Title

Analysis of exonic deletions in a large population study provides novel insights into NRXN1 pathology.

Journal

npj Genomic Medicine, 9(1)

Authors

Montalbano, Simone

Krebs, Morten

Rosengren, Anders

et al.

Publication Date

2024-12-19

DOI

10.1038/s41525-024-00450-8

Peer reviewed

npj | genomic medicine

Article

Published in partnership with CEGMR, King Abdulaziz University

https://doi.org/10.1038/s41525-024-00450-8

Analysis of exonic deletions in a large population study provides novel insights into NRXN1 pathology

Simone Montalbano[1,2], Morten Dybdahl Krebs[1,2], Anders Rosengren[1,2], Morteza Vaez[1,2], Kajsa-Lotta Georgii Hellberg[1,2], Preben B. Mortensen[2,3], Anders D. Børglum[2,4,5], Daniel H. Geschwind[6,7,8,9], iPSYCH Investigators*, Armin Raznahan[10], Wesley K. Thompson[2,11], Dorte Helenius[1,2], Thomas Werge[1,2,12] & Andrés Ingason[1,2]

The *NRXN1* locus is a hotspot for non-recurrent copy number variants and exon-disrupting *NRXN1* deletions have been associated with increased risk of neurodevelopmental disorders in case-control studies. However, corresponding population-based estimates of prevalence and disease-associated risk are currently lacking. Also, most studies have not differentiated between deletions affecting exons of different *NRXN1* splice variants nor considered intronic deletions. We used the iPSYCH2015 case-cohort sample to obtain unbiased estimates of the prevalence of *NRXN1* deletions and their associated risk of autism, schizophrenia, depression, and ADHD. Most exon-disrupting deletions affected exons specific to the alpha isoform, and almost half of the non-exonic deletions represented a previously reported segregating founder deletion. Carriage of exon-disrupting *NRXN1* deletions was associated with a threefold and twofold increased risk of autism and ADHD, respectively, whereas no significantly increased risk of depression or schizophrenia was observed. Our results highlight the importance of using population-based samples in genetic association studies.

Larger genomic deletions in the *NRXN1* locus have been associated with a highly increased risk of mental disorders and, in particular, schizophrenia. However, the locus is known to harbour highly heterogeneous CNVs (Copy Number Variations, deletions and duplications) and, moreover, no population-based estimates of risk are available. Here, we use the iPSYCH2015 case-cohort sample to investigate the population prevalence and phenotypic consequences of specific types of deletions within the locus.

Neurexins are a family of highly conserved transmembrane proteins strongly involved in the development and function of neuronal synapses[1]. Like all mammals, humans possess three genes encoding different neurexin proteins (NRXN1-3)[2]. All three genes encode two main protein isoforms, alpha and beta[1], and are almost exclusively expressed in neuronal tissue[3,4]. Notably, hundreds of splicing isoforms are expressed in humans and mice, many of which are specific to certain neuronal cell types[1,5,6]. Neurexin proteins are expressed by neurons at the presynaptic nerve terminal and their expression peaks around birth[1]. Among other ligands, neurexins bind to the calcium/calmodulin-dependent serine protein kinase (CASK) scaffolding molecules, contributing to the coupling of Ca^{2+} channels to synaptic release machinery[1,7].

NRXN1 is a 1.3 Mbp gene located on the short arm of chromosome 2 (GRCh38:49,918,503–51,225,575)[8]. Among the three neurexin genes, *NRXN1* is the most studied with respect to association with disease[9–11]. Multiple case-control studies have associated exonic deletions with increased risk of neurodevelopmental disorders, including schizophrenia

[1]Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Roskilde, Denmark. [2]The Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), Copenhagen and Aarhus, Denmark. [3]National Centre for Register-based Research, Aarhus University, Aarhus, Denmark. [4]Department of Biomedicine – Human Genetics and the iSEQ Center, Aarhus University, Aarhus, Denmark. [5]Center for Genomics and Personalized Medicine, Aarhus, Denmark. [6]Department of Neurology, University of California, Los Angeles, CA, USA. [7]Department of Human Genetics, University of California, Los Angeles, CA, USA. [8]Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, CA, USA. [9]Program in Neurobehavioral Genetics, Semel Institute, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA. [10]Section on Developmental Neurogenomics, Human Genetics Branch, National Institute of Mental Health Intramural Research Program, Bethesda, MD, USA. [11]Laureate Institute for Brain Research, Tulsa, OK, USA. [12]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. *A list of authors and their affiliations appears at the end of the paper. e-mail: andres.ingason@regionh.dk

(odds ratio (OR): 4.5; 95% CI: 2.0–10.9)[12], autism spectrum disorder (OR: 7.2; 95% CI: 0.9–326)[13], attention-deficit/hyperactivity disorder (OR: 4.68; CI95%: 1.82–10.64)[14], depression (OR: 2.01; CI95%: 1.18–3.19)[15], intellectual disability and/or developmental delay (OR: 8.14; 95% CI: 2.91–22.7)[16], epilepsy (OR: 9.91; 95% CI: 1.92–51.1)[17], and Tourette Syndrome (OR: 20.3; 95% CI: 2.6–156)[18]. Deletions in the 5′ end of the gene are more commonly observed compared to the rest of the gene[9]. To our knowledge, duplications and intronic deletions have not been strongly associated with disease risk in previous studies and, in general, appear to be far less studied than exonic deletions of *NRXN1*.

CNVs in the *NRXN1* locus are non-recurrent, meaning that CNVs result from unrelated de novo mutations which do not share fixed breakpoints, and their mutational mechanism is different from that observed in non-allelic homologous recombination (NaHR) mediated by low-copy repeat (LCR) sequence elements[19]. One possible explanation for this genomic instability is that the *NRXN1* locus, similarly to other large genes, is a late replicating region and therefore more prone to mutations resulting from stress-induced replication errors[20].

As was also the case for rare recurrent CNVs (such as 22q11.2 deletions and 16p11.2 duplication) *NRXN1* deletions were originally associated with high risk of disease from single case studies or small collections of cases[21–24], followed by larger case-control studies also based on highly selected samples (e.g., cases with severe or long-term illness and controls screened for any family history of mental illness)[11,13,16–18,25]. However, recent research on recurrent CNV loci in larger and more population-representative study samples suggests that associations obtained using selected case-control samples tend to be biased toward an overestimation of the disease risk, owing largely to an underestimation of the prevalence of recurrent CNVs in the general population[26–28].

In this study, we use the unique design of the iPSYCH2015[29] case-cohort study to provide population-representative estimates of the prevalence of *NRXN1* deletions, and the associated risk of attention-deficit/hyperactivity disorder (ADHD), major depressive disorder (MDD), schizophrenia spectrum disorder (SSD), autism spectrum disorder (ASD), and bipolar disorder (BPD). We assess the risk of any deletion in the *NRXN1* locus as well as that of different subgroups (including non-exonic ones). Moreover, we show that a significant proportion of intronic deletions in the locus is segregating in the population and may be associated with an increased risk of some psychiatric disorders.

