Short Title: Synaptic and Gene Regulatory Mechanisms in 22q11.2 CNVs

Synaptic and Gene Regulatory Mechanisms in Schizophrenia, Autism, and 22q11.2 CNV Mediated Risk for Neuropsychiatric Disorders

Authors

Jennifer K. Forsyth[1], Daniel Nachun[1], Michael J. Gandal[1,2], Daniel H. Geschwind[2,3,4], Ariana E. Anderson[1], Giovanni Coppola[1,3,4], Carrie E. Bearden*[,1,4,5]

Affiliations

[1]Department of Psychiatry & Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA 90095

[2]Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA 90095

[3]Department of Neurology, University of California, Los Angeles, Los Angeles, CA 90095

[4]Brain Research Institute, University of California, Los Angeles, Los Angeles, CA 90095

[5]Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095

*Corresponding Author

Correspondence to Carrie E Bearden, Ph.D.

A7-460 Semel Institute

Los Angeles, CA 90095

(310) 206-2983

cbearden@mednet.ucla.edu

Abstract Word Count: 244

Main Text Word Count: 4761

Number of Figures: 7

Number of Tables: 2

Supplements: 1 supplementary information file + 1 excel file containing 2 supplementary data tables

Keywords: 22q11.2 copy number variants, schizophrenia, autism spectrum disorders, genetic risk, functional genomics, brain development

Abstract

Background: 22q11.2 copy number variants (CNVs) are among the most highly penetrant genetic risk variants for developmental neuropsychiatric disorders such as schizophrenia (SCZ) and autism spectrum disorder (ASD). However, the specific mechanisms through which they confer risk remain unclear.

Methods: Using a functional genomics approach, we integrated transcriptomic data from the developing human brain, genome-wide association findings for SCZ and ASD, protein interaction data, and pathophysiological signatures of SCZ and ASD to: 1) organize genes into the developmental cellular and molecular systems within which they operate; 2) identify neurodevelopmental processes associated with polygenic risk for SCZ and ASD across the allelic frequency spectrum; and 3) elucidate pathways and individual genes through which 22q11.2 CNVs may confer risk for each disorder.

Results: Polygenic risk for SCZ and ASD converged on partially overlapping gene networks involved in synaptic function and transcriptional regulation, with ASD risk variants additionally enriched for networks involved in neuronal differentiation during fetal development. The 22q11.2 locus formed a large protein network that disproportionately affected SCZ- and ASD-associated neurodevelopmental networks, including loading highly onto synaptic and gene regulatory pathways. *SEPT5*, *PI4KA*, and *SNAP29* genes are candidate drivers of 22q11.2 synaptic pathology relevant to SCZ and ASD, and *DGCR8* and *HIRA* are candidate drivers of disease-relevant alterations in gene regulation.

Conclusions: The current approach provides a powerful framework to identify neurodevelopmental processes affected by diverse risk variants for SCZ and ASD and elucidate mechanisms through which highly penetrant multi-gene CNVs contribute to disease risk.

Introduction

Copy number variants (CNVs) at chromosome 22q11.2 are among the genetic variants most strongly associated with developmental neuropsychiatric disorders. The hemizygous deletion, 22q11.2 deletion syndrome (22q11DS), typically spans a gene-rich ~2.5 Mb region(1) and is the greatest known single genetic risk factor for schizophrenia (SCZ). Approximately 25% of 22q11DS patients develop a psychotic disorder by adulthood and ~0.5-1% of patients with SCZ have 22q11DS(2). Interestingly, the reciprocal duplication (22q11DupS) may protect against SCZ(3,4), whereas both 22q11DS and 22q11DupS increase risk for autism spectrum disorders (ASD;5). As psychiatry increasingly looks towards precision medicine approaches to therapeutic development, clarifying the mechanisms through which 22q11.2 CNVs confer risk for SCZ and ASD, and how these mechanisms converge with risk derived from broader variants associated with each disorder, is a critical goal.

