Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders.

Christelle M. Durand, Catalina Betancur, Tobias M. Boeckers, Juergen Bockmann, Pauline Chaste, Fabien Fauchereau, Gudrun Nygren, Maria Rastam, I Carina Gillberg, Henrik Anckarsäter, et al.

Mutations of the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders

Christelle M. Durand[1], Catalina Betancur[2], Tobias M Boeckers[3], Juergen Bockmann[3], Pauline Chaste[1], Fabien Fauchereau[1,4], Gudrun Nygren[5], Maria Rastam[5], I. Carina Gillberg[5], Henrik Anckarsäter[5], Eili Sponheim[6], Hany Goubran-Botros[1], Richard Delorme[1], Nadia Chabane[7], Marie-Christine Mouren-Simeoni[7], Philippe de Mas[8], Eric Bieth[8], Bernadette Rogé[9], Delphine Héron[10], Lydie Burglen[11], Christopher Gillberg[5,12], Marion Leboyer[2,13], Thomas Bourgeron[1,4*]

[1] *Human Genetics and Cognitive Functions, Institut Pasteur, Paris, France*

[2] *INSERM U513, Université Paris XII, Créteil, France*

[3] *Institute for Anatomy and Cell Biology, Ulm University, Ulm, Germany*

[4] *University Denis Diderot Paris 7, Paris, France*

[5] *Department of Child and Adolescent Psychiatry, Göteborg University, Göteborg, Sweden*

[6] *Centre for Child and Adolescent Psychiatry, University of Oslo, Oslo, Norway*

[7] *Service de Psychopathologie de l'Enfant et de l'Adolescent, Hôpital Robert Debré, Assistance Publique-Hôpitaux de Paris, Paris, France*

[8] *Purpan Hospital, Department of Medical Genetics, Toulouse, France*

[9] *Centre d'Etudes et de Recherches en PsychoPathologie, Université de Toulouse le Mirail, Toulouse, France*

[10] *Département de Génétique, Groupe Hospitalier Pitié-Salpêtrière, Paris, France*

[11] *Service de Génétique, Hôpital Trousseau, Assistance Publique-Hôpitaux de Paris, Paris, France*

[12] *Saint George's Hospital Medical School, London, United Kingdom*

[13] *Département de Psychiatrie, Groupe Hospitalier Henri Mondor et Albert Chenevier, Assistance Publique-Hôpitaux de Paris, Créteil, France*

* Correspondence should be addressed to Thomas Bourgeron, PhD, Human Genetics and Cognitive Functions, Institut Pasteur, 25 rue du Docteur Roux, 75015 Paris, France. Tel: 33-1-40-61-32-16. Fax: 33-1-40-61-39-53. E-mail: thomasb@pasteur.fr

Abstract

SHANK3/ProSAP2 regulates the structural organization of dendritic spines and is a binding partner of neuroligins, previously found to be mutated in autism and Asperger syndrome. Here, we report that a mutation of a single copy of *SHANK3* on chromosome 22q13 is sufficient to develop language and/or social communication disorders. These mutations concern only a small number of individuals, but shed light on one synaptic pathway, sensitive to gene dosage, involved in autism spectrum disorders.

Autism spectrum disorder (ASD) affects about 6/1000 children and is characterized by impairments in reciprocal social interaction and communication as well as restricted and stereotyped patterns of interests and activities[1]. ASD ranges from severe (autistic disorder with moderate or severe cognitive impairment) to a milder variant (Asperger syndrome [AS] with higher cognitive ability). Although the causative genes remain largely unknown[2], familial and twin studies indicate that ASD is one of the most genetic neuropsychiatric disorders. Standard karyotype analyses reveal chromosomal rearrangements in 3%-6% of cases, the most common being deletions/duplications on chromosomes 15q, 22q and 7q[3]. Among the most frequent rearrangements associated with cognitive deficits, the 22q13.3 microdeletion syndrome is characterized by neonatal hypotonia, global developmental delay, normal to accelerated growth, absent to severely delayed speech, autistic behavior, and minor dysmorphic features[4]. The loss of terminal 22q13.3 can be subtle and go undetected by routine chromosome analysis; fluorescence in situ hybridization (FISH) is often required to confirm the presence of this deletion.

Among the three genes (*ACR*, *RABL2B*, *SHANK3*) located in the minimal telomeric region[5], *SHANK3* (also known as ProSAP2) is the strongest candidate for the neurobehavioral symptoms observed in patients with 22q13 deletions. SHANK3 is a scaffolding protein found in excitatory synapses directly opposite to the presynaptic active zone. Shank proteins are believed to function as master organizers of the postsynaptic density (PSD), due to their ability to form multimeric complexes with postsynaptic receptors, signaling molecules and cytoskeletal proteins present in dendritic spines/PSDs[6,7]. SHANK3 can bind to the cell adhesion proteins neuroligins (NLGN)[8], which we previously found to be mutated in autism and Asperger syndrome[9]. *SHANK3* was disrupted by a *de novo* balanced translocation in a child with all the features of the 22q13.3 deletion syndrome[10]. In this paper, we report evidence showing that abnormal gene dosage of *SHANK3* is associated with severe cognitive deficits, including language/speech disorder and ASD.

We investigated chromosome 22q13 and *SHANK3* in patients with ASD by FISH analysis (n=97) and/or direct sequencing (n=227). We also sequenced all *SHANK3* exons in a minimum of 190 controls to ascertain the diversity of *SHANK3* non-synonymous variations in the general population. *SHANK3* spans 57 kb and contains 24 exons. Seven exons are alternatively spliced, including exon 18, which is mostly detected in the brain (Supplementary Fig. 1). During our screening, three families with ASD showed unambiguous alteration of 22q13/*SHANK3*. In family ASD 1, the proband with

autism, no language and moderate mental retardation, carried a *de novo* deletion of 22q13 (the clinical description of all patients is provided in Supplementary note 1). The deletion breakpoint was located in intron 8 of *SHANK3* and removed 142 kb of the terminal 22q13 (Fig. 1a). This deletion had been "repaired" by addition of telomeric repeats and was similar to the minimum deleted region described previously[5]. The occurrence of recurrent deletions in this region may be due to the quadruplex forming G-rich sequence (QGRS) surrounding the breakpoint (Supplementary Fig. 2), which provides a structural substrate for inappropriate telomere formation.

In family ASD 2, two brothers with autism were heterozygous for an insertion of a guanine nucleotide in exon 21 (Fig. 1b). Both brothers had severely impaired speech and severe mental retardation. The mutation was absent in an unaffected brother and the unaffected parents. Using fourteen informative single nucleotide polymorphisms (SNPs), we found that the mutation was located on the same maternal haplotype in the two affected brothers and that the unaffected brother did not have this haplotype (Supplementary Fig. 3). The mutation was absent in the DNA isolated from blood leukocytes and mouth cells of the mother. These results strongly suggest a germinal mosaicism in the mother. The guanine insertion creates a frame-shift at nucleotide 3680, modifying the C–terminal sequence of the protein (Fig1b). This putative truncated protein lacks several crucial domains involved in mGluR- and actin binding (Homer, AbP1, cortactin) and in the synaptic targeting and postsynaptic assembly of SHANK3 multimers[11,12]. Consistent with the loss of these domains, the mutant protein over-expressed in rat hippocampus neuronal cells showed no synaptic localization compared to the wild-type sequence (Supplementary Fig. 4).

In family ASD 3, we identified a terminal 22q deletion in a girl with autism and severe language delay and a 22qter partial trisomy in her brother with Asperger syndrome, who exhibited precocious language development and fluent speech (Fig. 1c). We demonstrated that these unbalanced cytogenetic abnormalities were inherited from a paternal translocation t(14;22)(p11.2;q13.33). The chromosome 14p11.2 breakpoint falls within the heterochromatic DNA sequence characteristic of acrocentric chromosomes and contains no putative transcripts or genes. On chromosome 22q13.33, using informative SNPs and quantitative PCR, we mapped the breakpoint between *ALG12* and *MLC1* (Fig. 1d). The deletion/duplication rearrangement observed in both siblings involved 25 genes, including *SHANK3*, located in the 800 kb terminal sequence of 22q13. No other *SHANK3* deletions or duplications were observed after screening by quantitative PCR 155 individuals (58 with autism, 38 with Asperger syndrome and 59 controls).

