Data supplement for Shimelis et al., Prevalence and Penetrance of Rare Pathogenic Variants in Neurodevelopmental Psychiatric Genes in a Health Care System Population. Am J Psychiatry (doi: 10.1176/appi.ajp.22010062)

Methods

Variants Quality Control (QC)

Single nucleotide variants (SNVs) and indels were identified after applying stringent quality metric parameters: allelic balance (AB) 0.2, genotype depth (DP) 20, and genotype quality (GQ) 80. Variants were excluded if located in segmental duplication/repetitive regions or outside of consistently covered exome regions (bases covered at >20x in 90% of VCRome samples and covered at >20x in 90% of xGen samples) between the two sequencing platforms (VCRome set and xGen set) used for sequencing of the first ~61K and ~31K samples (1). After applying the QC metrics above, LOF variants in NPD genes are distributed across both sequencing platforms (Figure S4). A total of 59,441 of the 90,595 samples that passed QC were in the VCRome set, while 31,154 were in the xGen set. To minimize the number of potentially false positive variants, all variants involving substitution or deletion of more than one nucleotide (indels) were confirmed by evaluation of IGV plots in at least two samples when observed in multiple individuals. We included only variants on autosomes. Single gene copy number variants (CNVs) were called from ES data and were confirmed by an array-based call (with 10 or more SNPs) using PennCNV or manual inspection of the signal intensity data (logR and B-allele frequency) within the pathogenic region as described previously (2, 3). Exome sequencing CNV calls involving two or more exons and validated by 10 or more SNPs were included in our analysis.

Rare Pathogenic Loss-Function-Variants Identification

Variants were annotated with Ensembl Variant Effect Predictor (VEP) version 95 based on ensemble gene build 96. Ensemble canonical transcript was selected. Based on the selected transcript (Table S1), loss-of-function (LOF) variants (frameshift, nonsense, essential (+/- 1,2) splice site, stop-loss, and start-loss variants); were functionally annotated with LOFTEE (4), to identify “high-confidence” and “low-confidence” variants. Rare variants were identified by exclusion of variants with minor allele frequency (MAF) of ≥0.001 in either the DiscovEHR cohort or gnomAD. Rare LOF variants and single gene CNVs were evaluated for evidence of pathogenicity following the general framework of ACMG-AMP guidelines (5–7). Only variants classified as pathogenic or likely pathogenic were included in prevalence, penetrance, and effect size estimates. *KANSL1* variants in exons 1-3 were excluded due to potential loss of function variants artefacts from copy-number polymorphic segmental duplications encompassing exons 1-3 (8).

NPD Genes Selection and Annotation Across External Sources

NPD genes were selected from high-confidence candidate genes (Tier 1) in the Developmental Brain Disorder (DBD) Gene Database (https://dbd.geisingeradmi.org/; accessed January 2021) (9). Of the 152 Tier 1 genes, we restricted our analysis to 94 high-confidence NPD genes that met the following three criteria: 1) autosomal dominant inheritance, 2) haploinsufficiency (LOF mechanism), and 3) strong evidence for NPD association, either by having the highest confidence ranking in at least one externally curated NPD gene database; or by being reported in six or more probands curated by

the DBD Gene Database. Genes on the X -chromosome were excluded from this study (see Figures S1 and S2 in the online supplement). Tier 1 genes were considered supported by an external database if ranked as having the highest confidence of association with NPD in the respective database: SFARI genes with an assigned gene score of 1 (SFARI Gene: https://gene.sfari.org/; accessed February 2021); Development Disorder Genotype-Phenotype (DDG2P) genes “Confirmed” to cause neurodevelopmental disorder (DDG2P: https://www.deciphergenomics.org/ddd/ddgenes; accessed February 2021); and denovo-db genes (https://denovo-db.gs.washington.edu/denovo-db/; accessed February 2021) with likely gene disruption (LGD) that met a genome-wide statistical significance threshold for association with a neurodevelopmental disorder (10).

We performed an intersection between the 94 NPD genes and three active NPD gene curation databases to support disease causality: SFARI (11), DDG2P (12), and denovo-db. Nearly all NPD genes (n=91, 97%,) are supported by at least one external database as they were categorized as high-confidence genes in the respective databases (see Table S1, Figure S2 in the online supplement). Consistent with the expectations of monogenic NPD genes, namely strong selective pressure and rarity of loss-of-function variants in the general population, our list of NPD genes is enriched for constraint (n=89; 95%) based on the probability of loss-of-function intolerance (pLI) metric in gnomAD (pLI >0.9). As of January 2021, the Clinical Genome Resource (ClinGen) has curated and scored 76 of the 94 NPD genes for haploinsufficiency: 67 are scored 3 (sufficient evidence), and 9 are scored 2 (emerging evidence). ClinGen

curated 42 of the 94 genes and classified 41 with definitive and 1 with moderate evidence of gene-disease relationship (see Table S1 in the online supplement).

Identification of LOF variants Observed in the gnomAD Non-Finnish-European (NFE) population

The gnomAD v2.1.1 all chromosomes sites VCF (https://gnomad.broadinstitute.org/downloads: Accessed May 2022) was used to identify LOF variants in NPD genes observed in the gnomAD population (13). Variants in NPD genes passing gnomAD QC filters (FILTER=PASS), without Segdup, LCR, or decoy flags were identified. To subset gnomAD variants to genomic regions included in exome sequences from our study, the genomic positions of gnomAD variants in NPD genes were intersected with consistently covered bases (described above). LOF variants in each of the 94 genes observed in gnomAD were identified after annotation of the VCF with VEP version 103 using the same transcripts listed in Table S1. Because DiscovEHR is largely White Europeans, we compared DiscovEHR's variant frequency to gnomAD NFE population. *KANSL1* variants in exons 1-3 were also excluded from the list of variants observed in the gnomAD NFE population. The gnomAD frequency across genes was calculated by dividing the total number LOF alleles by the median number of LOF alleles. The difference in LOF variant frequency between gnomAD and DiscovEHR datasets could be attributed to differences in the frequency of pathogenic LOF variants and/or variant quality thresholds used in gnomaAD (GQ>=20, DP>=10, and allele balance >0.2) and our study (GQ>=80, DP>=20, and allele balance >0.2). We also examined single gene CNVs in the gnomAD SV (structural variant) dataset from 10,738

unrelated genomes, however none of the CNVs included in our study have been observed in the gnomAD SV dataset.

NPD Diagnosis Extraction

NPD diagnoses were extracted from the participant's EHR ICD 9/10) *billing codes* (see Table S2 in the online supplement). Billing codes used here have been described by our group previously and were used to evaluate phenotypes associated with NPD in the DiscovEHR cohort (2). Selected ICD9/10 billing codes were intended to capture "neurodevelopmental disorders," "schizophrenia spectrum and other psychotic disorders," "bipolar and related disorders," "depressive disorders," "anxiety disorders," "obsessive-compulsive and related disorders," "epilepsy," and "cerebral palsy". The presence of a congenital anomaly, including cleft lip/palate, as well as structural malformations of the central nervous, cardiac, and urogenital systems, was defined as described in our previous analysis (2). Clinical diagnosis data were collected in December 2020.

Sensitivity Analyses

We conducted a sensitivity analysis to ensure that the odds ratio estimates for NPD disorders based on the entire dataset are consistent with OR estimates based on case-control samples from each of the two sequencing platforms. For this sensitivity analysis, the odds ratios were calculated using the European samples of the VCRome (n=54,597) or IDT (n=27,130) sets. We also performed association analysis using an unrelated subset of the cohort (n=64,914) (1), which included one sample from each 1st and 2nd degree related families in DiscovEHR, to account for potential effects of relatedness on

the association between pathogenic variants and NPD diagnosis. To account for the effects of pediatric patient participants (i.e., age group which may disproportionately have a high frequency of NPD diagnosis) and patients with intellectual disability (NPD outcome with the highest odds ratio) on the association between pathogenic variants and NPD diagnosis, case-control analysis was performed after excluding pediatric participants (age <18) and those with ID diagnosis from the case-control analysis. Because all case-control analyses were conducted using the DiscovEHR cohort's European subset, the effects of pathogenic variants in Black/African American and White Hispanic/Latino minority populations were also investigated in separate analysis. We conducted case-control study of the association between NPD and pathogenic variants or CNVs, classifying cases as those with the specific diagnosis and controls as those without any of the 14 NPD, accounting for the well-established phenotypic variability of NPDs because those with any of the NPD are considered affected cases and were not included in the control group for each test.

Chart Review Methodology

To supplement our analysis of NPD identification based on ICD codes, we performed a manual chart review of 27 individuals with a pathogenic variant in 11 of the 94 NPD genes expected to be fully penetrant for the associated disorder. To identify fully penetrant genes, penetrance of all 61 genes with ≥1 individual in DiscovEHR population cohort with a variant were systematically extracted from the “Penetrance” section in each GeneReviews (http://www.genereviews.org/) accessed April 2021. Of the 61 genes, 10 (*EFTUD2, TSC2, STXBP1, CHD2, GRIN2B, KMT2D, KANSL1, NSD1, TCF4*

and WAC) genes were listed as being fully penetrant for the respective genetic disorder. Individuals with *CHD8* pathogenic variant were also included in manual chart review, despite the fact that the gene is not included in GeneReviews, because variants in this gene have been reported to have high penetrance for ASD (10).

Systematic, manual chart reviews were conducted in Epic EHRs. Coded incidents of current and past diagnoses were reviewed, along with biometrics (head circumference, height, BMI), social history (educational attainment, occupational history) and medication use. The search feature was used to identify NPD phenotypes for each patient that were captured in visit notes or encounter documentation and were otherwise absent from ICD code lists. Scanned documents (not text searchable) were omitted from this analysis.

Through chart review, we captured:

- **Length of EHR (years)**
  - **Methods**:
    - Open chart Click to “Load All Records Now” Note date of most recent encounter Sort oldest to newest Note date of first documented encounter
    - Open chart Click to “Load All Records Now” Click Filters Click Encounter Department Review variety of departments and visit counts
  - **Justification**: short duration or seen only for labs/specialty care could mean that the lack of NPD evidence comes from minimal documentation as opposed to absence of phenotype

- **Educational attainment**
  - **Methods**:
    - Open chart Click History tab Click “Socioeconomic” subtab (within Social Determinants) Note Years of Education (if documented)
    - Key term search: /edu/ or educate or education or school or high school or “HS” or college or degree or “GED” or graduate or graduation or special education or learning support
  - **Justification**: school psychological testing, notes from teachers, classroom placement (special ed, autism support, emotional

support, with a 1:1 aide) and school therapies (speech, OT, PT) can all indirectly suggest NPD history

- **Occupation**
  - **Methods**:
    - Open chart Click Demographics tab Note Employment Information (Occupation, Employer) (if documented) Click to View Employer for more details (if available)
    - Key term search: work or works or job or occupation or employer or employment or retired or homemaker or disability or on disability
  - **Justification**: if on disability, clinicians may have documented qualifying incident(s), which may include seizure on the job, mood disorder interfering with ability to work, etc.

- **NPD Phenotypes**
  - **Methods**:
    - Open chart Click Snapshot tab Review diagnoses in problem list (should be captured by ICD code pull) Click History tab Review diagnoses in Past Medical History List
    - Key term search:
      - psych or psychology or psychiatry or psychologist or psychiatrist or psychosis or psychotic or mental health or neurodevelopment
      - gene or genetic or exome or genome or variant or mutation
      - cognition or cognitive or /cogn/ or intelligence or intelligence quotient or /intel/ or “IQ” or “FSIQ” or learning or “barriers to learning” or mental or mental retardation or “MR” or developmental delay or delay or remedial or regression
      - anxiety or anxious or phobia or panic or depression or depressed or mood or mood disorder or adjustment or adjustment disorder or bipolar or schizophrenia or /schiz/ or delusion
      - post-traumatic stress disorder or trauma or “PTSD” or /suicid/
      - autism or autism spectrum disorder or “ASD” or asperger or asperger syndrome or pervasive developmental disorder or “PDD” or “PDD-NOS” or social communication
      - seizure or epilepsy or convulsion or convulsive or “EEG” or “MRI” or migraine or neuro or neurology or neuropathy or neuroscience or neurologist
      - speech or language or speech delay or language delay or language disorder or speech disorder or

specific language disorder or symbolic dysfunction or phonological or speech therapy or communication disorder or communication deficit or verbal or nonverbal or hearing loss
- attention deficit or “ADHD” or impulsivity or impulsive or hyperactive or hyperkinetic or hyperkinesis or motor or motor delay or motor disorder or cerebral palsy or “CP” or stereotypy or coordination or coordination disorder or obsessive-compulsive disorder or “OCD” or compulsive or repetitive

- **Biometrics (height, weight, BMI, head circumference)**
    - Methods**:**
        - Open chart Review Recent Vitals for Height, Weight, BMI (if applicable)
        - Open chart Click Growth Chart tab (if available) Note Recent Height, Weight, BMI and relative percentiles by age
        - Key term search: head or head circumference or head size or Nellhaus or Nellhaus curve or normocephalic or “HEENT”

REFERENCES

1. Staples J, Maxwell EK, Gosalia N, et al.: Profiling and Leveraging Relatedness in a Precision Medicine Cohort of 92,455 Exomes. Am J Hum Genet 2018; 102:874–889

2. Martin CL, Wain KE, Oetjens MT, et al.: Identification of Neuropsychiatric Copy Number Variants in a Health Care System Population [Internet]. JAMA Psychiatry 2020; Available from: http://dx.doi.org/10.1001/jamapsychiatry.2020.2159

3. Wang K, Li M, Hadley D, et al.: PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res 2007; 17:1665–1674

4. Karczewski K: loftee [Internet]. Github[cited 2021 Jan] Available from: https://github.com/konradjk/loftee

5. Richards S, Aziz N, Bale S, et al.: 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 2015; 17:405–424

6. Tayoun AA, Pesaran T, DiStefano M, et al.: Recommendations for Interpreting the Loss of Function PVS1 ACMG/AMP Variant Criteria [Internet]. bioRxiv 2018; 313718[cited 2018 Sep 6] Available from: https://www.biorxiv.org/content/early/2018/05/09/313718

7. Riggs ER, Andersen EF, Cherry AM, et al.: Technical standards for the interpretation and reporting ofconstitutional copy-number variants: a joint consensus recommendation of the AmericanCollege of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource(ClinGen). Genet Med 2020; 22:245–257

8. Koolen DA, Pfundt R, Linda K, et al.: The Koolen-de Vries syndrome: a phenotypic comparison of patients with a 17q21.31 microdeletion versus a KANSL1 sequence variant. Eur J Hum Genet 2016; 24:652–659

9. Gonzalez-Mantilla AJ, Moreno-De-Luca A, Ledbetter DH, et al.: A Cross-Disorder Method to Identify Novel Candidate Genes for Developmental Brain Disorders. JAMA Psychiatry 2016; 73:275–283

