European Journal of Human Genetics (2013) 21, 310–316

www.nature.com/ejhg

ARTICLE

Prevalence of *SHANK3* variants in patients with different subtypes of autism spectrum disorders

Luigi Boccuto[1,2], Maria Lauri[2], Sara M Sarasua[1], Cindy D Skinner[1], Daniela Buccella[2], Alka Dwivedi[1], Daniela Orteschi[2], Julianne S Collins[1,*], Marcella Zollino[2], Paola Visconti[3], Barb DuPont[1], Danilo Tiziano[2], Richard J Schroer[1], Giovanni Neri[2], Roger E Stevenson[1], Fiorella Gurrieri[2] and Charles E Schwartz*[,1]

Autism spectrum disorders (ASDs) include three main conditions: autistic disorder (AD), pervasive developmental disorder, not otherwise specified (PDD-NOS), and Asperger syndrome. It has been shown that many genes associated with ASDs are involved in the neuroligin–neurexin interaction at the glutamate synapse: *NLGN3*, *NLGN4*, *NRXN1*, *CNTNAP2*, and *SHANK3*. We screened this last gene in two cohorts of ASD patients (133 patients from US and 88 from Italy). We found 5/221 (2.3%) cases with pathogenic alterations: a 106 kb deletion encompassing the *SHANK3* gene, two frameshift mutations leading to premature stop codons, a missense mutation (p.Pro141Ala), and a splicing mutation (c.1820-4 G>A). Additionally, in 17 patients (7.7%) we detected a c.1304+48C>T transition affecting a methylated cytosine in a CpG island. This variant is reported as SNP rs76224556 and was found in both US and Italian controls, but it results significantly more frequent in our cases than in the control cohorts. The variant is also significantly more common among PDD-NOS cases than in AD cases. We also screened this gene in an independent replication cohort of 104 US patients with ASDs, in which we found a missense mutation (p.Ala1468Ser) in 1 patient (0.9%), and in 8 patients (7.7%) we detected the c.1304+48C>T transition. While *SHANK3* variants are present in any ASD subtype, the SNP rs76224556 appears to be significantly associated with PDD-NOS cases. This represents the first evidence of a genotype–phenotype correlation in ASDs and highlights the importance of a detailed clinical-neuropsychiatric evaluation for the effective genetic screening of ASD patients.
European Journal of Human Genetics (2013) **21,** 310–316; doi:10.1038/ejhg.2012.175; published online 15 August 2012

Keywords: autism; *SHANK3*; PDD-NOS; screening; genotype–phenotype correlation

INTRODUCTION

Autism (MIM 209850) is a neurodevelopmental disorder that includes impairment of social interactions, impairment of verbal and nonverbal communications, restricted educational activities, abnormal interests, and stereotypic behaviors.[1] The phenotype includes a constellation of conditions that have been classified as autism spectrum disorders (ASDs), or pervasive developmental disorders. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) of the American Psychiatric Association,[1] the main ASD conditions are: (1) autistic disorder (AD), (2) Asperger syndrome, and (3) pervasive developmental disorder, not otherwise specified (PDD-NOS).

Overall, ASDs have a prevalence of approximately 1% in the US population under eight years of age,[2] with a male to female ratio of 4:1. Although the etiology is heterogeneous, the calculated heritability for ASDs is about 90%, with a 70% concordance for AD in monozygotic twins or 90% concordance if a broader definition of ASDs is used.[3–5] Recent papers report a genetic diagnosis in 10–40% of patients affected with ASDs.[6,7] New molecular techniques, such as array comparative genomic hybridization (CGH), seem to help increase the rate of genetic diagnosis, particularly in individuals where autism is accompanied by malformations or dysmorphic features.[8] Moreover, the deletions and duplications detected in ASD patients may point to new disease genes involved in non-syndromic cases[9] and it has been shown that functional *de novo* mutations occur in ASDs at higher frequency than expected.[10] ASD is sometimes associated with fragile X syndrome (MIM #300624), tuberous sclerosis (MIM #191100), and cytogenetically detectable chromosome abnormalities.[11–13] Overall, identified genetic causes of autism can be classified as cytogenetically visible chromosomal abnormalities (~5%), copy number variants (10–20%), and single-gene disorders (~5%).[13]

Molecular analyses have highlighted the role of point mutations in single genes located in some of these regions: *SHANK3* (*ProSAP2*) on chromosome 22q13,[14–17] *Neuroligin 3* (*NLGN3*) and *Neuroligin 4* (*NLGN4*) on the X chromosome,[18] *Neurexin 1* (*NRXN1*) on chromosome 2p16,[19–21] and *Contactin Associated Protein-like 2* (*CNTNAP2*) on chromosome 7q35.[22–24] Notably, both neuroligin and neurexin proteins have key roles in the formation and functioning of synapses, particularly the alignment and activation of glutamate and GABA synapses. SHANK3 is a scaffolding protein that participates in the postsynaptic density (PSD), a neuronal structure that is responsible for the proper alignment of the membrane proteins on the postsynaptic surface. The PSD has been identified in several glutamate synapses. *SHANK3* haploinsufficiency seems to be responsible for the major neurological symptoms in

[1]Greenwood Genetic Center, Greenwood, SC, USA; [2]Istituto di Genetica Medica, Università Cattolica del S. Cuore, Rome, Italy; [3]Ospedale Maggiore, Bologna, Italy
*Correspondence: Dr CE Schwartz, JC Self Research Institute, Greenwood Genetic Center, 113 Gregor Mendel Circle, Greenwood, SC 29646, USA. Tel: +1 864 941 8140; Fax: +1 864 388 1703; E-mail: ceschwartz@ggc.org
*Died 29 September 2011.
Received 26 April 2011; revised 10 July 2012; accepted 11 July 2012; published online 15 August 2012

22q13 deletion syndrome (MIM #606232),[25,26] a condition characterized by speech delay, neonatal hypotonia, and behavioral problems along with other features.