In the remaining individuals with ASD, we identified seven who had rare non-synonymous variations, which were not observed in controls (n=270-333; Fig. 2; Supplementary Table 1). However, all these variations were inherited from healthy parents, ruling out their direct involvement as dominant mutations in the disorder. Interestingly, for two substitutions modifying highly conserved amino acids (R12C and R300C; Supplementary Fig. 3), we observed that the over-expressed mutated GFP Shank3 fusion proteins clustered but showed significantly less co-localization with the presynaptic marker protein Bassoon, suggesting nonsynaptic clustering (Supplementary Fig. 4). These

observations might reflect posttranslational modifications or abnormal folding of the protein. Thus, although these genetic variations cannot be considered as causal mutations, they might nevertheless modify the synaptic scaffolding and represent risk factors for ASD in interaction with other susceptibility genes.

In this study, we show that a *SHANK3* heterozygous mutation can cause ASD. Interestingly, in the ASD 3 boy with Asperger syndrome, the presence of an additional copy of 22q13/*SHANK3* did not impair his language ability, but appears to have led to a severe impairment in social communication. These results, together with previous reports[13,14], highlight the importance of a fine gene dosage for the development of speech-language and/or social communication in humans.

The mutations identified in these patients are believed to have affected the function/localization of SHANK3 at PSD/dendritic spines. These results are consistent with the alterations of dendrites and the spine/PSD compartment in individuals with learning disabilities[15]. In mice, Shank-3 promotes the maturation and the enlargement of dendritic spine heads and is even able to induce spine formation in aspiny neurons[11]. In ASD, an abnormality of synapse formation/maintenance was first suggested by the identification of mutations in X-linked *NLGN3* and *NLGN4*[9], and next confirmed by functional studies of the causative mutations. Therefore, we hypothesize that the protein complex including NLGN/SHANK participates in the assembly of specialized postsynaptic structures required for the development of language and social communication.

References

1. Folstein, S.E. & Rosen-Sheidley, B. *Nat Rev Genet* **2**, 943-955 (2001).
2. Persico, A.M. & Bourgeron, T. *Trends Neurosci* **29**, 349-358 (2006).
3. Vorstman, J.A. et al. *Mol Psychiatry* **11**, 1, 18-28 (2006).
4. Manning, M.A. et al. *Pediatrics* **114**, 451-7 (2004).
5. Bonaglia, M.C. et al. *J Med Genet* **43**, 822-8 (2006).
6. Naisbitt, S. et al. *Neuron* **23**, 569-82 (1999).
7. Boeckers, T.M., Bockmann, J., Kreutz, M.R. & Gundelfinger, E.D. *J Neurochem* **81**, 903-10 (2002).
8. Meyer, G., Varoqueaux, F., Neeb, A., Oschlies, M. & Brose, N. *Neuropharmacology* **47**, 724-33 (2004).
9. Jamain, S. et al. *Nat. Genet.* **34**, 27-29 (2003).
10. Bonaglia, M.C. et al. *Am J Hum Genet* **69**, 261-8 (2001).
11. Roussignol, G. et al. *J Neurosci* **25**, 3560-70 (2005).
12. Baron, M.K. et al. *Science* **311**, 531-5 (2006).
13. Lai, C.S., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F. & Monaco, A.P. *Nature* **413**, 519-23. (2001).
14. Somerville, M.J. et al. *N Engl J Med* **353**, 1694-701 (2005).
15. Carlisle, H.J. & Kennedy, M.B. *Trends Neurosci* **28**, 182-7 (2005).

Acknowledgements

We thank the patients and their families for participating in this study and all the collaborators of the *Paris Autism Research International Sibpair Study (PARIS): Sweden*: Department of Child and Adolescent Psychiatry, Göteborg University, Göteborg: Christopher Gillberg, Maria Råstam, Carina Gillberg, Gudrun Nygren, Henrik Anckarsäter, Ola Ståhlberg; *France*: Department of Psychiatry, Groupe Hospitalier Albert Chenevier et Henri Mondor, Créteil: Marion Leboyer; INSERM U513, Université Paris XII, Créteil: Catalina Betancur; Service de Psychopathologie de l'Enfant et l'Adolescent, Hôpital Robert Debré, Paris: Catherine Colineaux, Deborah Cohen, Nadia Chabane, Marie-Christine Mouren-Siméoni; INSERM U679, Hôpital Pitié-Salpêtrière, Paris: Alexis Brice; *Norway*: Centre for Child and Adolescent Psychiatry, University of Oslo, Oslo: Eili Sponheim; Department of Pediatrics, Rikshospitalet, University of Oslo, Oslo: Ola H. Skjeldal; *USA*: Department of Pediatrics, Georgetown University School of Medicine, Washington D.C.: Mary Coleman; Children's National Medical Center, George Washington University School of Medicine, Washington, D.C.: Philip L. Pearl; New York State Institute for Basic Research in Developmental Disabilities, Staten Island, New York: Ira L. Cohen, John Tsiouris; *Italy*: Divisione di Neuropsichiatria Infantile, Azienda Ospedaliera Senese, Siena: Michele Zappella; *Austria*: Department of General Psychiatry, University Hospital, Vienna: Harald Aschauer ; *Belgium*: Centre de Génétique Humaine, Institut de Pathologie et de Génétique, Gerpinnes, Loverval: Lionel Van Maldergem

We also thank the DNA and cell bank of INSERM U679 (IFR des Neurosciences, Hôpital Pitié-Salpêtrière), the Centre d'Investigations Cliniques of the Hôpital Robert Debré, C. Bouchier and S. Duthoy for the use of sequencing facilities at the Génopole Pasteur, and A. Hchikat, L. Margarit, and G. Rouffet for technical assistance. This work was supported by the Pasteur Institute, INSERM, Assistance Publique-Hôpitaux de Paris, Fondation France Telecom, Cure Autism Now, Fondation de France, Fondation Biomédicale de la Mairie de Paris, Fondation pour la Recherche Médicale, EUSynapse European Commission FP6, AUTISM MOLGEN European Commission FP6, Fondation NRJ, the Swedish Science Council and the Deutsche Forschungsgemeinschaft DFG, SFB 497.

