

# Genomic variants and inferred biological processes in multiplex families with Tourette syndrome

Jakub P. Fichna, PhD\*; Małgorzata Borczyk, PhD\*; Marcin Piechota, PhD;  
Michał Korostynski, PhD; Cezary Zekanowski, PhD†; Piotr Janik, MD, PhD†

**Background:** Tourette syndrome is a developmental neuropsychiatric disorder. Its etiology is complex and elusive, although an important role of genetic factors has been established. The aim of the present study was to identify the genomic basis of Tourette syndrome in a group of families with affected members in 2 or 3 generations. **Methods:** Whole-genome sequencing was performed followed by co-segregation and bioinformatic analyses. Identified variants were used to select candidate genes, which were then subjected to gene ontology and pathway enrichment analysis. **Results:** The study group included 17 families comprising 80 patients with Tourette syndrome and 44 healthy family members. Co-segregation analysis and subsequent prioritization of variants pinpointed 37 rare and possibly pathogenic variants shared among affected individuals within a single family. Three such variants, in the *ALDH2*, *DLD* and *ALDH1B1* genes, could influence oxidoreductase activity in the brain. Two variants, in *SLC17A8* and *BSN* genes, were involved in sensory processing of sound by inner hair cells of the cochlea. Enrichment analysis of genes whose rare variants were present in all patients from at least 2 families identified significant gene sets implicated in cell–cell adhesion, cell junction assembly and organization, processing of sound, synapse assembly, and synaptic signalling processes. **Limitations:** We did not examine intergenic variants, but they still could influence clinical phenotype. **Conclusion:** Our results provide a further argument for a role of adhesion molecules and synaptic transmission in neuropsychiatric diseases. Moreover, an involvement of processes related to oxidative stress response and sound-sensing in the pathology of Tourette syndrome seems likely.

## Introduction

Tourette syndrome is a neurodevelopmental disorder characterized by motor and vocal/phonic tics persisting for longer than a year. The clinical phenotype of Tourette syndrome belongs to the spectrum of tic disorders.<sup>1</sup> Moreover, most patients with Tourette syndrome present a variety of additional symptoms due to psychiatric comorbidities, including attention-deficit/hyperactivity disorder (ADHD), obsessive–compulsive disorder (OCD), autism spectrum disorder (ASD), affective disorders, anxiety disorders, impulse control disorders and personality disorders, implying an overlapping etiology.<sup>2,3</sup> The tics usually emerge around the age of 4–6 years and are most severe after 5 years, albeit most cases improve during adolescence. The prevalence of Tourette syndrome in the general pediatric population ranges from 0.3% to 0.77%, and between 0.005% and 0.065% in adulthood.<sup>4</sup> Other tic disorders are more

common than Tourette syndrome and affect as much as 5% of the general population.<sup>3</sup>

Tourette syndrome and tic disorders in general have an important genetic component, with heritability estimated at 60%–80%.<sup>2,5</sup> However, the clinical phenotype may be influenced by environmental, prenatal and perinatal factors; hormonal disturbances; and psychosocial stressors interacting with multiple genes.<sup>6–9</sup> Based on the candidate gene approach and linkage studies, multiple genes have been suggested to be important in the etiology of Tourette syndrome.<sup>10,11</sup> The protein products of these genes are involved in neurotransmitter signalling, synapse development, organization and functioning, differentiation of axons, cell adhesion and mitochondrial activity.<sup>11,12</sup> However, the Human Phenotype Ontology database still lists only 2 genes, *HDC* and *SLITRK1*, as being involved in Tourette syndrome.<sup>13</sup>

Recent next-generation sequencing studies showed a complex genetic background of Tourette syndrome involving

**Correspondence to:** J.P. Fichna, Department of Neurogenetics and Functional Genomics, Mossakowski Medical Research Institute, Polish Academy of Sciences, Pawlinskiego 5 Street, 02-106 Warszawa, Poland; jfichna@imdik.pan.pl

\*Share first authorship.

†Share senior authorship.

Submitted Nov. 9, 2022; Revised Jan. 9, 2023; Accepted Mar. 13, 2023

**Cite as:** *J Psychiatry Neurosci* 2023 May 19;48(3). doi: 10.1503/jpn.220206

multiple interacting genes<sup>14</sup> and a role of the rare variant burden in tic disorders.<sup>15</sup> The available data suggest that de novo variants in approximately 400 genes contribute to Tourette syndrome risk in 12% of clinical cases.<sup>16</sup> It is increasingly evident that rare pathogenic variants in a single gene cannot be responsible for a substantial fraction of Tourette syndrome cases.

Results of genome-wide association studies (GWAS) explain as much as 21% of Tourette syndrome heritability by polymorphisms, with a minor allele frequency (MAF) between 0.1% and 5%.<sup>8</sup> The GWAS-based polygenic risk scores of tic disorders suggest that low-impact common variants, found also in the general population, contribute to the disease in a polygenic manner.<sup>17</sup> Other GWAS have shown that Tourette syndrome is correlated with OCD, ADHD and major depressive disorder, diseases known to have an overlapping and highly polygenic background.<sup>18</sup>

Moreover, epigenetic mechanisms and gene regulation by noncoding RNAs have been proposed to mediate the influence of environmental factors on the genetic background of Tourette syndrome. Also, the role of rare noncoding variants remains largely unexplored in tic disorders. Thus, owing to the complex and heterogeneous genetic architecture of tic disorders, with common and rare variants in different types of genes associated with numerous biological pathways, the identification of susceptibility genes has been challenging.

We hypothesize that most cases of familial Tourette syndrome in the Polish population can be explained by an oligogenic inheritance of multiple variants. To verify this hypothesis we analyzed a group of families comprising people with Tourette syndrome as well as healthy family members using whole-genome sequencing to identify ultra-rare, rare and uncommon variants associated with Tourette syndrome.

## Methods

### Study sample

All patients with tic disorders were recruited from a single outpatient clinic and assessed by the same clinician specialized in tic disorders according to *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR, DSM-5) criteria. Patients were systematically interviewed using a semistructured interview based on the TIC (Tourette syndrome International database Consortium) data entry form.<sup>19</sup> The prevalence of the most common comorbid disorders was evaluated and included ADHD; OCD; depression; anxiety disorders, including phobias, panic disorders, generalized anxiety disorder and separation anxiety disorder; oppositional defiant disorder; conduct disorder; and ASD. The list of obsessions and compulsions included in the Yale–Brown Obsessive Compulsive Scale (Y-BOCS) was used to establish the clinical spectrum of OCD. Each patient was carefully questioned about all the symptoms included in the DSM as the diagnostic criteria for the above-mentioned comorbid disorders. Children and adolescents were assessed using the M.I.N.I. International Neuropsychiatric Interview for Children and Adolescents. Previous diagnoses of mental disorders that had been made in psychiatric clinics

were accepted, and patients with severe psychiatric comorbidities were referred to a psychiatrist to confirm the diagnosis. All patients were from the Polish population, which is an ethnically homogeneous subgroup of White people.<sup>20</sup> The study was approved by the Ethics Committee of the Medical University of Warsaw (KB/2/2007, KB/53/A/2010, KB/63/A/2018) and was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All participants or their legal representatives gave informed consent before inclusion in the study.

In case of a positive family history of a tic disorder, DNA was collected from all available affected relatives and healthy members of the proband's family. Each patient was assigned to one of the following groups: Tourette syndrome, other tic disorders, or healthy controls.

### Whole-genome sequencing

Participants' DNA was extracted from peripheral blood leukocytes using a standard salting-out method<sup>21</sup> or from saliva (Oragene DNA Self Collection Kit and Prep IT L2P Purification Kit, DNA Genotek Inc.). Whole-genome sequencing was performed by Novogene (Beijing, China) according to the following protocol. Sequencing libraries were generated using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) following the manufacturer's recommendations. DNA was randomly fragmented to 350 bp on average with a Bioruptor, and DNA fragments were size-selected with sample purification beads. The selected fragments were then end-polished, A-tailed and ligated with a full-length adaptor. The fragments were then filtered with the beads again. Finally, the libraries were analyzed for size distribution on an Agilent 2100 Bioanalyzer, quantified using real-time polymerase chain reaction and paired-end sequenced on an Illumina high-throughput HiSeq X Ten sequencer.

