



# The Relevance of Variants With Unknown Significance for Autism Spectrum Disorder Considering the Genotype–Phenotype Interrelationship

Diogo V. Lovato<sup>1</sup>, Roberto R. Herai<sup>2,3</sup>, Graciela C. Pignatari<sup>1</sup> and Patricia C.B. Beltrão-Braga<sup>1,4\*</sup>

<sup>1</sup>Laboratory of Disease Modeling, Department of Microbiology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil, <sup>2</sup>Experimental Multiuser Laboratory, Graduate Program in Health Sciences, School of Medicine, Pontifícia Universidade Católica do Paraná, Curitiba, Brazil, <sup>3</sup>Lico Kaesemodel Institute (ILK), Curitiba, Brazil, <sup>4</sup>Laboratory of Disease Modeling, Scientific Platform Pasteur-USP, São Paulo, Brazil

## OPEN ACCESS

### Edited by:

Richard Eugene Frye,  
Phoenix Children's Hospital,  
United States

### Reviewed by:

Jinchen Li,  
Central South University, China  
Preeti Jacob,  
National Institute of Mental Health  
and Neurosciences, India

### \*Correspondence:

Patricia C.B. Beltrão-Braga  
patriciacbbbraga@usp.br

### Specialty section:

This article was submitted to  
Child and Adolescent Psychiatry,  
a section of the journal  
Frontiers in Psychiatry

**Received:** 03 March 2019

**Accepted:** 23 May 2019

**Published:** 07 June 2019

### Citation:

Lovato DV, Herai RR, Pignatari GC  
and Beltrão-Braga PCB (2019)  
The Relevance of Variants With  
Unknown Significance for Autism  
Spectrum Disorder Considering the  
Genotype-Phenotype Interrelationship.  
*Front. Psychiatry* 10:409.  
doi: 10.3389/fpsy.2019.00409

Several efforts in basic and clinical research have been contributing to unveiling the genetics behind autism spectrum disorders (ASD). However, despite these advancements, many individuals diagnosed with ASD and related neuropsychiatric conditions have been genetically investigated without elucidative results. The enormous genetic complexity of ASD-related conditions makes it a significant challenge to achieve, with a growing number of genes (close to a thousand) involved, belonging to different molecular pathways and presenting distinct genetic variations. Next-generation sequencing (NGS) is the approach most used in genetic research related to ASD, identifying *de novo* mutation, which is closely related to more severe clinical phenotypes, especially when they affect constrained and loss-of-function intolerant genes. On the other hand, *de novo* mutation findings contribute to a small percentage of the ASD population, since most of the cases and genetic variants associated with neuropsychiatric conditions are inherited and phenotypes are results of additive polygenic models, which makes statistical efforts more difficult. As a result, NGS investigation can sound vainly or unsuccessful, and new mutations on genes already related with ASD are classified as variants of unknown significance (VUS), hampering their endorsement to a clinical phenotype. This review is focused on currently available strategies to clarify the impact of VUS and to describe the efforts to identify more pieces of evidence throughout clinical interpretation and genetic curation process.

**Keywords:** autism spectrum disorders, variants of unknown significance, next-generation sequencing, *de novo* mutation, rare variants

## INTRODUCTION

Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder (ND), usually identified during early childhood according to clinical criteria, comprising impairment of the social abilities and cognitive competence, as well as typical behavior including stereotypic movements. Despite the criteria used to categorize someone under ASD's umbrella, it is well known that ASD is not homogeneous in terms of clinical signs and genetic aspect, being close to having one thousand

genes already implicated (1, 2). Considering that, it is quite evident that not all of the ASD subtypes had the same origin, and depending on the genetic background of the person and environmental factors involved (3, 4), clinical phenotype could be extremely different and could vary significantly, especially in terms of severity (5, 6).

Based on the challenging efforts to better characterize the ASD genetic background, several studies concentrate their energies to investigate and correlate the genetic profile of different ASD individuals. These investigations allowed the determination that only a small proportion of ASD cases have a clear-signed mendelian genetic background, while most of the current cases are caused by a complex combination of polygenic additive effect, comprising an enigma in terms of how these genetic variations are combined to manifest the autistic phenotypes (7, 8). Due to the genotypic diversity and enormous possibility of genetic variations among individuals within the spectrum, it is challenging to identify and correlate specific genetic variation for each and its relationship with ASD etiology, not only considering statistical issues but also how genetic mutations interfere in protein function and molecular pathways during early developmental stages (9).

