

Supplemental Data

Intellectual Disability Is Associated with Increased Runs of Homozygosity in Simplex Autism

Ece D. Gamsiz, Emma W. Viscidi, Abbie M. Frederick, Shailender Nagpal, Stephan J. Sanders, Michael T. Murtha, Michael Schmidt, Simons Simplex Collection Genetics Consortium Elizabeth W. Triche, Daniel H. Geschwind, Matthew W. State, Sorin Istrail, Edwin H. Cook Jr., Bernie Devlin, and Eric M. Morrow

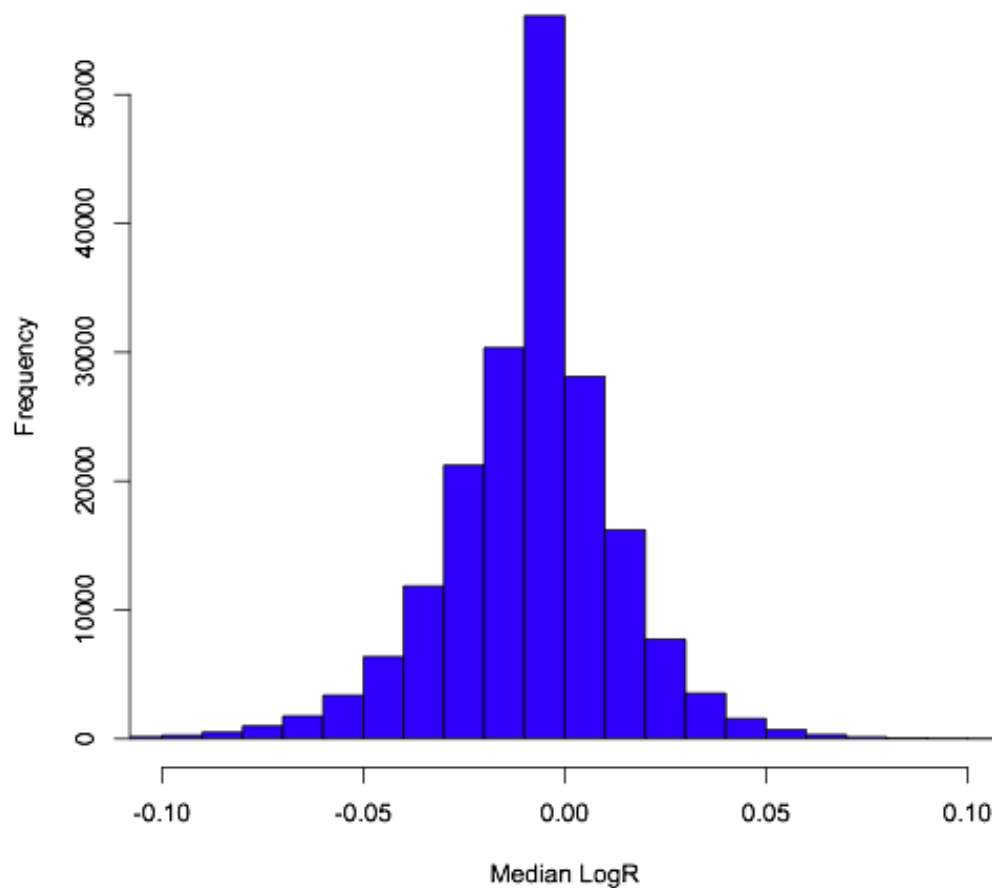


Figure S1. Distribution of LogR Data for the Probes Contained within All Blocks of ROH

Test for the presence of deletions within ROH blocks. LogR data regarding probe intensity for the probes contained within all blocks of ROH for the final segment size of 1000 kb or larger. Plotted are the number of probes (Y-axis) as compared to the LogR values reflecting the probe intensity values.

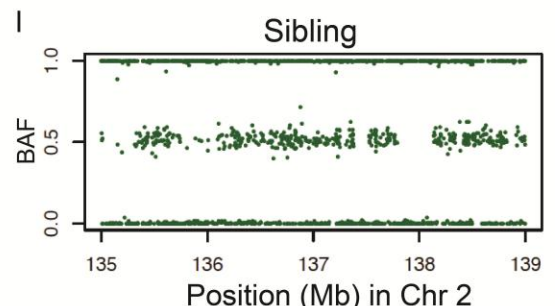
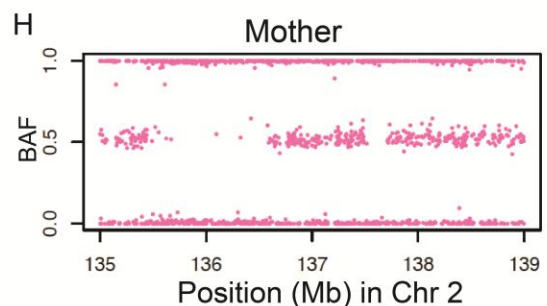
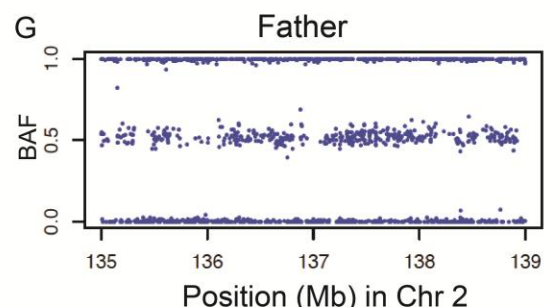
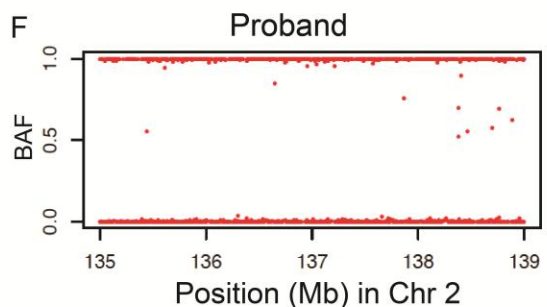
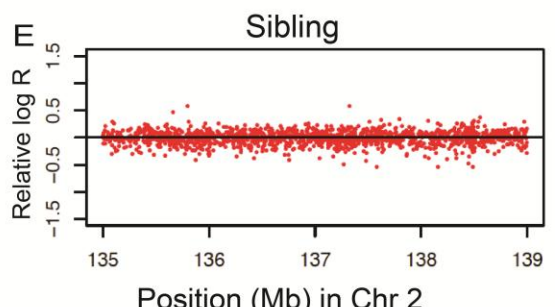
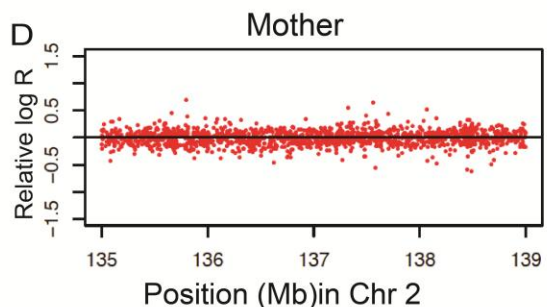
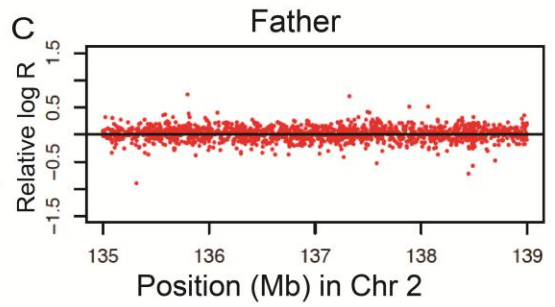
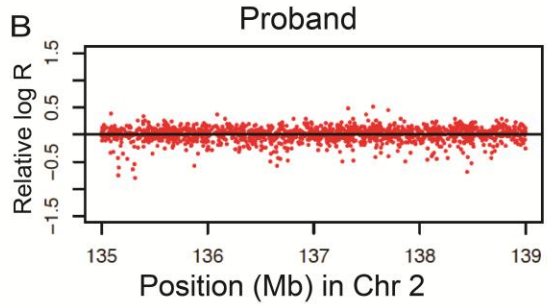
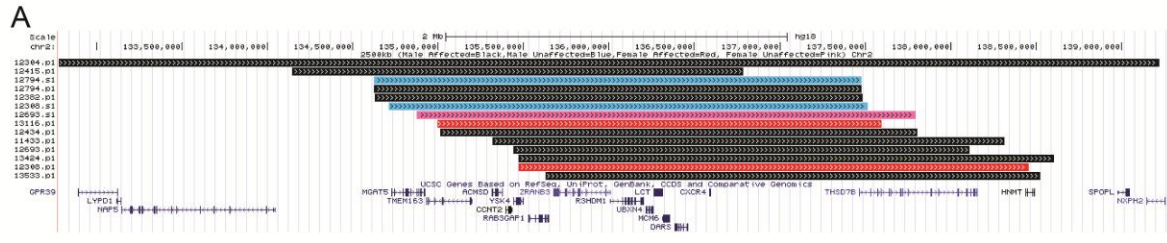


Figure S2. Overlapping ROH Blocks in Chr2: 135,000,000–139,000,000 and Relative LogR and B Allele Frequency for Family 12304

(A) Overlapping ROH blocks in chr2: 135,000,000–139,000,000 (hg18). (Male proband: black, female proband: red, male unaffected sibling: blue and female unaffected sibling: pink.)

(B–E) Relative logR for members of family 12304: (B) proband, (C) father, (D) mother and (E) unaffected sibling indicating no CNVs in the region for the proband.

(F–I) BAF for members of family 12304: (F) proband, (G) father, (H) mother and (I) unaffected sibling showing the homozygous block proband has in chr2: 132,782,514–139,222,178. Other family members are heterozygous in that region.

