

## Genetic variants in *DDX53* contribute to Autism Spectrum Disorder associated with the Xp22.11 locus.

Marcello Scala<sup>1,2</sup>, Clarrisa A. Bradley<sup>3,4,5</sup>, Jennifer L. Howe<sup>3,4</sup>, Brett Trost<sup>3,4</sup>, Nelson Bautista Salazar<sup>3,4</sup>, Carole Shum<sup>4</sup>, Miriam S. Reuter<sup>3,4</sup>, Jeffrey R. MacDonald<sup>3,4</sup>, Sangyoon Y. Ko<sup>5</sup>, Paul W. Frankland<sup>5,6</sup>, Leslie Granger<sup>7</sup>, George Anadiotis<sup>7</sup>, Verdiana Pullano<sup>8</sup>, Alfredo Brusco<sup>9,10</sup>, Roberto Keller<sup>11</sup>, Sarah Parisotto<sup>12</sup>, Helio F. Pedro<sup>12</sup>, Laina Lusk<sup>13,14,15</sup>, Pamela Pojomovsky McDonnell<sup>13,14,16</sup>, Ingo Helbig<sup>13,14,15,16</sup>, Sureni V. Mullegama<sup>17</sup>, Undiagnosed Diseases Network, Emilie D. Douine<sup>18</sup>, Bianca E. Russell<sup>18</sup>, Stanley F. Nelson<sup>18</sup>, Federico Zara<sup>1,2</sup>, and Stephen W. Scherer<sup>3,4,19,20</sup>

1 Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Genoa, Italy

2 UOC Genetica Medica, IRCCS Giannina Gaslini, Genoa, Italy

3 The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada

4 Genetics and Genome Biology Program, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada

5 Program in Neurosciences and Mental Health, The Hospital for Sick Children and Department of Physiology, University of Toronto, Toronto, Ontario, Canada.

6 Department of Psychology and Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada.

7 Department of Genetics and Metabolism, Randall Children's Hospital, Portland, OR 97227, USA

8 Department of Medical Sciences, University of Torino, Torino, Italy

9 Department of Neurosciences Rita Levi-Montalcini, University of Turin, 10126 Turin, Italy

10 Medical Genetics Unit, Città della Salute e della Scienza University Hospital, Torino, Italy

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

11 Adult Autism Centre DSM ASL Città di Torino, 10138 Turin, Italy

12 Center for Genetic and Genomic Medicine, Hackensack University Medical Center, Hackensack, New Jersey, USA

13 Epilepsy Neurogenetics Initiative, Children's Hospital of Philadelphia, Philadelphia, PA, USA

14 Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

15 Department of Biomedical and Health Informatics (DBHi), Children's Hospital of Philadelphia, Philadelphia, PA, USA

16 Department of Neurology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

17 GeneDx, Gaithersburg, MD, USA

18 Department of Human Genetics, David Geffen School of Medicine at University of California, Los Angeles, Los Angeles, CA, USA

19 Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada

20 McLaughlin Centre, Toronto, ON M5G 0A4, Canada

### **Correspondence:**

Stephen Scherer

The Hospital for Sick Children, Toronto, ON M5G 0A4, Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada.

Telephone: +1-416-813-6577

Electronic address: [stephen.scherer@sickkids.ca](mailto:stephen.scherer@sickkids.ca)

Marcello Scala

Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health University of Genoa, Genoa, Italy; UOC Genetica Medica, IRCCS Giannina Gaslini, Genoa, Italy

Telephone: +39-01056362758;

Electronic address: [mscala.md@gmail.com](mailto:mscala.md@gmail.com)

## Summary

Autism Spectrum Disorder (ASD) exhibits an ~4:1 male-to-female sex bias and is characterized by early-onset impairment of social/communication skills, restricted interests, and stereotyped behaviors. Disruption of the Xp22.11 locus has been associated with ASD in males. This locus includes the three-exon *PTCHD1* gene, an adjacent multi-isoform long noncoding RNA (lncRNA) named *PTCHD1-AS* (spanning ~1Mb), and a poorly characterized single-exon RNA helicase named *DDX53* that is intronic to *PTCHD1-AS*. While the relationship between *PTCHD1/PTCHD1-AS* and ASD is being studied, the role of *DDX53* has not been examined, in part because there is no apparent functional murine orthologue. Through clinical testing, here, we identified 6 males and 1 female with ASD from 6 unrelated families carrying rare, predicted-damaging or loss-of-function variants in *DDX53*. Then, we examined databases, including the Autism Speaks MSSNG and Simons Foundation Autism Research Initiative, as well as population controls. We identified 24 additional individuals with ASD harboring rare, damaging *DDX53* variations, including the same variants detected in two families from the original clinical analysis. In this extended cohort of 31 participants with ASD (28 male, 3 female), we identified 25 mostly maternally-inherited variations in *DDX53*, including 18 missense changes, 2 truncating variants, 2 in-frame variants, 2 deletions in the 3' UTR and 1 copy number deletion. Our findings in humans support a direct link between *DDX53* and ASD, which will be important in clinical genetic testing. These same autism-related findings, coupled with the observation that a functional orthologous gene is not found in mouse, may also influence the design and interpretation of murine-modelling of ASD.

**Keywords:** *DDX53*; RNA helicase; Xp22.11 locus; Autism; Autism spectrum disorder.

**Abbreviations:** ACMG-AMP = American College of Medical Genetics and Genomics and the Association for Molecular Pathology; ADHD = attention deficit-hyperactivity disorder; ASD = autism spectrum disorder; ES = exome sequencing; iPSC = induced pluripotent stem cell; NMD =

nonsense-mediated mRNA decay; SNVs = single nucleotide variants; CNV = copy number variant; SV = structural variant; SFARI - Simons Foundation Autism Research Initiative; EAGLE = evaluating evidence for ASD

Autism Spectrum Disorder (ASD) refers to a group of neurodevelopmental conditions characterized by a heterogeneous early-onset impairment of social interaction and communication, combined with the presence of restricted, repetitive, and stereotyped behaviors and interests<sup>1-4</sup>. Although cognitive development is variable in affected individuals, these disorders are the leading contributor to disability in preschool-age children<sup>3; 5; 6</sup>. The underlying abnormalities in brain development and functional connectivity are presumed to be secondary to interactions between the genetic and environmental factors, making ASD a complex and multifactorial condition<sup>2; 7</sup>.

ASD can occur in association with pathogenic sequence-level variants and copy number variations (CNVs; also, structural variations, SVs) in specific neuronal genes<sup>8-10</sup>. In a recent comprehensive annotation of all publicly available ASD genome sequence data, 134 ASD-associated genes were identified and ~15% of families studied had a clinically relevant rare variant. ASD-relevant gene lists are being curated by the Simons Foundation Autism Research Initiative (SFARI)<sup>11</sup> and the Evaluation of Autism Gene Link Evidence (EAGLE)<sup>12</sup> with both datasets available at <https://gene.sfari.org>.

Among the loci associated with ASD, chromosome Xp22.11 has been found to be disrupted in autistic males<sup>13-16</sup>. We present here a summary of the latest genotypic and phenotypic features mapping to this locus (Figure 1). It encompasses *DDX53*, *PTCHDI* (MIM \* 300828), and the long noncoding RNA (lncRNA) *PTCHDI-AS* (*PTCHDI* Antisense RNA (Head-To-Head))<sup>15</sup>. *Ptchd1* mutant mice show deficient attention and cognition, while variants in the human *PTCHDI* gene are linked to intellectual disability, impaired vision and more rarely, autistic features (Autism X-linked 4, AUTSX4 - MIM # 300830)<sup>17-19</sup>. However, many deletions detected in autistic participants also disrupt exons of the upstream lncRNA *PTCHDI-AS* and/or the protein-coding gene *DDX53*<sup>15; 18; 20</sup>

leaving uncertainty of a strict genotype-phenotype correlation. In our previous work, deletion mapping of males with ASD indicated that the exon 2-4 region of *PTCHDI-AS* (encompassing *DDX53*) represents a critical region and mouse modeling reaffirms a role for the lncRNA (unpublished).