Some recent studies identified *SHANK3* mutations in ~1% of ASD cases,[10,15–17] suggesting an important role for this gene in autism pathogenesis. Moreover, recent works in neuronal cell cultures[27,28] and mouse models[29–32] demonstrated a critical role for this gene in synaptic function, social interaction, and social communication, providing a functional link between SHANK3 deficiency and ASD behavioral features. Therefore, we screened this gene for mutations in two different groups of patients with ASD from the United States (US) and Italy.

MATERIALS AND METHODS

Patients and controls

We analyzed two cohorts of ASD patients: 133 South Carolina (SC) patients and 88 Italian patients. The SC group is composed of AD patients between the ages of 5 and 21 years from the South Carolina Autism Project (SCAP).[33] The diagnoses were established by the Autism Diagnostic Interview-Revised (ADI-R), the Autism Diagnostic Observation Schedule (ADOS), and/or the Childhood Autism Rating Scale (CARS) tests. This cohort contained 122 isolated and 11 familial cases (at least one affected sibling), with 101 male and 32 female patients (ratio 3.1:1). Seventy-six patients were Caucasian, fifty-three were African-American, and four were from other racial/ethnic backgrounds. This project was approved by institutional review board (IRB) and all families signed consent forms.

The Italian cohort was composed of Caucasian patients that met the ADI-R and/or ADOS criteria for ASD. The cohort included 81 isolated and 7 familial cases, with 67 males and 21 females (ratio 3.2:1). Seventeen patients were diagnosed with AD, sixty-eight with PDD-NOS, and three had Asperger syndrome. All the patients received thorough information about the research project and its consequences during a genetic counseling session and all of them signed consent forms.

We also screened the *SHANK3* gene in a second, independent replication cohort of 104 SC patients with ASD, 91 isolated and 13 familial cases, with 93 males and 11 females (ratio 8.4:1). Detailed diagnoses of the ASD sub-groups were not available on these patients. Seventy-four patients were Caucasian, thirteen were African-Americans, four were Hispanic, three were biracial, and ten were from other racial/ethnic backgrounds. All families and patients provided signed informed consent that was approved by IRB.

In all the cohorts, every patient received a clinical evaluation and had negative results for high-resolution chromosome analyses and *FMR1* molecular testing. Four SCAP patients carried balanced chromosomal rearrangements that did not involve any known autism locus.

We utilized 560 individuals from SC and 422 from Italy as control groups. The controls did not exhibit any behavioral problems and matched the patients for ethnic and background origin.

Molecular screening by denaturing high performance liquid chromatography (DHPLC) and direct sequencing

SHANK3 is composed of 23 exons according to Ensembl genome browser, using sequence ENSG00000251322 (NM_001080420.1 according to NCBI build GRCh37.1). However, some exons are particularly large, so 36 amplicons were utilized to amplify the coding sequences and about 100–150 base pairs deep in the intron (see Supplementary Table S1 for primer sequences and amplicon length). In order to amplify some of these amplicons, we used the primer sequences reported by Durand *et al*,[14] whereas we used Primer 3 Input (http://frodo.wi.mit.edu/primer3/) to design the primers for those amplicons that we had to adapt to DHPLC parameters. We were not able to amplify exons 1 and 11, because of their high-GC content. Similar problems have been reported in previous publications.[15,16] Several specific procedures, such as adding dimethyl-sulfoxide 10%, betaine 10%, or using polymerases specific for GC-rich amplicons, proved unsuccessful.

Molecular screening was performed by DHPLC analyses of the 36 amplicons, using the Transgenomic WAVE System (Transgenomics, Omaha, NE, USA). It is difficult to detect homozygous changes utilizing DHPLC approach. Therefore, to detect such variants, we decided to pool four samples for analysis. Each sample in a pool with abnormal peaks was sequenced using the DYEnamiv ET Dye Terminator Cycle Sequencing Kit on the MegaBACE 1000 Analysis System (Amersham Biosciences, Sunnyvale, CA, USA). Sequencing was performed in both forward and reverse directions.

Multiplex ligation-dependent probe amplification (MLPA) analysis

MLPA analysis of the chromosome 22q13 region was performed using the MRC-Holland SALSA P188 MLPA KIT 22q13. The kit contains four probes for exons 2, 9, 16, and 20 of the *SHANK3* gene. All samples were compared with seven negative controls (two samples from SC and five from Italy) and one positive control with a deletion of chromosome 22q13, encompassing *SHANK3* and several other genes.

Array comparative genomic hybridization

We performed array CGH in the 21 cases where a *SHANK3* sequence variant was found, in order to exclude other major deletions/duplications that may be responsible for the ASD phenotype. Array CGH was performed using Agilent oligonucleotide-array Kit 44B (Human Genome CGH Microarray Kit 44B; Agilent Technologies, Santa Clara, CA, USA), with an average resolution of about 75 kb, following the manufacturer's instructions.

Statistical Methods

The two-tailed *z*-test for two independent proportions was calculated using VassarStats (http://vassarstats.net)[34] to compare the proportions with *SHANK3* variants between cases and controls. When sample sizes were small ($np < 5$) Fisher's exact test was used using EpiInfo Statcalc. Given that the prevalence of SHANK3 variants could vary by racial ancestry background, comparisons between cases and controls were made separately for Caucasian and African-American subgroups as well as combined analysis. Adjustment for population stratification did not cause any significant finding to become non-significant, nor any non-significant finding to become significant.