Fastq files were processed using the Intelliseq Germline Pipeline (<https://gitlab.com/intelliseq/workflows>) built with Cromwell (<https://cromwell.readthedocs.io/en/stable/>) according to GATK best practices. Variants were called with the GATK (4.0.3) HaplotypeCaller to yield genomic variant calling files (gvcf). All the code for file preprocessing and analysis is available at the GitHub repository of the project (<https://github.com/ippas/imdik-zekanowski-gts>) in the "Burden and family" section. Genotypes for the study group were obtained via the Hail run\_combiner function (version 0.2.64, <https://hail.is/>).

Data analysis and filtering were performed in Hail (version 0.2.79). Multiallelic variants were split. The following variants were filtered out of the analysis: repeated and low-quality sequences (University of California, Santa Cruz RepeatMasker track  $\pm 2$  bp from each interval), loci with more than 90% gnomAD (v3) samples with a read depth (DP) of 1 or lower, variants with mean DP of 5 or lower, variants with mean genotype quality (GQ) of 50 or lower, variants that did not conform to the Hardy–Weinberg Equilibrium (HWE) ( $p < 0.05$ ), variants with 3 or more samples with DP below 3, and variants with 30 or more samples with GQ lower than 30. The variants were annotated with gnomAD

v3.1,<sup>22</sup> combined annotation dependent depletion (CADD) scores,<sup>23</sup> human phenotype ontology (HPO)<sup>13</sup> and Ensembl Variant Effect Predictor (VEP).<sup>24</sup> For each family, only variants that passed all quality filters for the whole group, had a CADD score above 10 and were within a gene locus based on Genecode v.32<sup>25</sup> were taken for analysis.

### Variant and enrichment analysis

The variants were analyzed according to their segregation pattern in each family (Appendix 1, available at [www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content](http://www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content)) and were accepted if all patients within the family (with Tourette syndrome or tic disorder) had at least 1 alternative allele while none of the noncarriers had it. Only variants with the MAF in non-Finnish Europeans (NFE; based on gnomAD v 3.1) lower than the selected threshold (0.001, 0.01 or 0.05) were analyzed. Finally, the results were filtered to remove variants within genes that were called in only a single family.

A functional enrichment analysis was performed to investigate the biological relevance of the candidate genes using Metascape software.<sup>26</sup> Lists of genes with variants cosegregating with the disease in at least 2–5 families were provided. A  $\text{Log}_{10}(q)$  score lower than  $-2$  was considered statistically significant.

For each observation, a probability of random occurrence was calculated (probability of observation;  $P(\text{obs})$ ). The chance of any variant(s) fulfilling the assumed criteria in any of the families was calculated using data from the NFE population in gnomAD (MAF or a probability of occurrence of ultra-rare/rare/uncommon variant in a given gene), combined with probabilities of observing given segregation patterns in a

given family or families. The results were adjusted for multiple testing using Bonferroni correction on the number of families, or number of combinations of 2–7 families (Appendix 2, available at [www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content](http://www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content)).

## Results

### Participants

Our study sample included 17 multiplex families comprising 80 patients (40 with Tourette syndrome and 40 with other tic disorders) and 44 healthy family members (Table 1, Appendix 1 and Appendix 3, available at [www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content](http://www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content)). Using whole-genome sequencing and variant analysis ultra-rare (MAF < 0.1%), rare (MAF < 1%) and uncommon (MAF < 5%) single nucleotide variants were identified.

### Cosegregation analysis

Analysis revealed 8282 uncommon variants present in all the patients from any single family (Appendix 4, available at [www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content](http://www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content)), with 289 ultra-rare variants segregating according to the applied schemas and located in genes called in more than 1 family. Most of those variants were noncoding, including 4 variants located in genes encoding lincRNAs (*LINC02306*, *LINC02763*, *LINC01414*, *LINC00298*; Appendix 5, available at [www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content](http://www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content)), with only 44 missense variants identified in protein-coding genes (Table 2).

**Table 1: Characteristics of included families**

Family code	Members, <i>n</i> [M/F] ( <i>n</i> = 124)	Tourette syndrome, <i>n</i> ( <i>n</i> = 40)	Other tic disorders, <i>n</i> ( <i>n</i> = 40)	Cosegregation probability <i>P</i> (seg)	Healthy controls, <i>n</i> ( <i>n</i> = 44)	Comorbidities
A	7 [4/3]	3	1	0.0625	3	ANX
B	8 [5/3]	3	2	0.0156	3	ADHD, ANX, MDD, OCD
C	9 [4/5]	2	4	0.0078	3	ANX
D	7 [5/2]	2	2	0.0156	3	ANX, MDD, OCD
E	7 [4/3]	2	2	0.0625	3	–
F	11 [7/4]	4	4	0.0078	3	–
G	4 [2/2]	1	2	0.25	1	ANX
H	6 [4/2]	2	3	0.125 or 0.125*	1	ADHD, ANX, OCD
I	6 [4/2]	2	2	0.0625	2	ANX, MDD, OCD
J	5 [2/3]	1	2	0.25	2	ANX, MDD, OCD
R	6 [4/2]	1	3	0.125	2	ANX
S	6 [2/4]	4	0	0.125	2	MDD, OCD
T	9 [4/5]	2	3	0.125 or 0.25*	4	OCD
U	6 [5/1]	2	2	0.0625	2	–
W	14 [7/7]	4	4	0.000976	6	–
X	5 [2/3]	2	1	0.125	2	ADHD
Y	8 [4/4]	3	3	0.03125	2	MDD, OCD

ADHD = attention-deficit and hyperactivity disorder; ANX = anxiety disorder; F = female; M = male; MDD = major depression disorder; OCD = obsessive-compulsive disorder.

\*In families H and T, 2 segregation patterns were analyzed (see Appendix 1, available at [www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content](http://www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content)).

**Table 2: Characteristics of ultra-rare, rare and uncommon variants\***

Location/type of variant	Number of variants		
	Ultra-rare	Rare	Uncommon
Intronic	206	288	776
Exonic/missense	44	24	54
Noncoding exon	6	5	25
5' UTR	8	4	9
3' UTR	11	7	20
Intergenic	5	6	15
Splicing	4	2	5
Exonic/synonymous	3	1	2
Upstream	2	0	2
Downstream	0	0	1
Nonsense	0	0	1

MAF = minor allele frequency; UTR = untranslated region.

\*Included are variants found in genes called in at least 2 families (ultra-rare variants, MAF < 0.1%), at least 3 families (rare variants, MAF < 1%), or at least 5 families (uncommon variants, MAF < 5%).

### Rare variant analysis

Analysis of rare variants segregating according to the applied schemas in at least 1 family revealed 97 variants with a CADD score above 20 (i.e., predicted to be within the 1% most pathogenic variants in the genome). There were 37 variants with a CADD score above 30 (i.e., predicted to be within the 0.1% most pathogenic variants in the genome (Table 3).

Three ultra-rare variants were identified in all patients from 2 families each (Table 4)

### Enrichment analysis

There were 121 genes with ultra-rare variants identified in at least 2 families; among them 13 genes with variants occurred in 3 or more families (Appendix 2). Of those 121 genes, 100 were amenable to gene enrichment analysis by Metascape, which showed 26 biological processes to be significantly enriched (Appendix 6, Supplementary table a, available at [www.jpnp.ca/lookup/doi/10.1503/jpn.220206/tab-related-content](http://www.jpnp.ca/lookup/doi/10.1503/jpn.220206/tab-related-content)). As many as 60 genes were related to those 26 categories, with 38 genes involved in at least 2 significantly enriched processes.

There were 71 genes with rare variants identified in at least 3 families (Appendix 2), and Metascape could analyze 63 of those genes. In this set, 20 biological processes were found to be significantly enriched (Appendix 6, Supplementary table d).

Uncommon variants were identified in at least 4 families in 108 genes (Appendix 2), and Metascape could analyze 90 of these genes. In this set, 28 biological processes were found to be significantly enriched (Appendix 6, Supplementary table e).

The top significantly enriched processes, pathways, or cellular compartments from the analysis of all 3 data sets are presented in Table 5 and Appendix 7, available at [www.jpnp.ca/lookup/doi/10.1503/jpn.220206/tab-related-content](http://www.jpnp.ca/lookup/doi/10.1503/jpn.220206/tab-related-content).