*DDX53* is an intron-less gene of 3,630 nucleotides located within intron 3 of the canonical transcript of *PTCHDI-AS* encoding a 631-amino acid DEAD-box RNA helicase<sup>21</sup>. Phylogenetic analysis suggests *DDX53* and *DDX43* are highly related and that *DDX53* appears to be a retrogene derived from the 16-coding exon *DDX43* progenitor on chromosome 6<sup>22</sup>. The majority of retrogenes appear to originate from X-linked progenitors, possibly as a result of hypothesized compensation mechanisms<sup>23; 24</sup>. *DDX43* (72.8 kDa) and *DDX53* (71.2 kDa) proteins are also similar in size. Moreover, *DDX53* is also described as being a cancer-associated gene (CAGE), since it was first detected in human gastric cancer cell lines<sup>21</sup>. It is overexpressed in human cancers, where it promotes cancer stem cell-like properties such as self-renewal, cell motility, tumor spheroid formation, and anticancer drug resistance<sup>25; 26</sup>. DEAD-box RNA helicases have crucial roles in mRNA metabolism, serving as ATP-dependent helicases and unwindases, chaperones, and mediators of association or dissociation with protein interactors (RNPs)<sup>27; 28</sup>. Additionally, RNA helicases act as transcriptional co-activators or co-repressors, driving target mRNAs either to protein synthesis or degradation<sup>28</sup>. As such, these enzymes contribute to coordinate gene expression programs within the cell<sup>28</sup>.

In this study, we investigated 7 participants (6 males and 1 female) with ASD, with or without impaired psychomotor development, in whom 6 distinct maternally-inherited variants in *DDX53* were identified. The study involving human participants was conducted according to the guidelines of the Declaration of Helsinki. Ethical review and approval were obtained locally through the Gaslini Children's Hospital Research Committee for families I-VI (Comitato Etico della Regione Liguria, n.163/2018) and The Hospital for Sick Children Research Ethics Board. Family VI was consented at UCLA within IRB#15-000766 of the Undiagnosed Diseases Network. For all

families, informed consent for the study and publication of research work was obtained from the participant or parents/ legal guardians of all the enrolled participants. After the identification of the index case, further individuals were included in the study cohort based on the identification of rare and potentially deleterious variants in *DDX53*, along with overlapping autistic and neurodevelopmental phenotypes. The initial participants were recruited through international collaboration involving clinical and research centers based in France, Italy, and the USA (further details available in the Supplementary Material), and also using GeneMatcher<sup>29</sup>. Enrolled participants were evaluated by pediatric neurologists, neuropsychiatrists, and neurogeneticists. For ASD assessment, thorough clinical examination and dedicated tests were employed, such as the Autism Diagnostic Interview-Revised (ADI-R) and Ritvo Autism Asperger Diagnostic Scale–Revised (RAADS–R) (Supplementary Material).

For genetic testing, trio-exome sequencing was performed for participant families I, II, and VI, and trio-whole genome sequencing was performed for VII. Participants #3, and #5 were investigated through the Next Generation Sequencing-based Autism/ID Xpanded Panel (GeneDx, Gaithersburg, USA) and participant #4 had targeted sequencing (GeneDx, Gaithersburg, USA) (Supplementary Material). Genomic DNA was extracted from peripheral blood lymphocytes of the probands and parents, with the sequencing being performed as described previously (Supplementary Material). Variants were filtered according to minor allele frequency  $\leq 0.001$  in Genome Aggregation Database (gnomAD), were assessed for levels of conservation (Genomic Evolutionary Rate Profiling – GERP), and functional impact was predicted according to different *in silico* tools (Supplementary Material), as well as the Variant Effect Predictor (VEP) pipeline from Ensembl. Candidate variants in known disease genes were classified according to ACMG guidelines<sup>30</sup>. Variants were mapped onto the protein structures and intolerance to variation of the affected amino acid residues was analyzed using the Metadome online software (<https://stuart.radboudumc.nl/metadome/method>)<sup>31</sup>. Copy number variation (CNV) assessment was performed either through array comparative genomic hybridization (aCGH) or whole

genome/exome-based CNV calling (Supplementary Material). The DECIPHER database and Database of Genomic Variants were used to interpret detected CNVs<sup>32;33</sup>. *DDX53* variants are reported according to the NM\_182699.4 transcript, corresponding to the unique isoform of the *DDX53* protein NP\_874358.2 (<https://www.ncbi.nlm.nih.gov/gene/168400>).

We identified six distinct variants in *DDX53* in the reported individuals: c.871A>G, p.(Arg291Gly) in #1; c.834G>A, p.(Met278Ile) in #2; c.567\_568insTGCAGGT, p.(Glu190Cysfs\*3) in #3 and #4; c.977G>A, p.(Ser326Asn) in #5; c.1736C>T, p.(Ala579Val) in #6; c.259G>T, p.(Gly87Trp) in #7. All variants were maternally inherited and segregated with the phenotypes in all families (Figure 2A). In family III, both participants harbored the same p.(Glu190Cysfs\*3) variant, although no information about a potential non-skewed X-inactivation in the female sibling (#4) was available. In family IV, the p.(Ser326Asn) variant identified in the proband (#5) was mosaic (4%) in the mother. In family V, the brother of participant #6 was reported to show autistic features, but segregation analysis could not be performed. In family VI, there were no other maternal male relatives for further segregation analysis available. All of the detected *DDX53* variants are rare (maximum allele frequency 0.000005534) and absent in the hemizygous state in gnomAD. The missense variants p.(Gly87Trp), p.(Met278Ile), p.(Arg291Gly), p.(Ser326Asn), and p.(Ala579Val) affect conserved residues (GERP scores ranging between 3.55 and 4.59) within or in close proximity to the KH, DEXDc, and HELICc domains (Figure 2B). These changes are predicted to be damaging according to *in silico* tools (Table S1). The frameshift variant p.(Glu190Cysfs\*3) is likely to lead to either nonsense-mediated mRNA decay (NMD) or the formation of a truncated transcript. Pathogenicity using the ACMG classifications could not be definitively assigned for this gene, as ACMG guidelines are only applicable for genes with already established gene-disease associations. The predicted effects of these variants are indicated in Tables 1 and 2.

The *DDX53* participants included in our cohort presented with marked ASD-related features, including poor social interaction, limited communication skills, self-stimulating behaviors,

weak visual problem-solving skills, and sensory hyperresponsiveness (Table 1). Eye contact avoidance was observed in all cases but #7. Motor and/or verbal stereotypies were present in four participants (#3, #5, #6, and #7), sometimes in association with ritualistic behaviors (Table S2). A form of high-functioning autism within the spectrum of Asperger syndrome was diagnosed in #2. Interestingly, while two participants from Family III presented with similar overlapping ASD phenotypes, female participant #4 showed a more severe phenotype including significant speech limitation, social avoidance, sensory hypersensitivity, and impaired visual problem-solving skills (no additional variants explaining this were found). Autistic features were associated with impaired psychomotor development in five out of seven participants (#3-#7), variably affecting speech development and/or motor abilities. However, this developmental delay only led to variable skills, generally in the mild to moderate range, in participant #7. Additional neuropsychiatric manifestations consisted of seizures (#1, #6, and #7), heterogeneous sleep disorders (#3-#7), and behavioral comorbidities (#3, #5-#7), such as attention-deficit hyperactivity disorder (ADHD), anxiety, and aggressiveness.

After the identification of the *DDX53* variants in the clinical cohorts, we searched the SFARI (<https://www.sfari.org/resource/sfari-gene/>) and MSSNG (<https://research.mss.ng>) databases. Within SFARI<sup>11; 34</sup>, we screened for *DDX53* variants in the Simons Simplex Collection (SSC) and Simons Foundation Powering Autism Research (SPARK) datasets. SSC consists of WGS data from 2,600 simplex families, each containing one proband with ASD as well as unaffected parents and siblings. SPARK (WES1, WES2) contains the whole exome sequence for 43,259 participants (including 21,900 with ASD) with phenotype information. Additionally, we screened the MSSNG WGS database<sup>8; 35-37</sup>, which at the time consisted of 11,312 individuals, including 5,100 with ASD.

By screening the SSC, SPARK, and MSSNG cohorts, we detected 19 additional interesting variations in *DDX53* in 24 individuals, including 22 males and 2 females, presenting with mostly isolated ASD. These included 13 missense changes, 1 stopgain, 2 in-frame variants, 2 distinct



deletions within the 3'UTR and one deletion CNV (Table 2). Fourteen of these variations were of maternal origin, whereas information about the segregation was not available in the remaining cases. Resembling those identified in our cohort, the *DDX53* variants detected in participants from the open-science autism datasets are exceptionally rare (max allele frequency of 0.0000563) and are absent in hemizygous state in gnomAD. Most missense changes affect conserved residues within different functional domains of the *DDX53* helicase (Figure 2B) and are predicted deleterious by *in silico* tools (Table S3). The novel truncating variant p.(Trp8\*) was predicted to exert a deleterious effect on protein expression, whereas the in-frame changes (p.(Gln591del) and p.(Met381\_Leu382delinsIleMet) resulted in the deletion or substitution of highly conserved amino acids, respectively. Of note, in the SPARK dataset, we also identified participants that were already part of our main cohort: subjects #3 (SP0039198), #4 (SP0042895), and #6 (SP0090623).