Current investigations, both in clinical routines and research, also involve the use of next-generation sequencing (NGS) technologies to extract and sequence the entire genome (whole genome sequencing—WGS) or exome (whole exome sequencing—WES) of ASD individuals. These approaches usually imply in the identification of several genetic modifications described as variants of unknown significance (VUS), which effectively represent mutations where the pathogenicity and the function of the gene involved is unclear, turning it into a critical issue since most patients remain without a genetic explanation for their condition (10, 11). Thus, these VUS are genetic variants which are genes already associated with ASD, but the specific altered regions do not have sufficient clinical evidence or functional shreds of evidence in order to be categorized in a pathogenic or benign variant. In other situations, some VUS seems to be relevant for clinical phenotype, but they are located at genes with few associations with ASD (12, 13). Even though in some cases researches could probably correlate some VUS and suggest pathogenicity, this is barely possible, claiming for a robust collection of evidence and data from the literature to support it, coming from research groups and laboratories around the world. This review presents the most recent knowledge regarding different strategies to obtain information from current VUS, during NGS curation and interpretation process, from ASD and related neuropsychiatric cases.

## GENETIC BEHIND ASD

Most ASD and associated developmental delay cases are a genetical combination of inherited common and rare genetic variants. On the other hand, non-inherited genetic mutations are the most clinically relevant alterations with higher morbidity when affecting constrained genes associated with neurodevelopmental biology (14–18).

During NGS analysis and interpretation, it is essential to consider all genes and molecular pathways involved in the early stages of neurodevelopment. As demonstrated in several groups, genes expressed during neurogenesis and involved pathways are the most critical genes affected in ASD etiology. Particular examples of genes that are well established to be associated with synaptic function include GRIN2B (glutamate receptor ionotropic NMDA type 2B) and SHANK3 (SH3 And Multiple Ankyrin Repeat Domains 3), and associated with chromatin regulation include CHD2 (Chromodomain Helicase DNA Binding Protein 2) and CHD8 (Chromodomain Helicase DNA Binding Protein 8) (19–21). These genes have been significantly studied and have had hundreds of genetic variants described. Even though, like in other ASD associated genes, the relevance of each VUS should be investigated and enlightened considering patient's phenotype, which would be very useful to proceed with research, clinical assessments, and genetic counseling. Additionally, the systematic identification of VUS could contribute to elucidate other NDs, including intellectual disabilities (IDs) and attention-deficit hyperactivity disorder (ADHD), schizophrenia, mood disorders and others, as well as ASD comorbidities, which could be related to genes and pathways that are strictly connected as a broader cognitive spectrum (22, 23). As many ASD associated genes are essential not only for nervous system development but also for other biological processes, many health conditions could escort ASD cases. Current genetic knowledge demonstrates the relevant overlap between several genes associated with ASD and some diseases (17, 18). For example, several genetic variants in voltage-gated ion channels associated with ASD also cause diseases, such as epileptic encephalopathies, different kinds of ataxias, heart diseases, and psychiatric disorders (24). In the genetic laboratory routine, the identification of VUS at such genes could significantly answer some clinical phenotypes when other health conditions are part of the patient's clinical phenotype.

To understand genetic differences and similarities found in ASD subcategories combined at DSM-5 (*The Diagnostic and Statistical Manual of Mental Disorders*), patients who have been previously diagnosed with autistic disorder, pervasive developmental disorder not otherwise specified (PDD-NOS), and Asperger's syndrome were genetically analyzed for *de novo* mutations (DNMs). Even when only considering the most damaging sort of genetic variants, the results demonstrated that DSM-4 ASD subcategories share many more similarities than differences, considering genes, gene expression, and molecular pathways (25). Similar data, regarding the genetic differences observed between ASD subcategories from inherited genetic variants, should be considered to train predictive tolls concerning clinical features, severity, and outcome in ASD as more genotyping–phenotyping data is being produced.

