Constitutional downregulation of SEMA5A expression in autism.
Malin Melin, Birgit Carlsson, Henrik Anckarsäter, Maria Rastam, Catalina Betancur, Anders Isaksson, Christopher Gillberg, Niklas Dahl

To cite this version:

HAL Id: inserm-00124743
http://www.hal.inserm.fr/inserm-00124743
Submitted on 2 Sep 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Constitutional down-regulation of SEMA5A expression in autism

M. Melin\textsuperscript{a}, B. Carlsson\textsuperscript{a}, H. Anckarsater\textsuperscript{b}, M. Rastam\textsuperscript{b}, C. Betancur\textsuperscript{c}, A. Isaksson\textsuperscript{a}, C. Gillberg\textsuperscript{b}, N. Dahl\textsuperscript{a}

\textsuperscript{a}Department of Genetics and Pathology, Uppsala University, Uppsala, Sweden
\textsuperscript{b}Department of Child and Adolescent Psychiatry, Göteborg University, Göteborg, Sweden
\textsuperscript{c}INSERM U513, Université Paris XII, Créteil, France

Correspondence should be addressed to Professor Niklas Dahl, Uppsala University, Department of Genetics and Pathology, Rudbeck Laboratory, 751 85 Uppsala, Sweden; e-mail: niklas.dahl@genpat.uu.se
Abstract

There is strong evidence for the importance of genetic factors in idiopathic autism. The results from independent twin and family studies suggest that the disorder is caused by the action of several genes, possibly acting epistatically. We have used cDNA microarray technology for the identification of constitutional changes in the gene expression profile associated with idiopathic autism. Samples were obtained and analyzed from six affected subjects belonging to multiplex autism families and from six healthy controls. We assessed the expression levels for approximately 7,700 genes by cDNA microarrays using mRNA derived from Epstein Barr virus (EBV)-transformed B-lymphocytes. The microarray data was analyzed in order to identify up- or down-regulation of specific genes. A common pattern with nine down-regulated genes was identified among samples derived from individuals with autism when compared to controls. Four of these nine genes encode proteins involved in biological processes associated with brain function or the immune system, and are consequently considered as candidates for genes associated with autism. Quantitative real-time PCR confirms the down-regulation of the gene encoding SEMA5A, a protein involved in axonal guidance. EBV should be considered as a possible source for altered expression but our consistent results make us suggest SEMA5A a candidate gene in the etiology of idiopathic autism.

Key Words

Autistic disorder, cDNA microarrays, Gene expression, Chromosome 7q31, SEMA5A

Introduction

Autistic disorder (AD) is a neurodevelopmental disorder characterized by impairments in reciprocal social interactions, such as verbal and non-verbal communication, and flexibility as regards interests and activities. The etiology as well as the clinical expression of the disorder is heterogeneous. In idiopathic autism, the symptoms appear during the first years of life. The prevalence of AD is approximately 0.1%, with a higher frequency among boys [1]. Genetic factors behind AD were suggested already by Kanner in 1943 [2], and this has been confirmed by subsequent studies. The average concordance rate for AD is 70% for monozygotic twins compared to 0%-5% for dizygotic twins according to three epidemiological same-sex twin studies [3-5]. Analysis of recurrence risks in families suggests that between two and ten gene loci, possibly more, may be involved [6]. Several full genome scans for autism have been performed, all of which show linkage to a number of susceptibility loci, supporting that several genes may contribute to the disorder [7-11]. Specific gene regions have also been suggested from the identification of patients with different chromosomal abnormalities [12]. Recently, mutations in the neuroligin genes NLGN3 and NLGN4 were reported in a subset of patients with AD [13], and mutations in NLGN4 has been confirmed in a family with X-linked mental
retardation with or without autism [14]. The neuroligin genes encode cell-adhesion proteins, suggesting abnormal synapse formation as a key mechanism in the pathogenesis of AD. Still, the genetic factors behind the majority of patients with idiopathic AD remain unknown.