Deletions within the 3'UTR of *DDX53* were identified in five participants. The [GRCh38] Xp22.11(23002082-23002458)x1 was detected in one individual (7-0111-003) and the [GRCh38]Xp22.11(23002148-23002522)x1 in four affected brothers (1-0627-003, 1-0627-005, 1-0627-006, and 1-0627-007). These deletions selectively involve a portion of the 3'UTR of the gene and are absent in the DECIPHER database. Overall, the identified *DDX53* variations are predicted to affect protein function through the alteration of conserved residues within functional domains or the loss of a functional transcript/part of the 3'UTR. Additionally, a maternally inherited, 21.8 kb deletion [GRCh38]Xp22.11(22987677-23009507)x1 impacting only *DDX53* and not exons of *PTCHD1-AS* was observed in one affected female. However, this individual also carries a 331 kb *de novo* deletion [GRCh38]7q11.22 (70441029-70771836)x1 impacting 5 exons of *AUTS2* (MIM\*607270), a gene strongly associated with ASD<sup>38</sup>. In the aforementioned individual, the *AUTS2* deletion is likely involved in the ASD, but the carrier status of the *DDX53* deletion may also need to be considered in male offspring.

In total, in our extended cohort of 31 individuals, including 28 males and 3 females, we identified 25 distinct *DDX53* variations. These included 18 missense changes, 2 truncating variants,

2 in-frame variants, 2 deletions of the 3'UTR, and one complete gene deletion. Four variants were found in more than one individual, including p.(Glu190Cysfs\*3) in two (#3 - SP0039198 and, #4 - SP0042895); p.(Asp575Gly) in two (SP0126624 and 1-0551-003); and p.(Gln591del) in two (SP0066816 and SP0066817). Of note, the [GRCh38]Xp22.11(23002148\_23002522)x1 deletion was also identified in four participants from the same family with a similar deletion found in another family with one affected male. Missense changes affected conserved residues within or in proximity to the KH, DEXDc, or HELICc functional domains (Figure 1B). *In silico* tools revealed that these variants are variably deleterious (Table S1 and Table S3) and mostly map to amino acid residues intolerant to variation according to Metadome (Figure 3 and Table S4). One of the in-frame variants included the indel p.(Met381\_Leu382delinsIleMet), causing the substitution of two conserved amino acids (GERP scores = 4.3) within the DEXDc domain, and p.(Gln591del), resulting in a deletion of a single conserved residue (GERP score = 4.46) in the C-terminus of the protein. The two truncating variants, which were identified in three participants, featured the stopgain p.(Trp8\*) and the frameshift p.(Glu190Cysfs\*3). These variants are likely to be deleterious based on the predicted severe impact on protein function and the loss of function intolerance of *DDX53* (probability of loss intolerance (pLI) of 0.66 and a loss of function observed/expected upper bound fraction (LOEUF) of 0.847). Of note, affected females (#4 - SP0042895, SP0003053 and 1-0485-003) either harbored truncating *DDX53* changes or a whole gene deletion. Although further studies are required to elucidate the X-inactivation pattern in these cases, this observation suggests that a severe *DDX53* loss of function may lead to a clinical phenotype in female carriers.

We scrutinized the mouse databases for *Ddx53* and based on the available mouse reference data at UCSC Genome Browser (GRCm39/mm39), there does not appear to be an orthologous transcript in the expected syntenic region (Figure 4)<sup>39</sup>. Instead, there is a gap in alignment between the mouse and human regions; however, that gap does not exist in the DNA of higher mammals. We designed primers in the repeat-free sequence on each side of the alignment gap to determine if there was any additional sequence in this gap not captured in the murine reference genome.

(forward: ttcaacctatttgtgtgcatc; reverse: tccagcacagtattaccagaaatagt). Using BAC clone RP23-405B14 from the C57BL/6 mouse strain as DNA template, PCR was performed with an expected product size of 4,029 bp. Based on gel electrophoresis the product appeared to be ~4 kb. This product was cloned into a Topo TA cloning vector (Invitrogen), and Sanger sequenced with the original primers, as well as walking primers, to generate the entire 4,029 bp cloned product. The Sanger sequence obtained corresponds with the reference sequence GRCm39/mm39. We repeated this PCR using template DNA from mouse genomic DNA (purchased from Cedarlane) and DNA extracted from internal C57BL/6 mouse brain P14. In each case, the expected product size was amplified and the Sanger sequencing agreed with the reference sequence. All data suggest that there is no additional sequence in this region, in this strain, and therefore a mouse *Ddx53* gene does not appear to be present in the syntenic region to *DDX53*. These observations are further validated because no obvious RNA transcripts were found for *Ddx53* in any database nor in our unpublished murine data. The rat ortholog (*Ddx43*), however, is annotated as a *DDX43/DDX53* homolog and appears to be functional (<https://www.genenames.org/tools/hcop/#!/>). The naked mole-rat model has a more specific *DDX53* homolog (<https://www.ncbi.nlm.nih.gov/gene/101703948/>). Other potentially functional homologs are present in rabbits and other small mammals (<https://www.ncbi.nlm.nih.gov/gene/168400/ortholog/?scope=9347&term=DDX53>) (Figure 4)<sup>40</sup>.

RNA helicases contribute in mRNA metabolism, mediating a range of biochemical activities not restricted to RNA unwinding, such as ribosome biogenesis, RNA degradation, and splicing<sup>41</sup>. While processive RNA helicases function by translocating along the nucleic acids, the non-processive helicases selectively unwind local secondary structures without translocating on their substrates<sup>28; 42</sup>. In humans, the two largest families of helicases are: the DEAD box (DDX) family, encompassing 41 members (<https://www.genenames.org/data/genegroup/#!/group/499>) ; and the DEAH box (DHX), including 16 members (<https://www.genenames.org/data/genegroup/#!/group/500>). The DEAD-box helicases belong to the non-processive group<sup>43</sup>. All DDX helicases share a conserved core consisting of two tandem RecA

domains, with recognizable sequence motifs involved in RNA binding, and ATP binding and hydrolysis<sup>41; 42</sup>. Additional C- and N-terminal extensions vary among the different members and consist of auxiliary domains implicated in RNA recognition and the regulation of ATPase activity<sup>41; 42</sup>.

Genetic variants affecting the function of DDX helicases have been shown to be involved in other neurodevelopmental disorders<sup>44</sup>. In particular, *de novo* variants in *DDX3X* (MIM \* 300160), mapping to chromosome Xp11.3-p11.23, cause a syndromic neural condition (MIM # 300958) with dysmorphism, congenital defects, and neurobehavioral comorbidities including autistic features, either in males or females with different disease severity<sup>45; 46</sup>. More recently, although in absence of supporting functional evidence, missense changes in *DDX54* (MIM \* 611665) have been associated with psychomotor development impairment in three unrelated participants, of whom one exhibited repetitive movements and behaviors<sup>44</sup>.

*DDX53* is a little characterized RNA helicase<sup>21</sup> whose main functional core consists of the DEAD-like Helicases superfamily (DEXDC) domain, which contains the ATP binding site and mediates the ATP-dependent RNA unwinding activity<sup>21; 47</sup>. Additionally, the protein has two auxiliary domains; the N-terminal K homology (KH) domain, which favors single-strand RNA recognition and binding, and the helicase conserved C-terminal (HELICc) domain, which interacts with the DEXDC domain for the helicase activity<sup>41; 47</sup>. So far, the *DDX53* helicase function has only been elucidated in cancer, where its overexpression appears to mediate the acquisition of oncogenic properties in tumoral cells<sup>21; 25; 26</sup>. Based on its elevated gene expression in testis, polymorphisms in *DDX53* may be associated with azoospermia, but this putative finding requires further testing<sup>48; 49</sup>.

