Experimental Report

New insights into candidate genes for autism spectrum disorder in 8p23.1 duplication syndrome

Maytza M. Corrêa 1,*, Thiago Corrêa 2, Cíntia B. Santos-Rebouças 3, Marino Miloca Rodrigues 4, Gisele Rozone de Luca 5, Louise Lapagesse de Camargo Pinto 5

1 Neurology Pediatric Unit, Universidade Federal do Paraná, Curitiba, PR, Brazil.
2 Estácio | IDOMED, Jaraguá do Sul, Brazil.
3 Department of Genetics, Institute of Biology Roberto Alcantara Gomes, State University of Rio de Janeiro, Rio de Janeiro, Brazil.
4 Children’s Hospital Jeser Amarante Faria, Joinville, Brazil.
5 Children’s Hospital Joana de Gusmão, Florianópolis, Brazil.

* Correspondence: maytzac@gmail.com.

Abstract: The 8p23.1 duplication syndrome is a rare condition, characterized by dysmorphisms, intellectual disability, congenital cardiac anomalies, and autism spectrum disorder (ASD). The current model for explaining the pathogenesis of this condition postulates that few dosage-sensitive genes within the duplication are sufficient for the core clinical features, although the molecular mechanisms leading to the ASD presentation remain to be solved. Herein, we described clinical and cytogenetic findings of an 8p23.1 duplication in a boy with mild facial dysmorphisms, cardiac anomalies and ASD. Therefore, we investigated the influence of duplicated genes on the pathophysiology of ASD in our patient. We identified four duplicated genes (BLK, GATA4, PINX1, TNKS) connected with proteins previously associated with ASD and involved in significant enriched pathways associated with human neurological conditions. Moreover, the candidate genes are highly expressed in brain regions associated to ASD, such as the hippocampus. Taken together, these results point out crucial interactions among BLK, GATA4, PINX1, and TNKS and genes associated with ASD. We indicate cellular networks perturbations encompassing neuronal development pathways related to our patient’s condition. Thus, these findings bring new insights into the genetic basis of ASD in patients with 8p23.1 duplication syndrome.

Keywords: 8p23.1 syndrome; Systems Biology; Duplication; ASD; Chromosome rearrangement.

1. Introduction

The 8p23.1 duplication syndrome is characterized by mild facial dysmorphisms, behavioral abnormalities, mild to moderate delayed speech and language development, intellectual disability, and congenital cardiac anomalies [1]. Additional clinical findings include cleft lip and/or palate, macrocephaly, seizures, attention deficit hyperactivity disorder (ADHD), autism spectrum disorder [ASD], ocular anomalies, gait imbalance, hypotonia, and hydrocele[2, 3]. Nonetheless, phenotypic features present variable penetrance, and can be compatible with independent adult life. The condition is rare, with an estimated prevalence of 1 in 58,000 births, resulting from a recurrent duplication at the chromosomal region 8p23.1, encompassing a minimal overlapping region of 3.66 Mb [1]. Moreover, the origin of the chromosome rearrangement can be de novo or maternally/paternally inherited.

The current pathogenesis model for 8p23.1 duplication syndrome indicates that such condition is caused by few dosage-sensitive genes, resulting from the duplicated region and their interactors [4]. Thus, three candidate dosage-sensitive genes have been
associated with the pathogenesis of the 8p23.1 duplication syndrome so far (SOX7, TNKS, and GATA4).

Herein, we reported a boy with a 3.7 Mb duplication involving the 8p23.1 region. The main clinical findings were mild facial dysmorphisms, hypotonia, cardiac anomalies, ASD and motor development. As ASD symptoms are poorly explored in the dup8p23.1 context, a network-based approach was applied to investigate the impact of the duplicated genes within the chromosome rearrangement on the presence of ASD in our patient.

2. Material and Methods

2.1 Chromosome microarray (CMA)

The CytoScan® 750 K platform was used for the detection of copy number variations (CNVs) in the proband. The results were analyzed with Chromosome Analysis Suite (ChAS) Affymetrix® software based on the human genome version of February 2009 [GRCh37/hg19].

2.2 Protein-protein interactions networks

The human interactome was recovered from the Human Integrated Protein-Protein Interaction Reference (HIPPIE) database (version 2.2) [http://cbdm.uni-mainz.de/hippie/][5]. Protein-protein interactions of coding genes were extracted from the duplicated region in the human interactome and their first neighbours were selected to expand and generate a subnetwork. To identify candidate genes that may contribute to ASD, a list of known ASD genes (DisGeNET; http://www.disgenet.org/) was used [6]. After that, only the proteins from the duplicated region that directly interact with proteins previously associated with ASD were filtered in Cytoscape V.3.7.0. software [7].

2.3 Expression and tissue-specific gene-network

The confidence value of the expression for each of these genes was identified in the central nervous system [CNS] using data from the HumanBase database (https://hb.flatironinstitute.org/) [8]. In the tissue-specific interactions, evidence from co-expression, protein-interaction, transcription factors binding, microRNA targets, and perturbations were retrieved.

2.4 Enrichment analysis

The webserver Enrichr (https://maayanlab.cloud/Enrichr/) [9] was used to identify significant pathways. The gene-set libraries used were BioPlanet, Kegg, Reactome, and WikiPathways. Only terms with a p-value <0.05 were retrieved.

2.5 Investigation of genes associated with ASD

Genes associated with ASD were identified by using the SFARI Gene web portal (https://gene.sfari.org/). For each node in the network, the score that reflects the strength of the evidence related to the development of ASD from peer-reviewed scientific and clinical studies was extracted [10]. Then, the scores were classified into three distinct categories: 1 (high confidence); 2 (strong candidate); 3 (suggestive evidence).

3. Results

The proband was born at 40 weeks of gestation (Apgar score of 7/9) as the third child of nonconsanguineous parents. His weight was 3,000 g (z-score ≥ −2 and < +2), length 48 cm (z-score ≥ −2 and < +2) and cephalic perimeter [z-score ≥ −1 and < +1]. Interventricular and atrial communication was identified during a morphological ultrasound yet in the 22th week of gestation. On the third day of life, the patient presented dyspnea when breast-feeding, hypotonia, and a systolic murmur. The echocardiogram identified a coarctation of the aorta and intraventricular communication, requiring a surgery in the thirteenth day
of life. During the procedure, an end-to-end anastomosis of the aorta and an intraventricular and atrial communication bandaging was performed without complications. Nevertheless, the patient exhibited a significant heart failure and interventricular dyssynchrony after the procedure.

During the first childhood, the proband presented delayed psychomotor development and his developmental hallmarks included social smile at two months, sat with support at 5 months, sat without support at 10 months, held his head at 10 months, crawled at 1 year-old, walked at 1 year and 6 months with physiotherapy assistance.

At 2 years and 7 months, the proband presented speech delay and did not emit any word. He did not make eye contact, manifested stereotypy, line up objects, little socialization, and did not play with toys in a playful way. The mother reports that he had aggressive moments with family and strangers. Moreover, the patient had feeding difficulties, weight gain, and mild facial dysmorphisms with a prominent forehead and arched eyebrows. After the diagnosis of secondary ASD, the patient was medicated with Risperidone 1mg/mL using 0.75ml daily, which attenuated his aggressiveness, self-aggressiveness, and improved his socialization. His mother and two sisters have no history of comorbidities, while the father had a delay in psychomotor and language development with subsequent learning disability. However, the father was unavailable to be tested.

Currently, at 3 years-old, the patient has appropriate weight and height for his age and undergoes treatment for hypothyroidism. The CMA revealed a 3.7 Mb pathogenic microduplication in chromosome 8, encompassing the region 8p23.1 (minimal interval - chr8:8107752-11845324; GRCh37/hg19) (Figure 1).