Our knowledge of the genetic architectures of both SCZ and ASD has advanced considerably in the last decade. It is now clear that SCZ and ASD are both highly polygenic and genetically heterogeneous disorders, with risk variants for each disorder spanning common (i.e., present in >1% of the population), rare (i.e., present in <1% of the population), and de novo variants (i.e., new mutations in offspring), in addition to CNVs(6). Despite this polygenicity, risk variants for SCZ and ASD each appear to converge on a subset of biological pathways. Neuronal and synaptic genes have been repeatedly implicated in SCZ, including the N-methyl-D-aspartate receptor (NMDAR) and activity-regulated cytoskeleton-associated protein (Arc) complexes, and targets of the fragile X mental retardation protein (FMRP;4,7–11). Similarly, neuronal and synaptic genes have been implicated in ASD, along with chromatin and transcription regulators, and FMRP targets(5,12–16). However, many genes in these gene-sets are developmentally regulated, and SCZ and ASD are both neurodevelopmental in origin. Brain development is tightly controlled through changes in transcription factor activity and chromatin condensation that alter gene expression patterns at precise times in development(17). These changes program the progression of stem cell proliferation, neuronal differentiation, and synaptogenesis during prenatal development, and myelination and synaptic refinement during postnatal development. Examining the convergence of SCZ and ASD risk variants on groups of developmentally co-expressed genes may therefore yield fundamental insights into the neurodevelopmental pathogenesis of each disorder (e.g.,15,16), and provide a powerful lens to understand how 22q11.2 CNVs confer risk for each disorder.

Here, we used network analyses of transcriptomic data from the developing human brain(17) to organize genes into the developmental cellular and molecular systems within which they operate(18). We then conducted a comprehensive examination of neurodevelopmental processes associated with risk variants for SCZ and ASD across the allelic frequency spectrum. Finally, we identified the brain developmental 22q11.2 protein network and developed a

framework to leverage this disease-informed lens to predict pathways and individual genes underlying 22q11.2 CNV-mediated risk for SCZ and ASD.

Methods and Materials

Defining Neurodevelopmental Co-Expression Networks

Networks of developmentally co-expressed genes were constructed using weighted gene co-expression network analysis (WGCNA) of BrainSpan data, which is the largest, publicly-available developmental brain transcriptome dataset derived from healthy human donors(17). Samples spanning the full period of brain development (i.e., ~6 post-conception weeks-30 years; 1061 samples) were included in the analysis. Briefly, correlation coefficients were computed between expression levels for each gene, and connectivity measures were calculated by summing the connection strength of each gene with every other gene. Genes were clustered based on connectivity to identify groups of co-expressed genes (i.e., modules), and the mean expression of the first principal component of each module (i.e., module eigengene) per developmental stage was used to summarize the developmental trajectory of each module. Genes in each module were annotated for enrichment of: 1) brain cell-type markers(19); 2) gene ontology (GO) using g:Profiler(20); and 3) genes expressed in specific human brain regions and developmental periods using the Specific Expression Analysis tool(21). See Supplementary Methods for details.

Identifying Neurodevelopmental Networks Associated with SCZ and ASD

Summary statistics, gene-lists, and/or variant lists from genome-wide association studies of SCZ and ASD were compiled and tested for module enrichment using methods appropriate to each data type. Briefly, when summary level statistics for common(8,11,22) and rare variants(23) were available, they were tested for enrichment using MAGMA(24). De novo mutations (DNMs) were compiled for SCZ(9,25–32) and ASD(33,34); the affected gene and predicted consequence of each DNM were annotated using Ensembl Variant Effect Predictor (VEP) and tested for enrichment using the DenovolyzeR package(35). Genes in SCZ-(4) and ASD-associated(14) CNV loci were tested for enrichment using logistic regression, controlling for gene length. Genes containing more deleterious ultra-rare variants in SCZ patients than controls(10); manually curated high-confidence genes implicated in ASD according to the SFARI gene database(36); and genes associated with ASD across rare inherited, de novo, and copy number variants(14) were formulated as hypergeometric distributions and tested for module enrichment using Fisher's exact tests with the GeneOverlap package(37). See Table 1 for summary of data sources and Supplementary Methods for details.

Characterizing the 22q11.2 Locus

To examine whether any BrainSpan module was over-represented among the 22q11.2 genes, given the number of genes in each module, or whether SCZ- and ASD-associated modules were over-represented overall, logistic regression controlling for gene length was conducted for each module, and for all disease-associated modules together, respectively.