We hypothesize that AD may be associated with constitutional abnormalities in the expression levels of several genes, and we present here a cDNA microarray analysis of 7,700 genes in samples from individuals with familial AD. Abnormalities in the expression profile in patient-derived lymphoblastoid cells are presented and discussed.

**Materials and methods**

Samples from six subjects with AD belonging to six sib-pair families included in the genome scan performed by Philippe et al [7], and co-segregating for chromosome 7q31, were selected in order to reduce the genetic heterogeneity and to facilitate the identification of common altered patterns in gene expression. Chromosome 7q31 has previously been reported to contain a susceptibility locus for autism. Blood samples were collected from the patients after written informed consent from the parents, and transformed into lymphoblastoid cell lines (LCLs) by Epstein Barr transfection. Diagnosis of AD/childhood autism was confirmed using the DSM-IV diagnostic criteria and Autism Diagnostic Interview-Revised algorithm [15,16]. Organic conditions associated with autism, such as tuberous sclerosis, epilepsy, fragile X syndrome, or chromosomal abnormalities were excluded in the subjects. Clinical features of the patients are presented in table 1. Six LCLs from healthy individuals (3 males and 3 females), without a family history of autism or other neuropsychiatric disorders, were used as controls. This study was approved by the regional ethical committee in Sweden.

The LCLs were cultured in RPMI with HEPES, 10% fetal calf serum, 4 µg/ml glutamine and 1000 E/ml penicillin-streptomycin. The cells were maintained at 37°C in a humidified atmosphere with 5% CO₂. Total RNA was prepared using TRIzol® (Invitrogen) according to the manufacturer’s protocol. The quality of the total RNA was controlled by capillary electrophoresis using a Bioanalyzer 2100 (Agilent Technologies). Only RNA without any sign of degradation was used. In-house produced cDNA microarrays with approximately 7,700 genes in duplicates were used to assess mRNA levels. The cDNA clones were obtained from the Research Genetics sequence-verified human cDNA collection (Invitrogen), and corresponding PCR-products were printed onto microscopic slides. Hybridizations with dye-swap replicates were made for each individual, where RNA from one of the control LCL was used as common reference in all experiments. Two different labeling and hybridization protocols were used: indirect labeling with Micromax™ TSA™ Labeling and Detection Kit (NEN® Life Science products) was used in 15 experiments and direct Cy-dNTP incorporation, using a modified CyScribe™ First Strand cDNA Labelling Kit (Amersham Biosciences) protocol was used in 9 experiments. Eight µg of total RNA were used for the Micromax™ TSA™ labeling and 25 µg of total RNA for the CyScribe labeling. The microarrays were scanned with a GenePix 4000B (Axon instruments) scanner and the fluorescent intensities were quantified with GenePix Pro 3.0 software (Axon Instruments). The data were normalized using a lowess algorithm, which is
implemented in the statistical software package R [17]. The normalization was intensity and print-tip dependent and no background subtraction was used. Intensity signals from spots not detected by the quantification software due to low intensities, or spots with artifacts on the arrays were removed in order to exclude irrelevant data. Ratios deviating more than three times between the two dye swap experiments were also removed. A mean value was calculated for each gene and missing values were imputed with a K-nearest neighbor method in SAM (Significance Analysis of Microarrays) [18]. A maximum of two imputations per group and gene was allowed. Two-class SAM was used for identifying candidate genes through an algorithm based on a modified t-test. A score, obtained from the difference between groups and variation within groups, was calculated for each gene and used for ranking. Q-values were produced from permutation of class labels and were used as significance estimates.

Quantitative real-time PCR (qPCR) analysis was performed for SEMA5A, HSPA1L, CCND2 and SAT on total RNA from the same 12 individuals. cDNA was synthesized using M-MLV enzyme (USB®) and random priming and for SEMA5A, HSPA1L, CCND2 and SAT were amplified and detected using TaqMan® Assays-on-Demand™ Gene Expression products (Applied Biosystems) and the ABI PRISM® 7000 Sequence Detection System (Applied Biosystems). mRNA levels of β-actin (ACTB) were used for normalization. The mRNA levels were detected in triplicates and mean values calculated after exclusion of outliers.