Methods

Study design, phenotypes, and genotyping

This study is based on the iPSYCH2015 case-cohort sample[29], an expanded version of iPSYCH2012, which has been previously described in detail[30]. In brief, the base population is defined as all 1,657,449 singleton births that occurred in Denmark between May 1, 1981, and Dec 31, 2008, who were alive and residing in Denmark on their first birthday and had a mother registered in the Danish Civil Registration System[31]. From the base population all persons who received a diagnosis of a major psychiatric disorder (as specified below) no later than Dec 31, 2015, were included in the case sample, $N = 92{,}531$ individuals. Then, a randomly selected population-representative cohort of $N = 50{,}615$ individuals was drawn from the base population, including 3030 who overlapped with the case sample. Individual diagnosis sample counts are as follow: SSD (ICD10 F20–F29; $n = 16{,}008$), MDD (ICD10 F32–F33 and ICD 8 296.09, 296.29, 298.09, and 300.49; $n = 37{,}555$), ASD (ICD10 F84; $n = 24{,}975$), or ADHD (ICD10 F90; $n = 29{,}668$).

We also assessed three other brain disorders; intellectual disability (ID), epilepsy, and Tourette syndrome (TS), with prior evidence of association with NRXN1 deletions[16–18], using information on hospital diagnoses that had been obtained through the Danish Psychiatric Central Research Register[32] and the Danish National Patient Registry[33] for other iPSYCH2015 studies. The diagnostic codes used to identify individuals with these disorders were as follows: ID (ICD10: F70-F79; ICD8: 311-315), epilepsy (ICD10: G40; ICD8: 345 (excluding 345.29)), TS (ICD10: F95.2).

Supplementary Table 2 provides carrier count for each diagnosis, as well as a subset by subcohort (iPSYCH2012 or iPSYCH2015i) and gender.

Genotyping was performed using Illumina microarrays and has been described elsewhere[30]. Notably, the genotyping was performed on dried blood spot samples taken at birth. iPSYCH2012 and the additional extension (iPSYCH2015i) were genotyped using two different arrays, PsychArray version 1.0 and Global Screening Array version 2 (GSA), respectively. B allele frequency (BAF) and logR ratio (LRR) values were extracted using GenomeStudio and samples with a genotyping call rate below 95% were excluded.

CNV calling and pre-processing

CNVs were called using PennCNV[34] as described in our previously published CNV calling and processing protocol[35]. All steps of the calling pipeline were run using the Singularity container provided in the protocol. In brief, the intensity files were filtered to include only biallelic autosomal SNPs mapping uniquely to the Haplotype Reference Consortium (HRC) hg19 reference map[36], with a minor allele frequency of at least 0.1%, which yielded 280,700 and 509,754 probes for the PsychArray and GSA, respectively. Next, PennCNV calls were obtained with the script *"detect_cnv.pl"* setting a minimum number of probes (*--minsnp*) at 5, and the minimum length (*--minlength*) at 1000 bp. We then merged adjacent calls, with the *PennCNV* script *"clean_cnv.pl"* using the settings *"--fraction 0.2 --bp"* whereby two calls are merged if the gap between them corresponds to less than 20% of the combined length (in base pairs) of the calls. After CNV calling, we excluded samples with high levels of noise from the analysis. Thus, samples were excluded if they had either a LRR standard deviation value ≥ 0.35, BAF drift ≥ 0.005 or |GCWF| ≥ 0.02.

The locus of interest was defined as the NRXN1 gene in Ensembl[8] GRCh37 (https://grch37.ensembl.org/Homo_sapiens/Info/Index) plus 0.5 Mbp upstream and downstream of the gene boundaries (chr2:49 645 643-51 759 674). Any CNV call overlapping the region by at least 0.1% of its length was selected for visual validation using the function "*select_stich_-calls()*" from the R package QCtreeCNV[35]; this step also removed CNV smaller than 10 SNPs. Visual inspection was performed independently by two analysts as already described[35]. The boundaries of true CNVs were manually adjusted if necessary and any discordant call between the analysts was re-evaluated in a final joint session.

CNV analysis

The genomic coordinates of *NRXN1* exons and transcripts were extracted using Ensembl[8] GRCh37 (https://grch37.ensembl.org/Homo_sapiens/Info/Index). We decided to focus on protein-coding transcripts only and thus selected all 9 transcripts with a protein match in UniProt[37] (https://www.uniprot.org/), yielding a total of 41 unique exons.

Under the assumption that exons mapping close to each other are likely to be deleted by the same CNVs, we investigated if any larger pattern was present at the level of the whole gene. We computed a genomically ordered correlation matrix across all exons, defined as an $N \times N$ matrix where N is the number of exons and the cell xy is the number of times a CNV affecting exon x also affects exon y.

CNVs are not equally distributed across the locus. We explored this topic using an IOU matrix, defined as an NxN matrix where N is the number of CNVs (381) and the cell xy is the IOU (Intersection Over the Union) score for CNVs x and y. IOU is 1 for two identical segments and ranges between 0 and 1 for any two overlapping segments, while non-overlapping segment pairs have an IOU range from 0 to approaching an asymptote at −1 the farther apart the two segments are. We then subgrouped exons in "alpha" and "beta" regions, based on Fig. 1d and previous literature[38], corresponding to exons ENSE00001682911 to ENSE00002460080 (beta), and exons ENSE00002453754 to ENSE00001547151 (alpha). For the purpose of the secondary analysis (Table 1), deletions affecting exons from both groups were assigned to "alpha".