10. Coe BP, Stessman HAF, Sulovari A, et al.: Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. Nat Genet 2019; 51:106–116

11. Abrahams BS, Arking DE, Campbell DB, et al.: SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). Mol Autism 2013; 4:36

12. Wright CF, Fitzgerald TW, Jones WD, et al.: Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. Lancet 2015; 385:1305–1314

13. Karczewski KJ, Francioli LC, Tiao G, et al.: The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020; 581:434–443

Figure S1. Selection of genes included in the study

Tier 1 genes (n=152)
TLK2 MYT1L STAG1 TBR1 PDHA1 WASF1 SCN2A USP7 CUL3 ASXL2 AHDC1 KIF11 CAMK2A KCNQ2 KMT2B CHD8 SRCAP KDM5B SON SATB2 PTEN CSDE1 EP300 TRIO RIMS1 PHIP ASXL3 SLC35A2 ITSN1 SETD2 SHANK2 ZMYND11 CAMK2B SETD1A PTCHD1 DLG4 CHAMP1 SETBP1 SMARCC2 NSD1 NIPBL SHANK3 POGZ GATAD2B HDAC8 KMT2D CHD3 ANK2 NRXN1 KAT6A NAA15 DYRK1A CREBBP RORA RALGAPB AUTS2 TRIP12 MED13L SYNGAP1 MBD5 EFTUD2 TSC2 NUS1 ARID1B DDX3X HNRNPU KDM6A ZBTB18 IL1RAPL1 NSD2 CASK MECP2 ZNF462 WAC GPC3 GRIN2B BRPF1 ZNF292 STXBP1 EHMT1 RAI1 GIGYF1 HIVEP2 FOXP1 STAG2 CTCF CHD7 WDFY3 TCF20 MYH10 MAPK8IP3 SMC1A TCOF1 NEXMIF KMT2C SOX5 CLTC DSCAM NFIX ANKRD11 SLC6A1 KDM5C CSNK2B EBF3 TCF4 PAX6 PPP3CA PPM1D KAT6B WDR45 H1-4 IRF2BPL DPP6 CTNNB1 PRR12 ARHGEF9 SLC2A1 NBEA PHF3 PHF21A KMT2A ZEB2 SOX11 HIVEP3 CTTNBP2 SPAST KANSL1 CDKL5 UPF3B ASH1L CDK13 SETD5 CIC TNRC6B FOXG1 ADNP USP9X IQSEC2 MEF2C CHD2 PCDH19 KMT5B SLC9A6 KDM6B PURA BCL11A NCKAP1 SCN1A ASXL1 DNMT3A MAGEL2 UBE3A

Exclude imprinted genes (n=2): *UBE3A MAGEL2*
Exclude genes associated with age related clonal hematopoiesis (n=3): *ASXL1 DNMT3A PPM1D*
Exclude genes on X-chromosome (n=22): *IL1RAPL1 CDKL5 UPF3B PTCHD1 WDR45 SMC1A NEXMIF CASK ARHGEF9 USP9X IQSEC2 PCDH19 MECP2 SLC9A6 GPC3 HDAC8 PDHA1 DDX3X KDM5C SLC35A2 KDM6A STAG2*
Exclude genes associated with syndromes ID/ASD is not primary features (n=3): *PAX6 SPAST TCOF1*

Exclude genes linked with both recessive and dominant inheritance (n=5): *KDM5B KDM6B NRXN1 SLC2A1 NUS1*

Exclude genes with ClinGen HI score 1, or known dominant negative (n=6): *ASXL2 HIVEP3 PHF3 RIMS1 SRCAP H1-4*

Exclude genes with dominant negative loss of function mechanism (n=2): *CDK13 and CHD3*

Autosomal dominant genes with haploinsufficiency disease mechanism (n=109)
ZBTB18 TLK2 NSD2 MYT1L STAG1 TBR1 ZNF462 WAC GRIN2B BRPF1 WASF1 SCN2A ZNF292 USP7 STXBP1 CUL3 RAI1 EHMT1 HIVEP2 GIGYF1 AHDC1 FOXP1 CTCF KIF11 CAMK2A KCNQ2 CHD7 WDFY3 TCF20 MYH10 KMT2B MAPK8IP3 CHD8 KMT2C SON SATB2 SOX5 PTEN CLTC DSCAM CSDE1 EP300 NFIX TRIO SLC6A1 PHIP ANKRD11 ASXL3 ITSN1 SETD2 CSNK2B EBF3 TCF4 SHANK2 ZMYND11 CAMK2B PPP3CA PPM1D KAT6B SETD1A DLG4 CHAMP1 SETBP1 IRF2BPL PRR12 CTNNB1 DPP6 SMARCC2 NBEA NSD1 PHF21A NIPBL SHANK3 KMT2A ZEB2 POGZ GATAD2B SOX11 CTTNBP2 KMT2D KANSL1 ANK2 KAT6A DYRK1A NAA15 CREBBP SETD5 ASH1L CIC TNRC6B RORA ADNP FOXG1 RALGAPB AUTS2 TRIP12 MED13L MEF2C CHD2 SYNGAP1 KMT5B MBD5 EFTUD2 TSC2 PURA ARID1B BCL11A NCKAP1 SCN1A HNRNPU

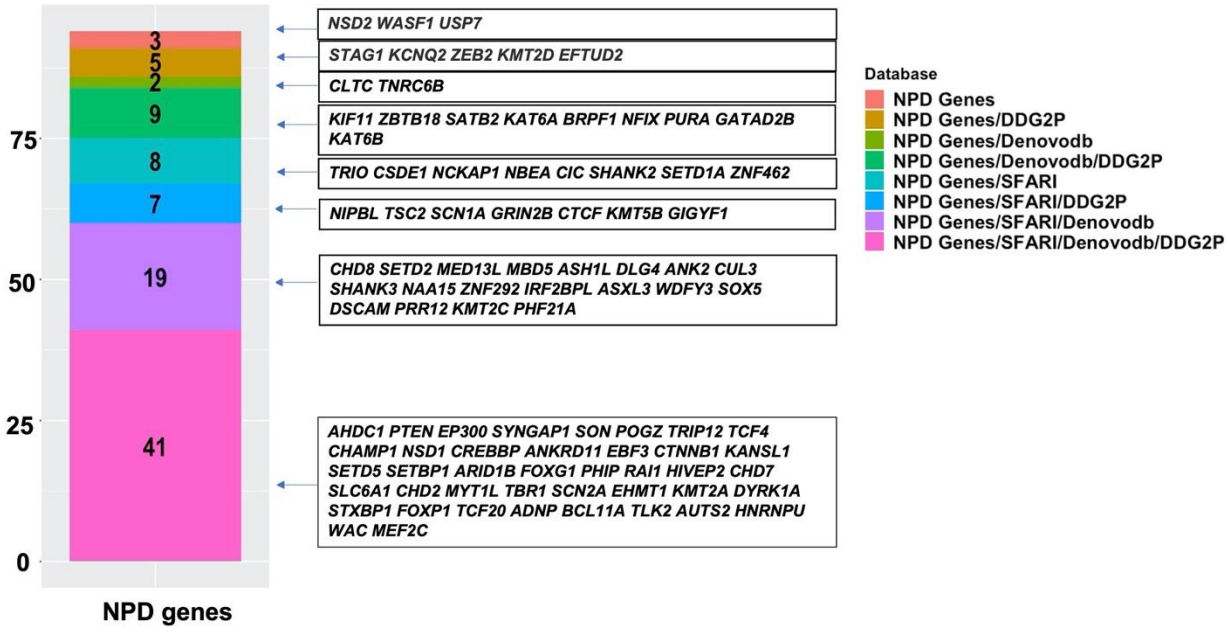
Exclude genes not supported by external NPD genes databases (SFARI, DDG2P, Denovodb) (n=18): *PPP3CA NSD2 CAMK2A MYH10 KMT2B MAPK8IP3 RORA DPP6 RALGAPB SOX11 WASF1 CTTNBP2 USP7 ITSN1 CSNK2B ZMYND11 CAMK2B.*

Autosomal dominant genes with haploinsufficiency disease mechanism supported by external database (n=91)
ZBTB18 TLK2 MYT1L STAG1 TBR1 ZNF462 WAC GRIN2B BRPF1 SCN2A ZNF292 STXBP1 CUL3 RAI1 EHMT1 HIVEP2 GIGYF1 AHDC1 FOXP1 CTCF KIF11 KCNQ2 CHD7 WDFY3 TCF20 CHD8 KMT2C SON SATB2 SOX5 PTEN CLTC DSCAM EP300 NFIX TRIO SLC6A1 PHIP ANKRD11 ASXL3 SETD2 EBF3 TCF4 SHANK2 KAT6B DLG4 CHAMP1 SETBP1 IRF2BPL PRR12 CTNNB1 NBEA NSD1 PHF21A NIPBL SHANK3 KMT2A ZEB2 POGZ GATAD2B KMT2D KANSL1 ANK2 KAT6A DYRK1A NAA15 CREBBP SETD5 ASH1L CIC TNRC6B ADNP FOXG1 AUTS2 TRIP12 MED13L MEF2C CHD2 SYNGAP1 KMT5B MBD5 EFTUD2 TSC2 PURA ARID1B BCL11A NCKAP1 SCN1A HNRNPU SETD1A

Include Genes with 6 or more *de novo* LOFs in the DBD genes database (n=3): Genes with sufficient genetic evidence; published reports of 6 or more *de novo* pLOF variants among unrelated NPD affected probands in the DBD genes database. *NSD2 USP7 WASF1*

Total number of genes included in the study: n=94

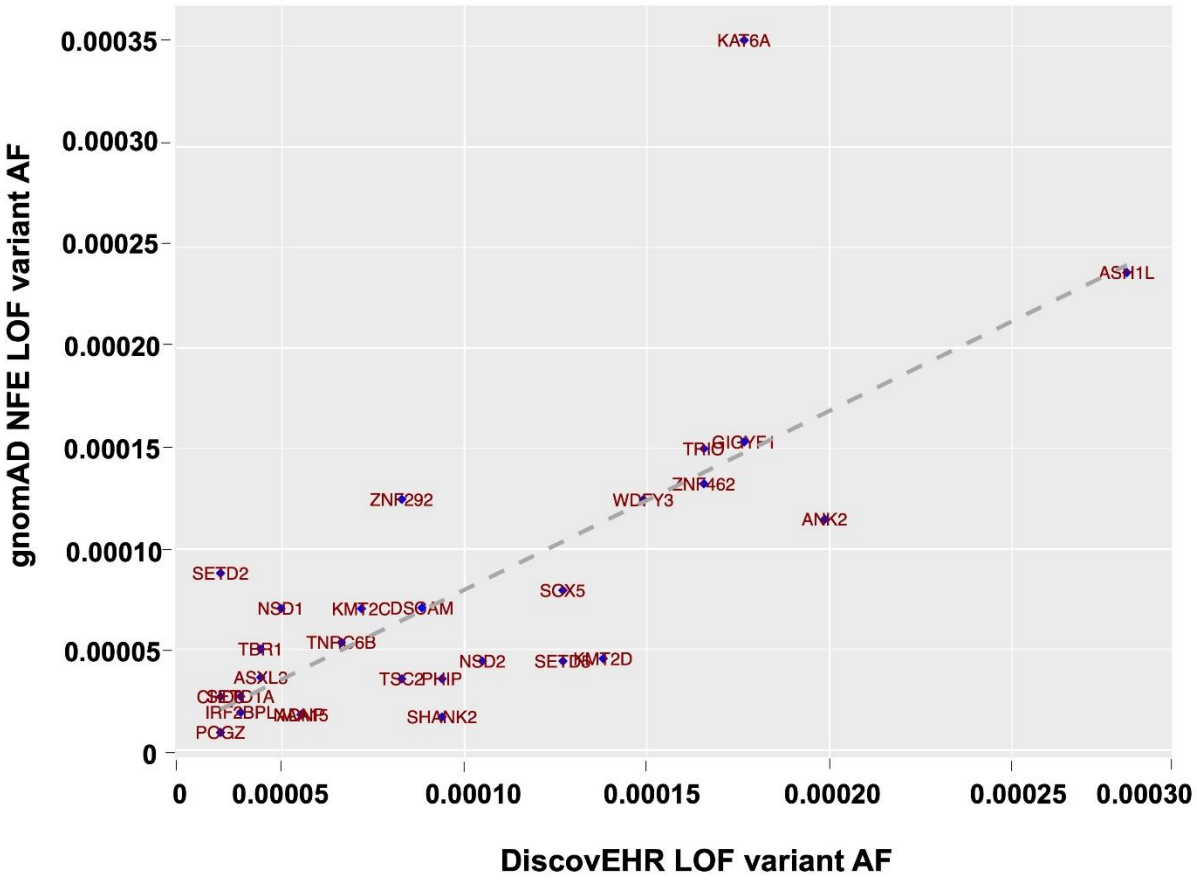
Figure S2. Ninety-seven percent (91/94) of NPD genes overlap with high-confidence genes in at least one other database



Ninety-four NPD genes were compared with SFARI genes with Score 1, DDG2P genes annotated as “Confirmed”, and Denovo-db genes with likely gene disruption (LGD) variants (10).
SFARI: Simons Foundation Autism Research Initiative
DDG2P: Developmental Disorder Genotype-phenotype database

Figure S3. LOF variants allele frequency across genes in the DiscovEHR and gnomAD NFE populations

(A) The LOF variant allele frequency of genes observed in five or more individuals in DiscovEHR. AF: Allele frequency



(B) LOF variant allele frequency of all genes with at least one LOF variant observed in both populations.