**Results**

After filtration, 2,101 genes remained for further analysis. The majority of the genes were removed due to low expression levels, with resulting signal intensities that could not be detected during the quantification process. Nine genes were identified with significantly altered levels in patient samples. All nine genes were down-regulated, ranging between -1.38 and -1.67 (Table 2). The q-values (0.095) suggest that one of the nine genes may be a false positive and erroneously identified as down-regulated. The magnitude of the changes is low, but in the same range as those found in the microarray study on autism performed by Purcell et al [19]. None of the genes is positioned in the 7q31 region. Table 3 presents the biological functions of the nine genes identified. Four of these genes, SEMA5A, HSPA1L, CCND2 and SAT, encode proteins involved in biological processes associated with brain function or the immune system, and are consequently candidate genes for idiopathic autism. Quantitative PCR confirms the array results for the SEMA5A gene, but not for the HSPA1L, CCND2 and SAT genes (Fig.1). One outlier in the patient group, case 2, reduced the statistical significance for SEMA5A, resulting in a group mean ratio of 0.96 compared to 2.82 without the outlier. The deviating measurement was well outside three standard deviations from the group mean and thus excluded from the analysis as an outlier [20]. The reason for the outlier in the qPCR experiments is unclear since it does not correspond to the results of the microarray analysis. In conclusion, the cDNA microarray analysis and the quantitative real-time PCR result are consistent and suggest a down-regulation of SEMA5A in the samples analyzed.
Discussion

In this report we have used cDNA microarray technology to analyze the expression pattern of 7,700 genes in LCLs from six non-related individuals with idiopathic autism. Each patient was selected from familial forms of AD (sib-pairs), in which at least one chromosome 7q-region segregated with the disorder. The analysis resulted in the identification of nine genes, all of which were down-regulated. None of these genes have been reported previously as associated with autism. Notably, one of the nine genes encodes semaphorin 5A, which is involved in axonal guidance [21]. The SEMA5A gene is positioned at chromosome 5p15, approximately 6 Mb from a marker indicating increased allele sharing in our previous genome-wide scan [7]. Rodent expression of Sema5A has been reported in cerebral cortex, basal ganglia, thalamus and hippocampus in embryonic and postnatal rat brains [22]. Semaphorin 5A is a bifunctional molecule, acting both as attracting and repulsing guidance cues [23,24]. Failure of Sema5A expression leads to abnormal development of the axonal connections in the forebrain of mice [25], which may affect the formation of functional synapses. Furthermore, it has been suggested that haploinsufficiency for SEMA5A is responsible for mental retardation in Cri-du-chat, a syndrome resulting from deletions of chromosome 5p [26]. Synaptic dysfunction of specific neurons is likely to be one etiologic mechanism in AD and this was supported by the identification of mutations in the NLGN3 and NLGN4 genes in a subgroup of patients with autism [13,14]. The NLGN3 and NLGN4 genes encode synaptic cell-adhesion molecules. A down-regulation of SEMA5A, with an effect on axonal guidance, may therefore be an analogous candidate mechanism behind AD in some patients.

The HSPA1L, CCND2 and SAT genes were also found to be down-regulated in our microarray study. These genes were of interest due to their involvement in the immune system or brain development [27-34], but differences in mRNA expression levels could not be confirmed by qPCR. This indicates that these genes may be false positives in our microarray analysis.

To our knowledge, no previous findings suggest association between the remaining five genes down-regulated in this study (CSDA, SARCOSIN, MGST1, RPN2 and SNX2) and mechanisms related to brain development or brain function.

Several genes in the 7q31 region, such as RAY1/ST7, WNT2, CORTB2 and RELN have been proposed as candidate genes for autism [35-38]. In addition, linkage and association to the β3 subunit of the GABA_3 receptor GABRB3 gene on chromosome 15q have been reported by several groups [39-41]. Unfortunately, the WNT2, RELN and GABRB3 genes were not expressed at levels high enough to detect true differences in our LCL samples. The results from a previous microarray analysis of brain specimens suggest that variation in the expression of EAAT1 and GluR1, two members of the glutamate system, are associated with autism [19]. We could not confirm changes in expression of these two candidate genes in our study. The RAY1/ST7, CORTB2, NLGN3 and NLGN4 genes were not present on the microarrays used in this study.

