List of genes having strong statistical support for association to mental illnesses

Authored by NIMH Staff  
9/21/2023  
Version 1.0

Table of Contents

[1. Introduction 2](#_Toc140932570)

[2. Purpose of the gene list 3](#_Toc140932571)

[3. Rare variation 3](#_Toc140932572)

[4. Common variation 4](#_Toc140932573)

[5. Intersection of rare and common variation 6](#_Toc140932574)

[6. NIMH Program guidance for investigators 6](#_Toc140932575)

[6a. Genes on this list 6](#_Toc140932576)

[6b. Genes on the list: Bonferroni versus 5% FDR 6](#_Toc140932577)

[6c. Genes not on the list (candidate genes) 7](#_Toc140932578)

[6d. Genes not on this list (but should be added) 7](#_Toc140932579)

[6e. Genes identified by techniques other than RVAS and GWAS 8](#_Toc140932580)

[6f. Genes of medical relevance that lack genome-wide support 8](#_Toc140932581)

[6g. Genes in copy number variant (CNV) or structural variation (SV) regions 9](#_Toc140932582)

[6h. Paralogs 9](#_Toc140932583)

[6i. Orthologs 9](#_Toc140932584)

[7. How the RVAS gene list was made 9](#_Toc140932585)

[8. Errors and omissions in the gene list 10](#_Toc140932586)

[9. Updates to the list 10](#_Toc140932587)

[10. Contact information 11](#_Toc140932588)

[11. References (this document) 11](#_Toc140932589)

[12. References (RVAS gene list) 15](#_Toc140932590)

# Introduction

The purpose of this document to provide guidance to investigators who are interested in applying to NIMH for funding of studies that involve particular genes. This document is accompanied by a list of human genes having statistically significant association to mental health traits.

We emphasize the following:

* This gene list includes examples of genes that have strong statistical support for association to mental illness.
* The emphasis of the list is on rare variation rather than common variation.
* The list is *not* restrictive. NIMH may support applications proposing to study genes that are not on this list, either because of basic science interest or with the inclusion of appropriate genome-wide evidence cited in the application.
* We encourage dialogue with the community about genes that investigators propose to study. We can be reached via NIMHgenes@nih.gov.

In 2018, the NIMH convened a Report of the National Advisory Mental Health Council Workgroup on Genomics. In previous years many researchers studied a favorite candidate gene, convinced of its importance in some mental health trait. The report emphasizes that NIMH should support research into genes that have appropriate, genome-wide evidence for association with a trait, whether the variation is rare or common. The [report](https://www.nimh.nih.gov/about/advisory-boards-and-groups/namhc/reports/report-of-the-national-advisory-mental-health-council-workgroup-on-genomics) and a [summary of recommendations](https://www.nimh.nih.gov/about/advisory-boards-and-groups/namhc/reports/namhc-genomics-workgroup-research-recommendations-summary) are available at the NIMH website.

This gene list was created by NIMH Program staff for the purpose of facilitating adherence to Council Workgroup recommendations. It is a representative (but not comprehensive) list of genes that do have genome-wide statistical support.

Examples of mental health illnesses for which NIMH provides funding support to investigators include anorexia nervosa (AN), attention deficit-hyperactivity disorder (ADHD), anxiety disorder, autism spectrum disorder (ASD), bipolar disorder (BPD), eating disorders, externalizing disorders (impulsive, disruptive conduct, substance use, and other addictive symptoms), intellectual disability (ID), internalizing disorders (anxiety, depression, and somatic symptoms), major depressive disorder (MDD), mood disorders, neurodevelopmental disorders (NDDs), obsessive-compulsive disorder (OCD), post-traumatic stress disorder (PTSD), schizophrenia (SCZ), suicidal thoughts and behaviors, and Tourette disorder (TD). NDDs are sometimes called neurodevelopmental psychiatric disorders (NPDs) involving a broad spectrum of psychiatric, behavioral, and cognitive symptoms (Jacquemont et al., 2022). NDDs (and NPDs) may encompass ASD and/or ID.

# Purpose of the gene list

(1) This gene list is used by NIMH Program staff as part of the process of evaluating grant applications.

(2) This gene list may be used by the biomedical research community

* to better understand NIMH priorities for genomics research
* to determine if particular genes have appropriate statistical support
* to study the properties of those genes having statistical support

If an investigator is interested in studying genes that lack rigorous genome-wide significance, we consider them candidate genes. As such, we might be interested in supporting their study from a basic science perspective, but not as genes relevant specifically to the trait of interest. We acknowledge that our rare variant gene list is not comprehensive.

There may be legitimate disagreement about what constitutes appropriate definitions of statistical significance. The gene list provides guidance but does not have a definitive role in decision-making about NIMH support for grant applications. NIMH Program welcomes feedback from investigators who are interested in studying genes that are not on the list.

Some gene lists have been prepared (by academic investigators, companies, and/or foundations) that report high confidence genes for a mental health condition. NIMH prioritizes support for those genes on such lists that have genome-wide statistical support for association to a trait or disease. This NIMH gene list can be compared to other lists in order to identify which are well supported.

Section 6 (below) provides more detailed guidance to investigators regarding the inclusion of genes in grant applications.

# Rare variation

Genetic variants may be classified as common or rare based on allele frequencies (with rare variants often defined as having a minor allele frequency [MAF] <1%). Two broad classes of experiment are rare variant association studies (RVAS), which seek association of a trait to rare variants, and genome-wide association studies (GWAS), which seek association of a trait to common variants.

For RVAS, each disease-associated variant is typically present only rarely, so it is routine for researchers to perform burden tests that aggregate risk across individual genes. Study design and statistical tests are discussed by Lee et al. (2014), Nicolae (2016), Zuk et al. (2014) with further extension to noncoding variants (Li et al., 2022). One tab of this spreadsheet reports RVAS results. Multiple test correction is applied to establish genome-wide significance, as reported in publications.

# Common variation

For GWAS, loci with significant association to a trait are reported and typically pass a threshold of p < 5x10-8 for genome-wide significance. A particular challenge for the understanding of GWAS results is that each locus may occur in a region of linkage disequilibrium (LD) that may contain dozens or hundreds of genes. One or more genes in an LD block may harbor a causal variant. A second challenge is sampling errors in test statistics used to identify causal genes.

There are resources that list genes possibly associated with mental health traits within genome-wide significant loci. For example, see the resource Open Targets Genetics (OTG; <https://genetics.opentargets.org/> ) (Ghoussaini et al., 2021; Mountjoy et al., 2021). For each genomic locus having genome-wide statistical support OTG lists the closest gene to that locus. Additionally, in some cases, OTG lists the most likely gene based on a “locus to gene” (L2G) statistical model. (The most likely gene and the closest gene are sometimes the same; in some cases, they differ; and in some cases no L2G gene is nominated.) Other examples of causal gene prioritization methods include combined SNP-to-gene (cS2G; Gazal et al., 2022), and BIobank-scale Gene-based association test via Knockoffs (BIGKnock; Ma et al., 2023). Additional GWAS resources are available, and other statistical approaches have been developed, sometimes producing a set of genes nominated from significant loci of GWAS studies.