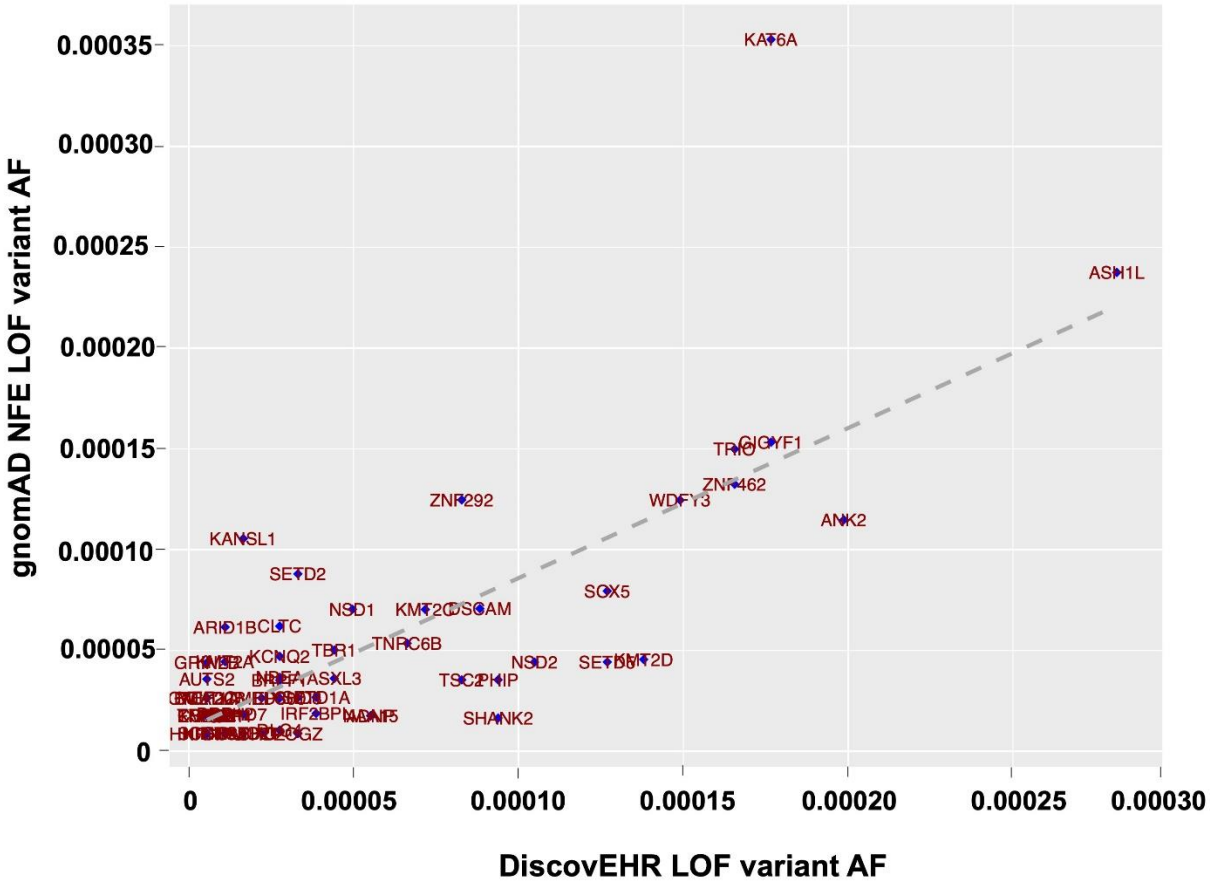
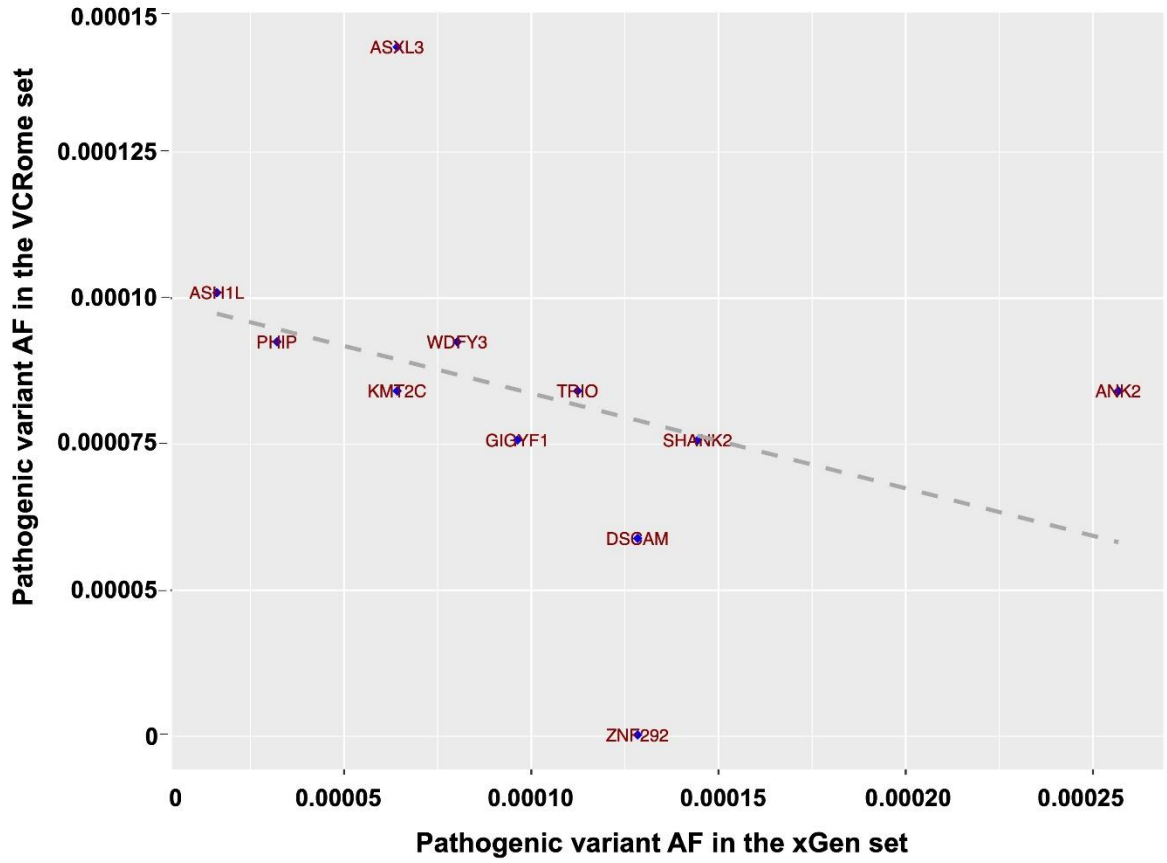


Figure S4. The frequency of pathogenic variants across genes in the VCRome and xGen sets



The frequency of pathogenic variants (including sequence variants and CNVs) across genes with the highest frequency (≥10 variants in the whole dataset) are shown for VCRome (n=59,441) and xGen (n=31,154) sets.

Table S1. Annotation of NPD genes across external databases

Gene	NCBI Reference sequence	gnomAD PLI	gnomAD LOEUF	ClinGen HI score	ClinGen Gene-disease validity; MONDO DISEASE ID: Classification	DDG2P category	SFARI gene-score[a]	Overlap with Denovo-db NDD genes[b]
ADNP	NM_001282531	1.00	0.123	3	MONDO_0014379 : Definitive	confirmed	1	Yes
AHDC1	NM_001029882	1.00	0.076	3	-	confirmed	1	Yes
ANK2	NM_001148	1.00	0.106	3	MONDO_0100038 : Definitive	possible	1	Yes
ANKRD11	NM_013275	1.00	0.107	3	MONDO_0007846 : Definitive	confirmed	1	Yes
ARID1B	NM_001346813	1.00	0.102	3	MONDO_0015452 : Definitive	confirmed	1	Yes
ASH1L	NM_018489	1.00	0.062	3	-	probable	1	Yes
ASXL3	NM_030632	1.00	0.199	3	-	probable	1	Yes
AUTS2	NM_015570	1.00	0.253	3	MONDO_0000508 : Definitive	confirmed	1	Yes
BCL11A	NM_022893	0.97	0.322	3	MONDO_0014914 : Definitive	confirmed	1	Yes
BRPF1	NM_004634	1.00	0.176	-	-	confirmed	-	Yes
CHAMP1	NM_032436	0.99	0.271	2	-	confirmed	1	Yes
CHD2	NM_001271	1.00	0.07	3	MONDO_0100038 : Definitive	confirmed	1	Yes
CHD7	NM_017780	1.00	0.076	3	MONDO_0008965 : Definitive	confirmed	1	Yes
CHD8	NM_001170629	1.00	0.082	3	MONDO_0005258 : Definitive	probable	1	Yes
CIC	NM_001304815	1.00	0.183	2	-	probable	1	NA
CLTC	NM_004859	1.00	0.104	-	-	probable	-	Yes
CREBBP	NM_004380	1.00	0.066	3	MONDO_0008393 : Definitive	confirmed	1	Yes
CSDE1	NM_001007553	1.00	0.144	-	-	-	1	NA
CTCF	NM_006565	1.00	0.148	2	-	confirmed	1	NA
CTNNB1	NM_001904	1.00	0.129	3	-	confirmed	1	Yes
CUL3	NM_003590	1.00	0.228	2	MONDO_0100038 : Definitive	possible	1	Yes
DLG4	NM_001365	1.00	0.238	-	-	probable	1	Yes
DSCAM	NM_001389	1.00	0.189	3	-	-	1	Yes
DYRK1A	NM_001347721	1.00	0.214	3	MONDO_0100038 : Definitive	confirmed	1	Yes
EBF3	NM_001005463	1.00	0.15	-	-	confirmed	1	Yes
EFTUD2	NM_004247	1.00	0.094	3	-	confirmed	-	NA
EHMT1	NM_024757	1.00	0.072	3	MONDO_0012455 : Definitive	confirmed	1	Yes
EP300	NM_001429	1.00	0.099	3	-	confirmed	1	Yes
FOXG1	NM_005249	0.94	0.329	3	MONDO_0100040 : Definitive	confirmed	1	Yes
FOXP1	NM_032682	1.00	0.175	3	MONDO_0013352 : Definitive	confirmed	1	Yes
GATAD2B	NM_020699	1.00	0.097	3	-	confirmed	-	Yes
GIGYF1	NM_022574	0.00	0.484	-	-	confirmed	1	NA
GRIN2B	NM_000834	1.00	0.061	3	MONDO_0100038 : Definitive	confirmed	1	NA
HIVEP2	NM_006734	1.00	0.08	3	-	confirmed	1	Yes
HNRNPU	NM_031844	1.00	0.114	-	-	confirmed	1	Yes

Gene	NCBI Reference sequence	gnomAD PLI	gnomAD LOEUF	ClinGen HI score	ClinGen Gene-disease validity; MONDO DISEASE ID: Classification	DDG2P category	SFARI gene-score[a]	Overlap with Denovo-db NDD genes[b]
IRF2BPL	NM_024496	0.84	0.405	-	-	probable	1	Yes
KANSL1	NM_015443	1.00	0.238	3	-	confirmed	1	Yes
KAT6A	NM_006766	1.00	0.069	3	-	confirmed	2	Yes
KAT6B	NM_012330	1.00	0.128	3	-	confirmed	3	Yes
KCNQ2	NM_172106	1.00	0.158	3	MONDO_0018614 : Definitive	confirmed	2	NA
KIF11	NM_004523	1.00	0.121	3	-	confirmed	-	Yes
KMT2A	NM_001197104	1.00	0.065	3	-	confirmed	1	Yes
KMT2C	NM_170606	1.00	0.122	3	-	probable	1	Yes
KMT2D	NM_003482	1.00	0.103	3	-	confirmed	-	NA
KMT5B	NM_017635	1.00	0.078	-	-	confirmed	1	NA
MBD5	NM_018328	1.00	0.136	3	MONDO_0100038 : Moderate	probable	1	Yes
MED13L	NM_015335	1.00	0.064	3	-	probable	1	Yes
MEF2C	NM_001364333	0.02	0.608	3	MONDO_0100038 : Definitive	confirmed	1	Yes
MYT1L	NM_001303052	1.00	0.087	3	MONDO_0014678 : Definitive	confirmed	1	Yes
NAA15	NM_057175	1.00	0.28	2	-	probable	1	Yes
NBEA	NM_015678	1.00	0.074	3	MONDO_0100038 : Definitive	probable	1	NA
NCKAP1	NM_013436	1.00	0.044	3	-	possible	1	NA
NFIX	NM_001271043	1.00	0.149	-	-	confirmed	S	Yes
NIPBL	NM_133433	1.00	0.032	3	MONDO_0016033 : Definitive	confirmed	1	NA
NSD1	NM_022455	1.00	0.095	3	MONDO_0019349 : Definitive	confirmed	1	Yes
NSD2	NM_001042424	1.00	0.119	-	-	probable	2	NA
PHF21A	NM_001352027	1.00	0.08	2	-	probable	1	Yes
PHIP	NM_017934	1.00	0.113	2	-	confirmed	1	Yes
POGZ	NM_015100	1.00	0.119	3	MONDO_0014606 : Definitive	confirmed	1	Yes
PRR12	NM_020719	1.00	0.051	-	-	probable	1	Yes
PTEN	NM_000314	0.26	0.507	3	MONDO_0017623 : Definitive	confirmed	1	Yes
PURA	NM_005859	0.94	0.335	3	MONDO_0100038 : Definitive	confirmed	-	Yes
RAI1	NM_030665	1.00	0.116	3	MONDO_0008434 : Definitive	confirmed	1	Yes
SATB2	NM_001172509	1.00	0.091	3	-	confirmed	3	Yes
SCN1A	NM_006920	1.00	0.071	3	MONDO_0018214 : Definitive	confirmed	1	NA
SCN2A	NM_001040142	1.00	0.127	3	MONDO_0100038 : Definitive	confirmed	1	Yes
SETBP1	NM_015559	1.00	0.11	3	MONDO_0100038 : Definitive	confirmed	1	Yes
SETD1A	NM_014712	1.00	0.136	3	-	probable	1	NA
SETD2	NM_014159	1.00	0.206	2	-	probable	1	Yes
SETD5	NM_001080517	1.00	0.227	3	-	confirmed	1	Yes
SHANK2	NM_012309	0.00	0.458	2	MONDO_0100038 : Definitive	probable	1	NA

Gene	NCBI Reference sequence	gnomAD PLI	gnomAD LOEUF	ClinGen HI score	ClinGen Gene-disease validity; MONDO DISEASE ID: Classification	DDG2P category	SFARI gene-score[a]	Overlap with Denovo-db NDD genes[b]
SHANK3	NM_033517	1.00	0.123	3	MONDO_0011652 : Definitive	probable	1	Yes
SLC6A1	NM_003042	1.00	0.15	3	-	confirmed	1	Yes
SON	NM_138927	1.00	0.118	3	-	confirmed	1	Yes
SOX5	NM_006940	1.00	0.188	3	-	probable	1	Yes
STAG1	NM_005862	1.00	0.119	-	-	confirmed	S	NA
STXBP1	NM_001032221	1.00	0.086	3	MONDO_0020071 : Definitive	confirmed	1	Yes
SYNGAP1	NM_006772	1.00	0.052	3	MONDO_0100038 : Definitive	confirmed	1	Yes
TBR1	NM_006593	1.00	0.194	3	MONDO_0100038 : Definitive	confirmed	1	Yes
TCF20	NM_005650	1.00	0.095	3	-	confirmed	1	Yes
TCF4	NM_001083962	1.00	0.217	3	MONDO_0012589 : Definitive	confirmed	1	Yes
TLK2	NM_006852	1.00	0.114	-	-	confirmed	1	Yes
TNRC6B	NM_001162501	1.00	0.173	3	MONDO_0100038 : Definitive	probable	2	Yes
TRIO	NM_007118	1.00	0.139	3	-	probable	1	NA
TRIP12	NM_001348323	1.00	0.06	3	-	confirmed	1	Yes
TSC2	NM_000548	1.00	0.074	3	MONDO_0001734 : Definitive	confirmed	1	NA
USP7	NM_003470	1.00	0.061	-	-	probable	2	NA
WAC	NM_016628	1.00	0.084	3	-	confirmed	1	Yes
WASF1	NM_003931	1.00	0.215	-	-	probable	S	NA
WDFY3	NM_014991	1.00	0.071	-	-	probable	1	Yes
ZBTB18	NM_001278196	1.00	0.179	3	-	confirmed	-	Yes
ZEB2	NM_014795	1.00	0.107	3	MONDO_0009341 : Definitive	confirmed	-	NA
ZNF292	NM_015021	1.00	0.14	-	MONDO_0100038 : Definitive	probable	1	Yes
ZNF462	NM_021224	1.00	0.085	3	-	probable	1	NA