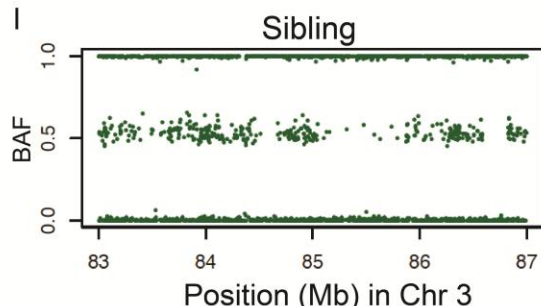
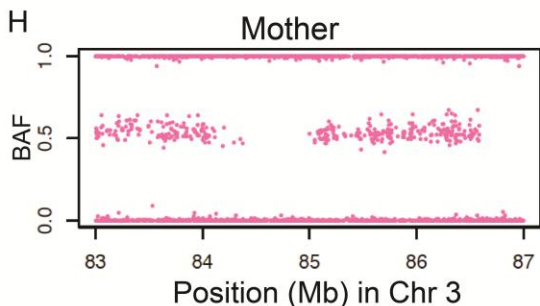
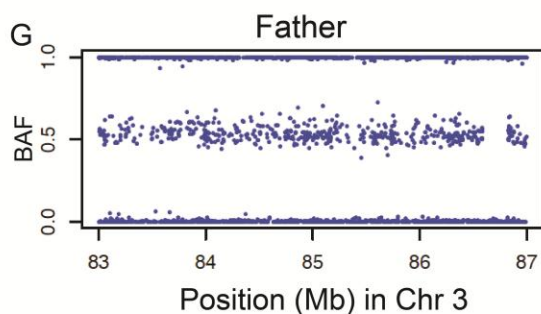
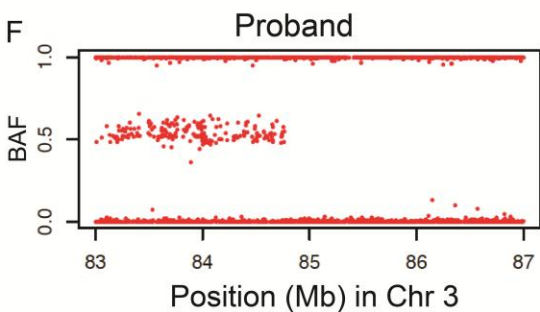
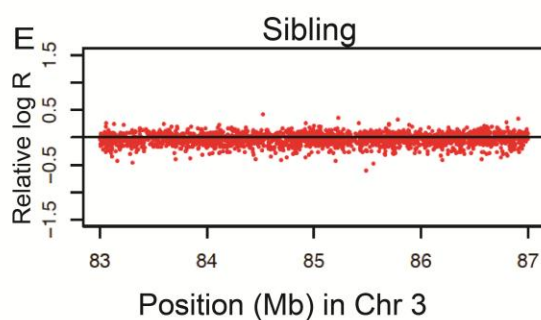
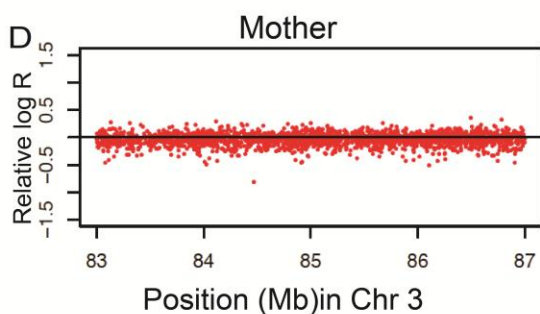
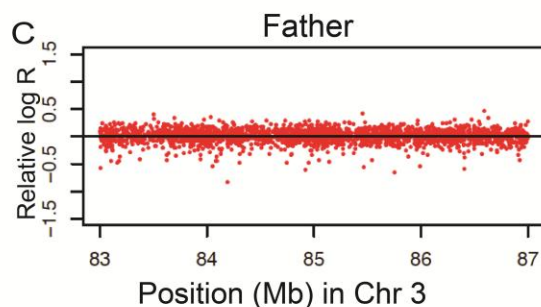
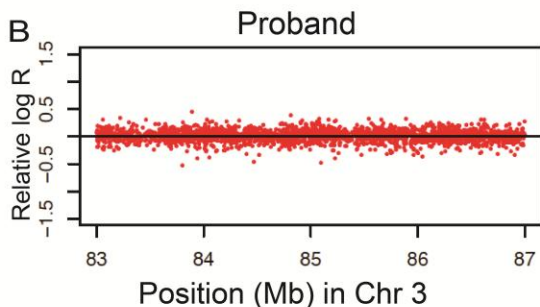
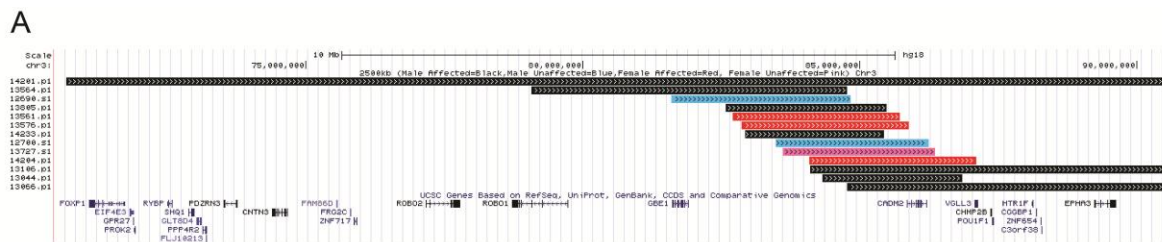


Figure S3. Overlapping ROH Blocks in Chr3: 83,000,000–87,000,000 and Relative LogR and B Allele Frequency for Family 13066

(A) Overlapping ROH blocks in chr3: 83,000,000–87,000,000 (hg18). (Male proband: black, female proband: red, male unaffected sibling: blue and female unaffected sibling: pink).

(B–E) Relative logR for members of family 13066: (B) proband, (C) father, (D) mother and (E) unaffected sibling indicating no CNVs in the region for the proband.

(F–I) BAF for members of family 13066: (F) proband, (G) father, (H) mother and (I) unaffected sibling showing the homozygous block proband has in chr3: 84,769,855–90,525,615. Other family members are heterozygous in that region.

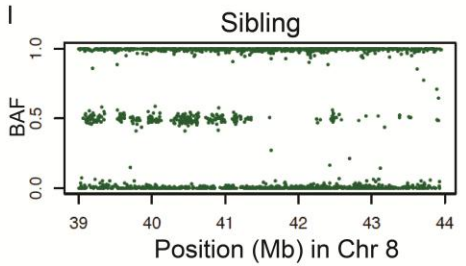
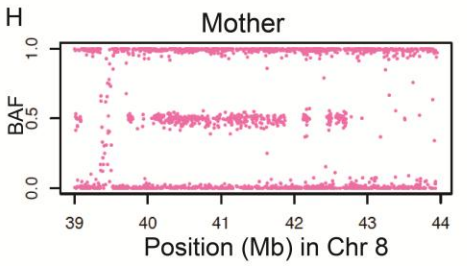
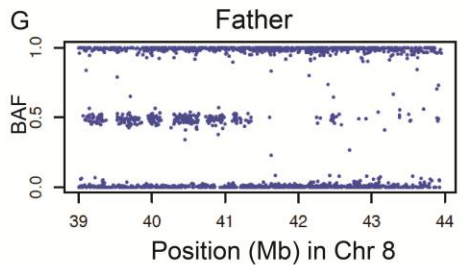
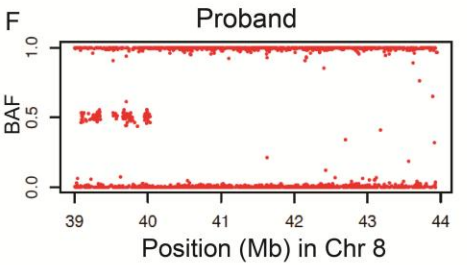
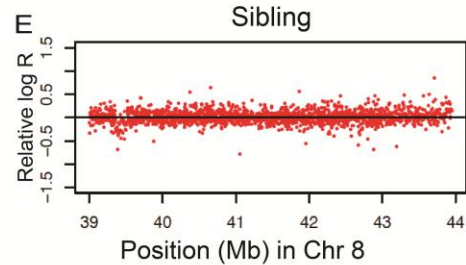
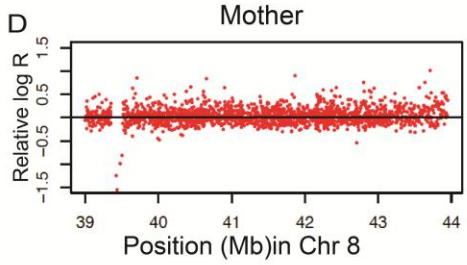
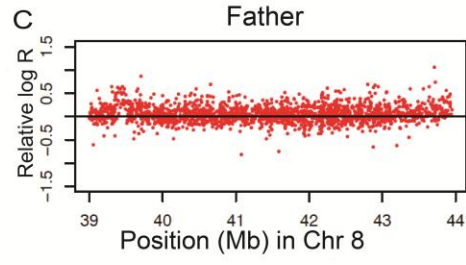
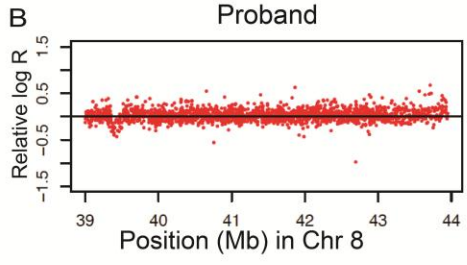
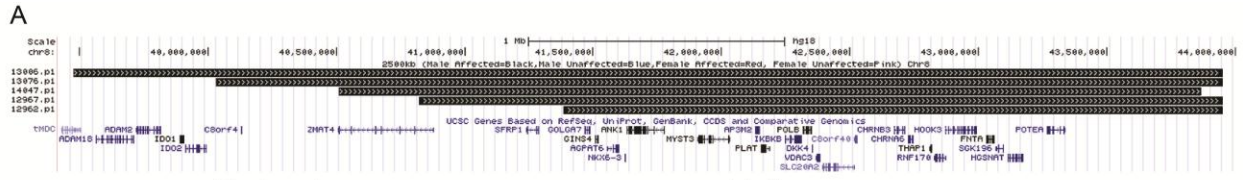


Figure S4. Overlapping ROH Blocks in Chr8: 39,000,000–44,000,000 and Relative LogR and B Allele Frequency for Family 13076

(A) Overlapping ROH blocks in chr8: 39,000,000–44,000,000 (hg18). (Male proband: black, female proband: red, male unaffected sibling: blue and female unaffected sibling: pink).

(B–E) Relative logR for members of family 13076: (B) proband, (C) father, (D) mother and (E) unaffected sibling indicating no CNVs in the homozygous region for the proband.

(F–I) BAF for members of family 13076: (F) proband, (G) father, (H) mother and (I) unaffected sibling showing the homozygous block proband has in chr8: 40,032,290–43,951,038. Other family members are heterozygous in that region.

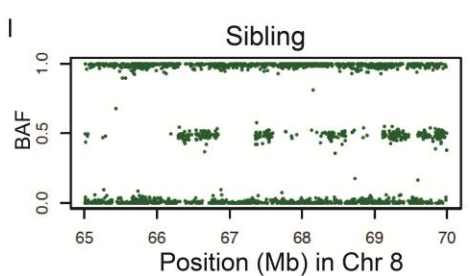
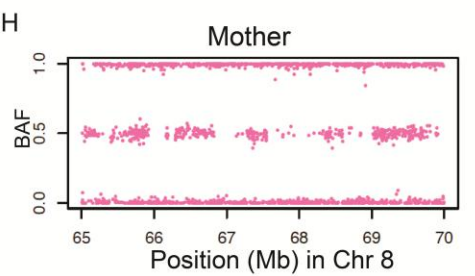
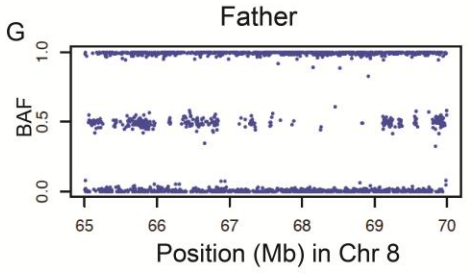
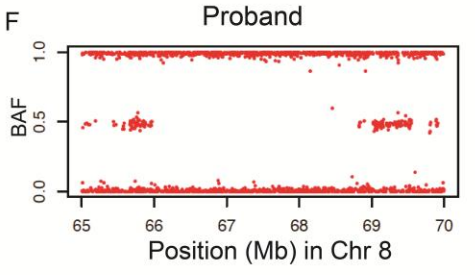
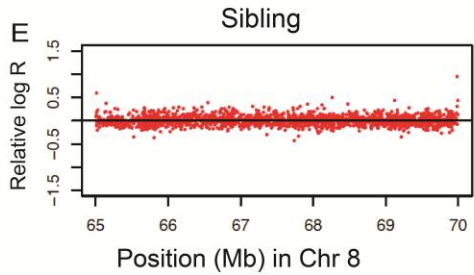
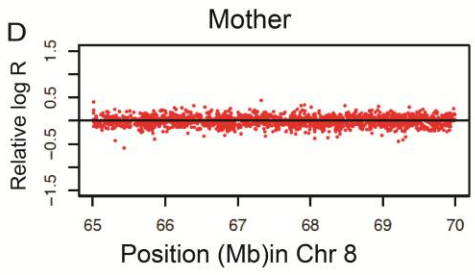
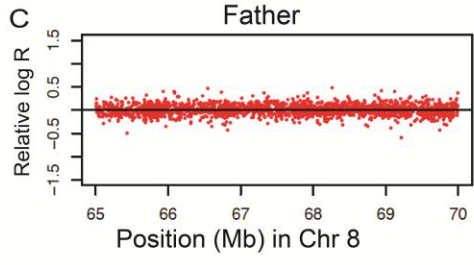
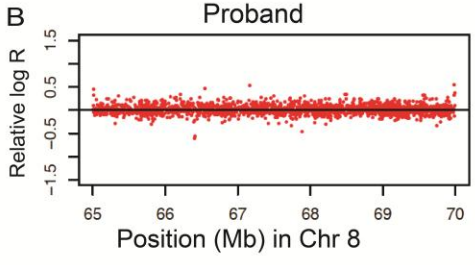
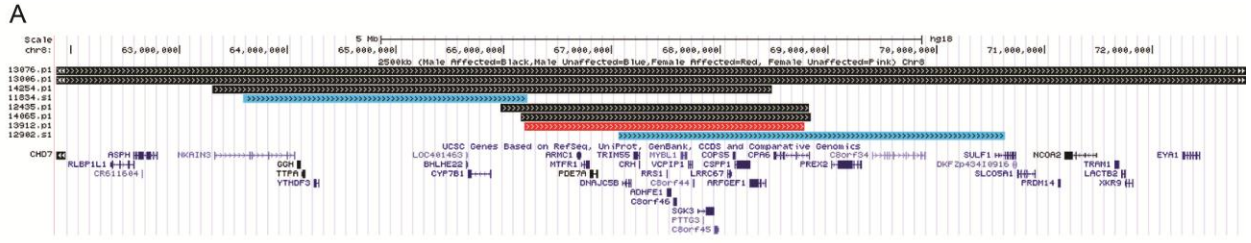


Figure S5. Overlapping ROH Blocks in Chr8: 65,000,000–70,000,000 and Relative LogR and B Allele Frequency for Family 12435

(A) Overlapping ROH blocks in chr8: 65,000,000–70,000,000 (hg18). (Male proband: black, female proband: red, male unaffected sibling: blue and female unaffected sibling: pink).

(B–E) Relative logR for members of family 12435: (B) proband, (C) father, (D) mother and (E) unaffected sibling indicating no CNVs in the region for the proband.

(F–I) BAF for members of family 12435: (F) proband, (G) father, (H) mother and (I) unaffected sibling showing the homozygous block proband has in chr8: 65,969,617–68,825,216. Other family members do not have that homozygous block.

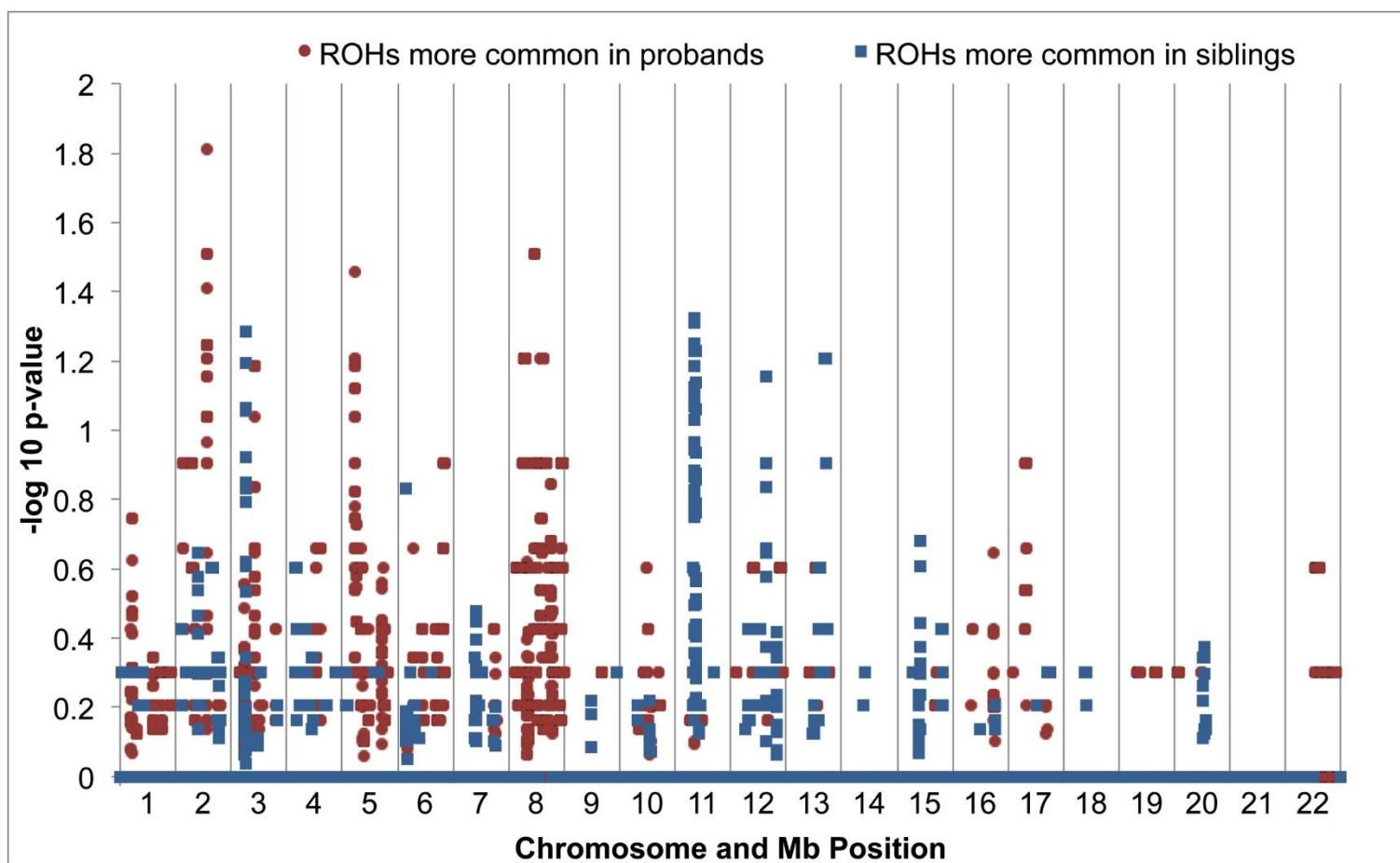
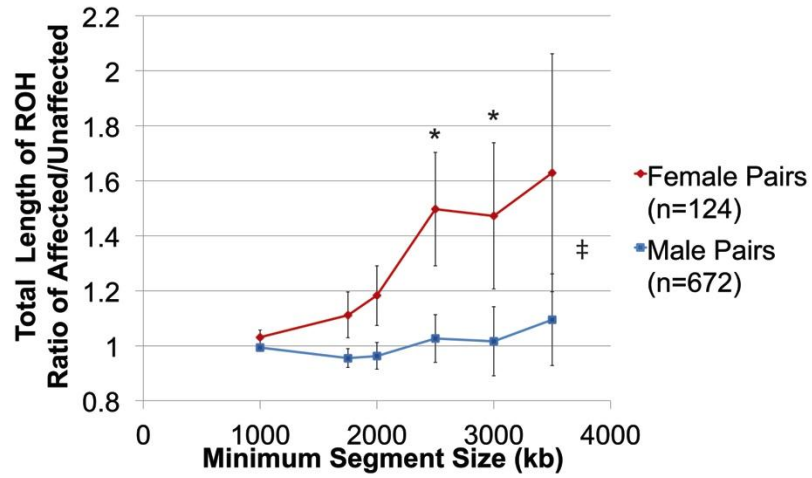


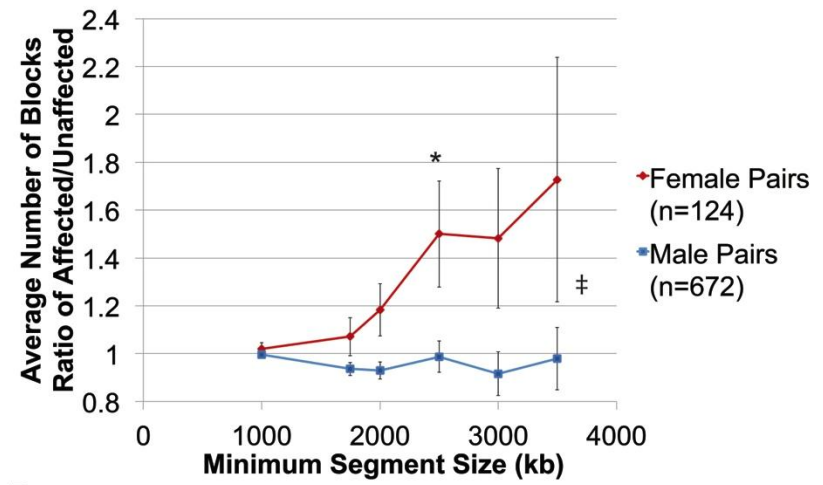
Figure S6. Overlapping ROHs in Probands and Designated Siblings at 50 kb Intervals along the Genome (for Minimum Final Segment Size of 2500 kb)

Plotted are the $-\log_{10} p$ values of overlapping ROH blocks based on the Fisher exact test in R that the distribution of overlap deviates from the null hypothesis of equal overlap between probands and siblings. Given the unusually low SNP density on the Omni Array in the region of 16p11 (31-35 Mbp), data from the Omni Array in this region was not included in the study.

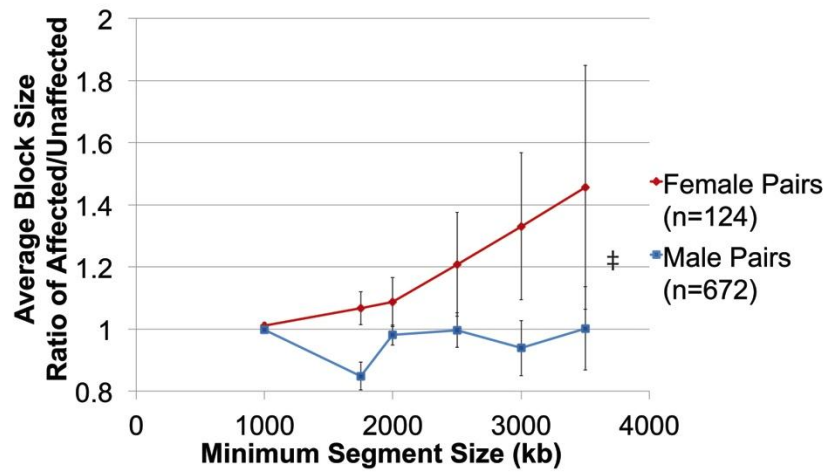
A



B



C



D

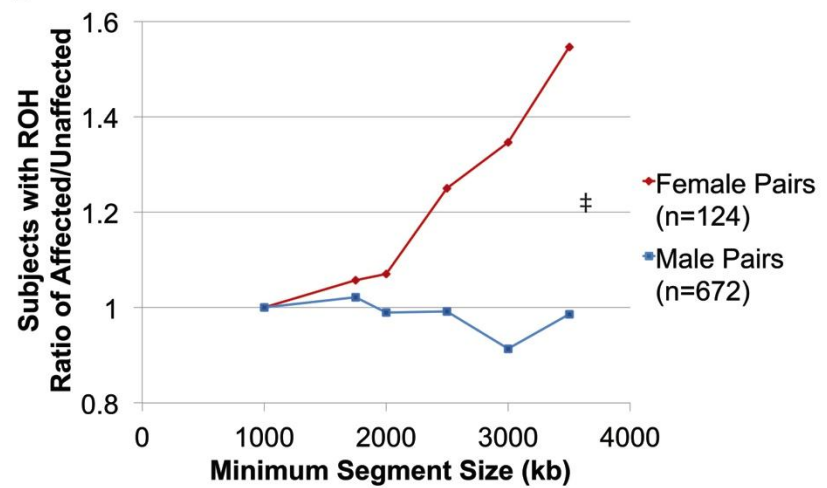


Figure S7. Autosomal ROH Burden Is Increased in Girls with Autism as Compared to Their Same-Sex Discordant Siblings in the SSC Data Set

Autosomal ROH burden is increased in girls with autism (red) as compared to their unaffected female siblings in SSC dataset. This effect is not observed in boys with autism (blue) as compared to their unaffected male siblings.

(A) Total length of ROH. The ratio of total length of ROH per individual for the proband as compared to the designated unaffected sibling is plotted. Each ratio is plotted for different minimum segment sizes used to determine a block of ROH.

(B) Average number of blocks. Ratio of average number of blocks per individual for the proband as compared to the designated unaffected sibling is plotted.

(C) Average block size. Ratio of the average block size per individual for the proband as compared to the designated unaffected sibling is plotted.

(D) Proportion of subjects with at least one block of ROH. Ratio of the proportion of probands with at least one block as compared to the designated unaffected sibling is plotted.

Statistically significant p-values are shown in the graph. Conditional logistic regression analysis in STATA version 11.1 was performed comparing probands to their designated unaffected siblings. Error bars represent standard error of the mean. Ratios were computed in STATA version 11.1 with standard errors. Statistical significance of proportion of subjects with ROH was obtained by chi-square test.

*Represent statistical significance, p-value < 0.05.

‡Represent statistical test if the curves for girls are different from curve for boys. p value <0.05. One-way ANOVA was performed to compare all data points for each curve. The p values for the ANOVAs were 0.0117 for A, 0.009 for B, 0.011 for C, and 0.025 for D. The numbers represented by n reflects the number of discordant sibling pairs contributed to the data.

Further information regarding this association analysis can be found in Table S6.

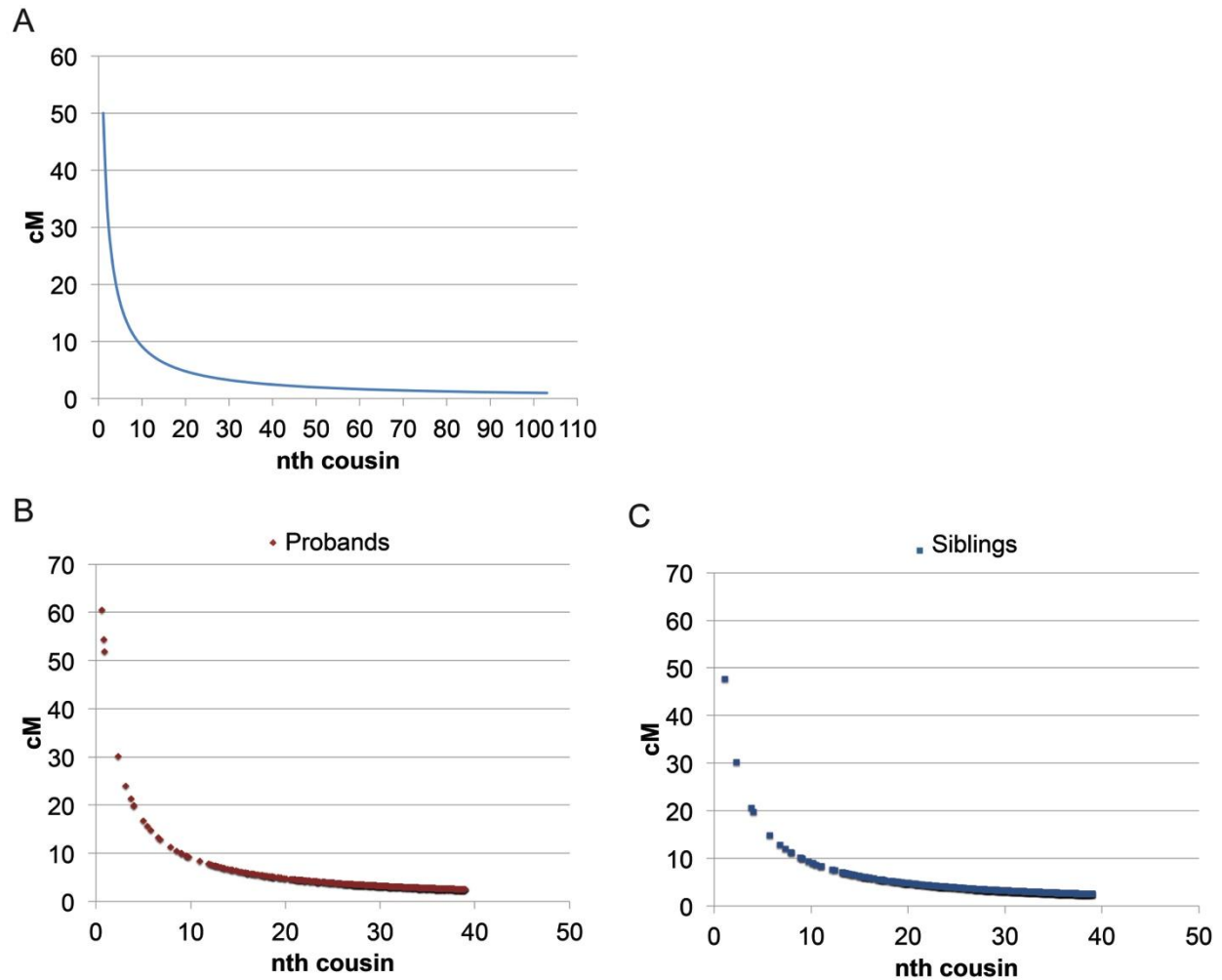


Figure S8. Generational Age of ROH Blocks

Generational age of autosomal ROH blocks of different length was estimated according to the methods described by Kong et al¹⁵.

(A) 1 centiMorgan (cM) in generic length was taken as 1Mb. If n is degree of cousins, average size of homozygous region shared is $200/(2n+2)$.

(B) Generational age of ROH blocks for probands for the minimum segment size of 2500 kb or larger. Each proband was represented by one ROH block. For probands with multiple ROH blocks, the longest block was chosen and the age of block was calculated.

(C) Generational age of ROH blocks for unaffected siblings for the minimum segment size of 2500 kb or larger. Each sibling was represented by one ROH block. For siblings with multiple ROH blocks, the longest block was chosen and the age of block was calculated.

Table S1. An Increase in Burden of Autosomal ROH in Autism Probands with IQ ≤ 70 as Compared to Unaffected Siblings

IQ ≤ 70					IQ >70			
Minimum Segment Size (kb)	Probands± SE (n = 509)	Designated Sibling ± SE (n = 509)	Odds Ratio	P Value	Probands ± SE (n = 1110)	Designated Sibling ± SE (n = 1110)	Odds Ratio	P Value
Average Total Length of ROH (Mb)								
1000	29.35±0.77	28.13±0.67	1.17	0.006	28.61±0.44	28.65±0.43	1.00	0.89
1750	7.14±0.5	6.20±0.35	1.25	0.008	6.12±0.31	6.20±0.30	0.97	0.64
2000	4.95±0.47	4.14±0.31	1.25	0.019	4.03±0.3	4.07±0.29	0.98	0.78
2500	2.89±0.46	2.19±0.27	1.27	0.034	2.11±0.29	2.10±0.28	1.01	0.89
3000	1.93±0.44	1.42±0.26	1.18	0.15	1.38±0.28	1.28±0.26	1.09	0.35
3500	1.45±0.43	0.94±0.24	1.22	0.15	1.1±0.27	0.95±0.26	1.15	0.17
Average Number of Blocks								
1000	20.82±0.44	20.36±0.42	1.03	0.055	20.76±0.24	20.80±0.24	1.00	0.80
1750	2.81±0.11	2.58±0.1	1.11	0.013	2.53±0.06	2.59±0.06	0.96	0.25
2000	1.63±0.08	1.47±0.07	1.15	0.021	1.40±0.05	1.45±0.05	0.95	0.26
2500	0.70±0.05	0.59±0.04	1.23	0.018	0.53±0.03	0.55±0.03	0.95	0.41
3000	0.34±0.04	0.30±0.03	1.15	0.24	0.26±0.03	0.25±0.03	1.04	0.71
3500	0.19±0.03	0.16±0.03	1.27	0.17	0.16±0.02	0.15±0.02	1.17	0.22
Average Block Length per Person (Mb)								
1000	1.39±0.01	1.37±0.008	8.18	0.071	1.36±0.008	1.36±0.009	1.03	0.97
1750	2.07±0.05	1.96±0.05	1.44	0.051	1.93±0.03	1.92±0.03	1.04	0.78
2000	1.97±0.07	1.89±0.07	1.15	0.27	1.71±0.04	1.76±0.04	0.89	0.28
2500	1.40±0.09	1.33±0.08	1.07	0.49	1.13±0.05	1.13±0.05	0.99	0.93
3000	0.89±0.09	0.88±0.08	1.01	0.91	0.72±0.05	0.72±0.05	0.99	0.94
3500	0.56±0.09	0.57±0.08	1.00	0.93	0.52±0.05	0.49±0.05	1.06	0.50
Proportion of Subjects with ROH								
1000	1:1	1:1		1	1:1	1:1		1
1750	0.88:1	0.84:1		0.056	0.86:1	0.85:1		0.67
2000	0.73:1	0.70:1		0.21	0.66:1	0.68:1		0.76
2500	0.41:1	0.39:1		0.57	0.34:1	0.34:1		0.79
3000	0.21:1	0.22:1		0.89	0.17:1	0.18:1		0.74
3500	0.10:1	0.11:1		0.61	0.10:1	0.10:1		0.67

ROH was calculated using PLINK with varying minimum final segments of 1000 to 3500 kb. Average total length of ROH (Mb), average number of blocks, average block length and number of genes in ROH were calculated for probands with IQ ≤ 70 and IQ >70 as well as their unaffected siblings regardless of gender. A conditional logistic regression model in STATA version 11.1 was used to calculate statistical significance for average total length of ROH, average number of blocks and average block length per person. Odds Ratios were calculated for total length of ROH and average block size based on 5 Mb and 2.5 Mb incremental increases respectively. Statistical significance of proportion of subjects with ROH was obtained by chi-square test. Probands with IQ ≤ 70 were observed to have higher average number of blocks and higher average total length of ROH compared to their designated siblings ($p < 0.05$) for final segments between 1000 and 2500 kb. No statistically significant difference was found in the comparison of probands with IQ >70 to their designated siblings in terms of ROH burden ($p > 0.05$).

Table S2. An Increase in Burden of Autosomal ROH in Autism Probands with IQ ≤ 55 and 55 ≤ IQ ≤ 70 as Compared to Unaffected Siblings

IQ ≤ 55					55 ≤ IQ ≤ 70			
Minimum Segment Size (kb)	Probands ± SE (n = 301)	Designated Sibling ± SE (n = 301)	Odds Ratio	P Value	Probands ± SE (n = 208)	Designated Sibling ± SE (n = 208)	Odds Ratio	P Value
Average Total Length of ROH (Mb)								
1000	30.11±1.03	28.90±0.95	1.20	0.012	28.26±1.16	27.10±0.90	1.11	0.18
1750	7.43±0.66	6.49±0.53	1.34	0.009	6.72±0.78	5.78±0.35	1.13	0.27
2000	5.15±0.61	4.57±0.48	1.23	0.076	4.66±0.74	3.52±0.28	1.29	0.12
2500	3.03±0.59	2.55±0.43	1.25	0.11	2.70±0.72	1.67±0.20	1.33	0.15
3000	1.98±0.57	1.78±0.42	1.11	0.43	1.84±0.71	0.89±0.16	1.41	0.18
3500	1.47±0.55	1.30±0.39	1.11	0.47	1.41±0.69	0.44±0.13	1.74	0.10
Average Number of Blocks								
1000	21.29±0.60	20.74±0.58	1.04	0.094	20.16±0.62	19.81±0.62	1.03	0.33
1750	2.93±0.15	2.60±0.14	1.18	0.005	2.62±0.16	2.56±0.15	1.03	0.62
2000	1.70±0.10	1.55±0.10	1.14	0.09	1.52±0.11	1.35±0.10	1.16	0.11
2500	0.74±0.07	0.63±0.06	1.25	0.07	0.64±0.08	0.51±0.06	1.22	0.14
3000	0.36±0.06	0.35±0.05	1.03	0.87	0.32±0.06	0.23±0.04	1.32	0.16
3500	0.20±0.04	0.20±0.04	0.97	0.90	0.18±0.05	0.09±0.03	1.90	0.06
Average Block Length per Person (Mb)								
1000	1.40±0.02	1.37±0.01	33.1	0.026	1.37±0.02	1.36±0.009	1.95	0.56
1750	2.13±0.06	2.02±0.07	1.46	0.12	2.00±0.09	1.88±0.06	1.40	0.23
2000	2.02±0.09	2.00±0.09	1.05	0.78	1.93±0.12	1.73±0.09	1.38	0.12
2500	1.48±0.11	1.43±0.11	1.06	0.67	1.28±0.14	1.19±0.11	1.09	0.58
3000	0.09±0.12	1.00±0.11	0.94	0.67	0.84±0.14	0.74±0.11	1.10	0.52
3500	0.60±0.12	0.73±0.11	0.84	0.24	0.51±0.13	0.33±0.09	1.26	0.24
Proportion of Subjects with ROH								
1000	1:1	1:1		1	1:1	1:1		1
1750	0.91:1	0.84:1		0.01	0.85:1	0.84:1		0.89
2000	0.75:1	0.72:1		0.35	0.71:1	0.67:1		0.40
2500	0.44:1	0.40:1		0.46	0.38:1	0.38:1		1.00
3000	0.22:1	0.23:1		0.85	0.20:1	0.20:1		1.00
3500	0.11:1	0.14:1		0.26	0.09:1	0.07:1		0.59

ROH was calculated using PLINK with varying minimum final segments of 1000-3500 kb. Average total length of ROH (Mb), average number of blocks, average block length and number of genes in ROH were calculated for probands with IQ ≤ 55 and 55 ≤ IQ ≤ 70 as well as their unaffected siblings regardless of gender. A conditional logistic regression model in STATA version 11.1 was used to calculate statistical significance for average total length of ROH, average number of blocks and average block length per person. Odds Ratios were calculated for total length of ROH and average block size based on per 5 Mb and 2.5 Mb incremental increases respectively. Statistical significance of proportion of subjects with ROH was obtained by chi-square test. Probands with IQ ≤ 55 were observed to have higher average total length of ROH compared to their designated siblings ($p < 0.05$) for final segments between 1000 and 1750 kb and higher average number of blocks for final segment of 1750 kb. No statistically significant difference was found in the comparison of probands with 55 ≤ IQ ≤ 70 to their designated siblings although there is an increase in the burden of ROH ($p > 0.05$).

Table S3. List of Genes within the Loci Recurrent in Autism with p Value Less Than 0.2 Shown in Figure S3

Chr #	Start (hg18)	End (hg18)	Interval Length (kb)	Number of Proband	Number of Siblings	Genes	Proband ID (Full-Scale IQ)
1	52650000	52900000	250	7	2	<i>PRPF38A, ZCCHC11, GPX7, FAM159A</i>	11601.p1(91), 12331.p1(115), 12715.p1(105), 13093.p1(167), 13096.p1(107) , 13101.p1(94), 13188.p1(78)
2	30450000	31100000	650	4	0	<i>LCLAT1, CAPN13, GALNT14</i>	12113.p1(66), 12356.p1(113), 12597.p1(70), 12956.p1(84)
2	73000000	74050000	1050	4	0	<i>EMX1, SFXN5, RAB11FIP5, SMYD5, PRADC1, CCT7, FBXO41, EGR4, ALMS1, NAT8, ALMS1P, NAT8B, TPRKB, DUSP11, C2orf78, STAMBP, ACTG2, DGUOK</i>	13741.p1(47) , 13781.p1(102), 13948.p1(44), 14227.p1(47)
2	135450000	138500000	3050	11	3	<i>MAP3K19, RAB3GAP1, ZRANB3, R3HDM1, UBXN4, LCT, MCM6, DARS, CXCR4, THSD7B, HNMT</i>	11433.p1(78) , 12304.p1(83) , 12308.p1(105) , 12382.p1(91), 12415.p1(94) , 12434.p1(117), 12693.p1(94), 12794.p1(96) , 13116.p1(35) , 13424.p1(114) , 13533.p1(47)
3	84650000	85450000	800	10	3	<i>CADM2</i>	14233.p1(13), 14204.p1(61), 14201.p1(43), 13805.p1(93), 13576.p1(40) , 13561.p1(51), 13564.p1(117) , 13106.p1(100), 13066.p1(120), 13044.p1(112)
5	41400000	41650000	250	5	0	<i>PLCXD3</i>	12582.p1(57) , 12642.p1(99) , 13750.p1(126), 14084.p1(86), 14484.p1(72),
5	42450000	44100000	1650	17	8	<i>GHR, CCDC152, SEPP1, ANXA2R, ZNF131, HMGCS1, CCL28, C5orf28, C5orf34, PAIP1, NNT</i>	12044.p1(72), 12073.p1(110) , 12086.p1(118) , 12240.p1(91), 12339.p1(43), 12340.p1(26) , 12582.p1(57) , 12603.p1(23) , 12642.p1(99) , 12689.p1(116) , 13023.p1(115), 13416.p1(79) , 13750.p1(126), 13830.p1(7), 13882.p1(74), 14084.p1(86), 14484.p1(72)

5	44250000	44450000	200	14	7	<i>FGF10</i>	12044.p1(72), 12073.p1(110) , 12086.p1(118), 12240.p1(91), 12339.p1(43), 12340.p1(26) , 12582.p1(57) , 12603.p1(23) , 12642.p1(99) , 12689.p1(116) , 13023.p1(115), 13416.p1(79) , 13830.p1(7), 13882.p1(74)
6	142250000	142900000	650	4	0	<i>NMBR, VTA1, GPR126</i>	13166.p1(46) , 12875.p1(64) , 13116.p1(35) , 12219.p1(121)
8	35050000	35450000	400	4	0	<i>UNC5D</i>	12289.p1(101), 13006.p1(56), 13083.p1(52) , 14334.p1(58)
8	40800000	44000000	3200	5	0	<i>ZMAT4, SFRP1, GOLGA7, GINS4, AGPAT6, ANK1, KAT6A, AP3M2, PLAT, IKBKB, POLB, DKK4, VDAC1P5, SLC20A2, SMIM19, CHRNB3, CHRNA6, THAP1, RNF170, HOOK3, FNTA, HGSNAT, POTEA</i>	12962.p1(79) , 12967.p1(88), 13006.p1(56), 13076.p1(89) , 14047.p1(91)
8	66150000	68500000	2350	6	1	<i>ARMC1, MTFR1, PDE7A, DNAJC5B, TRIM55, CRH, RRS1, ADHFE1, C8orf46, MYBL1, VCPIP1, C8orf44, SGK3, PTTG3P, MCMD2, SNHG6, SNORD87, PPP1R42, COPS5, CSPP1, ARFGEF1, CPA6</i>	12435.p1(47) , 13006.p1(56), 13076.p1(89) , 13912.p1(95) , 14065.p1(27), 14254.p1(17)
8	80900000	81100000	200	4	0	<i>MRPS28, TPD52</i>	12655.p1(48) , 13006.p1(56), 13076.p1(89) , 13741.p1(47)
8	84050000	85250000	1200	7	2		12463.p1(84) , 12655.p1(48), 12962.p1(79) , 13006.p1(56), 13076.p1(89) , 13575.p1(103) , 13741.p1(47)
8	87350000	87700000	350	7	2	<i>WWP1, RMDN1, CPNE3, CNGB3</i>	11834.p1(76) , 12655.p1(48) , 12962.p1(79) , 13006.p1(56), 13076.p1(89) , 13741.p1(47) , 14241.p1(24)
8	90450000	90900000	450	5	0	<i>RIPK2</i>	11834.p1(76) , 12655.p1(48) , 13006.p1(56), 13076.p1(89) , 13741.p1(47)

8	94000000	99450000	5450	4	0	<i>TRIQQ, C8orf87, FAM92A1, RBM12B, TMEM67, PDP1, CDH17, GMNN, RAD54B, FSBP, KIAA1429, ESRP1, DPY19L4, INTS8, CCNE2, NDUFAF6, TP53INP1, PLEKHF2, C8orf37, GDF6, UQCRB, MTERFD1, PTDSS1, SDC2, CPQ, TSPYL5, MTDH, LAPT4B, MATN2, RPL30, SNORA72, C8orf47, HRSP12, BVES, NIPAL2</i>	12655.p1(48) , 13006.p1(56), 13076.p1(89) , 13741.p1(47)
8	112300000	112450000	150	12	5		11491.p1(53), 11573.p1(68), 11765.p1(34), 12368.p1(47) , 12655.p1(48) , 13006.p1(56), 13277.p1(101), 13582.p1(83) , 13587.p1(78) , 13682.p1(71) , 13706.p1(67) , 14108.p1(115)
8	139900000	144000000	4100	4	0	<i>COL22A1, KCNK9, TRAPPC9, CHRAC1, AGO2, PTK2, DENND3, SLC45A4, GPR20, PTP4A3, MROH5, TSNARE1, BAI1, ARC, JRK, PSCA, LY6K, THEM6, SLURP1, LYPD2, LYNX1, LY6D, GML, CYP11B1, CYP11B2</i>	12170.p1(80), 12655.p1(48) , 13006.p1(56), 13076.p1(89)
17	24200000	24400000	200	6	1	<i>ERAL1, FLOT2, DHRS13, PHF12, SEZ6, PIPOX</i>	12228.p1(108) , 12597.p1(70), 13005.p1(117) , 13393.p1(77), 13774.p1(33) , 14249.p1(78)
17	26100000	26700000	600	6	1	<i>SUZ12P1, CRLF3, ATAD5, TEFM, ADAP2, RNF135, DPRXP4, NF1, OMG, EVI2B, EVI2A</i>	12228.p1(108) , 12597.p1(70), 13005.p1(117) , 13393.p1(77), 13774.p1(33) , 14249.p1(78)

Genome was divided into 50 kb segments and in each segment number of probands and siblings with ROH (minimum segment > 2500 kb) were calculated. P-value was calculated via Fisher exact test in R. Regions with p-value less than 0.2 were shown. The Proband ID column lists all SSC probands that had a block of ROH across the indicated region. Bolding signifies the subjects for whom we had whole exome sequencing data available in the publically available exome sequencing dataset^{18;19;20}. The participants full scale IQ is shown in brackets.

Table S4. List of Variants Found in Whole-Exome Data Analysis under the ROH Regions More Common in Proband Compared to Siblings

Gene	Subject ID	Position	Build	REF/ALT	AA change	Prediction (Polyphen or EVS)	Existing SNP ID (Y/N)	MAF% (EA/AA/AAL;European/American/African American/All)	SNP ID	Experimental Validation (Validated=V, Not Validated=NV, Not Tested=NT)	Proband-VCF data (GT:GQ:DP)	Mother-VCF data (GT:GQ:DP)	Father-VCF data (GT:GQ:DP)	Designated Sibling-VCF data (GT:GQ:DP)	Full IQ	Gender
ADAM8	13076.p1	chr10:135085088	hg19	C/T	V/M	missense, probably damaging	Y	0.49	rs36054052	V	1/1:84:38	0/1:109:21		1/1:60:26	89	male
ADH1C	13039.p1	chr4:100268190	hg19	A/C	*/G	stop-lost	Y	1.1636/0.2043/0.8385	rs283413	NT	1/1:184:102	1/1:182:78	1/1:188:92	1/1:180:99	121	male
ADH1C	13418.p1	chr4:100268190	hg19	A/C	*/G	stop-lost	Y	1.1636/0.2043/0.8385	rs283413	NT	1/1:209:106	1/1:194:107	1/1:198:110	1/1:189:135	135	male
BMP3	13076.p1	chr4:81967188	hg19	C/T	F/L	missense, possibly damaging	Y	0.0/0.3177/0.1076	rs6831040	V	1/1:220:87	1/1:114:44	1/1:60:23	1/1:141:55	89	male
BMP5	13175.p1	chr6:55620456	hg19	C/T	V/I	missense, probably damaging	N	0		NT	1/1:225:91	0/1:228:99	0/1:228:106	0/1:228:109	63	male
C1orf27	13076.p1	chr1:186360840	hg19	C/T	R/C	missense, probably damaging	N	0		NT	1/1:225:159	0/1:228:91	0/1:228:34		89	male
C2orf47	13618.p1	chr2:200532750	hg18	C/T	T/I	missense, possibly damaging	Y	0.0465/0.0227/0.0384	rs150048783	NT	1/1:225:96	0/1:122:132	0/1:81:116	0/1:152:116	44	female
C5orf34	12689.p1	chr5:43509348	hg19	T/A	S/T	missense, possibly damaging	Y	0.0/1.5887/0.5382	rs6872851	V	1/1:81:30	1/1:105:39	1/1:87:31	1/1:96:35	116	female
C5orf34	12073.p1	chr5:43509348	hg19	T/A	S/T	missense, possibly damaging	Y	0.0/1.5887/0.5382	rs6872851	V	1/1:225:94	1/1:225:122	1/1:225:116		110	female
C5orf34	12603.p1	chr5:43509348	hg19	T/A	S/T	missense, possibly damaging	Y	0.0/1.5887/0.5382	rs6872851	V	1/1:225:104	1/1:120:44	1/1:214:76		23	male
C5orf34	12582.p1	chr5:43509348	hg19	T/A	S/T	missense, possibly damaging	Y	0.0/1.5887/0.5382	rs6872851	V	1/1:108:42	1/1:126:46	1/1:120:44	1/1:138:52	57	male
C5orf34	12642.p1	chr5:43509348	hg19	T/A	S/T	missense, possibly damaging	Y	0.0/1.5887/0.5382	rs6872851	V	1/1:78:29	1/1:72:31	1/1:81:31	1/1:102:40	99	male
C5orf34	12340.p1	chr5:43545105	hg18	T/A	S/T	missense, possibly damaging	Y	0.0/1.5887/0.5382	rs6872851	NT	1/1:120:43	1/1:225:86	1/1:157:57	1/1:225:87	26	female
C5orf34*	12689.p1	chr5:43509463	hg19			C insertion in 5'-UTR	N	0		V					116	female
C5orf34*	12073.p1	chr5:43509463	hg19			C insertion in 5'-UTR	N	0		V					110	female
C5orf34*	12603.p1	chr5:43509463	hg19			C insertion in 5'-UTR	N	0		V					23	male
C5orf34*	12582.p1	chr5:43509463	hg19			C insertion in 5'-UTR	N	0		V					57	male
C5orf34*	12642.p1	chr5:43509463	hg19			C insertion in 5'-UTR	N	0		V					99	male
CENPC1	13076.p1	chr4:68378287	hg19	C/A	V/G	missense-near-splice, probably-damaging	Y	0.0247/1.2291/0.3941	rs141379239	V	1/1:225:277	0/1:193:159	0/1:91:86	1/1:225:227	89	male
COL7A1	12926.p1	chr3:48608103	hg19	G/C	P/R	missense, possibly damaging	Y	0.46	rs185142403	NT	1/1:51:21	0/1:39:15	0/1:65:13			male
CSMD3	13582.p1	chr8:113299353	hg19	A/G	S/P	missense, probably damaging	Y	0.0581/0.1816/0.1	rs145027071	NT	1/1:225:83	1/1:225:78	0/1:163:82	1/1:137:73	83	male
CYP39A1	13175.p1	chr6:46604145	hg19	G/C	S/C	missense, possibly damaging	Y	0.6059/0.0454/0.4159	rs147866724	NT	1/1:181:72	0/1:228:81	0/1:228:91		63	male