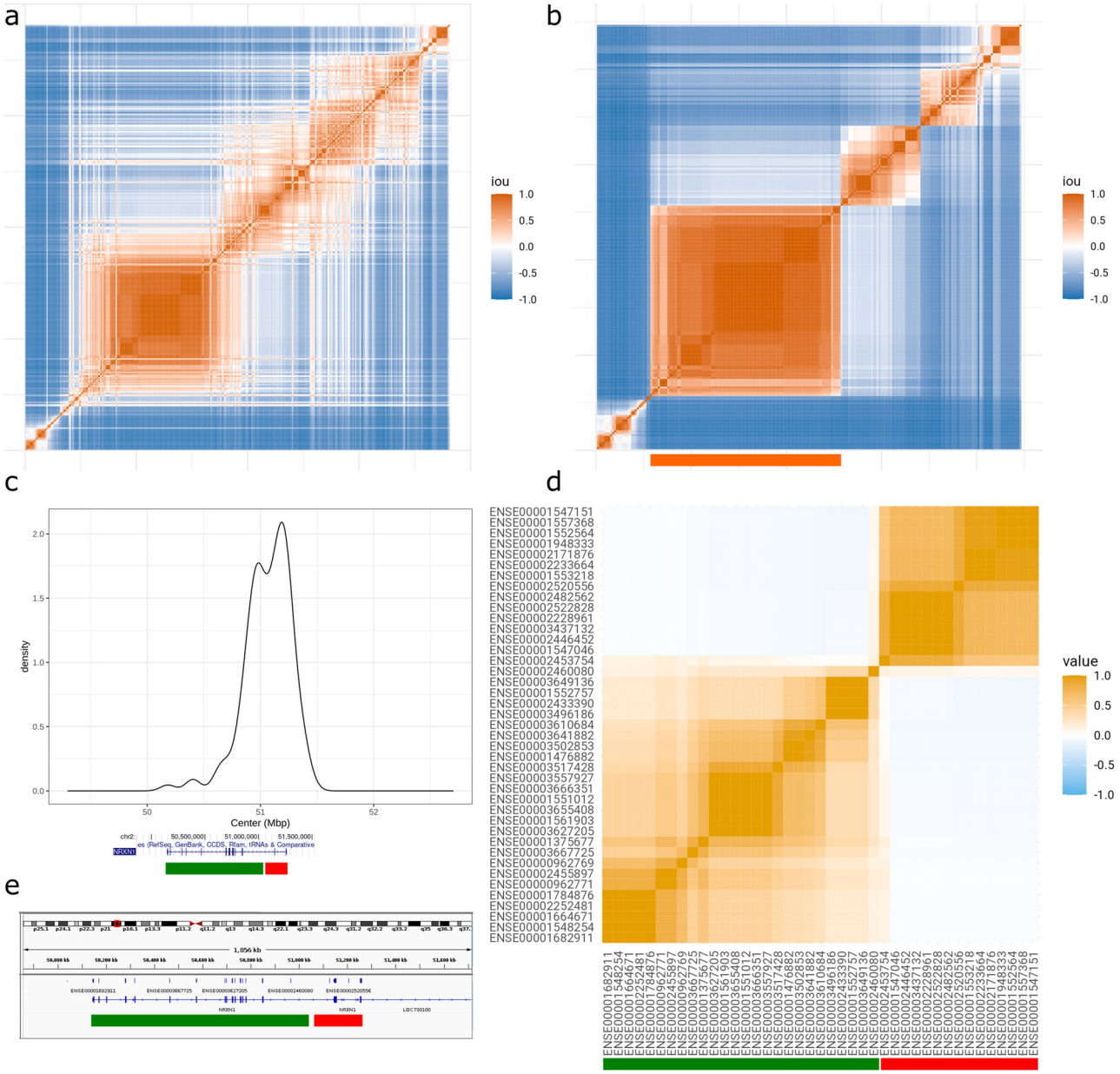

Fig. 1 | *NRXN1* deletions similarity matrices, and NRXN1 correlation matrix. Note that the *NRXN1* gene is encoded on the reverse strand, meaning the alpha promoter region (5′ of the gene) is shown on the right in this figure (see panel c for a breakdown of the gene structure). a Similarity heatmap for all deletions in the neurexin locus. Similarity is measured as IOU (intersection over the union), as described in the methods. Each row represents a deletion. Deletions are ordered on the x-axis based on the genomic position of the respective centre. Note that the scale is not linear as CNVs are not distributed equally across the locus. **b** Positional similarity for intronic deletions. This makes more evident the large group of very homogeneous deletions (marked with the orange bar on the x-axis). This group is referenced as segregating in the main text. **c** Distribution of the centre position for all exonic deletions in the NRXN1 gene locus. A schematic of the main gene isoform is aligned below the x-axis. The green and red bars mark the two exonic groups described in (**d**). **d** Exon correlation matrix. Exons are ordered based on genomic location. Note that the scale is not linear as exons are not distributed equally across the locus (see **c**). The red bar marks the exons in the "alpha" group and the green in the "beta" group. **e** A different view on the NRXN1 gene, the top blue graph shows all exons used in the study, while the bottom shows the top isoform.

Segregating deletion analysis

The coordinates of the segregating *NRXN1* deletion found in Rujescu et al.[25] were lifted over from hg18 to hg19 using the online tool LiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver).

To identify SNPs in high linkage disequilibrium (LD) with the segregating deletion, we performed an association analysis (using the "*--assoc*" command in PLINK[39,40] with default settings, Supplementary Fig. 2) where we compared the 100 identified carriers with 5000 randomly drawn non-carriers, across all SNPs with MAF > 0.01 and info >0.95 mapping on the entire chromosome 2, using an imputed genotype dataset of the iPSYCH2015[41]. We then pruned the resulting SNPs with the following settings *--clump-p1 0.00001 --clump-r2 0.8 --clump-kb 1000000*.

The phased genotypes of the top 10 SNPs (shown in Supplementary Table 1) were imported in R. Here we constructed all possible haplotypes of length between two and five SNPs and tested their association with the deletion carriers using the R function *fisher.test()*. The haplotypes with an OR ≥ 2 and a *p*-value ≤ 0.0001 were further tested using the function *roc()* from the R package pROC[42] to get the AUC (Area Under the Curve) value.

Table 1 | *NRXN1* deletions and associated risk of psychiatric disorders

Exposure	Outcome[a]	N_{aff}[b]	OR[c]	CI95%[c]	P[d]	P_{FDR}[d]
Main analysis (all exonic vs all non-exonic deletions)						
Exonic	Any	108	2.13	1.39–3.26	0.00048	0.0067
Exonic	ADHD	41	2.01	1.23–3.31	0.0057	0.040
Exonic	ASD	52	3.05	1.87–4.97	7.4×10^{-6}	0.00031
Exonic	MDD	31	1.46	0.83–2.56	0.19	0.53
Exonic	SSD	12	1.41	0.69–2.90	0.35	0.80
Exonic	SCZ	8	1.88	0.81–4.33	0.14	0.41
Non-exonic	Any	147	1.05	0.79–1.39	0.74	0.86
Non-exonic	ADHD	51	1.04	0.72–1.51	0.82	0.86
Non-exonic	ASD	41	1.06	0.71–1.58	0.79	0.86
Non-exonic	MDD	49	0.91	0.61–1.35	0.64	0.86
Non-exonic	SSD	37	1.48	0.96–2.29	0.078	0.27
Non-exonic	SCZ	23	1.90	1.13–3.1907	0.0200	0.090
Secondary analysis (alpha vs beta exonic, and segregating vs non-segregating non-exonic deletions)						
Exonic alpha	Any	68	2.83	1.56–5.13	0.0006	0.0067
Exonic alpha	ADHD	28	3.02	1.54–5.94	0.0013	0.011
Exonic alpha	ASD	29	3.75	1.88–7.47	0.00020	0.0036
Exonic alpha	MDD	20	2.30	1.02–5.18	0.045	0.19
Exonic alpha	SSD	>7	2.29	0.88–5.97	0.091	0.30
Exonic beta	Any	40	1.49	0.81–2.75	0.20	0.53
Exonic beta	ADHD	13	1.15	0.53–2.47	0.72	0.86
Exonic beta	ASD	23	2.45	1.22–4.90	0.011	0.068
Exonic beta	MDD	11	0.89	0.39–2.03	0.78	0.86
Exonic beta	SSD	<5	0.78	0.24–2.46	0.67	0.86
Non-exonic segregating	Any	68	1.19	0.77–1.82	0.43	0.80
Non-exonic segregating	ADHD	23	1.22	0.70–2.13	0.49	0.80
Non-exonic segregating	ASD	18	1.23	0.66–2.28	0.51	0.80
Non-exonic segregating	MDD	24	1.06	0.57–1.96	0.86	0.86
Non-exonic segregating	SSD	20	2.19	1.15–4.16	0.017	0.080
Non-exonic other	Any	79	0.95	0.66–1.38	0.79	0.86
Non-exonic other	ADHD	28	0.93	0.57–1.51	0.77	0.86
Non-exonic other	ASD	23	0.95	0.56–1.61	0.84	0.86
Non-exonic other	MDD	25	0.81	0.47–1.38	0.44	0.80
Non-exonic other	SSD	17	1.06	0.58–1.95	0.85	0.86
Tertiary analysis (segregating non-exonic deletion vs other carriers of the underlying haplotype)						
Segregating deletion carriers	Any	68	1.19	0.77–1.82	0.43	0.80
Segregating deletion carriers	ADHD	23	1.22	0.70–2.13	0.49	0.80
Segregating deletion carriers	ASD	18	1.23	0.66–2.28	0.52	0.80
Segregating deletion carriers	MDD	24	1.06	0.57–1.96	0.86	0.86
Segregating deletion carriers	SSD	20	2.19	1.15–4.16	0.017	0.080
Other top haplotype carriers	Any	1512	0.96	0.88–1.05	0.37	0.80
Other top haplotype carriers	ADHD	530	0.99	0.88–1.10	0.80	0.86
Other top haplotype carriers	ASD	442	0.97	0.86–1.10	0.69	0.86
Other top haplotype carriers	MDD	592	0.94	0.84–1.06	0.35	0.80
Other top haplotype carriers	SSD	250	0.95	0.81–1.10	0.49	0.80
Quaternary analysis (same as above for SSD, but only in unrelated subjects of European ancestry)						
Segregating deletion carriers	SSD	14	1.93	0.94–3.97	0.074	0.27
Other top haplotype carriers	SSD	190	0.89	0.75–1.06	0.20	0.53