## TOOLS FOR UNVEILING THE GENETIC OF AUTISM SPECTRUM DISORDER

As most clinical cases are associated with a broad spectrum of developmental delay and autistic features, it is also critical to apply a comprehensive genetic analysis approach to increase

performance and diagnostic yield. One of the main reasons for the unfeasibility of the genetic panels for ASD is because of the variety of genetic profiles and phenotype diversity in the spectrum, contributing to the popularization of WES and WGS (26). WGS is an ideal option to analyze ASD genotypical variety and complexity, not only due to the possibility of investigating the important regulatory intergenic region (27, 28), but also because of improved sequence coverage comparing to WES low coverage regions affecting essential ASD-related genes (16, 29, 30). Besides, WGS becomes more elucidative if realized in parallel with WGS of the proband's parents. Despite some criticism regarding cost-effectiveness of the utilization of NGS as an auxiliary tool to support the diagnostic of ASD and neuropsychiatric conditions, recent investigations have demonstrated that even when there are negative or unknown results associated with WES, patients' clinical trajectories are significantly better, further treatment and healthcare costs are reduced even after unclear genetic investigation (31–35). As the NGS technology costs are significantly decreasing and genetic databases are improving as more patients are sequenced, the cost–benefits relationship is tending to become even better (36).

The genetic variants found are matched with disease-causing mutations from specific databases and are also analyzed using a set of distinct computational tools to predict potentially deleterious DNMs. Although the selection of optimal methods could accelerate the identification of deleterious variants, the absence of a gold-standard approach makes the task of investigating these variations a challenging process. Considering that, some methods rely on the analysis of the effects caused by the mutation in the protein structure, such as SNPmuSiC (37). The approach is a stability-driven knowledge-based classifier that uses protein structure, artificial neural network and solvent accessibility-dependent combinations of statistical potentials to predict whether destabilizing or stabilizing mutations are disease-causing.

Another approach, PolyPhen/PolyPhen-2 (Polymorphism Phenotyping v2), predicts the possible impact of amino acid substitutions on the stability and function of human proteins using structural and comparative evolutionary considerations by annotating the mutation using distinct databases. The tool combined all information and used a prediction method that employs machine-learning classification to estimate the probability of the missense mutation being damaging (38).

Other similar approaches were also developed to annotate and analyze variants, such as SnpEff, which predicts and classifies genetic variants accordingly, i.e., to a synonymous or non-synonymous amino acid replacement, start codon gains or losses, stop codon gains or losses, or frameshifts (39). Another tool, MutationTaster, investigates and scores functional consequences of amino acid substitutions, short insertion or deletion (indel) mutations, variants spanning intron–exon borders, intronic, and synonymous alterations (40, 41).

Another tool, SIFT (Sorting Intolerant From Tolerant), predicts whether an amino acid substitution affects protein function. The tool can distinguish between functionally neutral and deleterious amino acid changes, which was previously validated on mutagenesis studies and human polymorphisms (42).

Other computational tools like SIFT and PolyPhen are based on a combination of methods and are used in prioritizing changes that are likely to cause a loss of protein function. However, their low specificity requires further evidence to support or refute pathogenicity that should be sought before reporting novel missense changes (43).

Recently, 23 distinct computational methods were analyzed to evaluate the performance measures using independent benchmark datasets correlated with disease-causing genetic variants (44). The analysis demonstrated that some methods showed different performances under different conditions, and the specificities were lower than the sensitivities for most of them. It was found that the combination of the software REVEL (45) and VEST3 (46) (i.e., ReVe) showed the best overall performances with all others being the benchmark data, including DNMs.

Although several computational tools have been widely used to predict potentially deleterious DNMs, and the previous comparative analysis suggests that more rigorous analysis is necessary to distinguish pathogenic variants, and consequently can accelerate the better identification of deleterious variants.

## GENETIC DATABASES FOR AUTISM SPECTRUM DISORDER

The elucidation of the genetics behind ASD has been contributing significantly to phenotype elucidation. The last decade produced a precious amount of valuable information, which had to be systematized in ASD genetic research and public databases. There are three public databases specialized in genetic variants associated with ASD. SFARI is currently the main source for curated data categorizing associated risk genes with different scores and storing several genetic variants described in the scientific literature (1). AutDb (47) and AutismKB (48) also represent relevant data collected to help genetic curation in ASD. Each database uses different categories and rates evidence in diverse strategies to interpret the relevance of several genes. Considering all databases and recently published articles, there are close to a thousand genes already associated with ASD and related neuropsychiatric conditions.