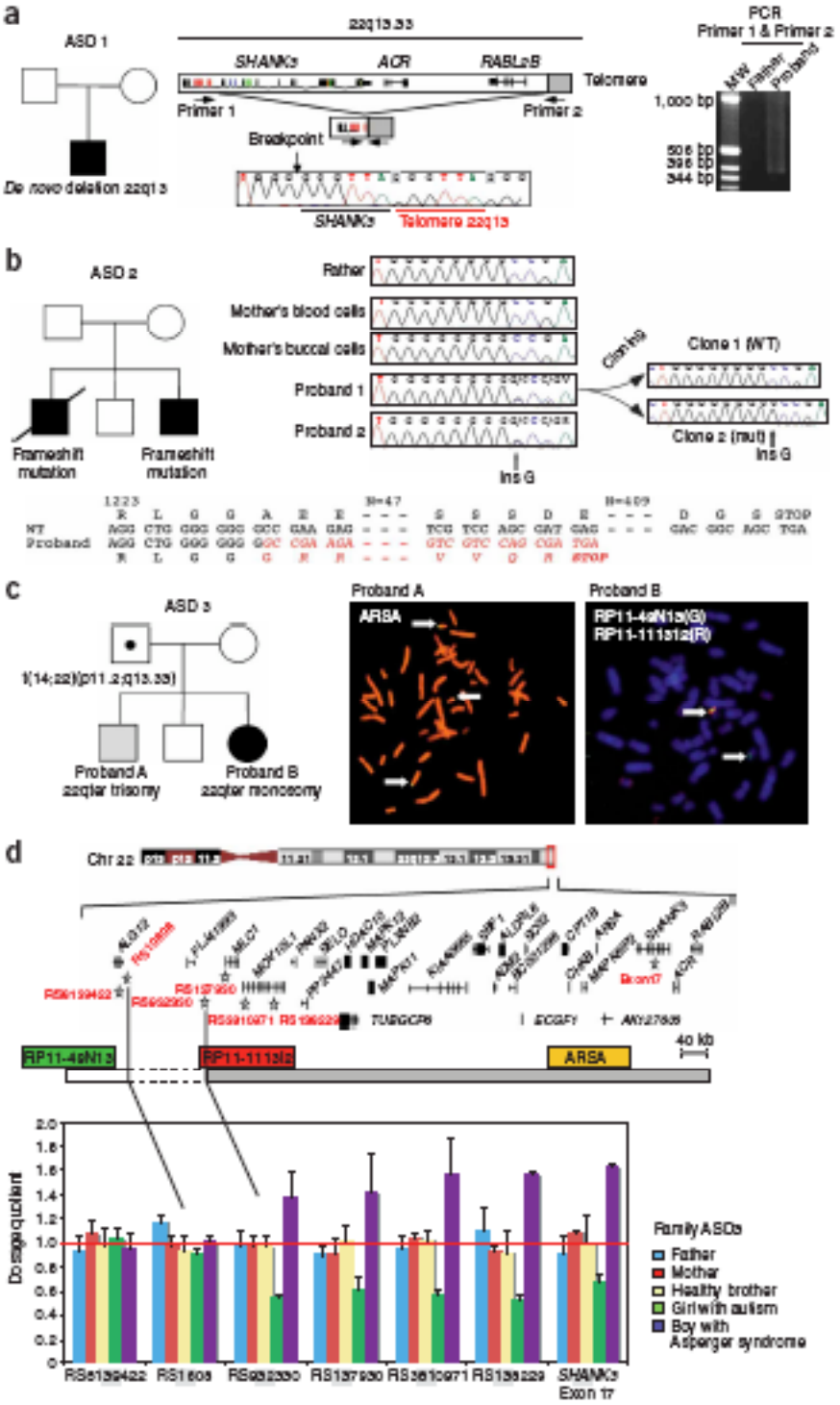


Figure 1. Genetic analyses of three families with ASD and *SHANK3* mutations. **(a)** In family ASD 1, the proband carries a *de novo* terminal deletion of the paternal chromosome 22q13. The deletion breakpoint is located in intron 8 of *SHANK3*. The breakpoint was sequenced after amplification of the proband DNA using primer 1 in *SHANK3* and primer 2 in the telomeric repeats. The heterogenous smear in the proband is likely due to the difference in telomere length from chromosome to chromosome and/or priming at different locations by the telomeric primer. **(b)** In family ASD 2, the two probands carry the same *de novo SHANK3* frame-shift mutation on the maternal chromosome 22q13. The mutation is absent from the mother blood and buccal cells, suggesting a germinal mosaicism. The guanine insertion is located in exon 21 of *SHANK3,* leading to a premature truncated protein. **(c)** In family ASD 3, the father carries a balanced translocation t(14,22)(p11.2;q13.33), proband A (Asperger syndrome) presents a partial 22qter trisomy and proband B (autism) has a 22qter deletion. **(d)** Using quantitative fluorescent PCR, we mapped the breakpoint between the genes *ALG12* and *MLC1*. The dosage quotient has a theoretical value of 0.5 for a deletion and 1.5 for a duplication.

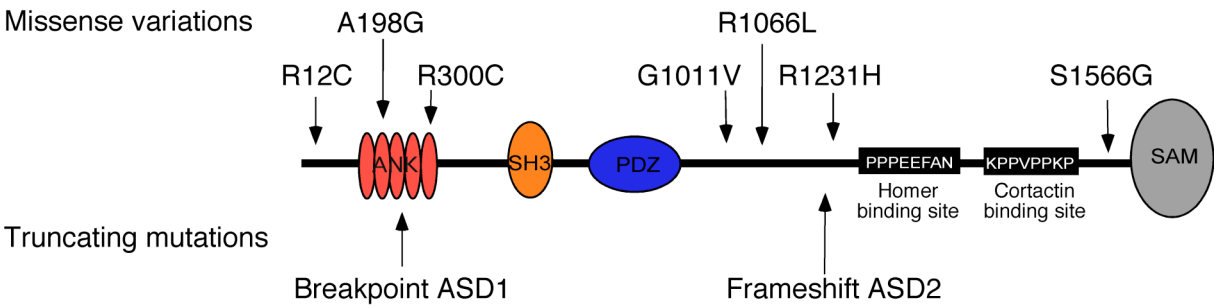


Figure 2. Localization of rare nonsynonymous variations or truncating *SHANK3* mutations identified in families with ASD. ANK: ankyrin repeats; SH3: Src homology 3 domain; PDZ: postsynaptic density 95/Discs large/zona occludens-1 homology domain; SAM: sterile alpha motif domain.

SUPPLEMENTARY MATERIAL

Materials and Methods

Families

Families (n=226) with at least one child with ASD were recruited by the PARIS (Paris Autism Research International Sibpair) study at specialized clinical centers in seven countries (France, Sweden, Norway, Italy, Belgium, Austria, and the United States); 163 families had one child with ASD and 63 families had at least two children with ASD. Diagnosis was based on clinical evaluation by experienced clinicians, DSM-IV criteria[1], and the Autism Diagnostic Interview-Revised (ADI-R)[2]. In Sweden, the Diagnostic Interview for Social and Communication Disorders (DISCO-10)[3] was used instead of the ADI-R in some cases. Like the ADI-R, the DISCO-10 utilizes algorithms based on ICD-10 and DSM-IV research criteria to diagnose autism. However, it can also be used to diagnose other pervasive developmental disorders, such as pervasive developmental disorder not otherwise specified (PDD-NOS) and Asperger syndrome. The diagnosis of Asperger syndrome was confirmed using the Asperger Syndrome Diagnostic Interview (ASDI)[4]. The sample consisted of 177 males and 50 females; there were 194 patients with autistic disorder and 29 with Asperger syndrome; 4 individuals narrowly missed the criteria for autistic disorder and were considered to have atypical autism (PDD-NOS). Laboratory tests to rule out medical causes of autism included standard karyotyping, fragile-X testing, and metabolic screening; brain imaging and EEG were performed when possible. Patients diagnosed with medical disorders, such as fragile X syndrome or chromosomal anomalies, were excluded from the study. The local research ethics boards reviewed and approved the study. Informed consent was obtained from all families. There were 206 Caucasian, 7 Black, 1 Asian and 13 families of mixed ethnicity.

The control sample (n=270) comprised 120 Caucasians from France and 150 Caucasians living in Sweden. An additional control group of 63 individuals from China was screened for rare variants because one proband carrying the R300C mutation had parents originating from this country. For exons with no mutation in ASD (2, 3, 9, 11, 13-17), 190 controls (95 from France and 95 from Sweden) were sequenced. Using this minimum sample size (n = 380 chromosomes), we can detect a polymorphism with a frequency of 1% with 95% power[5]. The controls from Sweden were part of two population-based cohorts living in Göteborg, originally recruited for a study of obesity and body fat distribution[6,7], who volunteered to provide a blood sample for genetic studies. Women (n=30) were born in 1956 and men (n=65) were born in 1944; subjects with one or two parents probably or certainly being non-Caucasian were excluded.

The controls from France (40 females and 55 males) were healthy volunteers, between 19 and 65 years old, interviewed with the Diagnostic Interview for Genetic Studies (DIGS) and the Family Interview for Genetic Studies (FIGS) to confirm the absence of both personal and family history of psychiatric disorders in first- and second-degree relatives.

Genomic structure and RT-PCR analyses of *SHANK3*

The genomic structure of *SHANK3* was deduced using the published data from Wilson *et al.* [8], the Expressed Sequence Tags (ESTs), and the rat cDNA (AJ133120) published in databases. *SHANK3* transcripts were detected in human brain regions from three independent controls (two females and one male) and in human tissues using the Clontech cDNA panel (Clontech Laboratories Inc). Total

RNA was isolated from brain tissues by the acid-guanidium thiocyanate phenol chlorophorm method and reverse transcribed by oligodT priming using SuperScript™ II Reverse Transcriptase (Invitrogen). Primer sequences for RT-PCR are indicated in Supplementary Table 2. Before sequencing the alternatively spliced exons, the breakpoint in family ASD 1 and the frame-shift mutation in family ASD 2, the PCR products were cloned with TOPO–TA cloning kit (Invitrogen). Prediction of the Quadruplex forming G-Rich Sequences (QGRS) at chromosome 22q13.3 was performed using the GQRS mapper software (http://bioinformatics.ramapo.edu/QGRS/index.php)[9].