[a]The risk associated with different classes of deletions (and for carriers of a haplotype underlying a segregating founder deletion) was assessed separately for; attention-deficit/hyperactivity disorder (ADHD), autism spectrum disorder (ASD), major depressive disorder (MDD), and schizophrenia spectrum disorder (SSD), as well as combined (i.e., being affected with any of those disorders; Any). In the main analysis (top) we also assessed the risk associated with schizophrenia, narrowly defined (ICD10:F20), and in the quaternary analysis (bottom) we assessed the risk associated with SSD for the founder deletion and the underlying haplotype in a subset of unrelated European-ancestry samples only.

[b]For each test we provide the number of affected carriers (Naff); for the alpha and beta subclasses of exonic deletions, the exact number of carriers with SSD cannot be disclosed due to legislation regarding the protection of personal-level data in the research of nationwide registers and biobanks in Denmark.

[c]The odds ratios (OR) and 95% confidence intervals (CI95%) were in all instances derived from a logistic regression model including sex (as assigned at birth), age (at end of follow-up) and genotyping array (PsychArray or GSA) as covariates.

[d]The associated *p*-values (P) were subsequently corrected for multiple testing using false discovery rate adjustment (P_{FDR}).

Statistical analysis

We derived population-based prevalence (with CI95%) for the different subgroups of *NRXN1* deletions using the svydesign() and svyciprop() functions from the R package survey[43], with finite population correction (FPC) to account for oversampling of cases in iPSYCH2015.

Briefly, we divided the post-QC number of cases (77,655) and individuals from the random population subcohort (43,311) with the total number of corresponding individuals in the source population (90,218 and 1,657,449) to derive the sampled population fractions; 0.85068 (100% of cases minus the ones failing genotype or excluded in QC) and 0.02613, respectively. Samples from overlapping individuals (cases-in-subcohort) were assigned the case population fraction (0.85068).

We calculated the corresponding prevalence of exonic *NRXN1* deletions in the UKB directly from carrier counts provided by Crawford et al.[44] and derived CI95% as follows (R pseudocode): *CI95% = qbeta(c(0.05/2,1-0.05/2), nCarrier + 0.5, nTotal-nCarrier + 0.5)*, where nCarrier and nTotal indicate the number of carriers and the total number of assessed samples (421,268), respectively.

We compared the prevalence of exonic deletions in iPSYCH2015 and UKB with Welch's test of the difference between two means assuming unequal variance. Briefly, we defined the difference; $d = (|log(p_{iPSYCH}/p_{UKB})|)$, the standard error of the difference; $SEd = \sqrt{(SE_{iPSYCH}^2 + SE_{UKB}^2)}$, and the p-value; $P = 2*(1\text{-}pnorm(d/SEd))$, where p_{iPSYCH} and SE_{iPSYCH}, and p_{UKB} and SE_{UKB}, indicate the prevalence and standard error of prevalence in iPSYCH2015 and UKB, respectively.

To estimate the risk of index psychiatric disorders associated with *NRXN1* deletions we ran a logistic regression analysis using *gam()* from the R package mgcv[45]. We used age, sex (at birth) and SNP array type as covariates, with a smoothed function to model the effect of age using the mgcv function s(). In each association, we included all cases for the phenotype of interest and all controls, defined as individuals not having any of the index diagnoses. For the later-onset disorders SSD, MDD and SCZ, we only included those controls who were at least as old as the youngest case. Multiple testing correction was applied to the table containing the results of all three analyses (Table 1) using the R function *p.adjust(method = "fdr")*. We then compared risk estimates with those reported in published case-control studies (in each case the study applying the largest case-control sample size for the respective disorder; only considering studies that controlled for genotyping array, when including samples genotyped on different arrays) using a Welch's test in a similar way as described above for prevalence estimate comparison. We performed two additional sensitivity analyses, we ran the first model on the phenotype schizophrenia (ICD10, F20) instead of SSD, and we ran the last model on the European unrelated subset of iPSYCH2015[41].

To estimate the risk of the three other brain disorders associated with *NRXN1* deletions we fitted a logistic regression model using case status for each of the four iPSYCH disorders (ADHD, ASD, MDD and SSD) as covariates in addition to age, sex (at birth) and SNP array type.

Software

All analyses were performed on HPC running CentOS Linux 7. PLINK[39,40] version 190b6.21, R[46] version 4.0.5 and VCFtools[47] 0.1.17 were installed via the conda package manager (https://anaconda.org/). PennCNV[34] version 1.0.5, bcftools[48] version 1.14, htslib[49] 1.14 are a part of the container we used for the CNV calling described in the previous section, available on Docker Hub (https://hub.docker.com/r/sinomem/docker_cnv_protocol). For the analysis and the figures, we used the following R packages: data.table[50], pROC[42], survey[43], mgcv[45] and ggplot2[51].

Ethics statement

This study is in full compliance with all relevant ethical regulations including the Declaration of Helsinki. Access to the data and its use for research purposes was granted by The Danish Scientific Ethics Committee, the Danish Health Data Authority, the Danish Data Protection Agency, and the DNSB Steering Committee. For this study, the Danish Scientific Ethics Committee has, in accordance with the Act on Research Ethics Review of Health Research Projects (in Danish: *Komitéloven*), waived the need for informed consent in biomedical research based on existing biobanks.

Results

Descriptive statistics and prevalences

After quality control, our sample consisted of 77,655 cases of the four disorders ascertained in iPSYCH2015 (22,167 ASD, 26,186 ADHD, 31,622 MDD, 13,126 SSD) and a population-representative random cohort of 43,311 samples, for a total of 118,427 unique samples. Given the structure of the sample, there is a small overlap between the two groups. Moreover, a given case can be diagnosed with more than one of the index disorders. We called CNVs in the larger *NRXN1* locus (*NRXN1* gene plus 0.5 Mbp upstream and downstream) and performed visual validation as described in the methods. In total 1387 calls were evaluated, of those 378 were deemed as true CNVs, 573 as false calls, and 436 as unknown (meaning no definitive judgement was possible, most often due to the small number of markers available). Given the small proportion of duplications (21 out of 378) and the low reliability of validating small duplications, we discarded duplications from all subsequent analyses and focused on deletions only. This resulted in a total of 357 carriers (255 cases, 102 controls) of which 135 (108 cases, 27 controls) were exonic, i.e., overlapping at least one exon.