Several other databases are used to support the classification and clinical impact of genetic variations found within an individual with ASD. The Human Gene Mutation Database (HGMD®) is composed of a set of germline mutations correlated with human inherited diseases (49). Another interest databank is denovo-db (<http://denovo-db.gs.washington.edu>), which provides standardization of annotation and improves accessibility for independent studies reported by the literature. This bank comprised more the 23 thousands of trios, and detailed information about variant information (chromosome location, change, type); detailed annotation at the transcript and protein levels; severity scores; frequency; validation status; and, most importantly, the phenotype of the individual with the variant (16). Another relevant resource is the databank, NPdenovo, composed by DNMs obtained from many studies on neuropsychiatric disorders, that have utilized massive trio-based WES and WGS (50). NPdenovo has 17,104 DNMs from

3,555 trios across four neuropsychiatric disorders: ASD, epileptic encephalopathy, intellectual disability, and schizophrenia, in addition to unaffected siblings (control) from 36 studies by WES/WGS. Another recent database for ASD research is the National Database for Autism Research (NDAR) that contains an extensive collection of clinical and behavioral assessments and health outcomes from novel interventions. NDAR has a global, unique patient identifier that can be linked to aggregated individual-level data, including genetic data, for hypothesis generation and testing, and for replicating research findings (51).

From a different perspective, dbMDEGA databank was created based on a meta-analysis investigation of brain gene expression profiles from reported human ASD expression datasets and knock-out mouse ASD model expression datasets. The databank allows accessing of the differentially expressed genes in the brain of individuals with ASD (52).

Other databanks in ASD have also showed their value when used in combination to help interpret specific mutations of individuals with ASD, such as the interactive autism network (IAN) (53), the Autism Treatment Network (ATN) (54), and PEDSnet (55). They have data from large cohorts of children with ASD but store different information. IAN has dedicated to saving patient-reported measures/phenotyping, ATN has data from clinical characterization, and PEDSnet has data from health care encounters and electronic medical record data. Although presenting strengths and weaknesses, these three databases can be better exploited when used together to interpret a genetic mutation in an individual with ASD (56).

## AUTISM SPECTRUM DISORDER GENETIC INTERPRETATION AND CURATION

Based on databases generated from several research groups, it is possible to incorporate gene expression data from brain transcriptomes of psychiatric patients to gene pathways associated with ASD, in order to produce more robust evidences and valuable information, to curate and interpret respective VUS in routine NGS analysis (9, 57–59).

Other critical features to analyze during gene curation and interpretation in ASD and neuropsychiatric conditions are genetic constraints, haploinsufficiency, evolutionary conservation and population frequency of genetic variants. Considering population frequency, when genetic variants are extremely rare or even never reported at current databases, such as gnomAD (The Genome Aggregation Database) (60), this is an important indication that might suggest pathogenicity in neurodevelopmental and psychiatric disorders, since low reproductive fitness is highly associated with these disorders. Different research groups demonstrated genetic intolerance to loss of function, such as RVIS (Residual Variation Intolerance Score) and pLI (Probability of Loss-of-function Intolerance), and sequence constraint are significantly associated with ASD and neuropsychiatric conditions like intellectual disability, ADHD, schizophrenia, and bipolar disorder (61–63). Not only genetic variants, related to loss of function intolerance, are significant in constrained genes, but also extremely rare variants never reported before in

any public database, which might suggest a similar effect as loss of function alterations. Additionally, a significant association between constrained and loss of function intolerant genes with more severe comorbidities was already demonstrated, being most of these genes described previously as ASD risk genes and sensible to DNM (26).

It is also important to mention that more severe clinical conditions and comorbidities tend to carry a higher genetic burden with several additional hits, including CNVs or rare inherited genetic variants, which are most of the time identified as VUS in neurotypical individuals, but in a significantly lower number (64). Another consistent strategy about genetic constraint is the identification of a specific constrained region in coding genome, not only in loss of function intolerant genes but conserved coding regions and protein domains critical to protein structure and function (65). All these strategies considering genetic constraint data could be combined to elucidate and distinguish VUS, with clinical relevance for ASD and related neuropsychiatric conditions during the genetic curation process.

Last, it is crucial to consider that even sophisticated strategies used to reclassify VUS are not enough for some inconclusive cases, due to multiple and complex genetic profiles recently discovered. Genetic alterations at noncoding regions correspond to a fraction of ASD patients with inconclusive test results. Recent researches identified the relevance of noncoding DNMs at promoter regions (27) and inherited cis-regulatory variants (28), affecting ASD risk genes, which are loss of function intolerant ( $pLI > 0.9$ ). Unfortunately, the identification of such variants is still far from clinical routine, since it requires WGS with high genome coverage and sophisticated software analysis. Other challenge for clinical routine and relevant genetic events associated with ASD are postzygotic mosaic DNMs, which represent 5% to 20% of total DNMs according to recent data (66–68).