RESULTS

We found five cases with deleterious *SHANK3* alterations (Table 1) out of 221 patients screened (2.3%). All the changes were heterozygous. Three alterations occurred in patients with AD (3 out of 150, 2%), one alteration occurred in a patient with PDD-NOS (1 out of 68 cases, 1.5%), and one in a patient with Asperger syndrome (1 out of 3 cases). Overall, three variants were detected in the 133 US patients (2.2%) and two in the 88 Italian patients (2.3%), while three variants were identified in 168 male patients (1.8%) and two in 53 female patients (3.8%).

Whole-genome SNP array analysis performed on a Caucasian US patient with AD and speech delay (14 878, clinical features in the Supplementary data) revealed a copy number loss of approximately 106 kb on chromosome band 22q13.33 between positions 49 475 238–249 581 309 (NCBI36/hg18). This deletion encompassed the *SHANK3* gene. Further characterization of this deletion by SNP analysis indicated the rearrangement was not inherited from one of the parents and revealed the loss of the paternal allele in the proband (SNP rs13055562; c.1304 + 42G > A; proband-G; mother-G/G; father-A/A), and the preservation of the adjacent *ACR* and *RABL2B* genes, thereby confirming the microarray results.

We identified a c.3931delG mutation in one Italian patient with PDD-NOS and severe intellectual disability (CMS14055, clinical features in the Supplementary data). This deletion causes a frameshift, starting at amino acid 1311 (p.Glu1311fs), which leads to a premature stop codon after 91 residues (fs1402X). The patient's mother did not carry the mutation. Unfortunately, DNA from the father was not available.

Table 1 Sequence variants in the *SHANK3* gene

Patient numbers	*Race*	*Sex*	*Nationality*	*Diagnosis*	*Variant*	*Parents tested*	*Effect*
14878	**C**	**M**	**SC**	**AD**	**106 kb del of *SHANK3***	**Neg (both)**	***SHANK3*loss of function**
17160[a]	**C**	**F**	**SC**	**AD**	**c.1339_1340insG (Ex 11)**	**Pos (Fa)**	**p.Ala447 fsX503**
CMS14055[b]	**C**	**M**	**Italian**	**PDD-NOS**	**c.3931delG (Ex 22)**	**Neg (Mo)**	**p.Glu1311 fsX1401**
12718[c]	**A**	**F**	**SC**	**AD**	**c.421 C>G (Ex 4)**	**Neg (both)**	**p.Pro141Ala**
CMS14218	**C**	**F**	**Italian**	**Asperger**	**c.1820-4 G>A (Int 15)**	**Pos (Mo)**	**Affects splice acceptor site[d]**
16309[e]	C	F	SC	AD		Neg (Mo)	
6477	C	M	SC	AD		Pos (Fa)	
12718[c]	A	F	SC	AD		Pos (Fa)	
16453	A	F	SC	AD		Pos (Mo)	
CMS15799	C	M	Italian	AD		Pos (Mo)	
CMS14081	C	M	Italian	PDD-NOS		Pos (Fa)	
CMS14090	C	M	Italian	PDD-NOS		Pos (Mo)	
CMS14067[f]	C	M	Italian	PDD-NOS	c.1304+48 C>T (Int 10)	Pos (Fa)	Loss of 1 CpG
CMS14056[g]	C	M	Italian	PDD-NOS		Pos (Mo)	
CMS14071[h]	C	M	Italian	PDD-NOS		Pos (Fa)	
CMS14077	C	M	Italian	PDD-NOS		Pos (Mo)	
CMS14092	C	M	Italian	PDD-NOS		Pos (Fa)	
CMS14093	C	M	Italian	PDD-NOS		Pos (Fa)	
CMS15702[i]	C	F	Italian	PDD-NOS		Pos (Fa)	
CMS17464	C	F	Italian	PDD-NOS		Pos (Mo)	
CMS17818	C	M	Italian	PDD-NOS		No parents available	
CMS17095[j]	C	M	Italian	PDD-NOS		Pos (Mo)	

Abbreviations: A, African-Americans; AD, autistic disorder; C, Caucasian, F, female; Fa, father, M, male; Mo, mother; SC, South Carolina; Neg, negative – the change was not found in the tested parent; PDD-NOS, pervasive developmental disorder, not otherwise specified; Pos, positive – the change has been detected in at least one of the parents.
[a]The patient also carries a missense mutation on the *NRXN1α* (c.83C>G; p.Gly28Ala).
[b]Father not available.
[c]The patient carries both the c.421 C>G and the c.1304+48 C>T changes.
[d]According to Splice Site Calculator.
[e]The patient also carries a chromosome 18p11.31 microduplication, not detected in the mother, father not available.
[f]The patient also carries a chromosome 10q11.23 *de novo* microdeletion, his affected twin sister carries the *SHANK3* variant but not the microdeletion.
[g]The affected sister carries the same variant.
[h]The patient's affected sister carries the same variant.
[i]The patient's affected brother carries the same variant.
[j]The patient also carries a chromosome 10q11.21 *de novo* microduplication and a chromosome 17q13.3 microduplication inherited from the father.
Deleterious mutations are indicated in bold.

Another frameshift mutation, c.1339_1340insG, was detected in a Caucasian US patient with AD and speech delay (17 160, clinical features in the Supplementary data). The insertion leads to a substitution of the alanine in position 447 with a glycine and introduces a premature stop codon at the residue 503 (p.A447fsX503). The same change was found in the patient's father, who is affected with learning problems and attention deficit disorder, and has a daughter with attention deficit hyperactivity disorder from a different relationship. The patient also carries a missense mutation in the *NRXN1* gene, c.83G>C; p.Gly28Ala. We were not able to test the parents for this variant. However, the bioinformatic analysis suggested the change can be considered as non-pathogenic (Supplementary Table S2).