HI: haploinsufficiency

[a]SFARI gene score 1, 2, 3, and “S” represents High Confidence; Strong Candidate; Suggestive Evidence; and syndromic, respectively

[b]145 NDD genes with excess of LGDs (likely gene disruptive only, excluding missense) described in Coe et al., Nat. Genet. (2018)

LOEUF: gnomAD loss-of-function observed/expected upper bound fraction

MONDO: The Mondo Disease Ontology

Table S2. ICD-9 and ICD-10 diagnoses codes for neurodevelopmental psychiatric disorders and congenital anomalies[a]

A. Neurodevelopmental psychiatric disorders (NPD)		
	ICD-9 Codes	**ICD-10 Codes**
Intellectual Disabilities	317, 318.0, 318.1, 318.2, 319	F70, F71, F72, F73, F78, F79
Communication Disorders	315.3, 315.31, 315.32, 315.34, 315.35, 315.39, 307.9	F80.0, F80.1, F80.2, F80.4, F80.8, F80.81, F80.82, F80.89, F80.9,
Autism Spectrum Disorder	299.00, 299.80	F84.*
ADHD	314, 314.0, 314.00, 314.01, 314.1, 314.2, 314.8, 314.9	F90, F90.0, F90.1, F90.2, F90.8, F90.9
Specific Learning Disorder	315.0*, 315.1, 315.2	F81, F81.0, F81.2, F81.8, F81.81, F81.89, F81.9
Motor Disorders	315.4, 307.2, 307.20, 307.21, 307.22, 307.23, 307.3, 781.3	F82, F98.4, F95, F95.0, F95.1, F95.2, F95.8, F95.9, R27.0, R27.8, R27.9
Other Neurodevelopmental Disorders	315.5, 315.8, 315.9	F88, F89
Schizophrenia Spectrum and Other Psychotic Disorders	295.*, 297.*, 298.*, 301.22	F20.*, F21, F22, F23, F24, F25.*, F28, F29, F06.1
Bipolar and Related Disorders	296.0*, 296.1*, 296.4*, 296.5*, 296.6*, 296.7, 296.8*	F30.*, F31.*, F34.0
Depressive Disorders	296.2*, 296.3*, 311	F32.*, F33.*, F34.1
Anxiety Disorders	300.0, 300.00, 300.01, 300.02, 300.09, 300.2*, 309.21, 312.23	F40.*, F41.*, F93.0, F94.0
Obsessive-Compulsive and Related Disorders	300.3, 300.7, 312.39	F42.*, F45.22, F63.3
Epilepsy	345.*	G40.*
Cerebral Palsy	343.*	G80.*

*Indicates wildcard character (none or any number)
[a]ICD-9/10diagnosis codes have been described previously(2)

B. Congenital Anomaly (CA)		
	ICD-9 Codes	**ICD-10 Codes**
Central Nervous System	740.*, 741.*, 742.*	Q00.*, Q01.*, Q02, Q03.*, Q04.*, Q05.*, Q06.*, Q07.*

Cardiac	745.*, 746.*, 747..0, 747.1, 747.2, 747.3, 747.4*	Q20.*, Q21.*, Q22.*, Q23.*, Q24.*, Q25.*, Q26.*
Cleft lip/palate	749.*	Q35.*, Q36.*, Q37.*
Genital	752.*	Q50.*, Q51.*, Q52.*, Q53, Q53.1*, Q53.0*, Q53.2*, Q54.*, Q55.*, Q56.*
Renal/Urinary	753.*	Q60.*, Q61.*, Q62.*, Q63.*, Q64.*

*Indicates wildcard character (none or any number)
[a]ICD-9/10 diagnosis codes have been described previously (2)

Table S3. Characteristics of participants with or without a variant or recurrent CNV

Demography	Participants with a variant in NPD genes (n=312)	Participants with a recurrent CNV (n=708)	Participants without variants in NPD genes or recurrent CNV (n=89,577)
Sex			
Female	179 (57.3)	436 (61.6)	54,288 (60.6)
Age			
Median age (IQR) [range]	56.6 (42-72) [6-97]	54 (39-66) [5-96]	63 (48-74) [5-103]
Aged <18	16 (5.1%)	27 (3.8%)	795 (0.9%)
Aged >=18	296 (94.9%)	681 (96.2%)	88,782 (99.1%)
Race/ethnicity			
Non-Hispanic White	291 (93.3)	679 (95.9)	86,144 (96.2)
Black or African American	9 (2.9)	17 (2.4)	1,504 (1.7)
Asian	1 (0.3)	3 (0.4)	256 (0.3)
Hispanic or Latino	10 (3.2)	6 (0.8)	1,383 (1.5)
Other	0 (0.0)	3 (0.4)	146 (0.2)
Unknown	1 (0.3)	0 (0)	144 (0.2)

Table S4. Prevalence of pathogenic variants in each gene in DiscovEHR cohort

Gene	SNV/indel	NPD gene CNV	Total	Prevalence (%) in DiscovEHR N=90,595	Prevalence in the 1st or 2nd degree unrelated subset (n=63,632)
ADNP	1	-	1	0.001	0.002
ANK2	26	-	26	0.029	0.024
ANKRD11	2	-	2	0.002	0.003
ARID1B	1	-	1	0.001	0.002
ASH1L	13	-	13	0.014	0.011
ASXL3	4	17	21	0.023	0.014
AUTS2	-	1	1	0.001	0.002
BRPF1	4	-	4	0.004	0.006
CHD2	2	-	2	0.002	0.003
CHD7	3	-	3	0.003	0.005
CHD8	5	-	5	0.006	0.005
CIC	4	-	4	0.004	0.003
CLTC	5	-	5	0.006	0.008
CSDE1	1	-	1	0.001	0.002
CTCF	1	-	1	0.001	0.002
CUL3	2	-	2	0.002	0.003
DSCAM	15	-	15	0.017	0.019
EFTUD2	1	-	1	0.001	0.000
EHMT1	1	1	2	0.002	0.003
EP300	4	-	4	0.004	0.006
FOXP1	2	1	3	0.003	0.002
GIGYF1	15	-	15	0.017	0.017
GRIN2B	1	-	1	0.001	0.002
IRF2BPL	7	-	7	0.008	0.008
KANSL1	1	-	1	0.001	0.002
KCNQ2	1	-	1	0.001	0.002
KMT2A	1	-	1	0.001	0.000
KMT2C	13	1	14	0.015	0.014
KMT2D	4	-	4	0.004	0.005
KMT5B	1	-	1	0.001	0.002
MBD5	4	-	4	0.004	0.006
NAA15	6	-	6	0.007	0.008
NBEA	4	-	4	0.004	0.003
NCKAP1	2	-	2	0.002	0.003
NIPBL	2	-	2	0.002	0.003
NSD1	3	-	3	0.003	0.005

Gene	SNV/indel	NPD gene CNV	Total	Prevalence (%) in DiscovEHR N=90,595	Prevalence in the 1st or 2nd degree unrelated subset (n=63,632)
NSD2	2	-	2	0.002	0.003
PHIP	13	-	13	0.014	0.017
POGZ	5	-	5	0.006	0.005
PTEN	1	-	1	0.001	0.002
RAI1	2	-	2	0.002	0.002
SATB2	2	-	2	0.002	0.002
SCN1A	1	-	1	0.001	0.002
SCN2A	1	-	1	0.001	0.000
SETD1A	3	-	3	0.003	0.003
SETD2	5	-	5	0.006	0.005
SETD5	8	-	8	0.009	0.006
SHANK2	17	1	18	0.020	0.017
STXBP1	1	-	1	0.001	0.002
TCF20	1	-	1	0.001	0.002
TCF4	2	1	3	0.003	0.005
TLK2	1	-	1	0.001	0.002
TNRC6B	8	-	8	0.009	0.011
TRIO	17	-	17	0.019	0.019
TRIP12	1	-	1	0.001	0.002
TSC2	3	-	3	0.003	0.003
WAC	3	-	3	0.003	0.002
WASF1	-	1	1	0.001	0.002
WDFY3	16	-	16	0.018	0.016
ZNF292	11	-	11	0.012	0.013
ZNF462	2	-	2	0.002	0.000
Any NPD Gene	288	24	312	0.344	0.339