## CONCLUSIONS

Considering all these pieces of evidence, they corroborate with the idea that genetic variants associated with neurodevelopmental and psychiatric conditions have a significant impact on reproductive fitness, not only due to rarity and constraint, but also when multiple hits are observed (7). ASD is a complex ND counting with 70% to 90% of heritability rating. In the genetic interpretation process for Mendelian conditions, keeping a focus on the combination of phenotyping and genetic curation could improve diagnostics and provide more relevant results, which will be returned to clinic practitioners (69). A large number of ASD affected individuals tend to have not only extremely rare variants in ASD associated genes, but additionally, some patients with severe phenotypes could have multiple hits accumulating and additively contributing to genetic burden (63, 64). Future investigations and development of novel approaches to identify and calculate such genetic burden, perhaps something similar to the polygenic risk score, but considering rare and ultra-rare variants, would be a breakthrough for the genetics of



ASD. For ASD and related neuropsychiatric conditions, the strategy of clinical phenotyping and collaboration during the curation process could provide even more information relevant to genetic complexity and could provide a strong background for patient stratification.

## REFERENCES

- Abrahams BS, Arking DE, Campbell DB, Mefford HC, Morrow EM, Weiss Lauren A, et al. Knowledgebase for the autism spectrum disorders (ASDs). *Mol Autism* (2013) 4(4):36. doi: 10.1186/2040-2392-4-36
- Bourgeron T. Current knowledge on the genetics of autism and propositions for future research. *C R Biol* (2016) 339(7–8):300–7. doi: 10.1016/j.crv.2016.05.004
- Tordjman S, Somogyi E, Coulon N, Keramarrec S, Cohen D, Bronsard G, et al. Gene × environment interactions in autism spectrum disorders: role of epigenetic mechanisms. *Front Psychiatry* (2014) 5:1–18. doi: 10.3389/fpsy.2014.00053
- Ng M, de Montigny J, Ofner M, Do M. Environmental factors associated with autism spectrum disorder: a scoping review for the years 2003–2013. *Health Promot Chronic Dis Prevent Can* (2017) 37(1):1–23. doi: 10.24095/hpcdp.37.1.01
- Kohane IS, Murphy S, Bickel J, Weber G, Kunkel L, McMurry A, et al. The co-morbidity burden of children and young adults with autism spectrum disorders. *PLoS ONE* (2012) 7(4):e33224. doi: 10.1371/journal.pone.0033224
- Alexeeff Stacey E, Qian Y, Croen Lisa A, Davis R, Davignon M, Crawford P, et al. Medical conditions in the first years of life associated with future diagnosis of ASD in children. *J Autism Dev Disord* (2017) 47(7):2067–79. doi: doi.org/10.1007/s10803-017-3130-4
- Crawford S. On the origins of autism: the quantitative threshold exposure hypothesis. *Med Hypotheses* (2015) 85(6):798–806. doi: 10.1016/j.mehy.2015.10.006
- Weiner DJ, Bader JS, Zwaigenbaum L, Iliadou B, Kolevzon A, Martsenkovsky I, et al. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nat Genet* (2017) 49(7):978–85. doi: 10.1038/ng.3863
- Chen S, Devlin B, Wang J, Roeder K, Fragoza R, Klei L, et al. An interactome perturbation framework prioritizes damaging missense mutations for developmental disorders. *Nat Genet* (2018) 50(7):1032–40. doi: 10.1038/s41588-018-0130-z
- Sue Richards NA, Bale S, Bick D, Das S, Gastier-Foster J. HHS Public Access. *Genet Med* (2015) 17(5):405–24. doi: 10.1038/gim.2015.30
- Tammimies K, Marshall CR, Walker S, Kaur G, Thiruvahindrapuram B, Lionel AC, et al. Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. *JAMA J Am Med Assoc* (2015) 314(9):959–903. doi: 10.1001/jama.2015.10078
- Bertier G, Sénécal K, Borry P, Vears DF. Unsolved challenges in pediatric Whole-Exome sequencing: a literature analysis. *Crit Rev Clin Lab Sci* (2017) 54(2):134–42. doi: 10.1080/10408363.2016.1275516
- Vears DF, Niemiec E, Howard HC, Borry P. Analysis of vus reporting, variant reinterpretation and recontact policies in clinical genomic sequencing consent forms. *Eur J Hum Genet* (2018) 26(12):1743–51. doi: 10.1038/s41431-018-0239-7