To characterize the total protein-coding gene content of the 22q11.2 locus and its gene content per module relative to all regions of the genome, the 22q11.2 locus was compared to a distribution generated by simulating CNVs of the same size (~2.57 mb) in 250 kb sliding steps across the genome (i.e., 12,149 genome-wide regions), using the GenomicRanges package(38). The percentile ranks of the 22q11.2 locus overall, and per module, were calculated relative to this distribution. A Mann-Whitney U test examined whether the 22q11.2 locus percentile ranks were higher for SCZ- and ASD-associated modules, overall, than for non-associated modules.

Characterizing the 22q11.2 Brain Development PPI Network

Genes form protein-protein interaction (PPI) networks that fulfill critical biological roles, and perturbing protein networks is a common characteristic of a majority of disease-associated mutations(39). However, protein interactions are dependent on the tissue and temporal availability of each protein(40). We therefore identified the brain development protein-protein interaction network (BD-PPI network) for the 22q11.2 locus (i.e., $22q_{BD\text{-}PPI}$ network) and for: 1) all ~2.57 Mb regions of the genome, as described above (i.e., $\text{Genomic-Region}_{BD\text{-}PPI}$ networks); and 2) 10,000 random gene-sets, equal in number of seed BrainSpan genes to the 22q11.2 locus (i.e., $\text{Gene-Set}_{BD\text{-}PPI}$ networks).

The $22q_{BD\text{-}PPI}$ network was first tested for over-representation in any BrainSpan modules, given the number of genes in each module, using hypergeometric overlap tests. To characterize the 22q11.2 region relative to all regions of the genome, and more specifically, the connectivity of 22q11.2 genes within protein networks relative to all BrainSpan genes, we then calculated the $22q_{BD\text{-}PPI}$ network percentile rank per module relative to the $\text{Genomic-Region}_{BD\text{-}PPI}$ and $\text{Gene-Set}_{BD\text{-}PPI}$ networks, respectively. Mann-Whitney U tests examined whether the $22q_{BD\text{-}PPI}$ network percentile ranks were higher overall for SCZ- and ASD-associated modules than for non-associated modules.

To generate BD-PPI networks, protein pairs were compiled from the BioGRID 3.4(41) and InWeb3 PPI databases(42). PPI pairs were thresholded by BrainSpan-derived brain development co-expression pairs (Spearman's $\rho>0.7$;16) to identify high-confidence PPI pairs likely to exist during brain development. BD-PPI networks were identified by extracting the seed genes and corresponding BD-PPI pairs of each genomic region or gene-set. See Supplementary Methods for details.

Prioritizing Disease-Relevant 22q11.2 Genes

To prioritize individual 22q11.2 genes likely to drive risk for SCZ and ASD, the number of genes per BrainSpan module in each individual 22q11.2 gene's BD-PPI network was compared to that for every other BrainSpan gene. BD-PPI networks of genes above the 95th percentile for SCZ- or ASD-associated modules were examined for GO term enrichment using g:Profiler.

Integrating SCZ and ASD Pathophysiological Signatures

Significantly down- and up-regulated genes in SCZ and ASD post-mortem cortex were obtained from (43) and first tested for BrainSpan module enrichment using hypergeometric overlap tests.

The $22q_{BD\text{-}PPI}$ network overlapping down-regulated genes in SCZ and ASD was then characterized relative to the above described $\text{Genomic-Region}_{BD\text{-}PPI}$ and $\text{Gene-Set}_{BD\text{-}PPI}$ networks. Similarly, the number of genes in each individual 22q11.2 gene's BD-PPI network overlapping down-regulated genes in SCZ and ASD was compared to that for every other BrainSpan gene.

Results

Genetic Risk for SCZ and ASD Converge on Partially Overlapping Neurodevelopmental Processes

Eighteen BrainSpan modules were identified across 17,216 genes (Fig. S1A; Table S1; 4 genes were not assigned to any module) with distinct developmental expression trajectories, cell-type enrichment patterns (Fig. S1B), and functional characteristics (Table S2), capturing broad aspects of neurotypical development.