## Discussion

To obtain insight into the genetic architecture of Tourette syndrome we studied 17 large families from the Polish population at risk for Tourette syndrome by performing whole-genome sequencing followed by bioinformatic analyses. We posit that variants in genes present in all affected family members in at least 2 families are highly likely to have a role in the disease etiology. Most of the discovered variants were in noncoding regions, mostly intronic and some in 3' and 5' untranslated regions (UTRs). These variants may affect the gene in which they reside and may modify regulatory processes (e.g., by modulating noncoding RNAs) or their target regions.<sup>27</sup> The CADD score, used here to prioritize genetic variants, gathers information about each variant from multiple sources, including conservation and epigenetic data.<sup>23</sup> Therefore, a high CADD score indicates a likely high deleteriousness of a variant, including impactful mutations in noncoding, albeit possibly regulatory, regions. The enrichment analysis of the obtained gene lists revealed their association with several processes and structures, including cell adhesion and synaptic signalling. Some of them have been previously reported to be involved in Tourette syndrome etiology.

### Variants cosegregating in 2 families

Three ultra-rare variants in genes not previously linked to Tourette syndrome (*NLRP12*, *COL25A1* and *IQGAP2*) segregated with the disease in 2 families. None of the genes is pathogenic according to the American College of Medical Genetics and Genomics (ACMG) criteria. However, the identified genes could be considered valid candidate genes in Tourette syndrome owing to their involvement in neurologic processes. *NLRP12* and *COL25A1* have been associated with neurologic disorders and the *IQGAP2*-encoded protein could, through interaction with other proteins, influence calcium sensing and synaptic plasticity.

A missense variant in the *NLRP12* gene was found in families G and I, one of the smallest in the study group (Table 1). The *NLRP12* protein is involved in the inflammasome assembly and signal transduction; NLRP and its inflammasomes were upregulated in depressed and suicidal people.<sup>28</sup> Polymorphisms in *NLRP12* were associated with alcohol misuse and depression,<sup>29,30</sup> and rare variants appeared to be causative of cold inflammatory syndrome<sup>31</sup> and multiple sclerosis.<sup>32</sup>

The brain-specific membrane-bound collagen gene *COL25A1* is expressed mainly in the hippocampus and the occipital lobe, with moderate expression in the frontal lobe and optic cranial nerve and low expression in the left cerebellum. The protein was first detected in amyloid preparations from Alzheimer disease brain and subsequently implicated in the pathogenesis of this disease.<sup>33,34</sup> Polymorphisms in the *COL25A1* gene were associated with antisocial personality disorder and substance dependence.<sup>35</sup> Rare pathogenic variants in *COL25A1* were reported in cases of congenital fibrosis of extraocular muscles.<sup>36</sup>

The *IQGAP2* gene is a liver-specific member of the IQGAP scaffold proteins family. These proteins facilitate the formation of complexes regulating cytokinesis, cytoskeletal dynamics, intracellular signalling, cell proliferation and migration.<sup>37</sup> *IQGAP2* binds to calmodulin,<sup>38</sup> and together with Munc13, forms a calcium sensor complex that controls short-term synaptic plasticity. Mutations in calcium/calmodulin-dependent serine protein kinase (CASK) were reported in autism-spectrum disorders.<sup>39</sup>

### Variants with a high CADD score cosegregating in families with Tourette syndrome

Among the variants with a high CADD score (> 30), 2 in aldehyde dehydrogenase genes and 1 in dihydrolipoamide dehydrogenase were present in 3 separate families. The aldehyde dehydrogenase (ALDH) superfamily plays an important role in pathways associated with development and detoxification; ALDH2 is crucial in the oxidation of

**Table 3: Rare and ultra-rare variants with the CADD score > 30 segregating with the disease in any single family**

Position and change	Gene	Type of variant	CADD	MAF, %	ACMG classification (ACMG criteria met)	Family, P(obs)*
chr7:106112481 G/A	<i>SYPL1</i>	Stop gain	41.0	0	Likely pathogenic (PP3, PM2)	Y, 0.00278
chr9:38396880 C/T	<i>ALDH1B1</i> †	Stop gain	40.0	0.203	VUS with pathogenic evidence (PM2, PP3)	T, 0.02225
chr18:12702512 G/A	<i>CEP76 (PSMG2)</i>	Stop gain	40.0	0	VUS with pathogenic evidence (PM2, PP3)	X, 0.01113
chr12:2855156 C/T	<i>TEX52 (ITFG2)</i>	Stop gain	37.0	0.166	VUS with pathogenic evidence (PM2, PP3)	H, 0.01113
chr11:113697354 C/T	<i>TMPPRSS5</i>	Stop gain	36.0	0.031	Likely pathogenic (PVS1, PM2)	U, 0.00556
chr1:117120939 G/T	<i>TRIM45</i>	Stop gain	36.0	0.182	Benign (BS1, BS2, PVS1)	J, 0.02225
chr3:149073278 C/A	<i>HLTF</i>	Stop gain	36.0	0.001	VUS with pathogenic evidence (PM2, PP3)	I, 0.002781
chr4:81459515 C/T	<i>RASGEF1B</i>	Splicing	35.0	0.030	Pathogenic (PVS1, PM2)	J, 0.02225
chr14:96264526 C/T	<i>BDKRB1</i>	Stop gain	35.0	0.250	Benign (BS1, BS2, BP4)	I, 0.002781
chr3:53855341 G/A	<i>IL17RB</i>	Missense	35.0	0.726	Benign (BS1, BS2, BP6)	A, 0.00556
chr7:2569000 G/A	<i>IQCE</i>	Splicing	34.0	0.024	Likely pathogenic (PVS1, PM2)	S, 0.01113
chr10:94687805 T/A	<i>CYP2C18</i>	Stop gain	34.0	0.316	Benign (BA1, BP4)	J, 0.02225
chr4:88417579 G/T	<i>HERC6</i>	Missense	34.0	0.396	Benign (BS1, BS2, BP1, PP3)	J, 0.02225
chr17:69026976 C/T	<i>ABCA9</i>	Missense	34.0	0	VUS with pathogenic evidence (PM2, PP3)	U, 0.00556
chr16:2317763 T/A	<i>ABCA3</i>	Missense	33.0	0.501	Benign (BS1, BS2, PP3, PP5)	T, 0.02225
chr3:132601106 C/T	<i>ACKR4</i> ( <i>ACAD11, NPHP3-ACAD11</i> )	Stop gain	33.0	0.003	Uncertain significance (PM2, BP4)	I, 0.002781
chr12:100396328 A/C	<i>SLC17A8</i> ‡	Splicing	33.0	0	Pathogenic (PVS1, PM2)	T, 0.02225
chr12:101319592 G/A	<i>UTP20</i>	Missense	32.0	0	Uncertain significance (PM2, PP3, BP1)	G, 0.01113
chr12:1885984 G/A	<i>CACNA2D4</i>	Missense	32.0	0	VUS with pathogenic evidence (PM2, PP3)	H, 0.01113
chr3:49004877 C/T	<i>P4HTM</i>	Missense	32.0	0.013	Uncertain significance (PM2, BP4)	T, 0.02225
chr17:10631666 T/C	<i>MYH3</i>	Missense	32.0	0.025	VUS with pathogenic evidence (PM2, PP3)	J, 0.02225
chr15:101324994 G/A	<i>PCSK6</i>	Missense	32.0	0.059	Uncertain significance (PM2, BP1)	J, 0.02225
chr16:4883689 G/A	<i>PPL</i>	Missense	32.0	0.156	Uncertain significance (PM2, PP3, BP1)	J, 0.02225
chr17:28385146 T/C	<i>SARM1</i>	Missense	32.0	0.291	VUS with pathogenic evidence (PM2, PP3)	G, 0.01113
chr16:89727323 C/T	<i>ZNF276</i>	Missense	32.0	0.969	Benign (BS1, BS2, BP6)	R, 0.01113
chr12:101677317 G/A	<i>MYBPC1</i>	Missense	32.0	0.001	VUS with pathogenic evidence (PM2, PP3)	U, 0.00556
chr12:111803962 G/A	<i>ALDH2</i> †	Missense	32.0	0.003	Benign (BA1, BP6, BS4, PP3)	U, 0.00556
chr12:94279637 G/A	<i>PLXNC1 (CEP83)</i>	Missense	32.0	0	Uncertain significance (PM2, PP3, BP1)	G, 0.01113
chr12:57244143 G/A	<i>STAC3</i>	Missense	31.0	0.024	Uncertain significance (PM2)	G, 0.01113
chr6:39861001 A/G	<i>DAAM2</i>	Missense	31.0	0.028	Likely pathogenic (PM2, PM1, PP3)	H, 0.01113
chr8:99999384 G/A	<i>RGS22</i>	Missense	31.0	0.046	Benign (BA1, BP6, PP3)	S, 0.01113
chr21:46114063 G/A	<i>COL6A2</i>	Missense	31.0	0.090	VUS with pathogenic evidence (PM1, PM2)	H, 0.01113
chr2:108499560 A/G	<i>GCC2</i>	Missense	31.0	0.876	Benign (BA1, BP1, PM2, PP3)	H, 0.01113
chr3:49652822 G/C	<i>BSN</i> ‡	Missense	31.0	0	VUS with pathogenic evidence (PM2, PP3)	D, 0.00139
chr16:80549649 C/G	<i>DYNLRB2</i> ( <i>AC105411.1AC108097.1</i> )	Missense	31.0	0	VUS with pathogenic evidence (PM2, PP3)	I, 0.002781
chr7:107916945 T/C	<i>DLDT</i> †	Missense	30.0	0	VUS with pathogenic evidence (PM2, PP3)	J, 0.02225
chr9:37541696 G/A	<i>FBXO10 (AL513165.2)</i>	Missense	30.0	0	VUS with pathogenic evidence (PM2, PP3)	T, 0.02225