- Sanders SJ, Roeder K, Meyer KA, Murdoch JD, DiLullo NM, Šestan N, et al. De Novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* (2012) 485(7397):237–41. doi: 10.1038/nature10945
- Gaugler T, Sklar P, Gaugler T, Manaa D, Bodea CA, Sanders SJ, et al. Most genetic risk for autism resides with common variation. *Nat Genet* (2014) 46(8):881–5. doi: 10.1038/ng.3039
- Turner TN, Nelson J, Pennacchio LA, Darnell RB, Kronenberg ZN, Hoekzema K, et al. Genomic patterns of De Novo mutation in simplex autism. *Cell* (2017) 171(3):710–722.e12. doi: 10.1016/j.cell.2017.08.047
- Rubeis SD, Watson M, Holman DM, Fox KA, Guy Jr GP, Seidenberg AB, et al. Synaptic, transcriptional, and chromatin genes disrupted in autism. *Nature* (2014) 515(7526):209–15. doi: 10.1038/nature13772
- Iossifov I, Murtha MT, Sullivan CA, Leotta A, Troge J, Ye K, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* (2014) 515(7526):216–21. doi: 10.1038/nature13908
- Willsey AJ, Sanders SJ, Li M, Dong S, Tebbenkamp AT, Muhle RA, et al. XCoexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* (2013) 155(5):997. doi: 10.1016/j.cell.2013.10.020
- Parikshak NN, Luo R, Zhang A, Won H, Lowe JK, Chandran V, et al. Xintegrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* (2013) 155(5):1008. doi: 10.1016/j.cell.2013.10.031
- Sanders SJ, Daly MJ, Smith L, State MW, Klei L, Teran NA, et al. Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron* (2015) 87(6):1215–33. doi: 10.1016/j.neuron.2015.09.016
- Hollegaard MV, Sanders SJ, Skuse D, Daly MJ, Neale BM, Werge, et al. Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nat Genet* (2016) 48(5):552–5. doi: 10.1038/ng.3529
- Lecavalier L, McCracken CE, Aman MG, McDougle CJ, McCracken JT, Tierney E, et al. An exploration of concomitant psychiatric disorders in children with autism spectrum disorder. *Compr Psychiatry* (2019) 88:57–64. doi: 10.1016/j.comppsy.2018.10.012
- Daghni M, Rima M, Fajloun Z, Ronjat M, Brusés JL, M'rad R, et al. Autism throughout genetics: perusal of the implication of ion channels. *Brain Behav* (2018) 8(8):1–18. doi: 10.1002/brb3.978
- Li J, He X, Zhang Y, Sun Z, Xia K, Shi L, et al. A comparative study of the genetic components of three subcategories of autism spectrum disorder. *Mol Psychiatry* (2018). doi: 10.1038/s41380-018-0081-x
- David MM, Enard D, Ozturk A, Daniels J, Jung JY, Diaz-Beltran L, et al. Comorbidity analysis of genes associated with autism spectrum disorders reveals differential evolutionary constraints. *PLoS ONE* (2016) 11(7):1–15. doi: 10.1371/journal.pone.0157937
- An JY, Lin K, Zhu L, Werling DM, Dong S, Brand H, et al. Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder. *Science* (2018) 362(6420):1–8. doi: 10.1126/science.aat6576
- Bandler WM, Muotri AR, Venter JC, Wang Z, Kaur G, Sebat J, et al. Paternally inherited cis-regulatory structural variants are associated with autism. *Science* (2018) 360(6386):327–31. doi: 10.1126/science.aan2261
- Guo H, Duyzend MH, Coe BP, Baker C, Hoekzema K, Gerds J, et al. Genome sequencing identifies multiple deleterious variants in autism patients with more severe phenotypes. *Genet Med* (2018) 0(0):1–10. doi: 10.1038/s41436-018-0380-2
- Lionel AC, Sondheimer N, Silverman E, Schulze A, Carter MT, Bikangaga P, et al. Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet Med* (2017) 20(4):435–43. doi: 10.1038/gim.2017.119
- Valencia CA, Harley JB, Kaufman KM, Tan T, Valencia C, Husami A, et al. Clinical impact and cost-effectiveness of whole exome sequencing as a diagnostic tool: a pediatric center's experience. *Front Pediatr* (2015) 3:1–15. doi: 10.3389/fped.2015.00067
- Rossi M, Tang S, Farwell Hagman KD, Powis Z, Black MH, Rossi M, et al. Outcomes of diagnostic exome sequencing in patients with diagnosed or suspected autism spectrum disorders. *Pediatr Neurol* (2017) 70:34–43.e2. doi: 10.1016/j.pediatrneurol.2017.01.033
- Du X, Gao X, Wang L, Du X, Gong Z, Yu Y, et al. Genetic diagnostic evaluation of Trio-Based whole exome sequencing among children with

## AUTHOR CONTRIBUTIONS

We state that all authors have read the paper and approved the data and the conclusions presented therein, and all authors have contributed to the present manuscript with equal effort.