It is noteworthy that in our own recent study, Neurogenin 2 (NGN2)-derived cortical excitatory neurons harboring termination codons in *DDX53* showed a comparable activity in the multi-electrode array to isogenic control lines<sup>50</sup>. The lack of a distinctive electrophysiological phenotype and the ability of the knockout induced pluripotent stem cell (iPSC) lines of generating

excitatory cortical neurons suggested that loss of *DDX53* does not lead to a glutamatergic dysregulation phenotype, which may argue against an overt role in ASD<sup>50</sup>. However, the isogenic iPSC-derived control neurons investigated in this study showed low levels of *DDX53* transcripts, resulting in protein levels below the threshold for western blot detection<sup>50</sup>. Thus, it is possible that these cells are not the most appropriate model to investigate *DDX53* function; a conjecture further supported through the expression of the *DDX53* gene in the human brain, where significant RNA expression is limited to the cerebellum<sup>15</sup> (<https://www.proteinatlas.org/ENSG00000184735-DDX53/tissue>). The cerebellum is responsible for the maturation of nonmotor neural circuitry, thus having a crucial role in guiding cognitive development<sup>51</sup>. Furthermore, specific cerebellar subzones modulate the activity of neocortical substrates contributing to the determination of social interaction patterns during developmentally sensitive periods<sup>51</sup>.

In summary, our genetic data suggest a role for *DDX53* in ASD. Interestingly, at the same Xp22.11 locus, there is also supporting evidence based on non-overlapping microdeletions, for the *PTCHD1-AS* lncRNA, which encompasses *DDX53*, to be independently involved in ASD<sup>15; 20</sup> (Figure 1). Moreover, at the same locus, for the protein-coding *PTCHD1*, non-synonymous variants have also been found to cause intellectual disability, and in a few individuals, an accompanying diagnosis of autism was present<sup>18</sup>. In murine models, *Ptchd1* deficiency leads to cognitive dysfunction and abnormal behavior with significant hyperactivity; however, despite these behavioral abnormalities they do not strictly resemble the autism spectrum<sup>17; 19</sup>. *Ptchd1-as* cell<sup>20</sup> and mouse models (unpublished) are being generated. As additional participants with ASD that have sequence-level or deletion variants impacting *DDX53*, *PTCHD1-AS* and/or *PTCHD1* are identified, we anticipate an increasingly complex social-behavioral and clinical presentation when two or more of these co-located genes are impacted. Given the newly found role for *DDX53* in ASD, and that there is no apparent functional ortholog (*Ddx53*) in mouse, generating a ‘humanized’ *Ddx53* murine model may be important for future mouse-modelling of autism.

## **Acknowledgements**

We acknowledge technical support from Bhooma Thiruvahindrapuram, and Sylvia Lamoureux. We thank the research participants and their families (including those participating in MSSNG) as well as the generosity of the donors who support this program. We acknowledge the resources of Autism Speaks and The Centre for Applied Genomics. We are grateful to all of the families described in this paper including those involved in the Autism Speaks MSSNG Project, as well as the SSC and SPARK sites including the principal investigators (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, R. Goin-Kochel, E. Hanson, D. Grice, A. Klin, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, K. Pelphrey, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, Z. Warren, and E. Wijsman). We appreciate obtaining access to genetic and phenotypic data on SFARI Base.

## **Funding**

Funding was provided by the University of Toronto McLaughlin Centre, Autism Speaks, Autism Speaks Canada, Ontario Brain Institute, SickKids Foundation and the Italian Ministry for Education, University and Research (Ministero dell'Istruzione, dell'Università e della Ricerca - MIUR) PRIN2020 code 20203P8C3X (Alfredo Brusco). Additional funding was partially provided by NIH award U01NS134356 and supported by the California Center for Rare Diseases within the Institute for Precision Health at UCLA. S.F.N holds the Dr. Allen and Charolotte Ginsburg Endowed Chair in Translational Genomics at the David Geffen School of Medicine. S.W.S. holds the Northbridge Chair in Pediatric Research at The Hospital for Sick Children and the University of Toronto.

## **Author Contributions**

Conceptualization: MS, SWS; Formal Analysis: MS, JLH, NBS, CS, MSR, JRM, SWS; Funding Acquisition: MS, SFN; SWS; Investigation: MS, CAB, JLH, BT, CS, MSR, JRM, SYK, FZ, SWS;

Methodology: MS, SWS; Project Administration: MS, JLH, SWS; Resources: MS, PWF, LG, GA, VP, AB, RK, SP, HFP, LL, PPM, IH, SVM, ED-B, BER, SFN, SWS; Supervision: MS, SWS; Visualization: MS, JLH, CS, MSR, SWS; Writing: MS, JLH, NBS, CS, JRM, SWS

## Web resources

CADD; <https://cadd.gs.washington.edu>

ClinVar; <https://www.ncbi.nlm.nih.gov/clinvar>

Combined Annotation Dependent Depletion (CADD); <http://cadd.gs.washington.edu>

Database of Genomic Variants (DGV); <http://dgv.tcag.ca/dgv/app/home>

DECIPHER; <https://decipher.sanger.ac.uk>

Gene Cards; <https://www.genecards.org>

Gene Matcher; <http://www.genematcher.org>

Genome Aggregation Database (GnomAD); <http://gnomad.broadinstitute.org>

Genomic Evolutionary Rate Profiling (GERP);

<http://mendel.stanford.edu/SidowLab/downloads/gerp/>

gnomAD; <https://gnomad.broadinstitute.org>

Metadome online software; <https://stuart.radboudumc.nl/metadome/method>

MSSNG; <https://research.mss.ng>

Online Mendelian Inheritance in Man; <https://www.ncbi.nlm.nih.gov/Omim>

Protein Atlas; <https://www.proteinatlas.org/ENSG00000184735-DDX53/tissue>

Polyphen-2; <http://genetics.bwh.harvard.edu/pph2/>

PubMed; <https://www.ncbi.nlm.nih.gov/pubmed>

RefSeq; <https://www.ncbi.nlm.nih.gov/refseq>

SFARI; <https://www.sfari.org/resource/sfari-gene/>

SIFT; <https://sift.bii.a-star.edu.sg>

UniProt; <https://www.uniprot.org>

UCSC Human Genome Database; <https://www.genome.ucsc.edu>

Varsome; <https://varsome.com>

Variant Effect Predictor (VEP) - Ensembl; <https://www.ensembl.org/info/docs/tools/vep/index.html>

### **Data availability**

The data that support the findings of this study are available from the corresponding authors upon request or included in the databases mentioned and linked in the manuscript.

### **Competing interests**

At the time of this study and its publication, S.W.S. served on the Scientific Advisory Committee of Population Bio. Intellectual property from aspects of his research held at The Hospital for Sick Children are licensed to Athena Diagnostics and Population Bio. These relationships did not influence data interpretation or presentation during this study but are disclosed for potential future considerations. SVM is an employee of GeneDx, LLC. HFP is on the research advisory boards and speaker bureau for Takeda Pharmaceutical, AvroBio, Amicus Therapeutics, Sanofi, Alexion Therapeutics, Denali Therapeutics and Acer Therapeutics. All other authors declare no conflict of interest.

### **References**

1. Lai, M.C., Lombardo, M.V., and Baron-Cohen, S. (2014). Autism. *Lancet* 383, 896-910.
2. Fakhoury, M. (2015). Autistic spectrum disorders: A review of clinical features, theories and diagnosis. *Int J Dev Neurosci* 43, 70-77.
3. Hirota, T., and King, B.H. (2023). Autism Spectrum Disorder: A Review. *JAMA* 329, 157-168.
4. Anagnostou, E., Zwaigenbaum, L., Szatmari, P., Fombonne, E., Fernandez, B.A., Woodbury-Smith, M., Brian, J., Bryson, S., Smith, I.M., Drmic, I., et al. (2014). Autism spectrum disorder: advances in evidence-based practice. *CMAJ* 186, 509-519.