A c.421C>G transversion was detected in an African-American US patient affected with AD (12718, clinical features in the Supplementary data). The variant caused the substitution of a proline with an alanine at the residue 141 of the amino acid sequence (p.Pro141Ala) within the N-terminal ankyrin repeats (α-fodrin) domain, which is critical for SHANK3 connection to the membrane-associated cytoskeleton. Sequence analysis of the patient's parents failed to detect the c.421C>G transversion, indicating that the variant occurred *de novo* in the proband. Moreover, the change is not reported in the NCBI SNP and 1000 genome database. Bioinformatic analysis suggested a potential destabilizing effect on the SHANK3 protein (Supplementary Table S2). The patient also carries another *SHANK3* variant (c.1304+48 C>T), described later in the results.

We found a c.1820-4G>A transition in an Italian patient affected by Asperger syndrome (CMS14218, clinical features in the Supplementary data). The substitution is inherited from the apparently unaffected mother and, according to bioinformatic analyses (Supplementary Table S2), it is hypothesized to disrupt the splicing function at the intron 15/exon 16 boundary, as it is located in the highly conserved splice acceptor site. Exons 15 and 16 constitute the core of the PDZ domain of the SHANK3 protein. We were unable to confirm a disruption of splicing as it was not possible to obtain fresh blood samples. However, we screened 776 normal Italian chromosomes for this mutation and failed to detect this change, which would indicate the alteration is not very common.

A c.1304+48C>T transition was detected in seventeen cases (7.7%): 4/133 SCAP (3.0%) patients and 13/88 (14.8%) Italian patients (see clinical features in the Supplementary data). Five of these 17 patients were affected by AD and 12 by PDD-NOS (See Table 1 for more details). In 8/17 cases (47%) the variant was inherited from the father, in 7/17 from the mother (41.1%). In one case, the only available parent was normal and in one case there were no parents available. Four of these 17 patients (23.5%) had an affected sibling, who also carried the variant. In all of these cases, the siblings also had the same PDD-NOS diagnosis as the probands. We screened normal individuals from both the US and Italian populations and we found the variant in 5/560 (0.9%) US and 9/422 (2.1%) Italian controls. The variant is reported as SNP rs76224556 with a frequency in the normal population of 3.8% (1000 genome database).

Table 2 Statistical analyses of *SHANK3* variants

SHANK3 *variants*					
AD vs *PDD-NOS (mutations)*			*AD* vs *PDD-NOS (mutations + c.1304 + 48C > T)*		
3/150 (2%) *vs* 1/68 (1.5%) $P = 1.0^a$			7/150 (4.6%) *vs* 13/68 (19.1%) $P =$ **0.0006**		
c.1304 + 48 C > T (SNP rs 76224556)					
				It patients vs *It controls*	
AD vs *PDD*	*AD* vs *controls (all)*	*PDD-NOS* vs *controls (all)*	*US patients* vs *US controls*		*ASD (all)* vs *controls (all)*
5/150 (3.3%) *vs* 12/68 (17.6%) $P =$ **0.0003**	5/150 (3.3%) *vs* 14/982 (1.4%) $P = 0.0903$	12/68 (17.6%) *vs* 14/982 (1.4%) $P <$ **0.0002**	4/133 (3%) *vs* 5/560 (0.9%) $P = 0.0740^a$	13/88 (14.8%) *vs* 9/422 (2.1%) $P <$ **0.0002**	17/221 (7.7%) *vs* 14/982 (1.4%) $P <$ **0.0002**
AD vs *1000 genomes, ALL*	*PDD-NOS* vs *1000 genomes, ALL*		*Italian patients vs 1000 genomes, EUR*		*ASD (all)* vs *1000 genomes, ALL*
5/150 (3.3%) *vs* 24/629 (3.8%) $P = 0.7795$	12/68 (17.6%) *vs* 24/629 (3.8%) $P <$ **0.0002**		13/88 (14.8%) *vs* 22/283 (7.8%) $P =$ **0.0498**		17/221 (7.7%) *vs* 24/629 (3.8%) $P =$ **0.0207**

Abbreviations: AD, autistic disorder; ASD, autism spectrum disorder; It, Italian; PDD-NOS, pervasive developmental disorder not otherwise specified; US, United States.
z-test calculated using Vassar Stats.
[a]Fisher's Exact (EpiInfo) used when np < 5 (rare outcome).
Significant *P* values (<0.05) are indicated in bold.

The statistical analyses reported in Table 2 show that the frequency of this SNP is significantly higher in our ASD cohort as compared with both the US and Italian controls we tested (P value < 0.0002) and the 1000 genome population (P value $= 0.0207$). The difference was significant after considering stratification by ethnic background for the US population (data not shown). These data show the variant is significantly more frequent in PDD-NOS cases (12/68, 17.6%) than in AD patients (5/150, 3.3%, $z = 3.651$, P value $= 0.00026$). Overall, the variant is more frequent in the ASD cohort than in the combined control population (7.7% *vs* 1.4%, $z = 5.312$, P value < 0.0002). This change occurs in a highly CG-rich region and actually causes the loss of a CpG dinucleotide. Bisulfite analysis revealed that the cytosine in the CpG dinucleotide is usually methylated and is located in the middle of a CpG island spanning exon 10 and intron 10 (Supplementary Figure S1). This region occurs immediately before two alternatively spliced exons, which code for the SH3 domain of the protein.

We performed MLPA analysis in the two cohorts to identify exon deletions or small chromosomal rearrangements, which may have been missed by sequencing, DHPLC, or high-resolution cytogenetic analyses. No major alterations were detected.

