SOBP Is Mutated in Syndromic and Nonsyndromic Intellectual Disability and Is Highly Expressed in the Brain Limbic System

Efrat Birk,1 Adi Har-Zahav,1 Chiara M. Manzini,2 Metsada Pasmanik-Chor,3 Liora Kornreich,4 Christopher A. Walsh,2 Konrad Noben-Trauth,5 Adi Albin,6 Amos J. Simon,7 Laurence Colleaux,8 Yair Morad,1,9 Limor Rainshtein,6 David J. Tischfield,2 Peter Wang,2 Nurit Magal,6 Idit Maya,6 Noa Shoshani,1 Gideon Rechavi,1,7 Doron Gothelf,1,10 Gal Maydan,11 Mordechai Shohat,1,6,12 and Lina Basel-Vanagaite1,6,13,*

Intellectual disability (ID) affects 1%–3% of the general population. We recently reported on a family with autosomal-recessive mental retardation with anterior maxillary protrusion and strabismus (MRAMS) syndrome. One of the reported patients with ID did not have dysmorphic features but did have temporal lobe epilepsy and psychosis. We report on the identification of a truncating mutation in the SOBP that is responsible for causing both syndromic and nonsyndromic ID in the same family. The protein encoded by the SOBP, siren oculis binding protein ortholog, is a nuclear zinc finger protein. In mice, Sobb (also known as Jxc1) is critical for patterning of the organ of Corti; one of our patients has a subclinical cochlear hearing loss but no gross cochlear abnormalities. In situ RNA expression studies in postnatal mouse brain showed strong expression in the limbic system at the time interval of active synaptogenesis. The limbic system regulates learning, memory, and affective behavior, but limbic circuitry expression of other genes mutated in ID is unusual. By comparing the protein content of the +/jc tojcjc mice brains with the use of proteomics, we detected 24 proteins with greater than 1.5-fold differences in expression, including two interacting proteins, dynamin and pacsin1. This study shows mutated SOBP involvement in syndromic and nonsyndromic ID with psychosis in humans.
be fragmentation of his thought processes. He became easily frustrated and exhibited recurrent aggressive outbursts. Psychiatric symptoms are known to be a part of several syndromes, including velo-cardio-facial syndrome (MIM 192430), fragile X syndrome (MIM 300624), MRX30 (MIM 300558), and others.

In order to identify the causative mutation in this family, after informed consent was obtained from all family members or their legal guardians (in accordance with a protocol approved and reviewed by the National Committee for Genetic Studies, Israel Ministry of Health), we genotyped 400 microsatellite markers from the ABI PRISM linkage mapping set, version 2.5 (Applied Biosystems, Foster City, CA, USA). All of the patients exhibited a large continuous segment of homozygosity on chromosomal region 6q21. A common homozygous disease-bearing haplotype was constructed; informative recombinations delineated a critical region of 9.8 Mb between the polymorphic markers D6S449 and D6S404 (Figure 1A). Individual IV-5, with nonsyndromic ID, showed the same disease-associated haplotype. Through SNP genotyping using Affymetrix Human Mapping 50k Xba 240 arrays on individuals IV-1, IV-2, IV-4, and IV-7, the candidate interval was reduced to the region between markers D6S449 and rs2235989. There was only one homozygosity region shared by all of the patients; there were no homozygosity regions shared by all individuals apart from individual IV-5 (Figure S1, available online). The two-point LOD score calculations were performed with the program SUPERLINK and yielded a $Z_{\text{max}}$ score of 4.45 at recombination fraction (0) 0.00 (Table S1). The candidate region contained 36 known or predicted genes (Figure 1B), as identified with the use of available databases (NCBI and UCSC Genome Browser, March 2006 version). All exons including exon-intron junctions of the following candidate genes were sequenced: GRIK2 (MIM 138244), LIN28B (MIM 610166), BVES (MIM 604577), POPDC3 (MIM 605824), PREP (MIM 600400), ATG5 (MIM 604261), RTN4IP1 (MIM 610502), PDSS2 (MIM 610564), SCML4, OSTM1 (MIM 607649), NR2E1 (MIM 603849), SNX3 (MIM 605930), FOXO3A (MIM 602681), ARMC2, SESN1 (MIM 606103), CD164 (MIM 603356), LACE1, PPIL6, SMPD2 (MIM 603498), ZBTB24, C6ORF199, GPR6 (MIM 600553), WASF1 (MIM 605035), CDC40 (MIM 605585), and DDO (MIM 124450).