- diagnosed or suspected autism spectrum disorder. *Front Genet* (2018) 9:1–8. doi: 10.3389/fgene.2018.00594
34. Vrijenhoek T, van Gassen KLI, Geenen JW, van Amstel HKP, Frederix GWJ, Knoers NV, et al. Whole-exome sequencing in intellectual disability: cost before and after a diagnosis. *Eur J Hum Genet* (2018) 26(11):1566–71. doi: 10.1038/s41431-018-0203-6
  35. Yuen T, Carter MT, Szatmari P, Ungar WJ. Cost-effectiveness of genome and exome sequencing in children diagnosed with autism spectrum disorder. *Appl Health Econ Health Policy* (2018) 16(4):481–93. doi: 10.1007/s40258-018-0390-x
  36. Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). Available at: [www.genome.gov/sequencingcostsdata](http://www.genome.gov/sequencingcostsdata). Accessed [April 10, 2019]. <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>. Last updated: April 25, 2018
  37. Ancien F, Pucci F, Godfroid M, Rooman M. Prediction and interpretation of deleterious coding variants in terms of protein structural stability. *Sci Rep* (2018) 8(1):1–11. doi: 10.1038/s41598-018-22531-2
  38. Adzhubei I, Jordan M, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* (2013) 7.20.1-7.20.41. doi: 10.1002/0471142905.hg0720s76
  39. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* (2012) 6(2):80–92. doi: 10.4161/fly.19695
  40. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* (2010) 7(8):575–6. doi: 10.1038/nmeth0810-575
  41. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* (2014) 11(4):361–2. doi: 10.1038/nmeth.2890
  42. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* (2003) 31(13):3812–4. doi: 10.1093/nar/gkg509
  43. Flanagan SE, Patch AM, Ellard S. Using SIFT and polyphen to predict loss-of-function and gain-of-function mutations. *Genet Test Mol Biomarkers* (2010) 14(4):533–7. doi: 10.1089/gtmb.2010.0036
  44. Li J, Zhao T, Zhang Y, Zhang K, Shi L, Chen Y, et al. Performance evaluation of pathogenicity-computation methods for missense variants. *Nucleic Acids Res* (2018) 46(15):7793–804. doi: 10.1093/nar/gky678
  45. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet* (2016) 99(4):877–85. doi: 10.1016/j.ajhg.2016.08.016
  46. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R. Identifying mendelian disease genes with the variant effect scoring tool. *BMC Genomics* (2013) 14(Suppl 3):S3. doi: 10.1186/1471-2164-14-S3-S3
  47. Peraanu W, Larsen EC, Das I, Estévez MA, Sarkar AA, Spring-Pearson S, et al. AutDB: a platform to decode the genetic architecture of autism. *Nucleic Acids Res* (2018) 46(D1):D1049–54. doi: 10.1093/nar/gkx1093
  48. Yang C, Li J, Wu Q, Yang X, Yue Huang A, Zhang J, et al. AutismKB 2.0: a knowledgebase for the genetic evidence of autism spectrum disorder. *Database* (2018) 2018(2018):1–8. doi: 10.1093/database/bay106
  49. Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, et al. The human gene mutation database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet* (2017) 136(6):665–77. doi: 10.1007/s00439-017-1779-6
  50. Li J, Cai T, Jiang Y, Chen H, He X, Chen C, et al. Genes with de novo mutations are shared by four neuropsychiatric disorders discovered from npdenovo database. *Mol Psychiatry* (2016) 21(2):290–7. doi: 10.1038/mp.2015.40
  51. Payakachat N, Tilford JM, Ungar WJ. National database for autism research (NDAR): big data opportunities for health services research and health technology assessment. *Pharmacoeconomics* (2016) 34(2):127–38. doi: 10.1007/s40273-015-0331-6
  52. Zhang S, Deng L, Jia Q, Huang S, Gu J, Zhou F, et al. DbMDEGA: a database for meta-analysis of differentially expressed genes in autism spectrum disorder. *BMC Bioinformatics* (2017) 18(1):1–8. doi: 10.1186/s12859-017-1915-2