The prevalence of *NRXN1* deletions in the general Danish population is 2.55 (95% CI: 2.13–3.04) per 1000 individuals and 0.70 (95% CI: 0.50–0.98) when restricting to exonic deletions. This is almost two times higher than what was previously reported in UKB[44], 0.70 vs 0.39 per 1000 individuals (*p*-value 0.0014, Welch's test). Subgrouping by subcohort (iPSYCH2012 and the extension iPSYCH2015i respectively) the prevalence estimates are 2.20 (95% CI: 1.72–2.81) and 3.07 (95% CI: 2.37–3.98) for any deletion, and 0.78 (95% CI: 0.52–1.17) and 0.58 (95% CI: 0.32–1.05) for exonic deletions only. Supplementary Table 3 provides a prevalence breakdown per gender.

NRXN1 deletions subgrouping

Neither exonic nor non-exonic deletions are distributed uniformly across the locus (Supplementary Fig. 1). In order to disentangle the risk signal in *NRXN1* CNVs further than exonic/non-exonic deletions, we created a set of subgroups. We used a similarity matrix of all CNV pairs (Fig. 1a, b) and a correlation matrix of the deleted exons (Fig. 1c) as described in the methods. Regarding non-exonic CNVs, we identified a clear subgroup of 100 very similar CNVs (IOU > 80%) corresponding to those between exons ENSE00003649136 and ENSE00002460080 (Fig. 1a, b, Supplementary Fig. 1d). The average boundaries of this group of deletions correspond to a deletion previously found segregating in several European populations (Chr2:50,882,153–50,945,699 in Rujescu et al. and Chr2:50,882,111–50,947,645 in this study)[25]. The prevalence of this segregating intronic deletion is 0.77 (95% CI: 0.55–1.06) per 1000 individuals.

Regarding exonic CNVs, the correlation plot (Fig. 1d) shows that exons are affected by deletions essentially in two blocks, exons ENSE00001682911 to ENSE00002460080 (roughly corresponding to the 3′ end of the gene to the group of exons where the promoter of the beta isoform is located, referred to as beta region from now on), and exons ENSE00002453754 to ENSE00001547151 (roughly corresponding to said group of exons to the 5′ end of the gene, referred to as alpha promoter region from now on). See also Supplementary Table 4 and Supplementary Fig. 3 for more details on exonic deletions. The number of carriers in each group was 81 and 54, for the alpha and beta promoter regions, respectively. While smaller clusters are observed within both large groups, further subgrouping of these two main clusters resulted in limited study power, thus we only used these two main clusters for further analysis.

NRXN1 deletions and associated risk of psychiatric disorders

To estimate the association between *NRXN1* deletions and the risk of the four index psychiatric disorders (ADHD, ASD, MDD, SSD) we conducted three separate analyses based on the deletion subgroups described above. As

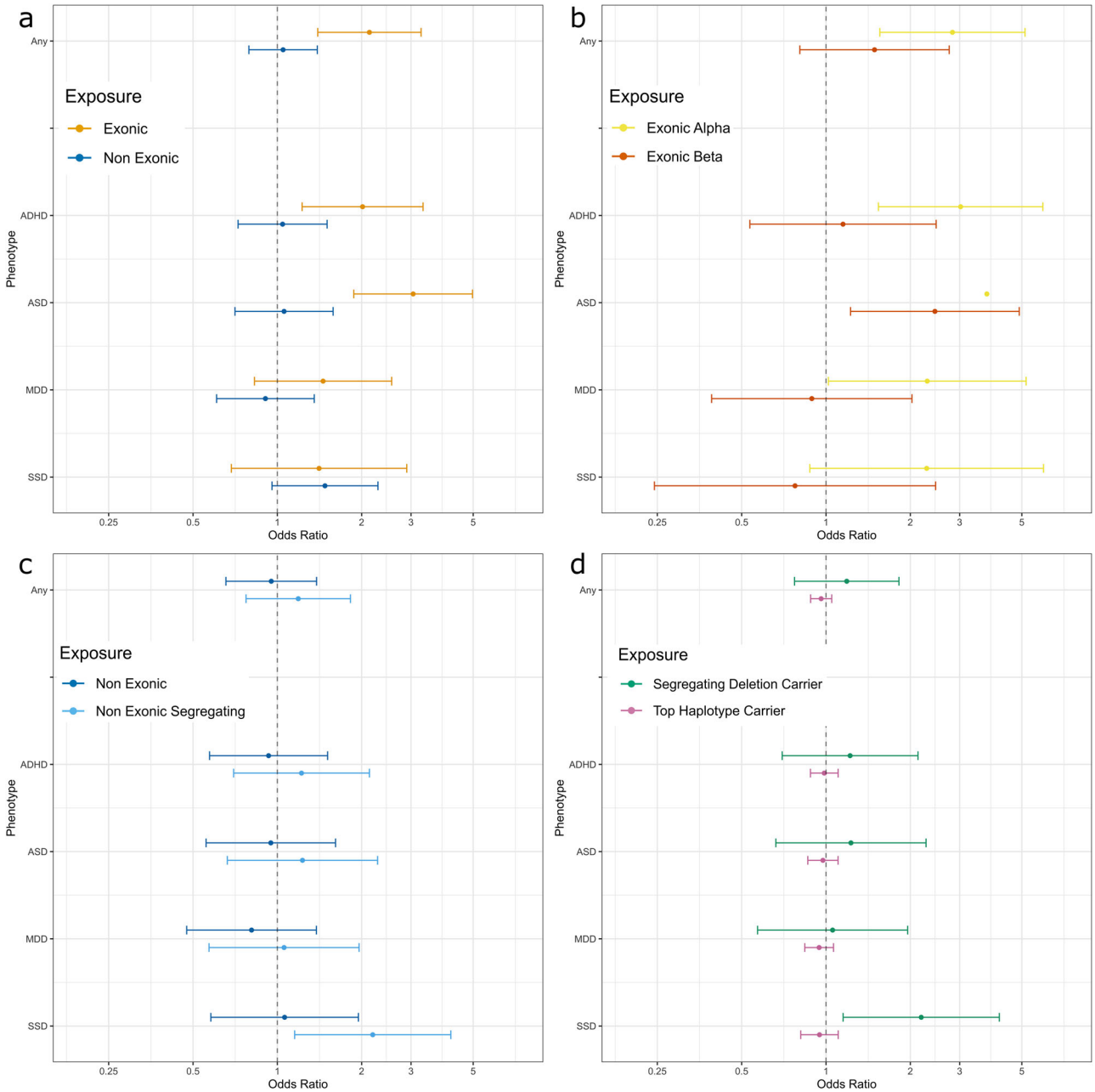

Fig. 2 | Forest plots showing the ORs resulting from three logistic regression analyses on four neurodevelopmental disorders. a First model, ORs for exonic and non-exonic deletions in the *NRXN1* locus. **b**, **c** Second model, exonic deletion is divided into three subgroups based on the exons they overlap (alpha promoter region, beta promoter region, at least one of both) and non-exonic are divided into two subgroups (those belonging to the segregating deletion and all the rest). Note that the scale of (**b**) differs from the rest. **d** Third model, ORs for being a carrier of the segregating deletion or of the haplotype associated with the deletion but without such deletion.

described in the methods, we used a logistic model adjusting for age, SNP array type and sex. The resulting OR estimates and carrier counts are summarised in Fig. 2 and Table 1. Overall, we see an increased risk of ADHD and ASD associated with carriage of exonic deletions, but not of SSD (also when running the analysis on the stricter schizophrenia phenotype, OR: 1.87, 95% CI: 0.81–4.33) or MDD.