The NIMH gene list does not include genes implicated from common variation.

* There is yet not a consensus in the field as to which specific genes from GWAS are relevant to the disease.
* Some researchers may choose to study genes implicated by GWAS.
* In this section of the documentation we list methods (such as fine-mapping) that may help provide evidence supporting the roles of specific genes in mental illness.
* We encourage investigators to discuss with NIMH program staff the study of genes implicated by common variants linked to mental illness.

A further series of approaches refine lists of significant GWAS hits by incorporating fine-mapping, transcriptomics (expression) data, and other functional data. Such methods routinely identify genes that are not the nearest annotated gene to the most statistically significant GWAS locus SNP. We describe the use of these methods with selected results from publications reporting GWAS for ADHD (Demontis et al., 2023), MDD (Levey et al., 2021), and SCZ (Trubetskoy et al., 2022). NIMH encourages such approaches, which include the following.

(1) Statistical fine-mapping is the process by which relevant genes (or other genomic elements) are identified within a locus. There is no broad consensus on which fine-mapped genes are causally associated with the trait or disease. In one approach, the gene closest to the locus having significant association is nominated as the causal gene. It is currently unknown how often the closest gene is a true positive.

* As an example, Trubetskoy et al. (2022) performed a GWAS of ~76,755 individuals with schizophrenia and 243,649 control individuals, identifying common variant associations at 287 distinct genomic loci. Fine-mapping with a Bayesian approach implemented in FINEMAP (Benner et al., 2016) identified a broad set of 628 genes containing at least one credible SNP, and several smaller sets of genes based on additional criteria such as those including non-synonymous or untranslated region (UTR) variants.

(2) Functional mapping and annotation of GWAS loci has been performed using methods such as FUMA (Watanabe et al., 2017). FUMA first annotates SNPs and maps them to genes based on positional, expression quantitative trait locus (eQTL) associations, and 3D chromatin interactions. It then analyzes differentially expressed genes, overrepresentation in gene sets, and provides other functional annotation.

* As an example, Demontis et al. (2023) performed a GWAS of ADHD, identified 27 genome-wide significant risk loci, and used FUMA to nominate 76 plausible risk genes.

(3) Transcriptome-wide association studies (TWAS) have been performed (Gusev et al., 2016; Wainberg et al., 2019). A predictive model of gene expression is trained using a reference panel such as GTEx. This model is used to predict gene expression for participants in a GWAS cohort, and association analysis is performed between the predicted expression and the trait.

* Demontis et al. (2023) used expression data from >900 samples from dorsolateral prefrontal cortex to perform TWAS and identified 15 genes. The investigators also identified 18 isoforms, yielding a total of 23 genes.
* Levey et al. (2021) performed TWAS using GWAS association statistics and eQTL data (from brain and blood using GTEx). They performed fine-mapping of 178 risk loci, annotated SNPs with relatively high causal posterior probability to identify those likely to be pathogenic, and then further identified 17 genes overlapping with significantly associated genes from TWAS.

(4) Gene expression effects have been assessed using summary-based Mendelian randomization (SMR; Zhu et al., 2016). GWAS summary-level data (such as effect sizes or test statistics) are integrated with eQTL data to identify genes associated with a complex trait because of pleiotropy (the same genetic variant is associated with both gene expression and complex trait). As with TWAS, associations detected with an SMR test may be due to linkage among nearby SNPs, so a heterogeneity test can be applied to help distinguish linkage from pleiotropy.

* Trubetskoy et al. (2022) used SMR to identify 101 genes (rejecting co-localizations due to linkage). They further prioritized these genes using three approaches: (1) n=32 of these genes were the only ones (or most likely ones) at the locus; (2) for n=16 genes the putatively causal eQTLs captures >50% of the FINEMAP posterior probability; (3) for n=29 genes there was supportive chromatin conformation (Hi-C) data. Combining all approaches (FINEMAP and SMR), n=120 genes (n=106 protein-coding) were prioritized.

(5) Multi-marker Analysis of GenoMic Annotation (MAGMA) has been used for gene and gene-set analyses (de Leeuw et al., 2015). MAGMA uses multiple regression to incorporate LD between markers and to identify multi-marker effects aggregated with nearby genes.

* Levey et al. (2021) reported 426 significantly associated genes, passing Bonferroni correction for 16,038 protein-coding genes, by using MAGMA to perform a genome-wide gene-based association study (GWGAS).

(6) Genomic structural equation modeling (genomic SEM; Grotzinger et al., 2019) is a multivariate method for analyzing the joint genetic architecture of complex traits. Genomic SEM has been used to model genetic associations among complex phenotypes and to identify variants that affect cross-trait liability. As an extension of this method, transcriptome-wide SEM (T-SEM; Grotzinger et al., 2022) is a multivariate method that estimates the effects of tissue-specific gene expression in applying multiple GWAS of traits or diseases. T-SEM also offers *Q*gene, a heterogeneity statistic that reports tissue-specific patterns of gene expression associated with a subset of traits or diseases. Measures such as these, which identify genes associated with traits and diseases and that pass Bonferroni corrections, are also considered well supported.

# Intersection of rare and common variation

It is of interest to identify genes that are implicated in diseases or traits through studies of both common and rare variation. For example, Trubetskoy et al. (2022) noted that a set of 32 genes, previously identified as harboring ultra-rare, deleterious mutations by the Schizophrenia Exome Sequencing Meta-Analysis (SCHEMA) consortium, were enriched for common variant associations. Two of these genes were prioritized from fine-mapping studies (*GRIN2A*, *SP4*). Demontis et al. (2023) similarly assessed overlap between common and rare variant studies in ADHD.

# NIMH Program guidance for investigators

## Genes on this list

NIMH encourages the study of genes that are present on this list (noting the caveats that the list is not comprehensive, and that it may contain false positive findings).

## Genes on the list: Bonferroni versus 5% FDR

All of the genes on the list are significant at the relatively lenient threshold of passing a 5% false discovery rate (FDR, with values obtained from publications). Some of the genes are also significant at the stringent threshold of passing a Bonferroni correction. Please note that any gene on the gene list may represent a false positive. Replication studies are particularly helpful to provide further evidence that genes are relevant to mental illness. In general, studies could suffer from potential artifacts (such as cases and controls coming from separate populations, or cases and controls sequenced using different methods, or other nuisance variables). Independent replication studies can help increase confidence in such findings.