Table S5. Unique variants in NPD genes included in the study

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
20	50894171	CCTAA	C	*ADNP*	NM_001282531.3	c.539_542del	p.Val180GlyfsTer17	frameshift_variant	PVS1_strong, PM2	LP
4	113354311	CAGAA	C	*ANK2*	NM_001148.6	c.5697_5700del	p.Glu1899AspfsTer94	frameshift_variant	PVS1, PM2	LP
4	113354418	GA	G	*ANK2*	NM_001148.6	c.5805del	p.Lys1935AsnfsTer59	frameshift_variant	PVS1, PM2	LP
4	113355863	C	CCGAGATAG	*ANK2*	NM_001148.6	c.7249_7256dup	p.Val2420IlefsTer13	frameshift_variant	PVS1, PM2	LP
4	113356083	CT	C	*ANK2*	NM_001148.6	c.7467del	p.Pro2490LeufsTer10	frameshift_variant	PVS1, PM2	LP
4	113356619	AC	A	*ANK2*	NM_001148.6	c.8003del	p.Pro2668LeufsTer19	frameshift_variant	PVS1, PM2	LP
4	113356972	C	G	*ANK2*	NM_001148.6	c.8354C>G	p.Ser2785Ter	stop_gained	PVS1, PM2	LP
4	113357877	AC	A	*ANK2*	NM_001148.6	c.9263del	p.Pro3088GlnfsTer24	frameshift_variant	PVS1, PM2	LP
4	113358741	C	T	*ANK2*	NM_001148.6	c.10123C>T	p.Gln3375Ter	stop_gained	PVS1, PM2	LP
4	113359097	ACAGT	A	*ANK2*	NM_001148.6	c.10483_10486del	p.Ser3495ArgfsTer6	frameshift_variant	PVS1, PM2	LP
4	113363385	C	T	*ANK2*	NM_001148.6	c.10804C>T	p.Arg3602Ter	stop_gained	PVS1, PM2	LP
4	113369650	G	T	*ANK2*	NM_001148.6	c.11455G>T	p.Glu3819Ter	stop_gained	PVS1, PM2	LP
16	89279859	TC	T	*ANKRD11*	NM_013275.6	c.6682del	p.Glu2228LysfsTer109	frameshift_variant	PVS1, PM2	LP
16	89284375	CTT	C	*ANKRD11*	NM_013275.6	c.2165_2166del	p.Lys722ArgfsTer19	frameshift_variant	PVS1, PM2	LP
6	157181180	T	C	*ARID1B*	NM_001346813.1	c.3465+2T>C		splice_donor_variant	PVS1_moderate, PS2_moderate, PM2	LP
1	155338161	G	A	*ASH1L*	NM_018489.3	c.8731C>T	p.Arg2911Ter	stop_gained	PVS1_strong, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
1	155338203	CCT	C	*ASH1L*	NM_018489.3	c.8687_8688del	p.Glu2896GlyfsTer3	frameshift_variant	PVS1_strong, PM2	LP
1	155338208	GT	G	*ASH1L*	NM_018489.3	c.8683del	p.Thr2895GlnfsTer44	frameshift_variant	PVS1_strong, PM2	LP
1	155338208	GT	GTT	*ASH1L*	NM_018489.3	c.8683dup	p.Thr2895AsnfsTer5	frameshift_variant	PVS1_strong, PM2	LP
1	155521360	G	A	*ASH1L*	NM_018489.3	c.160C>T	p.Arg54Ter	stop_gained	PVS1_strong, PM2	LP
18	33671845	C	T	*ASXL3*	NM_030632.3	c.694C>T	p.Arg232Ter	stop_gained	PVS1, PM2	LP
18	33683514	A	AC	*ASXL3*	NM_030632.3	c.827dup	p.Gln277SerfsTer16	frameshift_variant	PVS1, PM2	LP
18	33739907	C	T	*ASXL3*	NM_030632.3	c.2503C>T	p.Gln835Ter	stop_gained	PVS1, PM2	LP
18	33744703	C	T	*ASXL3*	NM_030632.3	c.4855C>T	p.Gln1619Ter	stop_gained	PVS1_strong, PM2	LP
3	9734293	A	AC	*BRPF1*	NM_004634.2	c.158dup	p.Gln54ThrfsTer35	frameshift_variant	PVS1, PM2	LP
3	9739314	C	A	*BRPF1*	NM_004634.2	c.915C>A	p.Tyr305Ter	stop_gained	PVS1, PM2	LP
3	9743701	C	CAA	*BRPF1*	NM_004634.2	c.2418_2419insAA	p.Arg807AsnfsTer6	frameshift_variant	PVS1, PM2	LP
3	9745904	C	T	*BRPF1*	NM_004634.2	c.3280C>T	p.Arg1094Ter	stop_gained	PVS1, PM2	LP
15	92981420	C	G	*CHD2*	NM_001271.4	c.3029C>G	p.Ser1010Ter	stop_gained	PVS1, PM2	LP
15	92997041	AG	A	*CHD2*	NM_001271.4	c.3682del	p.Glu1228SerfsTer21	frameshift_variant	PVS1, PM2	LP
8	60865506	ACT	A	*CHD7*	NM_017780.4	c.8570_8571del	p.Ser2857Ter	frameshift_variant	PVS1_strong, PM2	LP
8	60865509	C	CT	*CHD7*	NM_017780.4	c.8571dup	p.Glu2858Ter	frameshift_variant	PVS1_strong, PM2	LP
14	21386177	C	T	*CHD8*	NM_001170629.1	c.7183-1G>A		splice_acceptor_variant	PVS1_strong, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
14	21397857	G	A	*CHD8*	NM_001170629.1	c.5017C>T	p.Arg1673Ter	stop_gained	PVS1, PM2	LP
14	21402425	T	A	*CHD8*	NM_001170629.1	c.3793A>T	p.Arg1265Ter	stop_gained	PVS1, PM2	LP
14	21428046	C	CG	*CHD8*	NM_001170629.1	c.1423dup	p.Arg475ProfsTer3	frameshift_variant	PVS1, PM2	LP
16	67610883	TAAAG	T	*CTCF*	NM_006565.4	c.58_62del	p.Lys20GlufsTer28	frameshift_variant	PVS1_strong, PM2	LP
2	224503813	G	A	*CUL3*	NM_003590.5	c.1216C>T	p.Gln406Ter	stop_gained	PVS1, PM2	LP
2	224535588	AT	A	*CUL3*	NM_003590.5	c.317del	p.Asn106IlefsTer12	frameshift_variant	PVS1, PM2	LP
21	40078860	T	TA	*DSCAM*	NM_001389.5	c.4537dup	p.Tyr1513LeufsTer25	frameshift_variant	PVS1, PM2	LP
21	40080150	AC	A	*DSCAM*	NM_001389.5	c.4420+1del		splice_donor_variant	PVS1, PM2	LP
21	40080152	C	A	*DSCAM*	NM_001389.5	c.4420G>T	p.Glu1474Ter	stop_gained&splice_region_variant	PVS1, PM2	LP
21	40083982	G	GT	*DSCAM*	NM_001389.5	c.4156dup	p.Thr1386AsnfsTer36	frameshift_variant	PVS1, PM2	LP
21	40087279	G	A	*DSCAM*	NM_001389.5	c.3859C>T	p.Arg1287Ter	stop_gained	PVS1, PM2	LP
21	40093769	C	A	*DSCAM*	NM_001389.5	c.3802G>T	p.Gly1268Ter	stop_gained	PVS1, PM2	LP
21	40134010	C	T	*DSCAM*	NM_001389.5	c.3407-1G>A		splice_acceptor_variant	PVS1, PM2	LP
21	40142587	ACT	A	*DSCAM*	NM_001389.5	c.3375_3376del	p.Arg1125SerfsTer14	frameshift_variant	PVS1, PM2	LP
21	40142687	CG	C	*DSCAM*	NM_001389.5	c.3276del	p.Glu1093LysfsTer6	frameshift_variant	PVS1, PM2	LP
21	40179053	G	A	*DSCAM*	NM_001389.5	c.2821C>T	p.Gln941Ter	stop_gained	PVS1, PM2	LP
21	40276151	CCTTGCAGAGGT	C	*DSCAM*	NM_001389.5	c.2291_2301del	p.Tyr764CysfsTer18	frameshift_variant	PVS1, PM2	LP
21	40347731	G	C	*DSCAM*	NM_001389.5	c.1149C>G	p.Tyr383Ter	stop_gained	PVS1, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
17	44876034	TG	T	*EFTUD2*	NM_004247.4	c.768del	p.Cys256Ter	frameshift_variant	PVS1, PM2, PS2	P
9	137752379	G	T	*EHMT1*	NM_024757.5	c.1219G>T	p.Glu407Ter	stop_gained	PVS1, PM2	LP
22	41125976	C	A	*EP300*	NM_001429.4	c.842C>A	p.Ser281Ter	stop_gained	PVS1, PM2	LP
22	41166633	C	T	*EP300*	NM_001429.4	c.3841C>T	p.Arg1281Ter	stop_gained	PVS1, PM2	LP
22	41168501	TC	T	*EP300*	NM_001429.4	c.3928del	p.Leu1310Ter	frameshift_variant	PVS1, PM2	LP
22	41168587	GC	G	*EP300*	NM_001429.4	c.4014del	p.Met1339Ter	frameshift_variant	PVS1, PM2	LP
3	70972637	CA	C	*FOXP1*	NM_032682.5	c.1569del	p.Phe523LeufsTer4	frameshift_variant	PVS1, PM2	LP
3	71046941	C	T	*FOXP1*	NM_032682.5	c.664+1G>A		splice_donor_variant	PVS1, PM2	LP
12	13563307	G	A	*GRIN2B*	NM_000834.4	c.3931C>T	p.Gln1311Ter	stop_gained	PVS1_strong, PM2	LP
17	46032106	G	A	*KANSL1*	NM_015443.3	c.3031C>T	p.Arg1011Ter	stop_gained	PVS1, PM2	LP
20	63414901	A	C	*KCNQ2*	NM_172106.2	c.1471+2T>G		splice_donor_variant	PVS1, PM2	LP
11	118519731	G	T	*KMT2A*	NM_001197104.1	c.11260G>T	p.Glu3754Ter	stop_gained	PVS1, PM2	LP
7	152145160	CA	C	*KMT2C*	NM_170606.3	c.14166del	p.Phe4722LeufsTer3	frameshift_variant	PVS1, PM2	LP
7	152145252	C	CCGTAT	*KMT2C*	NM_170606.3	c.14070_14074dup	p.Gly4692AspfsTer26	frameshift_variant	PVS1, PM2	LP
7	152145254	GTA	G	*KMT2C*	NM_170606.3	c.14071_14072del	p.Tyr4691ArgfsTer13	frameshift_variant	PVS1, PM2	LP
7	152148803	C	CA	*KMT2C*	NM_170606.3	c.13123dup	p.Cys4375LeufsTer2	frameshift_variant	PVS1, PM2	LP
7	152151440	A	G	*KMT2C*	NM_170606.3	c.12666+2T>C		splice_donor_variant	PVS1, PM2	LP
7	152152909	G	A	*KMT2C*	NM_170606.3	c.12322C>T	p.Arg4108Ter	stop_gained	PVS1, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
7	152156009	TGGGGTCTGAAG	T	*KMT2C*	NM_170606.3	c.11850_11860del	p.Phe3951GlyfsTer16	frameshift_variant	PVS1, PM2	LP
7	152156034	GT	G	*KMT2C*	NM_170606.3	c.11835del	p.Lys3945AsnfsTer42	frameshift_variant	PVS1, PM2	LP
7	152156059	T	C	*KMT2C*	NM_170606.3	c.11813-2A>G		splice_acceptor_variant	PVS1, PM2	LP
7	152162957	AG	A	*KMT2C*	NM_170606.3	c.10619del	p.Ser3540LeufsTer10	frameshift_variant	PVS1, PM2	LP
7	152163058	G	A	*KMT2C*	NM_170606.3	c.10519C>T	p.Arg3507Ter	stop_gained	PVS1, PM2	LP
7	152163652	G	A	*KMT2C*	NM_170606.3	c.9925C>T	p.Gln3309Ter	stop_gained	PVS1, PM2	LP
7	152176412	CCA	C	*KMT2C*	NM_170606.3	c.9039_9040del	p.Gly3014AlafsTer10	frameshift_variant	PVS1, PM2	LP
7	152315139	TTCTC	T	*KMT2C*	NM_170606.3	c.585_588del	p.Arg196LysfsTer9	frameshift_variant&splice_region_variant	PVS1, PM2	LP
12	49022844	T	TA	*KMT2D*	NM_003482.3	c.16083dup	p.Lys5362Ter	frameshift_variant	PVS1, PM2	LP
12	49027256	G	A	*KMT2D*	NM_003482.3	c.14710C>T	p.Arg4904Ter	stop_gained	PVS1, PM2	LP
12	49038315	AG	A	*KMT2D*	NM_003482.3	c.9040del	p.Leu3014TrpfsTer15	frameshift_variant	PVS1, PM2	LP
12	49038613	G	A	*KMT2D*	NM_003482.3	c.8743C>T	p.Arg2915Ter	stop_gained	PVS1, PM2	LP
11	68175166	CCTTT	C	*KMT5B*	NM_017635.5	c.391_394del	p.Lys131GlufsTer6	frameshift_variant	PVS1_strong, PM2	LP
2	148483194	G	GC	*MBD5*	NM_018328.4	c.2604dup	p.Val869ArgfsTer20	frameshift_variant	PVS1_strong, PM2	LP
2	148485939	C	T	*MBD5*	NM_018328.4	c.3043C>T	p.Gln1015Ter	stop_gained	PVS1_strong, PM2	LP
2	148489973	AT	A	*MBD5*	NM_018328.4	c.3643del	p.Leu1216Ter	frameshift_variant	PVS1_strong, PM2	LP
2	148490571	C	T	*MBD5*	NM_018328.4	c.4240C>T	p.Gln1414Ter	stop_gained	PVS1_strong, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
4	139336946	CAT	C	*NAA15*	NM_057175.5	c.239_240del	p.His80ArgfsTer17	frameshift_variant	PVS1, PM2	LP
4	139341049	C	T	*NAA15*	NM_057175.5	c.382C>T	p.Arg128Ter	stop_gained	PVS1, PM2	LP
4	139344318	CT	C	*NAA15*	NM_057175.5	c.672del	p.Ala225LeufsTer2	frameshift_variant	PVS1, PM2	LP
4	139349476	C	T	*NAA15*	NM_057175.5	c.706C>T	p.Gln236Ter	stop_gained	PVS1, PM2	LP
4	139354024	A	C	*NAA15*	NM_057175.5	c.1015-2A>C		splice_acceptor_variant	PVS1, PM2	LP
13	35159100	A	T	*NBEA*	NM_015678.4	c.2929A>T	p.Lys977Ter	stop_gained	PVS1, PM2	LP
13	35654928	T	A	*NBEA*	NM_015678.4	c.8046T>A	p.Cys2682Ter	stop_gained	PVS1_strong, PM2	LP
13	35667499	C	T	*NBEA*	NM_015678.4	c.8527C>T	p.Arg2843Ter	stop_gained	PVS1_strong, PM2	LP
13	35667505	C	T	*NBEA*	NM_015678.4	c.8533C>T	p.Arg2845Ter	stop_gained	PVS1_strong, PM2	LP
2	182957598	T	C	*NCKAP1*	NM_013436.5	c.1882-2A>G		splice_acceptor_variant	PVS1, PM2	LP
2	182989028	AC	A	*NCKAP1*	NM_013436.5	c.947+1del		splice_donor_variant	PVS1, PM2	LP
5	37064721	G	A	*NIPBL*	NM_133433.4	c.8244G>A	p.Trp2748Ter	stop_gained	PVS1_strong, PM2, PP1	LP
5	37064854	C	T	*NIPBL*	NM_133433.4	c.8377C>T	p.Arg2793Ter	stop_gained	PVS1_strong, PM2, PP1	LP
5	177294239	C	T	*NSD1*	NM_022455.4	c.6871C>T	p.Gln2291Ter	stop_gained	PVS1_strong, PM2	LP
5	177294949	G	A	*NSD1*	NM_022455.4	c.7581G>A	p.Trp2527Ter	stop_gained	PVS1_strong, PM2	LP
5	177295136	C	CT	*NSD1*	NM_022455.4	c.7768_7769insT	p.Pro2590LeufsTer12	frameshift_variant	PVS1_strong, PM2	LP
4	1900841	C	T	*NSD2*	NM_001042424.3	c.187C>T	p.Gln63Ter	stop_gained	PVS1, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
4	1939695	C	T	*NSD2*	NM_001042424.3	c.1798C>T	p.Arg600Ter	stop_gained	PVS1, PM2	LP
6	78954945	T	A	*PHIP*	NM_017934.7	c.3922A>T	p.Arg1308Ter	stop_gained	PVS1, PM2	LP
6	78963242	AG	A	*PHIP*	NM_017934.7	c.3389del	p.Pro1130LeufsTer4	frameshift_variant	PVS1, PM2	LP
6	78970138	A	T	*PHIP*	NM_017934.7	c.3033T>A	p.Tyr1011Ter	stop_gained	PVS1, PM2	LP
6	78970815	GGAT	G	*PHIP*	NM_017934.7	c.2959_2962del	p.Asn987ProfsTer17	frameshift_variant	PVS1, PM2	LP
6	78978627	G	A	*PHIP*	NM_017934.7	c.2854C>T	p.Arg952Ter	stop_gained	PVS1, PM2	LP
6	78990898	TG	T	*PHIP*	NM_017934.7	c.2288del	p.Pro763GlnfsTer50	frameshift_variant	PVS1, PM2	LP
6	78990938	GTCTT	G	*PHIP*	NM_017934.7	c.2245_2248del	p.Lys749LeufsTer11	frameshift_variant	PVS1, PM2	LP
6	79001932	G	A	*PHIP*	NM_017934.7	c.1846C>T	p.Gln616Ter	stop_gained	PVS1, PM2	LP
6	79025555	CA	C	*PHIP*	NM_017934.7	c.886del	p.Cys296ValfsTer12	frameshift_variant	PVS1, PM2	LP
6	79042864	TC	T	*PHIP*	NM_017934.7	c.578del	p.Arg193GlnfsTer15	frameshift_variant	PVS1, PM2	LP
6	79042911	G	A	*PHIP*	NM_017934.7	c.532C>T	p.Arg178Ter	stop_gained	PVS1, PM2	LP
6	79043004	C	T	*PHIP*	NM_017934.7	c.440-1G>A		splice_acceptor_variant	PVS1, PM2	LP
6	79060767	G	A	*PHIP*	NM_017934.7	c.241C>T	p.Arg81Ter	stop_gained	PVS1, PM2	LP
1	151423514	G	A	*POGZ*	NM_015100.4	c.1561C>T	p.Gln521Ter	stop_gained	PVS1, PM2	LP
1	151427985	G	A	*POGZ*	NM_015100.4	c.916C>T	p.Arg306Ter	stop_gained	PVS1, PM2	LP
10	87933147	C	T	*PTEN*	NM_000314.7	c.389C>T	p.Arg130Ter	stop_gained	PVS1, PM2	LP
17	17795085	A	AC	*RAI1*	NM_030665.4	c.2140dup	p.Leu714ProfsTer14	frameshift_variant	PVS1, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
17	17796391	C	G	*RAI1*	NM_030665.4	c.3443C>G	p.Ser1148Ter	stop_gained	PVS1, PM2	LP
2	199308842	TC	T	*SATB2*	NM_001172509.2	c.1657del	p.Asp553MetfsTer71	frameshift_variant	PVS1, PM2	LP
2	199455954	TG	T	*SATB2*	NM_001172509.2	c.83del	p.Pro28GlnfsTer2	frameshift_variant	PVS1, PM2	LP
2	165991279	TG	T	*SCN1A*	NM_006920.5	c.5962del	p.Gln1988LysfsTer13	frameshift_variant	PVS1_strong, PM2	LP
2	165388682	C	T	*SCN2A*	NM_001040142.2	c.4876C>T	p.Arg1626Ter	stop_gained	PS2, PM2	LP
16	30965117	T	TC	*SETD1A*	NM_014712.3	c.1375_1376insC	p.Arg461ProfsTer18	frameshift_variant	PVS1, PM2	LP
16	30967499	A	G	*SETD1A*	NM_014712.3	c.2683-2A>G		splice_acceptor_variant	PVS1, PM2	LP
3	47056931	G	A	*SETD2*	NM_014159.6	c.6853C>T	p.Gln2285Ter	stop_gained	PVS1_strong, PM2	LP
3	47056937	G	A	*SETD2*	NM_014159.6	c.6847C>T	p.Gln2283Ter	stop_gained	PVS1_strong, PM2	LP
3	47086237	C	CT	*SETD2*	NM_014159.6	c.5354dup	p.Leu1786AlafsTer3	frameshift_variant	PVS1, PM2	LP
3	47106044	G	A	*SETD2*	NM_014159.6	c.4792C>T	p.Arg1598Ter	stop_gained	PVS1, PM2	LP
3	47124421	TTC	T	*SETD2*	NM_014159.6	c.213_214del	p.Lys72GlyfsTer15	frameshift_variant	PVS1, PM2	LP
3	9433933	C	T	*SETD5*	NM_001080517.2	c.160C>T	p.Arg54Ter	stop_gained	PVS1, PM2	LP
3	9434437	CT	C	*SETD5*	NM_001080517.2	c.283del	p.Cys95ValfsTer18	frameshift_variant	PVS1, PM2	LP
3	9453775	C	T	*SETD5*	NM_001080517.2	c.2383C>T	p.Gln795Ter	stop_gained	PVS1, PM2	LP
3	9474540	CCCTCTCCCACATGGGAGAGT	C	*SETD5*	NM_001080517.2	c.3590_3609del	p.Pro1197GlnfsTer24	frameshift_variant	PVS1, PM2	LP
11	70485387	G	A	*SHANK2*	NM_012309.4	c.4906C>T	p.Arg1636Ter	stop_gained	PVS1, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
11	70490294	G	A	*SHANK2*	NM_012309.4	c.2533C>T	p.Arg845Ter	stop_gained	PVS1, PM2	LP
11	71092580	C	A	*SHANK2*	NM_012309.4	c.754G>T	p.Glu252Ter	stop_gained	PVS1, PM2	LP
11	71094563	G	A	*SHANK2*	NM_012309.4	c.718C>T	p.Arg240Ter	stop_gained	PVS1, PM2	LP
11	71113316	G	A	*SHANK2*	NM_012309.4	c.460C>T	p.Gln154Ter	stop_gained	PVS1, PM2	LP
11	71113361	G	A	*SHANK2*	NM_012309.4	c.415C>T	p.Arg139Ter	stop_gained	PVS1, PM2	LP
11	71118833	AG	A	*SHANK2*	NM_012309.4	c.406del	p.Leu136TrpfsTer29	frameshift_variant	PVS1, PM2	LP
11	71118917	C	CTGCA A	*SHANK2*	NM_012309.4	c.322_323insTTG CA	p.Ser108IlefsTer59	frameshift_variant	PVS1, PM2	LP
11	71118924	GC	G	*SHANK2*	NM_012309.4	c.315del	p.Gln105HisfsTer60	frameshift_variant	PVS1, PM2	LP
9	127673250	C	T	*STXBP1*	NM_001032221.3	c.1099C>T	p.Arg367Ter	stop_gained	PVS1, PM2	LP
22	42213768	G	GAC	*TCF20*	NM_005650.3	c.1536_1537dup	p.Ser513CysfsTer8	frameshift_variant	PVS1, PM2	LP
18	55279573	G	GAT	*TCF4*	NM_001083962.1	c.632_633insAT	p.Phe211LeufsTer24	frameshift_variant	PVS1, PM2	LP
18	55403453	C	A	*TCF4*	NM_001083962.1	c.369+1G>T		splice_donor_variant	PVS1, PM2	LP
17	62596650	AT	A	*TLK2*	NM_006852.3	c.1528del	p.Tyr510ThrfsTer13	frameshift_variant	PVS1, PM2	LP
22	40265582	AAG	A	*TNRC6B*	NM_001162501.2	c.1361_1362del	p.Glu454GlyfsTer17	frameshift_variant	PVS1, PM2	LP
22	40265617	C	T	*TNRC6B*	NM_001162501.2	c.1387C>T	p.Gln463Ter	stop_gained	PVS1, PM2	LP
22	40265633	CA	C	*TNRC6B*	NM_001162501.2	c.1409del	p.Asn470MetfsTer33	frameshift_variant	PVS1, PM2	LP
22	40265717	G	A	*TNRC6B*	NM_001162501.2	c.1487G>A	p.Trp496Ter	stop_gained	PVS1, PM2	LP
22	40265813	C	CA	*TNRC6B*	NM_001162501.2	c.1586dup	p.Asn529LysfsTer6	frameshift_variant	PVS1, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
22	40280030	C	T	*TNRC6B*	NM_001162501.2	c.3298C>T	p.Arg1100Ter	stop_gained	PVS1, PM2	LP
22	40280144	G	A	*TNRC6B*	NM_001162501.2	c.3411+1G>A		splice_donor_variant	PVS1, PM2	LP
22	40315470	CAG	C	*TNRC6B*	NM_001162501.2	c.4867_4868del	p.Arg1623GlyfsTer17	frameshift_variant	PVS1, PM2	LP
5	14270830	C	T	*TRIO*	NM_007118.4	c.163C>T	p.Arg55Ter	stop_gained	PVS1, PM2	LP
5	14336650	C	T	*TRIO*	NM_007118.4	c.1969C>T	p.Gln657Ter	stop_gained	PVS1, PM2	LP
5	14336699	C	A	*TRIO*	NM_007118.4	c.2018C>A	p.Ser673Ter	stop_gained	PVS1, PM2	LP
5	14394050	C	T	*TRIO*	NM_007118.4	c.4231C>T	p.Arg1411Ter	stop_gained	PVS1, PM2	LP
5	14398986	C	CAG	*TRIO*	NM_007118.4	c.4530_4531insAG	p.Leu1511SerfsTer7	frameshift_variant	PVS1, PM2	LP
5	14405925	GGGAGCCCTGAA	G	*TRIO*	NM_007118.4	c.4802_4812del	p.Leu1601HisfsTer5	frameshift_variant	PVS1, PM2	LP
5	14461069	T	TC	*TRIO*	NM_007118.4	c.5257dup	p.Leu1753ProfsTer136	frameshift_variant	PVS1, PM2	LP
5	14465610	TG	T	*TRIO*	NM_007118.4	c.5734del	p.Ala1912GlnfsTer6	frameshift_variant	PVS1, PM2	LP
5	14476937	A	T	*TRIO*	NM_007118.4	c.6127A>T	p.Lys2043Ter	stop_gained	PVS1, PM2	LP
5	14479312	G	GA	*TRIO*	NM_007118.4	c.6206dup	p.His2070AlafsTer8	frameshift_variant	PVS1, PM2	LP
5	14485209	C	CA	*TRIO*	NM_007118.4	c.6800dup	p.Asn2267LysfsTer174	frameshift_variant	PVS1, PM2	LP
5	14487650	TC	T	*TRIO*	NM_007118.4	c.7027del	p.Gln2343SerfsTer70	frameshift_variant	PVS1, PM2	LP
5	14487816	C	CCGAGG	*TRIO*	NM_007118.4	c.7188_7189insCGAGG	p.Ala2397ArgfsTer18	frameshift_variant	PVS1, PM2	LP
5	14492711	C	T	*TRIO*	NM_007118.4	c.7777C>T	p.Arg2593Ter	stop_gained	PVS1, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
5	14498579	C	A	*TRIO*	NM_007118.4	c.8271C>A	p.Tyr2757Ter	stop_gained	PVS1, PM2	LP
2	229880040	G	A	*TRIP12*	NM_001348323.1	c.40C>T	p.Arg14Ter	stop_gained	PVS1, PM2, PS2	P
16	2077613	A	AC	*TSC2*	NM_000548.5	c.2853_2854insC	p.Lys954GlnfsTer6	frameshift_variant	PVS1, PM2	LP
16	2080179	C	T	*TSC2*	NM_000548.5	c.3412C>T	p.Arg1138Ter	stop_gained	PVS1, PM2	LP
16	2081777	CCT	C	*TSC2*	NM_000548.5	c.3796_3797del	p.Leu1266AlafsTer55	frameshift_variant	PVS1, PM2	LP
10	28595945	C	T	*WAC*	NM_016628.4	c.823C>T	p.Gln275Ter	stop_gained	PVS1, PM2	LP
10	28610703	T	TA	*WAC*	NM_016628.4	c.1171dup	p.Thr391AsnfsTer15	frameshift_variant	PVS1, PM2	LP
10	28611772	A	T	*WAC*	NM_016628.4	c.1289-2A>T		splice_acceptor_variant	PVS1, PM2	LP
4	84690535	C	A	*WDFY3*	NM_014991.4	c.9334G>T	p.Glu3112Ter	stop_gained	PVS1, PM2	LP
4	84692993	G	A	*WDFY3*	NM_014991.4	c.8941C>T	p.Arg2981Ter	stop_gained	PVS1, PM2	LP
4	84693005	G	A	*WDFY3*	NM_014991.4	c.8929C>T	p.Arg2977Ter	stop_gained	PVS1, PM2	LP
4	84708953	G	A	*WDFY3*	NM_014991.4	c.8173C>T	p.Gln2725Ter	stop_gained	PVS1, PM2	LP
4	84709327	A	T	*WDFY3*	NM_014991.4	c.8063T>A	p.Leu2688Ter	stop_gained	PVS1, PM2	LP
4	84716910	T	A	*WDFY3*	NM_014991.4	c.7861A>T	p.Arg2621Ter	stop_gained	PVS1, PM2	LP
4	84753848	C	T	*WDFY3*	NM_014991.4	c.5588G>A	p.Trp1863Ter	stop_gained	PVS1, PM2	LP
4	84765852	G	A	*WDFY3*	NM_014991.4	c.5146C>T	p.Gln1716Ter	stop_gained	PVS1, PM2	LP
4	84772922	G	A	*WDFY3*	NM_014991.4	c.4762C>T	p.Gln1588Ter	stop_gained	PVS1, PM2	LP
4	84774979	TC	T	*WDFY3*	NM_014991.4	c.4594del	p.Glu1532LysfsTer8	frameshift_variant&splice_region_variant	PVS1, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
4	84780123	TC	T	*WDFY3*	NM_014991.4	c.4349del	p.Arg1450LysfsTer26	frameshift_variant	PVS1, PM2	LP
4	84831513	C	T	*WDFY3*	NM_014991.4	c.669G>A	p.Trp223Ter	stop_gained	PVS1, PM2	LP
9	106927627	C	T	*ZNF462*	NM_021224.6	c.3715C>T	p.Arg1239Ter	stop_gained	PVS1, PM2	LP
9	106929175	C	T	*ZNF462*	NM_021224.6	c.5263C>T	p.Gln1755Ter	stop_gained	PVS1, PM2	LP
17	59681478	G	GGT	*CLTC*	NM_004859.3	c.3249+1_3249+2dup		frameshift_variant&splice_region_variant	PVS1, PM2	LP
17	59683936	CTG	C	*CLTC*	NM_004859.3	c.4388_4389del	p.Val1463GlufsTer2	frameshift_variant	PVS1, PM2	LP
17	59685223	C	G	*CLTC*	NM_004859.3	c.4602C>G	p.Tyr1534Ter	stop_gained	PVS1, PM2	LP
1	114730525	C	CT	*CSDE1*	NM_001007553.3	c.1173dup	p.Glu392ArgfsTer7	frameshift_variant	PVS1, PM2, PS2	P
14	77026542	C	CA	*IRF2BPL*	NM_024496.4	c.1250dup	p.Leu417PhefsTer6	frameshift_variant	PVS1_strong, PM2	LP
14	77026685	T	TTG	*IRF2BPL*	NM_024496.4	c.1107_1108insCA	p.Ser370GlnfsTer49	frameshift_variant	PVS1_strong, PM2	LP
14	77026840	G	T	*IRF2BPL*	NM_024496.4	c.953C>A	p.Ser318Ter	stop_gained	PVS1_strong, PM2	LP
14	77027721	C	T	*IRF2BPL*	NM_024496.4	c.72G>A	p.Trp24Ter	stop_gained	PVS1_strong, PM2	LP
6	87233385	TCAAA	T	*ZNF292*	NM_015021.3	c.604_608del	p.Thr202SerfsTer14	frameshift_variant	PVS1, PM2	LP
6	87233499	C	G	*ZNF292*	NM_015021.3	c.713C>G	p.Ser238Ter	stop_gained	PVS1, PM2	LP
6	87255301	C	T	*ZNF292*	NM_015021.3	c.1672C>T	p.Arg558Ter	stop_gained	PVS1_strong, PM2	LP
6	87257020	A	T	*ZNF292*	NM_015021.3	c.3391A>T	p.Lys1131Ter	stop_gained	PVS1_strong, PM2	LP
6	87258043	C	T	*ZNF292*	NM_015021.3	c.4414C>T	p.Arg1472Ter	stop_gained	PVS1_strong, PM2	LP