ACMG = American College of Medical Genetics and Genomics; CADD = combined annotation dependent depletion; MAF = minor allele frequency; VUS = variant of unknown significance.

\*P(obs) = probability of cosegregation of a rare variant with a CADD > 30 in a given family, Bonferroni corrected to 17 families: P(obs) = P(seg) × (P(var) × 17). P(var) = 1 – Π<sup>n</sup>; (1 – MAF)<sup>n</sup> n = 67 – overall number of variants with CADD > 30.

†Genes encoding dehydrogenases influencing oxidoreductase activity.

‡Genes involved in sensory processing of sound by hair cells of the cochlea.

aldehydes in the brain<sup>40</sup> and is especially important in removing catecholaminergic metabolites (DOPAL and DOPEGAL) and the principal product of lipid peroxidation 4-HNE. Dihydrolipoamide dehydrogenase encoded by *DLD* catalyzes oxidative regeneration of lipoic acid cofactor. *ALDH2* has been linked to neurodegenerative disorders (Alzheimer disease and Parkinson disease), but not with neurodevelopmental ones. However, a possible link between *ADH/ALDH* and drugs used in the management of akinetic and dyskinetic (resembling tics) movement disorders has been described.<sup>41</sup> In addition, *ALDH2* activation is associated with significant attenuation of depressive and anxiety-like behaviours in prenatally stressed rats, probably via modulation of various processes associated with inflammation, oxidative stress and apoptosis.<sup>42</sup> In turn, pathogenic variants in *DLD* were reported in progressive neurologic deterioration.<sup>43</sup> Proteomic analysis showed both *ALDH2* and *DLD* as protein candidates that might be associated with susceptibility to stress-induced depression or anxiety and stress resilience.<sup>44</sup> It could be hypothesized that rare variants in *ALDH2* and *DLD* could hamper the aforementioned mechanisms and cause neurologic deterioration due to the sensitivity of the central nervous system to defects in oxidative metabolism.

### Enriched processes

Most of the genes pinpointed in the cosegregation analysis in the families with Tourette syndrome were involved in neurologic processes, as indicated by the enrichment analysis showing that 12 of the 25 most significantly enriched processes (Figure 1) were specifically related to the nervous system (including neuromuscular processes). Many of these processes

and compartments are interrelated (Appendix 8, available at [www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content](http://www.jpn.ca/lookup/doi/10.1503/jpn.220206/tab-related-content)). Moreover, many genes identified in our study that were associated with significantly enriched processes have previously been reported to be associated with other neuropsychiatric disorders (Table 6).

The top 2 enriched pathways with ultra-rare variants, cell junction assembly and cell–cell adhesion, although not neural-specific, relate closely with numerous genes already linked with Tourette syndrome (*NLGN4*, *CDH23*, *CNTNAP2/CASPR2*, and *DPP6*). Rare copy number variants (CNVs) in *NRXN1* — a gene coding for the presynaptic cell adhesion molecule neuroligin 1, which together with neuroligins,<sup>45</sup> including neuroligin 4 (encoded by *NLGN4*), is involved in glutamatergic and GABAergic neurotransmission and synaptogenesis<sup>46</sup> — have already been associated with Tourette syndrome.<sup>47,48</sup> Moreover, *NRXN1* deletions have been implicated in other neurodevelopmental and psychiatric disorders, including ASD and schizophrenia<sup>49,50</sup>

Variants in *CNTN6* (overlapping *CNTN4*) and polymorphisms in *TENM2* (rs147208935) and *BTBD9* (rs9296249) were already associated with risk for Tourette syndrome.<sup>7,48,51</sup> De novo variants in *FNI* have been described in cases of Tourette syndrome.<sup>16,52</sup> The genomic region encompassing *PKP4* has been identified in GWAS analyses of Tourette syndrome and ADHD.<sup>14,53</sup> *ROBO2* has been identified as strongly associated with Tourette syndrome,<sup>14</sup> and *CDH23* was reported to be a candidate gene in a multiplex Tourette syndrome family.<sup>15</sup> Our results are therefore in agreement with earlier analyses showing that cell adhesion and synaptic signalling is significantly related to Tourette syndrome etiology.<sup>14</sup>

**Table 4: Ultra-rare and rare variants identified in 2 families each, sorted by MAF**

Position change	Gene*	Location/ type of variant	MAF, %	CADD	ACMG classification (ACMG criteria met)	Fam, P(obs)†
19:53803962 G/A	<i>NLRP12</i>	Exonic/missense	0.01	23.3	Likely benign (BS2)	G, I, 0.000211
5:76482747 A/C	<i>IQGAP2</i>	Intronic	0.08	15.9	Benign (BS1, BS2)	E, X, 0.000842
4:109086168 T/A	<i>COL25A1</i>	Intronic	0.10	19.0	Likely benign (BS1)	H, J, 0.004208
1:223300619 G/A	<i>SUSD4</i>	Intronic	0.20	16.4	VUS with benign evidence (BP4)	J, X, 0.008415
12:120368947 C/A	<i>MSI1</i>	Intronic	0.33	16.4	Benign (BS1, BS2, BP4)	J, R, 0.013885
1:31727178 G/A	<i>ADGRB2</i>	3' UTR	0.60	13.2	Benign (BS1, BS2)	B, G, 0.003151
12:10723010 G/T	<i>YBX3</i>	Exonic/missense	0.72	22.2	Likely benign (BS1, BP4)	J, T, 0.030294
11:111020375 C/A	<i>AP003973.4</i>	Intergenic	0.73	14.0	Benign (BS1, BS2, BP4, BP7)	E, S, 0.007679
6:101375096 A/G	<i>GRIK2</i>	Intronic	0.73	20.5	Benign (BS1, BS2)	D, G, 0.003833
12:97272797 T/A	<i>LINC02409</i>	Intronic/splice	0.75	14.7	Benign (BA1, BP4, BP7)	G, X, 0.031556
7:713068 C/T	<i>PRKAR1B</i>	5' UTR	0.75	10.5	Benign (BS1, BS2, BP4)	A, J, 0.015778
1:121379326 G/A	<i>SRGAP2C</i>	Intronic	0.91	15.5	Benign (BA1, BP4)	J, X, 0.038288
10:74875378 T/C	<i>KAT6B</i>	Intronic	0.95	11.0	Benign (BS1, BS2, BP4)	J, X, 0.039972
6:17649263 G/A	<i>NUP153</i>	Exonic/missense	0.99	27.7	Benign (BS1, BS2, BP1, PP3)	C, J, 0.002599

ACMG = American College of Medical Genetics and Genomics; CADD = combined annotation dependent depletion; MAF = minor allele frequency; UTR = untranslated region; VUS = variant of unknown significance.

\*Several genes had more than 1 ultra-rare variant with putatively severe consequences. Seven genes (*BSN*, *FLG-AS1*, *FNI*, *NLRP12*, *PTPN14*, *RBL2*, and *SNX19*) had 2 missense variants each, with a third variant in *BSN* located in 3' UTR. Three out of 4 variants located in *TTN* were missense. Eight genes had 1 missense and 1 3' UTR or 5' UTR variant: *AC046130.1*, *GAPVD1*, *IGSF3*, *NOL4L*, *ROBO2*, *SEMA4A*, *TMEM63B*, *TPRG1*.