We also performed array CGH in those cases where a *SHANK3* sequence variant was found, in order to exclude other major deletions/duplications that may be responsible for the ASD phenotype. We found alterations in 3 of 17 cases with the c.1304 + 48C > T SNP (17.6%). A chromosome 18p11.31 microduplication was found in a SCAP female patient with AD that was not present in the mother (father was not available). A chromosome 10q11.23 *de novo* microdeletion was identified in an Italian male patient with PDD-NOS. His affected twin sister shows a milder phenotype and does not carry the microdeletion, indicating that maybe this rearrangement is not responsible for the shared ASDs, but may influence the severity of symptoms. A chromosome 10q11.21 *de novo* microduplication and a chromosome 17q13.3 microduplication inherited from the father were detected in an Italian male patient affected by PDD-NOS.

In a replication study, we detected a c.4402G > T transversion in 1 out of 104 SC patients with ASD (0.9%). This mutation leads to the substitution of an alanine with a serine at the position 1468 of the SHANK3 protein (p.Ala1468Ser). Unfortunately, no parents were available for this patient so we do not know if this change was *de novo*. Bioinformatic analysis of the change revealed a decreased stability in the mutant protein (Supplementary Table S2). We consider the variant as potentially deleterious for several reasons: it was not found in 376 US controls (752 chromosomes), neither it is in the NCBI SNP database nor the 1000 genome database, it causes the replacement of a neutral amino acid with a polar one, is located in a highly conserved region with a very important function (the cortarctin-binding domain) and it resides in a region of the gene that seems to be a hot spot for missense mutations.[23,24] In this same cohort we also identified eight individuals (7.7%) with the c.1304 + 48C > T transition (see Table 3 for details).

DISCUSSION

The SHANK3 protein is a crucial component of the PSD that takes part in the neuroligin-neurexin interaction at glutamate synapses. ASD patients have been shown to have mutations in other genes involved in this interaction: *NLGN3* and *NLGN4* on the X chromosome,[18] *NRXN1* on chromosome 2p16,[19–21] and *CNTNAP2* on chromosome 7q35.[22–24] Taken together, these data represent clear evidence for a link between the neuroligin–neurexin pathway and the etiology of autism.[16,19,35] Further clues have been provided by two different studies in mouse models, which showed that both *NLGN3*[36] and *NLGN4* mutations[37] were able to either reduce social interactions, impair communication, or produce aberrant behavior.

The role of SHANK3 in this pathway is apparently related to its scaffolding function in properly locating the protein receptors on the postsynaptic membrane. Recent studies on *Shank3*-deficient hippocampal neurons[27,28] and mouse models[29–32] showed how this gene has an important role in modeling the PSD at the striatum level, regulating basal neurotransmission at the AMPA glutamate receptors,

Table 3 Sequence variants in the *SHANK3* gene in the follow-up study

Patient number	*Race*	*Sex*	*Nationality*	*Diagnosis*	*Variant*	*Parents tested*	*Effect*
CMS 4296	**U**	**U**	**SC**	**ASD**	**c.4402 G>T (Ex 22)**	**Not tested**	**p.Ala1468Ser**
CMS 14120	C	M	SC	ASD		Neg (Mo)	
CMS 12214	H	M	SC	ASD		Pos (Mo)	
CMS 14795	C	M	SC	ASD		Neg (Mo)	
CMS 14806	C	M	SC	ASD	c.1304+48 C>T (Int 10)	Pos (Mo)	Loss of 1 CpG
CMS 9783	C	M	SC	ASD		Neg (Mo)	
CMS 4400	U	M	SC	ASD		Not tested	
CMS 15777	C	M	SC	ASD		Not tested	
CMS 12129	C	M	SC	ASD		Not tested	

Abbreviations: ASD, autism spectrum disorder; C, Caucasian; F, female; Fa, father; H, Hispanic; M, male; Mo, mother; Neg, negative – the change was not found in the tested parent; Pos, positive – the change has been detected in at least one of the parents; SC, South Carolina; U, unknown.
Deleterious mutations are indicated in bold.

and influencing long term potentiation and spine remodeling by organizing the AMPA receptor trafficking. Loss of one copy of *Shank3* correlates with behavioral issues reproducing most of the ASD traits, such as deficits in social interaction, social communication, and compulsive/repetitive actions.[27,29,31] These findings indicate that *SHANK3* loss-of-function mutations are associated with ASD features and neurodevelopmental disorders.

Several papers have reported *SHANK3* variants in ASD patients, consisting of both copy number variants[38] and point mutations.[15–17] These results indicate that *SHANK3* mutations have a frequency of about 1% in ASD cases, revealing a role for this gene as a possible major contributor in autism. The importance of the proper expression of the *SHANK3* gene in ASD is also supported by the finding that 7 out of 28 dysregulated microRNAs in ASD patients target this gene.[39] Several studies suggest that *SHANK3* has a critical role also in other neurodevelopmental disorders such as schizophrenia and intellectual disability.[10,40–42]

In the present study, we screened two cohorts of ASD patients for *SHANK3* mutations: 133 patients from SCAP[33] and 88 from Italy. We found five potentially pathogenic alterations (Table 1), resulting in a mutation rate of 2.3%, which is twice as high as the frequency of deleterious *SHANK3* mutations reported in previous studies.[15–17] In order to exclude intragenic deletion encompassing *SHANK3*, we performed MLPA analysis of the 22q13 region, but as the kit contains probe for just exons 2, 9, 16, and 20 of *SHANK3*, there is a possibility that very small copy number variants may have been missed.

Microarray analysis detected a *de novo* loss of 106 kb on chromosome 22q13.33 in a patient with AD and speech delay. The microdeletion and encompassed the whole *SHANK3* gene but did not involve the neighboring genes. Loss of one copy of *SHANK3* has been associated with ASDs in previous works.[15,38]

Two frameshift changes causing premature stop codons were detected. In the first case, a c.3931delG in a patient with PDD-NOS and severe intellectual disability was detected. We were not able to determine if the mutation was *de novo* because paternal DNA was not available. Nonetheless we considered the variant as pathogenic as it causes a frameshift and a premature stop codon at position 1401, leading to the loss of the C-terminal region of SHANK3: the Homer-binding, the Cortactin-binding and the Sterile Alpha Motif domains, which are crucial for SHANK3 interactions with other PSD proteins.