5. Chakrabarti, S., and Fombonne, E. (2005). Pervasive developmental disorders in preschool children: confirmation of high prevalence. *Am J Psychiatry* 162, 1133-1141.
6. Baxter, A.J., Brugha, T.S., Erskine, H.E., Scheurer, R.W., Vos, T., and Scott, J.G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychol Med* 45, 601-613.
7. Qiu, S., Qiu, Y., Li, Y., and Cong, X. (2022). Genetics of autism spectrum disorder: an umbrella review of systematic reviews and meta-analyses. *Transl Psychiatry* 12, 249.
8. Trost, B., Thiruvahindrapuram, B., Chan, A.J.S., Engchuan, W., Higginbotham, E.J., Howe, J.L., Loureiro, L.O., Reuter, M.S., Roshandel, D., Whitney, J., et al. (2022). Genomic architecture of autism from comprehensive whole-genome sequence annotation. *Cell* 185, 4409-4427 e4418.
9. Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.Y., Peng, M., Collins, R., Grove, J., Klei, L., et al. (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* 180, 568-584 e523.
10. Fu, J.M., Satterstrom, F.K., Peng, M., Brand, H., Collins, R.L., Dong, S., Wamsley, B., Klei, L., Wang, L., Hao, S.P., et al. (2022). Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. *Nat Genet* 54, 1320-1331.
11. Banerjee-Basu, S., and Packer, A. (2010). SFARI Gene: an evolving database for the autism research community. *Dis Model Mech* 3, 133-135.
12. Schaaf, C.P., Betancur, C., Yuen, R.K.C., Parr, J.R., Skuse, D.H., Gallagher, L., Bernier, R.A., Buchanan, J.A., Buxbaum, J.D., Chen, C.A., et al. (2020). A framework for an evidence-based gene list relevant to autism spectrum disorder. *Nat Rev Genet* 21, 367-376.
13. Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y., et al. (2008). Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 82, 477-488.

14. Pinto, D., Pagnamenta, A.T., Klei, L., Anney, R., Merico, D., Regan, R., Conroy, J., Magalhaes, T.R., Correia, C., Abrahams, B.S., et al. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 466, 368-372.
15. Noor, A., Whibley, A., Marshall, C.R., Gianakopoulos, P.J., Piton, A., Carson, A.R., Orlic-Milacic, M., Lionel, A.C., Sato, D., Pinto, D., et al. (2010). Disruption at the PTCHD1 Locus on Xp22.11 in Autism spectrum disorder and intellectual disability. *Sci Transl Med* 2, 49ra68.
16. Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., Thiruvahindrapuram, B., Xu, X., Ziman, R., Wang, Z., et al. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet* 94, 677-694.
17. Wells, M.F., Wimmer, R.D., Schmitt, L.I., Feng, G., and Halassa, M.M. (2016). Thalamic reticular impairment underlies attention deficit in *Ptchd1*(Y/-) mice. *Nature* 532, 58-63.
18. Chaudhry, A., Noor, A., Degagne, B., Baker, K., Bok, L.A., Brady, A.F., Chitayat, D., Chung, B.H., Cytrynbaum, C., Dyment, D., et al. (2015). Phenotypic spectrum associated with PTCHD1 deletions and truncating mutations includes intellectual disability and autism spectrum disorder. *Clin Genet* 88, 224-233.
19. Ung, D.C., Iacono, G., Meziane, H., Blanchard, E., Papon, M.A., Selten, M., van Rhijn, J.R., Montjean, R., Rucci, J., Martin, S., et al. (2018). *Ptchd1* deficiency induces excitatory synaptic and cognitive dysfunctions in mouse. *Mol Psychiatry* 23, 1356-1367.
20. Ross, P.J., Zhang, W.B., Mok, R.S.F., Zaslavsky, K., Deneault, E., D'Abate, L., Rodrigues, D.C., Yuen, R.K.C., Faheem, M., Mufteev, M., et al. (2020). Synaptic Dysfunction in Human Neurons With Autism-Associated Deletions in PTCHD1-AS. *Biol Psychiatry* 87, 139-149.
21. Cho, B., Lim, Y., Lee, D.Y., Park, S.Y., Lee, H., Kim, W.H., Yang, H., Bang, Y.J., and Jeoung, D.I. (2002). Identification and characterization of a novel cancer/testis antigen gene CAGE. *Biochem Biophys Res Commun* 292, 715-726.

22. Vinckenbosch, N., Dupanloup, I., and Kaessmann, H. (2006). Evolutionary fate of retroposed gene copies in the human genome. *Proc Natl Acad Sci U S A* 103, 3220-3225.
23. Marques, A.C., Dupanloup, I., Vinckenbosch, N., Reymond, A., and Kaessmann, H. (2005). Emergence of young human genes after a burst of retroposition in primates. *PLoS Biol* 3, e357.
24. Wang, P.J. (2004). X chromosomes, retrogenes and their role in male reproduction. *Trends Endocrinol Metab* 15, 79-83.
25. Shim, H., Shim, E., Lee, H., Hahn, J., Kang, D., Lee, Y.S., and Jeoung, D. (2006). CAGE, a novel cancer/testis antigen gene, promotes cell motility by activation ERK and p38 MAPK and downregulating ROS. *Mol Cells* 21, 367-375.
26. Kim, Y., Yeon, M., and Jeoung, D. (2017). DDX53 Regulates Cancer Stem Cell-Like Properties by Binding to SOX-2. *Mol Cells* 40, 322-330.
27. Schutz, P., Karlberg, T., van den Berg, S., Collins, R., Lehtio, L., Hogbom, M., Holmberg-Schiavone, L., Tempel, W., Park, H.W., Hammarstrom, M., et al. (2010). Comparative structural analysis of human DEAD-box RNA helicases. *PLoS One* 5.
28. Bourgeois, C.F., Mortreux, F., and Auboeuf, D. (2016). The multiple functions of RNA helicases as drivers and regulators of gene expression. *Nat Rev Mol Cell Biol* 17, 426-438.
29. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat* 36, 928-930.
30. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17, 405-424.

31. Wiel, L., Venselaar, H., Veltman, J.A., Vriend, G., and Gilissen, C. (2017). Aggregation of population-based genetic variation over protein domain homologues and its potential use in genetic diagnostics. *Hum Mutat* 38, 1454-1463.
32. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet* 84, 524-533.
33. MacDonald, J.R., Ziman, R., Yuen, R.K., Feuk, L., and Scherer, S.W. (2014). The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res* 42, D986-992.
34. Abrahams, B.S., Arking, D.E., Campbell, D.B., Mefford, H.C., Morrow, E.M., Weiss, L.A., Menashe, I., Wadkins, T., Banerjee-Basu, S., and Packer, A. (2013). SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Mol Autism* 4, 36.
35. Jiang, Y.H., Yuen, R.K., Jin, X., Wang, M., Chen, N., Wu, X., Ju, J., Mei, J., Shi, Y., He, M., et al. (2013). Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *Am J Hum Genet* 93, 249-263.
36. Yuen, R.K., Thiruvahindrapuram, B., Merico, D., Walker, S., Tammimies, K., Hoang, N., Chrysler, C., Nalpathamkalam, T., Pellecchia, G., Liu, Y., et al. (2015). Whole-genome sequencing of quartet families with autism spectrum disorder. *Nat Med* 21, 185-191.
37. Yuen, R., Merico, D., Bookman, M., Howe, J., Thiruvahindrapuram, B., Patel, R.V., Whitney, J., Deflaux, N., Bingham, J., Wang, Z., et al. (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat Neurosci* 20, 602-611.
38. Hori, K., Yamashiro, K., Nagai, T., Shan, W., Egusa, S.F., Shimaoka, K., Kuniishi, H., Sekiguchi, M., Go, Y., Tatsumoto, S., et al. (2020). AUTS2 Regulation of Synapses for Proper Synaptic Inputs and Social Communication. *iScience* 23, 101183.

39. Kolishovski, G., Lamoureux, A., Hale, P., Richardson, J.E., Recla, J.M., Adesanya, O., Simons, A., Kunde-Ramamoorthy, G., and Bult, C.J. (2019). The JAX Synteny Browser for mouse-human comparative genomics. *Mamm Genome* 30, 353-361.
40. Li, H., Coghlan, A., Ruan, J., Coin, L.J., Heriche, J.K., Osmotherly, L., Li, R., Liu, T., Zhang, Z., Bolund, L., et al. (2006). TreeFam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res* 34, D572-580.
41. Putnam, A.A., and Jankowsky, E. (2013). DEAD-box helicases as integrators of RNA, nucleotide and protein binding. *Biochim Biophys Acta* 1829, 884-893.
42. Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. *Trends Biochem Sci* 36, 19-29.
43. Linder, P., and Jankowsky, E. (2011). From unwinding to clamping - the DEAD box RNA helicase family. *Nat Rev Mol Cell Biol* 12, 505-516.
44. Paine, I., Posey, J.E., Grochowski, C.M., Jhangiani, S.N., Rosenheck, S., Kleyner, R., Marmorale, T., Yoon, M., Wang, K., Robison, R., et al. (2019). Paralog Studies Augment Gene Discovery: DDX and DHX Genes. *Am J Hum Genet* 105, 302-316.
45. Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M.R., Venselaar, H., Helsmoortel, C., Cho, M.T., Hoischen, A., et al. (2015). Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *Am J Hum Genet* 97, 343-352.
46. Scala, M., Torella, A., Severino, M., Morana, G., Castello, R., Accogli, A., Verrico, A., Vari, M.S., Cappuccio, G., Pinelli, M., et al. (2019). Three de novo DDX3X variants associated with distinctive brain developmental abnormalities and brain tumor in intellectually disabled females. *Eur J Hum Genet* 27, 1254-1259.
47. Floor, S.N., Condon, K.J., Sharma, D., Jankowsky, E., and Doudna, J.A. (2016). Autoinhibitory Interdomain Interactions and Subfamily-specific Extensions Redefine the Catalytic Core of the Human DEAD-box Protein DDX3. *J Biol Chem* 291, 2412-2421.