†P(obs) – probability of cosegregation of a particular variant in a given 2 families, Bonferroni corrected to 136 possible pairs of families: P(obs) = P(seg\_fam1) × P(anyvar) × P(seg\_fam2) × P(var) × 136. P(any var) = 0.99 – probability of a variant in family 1 with CADD > 10 and MAF < 0.01. P(var) = MAF – the probability of observing the particular variant in the second family P(var).

**Table 5: Top significantly enriched categories (LogP value < -4) with variants associated with Tourette syndrome**

Term ID	Description	LogP	Genes in which variants were found	MAF below, %*
GO:0098609	Cell-cell adhesion	-13.56	<i>CDH23</i> †, <i>NLGN1</i> †, <i>PTPRF</i> †, <i>ROBO2</i> †, <i>ADGRL3</i> †, <i>CD44</i> , <i>CLSTN2</i> , <i>CNTN4</i> †, <i>CTNNA3</i> , <i>FN1</i> †, <i>GNAS</i> , <i>LPP</i> , <i>NRXN3</i> , <i>PKHD1</i> †, <i>PKP4</i> , <i>PTPRD</i> , <i>SDK1</i> †, <i>TENM2</i> †	5.0
GO:0000902	Cell morphogenesis	-7.93	<i>CDH23</i> †, <i>EPHB1</i> †, <i>NLGN1</i> †, <i>ROBO2</i> †, <i>SEMA3A</i> †, <i>AUTS2</i> , <i>CNTN4</i> †, <i>FN1</i> †, <i>NCAM1</i> †, <i>NRG1</i> , <i>NRXN3</i> , <i>PKHD1</i> †, <i>PLXNC1</i> †, <i>USH2A</i> †	5.0
GO:0034329	Cell junction assembly	-7.75	<i>ERBB4</i> , <i>BSN</i> †, <i>PKP4</i> †, <i>PTPRD</i> †, <i>ADGRL3</i> †, <i>NRXN1</i> , <i>CDH22</i> , <i>EPB41L3</i> †, <i>FN1</i> †, <i>CLDN14</i>	0.1
GO:0034330	Cell junction organization	-7.70	<i>EPHB1</i> †, <i>NLGN1</i> †, <i>ADGRL3</i> †, <i>DLC1</i> †, <i>ERBB4</i> , <i>ERC2</i> , <i>FN1</i> †, <i>NRG1</i> , <i>PKHD1</i> †, <i>PKP4</i> , <i>PTPRD</i> , <i>SDK1</i> †	5.0
GO:0098609	Cell-cell adhesion	-7.57	<i>TENM2</i> , <i>CDH22</i> , <i>CDH23</i> †, <i>CNTN4</i> †, <i>ROBO2</i> †, <i>PTPRD</i> †, <i>ADGRL3</i> †, <i>CLDN14</i> , <i>PIK3CG</i> , <i>NRXN1</i> , <i>NRXN3</i> , <i>PKP4</i> †, <i>FN1</i> †	0.1
R-HSA-9662360	Sensory processing of sound by inner hair cells of the cochlea	-6.94	<i>KCNMA1</i> , <i>EPB41L3</i> †, <i>CDH23</i> †, <i>CACNA2D2</i> †, <i>BSN</i> †, <i>SLC17A8</i> †	0.1
GO:0007420	Brain development	-6.77	<i>SATB2</i> , <i>ERBB4</i> , <i>MEIS2</i> , <i>RARB</i> , <i>ROBO2</i> †, <i>ALK</i> †, <i>PRKG1</i> †, <i>CDH22</i> , <i>NDRG2</i> †, <i>ADGRL3</i> †, <i>EPHB1</i> , <i>ZNF148</i> , <i>SLC17A8</i> †, <i>CNTN4</i> †	0.1
GO:0016358	Dendrite	-6.56	<i>CACNA1C</i> †, <i>EPHB1</i> †, <i>NLGN1</i> †, <i>SEMA3A</i> †, <i>ANKS1B</i> , <i>CLSTN2</i> , <i>GIGYF2</i> †, <i>KCNIP4</i> , <i>MAGI2</i> , <i>SLC4A10</i> , <i>SLC8A1</i> , <i>TENM2</i> †	5.0
GO:0007420	Brain development	-6.50	<i>EPHB1</i> †, <i>ROBO2</i> †, <i>SEMA3A</i> †, <i>MACROD2</i> †, <i>ADGRL2</i> , <i>ADGRL3</i> †, <i>CNTN4</i> †, <i>DLC1</i> †, <i>ERBB4</i> , <i>NRG1</i> , <i>PLCB1</i> , <i>RARB</i> , <i>SLC4A10</i> , <i>SLC8A1</i>	5.0
GO:0008038	Neuron recognition	-6.44	<i>OPCML</i> †, <i>ROBO2</i> †, <i>NTM</i> †, <i>CNTN4</i> †	5.0
GO:0030424	Axon	-6.41	<i>EPHB1</i> †, <i>ROBO2</i> †, <i>SEMA3A</i> †, <i>ADGRL3</i> †, <i>AUTS2</i> , <i>CNTN4</i> †, <i>ERC2</i> , <i>NRG1</i> , <i>SLC4A10</i> , <i>SLC8A1</i> , <i>TENM2</i> †, <i>USH2A</i> †	5.0
GO:0000902	Cell morphogenesis	-6.06	<i>CDH22</i> , <i>CDH23</i> †, <i>EPHB1</i> , <i>KLF7</i> , <i>NLGN1</i> , <i>NRXN3</i> , <i>NTN4</i> , <i>PLXNC2</i> †, <i>ROBO2</i> †, <i>SEMA3A</i> †	1.0
GO:0098609	Cell-cell adhesion	-5.90	<i>CDH22</i> , <i>CDH23</i> †, <i>GNAS</i> †, <i>NLGN1</i> , <i>NRXN3</i> , <i>PKP4</i> , <i>PTPRD</i> , <i>ROBO2</i> †, <i>TENM2</i> †	1.0
GO:0051963	Regulation of synapse assembly	-5.90	<i>NRXN1</i> , <i>EPHB1</i> , <i>SEMA4A</i> †, <i>ROBO2</i> †, <i>PTPRD</i> †, <i>COLQ</i>	0.1
GO:0007626	Locomotor behaviour	-5.85	<i>BTBD9</i> †, <i>CDH23</i> †, <i>FGF12</i> , <i>GIGYF2</i> †, <i>NAV2</i> †, <i>NRG1</i> , <i>SLC4A10</i>	5.0
GO:0005509	Calcium ion binding	-5.85	<i>CDH23</i> †, <i>EYS</i> †, <i>ADGRL3</i> †, <i>CLSTN2</i> , <i>FSTL5</i> , <i>KCNIP4</i> , <i>LTBP1</i> †, <i>PLCB1</i> , <i>SLC8A1</i> , <i>STAB2</i> †, <i>TENM2</i> †, <i>TLL2</i> †	5.0
GO:0040007	Growth	-5.75	<i>EYS</i> †, <i>SEMA3A</i> †, <i>AUTS2</i> , <i>ERBB4</i> , <i>GIGYF2</i> †, <i>GNAS</i> , <i>MAGI2</i> , <i>RARB</i> , <i>SLC4A10</i>	5.0
GO:0007167	Enzyme-linked receptor protein signalling pathway	-5.72	<i>EPHB1</i> †, <i>PTPRF</i> †, <i>ANKS1B</i> , <i>ERBB4</i> , <i>FGF12</i> , <i>GIGYF2</i> †, <i>LTBP1</i> †, <i>MAGI2</i> , <i>NRG1</i> , <i>PLCB1</i> , <i>PTPRD</i>	5.0
GO:0007420	Brain development	-5.65	<i>ADGRL2</i> , <i>ALK</i> †, <i>CDH22</i> , <i>EPHB1</i> , <i>ERBB4</i> , <i>MACROD2</i> †, <i>PRKG1</i> , <i>RARB</i> †, <i>ROBO2</i> †, <i>SEMA3A</i> †, <i>SRGAP2C</i>	1.0
GO:0051962	Positive regulation of nervous system development	-5.47	<i>FN1</i> †, <i>SEMA4A</i> †, <i>MAP3K13</i> , <i>ROBO2</i> †, <i>PLXNC1</i> †, <i>PTPRD</i> †, <i>EPHB1</i> , <i>NRXN1</i>	0.1
GO:0120035	Regulation of plasma membrane bounded cell projection organization	-5.36	<i>ALK</i> †, <i>NLGN1</i> , <i>PLCE1</i> †, <i>PLXNC1</i> †, <i>PTPRD</i> , <i>ROBO2</i> †, <i>SEMA3A</i> †, <i>SRGAP2C</i> , <i>TENM2</i> †	1.0
GO:0098858	Actin-based cell projection	-5.34	<i>CDH23</i> †, <i>EPHB1</i> †, <i>NLGN1</i> †, <i>CD44</i> , <i>IQGAP2</i> †, <i>TENM2</i> †, <i>USH2A</i> †	5.0
GO:0034330	Cell junction organization	-5.27	<i>CDH22</i> , <i>EPHB1</i> , <i>ERBB4</i> , <i>ERC2</i> , <i>NLGN1</i> , <i>PKP4</i> , <i>PTPRD</i> , <i>SRGAP2C</i>	1.0
GO:0050804	Modulation of chemical synaptic transmission	-5.21	<i>BTBD9</i> †, <i>EPHB1</i> †, <i>NLGN1</i> †, <i>CLSTN2</i> , <i>CNTN4</i> †, <i>ERC2</i> , <i>PLCB1</i> , <i>PTPRD</i> , <i>SLC4A10</i>	5.0
GO:0016324	Apical plasma membrane	-4.89	<i>KCNMA1</i> †, <i>CD44</i> , <i>FN1</i> †, <i>GNAS</i> , <i>PARD3B</i> , <i>PKHD1</i> †, <i>SLC4A10</i> , <i>USH2A</i> †	5.0
GO:0005539	Glycosaminoglycan binding	-4.89	<i>COL25A1</i> †, <i>CD44</i> , <i>FN1</i> †, <i>HK1</i> †, <i>NAV2</i> †, <i>STAB2</i> †	5.0
GO:0050885	Neuromuscular process controlling balance	-4.67	<i>NRXN1</i> , <i>RBFOX1</i> , <i>CAMTA1</i> , <i>CDH23</i> †	0.1
GO:0016358	Dendrite development	-4.16	<i>EPHB1</i> , <i>KLF7</i> , <i>PRKG1</i> , <i>SEMA3A</i> †	1.0
GO:0045932	Negative regulation of muscle contraction	-4.11	<i>PRKG1</i> †, <i>KCNMA1</i> , <i>PIK3CDG</i>	0.1
WP2118	Arrhythmogenic right ventricular cardiomyopathy	-4.10	<i>CACNA1C</i> †, <i>CACNA2D3</i> †, <i>CTNNA3</i> , <i>SLC8A1</i>	5.0
GO:0098858	Actin-based cell projection	-4.07	<i>CDH23</i> †, <i>EPHB1</i> , <i>NLGN1</i> , <i>TENM</i> †, <i>WVWX</i>	1.0
GO:0043408	Regulation of MAPK cascade	-4.01	<i>EPHB1</i> , <i>ERBB4</i> , <i>ALK</i> †, <i>MAP3K13</i> , <i>PIK3CG</i> , <i>FN1</i> †, <i>NDRG2</i> †, <i>NLRP12</i> †, <i>PLCE1</i> †, <i>MAPKB1</i> †	0.1