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
6	87258737	AAG	A	*ZNF292*	NM_015021.3	c.5112_5113del	p.Asn1705PhefsTer6	frameshift_variant	PVS1_strong, PM2	LP
6	87259787	CAGA A	C	*ZNF292*	NM_015021.3	c.6159_6162del	p.Glu2054ValfsTer5	frameshift_variant	PVS1_strong, PM2, PS2	P
19	42292730	G	GT	*CIC*	NM_001304815.1	c.6067_6068insT	p.Ala2023ValfsTer37	frameshift_variant	PVS1, PM2, PS2	P
19	42293798	G	GC	*CIC*	NM_001304815.1	c.6734dup	p.Pro2246AlafsTer7	frameshift_variant	PVS1, PM2	LP
7	100682700	CTT	C	*GIGYF1*	NM_022574.5	c.2488_2489del	p.Lys830GlufsTer38	frameshift_variant	PVS1, PM2	LP
7	100683564	G	A	*GIGYF1*	NM_022574.5	c.2038C>T	p.Gln680Ter	stop_gained	PVS1, PM2	LP
7	100684867	G	A	*GIGYF1*	NM_022574.5	c.1318C>T	p.Gln440Ter	stop_gained	PVS1, PM2	LP
7	100685377	C	CG	*GIGYF1*	NM_022574.5	c.1158dup	p.Glu387ArgfsTer13	frameshift_variant	PVS1, PM2	LP
7	100685452	G	A	*GIGYF1*	NM_022574.5	c.1084C>T	p.Gln362Ter	stop_gained	PVS1, PM2	LP
7	100686367	A	AAATC	*GIGYF1*	NM_022574.5	c.757_760dup	p.Leu254Ter	stop_gained&frameshift_variant	PVS1, PM2	LP
7	100686727	CCT	C	*GIGYF1*	NM_022574.5	c.614_615del	p.Gln205ArgfsTer27	frameshift_variant	PVS1, PM2	LP
7	100686749	C	CCAGT T	*GIGYF1*	NM_022574.5	c.589_593dup	p.Trp198Ter	stop_gained&frameshift_variant	PVS1, PM2	LP
7	100686817	G	A	*GIGYF1*	NM_022574.5	c.526C>T	p.Arg176Ter	stop_gained&splice_region_variant	PVS1, PM2	LP
7	100686821	T	G	*GIGYF1*	NM_022574.5	c.524-2A>C		splice_acceptor_variant	PVS1, PM2	LP
7	100687323	G	A	*GIGYF1*	NM_022574.5	c.457C>T	p.Gln153Ter	stop_gained	PVS1, PM2	LP
7	100688019	G	A	*GIGYF1*	NM_022574.5	c.127C>T	p.Arg43Ter	stop_gained	PVS1_strong, PM2	LP
18	33607593-33746595	-	-	*ASXL3*	NM_030632.3	-	-	deletion	Overlap with HI gene	P