No pathogenic sequence changes were found. Sequencing of SOBP was performed with the use of ten primer pairs (Table S2). In this gene, we identified a homozygous C>T change in nucleotide 1981 of exon 6 (c.1981C>T) (NM_018013.3) (Figure 1C), which creates an early stop codon that causes the truncation of 212 residues of the 873 amino acids of the protein (p.Arg661X). SOBP mutation screening was performed on genomic DNA by PCR amplification with primer pair 7 (Table S2). The mutation segregated with the disease (Figure 1A) and was not observed in 300 control chromosomes of Israeli Arab ancestry.
individuals. mRNA from the lymphoblasts of one of the patients was successfully amplified, indicating that mutant mRNA is not completely degraded (data not shown).

In order to determine the frequency of \textit{SOBP} involvement, when mutated, in other families with autosomal-recessive ID, homozygosity at the 6q21 locus was tested by genotyping with the Affymetrix GeneChip Mapping 250K Array in 22 independent consanguineous patients presenting with isolated or mildly syndromic ID. However, no patient was found to be compatible with linkage to this locus.

The \textit{SOBP} structure was described with the use of the UCSC (GRCh37/hg19, February 2009), NCBI (build 37.1, August 2009), and Ensembl (GRCh37, February 2009) genome browsers (Figure 2A).\textsuperscript{13} DNA and RNA binding motifs were predicted by the following servers: RNABindR, BindN, PPRint, and MEME. The human \textit{SOBP} (sine oculis binding protein ortholog, also known as Jackson circler protein 1 \textit{[JXC1]; RefSeq NM_018013, GeneID: 55084}) encodes a zinc finger protein that is located on human chromosome 6q21 and covers 171,192 bp of genomic DNA. The mRNA size is 6227 bp. The \textit{SOBP} promoter consists of a distinct CpG island region; another CpG island region is clearly found in exon 6 and is predicted with high probability to be an additional promoter region. There are seven exons, the last of which is the untranslated

![Figure 2](image.png)
Genomic conservation analysis shows that exons are highly conserved among vertebrates, whereas introns are conserved mainly within primates (Figure 2A). Transcript variants encode the main 873 amino acid protein (transcript 1) (NP_060483) and a shorter variant of 232 amino acids (transcript 2) (Figure 2A). It contains two FCS-type zinc fingers (MYM-type Zinc finger with FCS sequence motif, PF06467) and two proline-rich (PR) regions (IPR000694) (Figure 2B). There are many species with orthologs of the human protein. In the UCSC database, the chimp SOBP protein is a 668 residue protein, and the human is an 874 residue protein. The chimp protein divides into two parts. The first section is 1–409, with 99% identity to the human protein. Then there is a missing section in the chimp, found only in the human, and then comes another fragment (chimp 410–668; human 696–874) with 87% identity. Human SOBP is expressed in various tissues. There is a clear nuclear localization signal (NLS) at residues 9–24 and another putative NLS around residue 750. The motifs found in the protein are depicted in Figure 2B. Predictions for RNA and DNA binding motifs by different servers show that SOBP contains RNA binding motifs around and including residue 661 (Figure 2B). Interestingly, a search for DNA binding motifs using various bioinformatics databases did not show such motifs in SOBP. Truncation of SOBP after residue 661 causes elimination of a PR2 domain as well as possible NLS motifs and an RNA binding region, which may be important for the functioning of the protein. Protein-conservation analysis by the ConSeq server showed high sequence conservation and functionality of the C-terminal part of the protein; the N-terminal part is less conserved and has fewer functional residues (Figure S2).