MAF = minor allele frequency.

\*Analyses were performed separately on gene sets with ultra-rare variants (MAF below 0.1%) occurring in at least 2 families, rare variants (MAF below 1%) occurring in at least 3 families, and uncommon variants (MAF below 5%) occurring in at least 4 families.

†Genes with variants found in all patients in at least  $n + 1$  families, where  $n$  is the threshold value selected for a given MAF (i.e., 2 for 0.1%, 3 for 1% and 4 for 5%) threshold.

‡Genes with missense and/or splicing variants.





*SLC17A8* encodes vesicular glutamate cotransporter vGlut3, which accumulates glutamate in the synaptic vesicles of the sensory IHCs before releasing it into receptors of auditory nerve terminals.<sup>54</sup> It is also present in striatal and serotonin neurons and is implicated in amphetamine-induced stereotypies in mice.<sup>55</sup> *BSN* encodes bassoon presynaptic cytomatrix protein, which in IHCs is anchoring presynaptic ribbons essential in synchronous auditory signalling and, consequently, in normal hearing.<sup>56</sup>

The auditory system comprises descending efferent neural projections joining the medial and the lateral nuclei of the superior olivary complex (MOC and LOC, respectively) and the inner ear complex structure. The unmyelinated axons of the LOC project to the dendrites of auditory nerve fibres near the IHC afferent synapses. The role of the LOC system is still unknown; however, the presence of several neurotransmitters and modulators in its terminals (dopamine, acetylcholine, and GABA) suggests that it has complex functions. Moreover, it is known from animal models that, during a critical period of postnatal development, IHCs transiently receive cholinergic innervation, driving neurons in the auditory pathway to respond. This process is important for the normal maturation of synapses and circuits of the entire auditory pathway.<sup>57</sup>

The role of MOC is mainly to inhibit cochlear responses by decreasing the gain of the OHC amplifier. Several lines of evidence suggest that the MOC system plays an important role in the protection from trauma produced by overly loud sounds. The activity of the MOC toward the IHC input is also inhibitory during this developmental period and would control the excitability of the hair cells.

An important element of sound processing is also the prepulse inhibition (PPI), which is a measure of sensorimotor gating in response to a stimulus (pulse) being diminished when the stimulus is preceded by a smaller stimulus (prepulse).<sup>58,59</sup> Sensorimotor gating is a normal protective

mechanism in the brain that functions to gate or filter irrelevant sensory stimulation. Deficits in sensorimotor gating may result in stimulus overload and misinterpretation of sensory information. The hypothesis that the premonitory urges present in Tourette syndrome cannot be properly filtered and thus can induce abnormal, semivoluntary (so-called involuntary) movements like tics is in line with deficient PPI found in people with Tourette syndrome.<sup>58</sup> Moreover, some tics can be triggered by diverse stimuli — auditory, visual, tactile or mental.<sup>60</sup> A hypersensitization to various stimuli is a common clinical symptom in people with ASD. We also found this feature in 44.8% (74/165) of patients with Tourette syndrome, and almost one-third (22/74) had an abnormal response to sound.<sup>61</sup>

We identified 8 families in which variants of genes associated with the processing of sound by cells of the cochlea were found. There is some degree of phenotype–genotype interplay concerning overreactivity to different sensory stimuli. Six of 7 probands were oversensitized to touch (families D, G, H, I, J and T), and 1 proband additionally showed hypersensitization to sound (family H). There were no clinical data regarding oversensitivity to stimuli in 1 proband (family R) and remaining members of all analyzed families, including affected and healthy people.

Patients with ASD, OCD, or Tourette syndrome and misophonia, selective sound sensitivity syndrome or auditory hypersensitivity have been described previously.<sup>62–64</sup> It has been proposed that these conditions may share pathophysiological development and etiological characteristics. Misophonia and Tourette syndrome could share abnormal activity of the limbic system, primary auditory cortex, and/or the autonomic nervous system.<sup>65,66</sup> However, to our knowledge, no genetic basis for such a link has been proposed so far. We hypothesize that abnormal processing of acoustic stimuli, associated with the processes in the cochlear hair cells and superior olivary complex, and/or prepulse inhibition in our enrichment analyses may link the above-mentioned mechanisms and Tourette syndrome.

**Table 7: Variants in genes involved in the sensory processing of sound by inner hair cells of the cochlea\***

Gene	Family D	Family E	Family G	Family H	Family I	Family J†	Family R	Family T
<i>BSN</i>	ur, 31, m	–	–	–	–	ur, 15, UTR ur, 11, m	–	–
<i>CACNA2D2</i>	–	–	–	–	–	ur, 19, i ur, 15, i	–	ur, 22, m
<i>CDH23</i>	ur, 22, i r, 17, i r, 17, i ur, 14, i	r, 20, i uc, 18, i uc, 16, i r, 13, i	–	uc, 21, i uc, 18, i	r, 23, m	ur, 21, i ur, 20, m r, 16, i ur, 15, i r, 14, i r, 11, i r, 10, i	–	–
<i>EPB41L3</i>	–	–	–	–	–	ur, 13, s	ur, 24, m	–
<i>KCNMA1</i>	–	uc, 12, i uc, 11, i	ur, 12, i	r, 12, i uc, 13, i	uc, 17, i uc, 16, i ur, 14, i	uc, 17, i uc, 13, s	–	–
<i>SLC17A8</i>	–	–	ur, 10, i	–	–	–	–	ur, 33, sp

CADD = combined annotation dependent depletion; i = intronic; m = missense; MAF = minor allele frequency; r = rare (MAF 0.1%–1%); s = synonymous; sp = splicing; uc = uncommon (MAF 1%–5%); ur = ultra-rare (MAF < 0.1%); UTR = untranslated region.