Functional studies have proved the critical role of these domains for the synaptic targeting signals, responsible for the proper localization and assembly of the PSD.[43] A similar frameshift mutation, causing a stop codon at position 1227 has been shown to affect dendritic spine development and morphology, as well as growth cone mobility in cultured hippocampal neurons.[27] Moreover, lack of the Homer-binding domain produces a gain-of-function effect, with up to 90% loss of SHANK3 at synapses due to downregulation of the wild-type copy expression, abnormal localization, and dimerization of SHANK3 at the PSD, combined with increased polyubiquitination and degradation of the protein.[31]

In the second case, a patient with AD and speech delay carries a c.1339_1340insG causing a premature stop codon at position 503, causing the disruption of the SH3 (Src homology 3) domain and the loss of the C-terminal half of the protein, including the PDZ (PSD-95/disk large/zonula occludens-1), Homer-binding, Cortactin-binding, and Sterile Alpha Motif domains. The patient presents with AD, mannerisms, and delayed speech development. She inherited the mutation from her father, who has learning problems and attention deficit disorder. Interestingly, Waga *et al*[17] reported an in-frame deletion involving the same region, c.1320-1338del; p.440-446delGPGPAP, in a Japanese boy presenting with ASD, speech and psychomotor delay. Also in this case the mutation was inherited from the mother, who had speech delay. Both *in vitro* and *in vivo* models reproducing the heterozygous loss of the C-terminal SHANK3 domains showed altered SHANK3 levels at synapses, reduced NMDA responses, abnormal dendritic spine development and morphology.[27,31] This patient also had a missense *NRXN1* mutation (c.83G>C; p.Gly28Ala). However, although this gene has been associated to ASD,[44] the bioinformatic analysis does not support a pathogenic role for this change.

It is noteworthy that all the three cases with *SHANK3* deletion and frameshift mutations showed some degree of speech or developmental delay in association with autistic traits: it is plausible that overall level of functional SHANK3 protein is lower than in those cases with different types of variants and the resulting disruption at the synaptic level may be more severe.

One AD patient carried a missense mutation (p.Pro141Ala) affecting the ankyrin repeats domain: missense mutations in this domain have already been proved to affect the colocalization of SHANK3 with other PSD proteins[15] and cause lower dendritic spine induction and maturation.[27]

We detected a c.1820-4G>A substitution both in an Asperger patient and in her apparently unaffected mother. The variant affects the highly conserved consensus sequence at the acceptor splice site, which appears to disrupt the proper splicing process, according to bioinformatic analysis, it has not been found in 776 chromosomes from ethnically matched controls, and the carrier mother was never clinically evaluated in order to exclude behavioral abnormalities: these

factors suggest that this change may actually have an important role in the pathogenesis of the observed phenotype.

We also detected a rare variant, c.1304 + 48C > T (SNP rs76224556) in 17 patients. The variant appears to be more frequent in patients with PDD-NOS as compared with AD (17.6% *vs* 3.3%, *P* value = 0.00013). It is interesting to note that deleterious *SHANK3* mutations have been detected in both AD and PDD-NOS cases with approximately the same rate, while the frequency of the rs76224556 SNP in our cohort is significantly higher in PDD-NOS patients than in AD patients (Table 2). In the 15/17 cases where both parents were available, 8 (53.3%) patients inherited the variant from the father and 7 (46.7%) from the mother, indicating that a parent-of-origin effect is not present.

SHANK3 tissue-specific expression is regulated in humans by DNA methylation, which particularly affects protein levels in hippocampal neurons.[45] It has also been shown that tissue-specific methylation of highly conserved intragenic regions in *SHANK3* has an important role in regulating intragenic promoter activity both *in vivo* and *in vitro*.[46] It is important to note that in a *Shank3*$^{-/-}$ mouse model, knocking down all the isoforms of the protein resulted in the synaptic disruption at the cortico-striatum level and the autistic-like behaviors. However, the knock-out model for only the α isoform (targeting the ankyrin repeats) showed minimal disruption and no behavioral issues.[28] On the other hand, deletion of the Homer-binding domain in one copy of *SHANK3* is sufficient to induce a 'gain-of-function' effect with severe consequences on SHANK3 levels and function at the PSD and on NMDA receptor activity.[31] The c.1304 + 48C > T transition affects a CpG dinucleotide in the middle of one of these highly conserved regions and bisulfate studies showed that this cytosine is supposed to be methylated and that the area surrounding the dinucleotide and encompassing part of exon 10 and part of intron 10 is methylated as well. Intronic mutations affecting an adjacent CpG island have been reported in ASD patients.[17] We hypothesize that the alteration of this CpG dinucleotide to a TpG dinucleotide may disrupt epigenetic regulation of *SHANK3* expression. According to recent studies, the region containing this CpG dinucleotide appears to be critical for the regulation of a shorter isoform of SHANK3.[28,46] The loss of this specific isoform seems to cause severe synaptic disorganization and ASD-like features in mice.[28]

In our study we detected some variants that were shared by ASD patients as well as unaffected parents and controls. However, we would like to highlight the fact that ASD is a very complex condition and the predisposing genetic background interacts with several other factors both genetic and environmental. It is possible that the same genetic change may lead to different phenotypes, and even that some of these phenotypes may fall in a very mild range and not even been recognized as a medical condition but rather as unusual personality traits. Also, the frequency of these variants between cases and controls is very significant and the test we used takes into account the size of the sample. This suggests that the reported changes may be able to increase the basal susceptibility to ASD that exists in the general population. For all these reasons, considering also their location within SHANK3 and the bioinformatic data we acquired about them, we believe that the mutations reported in this work may have an important role in altering, at some level, the proper function of the SHANK3 protein. Unfortunately, no cell lines are available from the patients carrying the variants mentioned above, so we could not perform functional or expression studies in order to validate the actual effect of these changes on the SHANK3 protein.