Mutation analysis

DNA was extracted from blood leukocytes or B lymphoblastoid cell lines with the phenol chlorophorm method[10]. In the mother of family ASD 2, DNA from buccal cells was extracted using the BuccalAmp DNA extraction kit (Tebu-bio). For mutation analysis, the 24 coding exons of *SHANK3* were amplified from genomic DNA with specific primers (Supplementary Table 2). Amplification was performed on 20 ng of DNA template with HotStar Taq polymerase (Qiagen) for all exons except for exons 1, 11, 21, 22, 22b, and 22c, for which amplification was performed with Taq polymerase from Eurobio and 10% GC melt (Clontech GC rich kit). Two PCR protocols were used: (i) Standard protocol: 95° for 15 min, followed by 35 cycles at 95°C for 30 sec, 55 to 65°C for 20 sec, 72°C for 30 sec to 1 min, with a final cycle at 72°C for 10 min; and (ii) Touchdown protocol: 95°C for 15 min followed by 20 cycles at 95°C for 30 sec, 70°-60° or 65-55°C for 30 sec, and 72°C for 30 sec, followed by 20 cycles at 95°C for 30 sec, 60° or 55°C for 10 sec, and 72°C for 30 sec, with a final cycle at 72°C for 10 min. Sequence analysis was performed by direct sequencing of the PCR products, using a 373A automated DNA sequencer (Applied Biosystems).

Detection of 22q13/*SHANK3* deletions and duplications

For quantitative analysis, the forward primer was labeled with fluorescent 6-carboxyfluorescein (6-FAM). PCR amplification (25 cycles) was as described above for standard protocol. Fluorescence-labeled PCR products were run on a 373A automated DNA sequencer (Applied Biosystems) with GENEFLO 625 DNA Ladder, ROX l (EurX). Following data collection, samples were analyzed with Genescan 3.7 software program. Two independent *SHANK3* PCRs (exon 9 and exon 17) were compared to two autosomal control genes *NLGN1* and *ANKRD15*, located on chromosome 3q26.31 and 9p24.3, respectively. The peak ratio of each PCR was used to calculate the dosage quotient (DQ) value. DQ= (PCR *SHANK3*/PCR in control region) in tested individual / (PCR *SHANK3*/PCR in control region) in control individual. Thus, DQ gives a theoretical value of 0.5 for a deletion and 1.5 for a duplication. A total of 155 individuals (58 with autism, 38 with Asperger, and 59 controls) were screened for deletion/duplication of *SHANK3*. All primers are indicated in supplementary table 2.

***In vitro* mutagenesis and transfection studies in hippocampal neurons**

Full-length rat *Shank3* cDNA (sequence AJ133120) was cloned into a pEGFP-C2 vector (Clontech). Mutagenesis was made using the QuickChange II XL site directed mutagenesis kit (Invitrogen) on 100 ng of wild type plasmid. Each clone was purified with Endofree Plasmid Maxi kit (Qiagen) and entirely sequenced to rule out additional mutations in the *Shank3* cDNA. Moreover, the GFP constructs were transfected into Cos cells and analyzed by Western Blot using a GFP antibody. The preparation of rat hippocampal cultures and the Shank3 localization experiments were performed

essentially as described previously[11,12]. Neurons were transfected after 14 days in culture. For co-localization of transfected Shank3 full length construct or the mutated constructs with the pre- and postsynaptic marker proteins SAP-90/PSD95, synaptophysin or Bassoon[13], cells were fixed on day 17 in 4% paraformaldehyde for 20 min at room temperature (20°C). Secondary antibodies conjugated to Cyanine 3 (Cy3) fluorophore were used (anti mouse Cy3 , anti-rabbit Cy3, Chemicon, Temecula, CA, USA), and the cells were visualized by fluorescence microscopy. All animal experiments were performed in compliance with the guidelines for the welfare of experimental animals issued by the Federal Government of Germany, the National Institutes of Health and The Max Planck Society.

Quantitative and qualitative measurement of GFP expression

Quantitative and qualitative measurement of GFP expression was performed essentially as described by Craven and Bredt[14]. In brief, transfected neurons were chosen at random from three independent transfections of each construct (4-5 cells per construct). The perimeter of the dendrites was traced (excluding spine heads/clusters), and the average pixel intensity was calculated. Similarly, the spine heads/clusters were traced, and the average pixel intensity was obtained. The ratio of the average pixel intensity in dendrites *vs*. spine heads/clusters was defined as the synaptic clustering ratio (SCR): a ratio of zero indicates complete synaptic clustering whereas a ratio of ≥ 0.7 indicates diffuse dendritic fluorescence with no clusters. Consecutively, we determined the synaptic localization of the clustering constructs by immunostaining with an antibody directed against the presynaptic marker protein Bassoon[13]. Between 250 and 300 synapses per construct were analyzed and the percentage of co-localization between Bassoon and the different GFP-constructs was determined (P/B in percentage).

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 4th Ed.*, (American Psychiatric Press, Washington D.C., 1994).
2. Lord, C., Rutter, M. & Le Couteur, A. *J Autism Dev Disord* **24**, 659-685 (1994).
3. Wing, L., Leekam, S.R., Libby, S.J., Gould, J. & Larcombe, M. *J. Child Psychol. Psychiatry* **43**, 307-25 (2002).
4. Gillberg, C., Rastam, M. & Wentz, E. *Autism* **5**, 57-66 (2001).
5. Collins, J.S. & Schwartz, C.E. *Am J Hum Genet* **71**, 1251-2 (2002).
6. Rosmond, R., Lapidus, L., Marin, P. & Bjorntorp, P. *Obes Res* **4**, 245-52 (1996).
7. Rosmond, R. & Bjorntorp, P. *Obes Res* **6**, 338-45 (1998).
8. Wilson, H.L. et al. *J Med Genet* **40**, 575-84 (2003).
9. Kikin, O., D'Antonio, L. & Bagga, P.S. *Nucleic Acids Res* **34**, W676-82 (2006).
10. Sambrook, J. & Russell, D.W. *Molecular Cloning, a Laboratory Manual*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2001).
11. Boeckers, T.M. et al. *J Neurochem* **92**, 519-24 (2005).
12. Dresbach, T. et al. *Mol Cell Neurosci* **23**, 279-91 (2003).
13. tom Dieck, S. et al. *J Cell Biol* **142**, 499-509 (1998).
14. Craven, S.E. & Bredt, D.S. *J Biol Chem* **275**, 20045-51 (2000).

Supplementary note 1. Clinical cases

In ***family ASD 1***, the proband is the only child of healthy unrelated parents. Family history is negative for mental retardation and ASD. He was born at 41 weeks after a pregnancy marked by bleedings. Birth growth parameters were large (weight 4850 g, length 54 cm, head circumference [HC] 36 cm). He exhibited neonatal hypotonia, poor suction and cried a lot as a baby. Early motor milestones were normal. He said a few words such as mama and dada but stopped using them around 30-36 months, when his parents noted a regression, both for language and play behavior. Recurrent otitis media was observed during early infancy and he had sleep difficulties. There was no history of seizures. Psychiatric evaluation at the age of 9 years showed total absence of language and moderate mental retardation. Non verbal communication was impaired, as well as social interaction, and he had motor stereotypies. The clinical diagnosis of autism was confirmed with the ADI-R. Physical examination at the age of 12 years showed no minor dysmorphic features except large ears and elbow extension limitation. Weight was 44 kg (+1.5 SD), height was 143 cm (–0.5 SD), and HC was 54.5 cm (average). Neurological examination was normal. A standard karyotype and fragile X testing were normal. Routine screening of the 22q13 region by FISH revealed a *de novo* terminal deletion. An MRI and EEG were performed and were normal.