48. Malcher, A., Stokowy, T., Berman, A., Olszewska, M., Jedrzejczak, P., Sielski, D., Nowakowski, A., Rozwadowska, N., Yatsenko, A.N., and Kurpisz, M.K. (2022). Whole-genome sequencing identifies new candidate genes for nonobstructive azoospermia. *Andrology* 10, 1605-1624.
49. Hardy, J., Pollock, N., Gingrich, T., Sweet, P., Ramesh, A., Kuong, J., Basar, A., Jiang, H., Hwang, K., Vukina, J., et al. (2022). Genomic testing for copy number and single nucleotide variants in spermatogenic failure. *J Assist Reprod Genet* 39, 2103-2114.
50. Faheem, M., Deneault, E., Alexandrova, R., Rodrigues, D.C., Pellecchia, G., Shum, C., Zarrei, M., Piekna, A., Wei, W., Howe, J.L., et al. (2023). Disruption of *DDX53* coding sequence has limited impact on iPSC-derived human NGN2 neurons. *BMC Med Genomics* 16, 5.
51. Wang, S.S., Kloth, A.D., and Badura, A. (2014). The cerebellum, sensitive periods, and autism. *Neuron* 83, 518-532.

## Legends

**Table 1. Clinical features of participants with of *DDX53* variants.**

**Table 2. Genetic and clinical features of ASD participants with *DDX53* variants in the SFARI and MSSNG databases.**

**Figure 1. Summary of the genotypic and phenotypic features of the ASD-related Xp22.11 locus.** The Xp22.11 locus includes *DDX53* (MIM \* 301079), *PTCHD1* (MIM \* 300828), and the long noncoding RNA (lncRNA) *PTCHD1-AS* (*PTCHD1* Antisense RNA (Head To Head)). A number of deletions identified in ASD participants were found to disrupt exons of the upstream lncRNA *PTCHD1-AS* and/or the *DDX53* gene. a = SFARI Gene with curated EAGLE gene scores: <https://gene.sfari.org/>; b = ClinGen <https://clinicalgenome.org/>; c = EAGLE gene curation <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10357004>.

**Figure 2. Pedigrees of the reported families and genetic findings in all *DDX53* participants. A)**

Pedigrees of the six families showing the segregation of the *DDX53* variants in affected individuals and the parents. All variants were inherited from unaffected mothers. The carrier status is indicated by small-filled circles. The mosaic status of the mother of the proband #5 is indicated by a small empty circle. **B)** Schematic representation of the total of the *DDX53* variants mapped to the unique protein isoform (NP\_874358.2). The variants identified in the described families and those identified in participants from the SFARI and MSSNG databases are reported in blue and black, respectively. DEXDc = DEAD-like Helicases superfamily domain; HELICc = helicase conserved C-terminal domain; KH = K homology domain.

**Figure 3. Graphic representation of intolerance to *DDX53* variants.** Using the Metadome software (<https://stuart.radboudumc.nl/metadome/>), we mapped the variants identified in our cohort and in the SFARI and MSSNG datasets to the *DDX53* protein. Most variants were found to affect amino acid residues that showed intolerance to variation according to their variation in the gnomAD dataset. While some amino acids were rarely impacted by genetic changes, variants in other residues are absent in gnomAD.

**Figure 4. A)** Graphical representation of human, mouse, and rat syntenic regions and breaks in chrX:10,445,310-35,803,753 (GRCh38). Syntenic alignment between human and mouse show inverted orientation of two corresponding segments between human and mouse. Syntenic alignment between human and rat is shown to highlight the mouse-specific inversion of this chromosomal segment, and the shared syntenic break located just upstream of the *PTCHD1* gene. **B)** Inset of human-mouse syntenic region at the *PTCHD1-AS* locus. **C)** Full phylogenetic tree of *DDX53* and *DDX43* (right panel) created by Treefam. Red triangle denotes a duplication event. Green circle denotes a speciation event. Inset of *DDX53* orthologs (left panel). **D)** Phylogenetic tree of *DDX53* and *DDX43* in common model organisms.

**Table S1. Extended *in silico* analysis of *DDX53* variants in our cohort.**

**Table S2. Extended clinical phenotype of participants with *DDX53* variants.**

**Table S3. Extended *in silico* analysis of all *DDX53* variants of participants from SFARI and MSSNG databases.**

**Table S4. Details of intolerance to variation for the amino acid residues affected by *DDX53* variants.**

**Table S5. UDN membership list.**



**Table 1. Clinical features of participants with *DDX53* variants.**

Patient ID (Family)	#1 (I)	#2 (II)	#3 (III)	#4 (III)	#5 (IV)	#6 (V)	#7 (VI)
<b>General information</b>							
Sex	M	M	M	F	M	M	M
Age	5-10 years	35-40 years	5-10 years	5-10 years	6-10 years	5-10 years	5-10 years
<b>Genetic data</b>							
Testing method	Trio-ES	Trio-ES	Autism/ID Xpanded Panel	Targeted variant testing	Autism/ID Xpanded Panel	Trio-ES	Trio-WGS and ES
cDNA (NM_182699.4)	c.871A>G	c.834G>A	c.567_568insTGCAG GT	c.567_568insTGCAG GT	c.977G>A	c.1736C>T	c.259G>T
Protein	p.(Arg291Gly)	p.(Met278Ile)	p.(Glu190Cysfs*3)	p.(Glu190Cysfs*3)	p.(Ser326Asn)	p.(Ala579Val)	p.(Gly87Trp)
Mutation type	Missense	Missense	Frameshift	Frameshift	Missense	Missense	Missense
Affected domain	DEXDc	DEXDc	-	-	DEXDc	Helicase C-terminal	KH
Predicted effect	Uncertain	Uncertain	Loss-of-function	Loss-of-function	Uncertain	Uncertain	Uncertain
Inheritance	Maternal	Maternal	Maternal	Maternal	Maternal (potentially mosaic, in 4% of 52 reads)	Maternal	Maternal
Other variants (pathogenic, likely pathogenic, or variants of uncertain significance)	None	None	None	None	<i>MAT1A</i> (NM_000429.2): c.529C>T, p.(Arg177Trp), homozygous – likely pathogenic	<i>RAD54L2</i> (NM_015106.4): c.2303_2304dup, p.(E769Pfs*18), <i>de novo</i> (mosaic 11%) – variant of uncertain significance	None
<b>Pregnancy and Growth</b>							
Abnormalities during pregnancy/delivery; feeding difficulties/FTT	None	None	None	None	None	None	None
<b>Developmental history</b>							
Developmental delay	No	No	Yes, mild	Yes, mild	Yes, mild	Yes, mild (plateaued)	Yes, mild (plateaued)
Intellectual disability	No	No	No	No	No	No	Yes, mild to moderate
Motor delay	No	No	Yes, mild	Yes, mild	No	Yes, mild	Yes, mild
Speech delay	No	No	Yes, mild	Yes, moderate	Yes, severe	Yes, moderate	Yes, mild
<b>ASD</b>							
ASD features	Limited communication skills, poor social interaction	Asperger syndrome spectrum	Limited communication skills, self-stimulating behavior	Limited communication and social skills, weak visual problem-solving skills	Severely impaired communication skills; weak visual problem-solving skills; sensory hyperresponsiveness	Impaired communication skills, echolalia, elopement, stimming	Echolalia, sensory stimming, repetitive play/water play, sensory hyperresponsiveness
Eye contact	Avoidance	Avoidance	Avoidance	Avoidance	Avoidance	Avoidance	Preserved
Stereotypies	No	No	Yes (motor)	No	Yes (verbal)	Yes (behavioral, verbal)	Yes (behavioral, motor)
<b>Neuropsychiatric features</b>							

Seizures	Yes	No	No	No	No	Yes	Yes (based on abnormal EEG; diagnosed with ESES)
Sleep disturbances	No	No	Yes	Yes	Yes	Yes	Yes
Other neurobehavioral features	None	None	ADHD	None	Hyperactivity	ADHD, anxiety, aggressivity	Hyperactivity, anxiety

Abbreviations: ASD = autism spectrum disorder; ES = exome sequencing; WGS = whole genome sequencing; ESES = electrical status epilepticus in sleep; F = female; FTT = failure to thrive; M = male.