Idiopathic autism is a genetically complex disorder, manifested in the brain. Since idiopathic autism is constitutional, altered transcript levels may be detected also in non-neuronal tissues, such as
lymphoblastoid cells, which are used as a model in this study. Consequently, the significance of our findings in the etiology of AD depends on the presence of constitutional changes in mRNA levels from genes expressed in both brain and lymphocytes. The small number of individuals included renders generalized conclusions hazardous, and the risk of separate false positives should be considered. Another limitation is the sensitivity of the microarray system. Small variations in expression levels may be difficult to detect and is further complicated for genes with low constitutional expression levels.

The significance of the down-regulated genes in this study is yet unclear. Further analyses, including a larger number of patient samples and preferably on brain tissue from AD patients, are needed to confirm a possible involvement of these transcripts in the disease process. The microarray technique still provides a useful tool for the investigation of mechanisms behind complex psychiatric disorders, and the method may contribute to clarify both primary and secondary events in these disorders.

Acknowledgement

This work was supported by grants from the Swedish Research Council, the Children’s Cancer Foundation of Sweden, the Swedish Cancer Society, the Swedish Medical Society, the Sävstaholm Society, and the Torsten and Ragnar Söderbergs Fund. The experiments comply with the current laws of Sweden.

References


42. Schopler E, Reichler R, Renner BR: The Childhood Autism Rating Scale (CARS) for diagnostic screening and classification of autism. Irvington 1986;New York:
Table 1. Clinical features of patients

<table>
<thead>
<tr>
<th>Case</th>
<th>Sex</th>
<th>Age $^a$</th>
<th>Ethnicity</th>
<th>Diagnosis</th>
<th>Onset Months</th>
<th>Regression</th>
<th>Language</th>
<th>ADI-R criteria</th>
<th>CARS</th>
<th>Cognitive level $^b$</th>
<th>Seizures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>f</td>
<td>7.0</td>
<td>Caucasian</td>
<td>autism</td>
<td>30</td>
<td>Yes (at 30 months)</td>
<td>non verbal</td>
<td>Yes</td>
<td>36.5</td>
<td>severe mental retardation</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>m</td>
<td>10.4</td>
<td>Caucasian</td>
<td>autism</td>
<td>24</td>
<td>No</td>
<td>non verbal</td>
<td>Yes</td>
<td>44</td>
<td>severe mental retardation</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>m</td>
<td>13.6</td>
<td>Caucasian</td>
<td>autism</td>
<td>35</td>
<td>Yes (at 35 months)</td>
<td>non verbal</td>
<td>Yes</td>
<td>37</td>
<td>mild mental retardation</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>f</td>
<td>6.9</td>
<td>Caucasian</td>
<td>autism</td>
<td>15</td>
<td>No</td>
<td>delayed (first phrases 60 months)</td>
<td>Yes</td>
<td>43</td>
<td>borderline intellectual functioning</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>m</td>
<td>10.5</td>
<td>Caucasian</td>
<td>autism</td>
<td>15</td>
<td>No</td>
<td>delayed (first words 25 months, phrases 37 months)</td>
<td>Yes</td>
<td>34</td>
<td>borderline intellectual functioning</td>
<td>–</td>
</tr>
<tr>
<td>6</td>
<td>m</td>
<td>4.0</td>
<td>Caucasian</td>
<td>autism</td>
<td>19</td>
<td>No</td>
<td>non verbal</td>
<td>Yes</td>
<td>44.5</td>
<td>mild mental retardation</td>
<td>–</td>
</tr>
</tbody>
</table>

ADI-R, Autism Diagnostic Interview-Revised [15]; CARS, Childhood Autism Rating Scale [42] (scores between 30 and 36.5 indicate mild to moderate autism while scores greater than 36.5 reflect severe autism).