  53. Lee H, Marvin AR, Watson T, Piggot J, Kiely Law J, Law PA, et al. Accuracy of phenotyping of autistic children based on internet implemented parent report. *Am J Med Genet Part B: Neuropsychiatric Genet* (2010) 153(6):1119–26. doi: 10.1002/ajmg.b.31103
  54. Murray DS, Fedele A, Shui A, Coury DL. The autism speaks autism treatment network registry data: opportunities for investigators. *Pediatrics* (2016) 137(Supplement):S72–78. doi: 10.1542/peds.2015-2851E
  55. Forrest CB, Margolis P, Seid M, Colletti RB. PEDSnet: how a prototype pediatric learning health system is being expanded into a national network. *Health Aff* (2014) 33(7):1171–7. doi: 10.1377/hlthaff.2014.0127
  56. Kuhlthau KA, Bailey LC, Baer BL, Coury DL, Kiely Law J, Murray DS, et al. Large databases for pediatric research on children with autism spectrum disorder. *J Dev Behav Pediatr* (2018) 39(2):168–76. doi: 10.1097/DBP.0000000000000534
  57. Pinto D, Delaby E, Merico D, Barbosa M, Merikangas A, Klei L, et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet* (2014) 94(5):677–94. doi: 10.1016/j.ajhg.2014.03.018
  58. Ballouz S, Gillis J. Strength of functional signature correlates with effect size in autism. *Genome Med* (2017) 9(1):1–14. doi: 10.1186/s13073-017-0455-8
  59. Gandal MJ, Leppa V, Won H, Parikshak NN, Geschwind DH. The road to precision psychiatry: translating genetics into disease mechanisms. *Nat Neurosci* (2016) 19(11):1397–407. doi: 10.1038/nn.4409
  60. Lek M, Weisburd B, Ruano-Rubio V, Daly MJ, Moonshine AL, Rivas MA, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* (2016) 536(7616):285–91. doi: 10.1038/nature19057
  61. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet* (2013) 9(8). doi: 10.1371/journal.pgen.1003709
  62. Samocha KE, Gibbs RA, McGrath LM, Stevens C, Boerwinkle E, Rehnström K, et al. A framework for the interpretation of de novo mutation in human disease. *Nat Genet* (2014) 46(9):944–50. doi: 10.1038/ng.3050
  63. Ganna A, Satterstrom FK, Zekavat SM, Das I, Kurki MI, Churchhouse C, et al. Quantifying the impact of rare and ultra-rare coding variation across the phenotypic spectrum. *Am J Hum Genet* (2018) 102(6):1204–11. doi: 10.1016/j.ajhg.2018.05.002
  64. Pizzo L, Galesi O, Zeesman S, Martin-Coignard D, Van Dijk A, Huber E, et al. Rare variants in the genetic background modulate cognitive and developmental phenotypes in individuals carrying disease-associated variants. *Genet Med* (2018) 0(0):819–25. doi: 10.1038/s41436-018-0266-3
  65. Havrilla JM, Pedersen BS, Layer RM, Quinlan AR. A map of constrained coding regions in the human genome. *Nat Genet* (2019) 51(1):88–95. doi: 10.1038/s41588-018-0294-6
  66. Freed D, Pevsner J. The contribution of mosaic variants to autism spectrum disorder. *PLoS Genet* (2016) 12(9):1–20. doi: 10.1371/journal.pgen.1006245
  67. Krupp DR, Barnard RA, Duffourd Y, Evans SA, Mulqueen RM, Bernier R, et al. Exonic mosaic mutations contribute risk for autism spectrum disorder. *Am J Hum Genet* (2017) 101(3):369–90. doi: 10.1016/j.ajhg.2017.07.016
  68. Lim ET, Hill RS, Zhang X, Kolevzon A, Penzol MJ, Carracedo A, et al. Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. *Nat Neurosci* (2017) 20(9):1217–24. doi: 10.1038/nn.4598
  69. Son JH, Xie G, Yuan C, Ena L, Li Z, Goldstein A, et al. Deep phenotyping on electronic health records facilitates genetic diagnosis by clinical exomes. *Am J Hum Genet* (2018) 103(1):58–73. doi: 10.1016/j.ajhg.2018.05.010

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Lovato, Herai, Pignatari and Beltrão-Braga. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.