\*Numbers indicate CADD score.

†The high number of variants found in Family J was partially due to the close kinship among all 3 affected family members (father and 2 siblings).

## Limitations

Although we have analyzed variants co-segregating in 17 multiplex families with at least 3 patients each with tic disorders, the number of participants, including unaffected noncarriers, can be considered a limitation of the study.

Most of the genomic variants identified here are located in regions not coding for protein products; therefore, it is difficult to assess their impact on the processes in which the gene products participate. The putative causal link with the disease was based mainly on the cosegregation with the clinical phenotype, and additionally on the rarity of variants. The assessment of likely pathogenicity of the variants was based solely on in silico predictions. Additional functional studies, especially in the case of intronic or intragenic variants, will be needed, as different prediction tools could give substantially divergent rankings of variant severity.

We identified enriched processes and functions using the Gene Ontology (GO) database, although the biological processes are interdependent and particular genes (proteins or noncoding RNAs) may have roles in multiple processes. Moreover, the GO annotation is supported by various sources, and the majority of genes are assigned to terms based on computational predictions.

The aim of our research was neither functional nor translational, and further work to understand the genetic contribution to the clinical phenotype of Tourette syndrome and underlying dysfunctions at the molecular level is needed.

## Conclusion

We identified putatively pathogenic genomic variants and molecular processes related to the etiology of Tourette syndrome in a group of Polish families. Three of these variants could influence oxidoreductase activity in the brain. Using enrichment analysis of the variant-bearing genes, we found evidence for a likely input of sensory processing of sound in the cochlea, in support of earlier reports of a hypersensitivity of a substantial fraction of patients with Tourette syndrome to diverse stimuli, including sonory ones. Other over-represented groups of genes with variants associated with Tourette syndrome were related to cell-cell adhesion, cell junction assembly and organization, synapse assembly and synaptic signalling, confirming earlier findings regarding the genetic basis of Tourette syndrome. Moreover, even if none of the identified variants is causal individually, our results support the concept of an oligogenic basis of Tourette syndrome and indicate that a burden of a large variety of rare and uncommon variants in genes implicated in various neurodevelopmental processes may be causally related to Tourette syndrome pathology. Further analyses using substantially larger groups of families, as well as individuals with sporadic Tourette syndrome and tic disorders, should be performed to confirm and expand our results.

**Affiliations:** From the Department of Neurogenetics and Functional Genomics (Fichna, Zekanowski), Mossakowski Medical Research Institute, Polish Academy of Sciences, Warsaw, Poland; the Department of Biological Sciences (Fichna), Purdue University, West Lafayette, Ind., USA; the Laboratory of Pharmacogenomics (Borczyk,

Piechota, Korostynski), Department of Molecular Neuropharmacology, Maj Institute of Pharmacology, Polish Academy of Sciences, Krakow, Poland; and the Department of Neurology (Janik), Medical University of Warsaw, Warsaw, Poland

**Funding:** This work was funded by the National Science Center, Poland (NCN) project UMO-2016/23/B/NZ2/03030, and supported by PLGrid Infrastructure. J. Fichna is supported by the Polish National Agency for Academic Exchange, Bekker programme PPN/BEK/2019/1/00452/U/00001.

**Competing interests:** None declared.

**Contributors:** C. Zekanowski designed the study. P. Janik acquired the data, which J. Fichna, M. Borczyk, M. Piechota and M. Korostynski analyzed. J. Fichna and M. Borczyk wrote the article, which all authors reviewed. All authors approved the final version to be published, agreed to be accountable for all aspects of the work and can certify that no other individuals not listed as authors have made substantial contributions to the paper.

**Content licence:** This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY-NC-ND 4.0) licence, which permits use, distribution and reproduction in any medium, provided that the original publication is properly cited, the use is noncommercial (i.e., research or educational use), and no modifications or adaptations are made. See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

## References

- Selvini C, Cavanna S, Cavanna A. Gilles de la Tourette syndrome. In: *Chromatin Signaling and Neurological Disorders*. Cambridge (MA): Academic Press; 2019:331-45. doi: 10.1016/B978-0-12-813796-3.00015-8.
- Robertson MM, Eapen V, Singer HS, et al. Gilles de la Tourette syndrome. *Nat Rev Dis Primers* 2017;3:16097.
- Robertson MM. Tourette syndrome, associated conditions and the complexities of treatment. *Brain* 2000;123:425-62.
- Levine JLS, Szejko N, Bloch MH. Meta-analysis: adulthood prevalence of Tourette syndrome. *Prog Neuropsychopharmacol Biol Psychiatry* 2019;95:109675.
- Paschou P. The genetic basis of Gilles de la Tourette syndrome. *Neurosci Biobehav Rev* 2013;37:1026-39.
- Tagwerker Gloor F, Walitza S. Tic disorders and Tourette syndrome: current concepts of etiology and treatment in children and adolescents. *Neuropediatrics* 2016;47:84-96.
- Yu D, Sul JH, Tsetsos F, et al. Interrogating the genetic determinants of Tourette's syndrome and other tic disorders through genome-wide association studies. *Am J Psychiatry* 2019;176:217-27.
- Davis LK, Yu D, Keenan CL, et al. Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. *PLoS Genet* 2013;9:e1003864.
- Mataix-Cols D, Isomura K, Pérez-Vigil A, et al. Familial risks of Tourette syndrome and chronic tic disorders. A population-based cohort study. *JAMA Psychiatry* 2015;72:787-93.
- Pagliaroli L, Veto B, Arányi T, et al. From genetics to epigenetics: new perspectives in Tourette syndrome research. *Front Neurosci* 2016;10:277.
- Qi Y, Zheng Y, Li Z, et al. Genetic studies of tic disorders and Tourette syndrome. *Methods Mol Biol* 2019;2011:547-71.
- Georgitsi M, Willsey AJ, Mathews CA, et al. The genetic etiology of Tourette syndrome: large-scale collaborative efforts on the precipice of discovery. *Front Neurosci* 2016;10:351.
- Köhler S, Gargano M, Matentzoglou N, et al. The human phenotype ontology in 2021. *Nucleic Acids Res* 2021;49(D1):D1207-17.
- Tsetsos F, Yu D, Sul JH, et al. Synaptic processes and immune-related pathways implicated in Tourette syndrome. *Transl Psychiatry* 2021;11:56.
- Cao X, Zhang Y, Abdulkadir M, et al. Whole-exome sequencing identifies genes associated with Tourette's disorder in multiplex families. *Mol Psychiatry* 2021;26:6937-51.
- Willsey AJ, Fernandez TV, Yu D, et al. De novo coding variants are strongly associated with Tourette disorder. *Neuron* 2017;94:486-499.e9.