We also attempted to replicate findings of previous studies linking exonic NRXN1 deletions to increased risk of ID[16], epilepsy[17] and TS[18], although these disorders had not been specifically targeted by the iPSYCH case-cohort design and as a consequence our estimates are not as well powered (or population-representative) as for the four index psychiatric disorders (Supplementary Table 5). As shown in Table 2, we replicate the previous reports for ID and epilepsy, but not for TS. In all instances, (both for the four index psychiatric disorders and the three other brain disorders) our risk estimates are lower than reported in the case-control studies that we draw comparisons with, although not significantly so except for SSD and TS (Table 2). When we used the stricter SCZ diagnosis (ICD:F20) the difference with the comparison study[12] was not significant ($P = 0.15$; Table 2).

When subgrouping CNVs, deletions in the alpha promoter region of the gene appear to carry the majority of the signal. This is in accordance with previous literature both based on case-control studies as well as in vitro studies[5,38].

While we observed no association between exonic deletions and risk of SSD (OR = 1.40; 95% CI: 0.68–2.89), this diagnosis group was the only one where we observed a significant increase in risk associated with intronic deletions. As shown in Fig. 2c, this association seems to be driven by the

Table 2 | Comparison of effect sizes for exon-disrupting NRXN1 deletions between iPSYCH2015 and published case-control studies

Psychiatric outcome	iPSYCH2015			Comparison study[a]			Welch's test[b]	
	OR (CI95%)	P	N[c]	OR (CI95%)	P	N[c]	d (se)	P_d
ADHD	2.01 (1.22–3.32)	0.0057	26,186 (0.15%) 40,626 (0.066%)	4.68 (1.82–10.6)	0.00093	8883 (0.1%) 180,809 (0.021%)	0.84 (0.52)	0.10
ASD	3.06 (1.88–4.95)	7.4×10^{-6}	22,167 (0.23%) 40,626 (0.066%)	7.24 (0.93–326)	0.036	2558 (0.27%) 2670 (0.037%)	0.81 (1.52)	0.57
MDD	1.46 (0.83–2.56)	0.19	31,622 (0.10%) 40,626 (0.066%)	2.01 (1.18–3.19)	0.0057	23,979 (0.079%) 383,095 (0.039%)	0.32 (0.38)	0.41
SSD	1.41 (0.69–2.90)	0.35	13,126 (0.091%) 40,626 (0.066%)	4.50 (2.03–10.9)	2.8×10^{-5}	20,403 (0.15%) 26,628 (0.034%)	1.16 (0.56)	0.040
ID	2.68 (1.65–4.34)	6.7×10^{-5}	5975 (0.38%) 40,626 (0.066%)	8.14 (2.91–22.7)	<0.0001	19,263 (0.21%) 15,264 (0.026%)	1.11 (0.58)	0.055
Epilepsy	1.94 (1.01–3.73)	0.046	3957 (0.25%) 40,626 (0.066%)	9.91 (1.92–51.1)	0.0049	1569 (0.32%) 6201 (0.032%)	1.63 (0.90)	0.070
TS	1.53 (0.65–3.56)	0.33	2222 (0.27%) 40,626 (0.066%)	20.3 (2.6–156)	5.9×10^{-5}	2434 (0.49%) 4093 (0.033%)	2.59 (1.13)	0.022

Comparison of effect sizes between iPSYCH2015 and published case-control studies

[a]Risk estimates for exonic deletions in iPSYCH2015 were compared with estimates from the largest available published case-control studies for attention-deficit/hyperactivity disorder (ADHD)[14], autism spectrum disorder (ASD)[13], major depressive disorder (MDD)[15], schizophrenia spectrum disorder (SSD)[12], intellectual disability (ID)[16], epilepsy[17], and Tourette syndrome (TS)[18].

[b]The comparison was done through a Welch's test, with d (se) denoting the absolute difference in estimates ($|\log(OR_1/OR_2)|$) and standard error thereof ($\sqrt{(SE_1^2 + SE_2^2)}$), and P_d indicating the significance of the difference (2*(1-pnorm(d/(se)))). The difference in risk estimates between iPSYCH2015 and Rees et al. (fourth row from top) was not significant when using iPSYCH2015 estimates for narrowly defined (ICD10;F20) schizophrenia (OR (CI95%) = 1.87 (0.81–4.33), d (se) = 0.88 (0.61), $P_d = 0.15$).

[c]Above; number of affected (% of affected with an exonic deletion in NRXN1 gene) – Below; number of unaffected (% of unaffected with an exonic deletion in NRXN1 gene).

segregating intronic deletion described above (*OR* = 2.20; 95% CI: 1.15–4.18). Since intronic deletions are usually not considered pathogenic, we hypothesised that the risk associated with the segregating deletion could be explained by another variation co-segregating with it. As described in the methods, we ran a simple association test between all SNPs in chromosome 2 and the recurrent deletion. Using the 10 most associated SNPs we constructed all two-to-five SNPs haplotypes, and we identified the most characteristic haplotype with an AUC of 0.94 (rs10205006-T, rs7608415-G, rs62140665-C, rs17041353-G). We then ran a final analysis grouping samples based on whether they were carriers of this haplotype or not. The results, shown in Fig. 2d, confirm that this deletion is only associated with an increased risk of SSD and, notably, that the associated risk is confined to the deletion ($n = 100$) and not observed among carriers of the underlying haplotype without the deletion ($n = 2341$). However, we do not observe a significantly increased risk of SSD associated with this deletion when we restrict the sample to the European unrelated subset (OR: 1.8, 95% CI: 0.8–3.8). Finally, given the high number of analyses we performed multiple testing corrections (FDR, adjusted *p*-values are provided in Table 1). As expected, the strongest association reported in this study, namely ASD and ADHD with exonic deletions in the NRXN1 locus, remains significant after the correction. However, the SSD association with the segregating intronic deletion did not remain significant after correction.

Discussion

Deletions affecting the *NRXN1* gene have been investigated for associations with psychiatric and developmental disorders for almost twenty years. CNVs in the *NRXN1* locus can be very heterogeneous, affecting one or more exons, besides occurring between two exons. Exonic deletions in particular have been associated with SDD[12,25], ADHD[14], MDD[15] and ASD[13]. However, most of the published studies have been limited to smaller case-control samples or meta-analyses of case-control samples. Moreover, intronic deletions are usually discarded from the analysis[11,25,44]. In this study, we attempt to disentangle the risk profile of exonic as well as intronic deletions defining subgroups of similar deletions. Using the population-representative case-cohort design of iPSYCH2015, we report unbiased estimates of the population prevalence and association of such subtypes of deletions with four core psychiatric disorders.