In order to establish whether Sobp expression in the developing and adult brain could provide some insight into the cause of cognitive impairment, we performed a complete developmental characterization of Sobp mRNA expression in the murine brain by in situ hybridization. For in situ hybridization, plasmid templates for riboprobe synthesis were generated by subcloning a fragment of Sobp PCR amplified from IMAGE cDNA clone 10083890 containing a 493 bp sequence corresponding to nucleotide positions 21–514 of the Sobp coding sequence. In situ hybridization was performed as described by Tucker et al.14 Sobp was present at embryonic day (E) 14 throughout the developing brain, with intense staining in the cortical plate (Figure 3). During postnatal (P) development (P0,
P10, and adult), Snhp was also expressed in all neurons, but a striking and intense expression pattern was especially evident in the limbic system, with the highest expression levels throughout layer V neurons in the cortex, the hippocampus, the pyriform cortex, the dorsomedial nucleus of the thalamus (Figures 3A and 3B), the amygdala, and the hypothalamus (Figure 3C). Cortical expression was strong throughout development, with no clear dorsoventral or rostrocaudal gradient, the highest levels at P10 in layers II/III and V and in the subplate (Figure 3D). Relatively strong expression was also observed in the mitral cells layer and anterior olfactory bulb (Figure 3B) and in the Purkinje cell layer in the cerebellum (not shown). We observed a striking and very specific expression of Snhp in the limbic system postnatally, corresponding to the time window of active synaptogenesis. It is known that on mouse P7 there is a period of significant synaptogenesis for neural-circuit formation and the expression of many genes encoding synaptic proteins and receptors.15 Another protein that has been found to be expressed in the limbic circuitry so strikingly is limbic system-associated membrane protein (LAMP).16 Because the circuit showing the highest Snhp expression regulates learning, memory, and affective behavior in humans, we suggest that disruption of SOBP in the limbic system during synaptogenesis leads to the abnormal cognition observed in the patients. The pathogenesis of strabismus in the patients is unclear. Because cortical Snhp expression is uniform, we cannot explain the origin of the strabismus on the basis of dynamic changes of Snhp expression levels in the visual cortex. Regulation of eye convergence may be affected at levels other than the cortex. Interestingly, in Drosophila, ectopic expression of Snhp leads to structural aberrations in the adult eye.17

In mice, Jxc1/Snhp is critical for cochlear growth, cell fate, and patterning of the organ of Corti.18,19 Mice mutants also exhibit erratic circling behavior and have no vestibular evoked potentials. In contrast to mutant mice, one of our patients has a subclinical cochlear hearing loss of 30–40 dB but does not show any vestibular abnormalities or gross cochlear abnormalities. In order to gain some insight into the cause of cognitive impairment, we compared the protein content of the +/jc to jc/jc mice brains by using proteomics. For the proteomics analysis, proteins from brain tissue from each of brains by using proteomics. For the proteomics analysis, 1.3-fold differences between +/jc and jc/jc samples (Table 1).

<table>
<thead>
<tr>
<th>Table 1. Expression with Greater than 1.5-fold Differences in Mus musculus Protein Name</th>
<th>NCBI Reference Sequence</th>
<th>Ratio of +/jc to jc/jc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamin</td>
<td>487851</td>
<td>1.55</td>
</tr>
<tr>
<td>Pacsin1</td>
<td>148690599</td>
<td>1.59</td>
</tr>
<tr>
<td>Dihydropyrimidinase-like 2</td>
<td>40254595</td>
<td>1.86</td>
</tr>
<tr>
<td>Syntaxin binding protein 1</td>
<td>17225417</td>
<td>1.82</td>
</tr>
<tr>
<td>Collapsin response mediator protein 1</td>
<td>40068507</td>
<td>1.52</td>
</tr>
<tr>
<td>Solute carrier family 25</td>
<td>27369581</td>
<td>1.51</td>
</tr>
<tr>
<td>NmrA-like family domain containing 1</td>
<td>24431937</td>
<td>1.64</td>
</tr>
<tr>
<td>Actin related protein 2/3 complex, subunit 2</td>
<td>112363072</td>
<td>1.58</td>
</tr>
<tr>
<td>Phosphoglycerate mutase 1</td>
<td>114326546</td>
<td>1.52</td>
</tr>
<tr>
<td>Carbonic anhydrase III</td>
<td>31982861</td>
<td>3.72</td>
</tr>
<tr>
<td>ATP synthase, H+ transporting, mitochondrial F1 complex, subunit d</td>
<td>21313679</td>
<td>3.79</td>
</tr>
<tr>
<td>Hypothetical protein LOC66469</td>
<td>27754130</td>
<td>1.66</td>
</tr>
<tr>
<td>Growth hormone</td>
<td>6679997</td>
<td>5.96</td>
</tr>
<tr>
<td>ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit, isofrom CRA_a</td>
<td>148699643</td>
<td>1.57</td>
</tr>
<tr>
<td>Unnamed protein product</td>
<td>74150634</td>
<td>1.96</td>
</tr>
<tr>
<td>Parvalbumin</td>
<td>53819</td>
<td>1.88</td>
</tr>
<tr>
<td>PREDICTED: similar to expressed in non-metastatic cells 1, protein (NM23A) (nucleoside diphosphate)</td>
<td>73963227</td>
<td>3.23</td>
</tr>
<tr>
<td>Beta-globin</td>
<td>156257681</td>
<td>1.56</td>
</tr>
<tr>
<td>Heat-responsive protein 12</td>
<td>40807498</td>
<td>2.06</td>
</tr>
</tbody>
</table>