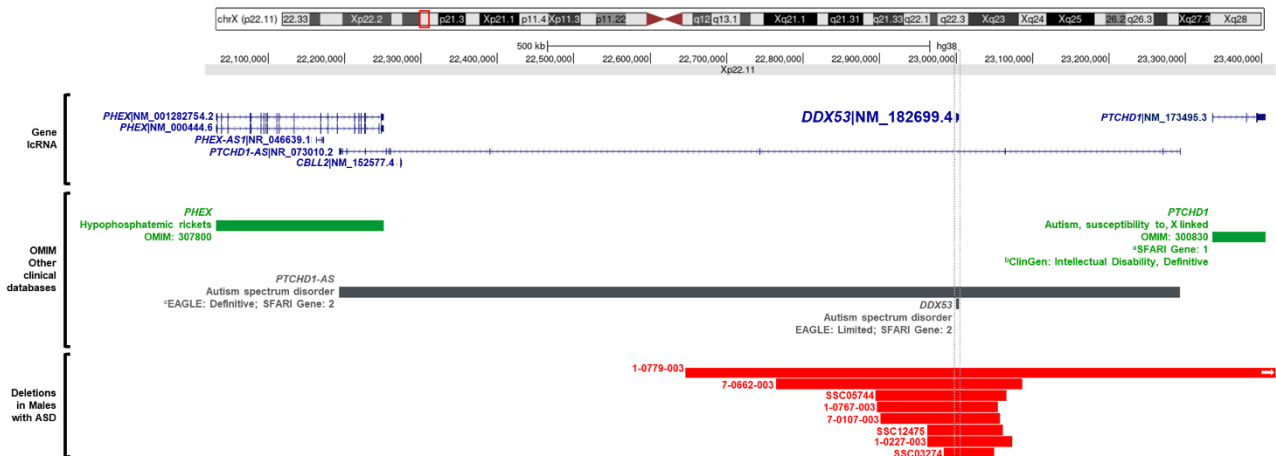
**Table 2. Genetic and clinical features of ASD participants with *DDX53* variants in the MSSNG and SFARI databases.**

Source	ID#	Sex	Variant (NM_182699.3); GRCh38	Inheritance	Consequence	AF (gnomAD); Decipher	Hemizygous in gnomAD	<i>In silico</i> pathogenic impact (deleterious vs benign)*	Predicted effect	Affected <i>DDX53</i> domain	Phenotype	Other pathogenic variants
SPARK	SP0003053	F	c.24G>A, p.(Trp8*)	Unknown	Stopgain	0.0000182	0	High (4 vs 1)	Loss-of-function	-	ASD	-
MSSNG	AU2440301	M	c.149C>A, p.(Pro50Gln)	Maternal	Missense	0.00000892	0	Medium (7 vs 11)	Uncertain	KH	ASD	-
SPARK	SP0039198†	M	c.567_568insTGCAGGT, p.(Glu190Cysfs*3)	Maternal	Frameshift	0	0	High (no benign)	Loss-of-function	-	ASD	-
SPARK	SP0042895†	F	c.567_568insTGCAGGT, p.(Glu190Cysfs*3)	Maternal	Frameshift	0	0	High (no benign)	Loss-of-function	-	ASD	-
SPARK	SP0035110	M	c.635A>G, p.(Glu212Gly)	Maternal	Missense	0	0	Medium (8 vs 11)	Uncertain	-	ASD	-
SPARK	SP0137474	M	c.812G>A, p.(Gly271Glu)	Maternal	Missense	0.00000894	0	High (16 vs 3)	Uncertain	DEXDc	ASD	-
SPARK	SP0032790	M	c.839G>C, p.(Gly280Ala)	Maternal	Missense	0.0000277	0	Moderate (8 vs 10)	Uncertain	DEXDc	ASD	-
SPARK	SP0115897	M	c.844A>T, p.(Ile282Phe)	Unknown	Missense	0.0000563	0	Medium (10 vs 9)	Uncertain	DEXDc	ASD	-
SPARK	SP0241833	M	c.1043T>C, p.(Ile348Thr)	Maternal	Missense	0	0	High (14 vs 5)	Uncertain	DEXDc	ASD	-
SPARK	SP0030523	M	c.1073A>G, p.(Asp358Gly)	Maternal	Missense	0	0	High (12 vs 7)	Uncertain	DEXDc	ASD	-
SPARK	SP0005644	M	c.1132G>A, p.(Ala378Thr)	Unknown	Missense	0	0	High (12 vs 6)	Uncertain	DEXDc	ASD	-
SPARK	SP0052296	M	c.1303C>A, p.(Leu435Ile)	Unknown	Missense	0	0	Medium (8 vs 11)	Uncertain	DEXDc	ASD	-
SSC	SSC10709	M	c.1334T>C, p.(Ile445Thr)	Maternal	Missense	0.00000892	0	Medium (9 vs 9)	Uncertain	-	ASD	-
SPARK	SP0038972	M	c.1342A>C, p.(Thr448Pro)	Maternal	Missense	0	0	Moderate (5 vs 13)	Uncertain	-	ASD	-
SSC	SSC10563	M	c.1672A>G, p.(Ile558Val)	Maternal	Missense	0	0	Medium (7 vs 12)	Uncertain	HELICc	ASD	-
SPARK	SP0126624	M	c.1724A>G, p.(Asp575Gly)	Maternal	Missense	0.0000109	0	Medium (7 vs 11)	Uncertain	C-terminal	ASD	-
MSSNG	1-0551-003	M	c.1724A>G, p.(Asp575Gly)	Maternal	Missense	0.0000109	0	Medium (7 vs 11)	Uncertain	C-terminal	ASD	-
SPARK	SP0090623‡	M	c.1736C>T, p.(Ala579Val)	Unknown	Missense	0.0000269	0	Medium (6 vs 12)	Uncertain	C-terminal	ASD	-
SPARK	SP0066816	M	c.1771_1773del, p.(Gln591del)	Unknown	Inframe deletion	0	0	High (no benign)	Uncertain	C-terminal	ASD	-
SPARK	SP0066817	M	c.1771_1773del, p.(Gln591del)	Unknown	Inframe deletion	0	0	High (no benign)	Uncertain	C-terminal	ASD	-
MSSNG	MSSNG00067-003	M	c.1143_1144delGCinsTA, p.(Met381_Leu382delinsIleMet)	Maternal	Inframe indel	0	0	Medium (1 vs 0)	Uncertain	DEXDc	ASD	<i>POGZ</i> (NM_015100.4) c.392dup, p.(Met132HisfsTer9), <i>de novo</i>
MSSNG	1-0485-003	F	Xp22.11 (22987677-23009507)x1	Maternal	Deletion (21.8kb)	Absent in Decipher	NA	NA	Loss-of-function	-	ASD	chr7:70441029-70771836, deletion (including <i>AUTS2</i> ), <i>de novo</i>
MSSNG	7-0111-003	M	Xp22.11 (23002082-23002458)x1	Unknown	Deletion within the 3'UTR	Absent in Decipher	NA	NA	Uncertain	NA	ASD	-
MSSNG	1-0627-003	M	Xp22.11 (23002148-23002522)x1	Maternal	Deletion within the 3'UTR	Absent in Decipher	NA	NA	Uncertain	NA	ASD, ADHD	-
MSSNG	1-0627-005	M	Xp22.11 (23002148-23002522)x1	Maternal	Deletion within the 3'UTR	Absent in Decipher	NA	NA	Uncertain	NA	ASD, ADHD	-