As in previous studies on the same cohort[26–28], we find the prevalence in the general population to be higher and the risk associations to be lower than previously reported. We observe exonic deletions to be associated with ASD and ADHD. When subgrouping deletions based on location in the gene, the association is driven by deletions in the alpha promoter region of the gene, while deletions in the other half of the gene are rarer and possibly associated with less increased risk of psychiatric disorders. Notably, CNVs in the alpha promoter region are known to be more frequent and indeed are in our sample as well. The association appears robust, suggesting a biological reason for the excess risk in one proportion of the gene. However, it may also be exacerbated by the difference in number of carriers. We also confirm the presence of a small segregating deletion that does not affect any exon and find it to be potentially linked to SSD. While this signal did not survive multiple testing corrections, we believe it can be taken as an indication that intronic CNVs should not be discarded a priori in this kind of analysis.

Notably, we do not find exonic deletions in the *NRXN1* locus to be associated with an increased risk of SSD, which at first glance seems in strong contrast with previous reports[11,12,25,52–55]. However, when we examine the methodology and timeline of these previous reports, a more conciliatory picture emerges. The first large-scale study of schizophrenia-associated risk with exon-disrupting NRXN1 deletions was that of Rujescu et al.[25], who reported an OR of 9 in a meta-analysis of European samples including ~3000 cases and >30,000 controls. Most subsequent studies derived their risk estimates either fully[11,52,55] or in part[53] by merging all schizophrenia cases and controls from previously published studies and performing a simple Fisher's exact test on the pooled sample. As a consequence, in all these studies a large fraction of the control individuals (40%–80%) are those from the original report by Rujescu et al.[25], whereas most case individuals are from other studies, most often applying denser arrays than the HumanHap300 array used in Rujescu et al.[25]. As NRXN1 deletions vary widely in size and breakpoints, the approach taken in these studies is very vulnerable to batch effects owing to differing resolution to detect exon-disrupting deletions across different genotyping platforms.

Since the initial report of Rujescu et al. only two other large-scale studies (Rees et al.[12], and Marshall et al.[54]) have been published that do not include the large control sample of Rujescu et al. Both these studies report slightly lower carrier rates in cases (0.15% and 0.11%) and higher carrier rates in controls (0.034% and 0.020%) than Rujescu et al. (0.24% in cases and

0.015% in controls), and when meta-analysing across genotyping platforms, both studies correspondingly report lower odds ratios (4.5 and 5.8, respectively). These estimates are still higher than we find in iPSYCH2015, as is also the case for the other three core iPSYCH2015 disorders. This could in part be due to case ascertainment; iPSYCH2015 relies on hospital-based diagnoses from national registers, without any further confirmation of case status. However, the carrier frequency among iPSYCH2015 cases is very similar to those reported by the largest previously published studies for each disorder. In contrast, the population-based prevalence of exon-disrupting NRXN1 deletions in iPSYCH2015 is twice as high as reported in UKB[15,44] and the control samples used in Rees et al.[12] and Girirajan et al.[13], and more than three times higher than among the controls of Gudmundsson et al.[14] This is in line with results of our previous CNV studies involving iPSYCH2015 and suggests that the overall tendency for lower CNV-associated risk estimates in iPSYCH2015 is in large part explained by the higher CNV prevalence in the general population compared to individuals used as controls in other studies.

The sample size is the major limitation of this study. Although *NRXN1* is a hotspot for non-recurrent CNVs, such events are rare. For this reason, we lacked the power to include duplications in the study or subgroup deletions beyond the two major groups. Also, both the relatively young age of participants and the specific focus on a limited number of psychiatric disorders in the iPSYCH case-cohort design limits our study power for the later-onset iPSYCH disorders (MDD and SSD) as well as other brain disorders not targeted by the study design (such as ID, epilepsy and TS). Some of the individuals from the random subcohort will later go on to develop MDD or SSD, which in the case of MDD, with its high lifetime prevalence of 10–15%, could have had an attenuating effect on the estimated OR, while it is unlikely to have had affected the risk estimate for SSD, with its much lower lifetime prevalence (1.0–1.5%). As for the brain disorders not targeted by the case-cohort design, the case sample sizes are relatively small and enriched with individuals with comorbid ADHD, ASD, MDD and/or SSD. To account for this enrichment, while also retaining the maximum case sample size, we fitted a logistic model that included each of the four iPSYCH disorders as covariates. While maximising study power, this approach probably leads to an overestimate of case carrier frequency but at the same time an underestimate of the associated OR for these disorders.

Notwithstanding these limitations, our results add important insight into the association between *NRXN1* deletions and the risk of psychiatric illness. Most importantly, we show that the risk is mainly driven by deletions disrupting exons specific to the alpha isoform of Neurexin 1. Also, we show that as with recurrent CNVs, previous case-control studies of *NRXN1* deletions have likely underestimated their population prevalence and consequently overestimated their associated risk. Finally, we characterise the haplotype background of a previously reported intronic deletion segregating at ~0.1% carrier frequency in the Danish population, and while inconclusive, our results warrant further study into its possible association with psychiatric and/or other cognitive/behavioural traits.

Data availability

Regarding access to study data (other than sensitive person-level data, which by requirement of the data custodian and Danish legislation cannot be shared) please contact the corresponding author.

Received: 4 June 2024; Accepted: 26 November 2024;
Published online: 19 December 2024

References

1. Südhof, T. C. Synaptic neurexin complexes: a molecular code for the logic of neural circuits. *Cell* **171**, 745–769 (2017).
2. Reissner, C., Runkel, F. & Missler, M. Neurexins. *Genome Biol.* **14**, 213 (2013).
3. GTEx Portal. https://gtexportal.org/home/.
4. Lonsdale, J. et al. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
5. Flaherty, E. et al. Neuronal impact of patient-specific aberrant NRXN1α splicing. *Nat. Genet.* **51**, 1679–1690 (2019).
6. Jenkins, A. K. et al. Neurexin 1 (NRXN1) splice isoform expression during human neocortical development and aging. *Mol. Psychiatry* **21**, 701–706 (2016).
7. Fuccillo, M. V. & Pak, C. Copy number variants in neurexin genes: phenotypes and mechanisms. *Curr. Opin. Genet. Dev.* **68**, 64–70 (2021).
8. Cunningham, F. et al. Ensembl 2022. *Nucleic Acids Res.* **50**, D988–D995 (2022).
9. Castronovo, P. et al. Phenotypic spectrum of *NRXN1* mono- and bi-allelic deficiency: a systematic review. *Clin. Genet.* **97**, 125–137 (2020).
10. Béna, F. et al. Molecular and clinical characterization of 25 individuals with exonic deletions of *NRXN1* and comprehensive review of the literature. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **162**, 388–403 (2013).
11. Kirov, G. et al. Neurexin 1 (NRXN1) deletions in schizophrenia. *Schizophr. Bull.* **35**, 851–854 (2009).
12. Rees, E. et al. Analysis of intellectual disability copy number variants for association with schizophrenia. *JAMA Psychiatry* **73**, 963–969 (2016).
13. Girirajan, S. et al. Refinement and discovery of new hotspots of copy-number variation associated with autism spectrum disorder. *Am. J. Hum. Genet.* **92**, 221–237 (2013).
14. Gudmundsson, O. O. et al. Attention-deficit hyperactivity disorder shares copy number variant risk with schizophrenia and autism spectrum disorder. *Transl. Psychiatry* **9**, 258 (2019).
15. Kendall, K. M. et al. Association of rare copy number variants with risk of depression. *JAMA Psychiatry* **76**, 818–825 (2019).
16. Lowther, C. et al. Molecular characterization of NRXN1 deletions from 19,263 clinical microarray cases identifies exons important for neurodevelopmental disease expression. *Genet. Med. J. Am. Coll. Med. Genet.* **19**, 53–61 (2017).
17. Møller, R. S. et al. Exon-disrupting deletions of NRXN1 in idiopathic generalized epilepsy. *Epilepsia* **54**, 256–264 (2013).
18. Huang, A. Y. et al. Rare copy number variants in NRXN1 and CNTN6 increase risk for tourette syndrome. *Neuron* **94**, 1101–1111 (2017).
19. Enggaard Hoeffding, L. K. et al. Sequence analysis of 17 NRXN1 deletions. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **165B**, 52–61 (2014).
20. Wilson, T. E. et al. Large transcription units unify copy number variants and common fragile sites arising under replication stress. *Genome Res.* **25**, 189–200 (2015).
21. Kirov, G. et al. Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum. Mol. Genet.* **17**, 458–465 (2008).
22. Zahir, F. R. et al. A patient with vertebral, cognitive and behavioural abnormalities and a de novo deletion of NRXN1alpha. *J. Med. Genet.* **45**, 239–243 (2008).
23. Kim, H.-G. et al. Disruption of neurexin 1 associated with autism spectrum disorder. *Am. J. Hum. Genet.* **82**, 199–207 (2008).
24. Marshall, C. R. et al. Structural variation of chromosomes in autism spectrum disorder. *Am. J. Hum. Genet.* **82**, 477–488 (2008).
25. Rujescu, D. et al. Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum. Mol. Genet.* **18**, 988–996 (2009).
26. Olsen, L. et al. Prevalence of rearrangements in the 22q11.2 region and population-based risk of neuropsychiatric and developmental disorders in a Danish population: a case-cohort study. *Lancet Psychiatry* **5**, 573–580 (2018).
27. Calle Sánchez, X. et al. Comparing copy number variations in a Danish Case Cohort of Individuals With Psychiatric Disorders. *JAMA Psychiatry* **79**, 59–69 (2022).
28. Vaez, M. et al. Population-Based Risk of Psychiatric Disorders Associated With Recurrent Copy Number Variants. *JAMA Psychiatry* **81**, 957–966 (2024).

