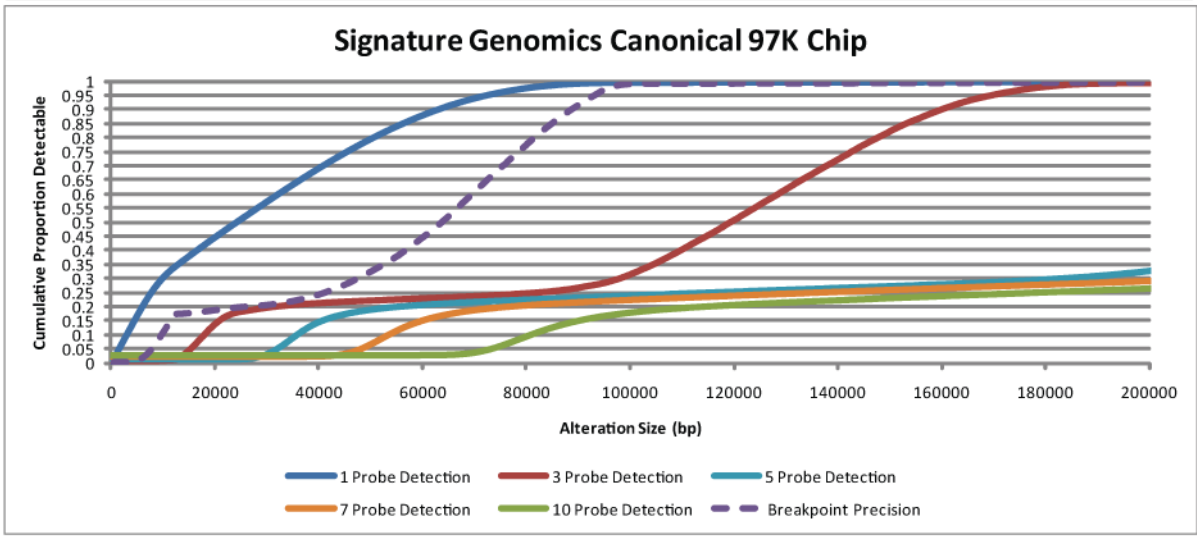
Among 15,767 individuals referred for clinical array CGH testing, six were identified as having the 15q25.2q25.3 deletion (Supplementary Figure 12). Interestingly, five of these cases carried a 660 kbp deletion (chr15:82,889,423-83,552,890) mediated by SDs (69.8 kbp, 98.6% identity) and two of these five were brothers. The critical deletion region contains eight RefSeq genes. Detailed clinical information was available for two probands, and the clinical features were varied (Supplementary Table 13). Clinical features were variable consisting of neurologic features and developmental delay; one female had only mild motor delays associated with a congenital myopathy and was cognitively normal. No dysmorphic features were reported. The two brothers with the deletion both had autism spectrum disorders, but additional family members were not tested. The deletion was also found in one mother whose only health complaint was migraines.



Nature Genetics: doi.10.1038/ng.909

Supplementary Figure 13. Resolution of the canonical 97K Signature Genomics platform. Shown is the probability of detecting an alteration of varying sizes based on requiring a minimum number of probes as calculated by ResCalc[1]. Curves are included for 1,3,5,7 and 10 probe requirements. The sigmoidal shape of the data is indicative of the targeted design of the platform used by Signature Genomics where portions of the genome are covered at high density, with uniform backbone coverage.