In ***family ASD 2***, the probands are the first and the third children of three. The unaffected brother meets DSM-IV criteria for ADHD but has no autism spectrum symptoms and no language delay. The mother meets DSM-IV criteria for social phobia. Family history is otherwise negative for mental retardation, language disorders, ASD and other psychiatric disorders. Both children were born from full term and uneventful pregnancies. Birth parameters were average. Neonatal history was limited to neonatal jaundice and no hypotonia was recorded. The eldest boy had a normal development during the first year but was very agitated. He started walking at 16 months. He spoke his first words at 5 years and first sentences at 6 years, but never acquired functional language. At the age of 20 years, after he was transferred to a new institution, a regression with severe outbursts was noted, he lost skills including cleanliness, and he developed anorexia and marked weight loss. At that time he had a seizure-induced aspiration, he was hospitalized and died a few days later. The second affected brother had normal measures at birth (weight 3420 g, length 50 cm, HC 35.5 cm) but he showed progressive macrocephaly until 9 months of age (+2 SD), followed by slowed growth (HC + 1 SD). He had limited eye contact since birth. He started walking at 18 months. He never acquired words, except mama and dada. At the age of 16 years he had an episode of aspiration with loss of consciousness, which necessitated hospitalization for a few days, and a second episode at the age of 20. Since then, he is fed a special diet with no liquids and soft food. When he was 16, after being transferred to the same institution where his brother had been, a regression was also noted, with loss of autonomy and cleanliness, and marked weight loss. He developed epilepsy at the age of 17 and was started on clonazepam. Physical examination at the age of 20 years showed growth parameters in the average (height 175 cm, HC 55.5 cm), strabismus and no dysmorphic features. Both brothers had severe language impairment, severe mental retardation and marked deficits in social interaction. The clinical diagnosis of autism was confirmed for both of them with the ADI-R.

In ***family ASD 3***, the probands are the first and the third born of three children. The unaffected brother has no history of psychiatric disorder. The mother fulfills DSM-IV criteria for agoraphobia; the father used to have symptoms of ADHD as a child but doesn't fulfill the criteria as an adult. There

is no family history of mental retardation or ASD, but the paternal grandfather has numerous rituals and routines and is socially isolated. The eldest boy was born full term (41 weeks of gestation) after an uneventful pregnancy, with large birth parameters (weight: 4040 g, length: 53 cm, HC: 38 cm). He had normal early motor milestones, said his first words at 12 months, first sentences at 18 months and already knew the letters at that time. He learned to read at 5 years, before attending primary school. He has impairment in social interactions, is often awkward in social situations and is quite isolated at school. He doesn't like to be observed by teachers and has marked difficulties to speak in class. He also has difficulties to identify social cues and to understand others people's reactions. His interests are focused on mechanical sports and he has great knowledge about this. He has excellent memory for car plates and is very rapid for mental calculation. He also has routines, especially when taking a shower. He has impairments in the interpretation of implied meanings. He has a small repertoire of facial expressions and limited use of gestures. Direct eye gaze is rare. His voice is often too loud, lacking modulation. He is quite clumsy and awkward. He has mood swings, hypersensitivity to noise, and phobia of chimney fire. Clinical evaluation at the age of 14.5 years confirmed the diagnosis of Asperger syndrome and showed a normal IQ. Physical examination at this age showed increased height (184 cm) and marfanoid habitus, without significant dysmorphic features. The youngest sister was born at 38 weeks of gestation, after an uneventful pregnancy. Birth parameters were low (weight: 2050 g, height: 47.5 cm, HC: 33 cm). She had discrete neonatal hypotonia but had normal early motor milestones. Language development was severely delayed with few words around 4 years and development of very short sentences before regression. Around 5 years old, she presented echolalia. She had motor stereotypies and exhibited some routines and need for sameness. She is now 12 years old and uses few words, with pronunciation errors. She communicates with pictograms. She went to school until 7 years old and now attends a day hospital. An IQ evaluation at 12.9 years showed a performance IQ of 44 (developmental age 5.8 years), and a verbal IQ of 15 (developmental age 24 months). Her non verbal communication is impaired, as well as her social interaction. At present she initiates activities by herself. She especially enjoys puzzles, she is very fast at them and can spend the whole day on this activity. She also plays with construction games and has a tendency to align the pieces. She still shows insistence on sameness. The diagnosis of autism was confirmed with the ADI-R. She has some self-aggressive behaviors, hitting her head. She has no hypersensitivity to noise and her reaction to pain is normal. She presents nocturnal enuresis. Fine and global motor skills are impaired. Physical examination at 10 years showed average growth parameters, without dysmorphic features. A brain MRI at 5 years of age was normal. The terminal 22q deletion in the girl was diagnosed by FISH when she was 8 years old. Study of other family members revealed a translocation t(14;22)(p11.2;q13.33) in the father and a 22qter partial trisomy in her brother.

The proband of ***family ASD 4*** carries the R12C SHANK3 mutation. He is the third child of healthy unrelated parents. His two brothers are healthy. Medical and psychiatric history of the maternal family includes Parkinson disease in both maternal grandparents, depression in the grandmother and thyroid disease in the grandmother, mother and several aunts. The mother had a major episode of depression during pregnancy and after the birth of the proband. In the father's family, two brothers of the grandmother committed suicide and two paternal uncles have a history of alcohol abuse. There is no history of mental retardation, language disorder or ASD. The proband was born after a full term and uneventful pregnancy. Delivery was unremarkable and birth parameters were in the average (weight 3750 g, length 51 cm, HC 35 cm). He didn't establish eye contact and early

milestones were delayed: he sat at 18 months and started walking at 4 years. EEG and brain imaging by tomodensitometry (TDM) were performed during his first year and were both normal. Vision was checked and was not impaired. Physical examination at the age of 23 years showed a narrow face, large ears, long philtrum, saddle nose, flattened midface, thin superior lip, and prominent supraorbital ridge. Height was 160 cm (-2 SD), weight 60 kg, and HC 55.3 cm. Neurological examination was normal and there was no history of seizures. There was a high pain threshold and low noise threshold. Psychiatric examination showed total absence of language and severe mental retardation. Non verbal communication was impaired, as well as social interaction, and he had motor stereotypies. There was no sleep disorder or hyperactivity. The clinical diagnosis of autism was confirmed with the ADI-R.

The proband of ***family ASD 5*** carries the R300C *SHANK3* mutation, which he inherited from his mother. He is the first of two children from non-consanguineous healthy parents of Chinese origin. His family history is non-contributory. The patient was born at term, after an uncomplicated pregnancy. The first symptoms were noticed at 2 years, when he lost some words (he said mama, dada, bye-bye, and could name certain things he wanted) and exhibited unusual eye contact. From this time, he showed no interest in other children and preferred to play on his own. He had sleeping problems during childhood. He showed repetitive behaviors and circumscribed interests, opening and closing doors, lining up objects, finding similar objects, leafing through certain books, or letting sand sift through his fingers. When evaluated at 42 months, he had no expressive verbal language, but used different sounds and song melodies. He had some receptive language; he could point out some familiar objects and was able understand some familiar sentences. He fulfilled the criteria for autistic disorder according to the ADI-R and had a Childhood Autism Rating Scale (CARS) score of 37.5, indicating severe autism. His developmental age at that time, evaluated with the Psycho-Educational Profile-Revised (PEP-R) was 18-20 months. No somatic pathology was detected at that time. A standard karyotype, fragile X testing and metabolic screening were normal. A brain CT scan performed at this time was also normal. Three EEGs, performed at 3, 7 and 10 years, did not show significant pathology. He started to speak at 5-6 years, but his language problems persisted. At the age of 10 years, his language was stereotyped and repetitive; his adaptive language function was below mental age of 3 years according to the Vineland Adaptive Behavior Scales. A Wechsler Intelligence Scale for Children (WISC) performed at 10 years showed a total IQ of 67, a verbal IQ of 63, and a performance IQ of 81. On examination at 11 years 4 months, his weight was 38 kg, his height 144.5 cm, and HC 55 cm. Neurological exam showed generalized hypotonia and brisk reflexes throughout all extremities. He had long arms, with narrow hands and long fingers. He had a supernumerary upper incisor (mesiodens) and he still had his milk teeth in his upper jaw except for the two front teeth. No obvious dysmorphic features were observed. His parents noted that the patient often took off his clothes and complained about being warm (patients with 22 deletion syndrome have a tendency to become overheated due to reduced perspiration[1]).

Reference

1. Phelan, M.C. et al. *Am J Med Genet* **101**, 91-9 (2001).

Supplementary Table 1. *SHANK3* variations identified in families with ASD and controls.