29. Bybjerg-Grauholm, J. et al. The iPSYCH2015 Case-Cohort sample: updated directions for unravelling genetic and environmental architectures of severe mental disorders. Preprint at *medRxiv* https://doi.org/10.1101/2020.11.30.20237768 (2020).
30. Pedersen, C. B. et al. The iPSYCH2012 case–cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Mol. Psychiatry* **23**, 6–14 (2018).
31. Pedersen, C. B. The Danish Civil Registration System. *Scand. J. Public Health* **39**, 22–25 (2011).
32. Mors, O., Perto, G. P. & Mortensen, P. B. The Danish Psychiatric Central Research Register. *Scand. J. Public Health* **39**, 54–57 (2011).
33. Schmidt, M. et al. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin. Epidemiol.* **7**, 449–490 (2015).
34. Wang, K. et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17**, 1665–1674 (2007).
35. Montalbano, S. et al. Accurate and effective detection of recurrent copy number variants in large SNP genotype datasets. *Curr. Protoc.* **2**, e621 (2022).
36. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
37. UniProt Consortium, The UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531 (2023).
38. Cosemans, N. et al. The clinical relevance of intragenic NRXN1 deletions. *J. Med. Genet.* **57**, 347–355 (2020).
39. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).
40. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
41. Appadurai, V. et al. Accuracy of haplotype estimation and whole genome imputation affects complex trait analyses in complex biobanks. *Commun. Biol.* **6**, 1–12 (2023).
42. Robin, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
43. Lumley, T. Analysis of complex survey samples. *J. Stat. Softw.* **9**, 1–19 (2004).
44. Crawford, K. et al. Medical consequences of pathogenic CNVs in adults: analysis of the UK Biobank. *J. Med. Genet.* **56**, 131–138 (2019).
45. Wood, S. N., Pya, N. & Saefken, B. Smoothing parameter and model selection for general smooth models (with discussion). *J. Am. Stat. Assoc.* **111**, 1548–1563 (2016).
46. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2021). https://www.R-project.org/.
47. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
48. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
49. Bonfield, J. K. et al. HTSlib: C library for reading/writing high-throughput sequencing data. *GigaScience* **10**, giab007 (2021).
50. Barrett, T. et al. data.table: Extension of 'data.frame'. https://CRAN.R-project.org/package=data.table (2024).
51. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, New York, 2016).
52. Dabell, M. P. et al. Investigation of NRXN1 deletions: clinical and molecular characterization. *Am. J. Med. Genet. A* **161A**, 717–731 (2013).
53. Rees, E. et al. Analysis of copy number variations at 15 schizophrenia-associated loci. *Br. J. Psychiatry* **204**, 108–114 (2014).
54. Marshall, C. R. et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat. Genet.* **49**, 27–35 (2017).
55. Hu, Z. et al. Genetic insights and neurobiological implications from NRXN1 in neuropsychiatric disorders. *Mol. Psychiatry* **24**, 1400–1414 (2019).

Acknowledgements

This work was supported by grants from the Lundbeck Foundation (R165–2013–15320, R155–2014–1724, and R248–2017–2003) and the NIH (R01 MH124789-01).

Author Contributions

Concept and design: Montalbano, Krebs, Rosengren, Vaez, Helenius, Ingason. Acquisition, analysis or interpretation of data: Montalbano, Ingason. Drafting of the manuscript: Montalbano, Ingason. Crtitical review of the manuscript for important intellectual content: Montalbano, Krebs, Rosengren, Vaez, Hellberg, Mortensen, Børglum, Geschwind, Raznahan, Thompson, Helenius, Werge, Ingason. Statistical analysis: Montalbano, Vaez, Thompson Helenius, Ingason. Obtained funding: Mortensen, Børglum, Raznahan, Werge, Ingason. Administrative, technical or material support: Montalbano, Rosengren, Werge, Ingason. Supervision: Werge, Ingason.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41525-024-00450-8.

Correspondence and requests for materials should be addressed to Andrés. Ingason.

iPSYCH Investigators

Anders D. Børglum[2,4,5], David M. Hougaard[13], Merete Nordentoft[14], Ole Mors[15], Preben B. Mortensen[2,3], Thomas Werge[1,2,12], Jakob Grove[2,4,5,16], Thomas D. Als[2,4,5], Alfonso Buil[1,2], Anders Rosengren[1,2], Andrés Ingason[1,2] ✉, Andrew J. Schork[1,2], Dorte Helenius[1,2], Jesper Gådin[1], Richard Zetterberg[1], Vivek Appadurai[1], Joeri Meijsen[1], Kajsa-Lotta Georgii Hellberg[1,2], Bjarni J. Vilhjálmsson[3,16], Carsten B. Pedersen[3], Esben Agerbo[3], Jakob Christensen[3], Liselotte V. Petersen[3], Marianne Giørtz Pedersen[3], Jonas Bybjerg-Grauholm[13] & Marie Bækvad-Hansen[13]

[13]Department for Congenital Disorders, Statens Serum Institute, Copenhagen, Denmark. [14]Mental Health Centre Copenhagen, Capital Region of Denmark, Copenhagen University Hospital, Copenhagen, Denmark. [15]Psychosis Research Unit, Aarhus University Hospital-Psychiatry, Aarhus, Denmark. [16]BiRC, Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark.