Figure S1. CNV deletion boundaries.
Deleted intervals according to microarray analysis: markers CN_559777 to SNP A_8609540 (Affymetrix 6.0) within the chromosomal region 70,184,472 – 70,065,899 were deleted in MR individual MO1812A011 and markers CN_559779 to rs557324 (Illumina Human 1M) spanning the chromosomal region 70,220,882 – 70,154,208 were deleted in ASD individual SK0217-003. Black boxes represent exons 5-7, respectively. CNV markers are indicated in red, microsatellite markers in green and additional SNP markers in blue. The deleted area is indicated by “-” and the undeleted region indicated by “+”. Genomic sequence position is according to the UCSC Human Genome Browser assembly March 2006. Microsatellite analysis of the deleted region in MO1812A011 revealed the loss of the maternal allele. The precise breakpoints for MO1812A011 were determined by sequencing and reside at sequence positions 70,185,030 and 70,065,205 covering exactly 119,825 bp. The breakpoints of SK0217-003 were further narrowed down by qPCR and reside within 1642 bp (position 70,149,963 - 70,148,322) at the centromeric and within 2633 bp (position 70,221,920 – 70,219,288) at the telomeric end resulting in a deletion of about 69 kb.
Figure S2. The SHANK2 isoform expressed in human brain does not contain the ankyrin domain.
The ankyrin repeats and the N-terminal region of SHANK2 are present in mRNA from kidney and other epithelial tissues like liver (SHANK2E; data not shown). The longest isoform expressed in the brain starts with the SH3 domain and has a unique first exon and 5' UTR (neuronal SHANK2_1), whereas the most frequent neuronal isoform starts with the N-terminal PDZ domain (neuronal SHANK2_2)\(^\text{\textsuperscript{16}}\). ARF, ADP-ribosylation factor 1, housekeeping gene; b, brain; k, kidney; g, genomic DNA; "-", negative control; Exon description of SHANK2E is according to NM_012309.3, neuronal SHANK2 is according to the mRNA sequence of PROSAP1 (AB208026.1). Both annotations are taken from the UCSC Human Genome Browser assembly March 2006.
Figure S3. Pedigree structure of families with ASD and families with MR showing amino acid changes in SHANK2 that were not found in 659 controls.

The respective parents transmitting the variant are undiagnosed for ASD but in some instances present a clinical family history of ASD and a predominance of psychiatric disorders like depression, anxiety, social difficulties, epilepsy as well as learning disability (see also Table S2). Siblings carrying the same variant as the ASD patient show a similar phenotype (e.g. SK0332-005) or milder symptoms like speech delay (SK0191-005, SK0332-004), attention difficulties (SK0305-003, SK0332-004), articulation difficulties (MM1206-004), anxiety or a broader autism phenotype (SK0191-004) (see also Table S2). Non-carrier siblings of the variant have a normal phenotype (SK0443-004). Black symbol, affected ASD; grey symbol, milder symptoms (see above); red symbol, affected MR; white symbol, unaffected with ASD or MR.
Figure S4. SHANK2 sequence variants of three ASD patients and their healthy family.

a, The R462X variant of patient SK0441-003 occurs de novo. b, The L1008_P1009 duplication in patient MM1206-003 is inherited from the mother and was also found in a brother with articulation difficulties. c-d, The male ASD patient SK0443-003 inherited the variant T1127M from his unaffected mother. His unaffected sister SK0443-004 does not carry the mutation in the highly conserved dynamin-2 binding motif of SHANK2.
<table>
<thead>
<tr>
<th>Nucleotide change</th>
<th>Amino acid change</th>
<th>Occurrence</th>
<th>Known SNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.76C&gt;T</td>
<td>p.R26W</td>
<td>1 0 0</td>
<td></td>
</tr>
<tr>
<td>c.92C&gt;T</td>
<td>p.T31M</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>c.106C&gt;T</td>
<td>p.R36W</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>c.157G&gt;A</td>
<td>p.D53N</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>c.204C&gt;T</td>
<td>p.P68P</td>
<td>0 0 2</td>
<td></td>
</tr>
<tr>
<td>c.243C&gt;T</td>
<td>p.A81A</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>c.404A&gt;G</td>
<td>p.E135G</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>c.438C&gt;T</td>
<td>p.A146A</td>
<td>0 1 0</td>
<td></td>
</tr>
<tr>
<td>c.467A&gt;G</td>
<td>p.K156R</td>
<td>1* 1 0</td>
<td></td>
</tr>
<tr>
<td>c.492C&gt;T</td>
<td>p.G164G</td>
<td>3 0 0</td>
<td></td>
</tr>
<tr>
<td>c.526+61G&gt;A</td>
<td>-</td>
<td>1 2 6</td>
<td></td>
</tr>
<tr>
<td>c.527-18C&gt;A</td>
<td>-</td>
<td>1 0 -</td>
<td></td>
</tr>
<tr>
<td>c.527-9C&gt;T</td>
<td>-</td>
<td>0 1 1</td>
<td></td>
</tr>
<tr>
<td>c.527-8G&gt;T</td>
<td>-</td>
<td>1 1 -</td>
<td></td>
</tr>
<tr>
<td>c.533G&gt;A</td>
<td>p.S178N</td>
<td>0 1 2</td>
<td></td>
</tr>
<tr>
<td>c.569G&gt;A</td>
<td>p.R190H</td>
<td>7 5 16</td>
<td></td>
</tr>
<tr>
<td>c.622C&gt;T</td>
<td>p.P208S</td>
<td>1 1 0</td>
<td></td>
</tr>
<tr>
<td>c.640+11C&gt;T</td>
<td>-</td>
<td>2 0 -</td>
<td></td>
</tr>
<tr>
<td>c.640+15C&gt;A</td>
<td>-</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>c.641-53C&gt;A</td>
<td>-</td>
<td>0 0 1 yes</td>
<td></td>
</tr>
<tr>
<td>c.645C&gt;G</td>
<td>p.T215T</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>c.692C&gt;A</td>
<td>p.S231Y</td>
<td>0 1 0</td>
<td></td>
</tr>
<tr>
<td>c.716+10G&gt;A</td>
<td>-</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>c.800-97G&gt;A</td>
<td>-</td>
<td>0 2 -</td>
<td></td>
</tr>
<tr>
<td>c.800-33C&gt;T</td>
<td>-</td>
<td>1 0 -</td>
<td></td>
</tr>
<tr>
<td>c.924+19G&gt;A</td>
<td>-</td>
<td>1 1 - yes</td>
<td></td>
</tr>
<tr>
<td>c.924+64G&gt;A</td>
<td>-</td>
<td>0 1 -</td>
<td></td>
</tr>
<tr>
<td>c.924+133G&gt;C</td>
<td>-</td>
<td>1 1 -</td>
<td></td>
</tr>
<tr>
<td>c.925-57C&gt;G</td>
<td>-</td>
<td>0 1 -</td>
<td></td>
</tr>
<tr>
<td>c.1061-81C&gt;T</td>
<td>-</td>
<td>1 0 -</td>
<td></td>
</tr>
<tr>
<td>c.1141+49G&gt;A</td>
<td>-</td>
<td>3 1 -</td>
<td></td>
</tr>
<tr>
<td>c.1148-138G&gt;A</td>
<td>-</td>
<td>0 1 -</td>
<td></td>
</tr>
<tr>
<td>c.1148-109G&gt;T</td>
<td>-</td>
<td>1 1 -</td>
<td></td>
</tr>
<tr>
<td>c.1201 A&gt;C</td>
<td>p.K401Q</td>
<td>1 0 4 yes</td>
<td></td>
</tr>
<tr>
<td>c.1250C&gt;A</td>
<td>p.T417N</td>
<td>- - 1</td>
<td></td>
</tr>
<tr>
<td>c.1263G&gt;A</td>
<td>p.R421R</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>c.1264C&gt;A</td>
<td>p.P422T</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>Variant</td>
<td>Amino Acid Change</td>
<td>Count +</td>
<td>Count 0</td>
</tr>
<tr>
<td>---------</td>
<td>------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>c.1284 G&gt;A p.P428P</td>
<td>1 0 0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>c.1302+14C&gt;A</td>
<td>1 2 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1302+35G&gt;A</td>
<td>1 0 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1303-120G&gt;A</td>
<td>0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1303-54 C&gt;T</td>
<td>2 0 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1303-21C&gt;A</td>
<td>0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1316G&gt;A p.R439H</td>
<td>1 2 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1337C&gt;T p.T446M</td>
<td>0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1384C&gt;T p.R462X</td>
<td>1 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1392G&gt;T p.T464T</td>
<td>1 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1414+108C&gt;A</td>
<td>0 0 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1415-141T&gt;C</td>
<td>0 2 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1415-107G&gt;A</td>
<td>0 1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1763A&gt;G p.Y588C</td>
<td>7 5 16</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>c.1923G&gt;A p.E641E</td>
<td>1 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.2052G&gt;A p.P684P</td>
<td>1 0 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.2823C&gt;T p.D941D</td>
<td>1 0 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3024_3029dup p.L1008_P1009dup</td>
<td>1 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3135C&gt;T p.P992P</td>
<td>1 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3142C&gt;T p.R1048W</td>
<td>0 1 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3150T&gt;C p.A1050S</td>
<td>0 0 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3229C&gt;A p.P1077T</td>
<td>0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3324C&gt;T p.A1108A</td>
<td>3 1 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3380C&gt;T p.T1127M</td>
<td>1 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3620C&gt;T p.P1207L</td>
<td>2 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3843-89G&gt;A</td>
<td>0 2 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3843-84C&gt;T</td>
<td>0 1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3843-12T&gt;C</td>
<td>1 0 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3854T&gt;C p.I1285T</td>
<td>0 0 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3949G&gt;A p.D1317N</td>
<td>0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.4014G&gt;A p.M1338I</td>
<td>0 1 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.4048G&gt;A p.A1350T</td>
<td>1 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.4152G&gt;A p.S1384S</td>
<td>0 1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.*54C&gt;T</td>
<td>0 1 -</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table S1. Summary of all novel nucleotide variants detected in the SHANK2 gene from 1239 analysed individuals. Variants were counterchecked with the NCBI dbSNP database (build 130). Twenty-five of the variants leading to amino acid changes were novel and not found in 659 analysed controls of European decent. Seventeen of those were also found in analysed controls (polymorphisms). *The K156R variant was observed in a Chinese patient with ASD. In 2/45 Chinese (HapMap) and 0/39 Japanese control samples used as additional controls, this variant was also detected and termed as polymorphism.
<table>
<thead>
<tr>
<th>MO18120A11</th>
<th>SK0217-003</th>
<th>SK0441-003</th>
<th>SK0443-003</th>
<th>MM1206-003</th>
<th>SK335-004</th>
<th>SK332-003</th>
<th>SK0351-003</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNV</td>
<td>CNV</td>
<td>R462X</td>
<td>T1127M</td>
<td>L1006_P1009dup</td>
<td>R26W</td>
<td>P208S</td>
<td>A1305T</td>
</tr>
<tr>
<td>Genetics</td>
<td>Loss of exon 7 Frameshift</td>
<td>Loss of exon 6 and 7 Frameshift</td>
<td>Non sense mutation</td>
<td>Truncation of the protein</td>
<td>Duplication of 6 nucleotides</td>
<td>Missense mutation</td>
<td>Missense mutation</td>
</tr>
<tr>
<td>Sex</td>
<td>f</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>Transmission</td>
<td>de novo</td>
<td>de novo</td>
<td>de novo</td>
<td>mother</td>
<td>mother</td>
<td>father</td>
<td>mother</td>
</tr>
<tr>
<td>Final Research Diagnosis</td>
<td>autism</td>
<td>autism</td>
<td>ASD</td>
<td>autism</td>
<td>ASD</td>
<td>autism</td>
<td>autism</td>
</tr>
<tr>
<td>DIGONOSTIC MEASURES</td>
<td>ADO</td>
<td>Module 2</td>
<td>Module 2</td>
<td>Module 3</td>
<td>Module 1</td>
<td>Module 3</td>
<td>Module 1</td>
</tr>
<tr>
<td>Communication</td>
<td>6</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Social Interaction</td>
<td>10</td>
<td>12</td>
<td>7</td>
<td>7</td>
<td>11</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>19</td>
<td>11</td>
<td>11</td>
<td>17</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Play</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Interests &amp; Behaviors</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>ASD diagnosis</td>
<td>autism</td>
<td>autism</td>
<td>ASD</td>
<td>ASD</td>
<td>autism</td>
<td>autism</td>
<td>autism</td>
</tr>
<tr>
<td>Severity Score</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>ADI or ADI-R</td>
<td>ADI-R</td>
<td>ADI-R</td>
<td>ADI-R</td>
<td>ADI-R</td>
<td>ADI-R</td>
<td>ADI-R</td>
<td>ADI-R</td>
</tr>
<tr>
<td>Social</td>
<td>23</td>
<td>12</td>
<td>21</td>
<td>19</td>
<td>26</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Communication</td>
<td>21</td>
<td>13</td>
<td>16</td>
<td>21</td>
<td>10 (non-verbal)</td>
<td>14 (non-verbal)</td>
<td>10</td>
</tr>
<tr>
<td>Behaviour</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Sibling</td>
<td>autism</td>
<td>autism</td>
<td>autism</td>
<td>autism</td>
<td>autism</td>
<td>autism</td>
<td>autism</td>
</tr>
<tr>
<td>IQ MEASURES</td>
<td>standard score (T)</td>
<td>WISC-R</td>
<td>WASI</td>
<td>WPPSI III</td>
<td>WASI</td>
<td>WASI</td>
<td>WASI</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>87 (19%)</td>
<td>81 (15%)</td>
<td>78 (17%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance IQ</td>
<td>77 (8%)</td>
<td>73 (8%)</td>
<td>75 (7%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Scale</td>
<td>&lt;50 (&lt;1%)</td>
<td>67 (1%)</td>
<td>60 (1%)</td>
<td>70 (2%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>45 (&lt;1%)</td>
<td>88 (19%)</td>
<td>81 (15%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily living skills</td>
<td>30 (&lt;1%)</td>
<td>83 (13%)</td>
<td>87 (15%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socialization</td>
<td>56 (11%)</td>
<td>98 (16%)</td>
<td>79 (14%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor skills</td>
<td>n/a</td>
<td>n/a</td>
<td>80 (13%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension/total</td>
<td>42 (1%)</td>
<td>87 (19%)</td>
<td>81 (15%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LANGUAGE MEASURES</td>
<td>standard score (T)</td>
<td>OWLS</td>
<td>OWLS</td>
<td>OWLS</td>
<td>OWLS</td>
<td>PL 3</td>
<td>PL 3</td>
</tr>
<tr>
<td>receptive</td>
<td>57 (1%)</td>
<td>64 (1%)</td>
<td>71 (2%)</td>
<td>80 (13%)</td>
<td>50 (1%)</td>
<td>50 (1%)</td>
<td>50 (1%)</td>
</tr>
<tr>
<td>expressive</td>
<td>50 (1%)</td>
<td>79 (5%)</td>
<td>57 (1%)</td>
<td>72 (3%)</td>
<td>50 (1%)</td>
<td>50 (1%)</td>
<td>50 (1%)</td>
</tr>
<tr>
<td>total</td>
<td>50 (1%)</td>
<td>70 (2%)</td>
<td>62 (1%)</td>
<td>74 (4%)</td>
<td>50 (1%)</td>
<td>50 (1%)</td>
<td>50 (1%)</td>
</tr>
<tr>
<td>PPVT or PPVT-4</td>
<td>standard score (T)</td>
<td>PPVT-4</td>
<td>PPVT-4</td>
<td>PPVT-4</td>
<td>PPVT-3</td>
<td>PPVT-3</td>
<td></td>
</tr>
<tr>
<td>Cognitive ability</td>
<td>90 (40%)</td>
<td>61 (2%)</td>
<td>61 (2%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPOTPP</td>
<td>standard score (T)</td>
<td>103 (38%)</td>
<td>103 (38%)</td>
<td>87 (19%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phono logical awareness</td>
<td>82 (12%)</td>
<td>70 (2%)</td>
<td>100 (10%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phono logical memory</td>
<td>85 (19%)</td>
<td>100 (20%)</td>
<td>85 (19%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAMILY HISTORY</td>
<td>Parental age at birth</td>
<td>M=37, F=38</td>
<td>M=34, F=37</td>
<td>M=30, F=32</td>
<td>M=26, F=30</td>
<td>M=28, F=33</td>
<td>M=32, F=30</td>
</tr>
<tr>
<td>Father</td>
<td>attention difficulties</td>
<td>depression?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental family</td>
<td>ADHD</td>
<td>Social difficulties</td>
<td>social difficulties, reported high IQ, ADI-R</td>
<td>depression, ADD, LD, substance abuse, seizures, dementia?</td>
<td>depression, anxiety, OCD; ADI-R, LD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>anxiety</td>
<td>Social difficulties</td>
<td>social difficulties, reported high IQ, ADI-R</td>
<td>depression, ADD, LD, substance abuse, seizures, dementia?</td>
<td>depression, anxiety, OCD; ADI-R, LD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal family</td>
<td>depression, social difficulties</td>
<td>depression, ADD, LD, substance abuse, seizures, dementia?</td>
<td>depression, ADD, LD, substance abuse, seizures, dementia?</td>
<td>anxiety, autism, depression; 2 miscarriages</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parenthood (other information)</td>
<td>normal growth and phenotypes, motor delay evident at 5 months, slow reactions and adaptability, mild mimic</td>
<td>bilateral clinodactyly, 5th digito, bilateral dysmorphic lines</td>
<td>exposure to pesticides during pregnancy</td>
<td>peripheral distinct, loss of language skills and social engagement at 18 mos. (PPVT)</td>
<td>mild hypotonia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibling</td>
<td>1 sibling, 4 half-siblings, all healthy</td>
<td>articulation difficulties</td>
<td>mild attention difficulties</td>
<td>mild attention difficulties, 1224</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ancestry</td>
<td>European</td>
<td>European</td>
<td>European</td>
<td>European</td>
<td>European</td>
<td>European</td>
<td>European</td>
</tr>
</tbody>
</table>

Table S2. Clinical description of ASD patients.

Two de novo CNV individuals and three patients with nonsense, missense and insertion mutations that were tested in functional assays are indicated in blue. Individuals with other missense mutations are indicated in white. ADOS: Autism Diagnostic Observation Schedule; ADI-R: Autism Diagnostic Interview-Revised; SON-R: Snijders-Oomen Nonverbal Intelligence Test; Leiter-R: Leiter International Performance Scale-Revised; WASI: Wechsler Abbreviated Scale of Intelligence; WPPSI III: Wechsler Preschool and Primary Scale of Intelligence - 3rd Edition; VABS: Vineland Adaptive Behaviour Scale (2nd Edition); OWLS: Oral and Written Language Scale; PLS-3: Preschool Language Scale-3; PPVT-3 (or 4): Peabody Picture Vocabulary Test (3rd or 4th Edition); CTOPP: Comprehensive Test of Phono logical Processing, f: female, m: male; ADHD: Attention-deficit hy peractivity disorder; ASD: Autism Spectrum Disorder; LD: Learning Disability; DD: Developmental Delay; OCD: Obsessive Compulsive Disorder; BAP: Broader Autism Phenotype; SLI: Specific Language Impairment; ART: Assisted Reproductive Technology.
Table S3: List of rare CNVs and overlapping genes in ASD probands identified with SHANK2 mutations

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Microarray</th>
<th>Ancestry¹</th>
<th>Cytoband</th>
<th>Length</th>
<th>CNV</th>
<th>Transmission</th>
<th>Genes</th>
<th>OMIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM1206-003</td>
<td>Illumina 1M</td>
<td>European</td>
<td>15q21.3</td>
<td>53,868</td>
<td>Loss maternal</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0191-003</td>
<td>Illumina 1M</td>
<td>European</td>
<td>3q22.3</td>
<td>21,324</td>
<td>Loss paternal</td>
<td>PCCB</td>
<td>PCCB</td>
<td></td>
</tr>
<tr>
<td>SK0305-004</td>
<td>Illumina 1M</td>
<td>European</td>
<td>6q27</td>
<td>17,584</td>
<td>Gain both</td>
<td>PSMB1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0332-003</td>
<td>Illumina 1M</td>
<td>West Indian</td>
<td>12q21.2</td>
<td>124,007</td>
<td>Loss paternal</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0332-003</td>
<td>Illumina 1M</td>
<td>West Indian</td>
<td>17q21.31</td>
<td>10,577</td>
<td>Loss paternal</td>
<td>VAT1, IFI35</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0441-003</td>
<td>Affymetrix 6.0</td>
<td>European</td>
<td>2p23.1</td>
<td>132,444</td>
<td>Loss maternal</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0441-003</td>
<td>Affymetrix 6.0</td>
<td>European</td>
<td>11q13.4</td>
<td>32,495</td>
<td>Gain unknown</td>
<td>FOLR3</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0441-003</td>
<td>Affymetrix 6.0</td>
<td>European</td>
<td>15q24.2</td>
<td>107,409</td>
<td>Gain unknown</td>
<td>IMP3, SNX33, SNUPN, CSPG4</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0441-003</td>
<td>Affymetrix 6.0</td>
<td>European</td>
<td>16q24.3</td>
<td>112,590</td>
<td>Gain unknown</td>
<td>CDK10, CPNE7, C16orf7, SPATA2L, CHMP1A, DPEP1, C16orf55</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0441-003</td>
<td>Affymetrix 6.0</td>
<td>European</td>
<td>20q13.33</td>
<td>55,228</td>
<td>Loss unknown</td>
<td>CDH4</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0443-003</td>
<td>Affymetrix 6.0</td>
<td>Admixed</td>
<td>6p12.3</td>
<td>62,586</td>
<td>Gain unknown</td>
<td>RUNX2, SUPT3H</td>
<td>RUNX2</td>
<td></td>
</tr>
<tr>
<td>SK0443-003</td>
<td>Affymetrix 6.0</td>
<td>Admixed</td>
<td>7p21.1</td>
<td>4,803</td>
<td>Loss unknown</td>
<td>HDAC9</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0217-003</td>
<td>Illumina 1M</td>
<td>European</td>
<td>6q21</td>
<td>98,376</td>
<td>Gain paternal</td>
<td>RFPL4B, LAMA4</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0217-003</td>
<td>Illumina 1M</td>
<td>European</td>
<td>11q13.3</td>
<td>66,175</td>
<td>Loss de novo</td>
<td>SHANK2</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SK0217-003</td>
<td>Illumina 1M</td>
<td>European</td>
<td>4q12</td>
<td>18,162</td>
<td>Loss maternal</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>HDMO1812A011</td>
<td>Affymetrix 6.0</td>
<td>European</td>
<td>11q13.3</td>
<td>118,573</td>
<td>Loss de novo</td>
<td>SHANK2</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>TUJS071228</td>
<td>Affymetrix 6.0</td>
<td>European</td>
<td>8p22</td>
<td>43,639</td>
<td>Loss paternal</td>
<td>DLC1</td>
<td>DLC1</td>
<td></td>
</tr>
<tr>
<td>TUJS071228</td>
<td>Affymetrix 6.0</td>
<td>European</td>
<td>8p21</td>
<td>585,330</td>
<td>Loss paternal</td>
<td>NEFM, NEFL</td>
<td>NEFL</td>
<td></td>
</tr>
<tr>
<td>TULB081863</td>
<td>Affymetrix 6.0</td>
<td>European</td>
<td>Xq13.2</td>
<td>20,793</td>
<td>Gain paternal</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

¹Ancestry determined using SNP genotypes as described in supplemental methods
²SK0443-003 is admixed 73% European, 13% East Asian, 14% African according to SNP genotype. Self reported ancestry is of Middle Eastern origin.
Methods

MR and ASD cases
CNV analysis was performed in 184 probands with well-defined Mental Retardation (MR) of unknown etiology and their unaffected parents were recruited as part of the German Mental Retardation Network (MRNET) study (data not shown). The same 184 individuals with MR, (21 individuals of the 184 had autistic features as reported from clinical assessment), were screened for mutations using Sanger-based sequencing. Laboratory tests to rule out known medical causes of MR were performed and included karyotyping, fragile X testing, screening of the subtelomeric regions. The ASD cohort was comprised of 396 unrelated individuals assessed using the Autism Diagnostic Interview-Revised (ADI-R)17 and Autism Diagnostic Observation Schedule (ADOS)18 that were ascertained at the Hospital for Sick Children in Toronto and McMaster University Hamilton and screened for CNVs (described in 19,20). These same 396 ASD cases were also screened using Sanger-based sequencing. Institutional ethical review board approvals were obtained for the study.

CNV Analysis
Affymetrix genome-wide human SNP Array 6.0 and Illumina Infinium 1 M single SNP arrays were used to assess CNVs in a cohort of 184 MR cases and 396 ASD cases and their respective parents. The CNV results relevant to the SHANK2 findings are described here. For CNV calling on the Affymetrix 6.0 array three different CNV calling algorithms were used: Birdsuite, Affymetrix Genotyping Console and our in-house developed algorithm iPattern (unpublished). For CNV calling on the Illumina 1M platform we used combinations of QuantSNP, iPattern and PennCNV. All CNVs required a minimum of five consecutive array probes detected and CNVs needed to be called by two or more algorithms. Both platforms have sufficient probe coverage for reliable detection of copy number changes to a resolution of 5-10 kb. The genomic region of the SHANK2 gene is interrogated by 397 SNP markers with a 1134 bp median intermarker distance on the Affymetrix 6.0 array and 224 SNP markers with a 2098 bp median intermarker distance on the Illumina 1M array.

Controls
For comparison to controls we used CNV data analyzed as described above from 5,023 controls of European ancestry (1,234 from Ottawa, Canada21, 1,123 from Northern Germany22 and 59 CEPH Hapmap controls run on the Affymetrix 6.0 array; 1,287 from the SAGE control project on the Illumina 1M platform; 1,320 from CHOP paediatric control study on the Illumina 550k array23). For DNA sequencing 659 control DNA samples of European ancestry were from Germany (374) or from the Ontario Population Genomics Project control collection (285) with a gender distribution of 378 males to 281 females.

SNP Ancestry Analysis
Affymetrix 6.0 genotypes from cases and controls were clustered using 1,120 unlinked SNPs by the program STRUCTURE24, assuming three ancestral populations from HapMap Phase 2 genotypes. European Ancestry was assigned using if the threshold of cluster membership score > 0.9. For cases genotyped with the Illumina 1M platform, ancestry was determined using 5,239 widely-spaced, independent SNPs that had a genotype completion rate of ≥99.9%. Software used was SpectralGEM19, which estimates 5 significant dimensions of ancestry25.
Fluorescence in situ hybridisation
Chromosome metaphases were prepared according to standard protocols from primary blood samples. FISH was performed as described elsewhere\textsuperscript{26} using the BAC DNA probes CTD-3053G20, CTD-3244D21 and CTD-2591D19 (Invitrogen).

Sequencing / Mutation Screening
DNA was extracted from blood leukocytes using phenol chloroform extraction\textsuperscript{27}. PCR primers were designed using the Primer 3 software (v. 0.4.0) to amplify each exon of \textit{SHANK2} and the flanking intron-exon boundaries. We analyzed the longest neuronal form of \textit{SHANK2}, \textit{SHANK2\_1}, that starts with the SH3 domain and has a unique first exon and 5' UTR. Exon numbering is according to reference sequence AB208026.1. PCRs were performed with Paq5000 polymerase (from Stratagene) using standard conditions, and products were purified and sequenced directly using the DYEnamic ET Terminator Cycle Sequencing Kit (GE Healthcare) and the MegaBACE 1000 DNA Analysis System (GE Healthcare). To narrow down the deletion breakpoints long-range PCR according to the Expand Long Template PCR System (Roche) and qPCR with the 7500 Fast Real-Time PCR System (Applied Biosystems) were performed.

RNA Analysis
RNA was prepared from a culture of dermal fibroblasts of patient MO1812A011 with the 120 kb \textit{SHANK2} deletion. Total RNA was isolated using Trizol reagent. cDNA was synthesized using the SuperScript III first-strand synthesis system for RT-PCR (Invitrogen). PCR primers were designed from exon 6 to exon 8, resulting in a 242 bp fragment.
List of primers:

Primer for long-range PCR & sequencing across the 120 kb deletion

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHANK2-LR8F</td>
<td>atgaggaccatctctcatctaccac</td>
</tr>
<tr>
<td>SHANK2-LR8R</td>
<td>ctcctatatctctgctgtgatt</td>
</tr>
</tbody>
</table>

Primer for sequencing the neuronal form of SHANK2

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHANK2_Ex1F</td>
<td>agacacccgccaccccttc</td>
</tr>
<tr>
<td>SHANK2_Ex1R</td>
<td>ggcaattttagagagaaatg</td>
</tr>
<tr>
<td>SHANK2_Ex2F</td>
<td>gaactcacagctccccctcg</td>
</tr>
<tr>
<td>SHANK2_Ex2R</td>
<td>gcttcgctcaacaccttcg</td>
</tr>
<tr>
<td>SHANK2_Ex3F</td>
<td>catctctctctctgcatca</td>
</tr>
<tr>
<td>SHANK2_Ex3R</td>
<td>ggaatctcattttccagtgga</td>
</tr>
<tr>
<td>SHANK2_Ex4F</td>
<td>ggaagtgcaagtccccaccaga</td>
</tr>
<tr>
<td>SHANK2_Ex4R</td>
<td>gatgaaggccagagagag</td>
</tr>
<tr>
<td>SHANK2_Ex5F</td>
<td>cagacccctccccccagacag</td>
</tr>
<tr>
<td>SHANK2_Ex5R</td>
<td>tccataattccagcccgta</td>
</tr>
<tr>
<td>SHANK2_Ex6F</td>
<td>caccctgcaagctctctgt</td>
</tr>
<tr>
<td>SHANK2_Ex6R</td>
<td>gggaacatattcccctcctg</td>
</tr>
<tr>
<td>SHANK2_Ex7F</td>
<td>dtggactagccacccgaat</td>
</tr>
<tr>
<td>SHANK2_Ex7R</td>
<td>cacaagccacccccagact</td>
</tr>
<tr>
<td>SHANK2_Ex8F</td>
<td>ctccacagcagcagcagcat</td>
</tr>
<tr>
<td>SHANK2_Ex8R</td>
<td>gtgggatcaactggtgagag</td>
</tr>
<tr>
<td>SHANK2_Ex9_10F</td>
<td>cacctgtccttctcccacat</td>
</tr>
<tr>
<td>SHANK2_Ex9_10R</td>
<td>ctcctctctctctccatct</td>
</tr>
<tr>
<td>SHANK2_Ex12F</td>
<td>tctatgtcattgtctcttg</td>
</tr>
<tr>
<td>SHANK2_Ex12R</td>
<td>gcactgagtgacggttaag</td>
</tr>
<tr>
<td>SHANK2_Ex13F</td>
<td>ccatactgagggctctc</td>
</tr>
<tr>
<td>SHANK2_Ex13R</td>
<td>gatgtctcaagtgtctcctac</td>
</tr>
<tr>
<td>SHANK2_E14F</td>
<td>cctcttgactctgacatg</td>
</tr>
<tr>
<td>SHANK2_E14R</td>
<td>attgcaagatggtgcttc</td>
</tr>
<tr>
<td>SHANK2_Ex15-1F</td>
<td>caccctgtccttctcccacat</td>
</tr>
<tr>
<td>SHANK2_Ex15-1R</td>
<td>ctcctctctctctccatct</td>
</tr>
<tr>
<td>SHANK2_Ex15-2F</td>
<td>aaccctgcaacacagagag</td>
</tr>
<tr>
<td>SHANK2_Ex15-2R</td>
<td>ctcctctctctctccatct</td>
</tr>
<tr>
<td>SHANK2_Ex15-3F</td>
<td>ctcctctctctctccatct</td>
</tr>
<tr>
<td>SHANK2_Ex15-3R</td>
<td>gatgagagagagagag</td>
</tr>
<tr>
<td>SHANK2_Ex15-4F</td>
<td>gctctgcaagctctctctc</td>
</tr>
<tr>
<td>SHANK2_Ex15-5F</td>
<td>gctctgcaagctctctctc</td>
</tr>
<tr>
<td>SHANK2_Ex15-5R</td>
<td>gctctgcaagctctctctc</td>
</tr>
<tr>
<td>SHANK2_Ex15-6F</td>
<td>caccctgtccttctcccacat</td>
</tr>
<tr>
<td>SHANK2_Ex15-6R</td>
<td>ctcctctctctctccatct</td>
</tr>
<tr>
<td>SHANK2_Ex16-1F</td>
<td>caccctgtccttctcccacat</td>
</tr>
<tr>
<td>SHANK2_Ex16-1R</td>
<td>ctcctctctctctccatct</td>
</tr>
<tr>
<td>SHANK2_Ex16-2F</td>
<td>caccctgtccttctcccacat</td>
</tr>
<tr>
<td>SHANK2_Ex16-2R</td>
<td>ctcctctctctctccatct</td>
</tr>
</tbody>
</table>

Primer to detect the deletion of Exon 7 on RNA level

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHANK2_RT1F</td>
<td>cttggttctggtcctcag</td>
</tr>
<tr>
<td>SHANK2_RT1R</td>
<td>ccacccactagccagaggt</td>
</tr>
</tbody>
</table>

Primer to detect the neuronal isoform of SHANK2

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHANK2_Ex1_5R</td>
<td>caagcgtgtctggaatggtag</td>
</tr>
<tr>
<td>SHANK2_Ex1_16F</td>
<td>ctacgtctcacaacacctctac</td>
</tr>
<tr>
<td>SHANK2_Ex5_8F</td>
<td>agctctcaccaggaagacaaat</td>
</tr>
<tr>
<td>SHANK2_Ex6_8R</td>
<td>atgtctctccccctgtaagacag</td>
</tr>
<tr>
<td>SHANK2_Ex7_10F</td>
<td>ccactgtctgccttccagat</td>
</tr>
<tr>
<td>SHANK2_Ex7_10R</td>
<td>cagacccctcccggggaagcag</td>
</tr>
<tr>
<td>SHANK2_Ex10_12F</td>
<td>cgctcaccagcagaaacaa</td>
</tr>
<tr>
<td>SHANK2_Ex10_12R</td>
<td>actcggccgaaacaccac</td>
</tr>
<tr>
<td>SHANK2_Ex14_16F</td>
<td>cttggattctgggtcctcag</td>
</tr>
</tbody>
</table>

Nature Genetics: doi:10.1038/ng.589
SHANK2_Ex14_16R  ccaccttaaggccaggtga
ARF For  gccaglgtcctccacccgtgc
ARF Rev  gcctcgtcacaagcctcctg

qPCR - Primer

centromeric breakpoint:

| 11q13.3_40F (not deleted) | atccccctcaatccattttcat |
| 11q13.3_41R (not deleted) | gaaagacgttggtggtcctt |
| 11q13.3_21F (deleted)     | aacctctgcctcctcttngc |
| 11q13.3_22R (deleted)     | cagtcagatatttggtggt |

telomeric breakpoint:

| 11q13.3_09F (not deleted) | tctttccacacatcagacaac |
| 11q13.3_10R (not deleted) | cattctcgcaacctctgctta |
| 11q13.3_36F (deleted)     | tctcgtcagttcccdcaa |
| 11q13.3_37R (deleted)     | aaggccagttatggacatct |

Supplementary notes/References