For investigators who propose to focus their efforts on a single gene (or a small number of genes), the Bonferroni threshold is appropriate when derived from a genome-wide genotyping and/or sequencing assay. For investigators who propose to study a large number of genes (such as assessing the function of a group of dozens or hundreds of genes relevant to NDDs), selecting genes that pass a 5% FDR threshold is reasonable. All genes on the list have passed genome-wide or exome-wide thresholds of significance.

## Genes not on the list (candidate genes)

NIMH does not support the study of candidate genes that lack appropriate genome-wide statistical support for association to mental illness. The reason is that many such candidate genes are false positives. (See section 2 above.)

*Example*. An investigator proposes to study the role of the *GRIN2C* gene (encoding glutamate ionotropic receptor NMDA type subunit 2C) in ASD. That gene is not on the gene list. NIMH has a strong interest in promoting basic science research, and would potentially encourage applications that focus on the role of that glutamate receptor subunit. However, because genome-wide evidence for association of that gene with ASD is lacking, investigators should not propose to study that gene as a known risk factor for ASD. They may propose to study it based as a basic science project rather than asserting it is an ASD gene.

*Example*. The current gene list includes *SCN1A*, *SCN2A*, *SCN3A*, *SCN4A*, and *SCN8A* encoding a series of sodium voltage-gated channel subunits. Each of these has significant association to NDDs. An investigator proposes to study *SCN5A* that is not on the list. NIMH considers this a candidate gene of lower priority. An investigator would be encouraged to provide NIMH with substantive evidence supporting the importance of studying *SCN5A* in NDD (e.g., its known association to non-mental health phenotypes; or the known biochemical function of its encoded protein; or the importance of studying its function from a basic science perspective rather than as an established NDD gene).

*Example*. Investigators perform a GWAS and report that a locus close to the gene *FKBP5* on chromosome 6 comes close to statistical significance but does not meet genome-wide criteria. NIMH program might encourage its study from a basic science perspective but not in association with mental illness. If transcriptomics studies reveal that FKBP5 mRNA is significantly regulated in a condition (such as ASD, PTSD, or schizophrenia) this does not imply significant association to the trait.

## Genes not on this list (but should be added)

The list of genes involving rare variation is not comprehensive. If you identify a gene (or set of genes) that should be added please contact us at [NIMHgenes@nih.gov](mailto:NIMHgenes@nih.gov). If you submit a grant application that focuses on the study of any genes, NIMH will review whether there is evidence of statistically significant association to the mental health trait(s) proposed to be studied.

## Genes identified by techniques other than RVAS and GWAS

Many techniques other than GWAS and RVAS have been developed to identify genes that have a significant association to mental health illnesses and traits.

* TWAS (introduced in section 4 above) shows association between predicted expression and a trait. As for GWAS, TWAS may produce multiple associated genes at a single locus.
* Epigenome-wide association studies (EWAS) have been used to identify associations using CpG dinucleotide methylation data.
* Other methods use GWAS summary data, such as case-case GWAS (CC-GWAS; e.g., Peyrot and Price, 2021) and multi-trait analysis of GWAS (MTAG; Turley et al., 2018). CC-GWAS and MTAG sometimes identify genome-wide significant loci that are not significant in GWAS.
* Protein studies have been used to identify genes and proteins associated with disease.

The use of alternative methods may introduce ambiguities regarding genome-wide statistical significance. For example, some EWAS studies have employed a threshold p value cutoff of 10-4. NIMH Program will evaluate arguments for statistical significance on a case-by-case basis.

## Genes of medical relevance that lack genome-wide support

The medical genetics (or clinical genetics) field has convincingly implicated single genes in many diseases for which genome-wide significance has not necessarily been established. Examples of resources listing such genes and diseases include Online Mendelian Inheritance in Man ([OMIM](https://www.omim.org/)) and [ClinGen](https://clinicalgenome.org/).

*Example*. Likely gene-disrputing mutations in *CSDE1* (encoding an RNA-binding protein) were identified in 17 probands having ASD and ID (Guo et al., 2019) and one proband with a milder phenotype (Krenn et al., 2022). *CSDE1* is one of 94 high confidence neuropsychiatric disorder genes studied by Shimelis et al. (2023) in the DiscovEHR cohort, and it is a high confidence ASD gene on the Simons Foundation [SFARI gene list](https://gene.sfari.org/database/human-gene/CSDE1). However, *CSDE1* is absent from the NIMH gene list because there does not appear to be genome-wide evidence for its role in ASD, ID, or NDD. If an applicant proposes to study the function of *CSDE1*, NIMH Program will consider its role in a particular syndrome (because of convincing medical genetics evidence), or its role encoding an RNA-binding protein (from a basic science perspective). However, NIMH would likely not support the study of its role in the broader phenotype of ASD or ID (for which statistically significant association has not been established).

For *CSDE1* or other syndromic forms of ID, ASD, or mental illnesses, investigators are encouraged to contact NIMH Program to discuss potential overlap of assignment between other NIH Institutes and Centers such as NINDS and NICHD.

*Example*. An investigator proposes a high throughput functional assay of 100 ASD genes, and proposes to select them from the [Simons Foundation SFARI list](https://gene.sfari.org/) of high confidence ASD genes. NIMH Program would note that about half the >200 genes on that list have genome-wide evidence for significant association to ASD and/or NDD, and would encourage the investigator to select those well-supported genes for further study.

## Genes in copy number variant (CNV) or structural variation (SV) regions

CNV regions and other SV regions (inversions, translocations, amplifications, deletions) often occur at loci harboring multiple genes. One class of analyses tests whether a particular structural variant is associated with a trait (e.g., 16p11.2 deletions and ASD) with genome-wide significance. Another class of analyses tests whether individual genes within a structural variant have genome-wide significant association with the trait being studied. Both types of analysis should be performed. Evidence suggests that some CNVs have significant association with disease but none of the individual genes within them have significant genome-wide association. This does not preclude the study of such CNVs.

We note that the *NRXN1* gene on chromosome 2p16.3 has genome-wide significance both as a single gene and and as a CNV.

## Paralogs

For paralogs, each gene that is proposed to be studied should have genome-wide statistical evidence. For example, *SCN1A*, *SCN2A*, *SCN3A*, and *SCN4A* are paralogs encoding sodium voltage-gated channel subunits. All have significant evidence for association to NDD from RVAS. However, *SCN5A* does not. An investigator may still propose to study *SCN5A* but should justify why this is appropriate.

## Orthologs

For orthologs, please note the NIMH position on the use of animal models in “[Notice](https://grants.nih.gov/grants/guide/notice-files/NOT-MH-19-053.html) of NIMH’s Considerations Regarding the Use of Animal Neurobehavioral Approaches in Basic and Pre-clinical Studies.” This document discusses the study of human risk genes in animal models.