Although these results should be further investigated and validated in additional studies, there are several interesting observations. Two of the differentially expressed proteins, pacsin1 and dynamin, may be involved in the same cellular pathway—the rat ortholog of pacsin1 was identified as a synaptic protein that interacts with dynamin.20 Dynamin is a neuron-specific guanosine triphosphatase that is important in endocytosis at the fission step of nascent clathrin-coated vesicles from the plasma membrane.21 In mice lacking dynamin, synaptic vesicle endocytosis is severely impaired during strong exogenous stimulation but resumes efficiently when the stimulus is terminated, suggesting that dynamin is necessary during high levels
of neuronal activity. Another differentially expressed protein, parvalbumin, encoded by pvalb, belongs to the family of calcium-binding proteins. Parvalbumin-positive neurons are observed in the hippocampal cells; they constitute an abundant subpopulation of cells that express GABA. During the development of the rat neocortex and hippocampus, Pvalb mRNA expression is visible after P6. As with Sobp, by P10 Pvalb expression is detected mainly in layer V of the neocortex and in the hippocampus; by P14 there is a marked increase of expression in the entire neocortex and in the hippocampus. Significant functional annotations of differentially expressed proteins, with mouse brain genes used as background, were for axon (GO: 0030424) and neuron (GO: 0043005) projection, axon guidance, synaptic vesicle traffic, 5HT3- and 5HT4-type receptor-mediated processes, the alpha and beta3 adrenergic receptors, the corticotropin releasing factor, and sugar metabolism (results not shown).

In summary, we describe a gene involved in human cognition, SOBP, that shows the highest expression in the limbic system postnatally and that, when mutated, causes syndromic and nonsyndromic autosomal-recessive ID.

Supplemental Data

Supplemental Data include two tables and two figures and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

This study was supported by the Israeli Ministry of Health Chief Scientist Foundation (grant no. 3-4963) and the Israeli Science Foundation (grant no. 558/09). The authors thank Gabrielle J. Halpern for her help with editing of the manuscript and the members of the family for their cooperation.

Received: August 24, 2010
Revised: October 5, 2010
Accepted: October 8, 2010
Published online: October 28, 2010

Web Resources

The URLs for data presented herein are as follows:

BindN, http://bioinfo.ggc.org/bindn
Database for Annotation, Visualization and Integrated Discovery (DAVID), http://david.abcc.ncifcrf.gov/home.jsp
GOTM, http://bioinfo.vanderbilt.edu/webgestalt
Homozygosity Mapper, http://www.homozygositymapper.org/
Multiple Em for Motif Elicitation (MEME), http://meme.sdsc.edu/meme4_4_0/cgi-bin/meme.cgi
Mouse Genome Informatics (MGI) database, http://www.informatics.jax.org/
PANTHER, http://www.pantherdb.org
PPRint, http://www.inttech.res.in/raghava/pprint/submit.html
RNABindR, http://bindr2.gdcb.iastate.edu/RNABindR/
UniProt server, http://www.uniprot.org/

References