Chr	Pos	ref	alt	Gene	Refseq	cDNA	Protein	Consequence	ACMG evidence	ACMG classification
18	33731967-33746596	-	-	*ASXL3*	NM_030632.3	-	-	deletion	Overlap with HI gene	P
7	70118131-70134572	-	-	*AUTS2*	NM_015570.3	-	-	deletion	PVS1, PM2	LP
9	137710966-137762821	-	-	*EHMT1*	NM_024757.5	-	-	deletion	Overlap with HI gene	P
3	70959246-71542089	-	-	*FOXP1*	NM_032682.5	-	-	deletion	Overlap with HI gene	P
7	152187261-152263131	-	-	*KMT2C*	NM_170606.3	-	-	deletion	Overlap with HI gene	P
11	70659827-71094688	-	-	*SHANK2*	NM_012309.4	-	-	deletion	PVS1, PM2	LP
18	55228224-55234684	-	-	*TCF4*	NM_001083962.1	-	-	deletion	Overlap with HI gene	P
6	110100521-110127602	-	-	*WASF1*	NM_003931.3	-	-	deletion	PVS1, PM2	LP

Chr: Chromosome; Pos: genomic position (GRCh38); Ref: Reference; Alt: alternate sequence; LP: Likely pathogenic; P: Pathogenic; HI: Haploinsufficient (ClinGen dosage sensitivity curated genes with HI score = 3)

Table S6. Neuropsychiatric developmental disorder diagnoses by gene

Gene	Total	ID	ASD	ADHD	EP	SCZ	BPD	CD	SLD	CP	OCD	OND	MD	Any NPD	Percent Affected
ADNP	1	1	1	0	0	0	0	0	0	0	0	1	0	1	100.0
ANK2	26	0	0	2	4	1	4	1	1	0	1	0	2	8	30.8
ANKRD11	2	0	0	1	0	0	0	1	0	0	0	0	0	1	50.0
ARID1B	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
ASH1L	13	1	0	0	3	2	0	0	1	0	0	0	2	6	46.2
ASXL3	21	2	0	1	2	0	3	0	1	1	0	1	2	6	28.6
AUTS2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
BRPF1	4	1	0	1	1	0	0	1	1	0	0	0	0	3	75.0
CHD2	2	1	0	0	1	0	0	0	0	0	0	0	0	1	50.0
CHD7	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
CHD8	5	0	0	0	1	0	0	0	0	0	0	0	0	1	20.0
CIC	4	0	0	2	0	1	1	0	0	0	0	0	0	2	50.0
CLTC	5	2	2	2	1	0	0	1	0	0	0	2	2	2	40.0
CSDE1	1	0	0	0	0	0	1	0	0	0	0	0	0	1	100.0
CTCF	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
CUL3	2	0	0	0	1	0	0	0	0	0	0	0	0	1	50.0
DSCAM	15	1	0	0	0	1	0	0	0	0	0	0	0	2	13.3
EFTUD2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
EHMT1	2	0	0	0	1	0	0	0	0	0	0	0	0	1	50.0
EP300	4	0	0	0	0	1	0	0	0	0	0	0	0	1	25.0
FOXP1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
GIGYF1	15	1	0	3	2	0	3	1	1	0	0	0	1	6	40.0
GRIN2B	1	0	0	0	1	0	0	0	0	0	0	0	0	1	100.0
IRF2BPL	7	0	0	1	0	1	1	0	0	0	0	0	0	3	42.9
KANSL1	1	0	0	0	0	0	1	0	0	0	0	1	0	1	100.0
KCNQ2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
KMT2A	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
KMT2C	14	0	0	0	0	1	1	0	0	0	0	0	0	1	7.1
KMT2D	4	0	0	0	0	1	0	0	0	0	0	0	0	1	25.0
KMT5B	1	0	0	1	0	0	0	0	0	0	0	0	0	1	100.0
MBD5	4	0	0	0	0	0	1	1	0	0	0	0	0	2	50.0
NAA15	6	0	0	0	2	2	2	0	0	0	0	0	0	3	50.0
NBEA	4	0	0	0	0	0	0	0	0	0	0	1	0	1	25.0
NCKAP1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
NIPBL	2	0	0	0	0	0	0	1	0	0	0	0	0	1	50.0
NSD1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
NSD2	2	1	0	1	0	0	0	0	0	0	0	1	1	1	50.0
PHIP	13	0	0	0	1	1	1	0	0	0	0	0	0	2	15.4
POGZ	5	0	0	2	0	1	0	1	1	0	0	1	0	2	40.0
PTEN	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
RAI1	2	0	0	0	1	0	0	0	0	0	0	0	0	1	50.0
SATB2	2	2	1	2	1	0	0	2	0	0	1	2	1	2	100.0
SCN1A	1	0	0	0	1	0	0	0	0	0	0	0	0	1	100.0
SCN2A	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
SETD1A	3	0	0	0	1	0	2	0	0	0	0	0	0	2	66.7
SETD2	5	0	0	0	1	1	1	0	1	0	0	0	0	1	20.0
SETD5	8	0	0	1	0	0	0	1	0	0	0	0	0	1	12.5
SHANK2	18	1	0	1	0	2	4	1	2	0	0	0	0	7	38.9

Gene	Total	ID	ASD	ADHD	EP	SCZ	BPD	CD	SLD	CP	OCD	OND	MD	Any NPD	Percent Affected
STXBP1	1	1	1	1	1	1	0	1	0	0	0	1	1	1	100.0
TCF20	1	1	0	1	0	1	0	0	0	0	0	0	0	1	100.0
TCF4	3	1	0	0	0	0	0	0	0	1	0	1	1	1	33.3
TLK2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
TNRC6B	8	0	0	1	0	1	0	0	0	0	0	0	0	2	25.0
TRIO	17	4	1	4	0	2	3	3	1	0	0	3	0	9	52.9
TRIP12	1	1	0	1	0	0	1	0	0	0	1	0	0	1	100.0
TSC2	3	0	0	0	1	0	1	0	0	0	0	0	0	1	33.3
WAC	3	1	1	1	1	1	1	0	0	0	0	0	0	2	66.7
WASF1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
WDFY3	16	0	2	2	1	2	2	1	1	0	1	1	0	6	37.5
ZNF292	11	0	0	2	1	0	1	1	1	0	1	0	0	5	45.5
ZNF462	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0
Any gene	312	23	9	34	31	24	35	18	12	2	5	16	13	107	34.3

ID: Intellectual Disability; CD: Communication Disorder; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; SLD: Specific Learning Disorder; MD: Motor Disorder; OND: Other Neurodevelopmental Disorder; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; OCD: Obsessive Compulsive Disorder; EP: Epilepsy; CP: Cerebral Palsy; NPD: Neuropsychiatric developmental Disorder

Table S7. Congenital anomaly diagnosis by gene

Gene	Total	CNS	Cleft Lip & Palate	Cardiac	Urinary	Genital	Any CA	Percent Affected
ADNP	1	0	0	0	0	0	0	0.0
ANK2	26	0	0	0	2	0	2	7.7
ANKRD11	2	0	0	0	0	0	0	0.0
ARID1B	1	0	0	0	0	0	0	0.0
ASH1L	13	1	0	0	0	0	1	7.7
ASXL3	21	1	0	0	0	0	1	4.8
AUTS2	1	0	0	0	0	0	0	0.0
BRPF1	4	1	0	0	0	0	1	25.0
CHD2	2	0	0	0	0	0	0	0.0
CHD7	3	0	0	0	0	0	0	0.0
CHD8	5	0	0	0	0	0	0	0.0
CIC	4	0	0	0	0	0	0	0.0
CLTC	5	0	0	1	0	0	1	20.0
CSDE1	1	0	0	0	0	0	0	0.0
CTCF	1	0	0	0	0	0	0	0.0
CUL3	2	0	0	0	0	0	0	0.0
DSCAM	15	0	0	1	0	0	1	6.7
EFTUD2	1	0	0	0	0	0	0	0.0
EHMT1	2	0	0	0	0	0	0	0.0
EP300	4	0	0	0	0	0	0	0.0
FOXP1	3	0	0	0	0	0	0	0.0
GIGYF1	15	0	0	1	1	0	2	13.3
GRIN2B	1	0	0	0	0	0	0	0.0
IRF2BPL	7	0	0	0	0	0	0	0.0
KANSL1	1	0	0	0	0	0	0	0.0
KCNQ2	1	0	0	0	0	0	0	0.0
KMT2A	1	0	0	0	0	0	0	0.0
KMT2C	14	0	0	0	1	0	1	7.1
KMT2D	4	0	0	1	0	0	1	25.0
KMT5B	1	0	0	0	0	0	0	0.0
MBD5	4	0	0	0	0	0	0	0.0
NAA15	6	0	0	0	1	0	1	16.7
NBEA	4	0	0	0	0	0	0	0.0
NCKAP1	2	0	0	0	1	0	1	50.0
NIPBL	2	0	0	0	0	0	0	0.0
NSD1	3	0	0	0	0	0	0	0.0
NSD2	2	1	0	0	0	0	1	50.0
PHIP	13	1	0	0	0	0	1	7.7
POGZ	5	0	0	0	0	1	1	20.0
PTEN	1	0	0	0	0	0	0	0.0
RAI1	2	0	0	0	0	0	0	0.0
SATB2	2	0	1	0	0	0	1	50.0
SCN1A	1	0	0	0	0	0	0	0.0
SCN2A	1	0	0	0	0	0	0	0.0
SETD1A	3	1	0	0	0	0	1	33.3
SETD2	5	0	0	0	0	0	0	0.0
SETD5	8	0	0	0	0	0	0	0.0
SHANK2	18	1	0	0	0	0	1	5.6

Gene	Total	CNS	Cleft Lip & Palate	Cardiac	Urinary	Genital	Any CA	Percent Affected
STXBP1	1	0	0	0	0	0	0	0.0
TCF20	1	0	0	0	0	0	0	0.0
TCF4	3	1	0	0	0	0	1	33.3
TLK2	1	0	0	0	0	0	0	0.0
TNRC6B	8	1	0	1	1	2	4	50.0
TRIO	17	2	0	0	1	1	3	17.6
TRIP12	1	0	0	0	1	0	1	100.0
TSC2	3	0	0	1	0	0	1	33.3
WAC	3	0	0	0	0	0	0	0.0
WASF1	1	1	0	0	0	0	1	100.0
WDFY3	16	1	0	1	0	1	3	18.8
ZNF292	11	0	0	1	0	0	1	9.1
ZNF462	2	0	0	0	0	0	0	0.0
Any NPD gene	312	13	1	8	9	5	34	10.9

CA: congenital anomaly; CNS: central nervous system

Table S8. Chart review of individuals with a variant in selected NPD genes

Gene	N[a]	ID	Age	Sex	Variant	NPD diagnosis captured by ICD 9/10 codes approach and confirmed by manual EHR review	NPD symptoms captured only by manual EHR review of unstructured text	Other genetic disorder-related phenotypes captured by manual EHR review
TCF4	3	PT-1	65	M	c.632_633insAT p.Phe211LeufsTer24	Depressive disorder		
		PT-2	72	F	c.369+1G>T		**Schizophrenia, bipolar disorder**	Short stature (4'11")
		PT-3	8	M	Chr18:55228224-55234684 deletion	Intellectual disability, motor disorder, cerebral palsy		Pitt-Hopkins Syndrome, stereotypy (head-shaking), exotropia
CHD8	5	PT-4	85	M	c.3793A>T (p.Arg1265Ter)			
		PT-5	42	F	c.5017C>T (p.Arg1673Ter)	Depressive disorder and anxiety disorder	**Mild mental retardation**	Migraine, tall stature: Height - 5'11"
		PT-6	65	F	c.5017C>T (p.Arg1673Ter)	Epilepsy, depressive disorder, and anxiety disorder		
		PT-7	48	M	c.1423dup (p.Arg475ProfsTer3)			Tall stature: Height - 6'9"
		PT-8	70	F	c.7183-1G>A	Depressive disorder		Suicidal thoughts
TSC2	3	PT-9	35	F	c.3412C>T (p.Arg1138Ter)	Epilepsy, bipolar disorder, depressive disorder and anxiety disorder	Learning disability	Tuberous Sclerosis

Gene	N[a]	ID	Age	Sex	Variant	NPD diagnosis captured by ICD 9/10 codes approach and confirmed by manual EHR review	NPD symptoms captured only by manual EHR review of unstructured text	Other genetic disorder-related phenotypes captured by manual EHR review
		PT-10	75	F	c.3796_3797del (p.Leu1266AlafsTer55)			
		PT-11	59	M	c.2853_2854insC (p.Lys954GlnfsTer6)	Anxiety disorder	**Mild dysarthria**	
GRIN2B	1	PT-13	61	M	c.3931C>T p.Gln1311Ter	Epilepsy, depressive disorder and anxiety disorder	Cognitive Impairment	Angry affect, hemiplegia and hemiparesis following unspecified cerebrovascular disease, Cortical dysfunction
KMT2D	4	PT-14	74	M	c.16083dup p.Lys5362Ter	Schizophrenia, depressive disorder, anxiety disorder		
		PT-15	54	F	c.14710C>T p.Arg4904Ter	Depressive disorder, anxiety disorder		
		PT-16	74	M	c.9040del p.Leu3014TrpfsTer15	Depressive disorder, anxiety disorder	**Altered mental status, reduced cognitive status**	Ptosis
		PT-17	82	M	c.8743C>T p.Arg2915Ter			Bilateral sensorineural hearing loss, Vertical Heterophoria, ptosis
STXBP1	1	PT-10	17	M	c.1099C>T p.Arg367Ter	Intellectual disability, ASD, epilepsy, communication disorder, motor disorder, other neurodevelopmental disorder, anxiety		

Gene	N[a]	ID	Age	Sex	Variant	NPD diagnosis captured by ICD 9/10 codes approach and confirmed by manual EHR review	NPD symptoms captured only by manual EHR review of unstructured text	Other genetic disorder-related phenotypes captured by manual EHR review
CHD2	2	PT-11	82	M	c.3029C>G p.Ser1010Ter	Depression, anxiety		
		PT-12	40	F	c.3682del p.Glu1228SerfsTer21	Intellectual disability, epilepsy		
EFTUD2	1	PT-9	85	M	c.768del (p.Cys256Ter)			
KANSL1	1	PT-18	28	F	c.3031C>T p.Arg1011Ter	Intellectual disability, bipolar disorder, anxiety disorder		Club foot, chronic headache
WAC	3	PT-25	90	F	c.823C>T p.Gln275Ter			
		PT-26	49	M	c.1171dup p.Thr391AsnfsTer15	Asperger syndrome, ADHD		
		PT-27	37	M	c.1289-2A>T	Intellectual disability, communication disorder, motor disorder, epilepsy, schizophrenia, bipolar disorder, anxiety disorder, depressive disorder		
NSD1	3	PT-19	25	F	c.6871C>T p.Gln2291Ter	Depressive disorder		Profound bilateral sensorineural hearing loss
		PT-20	54	F	c.7581G>A p.Trp2527Ter			

Gene	N[a]	ID	Age	Sex	Variant	NPD diagnosis captured by ICD 9/10 codes approach and confirmed by manual EHR review	NPD symptoms captured only by manual EHR review of unstructured text	Other genetic disorder-related phenotypes captured by manual EHR review
		PT-21	68	M	c.7768_7769insT p.Pro2590LeufsTer12	Depressive disorder, anxiety disorder		

[a]Number of individuals with variant in each gene
Individuals who did not have ICD codes for NPD diagnosis and were only identified by manual EHR review of unstructured text have their NPD symptoms bolded and underlined.
Additional NPD symptoms identified by chart review but not with ICD codes are underlined among those with various NPD diagnoses.