# How the RVAS gene list was made

7.1 We performed PubMed searches to publications reporting genes having genome-wide significant association to mental health illnesses (section 1 above). Data sources sometimes included archives ([bioRxiv](https://www.biorxiv.org/), [medRxiv](https://www.medrxiv.org/)) as indicated on the gene list annotations and in the references in sections 13 and 14 below.

7.2 For each study we downloaded relevant supplementary tables.

7.3 For RVAS we also searched online resources such as [SCHEMA](https://schema.broadinstitute.org/).

7.4 Data were wrangled (formatted consistently). We converted non-approved gene symbols and/or gene names to approved symbols and/or names following HUGO Gene Nomenclature Committee (HGNC) standards (<https://www.genenames.org/>).

As an example of symbol correction, in their list of 101 SMR genes, Trubetskoy et al. (2022) include three alias symbols (*EMB*, *IK*, *SF3B1*) for which approved symbols are added in the gene list (*XPO1*, *KCNN4*, and *SF3B2*, respectively). They also include 11 previous symbols for which the approved symbols are used here.

As another example of symbol correction, for a list of 76 genes in Supplementary Table 7 of Demontis et al. (2023) the following changes were made based on HGNC data:

| Reported symbol | Category | Approved symbol | Action |
| --- | --- | --- | --- |
| CYHR1 | Previous symbol | *ZFTRAF1* | Used approved symbol in gene list |
| FAM198A | Previous symbol | *GASK1A* | Used approved symbol in gene list |
| FAM203A | Previous symbol | *HGH1* | Used approved symbol in gene list |
| FAM212A | Previous symbol | *INKA1* | Used approved symbol in gene list |
| CTD-2330K9.3 | Unmatched | None | Excluded from gene list |
| RP11-6L6.2 | Unmatched | None | Excluded from gene list |

Trost et al. (2022) list the gene symbols *SRPR* for which the approved symbol is *SRPRA* and *SUV420H1* for which the approved symbol is *KMT5B*.

7.5 Genes were annotated. Studies of ASD, ID, NPD, and NDD were combined into the single category NDD.

# Errors and omissions in the gene list

This list consists of a sampling of genes reported in the literature. Several types of error may occur.

* False negatives due to literature omissions: a gene with genome-wide support is missing from the list. This may occur because we have not included relevant publications. Please contact us at [NIMHgenes@nih.gov](mailto:xxx@nih.gov) so that we may update the list.
* Challenges of complex statistical models may lead to differing interpretations of which genes have appropriate genome-wide evidence for support.
* False positives: a gene without appropriate genome-wide evidence appears on the list.
* Wrangling errors: a gene symbol may be ambiguous (e.g., *MST1* is both an approved symbol and an alias for a different gene, *STK4*) and the list may include the wrong gene.

# Updates to the list

We anticipate updates to the list as new studies are published and as additional data sources are identified. Each gene list is assigned a version and date.

# Contact information

For questions or comments, including suggestions for changes, please contact [NIMHgenes@nih.gov](mailto:NIMHgenes@nih.gov).

# References (this document)

Benner C, Spencer CC, Havulinna AS, Salomaa V, Ripatti S, Pirinen M. FINEMAP: efficient variable selection using summary data from genome-wide association studies. Bioinformatics. 2016 May 15;32(10):1493-501. doi: 10.1093/bioinformatics/btw018. Epub 2016 Jan 14. PMID: 26773131; PMCID: PMC4866522.

de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015 Apr 17;11(4):e1004219. doi: 10.1371/journal.pcbi.1004219. PMID: 25885710; PMCID: PMC4401657.

Gazal S, Weissbrod O, Hormozdiari F, Dey KK, Nasser J, Jagadeesh KA, Weiner DJ, Shi H, Fulco CP, O'Connor LJ, Pasaniuc B, Engreitz JM, Price AL. Combining SNP-to-gene linking strategies to identify disease genes and assess disease omnigenicity. Nat Genet. 2022 Jun;54(6):827-836. doi: 10.1038/s41588-022-01087-y. Epub 2022 Jun 6. PMID: 35668300; PMCID: PMC9894581.

Ghoussaini M, Mountjoy E, Carmona M, Peat G, Schmidt EM, Hercules A, Fumis L, Miranda A, Carvalho-Silva D, Buniello A, Burdett T, Hayhurst J, Baker J, Ferrer J, Gonzalez-Uriarte A, Jupp S, Karim MA, Koscielny G, Machlitt-Northen S, Malangone C, Pendlington ZM, Roncaglia P, Suveges D, Wright D, Vrousgou O, Papa E, Parkinson H, MacArthur JAL, Todd JA, Barrett JC, Schwartzentruber J, Hulcoop DG, Ochoa D, McDonagh EM, Dunham I. Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. Nucleic Acids Res. 2021 Jan 8;49(D1):D1311-D1320. doi: 10.1093/nar/gkaa840. PMID: 33045747; PMCID: PMC7778936.

Grotzinger AD, de la Fuente J, Davies G, Nivard MG, Tucker-Drob EM. Transcriptome-wide and stratified genomic structural equation modeling identify neurobiological pathways shared across diverse cognitive traits. Nat Commun. 2022 Oct 21;13(1):6280. doi: 10.1038/s41467-022-33724-9. PMID: 36271044; PMCID: PMC9586980.

Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, Ip HF, Marioni RE, McIntosh AM, Deary IJ, Koellinger PD, Harden KP, Nivard MG, Tucker-Drob EM. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav. 2019 May;3(5):513-525. doi: 10.1038/s41562-019-0566-x. Epub 2019 Apr 8. PMID: 30962613; PMCID: PMC6520146.

Guo H, Li Y, Shen L, Wang T, Jia X, Liu L, Xu T, Ou M, Hoekzema K, Wu H, Gillentine MA, Liu C, Ni H, Peng P, Zhao R, Zhang Y, Phornphutkul C, Stegmann APA, Prada CE, Hopkin RJ, Shieh JT, McWalter K, Monaghan KG, van Hasselt PM, van Gassen K, Bai T, Long M, Han L, Quan Y, Chen M, Zhang Y, Li K, Zhang Q, Tan J, Zhu T, Liu Y, Pang N, Peng J, Scott DA, Lalani SR, Azamian M, Mancini GMS, Adams DJ, Kvarnung M, Lindstrand A, Nordgren A, Pevsner J, Osei-Owusu IA, Romano C, Calabrese G, Galesi O, Gecz J, Haan E, Ranells J, Racobaldo M, Nordenskjold M, Madan-Khetarpal S, Sebastian J, Ball S, Zou X, Zhao J, Hu Z, Xia F, Liu P, Rosenfeld JA, de Vries BBA, Bernier RA, Xu ZD, Li H, Xie W, Hufnagel RB, Eichler EE, Xia K. Disruptive variants of CSDE1 associate with autism and interfere with neuronal development and synaptic transmission. Sci Adv. 2019 Sep 25;5(9):eaax2166. doi: 10.1126/sciadv.aax2166. PMID: 31579823; PMCID: PMC6760934.

Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, Jansen R, de Geus EJ, Boomsma DI, Wright FA, Sullivan PF, Nikkola E, Alvarez M, Civelek M, Lusis AJ, Lehtimäki T, Raitoharju E, Kähönen M, Seppälä I, Raitakari OT, Kuusisto J, Laakso M, Price AL, Pajukanta P, Pasaniuc B. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet. 2016 Mar;48(3):245-52. doi: 10.1038/ng.3506. Epub 2016 Feb 8. PMID: 26854917; PMCID: PMC4767558.

Jacquemont S, Huguet G, Klein M, Chawner SJRA, Donald KA, van den Bree MBM, Sebat J, Ledbetter DH, Constantino JN, Earl RK, McDonald-McGinn DM, van Amelsvoort T, Swillen A, O'Donnell-Luria AH, Glahn DC, Almasy L, Eichler EE, Scherer SW, Robinson E, Bassett AS, Martin CL, Finucane B, Vorstman JAS, Bearden CE, Gur RE; Genes to Mental Health Network. Genes To Mental Health (G2MH): A Framework to Map the Combined Effects of Rare and Common Variants on Dimensions of Cognition and Psychopathology. Am J Psychiatry. 2022 Mar;179(3):189-203. doi: 10.1176/appi.ajp.2021.21040432. PMID: 35236119; PMCID: PMC9345000.

Krenn M, Kepa S, Kasprian G, Riedhammer KM, Wagner M, Goedl-Fleischhacker U, Milenkovic I. A de novo truncating variant in CSDE1 in an adult-onset neuropsychiatric phenotype without intellectual disability. Eur J Med Genet. 2022 Mar;65(3):104423. doi: 10.1016/j.ejmg.2022.104423. Epub 2022 Jan 11. PMID: 35026469.

Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. Am J Hum Genet. 2014 Jul 3;95(1):5-23. doi: 10.1016/j.ajhg.2014.06.009. PMID: 24995866; PMCID: PMC4085641.

Li Z, Li X, Zhou H, Gaynor SM, Selvaraj MS, Arapoglou T, Quick C, Liu Y, Chen H, Sun R, Dey R, Arnett DK, Auer PL, Bielak LF, Bis JC, Blackwell TW, Blangero J, Boerwinkle E, Bowden DW, Brody JA, Cade BE, Conomos MP, Correa A, Cupples LA, Curran JE, de Vries PS, Duggirala R, Franceschini N, Freedman BI, Göring HHH, Guo X, Kalyani RR, Kooperberg C, Kral BG, Lange LA, Lin BM, Manichaikul A, Manning AK, Martin LW, Mathias RA, Meigs JB, Mitchell BD, Montasser ME, Morrison AC, Naseri T, O'Connell JR, Palmer ND, Peyser PA, Psaty BM, Raffield LM, Redline S, Reiner AP, Reupena MS, Rice KM, Rich SS, Smith JA, Taylor KD, Taub MA, Vasan RS, Weeks DE, Wilson JG, Yanek LR, Zhao W; NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium; TOPMed Lipids Working Group, Rotter JI, Willer CJ, Natarajan P, Peloso GM, Lin X. A framework for detecting noncoding rare-variant associations of large-scale whole-genome sequencing studies. Nat Methods. 2022 Oct 27. doi: 10.1038/s41592-022-01640-x. Epub ahead of print. PMID: 36303018.

Ma S, Wang C, Khan A, Liu L, Dalgleish J, Kiryluk K, He Z, Ionita-Laza I. BIGKnock: fine-mapping gene-based associations via knockoff analysis of biobank-scale data. Genome Biol. 2023 Feb 13;24(1):24. doi: 10.1186/s13059-023-02864-6. PMID: 36782330; PMCID: PMC9926792.

Mountjoy E, Schmidt EM, Carmona M, Schwartzentruber J, Peat G, Miranda A, Fumis L, Hayhurst J, Buniello A, Karim MA, Wright D, Hercules A, Papa E, Fauman EB, Barrett JC, Todd JA, Ochoa D, Dunham I, Ghoussaini M. An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. Nat Genet. 2021 Nov;53(11):1527-1533. doi: 10.1038/s41588-021-00945-5. Epub 2021 Oct 28. PMID: 34711957; PMCID: PMC7611956.

Nicolae DL. Association Tests for Rare Variants. Annu Rev Genomics Hum Genet. 2016 Aug 31;17:117-30. doi: 10.1146/annurev-genom-083115-022609. Epub 2016 Apr 21. PMID: 27147090.

Peyrot WJ, Price AL. Identifying loci with different allele frequencies among cases of eight psychiatric disorders using CC-GWAS. Nat Genet. 2021 Apr;53(4):445-454. doi: 10.1038/s41588-021-00787-1. Epub 2021 Mar 8. PMID: 33686288; PMCID: PMC8038973.

Shimelis H, Oetjens MT, Walsh LK, Wain KE, Znidarsic M, Myers SM, Finucane BM, Ledbetter DH, Martin CL. Prevalence and Penetrance of Rare Pathogenic Variants in Neurodevelopmental Psychiatric Genes in a Health Care System Population. Am J Psychiatry. 2023 Jan 1;180(1):65-72. doi: 10.1176/appi.ajp.22010062. Epub 2022 Dec 7. PMID: 36475376.

Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, Bryois J, Chen CY, Dennison CA, Hall LS, Lam M, Watanabe K, Frei O, Ge T, Harwood JC, Koopmans F, Magnusson S, Richards AL, Sidorenko J, Wu Y, Zeng J, Grove J, Kim M, Li Z, Voloudakis G, Zhang W, Adams M, Agartz I, Atkinson EG, Agerbo E, Al Eissa M, Albus M, Alexander M, Alizadeh BZ, Alptekin K, Als TD, Amin F, Arolt V, Arrojo M, Athanasiu L, Azevedo MH, Bacanu SA, Bass NJ, Begemann M, Belliveau RA, Bene J, Benyamin B, Bergen SE, Blasi G, Bobes J, Bonassi S, Braun A, Bressan RA, Bromet EJ, Bruggeman R, Buckley PF, Buckner RL, Bybjerg-Grauholm J, Cahn W, Cairns MJ, Calkins ME, Carr VJ, Castle D, Catts SV, Chambert KD, Chan RCK, Chaumette B, Cheng W, Cheung EFC, Chong SA, Cohen D, Consoli A, Cordeiro Q, Costas J, Curtis C, Davidson M, Davis KL, de Haan L, Degenhardt F, DeLisi LE, Demontis D, Dickerson F, Dikeos D, Dinan T, Djurovic S, Duan J, Ducci G, Dudbridge F, Eriksson JG, Fañanás L, Faraone SV, Fiorentino A, Forstner A, Frank J, Freimer NB, Fromer M, Frustaci A, Gadelha A, Genovese G, Gershon ES, Giannitelli M, Giegling I, Giusti-Rodríguez P, Godard S, Goldstein JI, González Peñas J, González-Pinto A, Gopal S, Gratten J, Green MF, Greenwood TA, Guillin O, Gülöksüz S, Gur RE, Gur RC, Gutiérrez B, Hahn E, Hakonarson H, Haroutunian V, Hartmann AM, Harvey C, Hayward C, Henskens FA, Herms S, Hoffmann P, Howrigan DP, Ikeda M, Iyegbe C, Joa I, Julià A, Kähler AK, Kam-Thong T, Kamatani Y, Karachanak-Yankova S, Kebir O, Keller MC, Kelly BJ, Khrunin A, Kim SW, Klovins J, Kondratiev N, Konte B, Kraft J, Kubo M, Kučinskas V, Kučinskiene ZA, Kusumawardhani A, Kuzelova-Ptackova H, Landi S, Lazzeroni LC, Lee PH, Legge SE, Lehrer DS, Lencer R, Lerer B, Li M, Lieberman J, Light GA, Limborska S, Liu CM, Lönnqvist J, Loughland CM, Lubinski J, Luykx JJ, Lynham A, Macek M Jr, Mackinnon A, Magnusson PKE, Maher BS, Maier W, Malaspina D, Mallet J, Marder SR, Marsal S, Martin AR, Martorell L, Mattheisen M, McCarley RW, McDonald C, McGrath JJ, Medeiros H, Meier S, Melegh B, Melle I, Mesholam-Gately RI, Metspalu A, Michie PT, Milani L, Milanova V, Mitjans M, Molden E, Molina E, Molto MD, Mondelli V, Moreno C, Morley CP, Muntané G, Murphy KC, Myin-Germeys I, Nenadić I, Nestadt G, Nikitina-Zake L, Noto C, Nuechterlein KH, O'Brien NL, O'Neill FA, Oh SY, Olincy A, Ota VK, Pantelis C, Papadimitriou GN, Parellada M, Paunio T, Pellegrino R, Periyasamy S, Perkins DO, Pfuhlmann B, Pietiläinen O, Pimm J, Porteous D, Powell J, Quattrone D, Quested D, Radant AD, Rampino A, Rapaport MH, Rautanen A, Reichenberg A, Roe C, Roffman JL, Roth J, Rothermundt M, Rutten BPF, Saker-Delye S, Salomaa V, Sanjuan J, Santoro ML, Savitz A, Schall U, Scott RJ, Seidman LJ, Sharp SI, Shi J, Siever LJ, Sigurdsson E, Sim K, Skarabis N, Slominsky P, So HC, Sobell JL, Söderman E, Stain HJ, Steen NE, Steixner-Kumar AA, Stögmann E, Stone WS, Straub RE, Streit F, Strengman E, Stroup TS, Subramaniam M, Sugar CA, Suvisaari J, Svrakic DM, Swerdlow NR, Szatkiewicz JP, Ta TMT, Takahashi A, Terao C, Thibaut F, Toncheva D, Tooney PA, Torretta S, Tosato S, Tura GB, Turetsky BI, Üçok A, Vaaler A, van Amelsvoort T, van Winkel R, Veijola J, Waddington J, Walter H, Waterreus A, Webb BT, Weiser M, Williams NM, Witt SH, Wormley BK, Wu JQ, Xu Z, Yolken R, Zai CC, Zhou W, Zhu F, Zimprich F, Atbaşoğlu EC, Ayub M, Benner C, Bertolino A, Black DW, Bray NJ, Breen G, Buccola NG, Byerley WF, Chen WJ, Cloninger CR, Crespo-Facorro B, Donohoe G, Freedman R, Galletly C, Gandal MJ, Gennarelli M, Hougaard DM, Hwu HG, Jablensky AV, McCarroll SA, Moran JL, Mors O, Mortensen PB, Müller-Myhsok B, Neil AL, Nordentoft M, Pato MT, Petryshen TL, Pirinen M, Pulver AE, Schulze TG, Silverman JM, Smoller JW, Stahl EA, Tsuang DW, Vilella E, Wang SH, Xu S; Indonesia Schizophrenia Consortium; PsychENCODE; Psychosis Endophenotypes International Consortium; SynGO Consortium; Adolfsson R, Arango C, Baune BT, Belangero SI, Børglum AD, Braff D, Bramon E, Buxbaum JD, Campion D, Cervilla JA, Cichon S, Collier DA, Corvin A, Curtis D, Forti MD, Domenici E, Ehrenreich H, Escott-Price V, Esko T, Fanous AH, Gareeva A, Gawlik M, Gejman PV, Gill M, Glatt SJ, Golimbet V, Hong KS, Hultman CM, Hyman SE, Iwata N, Jönsson EG, Kahn RS, Kennedy JL, Khusnutdinova E, Kirov G, Knowles JA, Krebs MO, Laurent-Levinson C, Lee J, Lencz T, Levinson DF, Li QS, Liu J, Malhotra AK, Malhotra D, McIntosh A, McQuillin A, Menezes PR, Morgan VA, Morris DW, Mowry BJ, Murray RM, Nimgaonkar V, Nöthen MM, Ophoff RA, Paciga SA, Palotie A, Pato CN, Qin S, Rietschel M, Riley BP, Rivera M, Rujescu D, Saka MC, Sanders AR, Schwab SG, Serretti A, Sham PC, Shi Y, St Clair D, Stefánsson H, Stefansson K, Tsuang MT, van Os J, Vawter MP, Weinberger DR, Werge T, Wildenauer DB, Yu X, Yue W, Holmans PA, Pocklington AJ, Roussos P, Vassos E, Verhage M, Visscher PM, Yang J, Posthuma D, Andreassen OA, Kendler KS, Owen MJ, Wray NR, Daly MJ, Huang H, Neale BM, Sullivan PF, Ripke S, Walters JTR, O'Donovan MC; Schizophrenia Working Group of the Psychiatric Genomics Consortium. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature. 2022 Apr;604(7906):502-508. doi: 10.1038/s41586-022-04434-5. Epub 2022 Apr 8. PMID: 35396580; PMCID: PMC9392466.

Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA, Nguyen-Viet TA, Wedow R, Zacher M, Furlotte NA; 23andMe Research Team; Social Science Genetic Association Consortium; Magnusson P, Oskarsson S, Johannesson M, Visscher PM, Laibson D, Cesarini D, Neale BM, Benjamin DJ. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nat Genet. 2018 Feb;50(2):229-237. doi: 10.1038/s41588-017-0009-4. Epub 2018 Jan 1. Erratum in: Nat Genet. 2019 Jul;51(7):1190. Erratum in: Nat Genet. 2019 Aug;51(8):1295. PMID: 29292387; PMCID: PMC5805593.

Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, Ermel R, Ruusalepp A, Quertermous T, Hao K, Björkegren JLM, Im HK, Pasaniuc B, Rivas MA, Kundaje A. Opportunities and challenges for transcriptome-wide association studies. Nat Genet. 2019 Apr;51(4):592-599. doi: 10.1038/s41588-019-0385-z. Epub 2019 Mar 29. PMID: 30926968; PMCID: PMC6777347.

Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017 Nov 28;8(1):1826. doi: 10.1038/s41467-017-01261-5. PMID: 29184056; PMCID: PMC5705698.

Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM, Yang J. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016 May;48(5):481-7. doi: 10.1038/ng.3538. Epub 2016 Mar 28. PMID: 27019110.

Zuk O, Schaffner SF, Samocha K, Do R, Hechter E, Kathiresan S, Daly MJ, Neale BM, Sunyaev SR, Lander ES. Searching for missing heritability: designing rare variant association studies. Proc Natl Acad Sci U S A. 2014 Jan 28;111(4):E455-64. doi: 10.1073/pnas.1322563111. Epub 2014 Jan 17. PMID: 24443550; PMCID: PMC3910587.

# References (RVAS gene list)

**Trait**: NDD  
**Source**:DDD\_2017  
**PubMed** **citation**:  
Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. Nature. 2017 Feb 23;542(7642):433-438. doi: 10.1038/nature21062. Epub 2017 Jan 25. PMID: 28135719; PMCID: PMC6016744.

**Trait**: NDD  
**Source**: ASD\_2019\_Feliciano  
**PubMed** **citation**:  
Feliciano P, Zhou X, Astrovskaya I, Turner TN, Wang T, Brueggeman L, Barnard R, Hsieh A, Snyder LG, Muzny DM, Sabo A; SPARK Consortium; Gibbs RA, Eichler EE, O'Roak BJ, Michaelson JJ, Volfovsky N, Shen Y, Chung WK. Exome sequencing of 457 autism families recruited online provides evidence for autism risk genes. NPJ Genom Med. 2019 Aug 23;4:19. doi: 10.1038/s41525-019-0093-8. PMID: 31452935; PMCID: PMC6707204.

**Trait**: NDD, SCZ, XLDD  
**Source**: DD\_2019\_ASD\_Coe  
**PubMed citation**: Coe BP, Stessman HAF, Sulovari A, Geisheker MR, Bakken TE, Lake AM, Dougherty JD, Lein ES, Hormozdiari F, Bernier RA, Eichler EE. Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. Nat Genet. 2019 Jan;51(1):106-116. doi: 10.1038/s41588-018-0288-4. Epub 2018 Dec 17. PMID: 30559488; PMCID: PMC6309590.

**Trait**: NDD  
**Source**: NDD\_2020\_Wang  
**PubMed** **citation**:  
Wang T, Hoekzema K, Vecchio D, Wu H, Sulovari A, Coe BP, Gillentine MA, Wilfert AB, Perez-Jurado LA, Kvarnung M, Sleyp Y, Earl RK, Rosenfeld JA, Geisheker MR, Han L, Du B, Barnett C, Thompson E, Shaw M, Carroll R, Friend K, Catford R, Palmer EE, Zou X, Ou J, Li H, Guo H, Gerdts J, Avola E, Calabrese G, Elia M, Greco D, Lindstrand A, Nordgren A, Anderlid BM, Vandeweyer G, Van Dijck A, Van der Aa N, McKenna B, Hancarova M, Bendova S, Havlovicova M, Malerba G, Bernardina BD, Muglia P, van Haeringen A, Hoffer MJV, Franke B, Cappuccio G, Delatycki M, Lockhart PJ, Manning MA, Liu P, Scheffer IE, Brunetti-Pierri N, Rommelse N, Amaral DG, Santen GWE, Trabetti E, Sedláček Z, Michaelson JJ, Pierce K, Courchesne E, Kooy RF; SPARK Consortium; Nordenskjöld M, Romano C, Peeters H, Bernier RA, Gecz J, Xia K, Eichler EE. Large-scale targeted sequencing identifies risk genes for neurodevelopmental disorders. Nat Commun. 2020 Oct 1;11(1):4932. doi: 10.1038/s41467-020-18723-y. Erratum in: Nat Commun. 2020 Oct 21;11(1):5398. PMID: 33004838; PMCID: PMC7530681.

**Trait**: NDD  
**Source**: ASD\_2020\_ASC\_Satterstrom  
**PubMed** **citation**:  
Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, Peng M, Collins R, Grove J, Klei L, Stevens C, Reichert J, Mulhern MS, Artomov M, Gerges S, Sheppard B, Xu X, Bhaduri A, Norman U, Brand H, Schwartz G, Nguyen R, Guerrero EE, Dias C; Autism Sequencing Consortium; iPSYCH-Broad Consortium; Betancur C, Cook EH, Gallagher L, Gill M, Sutcliffe JS, Thurm A, Zwick ME, Børglum AD, State MW, Cicek AE, Talkowski ME, Cutler DJ, Devlin B, Sanders SJ, Roeder K, Daly MJ, Buxbaum JD. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell. 2020 Feb 6;180(3):568-584.e23. doi: 10.1016/j.cell.2019.12.036. Epub 2020 Jan 23. PMID: 31981491; PMCID: PMC7250485.