![Figure 1](image)

The **TNKS** was identified as the main candidate gene in the duplicated region associated with the ASD (Table -1). To investigate the impact of TNKS and additional dosage-sensitive genes related to ASD in the 8p23.1 duplication, a gene network-based approach was applied in the human interactome context to identify ASD-related potentially disturbed interactions in the patient. Three proteins from the duplicated region other than TNKS (GATA4, PINX1, BLK) directly connecting with other proteins previously associated with ASD were identified in this study.
Table 1. Comparison between the molecular cytogenetic findings and neurological manifestations in the current patient and previously reported patients with the 8p23.1 duplication syndrome.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CMA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duplication size (Mbp)</td>
<td>3.7</td>
<td>1.8</td>
<td>0.44</td>
<td>1.13</td>
<td>3.5</td>
<td>3.6</td>
</tr>
<tr>
<td>Duplication location on chr:8 (GRCh37/hg19)</td>
<td>8,107,752-11,845,324</td>
<td>9,169,154-10,969,075</td>
<td>9,124,781-9,566,076</td>
<td>8,568,207-9,702,442</td>
<td>8,334,478-11,875,439</td>
<td>8,273,138-11,948,421</td>
</tr>
<tr>
<td>Sex</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Age (Year)</td>
<td>2</td>
<td>11</td>
<td>23</td>
<td>14</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Main protein-coding genes duplicated</td>
<td>BLK; GATA4; TNKS; PINX1; XKR6</td>
<td>SOX7; PINX1; PP1R3B; TNKS</td>
<td>MFHAS1; PINX1; TNKS</td>
<td>BLK; GATA4; PINX1; TNKS</td>
<td>BLK; GATA4; PINX1; TNKS</td>
<td></td>
</tr>
<tr>
<td>Main features</td>
<td>Facial dysmorphisms ASD; Motor development; ASD; ID</td>
<td>Cardiac anomalies; ASD; ID</td>
<td>ASD; ID</td>
<td>ASD; ID</td>
<td>Frontal bossing; Macrocephaly ASD; ID</td>
<td>Hearing impairment; Renal agenesis; ASD; ID</td>
</tr>
</tbody>
</table>

In an attempt to recognize functional relationships among these genes, a tissue-specific network (TSN) with interactions in the context of the CNS was generated (Figure 2).

3. Discussions

The 8p23.1 duplication syndrome is a rare condition with very few cases involving individuals with ASD. Therefore, the etiopathology of ASD in patients with the dup8p23.1 syndrome is poorly studied and remains unclear. It is worth mentioning that TNKS and GATA4 genes have already been described as contributing to the behavioral difficulties and cardiac anomalies, respectively [2, 3]. However, this is not the case of PINX1, which is involved in telomere integrity and in which variants were associated with late-onset Alzheimer disease [11]. In the same way, BLK was not previously related to ASD. It encodes a B lymphoid tyrosine kinase involved with behavioral problems, such as the development of neuroticism [12]. Interestingly, GATA4 acts as a transcription factor regulating the expression of brain natriuretic peptide, involved with levels of stress and anxiety when detected in the CNS [13, 14]. Besides, GATA4 was found altered in Huntington's disease [15]. Moreover, non-duplicated genes such as MET, GLRA2, MEF2C, APC, DISC1, UBE3A, BCL2, and NFI present in TSN are also involved in the ASD development. Several types of association involve these genes in the DisGeNET, such as genetic variation, biomarker, altered expression, or causal mutation, impacting crucial bioprocesses, such as neurogenesis, and developmental synaptic plasticity [6].