The difference in mutation frequencies between the two cohorts may reflect their patient composition. The SCAP patients have all been diagnosed with AD, while the Italian population includes 17 patients with AD, 68 with PDD-NOS, and 3 with Asperger syndrome. These different frequencies demonstrate that alterations in *SHANK3* are preferentially associated with cases with PDD-NOS and Asperger syndrome rather than AD (Table 2). These results are partially confirmed by recent studies that did not find any *SHANK3* alterations in cohorts of patients with AD.[47,48] Notably, no other gene variants have ever been found with a frequency as high as the one shown in PDD-NOS, making *SHANK3* the most relevant candidate gene for these cases.

In conclusion, *SHANK3* variants were detected in our cohorts in patients affected with any ASD subtype (AD, PDD-NOS, and Asperger syndrome). However, the SNP rs76224556 appears to be associated with a significant number of PDD-NOS cases. Our data strongly suggest this SNP may have a genotype/phenotype correlation in ASDs. Thus, perhaps testing of this gene, and particularly screening of the SNP rs76224556, should be considered in patients with a diagnosis of PDD-NOS confirmed by ADI-R, CARS, or ADOS testing. Furthermore, we think that this preferential frequency in patients with a specific ASD condition suggests that a careful clinical-neuropsychiatric evaluation is a very important tool for an effective genetic screening of ASD patients.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank all the patients and their families in Italy and SC who participated in the present study, and all the European colleagues who collaborated in collecting the samples and the clinical information: Giuseppe Gobbi from Ospedale Maggiore, Bologna, Maria Giulia Torrioli from Policlinico 'A. Gemelli', Università Cattolica del S. Cuore, Rome, Edith Said and Alfred Cuschieri from University of Malta, La Valletta. We thank Lynda Holloway and Raewyn Lowe for performing the DHPLC analysis and Dana Schultz and Melissa Cook for sequencing all samples. This work was supported in part by grants from Telethon to FG (N° GGP06170), from NICHD (HD26202) to CES and the South Carolina Department of Disabilities and Special Needs (SCDDSN). This work is dedicated to the memories of Ethan Francis Schwartz (1996–1998), Roland Carlsson (1949–2008), Giuseppe Coloca (1928–2008), and Carmela Romeo (1929–2010). We especially dedicate this paper to the memory of our colleague, Julianne S Collins (1969–2011), who contributed significantly to epidemiological studies at the Greenwood Genetic Center. Her contributions will be missed by all.

1 American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. Washington, DC: American Psychiatric Association, 2000.
2 Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008. Principal Investigators; Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders – Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. *MMWR Surveill Summ* 2012; **61**: 1–19.
3 Bailey A, Le Couteur A, Gottesman I *et al*: Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* 1995; **25**: 63–77.
4 Piven J, Palmer P, Jacobi D, Childress D, Arndt S: Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. *Am J Psychiatry* 1997; **154**: 185–190.
5 Chakrabarti S, Fombonne E: Pervasive developmental disorders in preschool children. *JAMA* 2001; **285**: 3093–3099.
6 Herman GE, Henninger N, Ratliff-Schaub K, Pastore M, Fitzgerald S, McBride KL: Genetic testing in autism: how much is enough? *Genet Med* 2007; **9**: 268–274.
7 Schaefer GB, Mendelsohn NJ: Clinical genetics evaluation in identifying the etiology of autism spectrum disorders. *Genet Med* 2008; **10**: 301–305.