	Variations		Sex of proband	Transmission and ethnicity			Additional family information	Occurrence	
	Exon	Change[a]		Parents	Siblings	Ethnicity		ASD (n=227)	Controls[b] (n=190-333)
ASD									
	E1	R12C	Male	Mother	Present in one unaffected brother and absent in one unaffected brother	European	The mother has social phobia traits; the father and the two brothers are healthy	1	0
	E5	A198G	Female	Mother	Absent in the affected brother	European	No history of psychiatric disorders in either parent	1	0
	E8	R300C	Male	Mother	Absent in the unaffected sister	Asian	Language delay in the father; the mother is healthy	1	0[c]
	E21a	G1011V	Female	Mother	(3 brothers, DNA not available)	European	No history of psychiatric disorders in any family member	2	0
	E21a	G1011V	Male	Father	Present in the affected monozygotic twin	European	No history of psychiatric disorders in either parent		
	E21a	R1066L	Male	Mother	Present in the unaffected sister	North African	The father has a depressive syndrome; the mother had depression during pregnancy	1	0
	E21a	ins(G)3680 at amino acid 1227	Male	*De novo* (germinal mosaicism in the mother)	Present in the two affected brothers and absent in the unaffected brother	European	No history of psychiatric disorders in either parent	1	0
	E21a	R1231H	Female	Mother	Absent in the affected brother	European	No history of psychiatric disorders in either parent	1	0
	E22b	S1566G	Male	Father	(2 unaffected siblings, DNA not available)	European	No history of psychiatric disorders in any family member	1	0
ASD and Controls									
	E21a	A963G	Female	Mother	(1 healthy brother, DNA not available)	European	No history of psychiatric disorders in any family member	2	1
	E21a	A963G	Male	Father	(1 brother, DNA not available)	European	The father has broad autism phenotype (very shy, some autistic traits); the younger brother has mental retardation and autistic traits (not autism)		
	E22c	P1654T	Male	Father	Absent in the affected brother	European	The father talked late and has stereotyped body movements; the mother also was late to talk.	2	1
	E22c	P1654T	Male	Nd	(1 sister, DNA not available)	European	No history of psychiatric disorders in either parent or the sister		
Controls									
	E6	A224T	Female	—	—	European		0	1
	E21a	V1333G	Male	—	—	European		0	2

[a] *SHANK3* cDNA sequence: NM_033517.1. [b] Exon 8 was sequenced in 333 controls; exons 1, 4-7, 10, 12, 18-20, 21a, and 22abc were sequenced in 270 controls; and exons 2, 3, 9, 11, and 13-17 were sequenced in 190 controls. Nd = not determined.

Supplementary Table 2. Primers used in this study

Primers for mutation screening

Forward primer (5'-3')		Reverse Primer (5'-3')		Size (bp)	Condition
SHANK3E1F	GCGCTCCGTTCCCCGGCGCGA	SHANK3E1R	CCTCCGCGAACCGCGGCCGAA	325	70°C + GC melt 10%
SHANK3E2F	GACCTGAGCTCACGAGCCCGCT	SHANK3E2R	CTGCCGTGCCCTTCACTGGTC	322	55°C
SHANK3E3F	TCCACTGTGGTAGTATGACTG	SHANK3E3R	TGGAACACATCACTGTACCAC	435	57°C
SHANK3E4F	GAGGAAGGCGGGTGATGTTCA	SHANK3E7R	AGTATATCCACACTCGGTGCA	1047	TD 66-56°C
SHANK3E8F	GTGTGCATTCCTGTGTGCGCA	SHANK3E8R	GGCTTCTGCACCCCTGCTGGT	485	TD 66-56°C
SHANK3E9F	ATCCCAGTTACAGACAAGAGT	SHANK3E9R	CAATGTTCACTCAACACAGGC	450	55°C
SHANK3E10F	GCCTGGGCAAACTGGACAAGT	SHANKE10R	TCCCCAACCAGGAAGCCCTAG	500	55°C
SHANK3E11F	GGCATCGCGTCCGTCACCTACGT	SHANK3E11R	AAGGTTCCTGCCGTGCGGGT	415	55°C + GC melt 10%
SHANK3E12F	AGCTGGGAGAAAGTGGGAAGG	SHANK3E12R	GTCACACACGTCCTATGTGTC	500	TD 65-55°C
SHANK3E13F	ACTGGTGACCAGCATGGGTGA	SHANK3E14R	GGCAGAAGCAAGAAGCTGAAG	560	TD 65-55°C
SHANK3E15F	CGCAGTCATCTCTTTCCTGTG	SHANK3E16R	GGCAGAAGCAAGAAGCTGAAG	589	55°C
SHANK3E17F	ACCTGAACAAGATCCTGGCAC	SHANK3E17R	CACCCATTCACCTCTGACCTG	390	55°C
SHANK3E18F	GCATGTACCAACTGACTCCAG	SHANK3E18R	CCATCACAGTCTCAGAGGGTC	500	55°C
SHANK3E19F	CTCTGTCAGCATCACGGGTG	SHANK3E20R	CTTCAAACCCAAGTCCACCCT	1010	61°C
SHANK3E21AF1	AAGGCTGGCCTCTGTGGGAGG	SHANK3E21AR1	ACGGACAGGAACACAGTGGAG	825	TD 70-60°C + DMSO 5%
SHANK3E21AF2	CAAGAGCCCCCTGGTGAAGCA	SHANK3E21AR2	GCTCTCGGGCAGCCAGGGCAA	300	TD 70-60°C + DMSO 5%
SHANK3E21AF3	TGCCCTGAAGCCGTTGGTCAG	SHANK3E21AR3	ACCTTCTCTGCCTCCCTGCGA	300	58°C
SHANK3E21AF4	CCTGTTTGTGGATGTACAGGCC	SHANK3E21AR4	ACCGTGGAGATGGTGCTTGTG	716	55°C + GC melt
SHANK3E21AF5	CTCAGGGAAGCCCAGCAGTGA	SHANK3E21AR5	AGAACAGACAAGAGGAATGAC	668	TD 70-60°C
SHANK3E21BF	GAGGAGCCCTTCGGGCCCGTG	SHANK3E21BR	GGCAGAGAAGAGCGGAGGGAG	641	TD 70-60°C
SHANK3E21CF	GTCTCGAAGGGAAACATGAAC	SHANKE21CR	CAGTGTCCATGTCTGACTTCC	431	55°C
SHANK3E22F1	CCGTAGGATCCCACCCTTTA	SHANK3E22R1	GCCTAGGTGGATGCTCTCCAG	500	60°C
SHANK3E22F2	CTTCGTGGTGCGCAGCGTGAG	SHANK3E22R2	ACAGCAAACAGGACGATTCA	485	TD 65-55°C + DMSO 5%

Primers for expression

Forward primer (5'-3')		Reverse Primer (5'-3')		Size (bp)	Condition
SHANKEX18F	CTGCGCTCCAAGTCCATGACA	SHANK3EX18R	GGCCCTGGCGTTCAAACAATG	219, 198	55°C
SHANK3EX21F	GCCTGAAGACGACAAACCAA	SHANK3EX21R	GAGCTGCAGCGGCTTCTGCTG	571, 403, 321	TD 70-60°C + DMSO 5%