17. Abdulkadir M, Mathews CA, Scharf JM, et al. Polygenic risk scores derived from a Tourette syndrome genome-wide association study predict presence of tics in the Avon Longitudinal Study of Parents and Children Cohort. *Biol Psychiatry* 2019;85:298-304.
18. Lee PH, Anttila V, Won H, et al. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* 2019;179:1469-82.e11.
19. Freeman RD, Fast DK, Burd L, et al. An international perspective on Tourette syndrome: selected findings from 3,500 individuals in 22 countries. *Dev Med Child Neurol* 2000;42:436-47.
20. Jarczak J, Grochowalski Ł, Marciniak B, et al. Mitochondrial DNA variability of the Polish population. *Eur J Hum Genet* 2019;27:1304-14.
21. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;16:1215.
22. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020;581:434-43.
23. Rentzsch P, Schubach M, Shendure J, et al. CADD-Splice: improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med* 2021;13:31.
24. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol* 2016;17:122.
25. Frankish A, Diekhans M, Ferreira AM, et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 2019;47:D766-73.
26. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019;10:1523.
27. Takata A. Estimating contribution of rare non-coding variants to neuropsychiatric disorders. *Psychiatry Clin Neurosci* 2019;73:2-10.
28. Pandey GN, Zhang H, Sharma A, et al. Innate immunity receptors in depression and suicide: upregulated NOD-like receptors containing pyrin (NLRPs) and hyperactive inflammasomes in the postmortem brains of people who were depressed and died by suicide. *J Psychiatry Neurosci* 2021;46:E538-47.
29. Edwards AC, Aliev F, Wolen AR, et al. Genomic influences on alcohol problems in a population-based sample of young adults. *Addiction* 2015;110:461-70.
30. Akosile W, Voisey J, Lawford B, et al. The inflammasome NLRP12 is associated with both depression and coronary artery disease in Vietnam veterans. *Psychiatry Res* 2018;270:775-9.
31. Jéru I, Duquesnoy P, Fernandes-Alnemri T, et al. Mutations in NALP12 cause hereditary periodic fever syndromes. *Proc Natl Acad Sci U S A* 2008;105:1614-9.
32. Vilarinho-Güell C, Zimprich A, Martinelli-Boneschi F, et al. Exome sequencing in multiple sclerosis families identifies 12 candidate genes and nominates biological pathways for the genesis of disease. *PLoS Genet* 2019;15:e1008180.
33. Hashimoto T, Wakabayashi T, Watanabe A, et al. CLAC: a novel Alzheimer amyloid plaque component derived from a transmembrane precursor, CLAC-P/collagen type XXV. *EMBO J* 2002;21:1524-34.
34. Forsell C, Björk BF, Lilius L, et al. Genetic association to the amyloid plaque associated protein gene COL25A1 in Alzheimer's disease. *Neurobiol Aging* 2010;31:409-15.
35. Li D, Zhao H, Kranzler HR, et al. Association of COL25A1 with comorbid antisocial personality disorder and substance dependence. *Biol Psychiatry* 2012;71:733-40.
36. Shinwari JMA, Khan A, Awad S, et al. Recessive mutations in COL25A1 are a cause of congenital cranial dysinnervation disorder. *Am J Hum Genet* 2015;96:147-52.
37. Hedman AC, Smith JM, Sacks DB. The biology of IQGAP proteins: beyond the cytoskeleton. *EMBO Rep* 2015;16:427-46.
38. Brill S, Li S, Lyman CW, et al. The Ras GTPase-activating-protein-related human protein IQGAP2 harbors a potential actin binding domain and interacts with calmodulin and Rho family GTPases. *Mol Cell Biol* 1996;16:4869-78.
39. Huang TN, Hsueh YP. Calcium/calmodulin-dependent serine protein kinase (CASK), a protein implicated in mental retardation and autism-spectrum disorders, interacts with T-Brain-1 (TBR1) to control extinction of associative memory in male mice. *J Psychiatry Neurosci* 2017;42:37-47.
40. Deza-Ponzio R, Herrera ML, Bellini MJ, et al. Aldehyde dehydrogenase 2 in the spotlight: the link between mitochondria and neurodegeneration. *Neurotoxicology* 2018;68:19-24.
41. Messiha FS. Tourette's medications: effect on minor oxidative and reductive pathways of biogenic amines. *Neurosci Biobehav Rev* 1988;12:215-8.
42. Stachowicz A, Głombik K, Olszanecki R, et al. The impact of mitochondrial aldehyde dehydrogenase (ALDH2) activation by Ald-1 on the behavioral and biochemical disturbances in animal model of depression. *Brain Behav Immun* 2016;51:144-53.
43. Hong YS, Kerr DS, Craigen WJ, et al. Identification of two mutations in a compound heterozygous child with dihydroliipoamide dehydrogenase deficiency. *Hum Molec Genet* 1996;5:1925-1930.
44. Tang M, Huang H, Li S, et al. Hippocampal proteomic changes of susceptibility and resilience to depression or anxiety in a rat model of chronic mild stress. *Transl Psychiatry* 2019;9:260.
45. Südhof TC. Neuroligins and neuexins link synaptic function to cognitive disease. *Nature* 2008;455:903-11.
46. Pak C, Danko T, Mirabella VR, et al. Cross-platform validation of neurotransmitter release impairments in schizophrenia patient-derived NRXN1-mutant neurons. *Proc Natl Acad Sci U S A* 2021;118:e2025598118.
47. Nag A, Bochukova EG, Kremeyer B, et al. CNV analysis in Tourette syndrome implicates large genomic rearrangements in COL8A1 and NRXN1. *PLoS One* 2013;8:e59061.
48. Huang AY, Yu D, Davis LK, et al. Rare copy number variants in NRXN1 and CNTN6 increase risk for Tourette syndrome. *Neuron* 2017;94:1101-1111.e7.
49. Woodbury-Smith M, Nicolson R, Zarrei M, et al. Variable phenotype expression in a family segregating microdeletions of the NRXN1 and MBD5 autism spectrum disorder susceptibility genes. *NPJ Genom Med* 2017;2:17.
50. Levy KA, Weisz ED, Jongens TA. Loss of *neurexin-1* in *Drosophila melanogaster* results in altered energy metabolism and increased seizure susceptibility. *Hum Mol Genet* 2022;31:3422-38.
51. Janik P, Berdyński M, Safranow K, et al. The BTBD9 gene polymorphisms in Polish patients with Gilles de la Tourette syndrome. *Acta Neurobiol Exp (Warsz)* 2014;74:218-26.
52. Wang S, Mandell JD, Kumar Y, et al. De novo sequence and copy number variants are strongly associated with Tourette disorder and implicate cell polarity in pathogenesis. *Cell Rep* 2018;24:3441-54.e12.
53. Yang Z, Wu H, Lee PH, et al. Investigating shared genetic basis across Tourette syndrome and comorbid neurodevelopmental disorders along the impulsivity-compulsivity spectrum. *Biol Psychiatry* 2021;90:317-27.
54. Seal RP, Akil O, Yi E, et al. Sensorineural deafness and seizures in mice lacking vesicular glutamate transporter 3. *Neuron* 2008;57:263-75.
55. Mansouri-Guilani N, Bernard V, Vigneault E, et al. VGLUT3 gates psychomotor effects induced by amphetamine. *J Neurochem* 2019;148:779-95.
56. Khimich D, Nouvian R, Pujol R, et al. Hair cell synaptic ribbons are essential for synchronous auditory signalling. *Nature* 2005;434:889-94.
57. Goutman JD, Elgoyhen AB, Gómez-Casati ME. Cochlear hair cells: the sound-sensing machines. *FEBS Lett* 2015;589:3354-61.
58. Swerdlow NR, Sutherland AN. Using animal models to develop therapeutics for Tourette syndrome. *Pharmacol Ther* 2005;108:281-93.
59. Swerdlow NR. Update: studies of prepulse inhibition of startle, with particular relevance to the pathophysiology or treatment of Tourette Syndrome. *Neurosci Biobehav Rev* 2013;37:1150-6.
60. Janik P, Milanowski L, Szejko N. Phenomenology and clinical correlates of stimulus-bound tics in Gilles de la Tourette syndrome. *Front Neurol* 2018;9:477.
61. Szejko N, Janik P. Stimulus sensitization in patients with Gilles de la Tourette syndrome [conference poster]. In: *11th European Conference on Tourette Syndrome & Tics Disorders*. P27. ESSTS; 2018:58.
62. Visser E, Zwiers MP, Kan CC, et al. Atypical vertical sound localization and sound-onset sensitivity in people with autism spectrum disorders. *J Psychiatry Neurosci* 2013;38:398-406.
63. Neal M, Cavanna AE. Selective sound sensitivity syndrome (misophonia) in a patient with Tourette syndrome. *J Neuropsychiatry Clin Neurosci* 2013;25:E01.
64. Webber TA, Johnson PL, Storch EA. Pediatric misophonia with comorbid obsessive-compulsive spectrum disorders. *Gen Hosp Psychiatry* 2014;36:231.e1-2.
65. Jastreboff PJ, Jastreboff MM. Tinnitus retraining therapy for patients with tinnitus and decreased sound tolerance. *Otolaryngol Clin North Am* 2003;36:321-36.
66. Palumbo DB, Alsalman O, De Ridder D, et al. Misophonia and potential underlying mechanisms: a perspective. *Front Psychol* 2018;9:953.