SCZ and ASD risk variant enrichment for BrainSpan modules is summarized in Fig. 1. SCZ-associated variants largely converged on 3 modules. SCZ-associated common variants and CNVs were significantly over-represented in M15 (FDR corrected <0.05 for number of modules), a neuronal marker-enriched module that increased in expression through adolescence and was enriched for genes expressed during postnatal development (Figs. 2A-C, S2M-N). GO analysis indicated that M15 was involved in synaptic transmission, modulation of neurotransmission, and learning and memory. SCZ common variants were additionally over-represented in M13, a neuronal module that increased in expression during prenatal development, with relatively stable postnatal expression (Fig. 2D-F). Consistent with this, M13 was enriched for genes expressed during late fetal and postnatal development (Fig S2K). M13 was enriched for synaptic transmission and regulation of the membrane potential, and its hub genes included *ATP1B1* and *KCNMA1*, which encode components of the Na+/K+ pump and a voltage-gated K+ channel, respectively, that critically regulate the electrochemical gradient(44,45). Thus, M13 appears to be involved in regulating basic neuronal excitability. Finally, SCZ common variants were significantly over-represented in M7, with rare variants showing nominal enrichment. M7 was involved in RNA processing and binding, and its hub genes included *POLR2B*, which encodes a major subunit of RNA polymerase II, the

central enzyme responsible for transcribing mRNA from DNA, and *CREB1*, which encodes the transcription factor CREB, known to regulate neuronal differentiation, survival, and plasticity(46; Fig. 2G-I). M7 showed its highest expression during fetal development but was minimally enriched for genes associated with any specific period of development (Fig. S2G). Together, this suggests that M7 regulates gene expression and RNA processing across development.

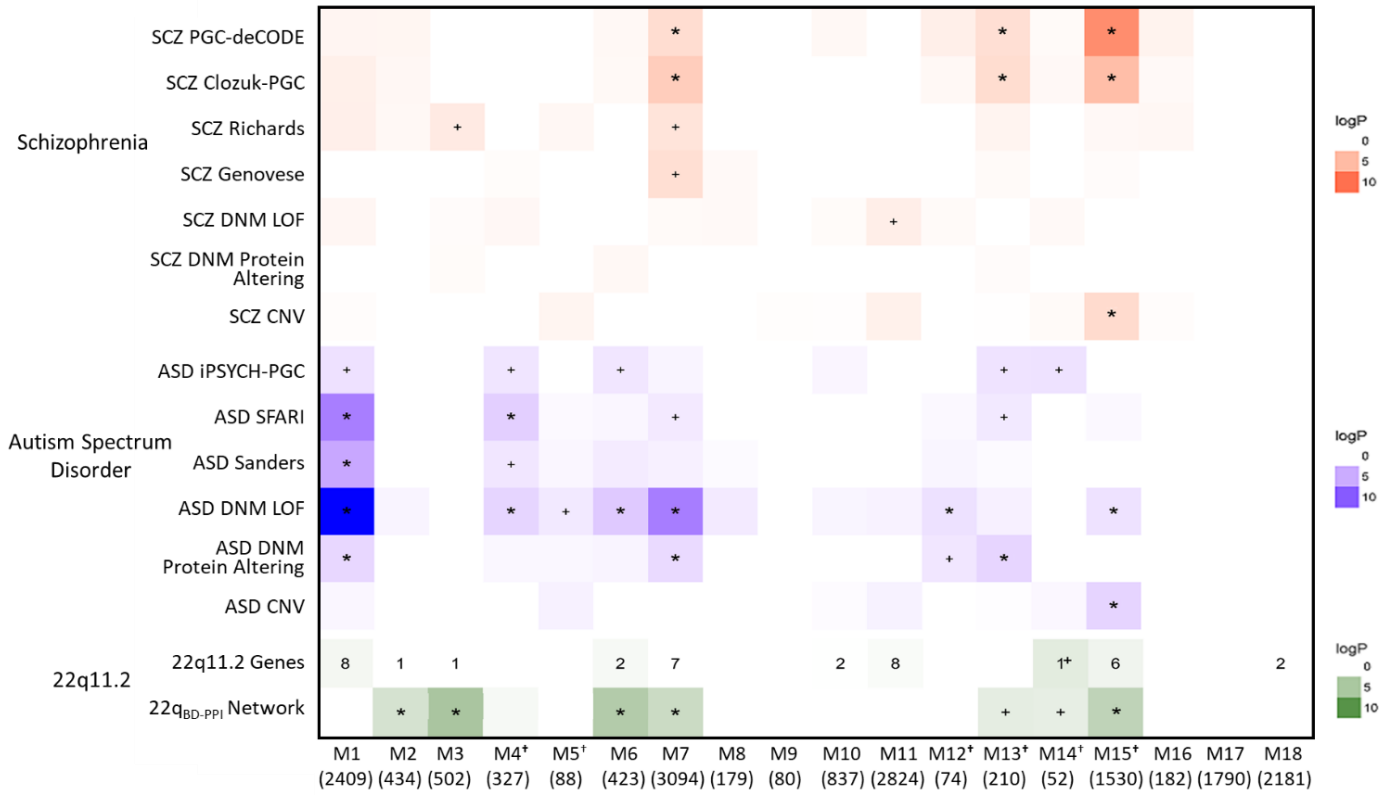
Risk variants for ASD showed overlapping enrichment for SCZ-associated modules, as well as enrichment for 4 distinct modules. Thus, protein-altering DNMs in ASD were over-represented in M13, with common variants and SFARI-defined ASD genes nominally enriched in M13. LoF DNMs in ASD were over-represented in M7, with protein-altering DNMs more broadly and SFARI-defined ASD genes also nominally enriched. LoF DNMs and ASD-associated CNVs (many of which are also associated with SCZ; Table S3) were also over-represented in M15.