The significant pathways identified in this study are related to neurological disorders in humans, when disturbed. Many terms identified in fig. 2b, such as L1CAM interactions, DCC mediated attractive signaling, nephrin interactions, and Wnt signaling pathway are known to play a fundamental role in axon guidance, an important biological process, in which the neurons extend axons and reach their targets to form synaptic junctions,
connecting the central and peripheral CNS during embryonic development [16, 17]. In this sense, the incorrect or incomplete synapse formation during the development of the neural circuit is a significant component that can contribute to the structural basis of ASD pathology [18, 19]. Furthermore, functional enrichment of other bioprocesses related to ASD in humans, including synaptic signaling pathways associated with autism, synaptic junction, serotonin receptor 2, and cholinergic synapse were found [20–23].

Figure 2. Network-based approach for identifying ASD candidate genes in the 8p23.1 duplication syndrome. (a) tissue-specific network (TSN) composed of 26 nodes/proteins and 213 interactions. Blue nodes are protein-coding genes from the duplicated region; dark grey nodes represent ASD-genes; light grey nodes were added by the database. (b) heat map of genes from TSN with the most significant biological processes. (c) heat map of expression of the blue and dark grey nodes of the TSN in the CNS. The confidence value was calculated between 0-1. (d) network with TNKS and its first neighbours. Nodes with yellow borders show scores associated with the development of ASD. Bellow, the scores were classified into: 1 (high confidence); 2 (strong candidate); 3 (suggestive evidence).

Additionally, the degree of expression in human tissues was interrogated for each candidate gene in TSN. The CNS was the tissue with the highest expression levels, supporting the idea that changes in the dosage of these genes may negatively impact neurological functions. The results include genes highly expressed in regions altered in the ASD
involved in cognitive functions, such as the hippocampus, cerebral cortex, amygdala, and basal ganglion [24] (Fig. 2c). Amongst the four candidate genes duplicated in the patient, TNKS obtained the most significant expression levels in these regions, indicating a prominent relevance in the maintenance of brain functions. TNKS is involved in diverse cellular functions, like Wnt signaling, telomere length, vesicle trafficking, and proteasome regulation [25–27], which is supported by a report considering it as the main candidate gene for the emergence of ASD in 8p23.1 duplication syndrome [3] (Tab. 1).

In TSN, besides the connections with the genes already associated with ASD, interactions between TNKS and genes associated with high scores to ASD from SFARI, such as TRIO, DST, SMARCA2, MAP1A, and SYNE1 were revealed (Fig. 2d). Interestingly, the genes TRIO (cri-du-chat region), SMARCA2, SYNE1, NF1 (NF1-microdeletion syndrome region), and DISC1 were found to be duplicated in several patients with ASD in the DE-CHIPER database [28]. DISC1 is involved in the control of many features in embryonic and adult neurogenesis [29]. Pathogenic variants in DISC1 are associated with neuropsychiatric diseases, including schizophrenia, major depression, bipolar disorder, ASD, and Asperger syndrome [30, 31]. Furthermore, two brothers with a DISC1 duplication in the 1q42.2 region exhibiting ASD and intellectual disability were reported [32].

4. Conclusions

We identified functional relationships from the human interactome between the duplicated genes in 8p23.1 region and genes associated with ASD with essential role in the neurological functioning. Disturbances involving components with significant expression in the CNS associated in the neuronal development pathways could be implicate to our patient's phenotype and suggest a relevant contribution to the genetic basis of ASD in 8p23.1 duplication syndrome.

Funding: None.

Research Ethics Committee Approval: The study has been carried out in accordance with the code of ethics of the world medical association. The study includes a statement on ethics approval by the Children’s Hospital Jeser Amarante Faria (approval number 4.944.246). Written informed consent was obtained from the parent/legal guardian of the patient for publication of the details of their medical case and any accompanying images.

Acknowledgments: None.

Conflicts of Interest: The authors declare no conflict of interest.

Supplementary Materials: The following supporting information can be downloaded at: https://bjcasereports.com.br/index.php/bjcr/new_insights_into_candidate_genes_sl, Table S1: Genes and pathways.

References