DNAH7	11055.p1	chr2:196651764	hg19	C/A	Q/H	missense, probably damaging	N	0		NT	1/1:225:86	0/1:228:127	0/1:228:105	28	male	
DYNC2H1	13741.p1	chr11:103104836	hg19	G/T	A/S	missense, probably damaging	N	0		V	1/1:225:112	0/1:215:103	0/1:206:146	47	male	
FAM63A	13135.p1	chr1:150972959	hg19	A/T	Y/N	missense, probably damaging	Y	1.0698/0.2043/0.7766	rs140386498	NT	1/1:120:41	0/1:146:48	0/1:180:73	1/1:135:48	80	male
FBF1	13076.p1	chr4:74282070	hg19	C/T	A/V	missense, possibly damaging	N	0		NT	1/1:225:88	0/1:179:54	0/1:124:25	1/1:225:84	89	male
FNIP1	12224.p1	chr5:131094624	hg18	C/A	G/C	missense, unknown	Y	0.4651/0.0454/0.3229	rs7730228	NT	1/1:96:34	1/1:120:43	1/1:51:20	1/1:105:42	80	male
FNIP1	13660.p1	chr5:131094624	hg18	C/A	G/C	missense, unknown	Y	0.4651/0.0454/0.3229	rs7730228	NT	1/1:36:13	1/1:78:28	1/1:69:24	1/1:132:49	53	male
FNIP1	13876.p1	chr5:131094624	hg18	C/A	G/C	missense, unknown	Y	0.4651/0.0454/0.3229	rs7730228	NT	1/1:225:146	1/1:225:154	1/1:225:146	1/1:114:40	40	male
GML	13076.p1	chr8:143922642	hg19	G/A	none	splice-5	Y	0.0116/2.9278/0.9995	rs114099899	V	1/1:225:235	0/1:228:124	0/1:228:66	0/1:228:185	89	male
GSTA2	13175.p1	chr6:52621101	hg19	C/T	E/K	missense, possibly damaging	Y	0.0233/1.4539/0.5077	rs75013911	NT	1/1:225:154	0/1:137:184	0/1:132:166	0/1:143:190	63	male
HTN3	13076.p1	chr4:70898922	hg19	T/A	Y/*	stop-gained	Y	0.0465/14.5483/4.9592	rs17147990	V	1/1:225:302	0/1:196:159	0/1:183:93	1/1:225:242	89	male
ITGAX	12198.p1	chr16:31382999	hg19	G/A	R/L	missense, probably damaging	Y	0.0465/0.0/0.0308	rs146647978	NT	1/1:160:61	0/1:194:34	0/1:180:30		117	male
KIAA0146	13076.p1	chr8:48641518	hg19	G/C	C/S	missense, probably damaging	N	0		NT	1/1:225:146	0/1:195:86	0/1:102:47	0/1:215:139	89	male
LEFTY1	13076.p1	chr1:226076583	hg19	C/T	M/V	missense, possibly-damaging	Y	0.0116/0.4766/0.1692	rs61995951	NT	1/1:145:57	0/1:170:21	1/1:39:15	1/1:111:42	89	male
LMBRD1	13076.p1	chr6:70500252	hg19	G/A	S/L	missense, probably damaging	N	0		NT	1/1:225:255	0/1:210:124	0/1:124:53	1/1:225:223	89	male
METTL8	12304.p1	chr2:172180771	hg19	A/G	*R	stop-lost	Y	0.0/0.7948/0.2409	rs10205459	NT	1/1:225:322	1/1:225:221	1/1:225:352		83	male
MKI67	13076.p1	chr10:129904253	hg19	G/T	I/L	missense, possibly-damaging	Y	0.0/2.1108/0.7151	rs34116632	NT	1/1:225:412	0/1:228:211	0/1:228:93	1/1:225:291	89	male
MROH5	12655.p1	chr8:142459814	hg19	C/T	E/G	missense, possibly damaging	Y	0.182	rs7003306	NT	1/1:24:11	1/1:18:10	1/1:60:28	1/1:48:19	48	male
MROH5	13076.p1	chr8:142459814	hg19	C/T	E/G	missense, possibly damaging	Y	0.18	rs7003306	NT	1/1:96:41	1/1:30:14	1/1:33:13	1/1:33:13	89	male
NBN	12655.p1	chr8:90967686	hg19	T/C	E/K	missense, possibly damaging	Y	0.0/0.8171/0.2768	rs34120922	V	1/1:102:37	0/1:43:27	0/1:89:49		48	male
NECAB1	13741.p1	chr8:91953077	hg19	G/T	A/S	missense, possibly-damaging	Y	0.012/1.1128/0.3707	rs115555424	V	1/1:205:71	0/1:228:60	0/1:228:89		47	male
OSBPL9	13188.p1	chr1:52231560	hg19	C/T	T/I	missense, possibly damaging	Y	0.9419/0.1135/0.6612	rs61739207	NT	1/1:225:166	0/1:228:65	0/1:228:159		78	male
PAMR1	13076.p1	chr11:35492221	hg19	G/A	L/F	missense, probably damaging	N	0		V	1/1:225:134	0/1:212:77	0/1:144:41		89	male
PCNXL2	13076.p1	chr1:233136061	hg19	T/C	R/K	missense, probably-damaging	Y	0.0/0.9981/0.3288	rs116467047	V	1/1:175:81	0/1:64:33	0/1:31:15	0/1:109:56	89	male
PKHD1	13175.p1	chr6:51512889	hg19	G/A	P/S	missense, probably damaging	Y	0.1744/0.0454/0.1307	rs41273722	NT	1/1:225:112	0/1:140:105	0/1:151:109	0/1:176:121	63	male

PKHD1	13175.p1	chr6:51512889	hg19	G/A	P/S	missense, probably damaging	Y	0.1744/0.0454/0.1307	rs41273722	NT	1/1:225:112	0/1:140:105	0/1:151:109	0/1:176:121	63	male
PROB1	13076.p1	chr5:138729609	hg19	C/T	V/M	missense, possibly damaging	N	0		V	1/1:102:40	0/1:42:22	0/1:83:12	0/1:196:28	89	male
PRSS48	12655.p1	chr4:152212475	hg19	T/C	P/L	missense, probably damaging	Y	0.012/0.0253/0.0163	rs117303916	V	1/1:225:82	0/1:147:58	0/1:167:101	1/1:225:111	48	male
SMTN	13612.p1	chr22:31487682	hg19	C/T	V/A	missense, probably damaging	Y	0.0233/3.7222/1.2763	rs11913728	NT	1/1:39:18	0/1:123:20	0/1:63:12	0/1:94:15	52	male
SMTN	13612.p1	chr22:31479234	hg19	G/A	E/G	missense	Y	0.0116/4.8409/1.6479	rs11913760	NT	1/1:105:41	0/1:91:34	0/1:116:27	0/1:134:38	52	male
TET1	12655.p1	chr10:70333081	hg19	T/G	C/F	missense, possibly damaging	Y	0.0349/0.0/0.0231	rs201095472	NT	1/1:54:22	0/1:43:14	0/1:50:22		48	male
TRAF4	13393.p1	chr17:24099461	hg18	G/A	A/T	missense, probably damaging	Y	0.7791/0.2043/0.5843	rs35932778	NT	1/1:225:134	0/1:228:94	0/1:228:82	0/1:228:80	77	female
TRAPPC9*	12655.p1	chr8:140742929	hg19	C/T		3'-UTR	Y	2.9	rs79408026	V					48	male
UGT2A1	13076.p1	chr4:70512825	hg19	C/T	E/K	missense, probably damaging	N	0		NT	1/1:138:149	0/1:67:65		1/1:145:88	89	male
UGT2B10	13076.p1	chr4:69687987	hg19	C/A	none	splice-3	Y	0.3023/37.1708/12.78 84	rs2942857	V	1/1:225:793	1/1:225:355	1/1:225:218	1/1:225:589	89	male
ZC3H3	13076.p1	chr8:144620293	hg19	G/A	L/P	missense, possibly- damaging	Y	0.0582/1.8157/0.6536	rs36008851	NT	1/1:220:93	0/1:111:44	0/1:59:18	0/1:111:64	89	male
ZNF107	13808.p1	chr7:63806454	hg18	G/A	none	frameshift	N	0		NT	1/1:172:65	1/1:24:11	1/1:181:64	1/1:111:47	72	male
ZNF138	13808.p2	chr7:63928768	hg18	T/A	F/Y	missense, unknown	Y	0.0314/4.9133/1.5112		NT	1/1:225:347	1/1:193:93	1/1:225:322	1/1:225:290	72	male
ZNF785	12198.p1	chr16:30594521	hg19	A/C	V/G	missense, probably damaging	Y	1.186/0.1593/0.8388	rs35215913	NT	1/1:24:9	0/1:45:12	0/1:9:13		117	male

Whole exome sequencing data available for the SSC dataset was analyzed. Regions more common in probands compared to siblings were chosen. A genome-wide analysis was performed to find number of probands and siblings with ROH in 1 Mb regions. In each region, if number of probands with ROH is higher compared to number of siblings, that region was investigated for variants. Polyphen 2 and exome variant server were used to filter the variants in those regions. Identified variants were experimentally validated depending on the availability of DNA sample. Twenty-four variants were confirmed by Sanger sequencing and all SNVs were deemed to be true call and we thereby estimate call efficacy in the range of 100%.