**Trait**: NDD, SCZ  
**Source**: SCZ\_SCHEMA\_2020  
**PubMed citation**:  
Singh T, Poterba T, Curtis D, Akil H, Al Eissa M, Barchas JD, Bass N, Bigdeli TB, Breen G, Bromet EJ, Buckley PF, Bunney WE, Bybjerg-Grauholm J, Byerley WF, Chapman SB, Chen WJ, Churchhouse C, Craddock N, Cusick CM, DeLisi L, Dodge S, Escamilla MA, Eskelinen S, Fanous AH, Faraone SV, Fiorentino A, Francioli L, Gabriel SB, Gage D, Gagliano Taliun SA, Ganna A, Genovese G, Glahn DC, Grove J, Hall MH, Hämäläinen E, Heyne HO, Holi M, Hougaard DM, Howrigan DP, Huang H, Hwu HG, Kahn RS, Kang HM, Karczewski KJ, Kirov G, Knowles JA, Lee FS, Lehrer DS, Lescai F, Malaspina D, Marder SR, McCarroll SA, McIntosh AM, Medeiros H, Milani L, Morley CP, Morris DW, Mortensen PB, Myers RM, Nordentoft M, O'Brien NL, Olivares AM, Ongur D, Ouwehand WH, Palmer DS, Paunio T, Quested D, Rapaport MH, Rees E, Rollins B, Satterstrom FK, Schatzberg A, Scolnick E, Scott LJ, Sharp SI, Sklar P, Smoller JW, Sobell JL, Solomonson M, Stahl EA, Stevens CR, Suvisaari J, Tiao G, Watson SJ, Watts NA, Blackwood DH, Børglum AD, Cohen BM, Corvin AP, Esko T, Freimer NB, Glatt SJ, Hultman CM, McQuillin A, Palotie A, Pato CN, Pato MT, Pulver AE, St Clair D, Tsuang MT, Vawter MP, Walters JT, Werge TM, Ophoff RA, Sullivan PF, Owen MJ, Boehnke M, O'Donovan MC, Neale BM, Daly MJ. Rare coding variants in ten genes confer substantial risk for schizophrenia. Nature. 2022 Apr;604(7906):509-516. doi: 10.1038/s41586-022-04556-w. Epub 2022 Apr 8. PMID: 35396579; PMCID: PMC9805802.

**Trait**: NDD, XLDD  
**Source**: DDDX\_2021  
**PubMed citation**:  
Martin HC, Gardner EJ, Samocha KE, Kaplanis J, Akawi N, Sifrim A, Eberhardt RY, Tavares ALT, Neville MDC, Niemi MEK, Gallone G, McRae J; Deciphering Developmental Disorders Study; Wright CF, FitzPatrick DR, Firth HV, Hurles ME. The contribution of X-linked coding variation to severe developmental disorders. Nat Commun. 2021 Jan 27;12(1):627. doi: 10.1038/s41467-020-20852-3. PMID: 33504798; PMCID: PMC7840967.

**Trait**: NDD  
**Source**: ASD\_2022\_ASC\_Fu  
**PubMed citation**:  
Fu JM, Satterstrom FK, Peng M, Brand H, Collins RL, Dong S, Wamsley B, Klei L, Wang L, Hao SP, Stevens CR, Cusick C, Babadi M, Banks E, Collins B, Dodge S, Gabriel SB, Gauthier L, Lee SK, Liang L, Ljungdahl A, Mahjani B, Sloofman L, Smirnov AN, Barbosa M, Betancur C, Brusco A, Chung BHY, Cook EH, Cuccaro ML, Domenici E, Ferrero GB, Gargus JJ, Herman GE, Hertz-Picciotto I, Maciel P, Manoach DS, Passos-Bueno MR, Persico AM, Renieri A, Sutcliffe JS, Tassone F, Trabetti E, Campos G, Cardaropoli S, Carli D, Chan MCY, Fallerini C, Giorgio E, Girardi AC, Hansen-Kiss E, Lee SL, Lintas C, Ludena Y, Nguyen R, Pavinato L, Pericak-Vance M, Pessah IN, Schmidt RJ, Smith M, Costa CIS, Trajkova S, Wang JYT, Yu MHC; Autism Sequencing Consortium (ASC); Broad Institute Center for Common Disease Genomics (Broad-CCDG); iPSYCH-BROAD Consortium; Cutler DJ, De Rubeis S, Buxbaum JD, Daly MJ, Devlin B, Roeder K, Sanders SJ, Talkowski ME. Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. Nat Genet. 2022 Sep;54(9):1320-1331. doi: 10.1038/s41588-022-01104-0. Epub 2022 Aug 18. PMID: 35982160; PMCID: PMC9653013.

**Trait:** NDD

**Source:** ASD\_2022\_MSSNG\_Trost  
**PubMed citation:**  
Trost B, Thiruvahindrapuram B, Chan AJS, Engchuan W, Higginbotham EJ, Howe JL, Loureiro LO, Reuter MS, Roshandel D, Whitney J, Zarrei M, Bookman M, Somerville C, Shaath R, Abdi M, Aliyev E, Patel RV, Nalpathamkalam T, Pellecchia G, Hamdan O, Kaur G, Wang Z, MacDonald JR, Wei J, Sung WWL, Lamoureux S, Hoang N, Selvanayagam T, Deflaux N, Geng M, Ghaffari S, Bates J, Young EJ, Ding Q, Shum C, D'Abate L, Bradley CA, Rutherford A, Aguda V, Apresto B, Chen N, Desai S, Du X, Fong MLY, Pullenayegum S, Samler K, Wang T, Ho K, Paton T, Pereira SL, Herbrick JA, Wintle RF, Fuerth J, Noppornpitak J, Ward H, Magee P, Al Baz A, Kajendirarajah U, Kapadia S, Vlasblom J, Valluri M, Green J, Seifer V, Quirbach M, Rennie O, Kelley E, Masjedi N, Lord C, Szego MJ, Zawati MH, Lang M, Strug LJ, Marshall CR, Costain G, Calli K, Iaboni A, Yusuf A, Ambrozewicz P, Gallagher L, Amaral DG, Brian J, Elsabbagh M, Georgiades S, Messinger DS, Ozonoff S, Sebat J, Sjaarda C, Smith IM, Szatmari P, Zwaigenbaum L, Kushki A, Frazier TW, Vorstman JAS, Fakhro KA, Fernandez BA, Lewis MES, Weksberg R, Fiume M, Yuen RKC, Anagnostou E, Sondheimer N, Glazer D, Hartley DM, Scherer SW. Genomic architecture of autism from comprehensive whole-genome sequence annotation. Cell. 2022 Nov 10;185(23):4409-4427.e18. doi: 10.1016/j.cell.2022.10.009. PMID: 36368308.

**Trait**: Suicide  
**Source**: Suicide\_2021  
**PubMed citation**:  
DiBlasi E, Shabalin AA, Monson ET, Keeshin BR, Bakian AV, Kirby AV, Ferris E, Chen D, William N, Gaj E, Klein M, Jerominski L, Callor WB, Christensen E, Smith KR, Fraser A, Yu Z, Gray D; PsychChip Investigators of the Psychiatric Genomics Consortium; Camp NJ, Stahl EA, Li QS, Docherty AR, Coon H. Rare protein-coding variants implicate genes involved in risk of suicide death. Am J Med Genet B Neuropsychiatr Genet. 2021 Dec;186(8):508-520. doi: 10.1002/ajmg.b.32861. Epub 2021 May 27. PMID: 34042246; PMCID: PMC9292859.