Primers for *in vitro* mutagenesis

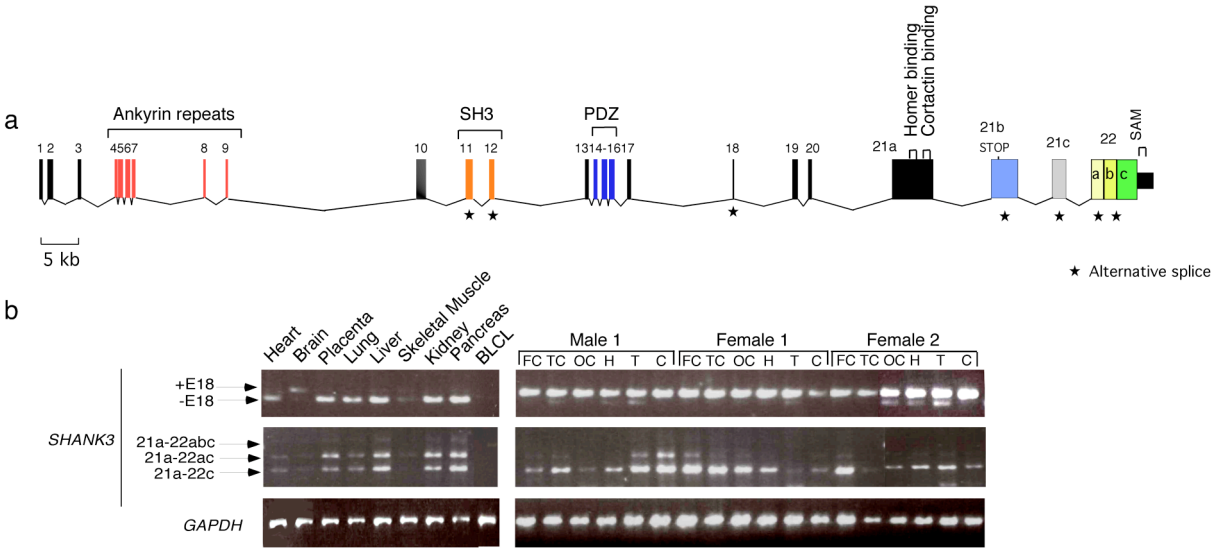
Forward primer (5'-3')		Reverse Primer (5'-3')	
SHANK3R12CF	GGGCCAGCGCCGTGGTCGTGTGCGTCGGCATCCCGGAC	SHANK3R12CR	TCCGGGATGCCGACGCACACGACCACGGCGCTGGCCC
SHANK3R300CF	ACCAGGAGAGCTGTGCCTGCGTCCTGCTTTTCCGTGG	SHANK3R300CR	CCACGGAAAAGCAGGACGCAGGCACAGCTCTCCTGGT
SHANK3INSF	GAGCCCAACAGGCTGGGGGGCTGAAGAGGAGCGCC	SHANK3INSR	GGCGCTCCTCTTCAGCCCCCCAGCCTGTTGGGCTC

Primers for telomere-*SHANK3* PCR

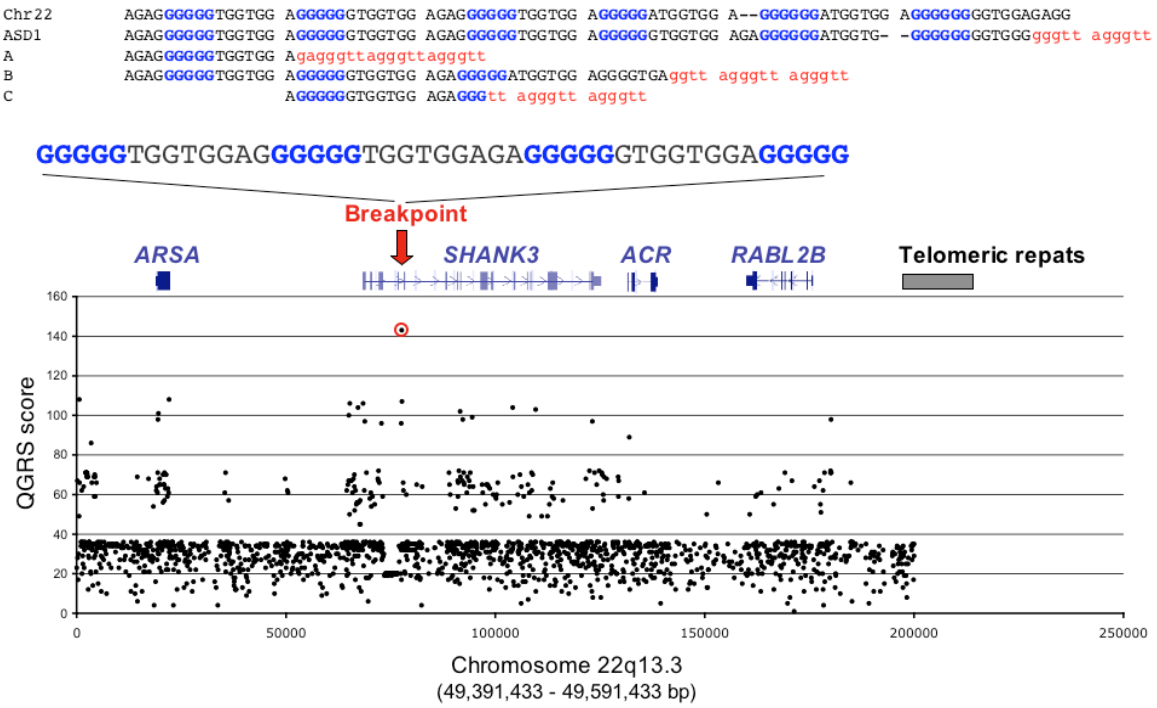
Forward primer (5'-3')		Reverse Primer (5'-3')	
BP22QF	GTGACTTGACTTCTCTGAACCTTGG	BP22QR	TATGGATCCCTAACCCTAACCCTAACCC

Primers for quantitative PCR

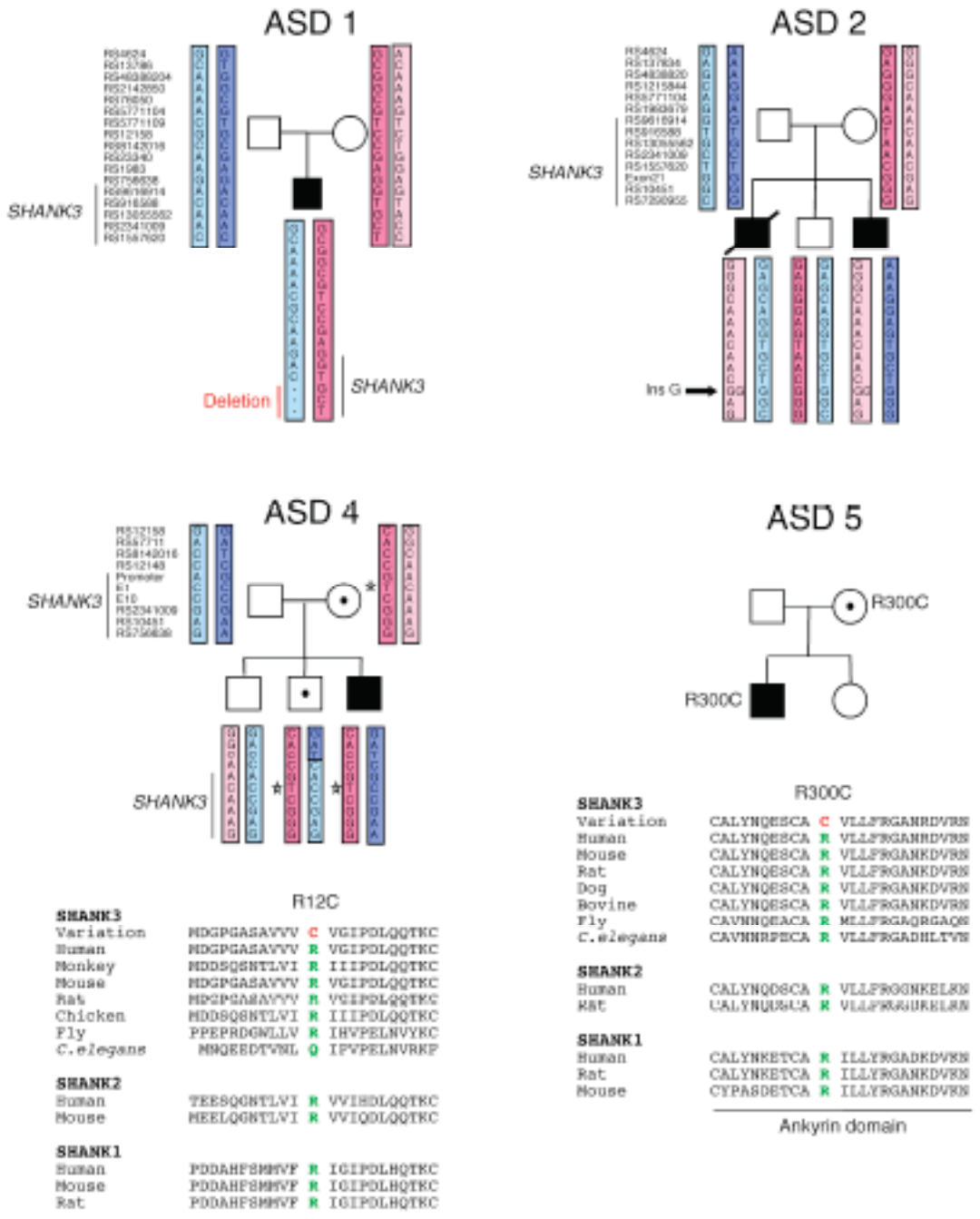
Forward primer (5'-3')		Reverse Primer (5'-3')	
SHANK3E9F-FAM	mATCCCAGTTACAGACAAGAGT	SHANK3E9R	CAATGTTCACTCAACACAGGC
SHANK3E17F-FAM	mACCTGAACAAGATCCTGGCAC	SHANK3E17R	CACCCATTCACCTCTGACCTG
NLGN1F-FAM	mCCTTGATTAATACAGGCTTCA	NLGN1R	ATACAGAGCATCACATACTAC
ANKRD15F-FAM	mGAAGAACTAACGACCACTTG	ANKRD15R	TCTCTGAGATGAGTCACAACC