95% Detection Thresholds:
1 Probe: 71.9 kbp
3 Probes: 169.1kbp
5 Probes: 316.5 kbp
7 Probes: 461.1 kbp
10 Probes: 674.1 kbp

Breakpoint Precision:
95% @ <92.9 kbp

Nature Genetics: doi.10.1038/ng.909

Supplementary Note

Detection and validation of large CNVs using array CGH

Whole-genome bacterial artificial chromosome (BAC) microarray chip (SignatureChipWG®) and an oligo-based (SignatureChipOS®) chip (either 105K custom designed by Signature and manufactured by Agilent Technologies or 135K custom designed by Signature and manufactured by Roche NimbleGen) were used for CNV detection. Microarray hybridizations were performed as described previously[2]. CNVs from the Signature collection were then rigorously scrubbed to eliminate potential size estimation errors associated with low probe densities, intensity noise resulting from high-copy duplications, rearrangements associated with immune genes, reference-sample CNVs, and other potential artifacts. We filtered CNVs according to the following criteria. CNVR count <158 (1% of individuals), >1 probe per 150 kbp (hg18), <10% of the CNV overlapping with an "artifact" list. To define CNVR frequency, we used a previously published approach[3]. We first sorted CNVs by size and, beginning with the largest, collapsed all smaller or equal-sized CNVs of the same type (deletion or duplication) and both with start and stop coordinates within 50 kbp of the initial CNV. This procedure then moves to the next largest CNV not already assigned to a CNVR and continues until all CNVs have been assigned. The list of artifactual loci was defined by regions with immune system genes prone to rearrangement, known reference sample CNVs, very large blocks of highly similar SDs with reduced SNP array probe densities, and artifacts identified as part of a batch effect. The artifact list includes: chr2: 88937989-89411302, chr2: 89589457-89897555, chr2:196517337-196847645, chr3:30618438-30728248, chr7:105609512-105811026, chr14:21159851-22090936, chr14:105065301-106352275, chr15:0-20060121, chr15:91157836-91364629, chr16:87299650-87418927, chr22:20602619-20926359, and chr22: 20715572-21595082. Additionally, all CNVs present at greater than 1% population frequency were removed from cases (but left in controls).

Validation experiments were performed in four steps (Supplementary Table 3). (1) Fluorescent *in situ* hybridization (FISH) using BAC probes targeted to specific regions. (2) Custom-designed arrays targeted to genomic hotspots. This array consisted of 135,000 probes tiled at one probe every 2.5 kbp within the genomic hotspots and genomic backbone with probe spacing every 36 kbp. (3) Custom-designed, high-density arrays targeted to genomic hotspots with 400,000 probes

Nature Genetics: doi.10.1038/ng.909

distributed to disease-associated hotspots with median probe spacing of 500 bp, median probe spacing of 8 kbp in the other genomic hotspots, and a genomic backbone with median probe spacing of 14 kbp (Girirajan and Eichler, unpublished). (4) Custom-designed arrays to test specific regions of interest: 16p11.2, 16p12.1, *AUTS2*, and *PARK2* CNVs with a median probe density of 125 bp. All microarray hybridization experiments were performed using a single unaffected male (GM15724 [Coriell]) as a reference.

Regions of excess CNV in control samples

We first examined controls in the context of gene disruptions and global patterns of CNVs >5 kbp. In an initial gene level analysis, we found, consistent with previous studies[4-6], that the most frequently affected genes include olfactory receptors, cytochrome p450 genes, and other SD-associated genes (Supplementary Table 13), supporting the consistency of our control samples with previous studies. To identify genes highly mutable by CNVs, we merged unique CNVs into CNVRs using a 50% reciprocal overlap criterion, and then used a sliding window approach to identify genes and regions with highly heterogeneous collections of CNVs. To determine if the regions of high CNVR load are due to population sequence diversity, or fragile regions, we correlated our results with regions of high SNP diversity[7] and the density of dbSNP entries. This analysis identified a subset of regions that demonstrated both high SNP diversity and CNVR load, including the *HLA* locus on 6p, while other regions demonstrated high CNVR load in the absence of high SNP diversity, such as the *SMN* locus on 5q and *PARK2* on 6q (Supplementary Figure 7). The identification of *PARK2* as a variant region in controls is of particular interest as it has been linked to recessive juvenile Parkinson's disease[8]. *PARK2* is the 21st most commonly disrupted gene, with at least 22 distinct exon-affecting deletions in controls, and furthermore overlaps the fifth most CNVRs of any gene. Other notable genes with high CNVR burden previously reported in individuals with neuropsychiatric disease include: *IMMP2L* (19 CNVRs, ranked 15th), which is associated with attention-deficit/hyperactivity disorder[9], autism[10,11], and Tourette syndrome [12]; and *CNTNAP2* (15 CNVRs, ranked 24th), which has been associated with autism[13] (Supplementary Table 13).

Nature Genetics: doi.10.1038/ng.909

Analysis to identify genomic disorder regions

Prior to intersection with the CNV data for both cases and controls, we defined genomic coordinates as critical, exact, or maximum based on published data. We identified 52 genomic disorder regions (Supplementary Table 5). Since these regions were interrogated for frequency of disease-associated CNVs as opposed to discovery, we utilized all the unfiltered CNV calls for the analysis.

Genomic hotspot analyses

Using the SD architecture of the human genome sequence assembly (build36), we curated a list of hotspot regions that potentially mediate rearrangements by NAHR. Restricting our analysis to only intrachromosomal regions, a nonredundant set of 111 hotspots were identified. Only 54/111 regions contained the predisposing directly orientated SD structure necessary for NAHR events while the remainder (57/111) carried a "protective architecture" (i.e. SDs in opposite orientations). We then counted CNVs in cases and controls that map within these genomic hotspots. While 46/54 predisposing hotspots regions harbored at least one CNV in cases or controls; eight regions were essentially "cold" lacking any events. Absence of hotspot CNVs in spite of a predisposing duplication structure suggests a possible extreme rarity of the event or incompatibility of the event with normal viability. In contrast, of the 57 regions with protective SD architecture, only 16 regions harbored at least one CNV, likely indicating that most of these loci are not prone to NAHR. The existence of CNVs despite a protective architecture suggests the presence of alternative SD structures distinct from the reference sequence assembly[14,15]. Notable examples include recently described pathogenic loci including 17q21.31 (*MAPT*)[16-18], 15q13.3 (*CHRNA7*)[19-22], 15q11.2 (*NIPA1*)[23,24], 2p13 (*NPHP1*), and 7q11.2 (*ZNF92*)[25]. Further, 41 protected regions did not carry any CNVs likely due to presence of protective SD architecture. We also note that the duplications flanking active hotspots were larger and with higher sequence identity compared to "cold" hotspots (Kolmogorov-Smirnov test, p=0.0022) (Supplementary Figure 5), further supporting the role of genome structure in relation to disease risk.

Nature Genetics: doi.10.1038/ng.909

Sliding window analysis

To identify genomic regions enriched for CNVs among cases, we performed a heuristic sliding window procedure. For 200 kbp windows and a step size of 50 kbp, we counted overlaps for all CNVs larger than 400 kbp and smaller than 10 Mbp in cases and controls; Signature CNVs spanning fewer than 10 SNP probes on the lowest density SNP array (~317,000 probes) were excluded, and all windows containing frequent CNVs (>1% of samples) were ignored. The CNV counts for each window were then subjected to a one-tailed Fisher's exact test. Windows with p-values <0.1 separated by <400 kbp were merged together, and all merged blocks with at least five windows with p <0.1 were flagged. These regions were subject to additional filters based on gap and SD content and manual analyses of the individual CNVs (not shown).

Below we summarize the numbers, sizes, ethnicities, and inheritance status for CNV carriers at each of the 15 novel CNV-loci identified as enriched (14 are represented in Table 2). Note that the counts presented below explain the "adjusted" counts in Table 2, and include carriers of CNVs that overlap any portion of the significant region (some of which do not overlap the single most significant window within the region). The coordinates in Table 2 correspond to the midpoints of the windows that bound the region, and as such do not precisely describe the actual CNV breakpoints (see Supplementary Table 8).

10p15.3 duplications:

35 carriers in Signature, median size of 429 kb, 1 of 1 tested CNV is inherited.

Of 28 with known ancestry, 6 are European, 19 are African-American, and 3 are other.

There are 7 controls with large (>400 kb) CNVs overlapping a portion of this interval. 2 are from the WTCCC data (British/European), 2 are from Itsara *et al* [26](Hispanic admixed population from Mexico City), 1 is a Biaka Pygmy individual from the HGDP (African), 1 is a Bedouin individual from the HGDP (Israel-Negev/Middle Eastern), and 1 is from the PARC study [3](individuals of European ancestry in the USA). We additionally found evidence for three individuals of African ancestry within the 1000 genomes study population [27].

These data include two siblings (total independent affected individual count of 34).

The enrichment of African-Americans among cases with 10p15.3 duplications (~68%) relative to all Signature cases studied here (~7.5%) and the observation of 3 additional control carriers

Nature Genetics: doi.10.1038/ng.909

among individuals of African descent in the 1000 Genomes project suggest that the apparent disease association here is an artifact of population stratification.

2q13 deletions and duplications:

21 carriers in Signature, median size of 1.7 Mb, 2 of 8 tested CNVs are *de novo*.

Of 18 with known ancestry, 15 are European, 1 is African-American, and 2 are other.

There is 1 control deletion from the InChianti study of Italians from Tuscany (European).

10p23.1 deletions:

13 carriers in Signature, median size of 7.2 Mb, 3 of 7 tested CNVs are *de novo*.

Of 10 with known ancestry, 8 are European, none are African-American, and 2 are other.

There data also include two siblings, and a mother with two children (total independent affected individual count of 10).

2p21 duplications:

9 carriers in Signature, median size of 542 kb, no inheritance data.

Of 8 with known ancestry, all are European.

4p16.1 duplications:

8 carriers in Signature, median size of 975 kb, no inheritance data.

Of 8 with known ancestry, 6 are European, 1 isAfrican-American, and 1 is other.

These data also include two siblings (total independent affected individual count of 7).

4q21.21q21.22 deletions:

8 carriers in Signature, median size of 2.0 Mb, 3 of 3 tested CNVs *de novo*.

Of 7 with known ancestry, 6 are European, 1 is African-American, and none are other.

2p25.3 duplications:

7 carriers in Signature, median size of 1.7 Mb, 1 of 1 tested CNV is inherited.

Of 4 with known ancestry, 3 are European, none are African-American, and 1 is other.

2q24.3 deletions:

Nature Genetics: doi.10.1038/ng.909

7 carriers in Signature, median size of 3.6 Mb, 4 of 4 tested CNVs are *de novo*.

Of 6 with known ancestry, 5 are European, none are African-American, and 1 is other.

21q21.1 deletions:

7 carriers in Signature, median size of 3.2 Mb, 2 of 2 tested CNVs are inherited.

Of 4 with known ancestry, 1 is European, 1 is African-American, and 2 are other.

These data also include two siblings (total independent affected individual count of 6).

8q11.23 duplications:

7 carriers in Signature, median size of 525 kb, 1 of 1 tested CNV is inherited.

Of 7 with known ancestry, 6 are European, none are African-American, and 1 is other.

1q24.3 deletions:

6 carriers in Signature, median size of 6.9 Mb, 1 of 1 tested CNV is inherited.

Of 5 with known ancestry, all are European.

12p13.31 duplications:

7 carriers in Signature, median size of 8.1 Mb, 1 of 1 tested CNV is *de novo*.

Of 7 with known ancestry, all are European.

15q25.2 deletions:

6 carriers in Signature, median size of 544 kb, 3 of 3 tested CNVs are inherited.

Of 5 with known ancestry, 1 is European, 2 are African-American, and 2 are other.

These data include two siblings (total independent affected individual count of 5).

We also assessed the ethnicity of the literature-reported affected individuals that carry CNVs at 15q25. There are 2 schizophrenia-affected individuals from ISC *Nature* 2008 [28]; all individuals in that study were of European ancestry. There is one autism case from Marshall *et al* [29] that is not specifically annotated; however, >90% of the studied cases were European, while <1% were African (5% were admixed European, 5% were Asian). An additional autism case was identified in Christian *et al* [30], and is listed as "unknown" ancestry (a label applied to ~25% of all cases studied therein). For those cases with known ancestry from Christian *et al*, ~80% were

Nature Genetics: doi.10.1038/ng.909

European, while only 1.3% were African. We conclude that enrichment for 15q25 deletions is unlikely to be an artifact of stratification, although it cannot be ruled out given the limited sample size.

6p22.3 deletions:

6 carriers in Signature, median size of 3.3 Mb, 2 of 2 tested CNVs are *de novo*.

Of 3 with known ancestry, 1 is European, 1 is African-American, and 1 is other.

Gene enrichment analyses

In addition to the sliding window–based analysis of the Signature CNVs, we performed a gene-focused analysis. Genes were flagged based on intersection with CNVs under 10 Mbp, and only coding disruptions were enumerated. We then utilized a one-tailed 2x2 Fisher's exact test to compare the disruption frequencies for genes gained and lost in less than 1% of cases or controls in the overall Signature set, as well as cases described by one of the following subphenotypes: autism, epilepsy, cardiovascular, craniofacial, or neurological. Regions were defined by close genomic proximity (750 kbp, or adjacent) of significantly disrupted genes, and assignment as subphenotype-specific was determined based on the maximum count of genes significant in the overall analysis and genes significant in the subphenotype analysis. Analysis of global functional disruption was performed using the functions feature in Ingenuity Pathway Analysis.

Clinical features of individuals with CNVs in *PARK2*

We found *PARK2* to contain at least six distinct exon-affecting deletions ranging in size from 118 to 315 kbp (Supplementary Figure 8, Figure 4), a phenomenon that also appears to hold true for control samples. While there appears to be clustering of CNVs in cases around the 3' end of the gene (Supplementary Figure 8) compared to control data, our analysis does not conclusively provide sufficient statistical support for this observation. Further analysis including more cases and controls is necessary.

Nature Genetics: doi.10.1038/ng.909

Validation of smaller deletions

We designed a custom microarray targeted to 19 Consensus Coding Sequences (CCDS) gene models, with high density within and near exons and regular spacing within intronic and up- or downstream:

1. ~60 bp spacing at exons and probes from previous experiments, extended 2 kbp both up and downstream;
2. ~1 kbp spacing between targeted sites from set 1 (~2 kbp spacing for genes >500 kbp);
3. ~1 kbp spacing for 100 kbp upstream of the first exon and downstream of the last exon;
4. ~10 kbp spacing for 1 Mbp up- and downstream of the boundaries set 3.

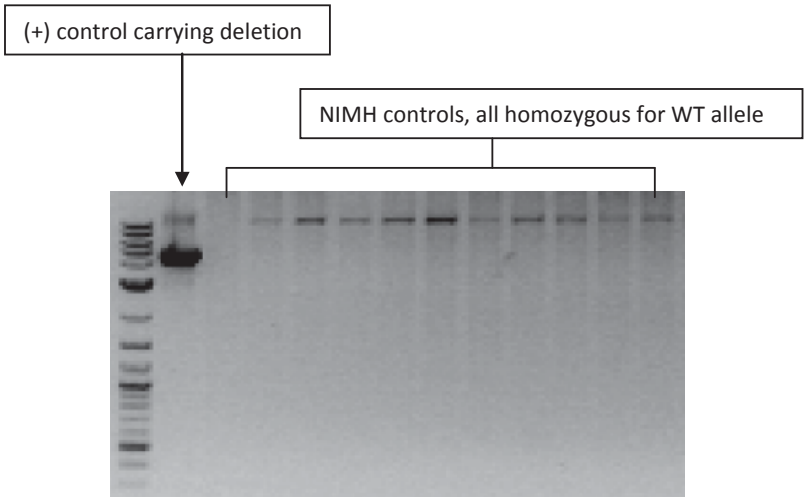
We also targeted three common CNVs, with 20 probes each in the up- and downstream flanks and 30 probes internal to the CNV. All microarray hybridization experiments were performed using a single unaffected male (GM15724 [Coriell]) as reference. The results of the validation experiments are shown in Table 3. Browser snapshot of the array validations for a representative set of predicted calls are shown in Supplementary Figure 7.

PCR validations in control individuals

We tested a subset of deletions for PCR validation; two genes were selected for large-scale screening in control individuals. The results are as follows:

FGF9	**Wt/Wt**	**Wt/Del**	**Del/Del**	**Total samples**
NIMH samples	612	0	0	612

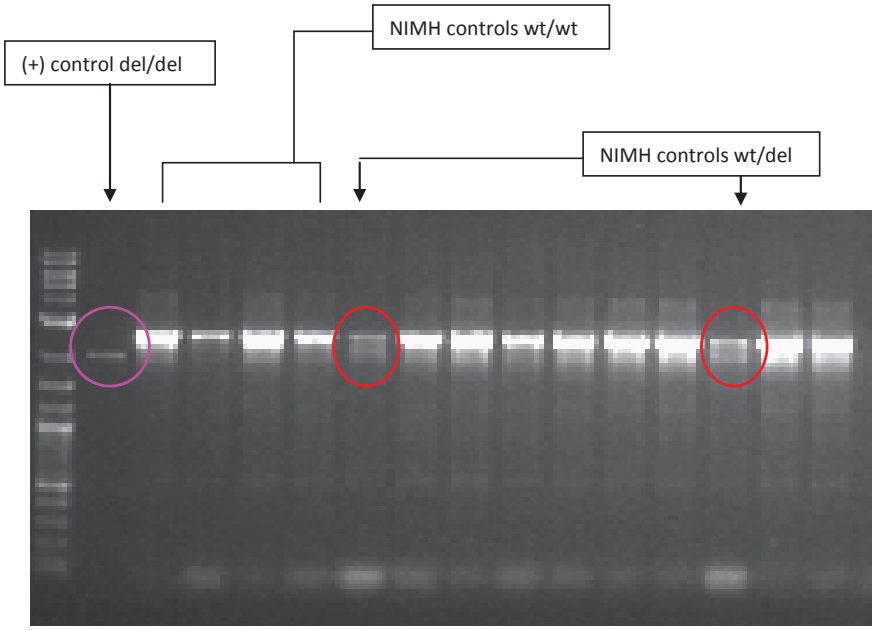
Alleles: Wild type = 1,224**;** Deletion allele = 0



Nature Genetics: doi.10.1038/ng.909

CHS	Wt/Wt	Wt/Del	Del/Del	Total samples
NIMH samples	179	8	0	187

Alleles: Wild type = 380; heterozygous deletion allele = 8; homozygous deletion allele = 0



Comparison with published data from Boone et al.

We compared our CNV morbidity map to the genes identified by Boone and colleagues using exon array CGH. Interestingly, we find that a number of the genes found to harbor exon-overlapping events in affected families show significant enrichment among subsets of the Signature samples that have similar phenotypic classifications. Two genes (*CCREBBP* and *SLC1A1*) were statistically enriched in the Signature set, included partial deletion events, and represented comparable phenotypes to those presented in Boone et al. Additionally, three genes were present within a known syndrome region (*RFC2*, *ALDOA*, and *EXT2*). A further 10 genes were represented in the Signature set with alterations in phenotypes matching those presented in Boone et al. but failed to rise to statistical significance due to rarity of the CNVs (*EP300*, *SCN2A*/*SCN3A*, *STXBP1*, *TCF4*, *TTC8*, *MSH6*, *KIF7*, *FOXP1* [partial deletions]), or rare control CNVs (*NRXN1*, *EYA1*). The remaining eight genes were not supported by our data, including detection of no events for two genes (*TEK* and *JAG1*) and a single control deletion corresponding to part of *PTEN*. Five genes demonstrate partial deletion/gain but were either not significantly enriched (*REEP3*, *KIF1B*, *CNTNAP2* [demonstrates a complex pattern of partial

Nature Genetics: doi.10.1038/ng.909

CNVs in cases and controls]) compared to controls or demonstrated increased partial gene deletions in controls (*CTNNA3* and *FHIT*).

Nature Genetics: doi.10.1038/ng.909

References

1. Coe, B.P. et al. Resolving the resolution of array CGH. *Genomics* **89**, 647-53 (2007).
2. Duker, A.L. et al. Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116 C/D box snoRNA cluster in Prader-Willi syndrome. *Eur J Hum Genet* **18**, 1196-201 (2010).
3. Itsara, A. et al. Population analysis of large copy number variants and their relationship to hotspots of human genetic disease. *American Journal of Human Genetics* (2009).
4. Conrad, D.F. et al. Origins and functional impact of copy number variation in the human genome. *Nature* **464**, 704-12 (2010).
5. 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* (2008).
6. McCarroll, S.A. et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* **40**, 1166-74 (2008).
7. Kidd, J.M. et al. Mapping and sequencing of structural variation from eight human genomes. *Nature* **453**, 56-64 (2008).
8. Kitada, T. et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* **392**, 605-8 (1998).
9. Elia, J. et al. Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol Psychiatry* **15**, 637-46 (2010).
10. Maestrini, E. et al. High-density SNP association study and copy number variation analysis of the AUTS1 and AUTS5 loci implicate the IMMP2L-DOCK4 gene region in autism susceptibility. *Mol Psychiatry* **15**, 954-68 (2010).
11. Pagnamenta, A.T. et al. Characterization of a family with rare deletions in CNTNAP5 and DOCK4 suggests novel risk loci for autism and dyslexia. *Biol Psychiatry* **68**, 320-8 (2010).
12. Petek, E. et al. Disruption of a novel gene (IMMP2L) by a breakpoint in 7q31 associated with Tourette syndrome. *Am J Hum Genet* **68**, 848-58 (2001).
13. Abrahams, B.S. et al. Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc Natl Acad Sci U S A* **104**, 17849-54 (2007).
14. Antonacci, F. et al. A large and complex structural polymorphism at 16p12.1 underlies microdeletion disease risk. *Nat Genet* **42**, 745-50 (2010).
15. Zody, M.C. et al. Evolutionary toggling of the MAPT 17q21.31 inversion region. *Nat Genet* **40**, 1076-83 (2008).
16. Koolen, D.A. et al. A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat Genet* **38**, 999-1001 (2006).
17. Sharp, A.J. et al. Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* **38**, 1038-42 (2006).
18. Shaw-Smith, C. et al. Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nat Genet* **38**, 1032-7 (2006).
19. Helbig, I. et al. 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. *Nat Genet* **41**, 160-2 (2009).
20. Pagnamenta, A.T. et al. A 15q13.3 microdeletion segregating with autism. *Eur J Hum Genet* **17**, 687-92 (2009).

Nature Genetics: doi.10.1038/ng.909

21. Sharp, A.J. et al. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet* **40**, 322-8 (2008).
22. Stefansson, H. et al. Large recurrent microdeletions associated with schizophrenia. *Nature* **455**, 232-6 (2008).
23. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237-41 (2008).
24. Mefford, H.C. et al. A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Res* (2009).
25. Rudd, M.K. et al. Segmental duplications mediate novel, clinically relevant chromosome rearrangements. *Hum Mol Genet* **18**, 2957-62 (2009).
26. Itsara, A. et al. De novo rates and selection of large copy number variation. *Genome Res* **20**, 1469-81 (2010).
27. Sudmant, P.H. et al. Diversity of human copy number variation and multicopy genes. *Science* **330**, 641-6 (2010).
28. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237-41 (2008).
29. Marshall, C.R. et al. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* **82**, 477-88 (2008).
30. Christian, S.L. et al. Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol Psychiatry* **63**, 1111-7 (2008).
31. Koolen, D.A. et al. Clinical and molecular delineation of the 17q21.31 microdeletion syndrome. *J Med Genet* **45**, 710-20 (2008).

Nature Genetics: doi.10.1038/ng.909