8 Jacquemont ML, Sanlaville D, Redon R *et al*: Array-based comparative genomic hybridisation identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. *J Med Genet* 2006; **43**: 843–849.
9 Sebat J, Lakshmi B, Malhotra D *et al*: Strong association of *de novo* copy number mutations with autism. *Science* 2007; **316**: 445–449.
10 Awadalla P, Gauthier J, Myers RA *et al*: Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. *Am J Hum Genet* 2010; **87**: 316–324.
11 Xu J, Zwaigenbaum L, Szatmari P, Scherer SW: Molecular cytogenetics of autism. *Curr Genomics* 2004; **5**: 347–364.
12 Veenstra-Vanderweele J, Christian SL, Cook EH Jr: Autism as a paradigmatic complex genetic disorder. *Annu Rev Genomics Hum Genet* 2004; **5**: 379–405.
13 Miles JH: Autism spectrum disorders–a genetics review. *Genet Med* 2011; **13**: 278–294.
14 Durand CM, Betancur C, Boeckers TM *et al*: Mutations in the gene encoding the synaptic scaffolding protein *SHANK3* are associated with autism spectrum disorders. *Nat Genet* 2007; **39**: 25–27.
15 Moessner R, Marshall CR, Sutcliffe JS *et al*: Contribution of *SHANK3* mutations to autism spectrum disorder. *Am J Hum Genet* 2007; **81**: 1289–1297.
16 Gauthier J, Spiegelman D, Piton A *et al*: Novel *de novo SHANK3* mutation in autistic patients. *Am J Med Genet* 2009; **150B**: 421–424.
17 Waga C, Okamoto N, Ondo Y *et al*: Novel variants of the *SHANK3* gene in Japanese autistic patients with severe delayed speech development. *Psychiatr Genet* 2011; **21**: 208–211.
18 Jamain S, Quach H, Betancur C *et al*: Mutations of the X-linked genes encoding neuroligins *NLGN3* and *NLGN4* are associated with autism. *Nat Genet* 2003; **34**: 27–29.
19 Chubykin AA, Liu X, Comoletti D, Tsigelny I, Taylor P, Sudhof TC: Dissection of synapse induction by neuroligins: effect of a neuroligin mutation associated with autism. *J Biol Chem* 2005; **280**: 22365–22374.
20 Feng J, Schroer R, Yan J *et al*: High frequency of neurexin 1beta signal peptide structural variants in patients with autism. *Neurosci Lett* 2006; **409**: 10–13.
21 Szatmari P, Paterson AD *et al*: Autism Genome Project Consortium (AGPC). Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 2007; **39**: 319–328.
22 Alarcón M, Abrahams BS, Stone JL *et al*: Linkage, association, and gene-expression analyses identify *CNTNAP2* as an autism-susceptibility gene. *Am J Hum Genet* 2008; **82**: 150–159.
23 Arking DE, Cutler DJ, Brune CW *et al*: A common genetic variant in the neurexin superfamily member *CNTNAP2* increases familial risk of autism. *Am J Hum Genet* 2008; **82**: 160–164.
24 Bakkaloglu B, O'Roak BJ, Louvi A *et al*: Molecular cytogenetic analysis and resequencing of *contactin associated protein-like 2* in autism spectrum disorders. *Am J Hum Genet* 2008; **82**: 165–173.
25 Bonaglia MC, Giorda R, Borgatti R *et al*: Disruption of the *ProSAP2* gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. *Am J Hum Genet* 2001; **69**: 261–268.
26 Wilson HL, Wong AC, Shaw SR *et al*: Molecular characterization of the 22q13 deletion syndrome supports the role of haploinsufficiency of *SHANK3/PROSAP2* in the major neurological symptoms. *J Med Genet* 2003; **40**: 575–584.
27 Durand CM, Perroy J, Loll F *et al*: *SHANK3* mutations identified in autism lead to modification of dendritic spine morphology via an actin-dependent mechanism. *Mol Psychiatry* 2012; **17**: 71–84.
28 Verpelli C, Dvoretskova E, Vicidomini C *et al*: Importance of Shank3 protein in regulating metabotropic glutamate receptor 5 (mGluR5) expression and signaling at synapses. *J Biol Chem* 2011; **286**: 34839–34850.
29 Bozdagi O, Sakurai T, Papapetrou D *et al*: Haploinsufficiency of the autism-associated Shank3 gene leads to deficits in synaptic function, social interaction, and social communication. *Mol Autism* 2011; **1**: 1–15.
30 Peça J, Feliciano C, Ting JT *et al*: Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* 2011; **472**: 437–442.
31 Bangash MA, Park JM, Melnikova T *et al*: Enhanced polyubiquitination of Shank3 and NMDA receptor in a mouse model of autism. *Cell* 2011; **145**: 758–772.
32 Wang X, McCoy PA, Rodriguiz RM *et al*: Synaptic dysfunction and abnormal behaviors in mice lacking major isoforms of *Shank3*. *Hum Mol Genet* 2011; **20**: 3093–3108.
33 Schroer RJ, Phelan MC, Michaelis RC *et al*: Autism and maternally derived aberrations of chromosome 15q. *Am J Med Genet* 1998; **76**: 327–336.
34 Dean AG, Dean JA, Burton AH, Dicker RC: Epi Info: a general-purpose microcomputer program for public health information systems. *Am J Prev Med* 1991; **7**: 178–182.
35 Garber K: Neuroscience. Autism's cause may reside in abnormalities at the synapse. *Science* 2007; **317**: 190–191.
36 Tabuchi K, Blundell J, Etherton MR *et al*: A *neuroligin-3* mutation implicated in autism increases inhibitory synaptic transmission in mice. *Science* 2007; **318**: 71–76.
37 Jamain S, Radyushkin K, Hammerschmidt K *et al*: Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism. *Proc Natl Acad Sci USA* 2008; **105**: 1710–1715.
38 Marshall CR, Noor A, Vincent JB *et al*: Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008; **82**: 477–488.
39 Abu-Elneel K, Liu T, Gazzaniga FS *et al*: Heterogeneous dysregulation of microRNAs across the autism spectrum. *Neurogenetics* 2008; **9**: 153–161.
40 Gauthier J, Champagne N, Lafrenière RG *et al*: De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia. *Proc Natl Acad Sci USA* 2010; **107**: 7863–7868.
41 Hamdan FF, Gauthier J, Araki Y *et al*: Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am J Hum Genet* 2011; **88**: 306–316.
42 Gong X, Jiang YW, Zhang X *et al*: High proportion of 22q13 deletions and SHANK3 mutations in Chinese patients with intellectual disability. *PLoS One* 2012; **7**: e34739.
43 Boeckers TM, Liedtke T, Spilker C *et al*: C-terminal synaptic targeting elements for postsynaptic density proteins ProSAP1/Shank2 and ProSAP2/Shank3. *J Neurochem* 2005; **92**: 519–524.
44 Kim H-G, Kishikawa S: Higgins *et al:* Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet* 2008; **82**: 199–207.
45 Beri S, Tonna N, Menozzi G, Bonaglia MC, Sala C, Giorda R: DNA methylation regulates tissue-specific expression of *Shank3*. *J Neurochem* 2007; **101**: 1380–1391.
46 Maunakea AK, Nagarajan RP, Bilenky M *et al*: Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 2010; **466**: 253–257.
47 Qin J, Jia M, Wang L *et al*: Association study of *SHANK3* gene polymorphisms with autism in Chinese Han population. *BMC Med Genet* 2009; **10**: 61.
48 Sykes NH, Toma C, Wilson N *et al*: Copy number variation and association analysis of *SHANK3* as a candidate gene for autism in the IMGSAC collection. *Eur J Hum Genet* 2009; **17**: 1347–1353.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)