$^a$Age in years at which individual was evaluated for inclusion in the study.

$^b$Assessed using age and developmentally appropriate instruments, including the Psychoeducational Profile, the Wechsler Intelligence Scale for Children-Revised, and the Leiter.
Table 2. Differentially expressed genes identified by SAM analysis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr. Location</th>
<th>Score</th>
<th>Fold change</th>
<th>q value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSDA</td>
<td>12p13</td>
<td>-2.96</td>
<td>-1.67</td>
<td>0.095</td>
</tr>
<tr>
<td>SARCOSIN</td>
<td>2q31</td>
<td>-2.53</td>
<td>-1.37</td>
<td>0.095</td>
</tr>
<tr>
<td>SEMA5A</td>
<td>5p15</td>
<td>-2.43</td>
<td>-1.45</td>
<td>0.095</td>
</tr>
<tr>
<td>MGST1</td>
<td>12p12</td>
<td>-2.27</td>
<td>-1.50</td>
<td>0.095</td>
</tr>
<tr>
<td>SAT</td>
<td>Xp22</td>
<td>-2.19</td>
<td>-1.42</td>
<td>0.095</td>
</tr>
<tr>
<td>CCND2</td>
<td>12p13</td>
<td>-2.17</td>
<td>-1.65</td>
<td>0.095</td>
</tr>
<tr>
<td>HSPA1L</td>
<td>6p21</td>
<td>-2.16</td>
<td>-1.66</td>
<td>0.095</td>
</tr>
<tr>
<td>RPN2</td>
<td>20q12-q13</td>
<td>-2.12</td>
<td>-1.40</td>
<td>0.095</td>
</tr>
<tr>
<td>SNX2</td>
<td>5q23</td>
<td>-2.06</td>
<td>-1.38</td>
<td>0.095</td>
</tr>
</tbody>
</table>

Score, fold change and q values are derived from SAM.

Table 3. Biological mechanisms related to nine down-regulated genes

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
<th>Molecular function</th>
<th>Biological process</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSDA</td>
<td>Cold shock domain protein A</td>
<td>Transcription co-repressor activity</td>
<td>Cold response</td>
</tr>
<tr>
<td>SARCOSIN</td>
<td>Sarcomeric muscle protein</td>
<td>Protein binding activity</td>
<td>Striated muscle contraction</td>
</tr>
<tr>
<td>SEMA5A</td>
<td>Semaphorin 5A</td>
<td>Receptor activity</td>
<td>Neurogenesis</td>
</tr>
<tr>
<td>MGST1</td>
<td>Microsomal glutathione S-</td>
<td>Glutathione transferase activity</td>
<td>Cellular defense (against oxidative stress)</td>
</tr>
<tr>
<td></td>
<td>transferase 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT</td>
<td>Spermidine/spermine N(1)-</td>
<td>Acetyltransferase activity</td>
<td>Polyamine catabolism</td>
</tr>
<tr>
<td></td>
<td>acetyltransferase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCND2</td>
<td>Cyclin D2</td>
<td>Forms serine/threonine kinase holoenzyme complex</td>
<td>Cell cycle regulation</td>
</tr>
<tr>
<td>HSPA1L</td>
<td>Heat shock 70kDa protein 1-like</td>
<td>Heat shock protein activity</td>
<td>Heat shock response</td>
</tr>
<tr>
<td>RPN2</td>
<td>Ribophorin II</td>
<td>Transferase activity</td>
<td>Protein modification</td>
</tr>
<tr>
<td>SNX2</td>
<td>Sorting nexin 2</td>
<td>Protein transporter activity</td>
<td>Intracellular trafficking</td>
</tr>
</tbody>
</table>
Fig. 1. Relative down-regulation of SEMA5A, HSPA1L, CCND2 and SAT mRNA levels in Epstein-Barr virus-transformed cells derived from patients with AD compared to healthy controls. Fold-changes are based on microarray (■) and qPCR (■) analysis. Error bars indicate standard deviations.