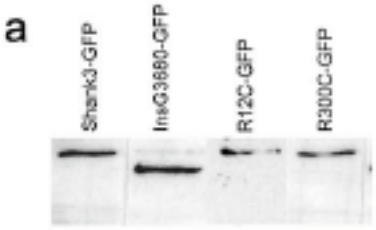
Supplementary Figure 1. Genomic structure and mRNA expression of the human *SHANK3* gene. **(a)** Genomic structure of the *SHANK3* gene, which is located on chromosome 22q13.3, spans 57 kb and contains 24 exons. The mouse gene on chromosome 15qE3, the rat gene on chromosome 7q34, and the human gene share the same genomic structure. **(b)** Specific RT-PCRs were performed on total RNA from different brain regions and cDNA libraries (Clontech) to detect *SHANK3* mRNA and the alternative splice of exon 18, exon 22a, b and c. The age of the male and the two females studied were 74, 55, and 36 years, with a post-mortem delay of 10, 24 and 2 h, respectively. FC, frontal cortex; TC, temporal cortex; OC, occipital cortex; H, hippocampus; T, thalamus; C, cerebellum. Normal control human brains were obtained at autopsy under guidelines approved by the ethics committee. BLCL: B lymphoblastoid cell lines.



Supplementary Figure 2. Genomic sequence of the deletion breakpoint in family ASD 1 and prediction of quadruplex forming G-rich sequences (QGRS) at the terminal end of chromosome 22q13. The breakpoint identified in family ASD 1 is similar to those previously identified in probands A, B, C from Bonaglia *et al.*[5] The sequence surrounding the breakpoint is recognized as a QGRS and has the best prediction score (143, indicated by a red circle) out of the 1700 GQRS identified. This structure may represent a substrate for telomere formation, thereby increasing the risk of recurrent deletions in this region. Sequence was extracted from the Human March 2006 (hg18) assembly. Prediction was performed using the QGRS mapper software (http://bioinformatics.ramapo.edu/QGRS/index.php).

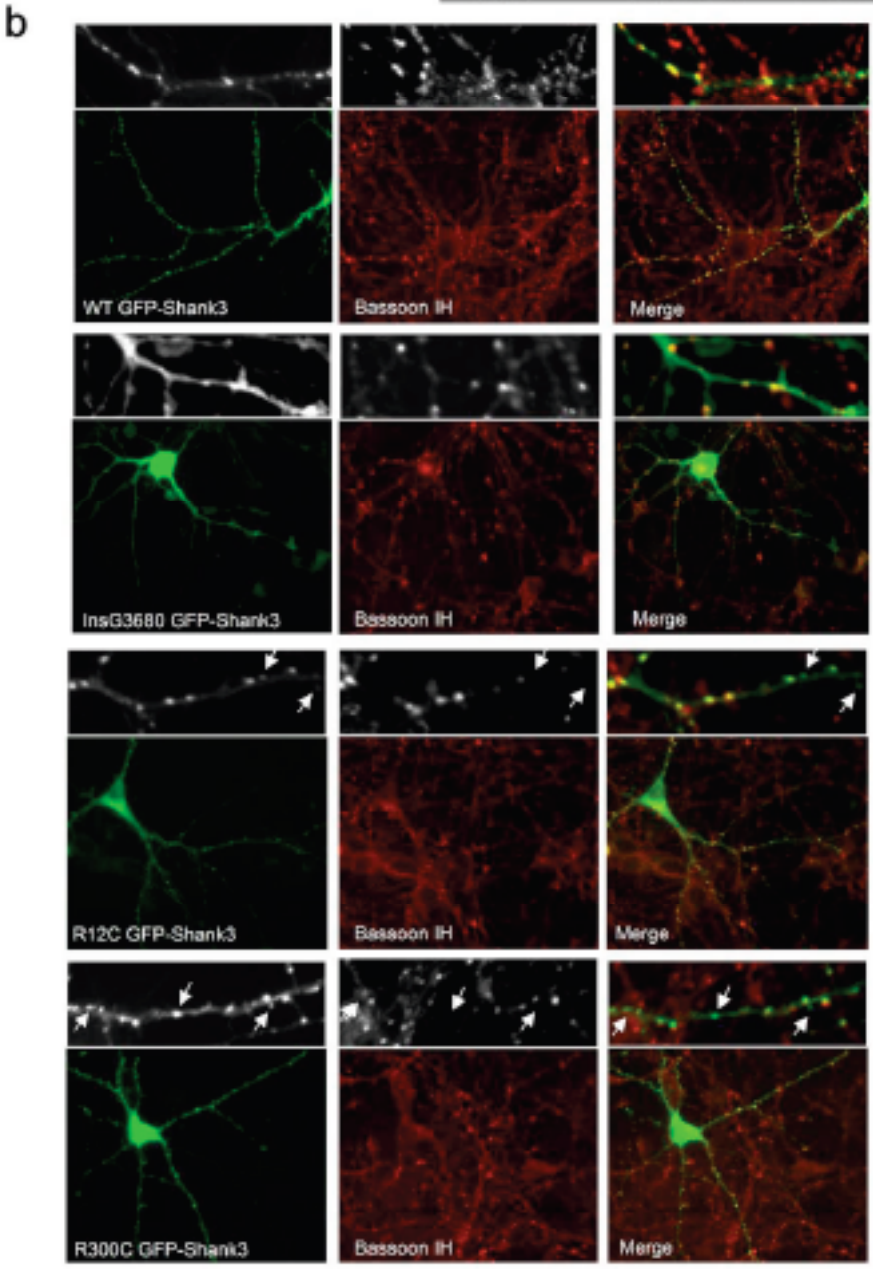


Supplementary Figure 3. Pedigree structure, haplotype analyses and conservation of the *SHANK3* mutations and variants identified in individuals with autism. In family ASD 1, the proband carries a de novo 22q13 deletion on the paternal chromosome. In family ASD 2, the two affected siblings carry a G insertion on the maternal chromosome, originating from a germinal mosaicism. The insertion in exon 21 of *SHANK3* leads to a premature truncated protein. The proband of family ASD 4 carries the R12C *SHANK3* mutation, transmitted by the mother and shared with his healthy brother. The study of 10 SNPs revealed that the two brothers carrying the R12C variation don't share the same paternal allele of *SHANK3*. The proband of family ASD 5 carries the R300C *SHANK3* mutation, transmitted by the mother, located in the ankyrin domain. The promoter region, the 5'UTR and the 3'UTR of *SHANK3* were sequenced in the patients ASD 4 and ASD 5, but no additional variations were identified.



c

Shank3	SCR	S/B
WT	0.21 ± 0.09	87.6 ± 5.7%
InsG3680	0.8 ± 0.08	no localization
R12C	0.32 ± 0.1	69.4 ± 16.5%***
R300C	0.35 ± 0.05	64.4 ± 10.7%***



Supplementary Figure 4. Analyses of SHANK3 mutations in rat hippocampal neuronal cultures. **(a)** Western Blot analysis of the GFP constructs revealed similar sizes of Shank3 WT and fusion proteins carrying point mutations (R12C, R300C). The frame-shift mutation (InsG3680) results in a truncation of the protein. **(b, c)** Compared to the WT protein, the truncated Shank3 molecule (missing the C-terminal SAM domain), is evenly distributed in dendrites and does not cluster at synapses (synaptic clustering ratio, SCR, 0.8) as revealed by immunostaining (IH) against the presynaptic marker molecule Bassoon. In contrast, the two other mutated Shank3 proteins have the potential to cluster (SCR ≤ 0.35). However, only about one third of these clusters are Bassoon positive (S/B = Shank3/Bassoon ratio, arrows, ***≤ 0.001).