Table S9. Association between recurrent NPD-associated CNVs and NPD diagnosis in DiscovEHR cohort

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Uncorrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	**8.77**	6.05	12.38	<0.0001	<0.0001	38	423	610	81304
ASD	**3.30**	1.64	6.19	0.019	1.33×10^{-3}	13	229	635	81498
ADHD	**1.39**	0.98	1.93	0.914	0.065	41	2397	607	79330
EP	**2.19**	1.65	2.84	3.72×10^{-6}	2.65×10^{-7}	59	3311	589	78416
SCZ	**1.76**	1.17	2.53	0.114	8.14×10^{-3}	27	2234	621	79493
BPD	**1.51**	1.12	1.99	0.114	8.13×10^{-3}	53	3636	595	78091
CD	**2.79**	1.84	4.07	1.00×10^{-4}	7.45×10^{-6}	28	901	620	80826
SLD	**3.30**	1.97	5.23	3.85×10^{-4}	2.75×10^{-5}	19	447	629	81280
CP	**1.91**	0.53	4.79	>0.99	0.280	3	161	645	81566
OCD	**1.07**	0.54	1.87	>0.99	0.838	10	966	638	80761
OND	**6.13**	3.82	9.51	1.52×10^{-10}	1.08×10^{-11}	30	406	618	81321
MD	**2.23**	1.41	3.32	0.014	0.001	22	1196	626	80531
AX	**1.14**	0.97	1.33	>0.99	0.110	308	34611	340	47116
DP	**1.25**	1.07	1.46	0.0812	0.006	298	32057	350	49670

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. Both Bonferroni corrected and uncorrected p-values are presented

Table S10. Association between LOF variants and NPD diagnosis in the European subset of participants sequenced on the VCRome platform (n=54,597)

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	**6.20**	2.27	13.69	0.018	5	255	165	54342
ASD	**10.77**	2.47	34.06	0.048	3	116	167	54481
ADHD	**2.96**	1.52	5.30	0.032	12	1352	158	53245
EP	**2.18**	1.24	3.56	0.123	15	2310	155	52287
SCZ	**2.79**	1.52	4.69	0.024	13	1700	157	52897
BPD	**1.75**	0.92	3.04	>0.99	12	2288	158	52309
CD	**2.11**	0.69	4.89	>0.99	4	591	166	54006
SLD	**3.37**	0.90	8.92	0.938	3	284	167	54313
CP	**3.91**	0.44	14.54	>0.99	1	103	169	54494
OCD	**0.24**	0.00	1.63	>0.99	0	622	170	53975
OND	**6.03**	1.76	16.16	0.095	4	247	166	54350
MD	**1.72**	0.57	3.93	>0.99	4	803	166	53794
AX	**1.11**	0.81	1.52	>0.99	75	23179	95	31418
DP	**1.43**	1.05	1.94	0.340	83	22376	87	32221

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 14 tests

Table S11. Association between LOF variants and NPD diagnosis in the European subset of samples sequenced on the IDT platform (n=27,130)

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	**11.65**	5.30	23.38	1.69×10^{-6}	10	168	95	26962
ASD	**2.52**	0.56	8.76	>0.99	3	113	102	27017
ADHD	**2.91**	1.52	5.25	0.025	15	1045	90	26085
EP	**2.96**	1.51	5.28	0.037	11	1001	94	26129
SCZ	**3.33**	1.35	6.85	0.170	6	534	99	26596
BPD	**3.24**	1.85	5.39	0.002	17	1348	88	25782
CD	**5.38**	2.32	11.27	0.004	9	310	96	26820
SLD	**9.39**	3.91	20.11	1.26×10^{-4}	8	163	97	26967
CP	**1.57**	0.01	11.20	>0.99	0	58	105	27072
OCD	**3.49**	1.29	7.64	0.245	5	344	100	26786
OND	**6.67**	2.29	17.16	0.013	7	159	98	26971
MD	**2.88**	0.95	6.67	0.850	4	393	101	26737
AX	**1.20**	0.81	1.78	0.113	51	11432	54	15698
DP	**1.61**	1.09	2.37	0.248	50	9681	55	17449

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 14 tests

Table S12. Association between NPD genes pathogenic variants and NPD diagnosis in 1st and 2nd degree unrelated subset

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	**8.65**	4.17	16.20	9.08×10^{-6}	10	273	176	56265
ASD	**6.39**	2.05	16.75	0.04	5	147	181	56391
ADHD	**3.33**	1.97	5.38	3.46×10^{-4}	21	1625	165	54913
EP	**3.01**	1.88	4.59	2.97×10^{-4}	22	2326	164	54212
SCZ	**3.24**	1.81	5.37	0.004	14	1522	172	55016
BPD	**2.29**	1.39	3.60	0.026	20	2584	166	53954
CD	**4.32**	2.11	7.95	0.003	10	560	176	55978
SLD	**8.70**	4.11	16.71	1.24×10^{-5}	10	266	176	56272
CP	**3.28**	0.37	12.14	>0.99	1	102	185	56436
OCD	**1.82**	0.60	4.15	>0.99	4	653	182	55885
OND	**5.52**	2.10	12.77	0.016	7	249	179	56289
MD	**1.62**	0.54	3.68	>0.99	4	807	182	55731
AX	**1.32**	0.98	1.77	>0.99	89	23373	97	33165
DP	**1.57**	1.17	2.10	0.038	90	21480	96	35058

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 14 tests

Table S13. Association between NPD genes pathogenic variants and NPD diagnosis after exclusion of individuals age younger than 18 years

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	**9.75**	5.14	16.89	1.18×10^{-7}	12	349	256	80774
ASD	**9.08**	3.07	22.30	4.65×10^{-3}	5	122	263	81001
ADHD	**2.88**	1.78	4.46	6.67×10^{-4}	23	2211	245	78912
EP	**2.32**	1.49	3.45	5.56×10^{-3}	24	3248	244	77875
SCZ	**2.91**	1.77	4.50	0.001	19	2230	249	78893
BPD	**2.33**	1.53	3.43	0.002	28	3629	240	77494
CD	**4.12**	2.00	7.50	0.007	9	663	259	80460
SLD	**6.84**	3.35	12.57	6.02×10^{-5}	10	383	258	80740
CP	**2.52**	0.29	9.17	>0.99	1	150	267	80973
OCD	**1.27**	0.42	2.87	>0.99	4	953	264	80170
OND	**8.36**	3.67	16.77	1.42 x 10-4	8	240	260	80883
MD	**1.84**	0.76	3.70	>0.99	6	1080	262	80043
AX	**1.17**	0.92	1.50	>0.99	125	34463	143	46660
DP	**1.52**	1.19	1.95	0.011	133	32022	135	49101

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 14 tests

Table S14. Association between NPD genes pathogenic variants and NPD diagnosis after exclusion of individuals with intellectual disability

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ASD	**4.53**	1.05	14.53	0.567	3	152	257	81152
ADHD	**2.38**	1.40	3.83	0.026	19	2288	241	79016
EP	**2.15**	1.34	3.26	0.030	21	3187	239	78117
SCZ	**2.94**	1.76	4.59	0.002	18	2159	242	79145
BPD	**2.21**	1.42	3.30	0.010	25	3573	235	77731
CD	**3.26**	1.56	6.01	0.039	9	802	251	80502
SLD	**7.49**	3.60	14.02	3.45×10^{-5}	10	383	250	80921
CP	**1.02**	0.01	7.02	>0.99	0	132	260	81172
OCD	**1.05**	0.29	2.61	>0.99	3	938	257	80366
OND	**5.86**	2.09	13.81	0.022	6	275	254	81029
MD	**1.46**	0.55	3.09	>0.99	5	1145	255	80159
AX	**1.17**	0.91	1.50	>0.99	120	34391	140	46913
DP	**1.48**	1.15	1.90	0.027	126	31861	134	49443

ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 13 tests

Table S15. Association between NPD genes pathogenic variants and NPD diagnosis in European ancestry and all non-Europeans

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	**10.63**	6.51	16.61	4.97×10^{-14}	23	501	289	90094
ASD	**3.46**	1.44	7.62	0.097	9	309	303	90286
ADHD	**2.79**	1.84	4.11	6.55×10^{-5}	34	2783	278	87812
EP	**2.44**	1.65	3.47	3.79×10^{-4}	31	3723	281	86872
SCZ	**3.28**	2.11	4.86	2.31×10^{-5}	24	2454	288	88141
BPD	**2.27**	1.56	3.22	7.09×10^{-4}	35	4172	277	86423
CD	**3.55**	2.07	5.76	3.04×10^{-4}	18	1066	294	89529
SLD	**4.19**	2.14	7.52	0.002	12	517	300	90078
CP	**2.78**	0.58	8.03	>0.99	2	185	310	90410
OCD	**1.24**	0.46	2.62	>0.99	5	1064	307	89531
OND	**5.31**	2.71	9.86	7.53×10^{-5}	16	502	296	90093
MD	**2.77**	1.52	4.61	0.023	13	1339	299	89256
AX	**1.09**	0.86	1.37	>0.99	139	38125	173	52470
DP	**1.41**	1.13	1.77	0.041	147	35301	165	55294

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 14 tests

Table S16. Associations between variants in NPD genes and NPD diagnosis in Black or African Americans

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	93.29	10.07	918.13	0.001	3	19	6	1507
ASD	-	-	-	-	2	16	7	1510
ADHD	-	-	-	-	1	62	8	1464
EP	-	-	-	-	0	84	9	1442
SCZ	-	-	-	-	1	46	8	1480
BPD	-	-	-	-	1	113	8	1413
CD	-	-	-	-	1	26	8	1500
SLD	-	-	-	-	0	14	9	1512
CP	-	-	-	-	0	4	9	1522
OCD	-	-	-	-	0	7	9	1519
OND	-	-	-	-	2	27	7	1499
MD	-	-	-	-	0	21	9	1505
AX	-	-	-	-	2	568	7	958
DP	**1.37**	0.36	4.95	>0.99	4	569	5	957

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 14 tests. (-) Odds ratio could not be calculated for all disorders due to the small number of cases with a pathogenic variant

Table S17. Associations between variants in NPD genes and NPD diagnosis in Hispanic/Latino individuals

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	-	-	-	-	2	16	7	1282
ASD	-	-	-	-	0	20	9	1278
ADHD	-	-	-	-	1	61	8	1237
EP	-	-	-	-	0	71	9	1227
SCZ	-	-	-	-	0	37	9	1261
BPD	-	-	-	-	0	107	9	1191
CD	-	-	-	-	0	29	9	1269
SLD	-	-	-	-	1	8	8	1290
CP	-	-	-	-	1	4	8	1294
OCD	-	-	-	-	0	17	9	1281
OND	-	-	-	-	2	18	7	1280
MD	**15.94**	3.14	68.79	0.025	3	26	6	1272
AX	-	-	-	-	1	563	8	735
DP	-	-	-	-	1	526	8	772

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 14 tests. (-) odds ratio could not be calculated for all disorders due to the small number of cases with a pathogenic variant

Table S18. Association between NPD genes pathogenic variants and NPD diagnosis based on control group without any NPD diagnoses

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	12.23	6.31	22.42	1.86×10^{-9}	15	423	85	34590
ASD	10.71	3.25	32.09	0.003	6	229	85	34590
ADHD	3.41	1.99	5.73	2.27×10^{-4}	27	2397	85	34590
EP	3.09	1.95	4.75	9.65×10^{-5}	26	3311	85	34590
SCZ	3.65	2.16	5.86	1.16×10^{-4}	19	2234	85	34590
BPD	3.32	2.07	5.17	2.82×10^{-5}	29	3636	85	34590
CD	4.25	2.14	7.87	1.53×10^{-3}	13	901	85	34590
SLD	9.07	4.14	18.43	1.05×10^{-5}	11	447	85	34590
CP	3.47	0.39	13.14	>0.99	1	161	85	34590
OCD	2.19	0.80	4.87	>0.99	5	966	85	34590
OND	9.4	3.65	22.88	1.52×10^{-4}	11	406	85	34590
MD	2.79	1.27	5.36	0.187	8	1196	85	34590
AX	1.47	1.11	1.96	0.109	126	34611	85	34590
DP	1.72	1.30	2.28	0.002	133	32057	85	34590

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 14 tests

Table S19. Association between NPD-associated CNVs and NPD diagnosis based on control group without any NPD diagnoses

Disorder	Odds ratio	95% CI lower	95% CI upper	Corrected p-value	Number of cases with a variant	Total cases	Number of controls with a variant	Total controls
ID	12.29	8.18	18.07	<0.0001	38	423	212	34590
ASD	6.01	2.63	12.88	0.001	13	229	212	34590
ADHD	1.89	1.28	2.74	0.024	41	2397	212	34590
EP	2.63	1.94	3.50	5.34×10^{-8}	59	3311	212	34590
SCZ	2.08	1.36	3.05	0.015	27	2234	212	34590
BPD	2.13	1.54	2.91	1.83×10^{-4}	53	3636	212	34590
CD	3.39	2.16	5.13	1.00×10^{-5}	28	901	212	34590
SLD	4.91	2.79	8.22	6.64×10^{-6}	19	447	212	34590
CP	2.86	0.79	7.25	>0.99	3	161	212	34590
OCD	1.53	0.76	2.74	>0.99	10	966	212	34590
OND	8.22	4.67	14.13	1.27×10^{-10}	30	406	212	34590
MD	2.74	1.72	4.17	1.09×10^{-3}	22	1196	212	34590
AX	1.34	1.12	1.60	0.022	308	34611	212	34590
DP	1.41	1.18	1.69	0.0026	298	32057	212	34590

ID: Intellectual Disability; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit and Hyperactivity Disorder; EP: Epilepsy; SCZ: Schizophrenia and Related Disorder; BPD: Bipolar Disorder; CD: Communication Disorder; SLD: Specific Learning Disorder; CP: Cerebral Palsy; OCD: Obsessive Compulsive Disorder; OND: Other Neurodevelopmental Disorder; MD: Motor Disorder; AX: Anxiety; DP: Depressive disorder. P-values are Bonferroni corrected for 14 tests