*Represent variants found by experimental validation only, not in the whole exome data analysis.

Table S5. Phenotype-Genotype Analysis of Female and Male Probands with and without ROH for Cognition, Adaptive Functioning, ADOS, and ADI-R

	Entire Sample (n = 2066)	Females with ROH (n = 102)	Females without ROH (n = 175)	Males with ROH (n = 648)	Males without ROH (n = 1141)	p Values	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Females ROH to Females No ROH	Males ROH to Males No ROH
Cognition							
Full-scale IQ	81.9 (27.7)	74.1(31.2)	76(25.7)	80.2(28.6)	84.4(26.9)	0.57	0.002
Verbal IQ	78.9 (30.9)	73.2(34.7)	75.2(30)	76.8(31.8)	81.1(30)	0.59	0.004
Non-verbal IQ	85.2 (25.9)	76.4(28.7)	78.6(24)	83.8(26.5)	87.8(25.2)	0.49	0.001
Verbal mental age	90.0 (47.7)	75.9(51.1)	76.4(46.8)	81.2(52.3)	84.8(51.7)	0.75	0.11
Non-verbal Mental age	82.5 (51.5)	79.9(44.6)	81(41.6)	89.3(48.3)	92.7(48.2)	0.84	0.15
Adaptive Functioning							
Composite score	73.6 (11.9)	71.3(13.4)	71.4(10.9)	72.8(12.1)	74.6(11.8)	0.91	0.003
Communication	77.7 (14.5)	75.1(16.3)	75.3(12.7)	76.5(14.9)	78.9(14.2)	0.89	0.001
Social	71.4 (12.4)	69.3(13.8)	70.3(11.8)	70.6(12.3)	72.2(12.4)	0.52	0.01
Daily living skills	76.8 (13.7)	73.6(14.7)	74.7(12.6)	75.9(13.9)	77.9(13.6)	0.52	0.004
Motor skills	81.8 (12.8)	82.1(15.3)	78.2(13.5)	82(12.4)	82.2(12.6)	0.13	0.88
ADOS							
Calibrated severity score	7.4 (1.7)	7.3(1.6)	7.4(1.8)	7.5(1.7)	7.4(1.7)	0.54	0.69
Social affect	11.0 (4.0)	11.2(4.1)	11.6(4.3)	11.1(4)	10.9(4)	0.45	0.24
Restricted/ repetitive behavior	4.0 (2.1)	4(2.3)	3.8(2.1)	4.1(2.1)	3.9(2)	0.41	0.12
Communication	4.5 (2.1)	4.6(2.2)	4.7(2.3)	4.4(2.1)	4.4(2.1)	0.88	0.69
Reciprocal social	8.2 (2.6)	8.3(2.7)	8.5(2.8)	8.3(2.6)	8(2.5)	0.56	0.07
ADI-R							
Social	20.2 (5.7)	20.1(5.8)	20.3(6.1)	20.4(5.8)	20.1(5.6)	0.85	0.30
Restricted/ repetitive behavior	6.5 (2.5)	6.1(2.3)	6.2(2.3)	6.6(2.5)	6.6(2.5)	0.78	1.00
Communication Non-verbal	9.2 (3.5)	9.2(3.5)	9.1(3.7)	9.3(3.5)	9.2(3.4)	0.97	0.62
Communication verbal	16.4 (4.3)	16.3(4.2)	16.7(4.3)	16.4(4.4)	16.4(4.2)	0.50	0.70
Social	20.2 (5.7)	20.1(5.8)	20.3(6.1)	20.4(5.8)	20.1(5.6)	0.85	0.30

Minimum final segment size used was 2500 kb. Phenotype-genotype analysis was performed to compare female and male probands with and without ROH in terms of IQ measures, the Vineland Adaptive Scales, and autism measures as reported by the ADI-R, the ADOS, the calibrated ADOS severity score. All tests were conducted in SAS version 9.2. Generalized Linear Model (GLM) was used to compute least squares means for each group.

Table S6. An Increase in Burden of Autosomal ROH in Females with Autism as Compared to Unaffected Female Siblings

Female Pairs					Male Pairs			
Minimum Segment Size (kb)	Probands± SE (n = 124)	Designated Sibling ± SE (n = 124)	Odds Ratio	P Value	Probands ± SE (n = 672)	Designated Sibling ± SE (n = 672)	Odds Ratio	P Value
Average Total Length of ROH (Mb)								
1000	28.62±1.06	27.76±0.94	1.14	0.24	28.26±0.51	28.46±0.48	0.97	0.55
1750	6.59±0.47	5.93±0.4	1.28	0.17	5.93±0.3	6.21±0.25	0.91	0.20
2000	4.49±0.38	3.80±0.33	1.47	0.080	3.91±0.29	4.06±0.23	0.94	0.45
2500	2.41±0.31	1.61±0.23	2.31	0.008	2.06±0.27	2.00±0.20	1.03	0.77
3000	1.31±0.23	0.89±0.17	2.35	0.048	1.24±0.25	1.22±0.19	1.01	0.90
3500	0.77±0.20	0.47±0.15	2.62	0.09	0.90±0.25	0.81±0.18	1.06	0.58
Average Number of Blocks								
1000	20.68±0.75	20.28±0.67	1.03	0.42	20.61±0.31	20.69±0.32	0.99	0.69
1750	2.80±0.19	2.62±0.17	1.08	0.36	2.47±0.07	2.64±0.08	0.91	0.02
2000	1.67±0.13	1.48±0.13	1.16	0.18	1.38±0.06	1.49±0.06	0.90	0.06
2500	0.73±0.09	0.48±0.07	1.65	0.01	0.55±0.037	0.55±0.035	0.98	0.84
3000	0.32±0.052	0.22±0.039	1.84	0.06	0.24±0.025	0.26±0.024	0.90	0.38
3500	0.15±0.04	0.09±0.03	2.40	0.09	0.137±0.02	0.14±0.02	0.97	0.87
Average Block Length per Person (Mb)								
1000	1.38±0.013	1.36±0.01	35.22	0.26	1.36±0.008	1.36±0.007	0.70	0.73
1750	2.07±0.08	1.94±0.09	1.73	0.20	1.97±0.04	2.31±0.12	0.83	0.004
2000	1.96±0.12	1.81±0.13	1.41	0.26	1.78±0.05	1.81±0.06	0.93	0.55
2500	1.47±0.16	1.22±0.17	1.37	0.17	1.22±0.07	1.23±0.07	0.99	0.95
3000	1.14±0.18	0.86±0.16	1.46	0.11	0.79±0.07	0.84±0.07	0.94	0.50
3500	0.69±0.17	0.47±0.15	1.49	0.17	0.53±0.07	0.53±0.06	1.00	0.99
Proportion of Subjects with ROH								
1000	1:1	1:1		1	1:1	1:1		1
1750	0.90:1	0.85:1		0.26	0.86:1	0.85:1		0.35
2000	0.74:1	0.69:1		0.40	0.68	0.69		0.77
2500	0.44:1	0.35:1		0.15	0.37:1	0.37:1		0.97
3000	0.28:1	0.21:1		0.18	0.19:1	0.21:1		0.41
3500	0.14:1	0.09:1		0.23	0.10:1	0.11:1		0.93

ROH was calculated using PLINK for a minimal final segment of 1000-3500 kb. Including those without ROH, average total length of ROH, average number of blocks and average block length were compared for the same gender proband and designated sibling pairs. A conditional logistic regression model in STATA version 11.1 was used to calculate statistical significance for average total length of ROH, average number of blocks and average block length per person. Odds Ratios were calculated for total length of ROH and average block size based on per 5 Mb and 2.5 Mb incremental increase respectively. Statistical significance of proportion of subjects with ROH was obtained by chi-square test. Female probands were observed to have higher average number of blocks and higher average total length of ROH compared to their designated siblings ($p < 0.05$) for final segment of 2500 kb. No statistically significant difference was found in the comparison of male probands to their designated siblings in terms of ROH burden ($p > 0.05$).

Table S7. Logistic Regression Analysis of Gender, ID, and ROH (Defined as Having ROH or no ROH) (Final segment size of 2500 kb) in Simplex Autism

Predictors	Unadjusted			Adjusted		
	OR	95% CI	p Value	OR	95% CI	p Value
Modeling Probability of ID in Simplex Autism						
Gender (female vs. male)	1.85	1.43 2.40	<0.0001	1.85	1.43 2.40	<0.0001
ROH (yes/no)	1.25	1.03 1.52	0.02	1.26	1.04 1.52	0.02
Modeling Probability of ROH in Simplex Autism						
Gender (female vs. male)	0.98	0.76 1.28	0.90	0.95	0.73 1.24	0.72
ID (ID vs. non-ID)	1.25	1.03 1.52	0.02	1.26	1.04 1.52	0.02

Logistic regression analyses in SAS version 9.2 were run to predict the association between gender, ID and ROH in affected probands. We fit simple logistic models to obtain unadjusted odds ratios and 95% confidence intervals and adjusted models to obtain adjusted odds ratios and 95% confidence intervals. We modeled the probability of having ID by gender and ROH; and the probability of having ROH by gender and ID.

¹Adjusted model includes both gender and ROH as opposed to the unadjusted model, which includes each risk factor individually.

²Adjusted model includes both gender and ID as opposed to the unadjusted model, which includes each risk factor individually.