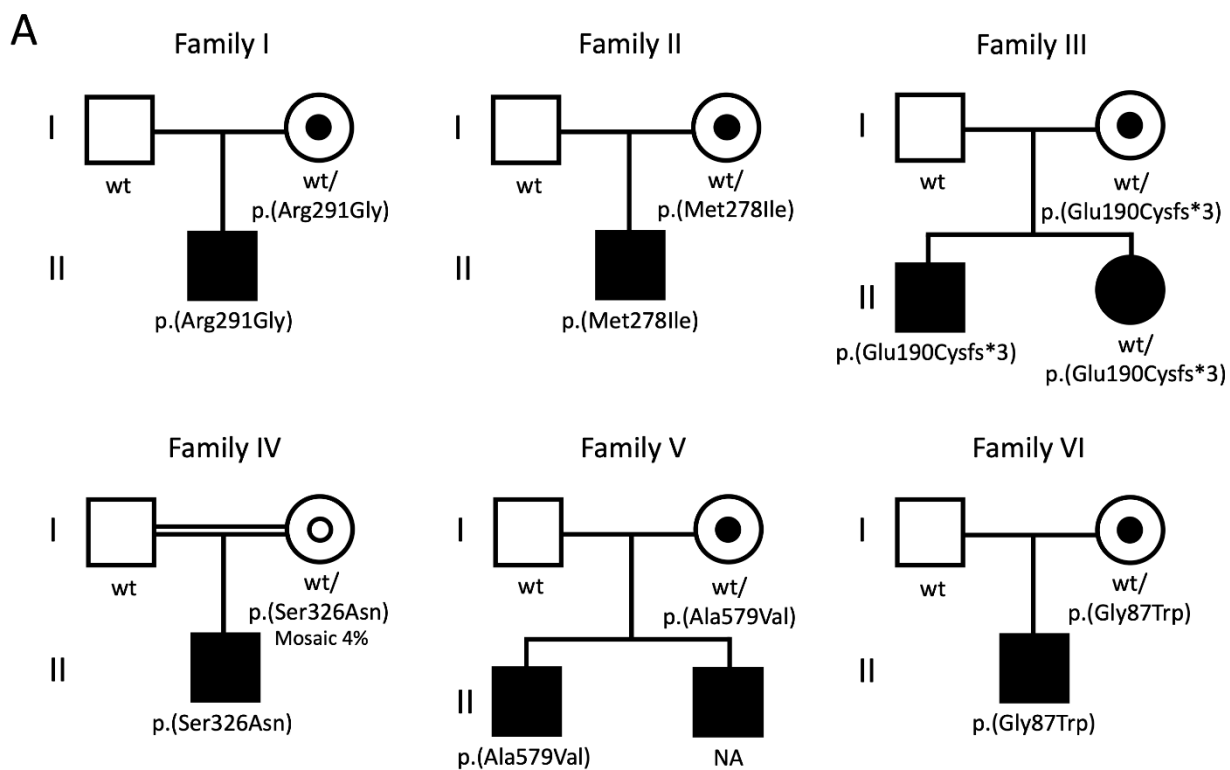
MSSNG	<b>1-0627-006</b>	M	Xp22.11 (23002148-23002522)x1	Maternal	Deletion within the 3'UTR	Absent in Decipher	NA	NA	Uncertain	NA	ASD, ADHD	-
MSSNG	<b>1-0627-007</b>	M	Xp22.11 (23002148-23002522)x1	Maternal	Deletion within the 3'UTR	Absent in Decipher	NA	NA	Uncertain	NA	ASD, ADHD	chr16:83861601-85760600, deletion, <i>de novo</i>

Abbreviations: ADHD = attention-deficit/hyperactivity disorder; ASD = autism spectrum disorder; F = female; M = male; NA = not available. Genome Build GRCh38. \* According to *in silico* tools predictions available in Varsome (<https://varsome.com>) † These patients correspond to subjects #3 (SP0039198) and #4 (SP0042895) from the main cohort, respectively. ‡ This patient (SP0090623) corresponds to subject #6 from the main cohort. ACMG guidelines could not be applied to determine pathogenicity of DDX53 variants as these should only be used for genes with established gene disease/disorder associations and not novel/potential disease/disorder genes.

Figure 1:



**Figure 2:**



**B**

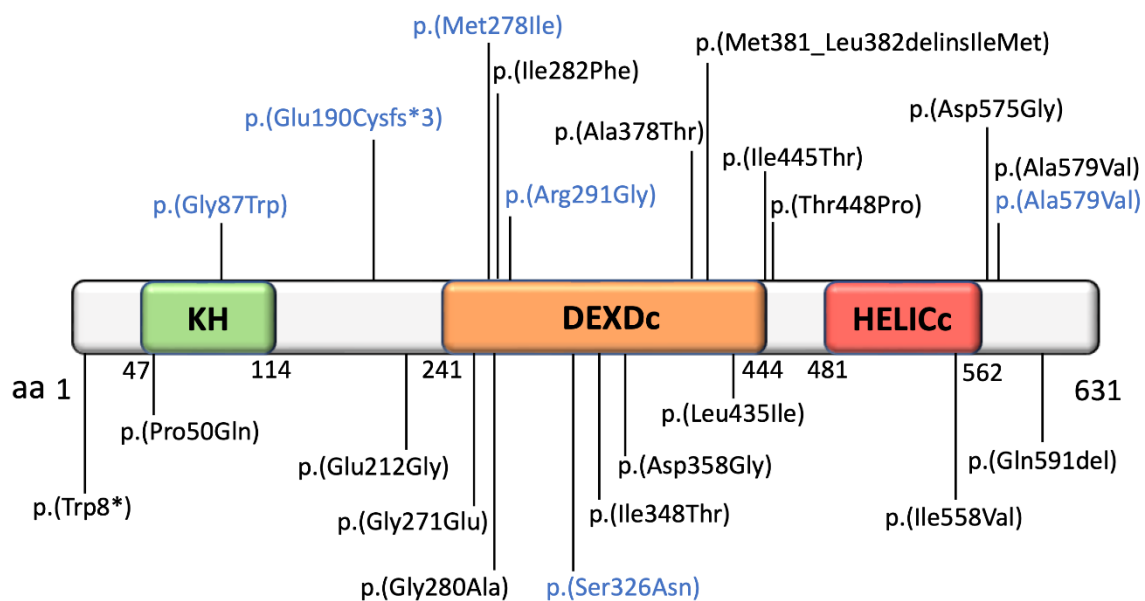
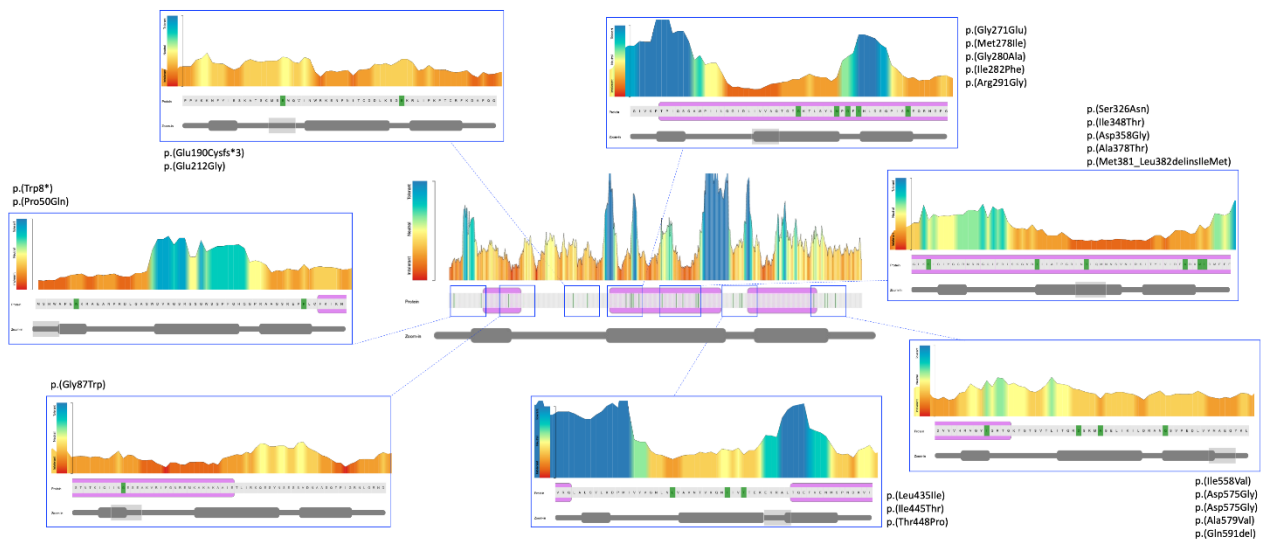


Figure 3:



**Figure 4**