Table 1. Genome-wide schizophrenia (SCZ) and autism spectrum disorder (ASD) datasets tested for BrainSpan module enrichment.

Data	Reference	Variant Type	Description	Patient N	Control N	Analysis
SCZ						
SCZ PGC-deCODE*	11	Common	PGC + deCODE meta-analysis	36,989	113,075	MAGMA
SCZ CLOZUK-PGC*	8	Common	CLOZUK + PGC meta-analysis	40,675	64,643	MAGMA
SCZ Richards	23	Rare	Rare variants (MAF<1%)	5,585	8,103	MAGMA
SCZ Genovese	10	Ultra-rare	43 genes enriched for deleterious ultra-rare variants (p<0.01)	4,877	6,242	Hypergeometric Test
SCZ DNM	9, 25-32	De novo	Compilation of DNMs	1,139	NA	DenovolyzeR
SCZ CNV	4	CNV	13 unique loci (FDR<0.05)	21,094	20,227	Logistic Regression
ASD						
ASD iPSYCH-PGC	22	Common	iPSYCH + PGC meta-analysis	18,381	27,969	MAGMA
ASD SFARI	www.sfari.org	Gene list	80 literature-curated genes (score 1 or 2)	NA	NA	Hypergeometric Test
ASD Sanders	14	Rare	65 genes (FDR<0.1)	3,982	1,911	Hypergeometric Test
ASD DNM	33, 34	De novo	Compilation of DNMs	4,424	NA	DenovolyzeR
ASD CNV	14	CNV	9 unique loci, +3 nested loci, in SSC and AGP cohorts (FDR<0.05)	4,687	2,100	Logistic Regression

*The PGC-deCODE meta-analysis includes 1,513 SCZ cases and 36,989 controls from the deCODE study not in the CLOZUK-PGC study and the SCZ CLOZUK-PGC meta-analysis includes 5,220 SCZ cases and 18,823 controls not in the SCZ PGC-deCODE study.

Figure 1. BrainSpan module enrichment for SCZ (red) and ASD (blue) risk variant datasets, 22q11.2 genes, and the 22q11.2 brain-development PPI network ($22q_{BD\text{-}PPI}$ network; green). Log p values are indicated by color for all datasets with enrichment odds ratios >1.0. To show potential convergence of enrichment across datasets, all enrichment values for p<0.05 are annotated (+p<0.05; *FDR corrected p<0.05 for number of modules). Numerical values denote number of 22q11.2 genes in each module. ^Module enriched for neuronal cell-type markers.



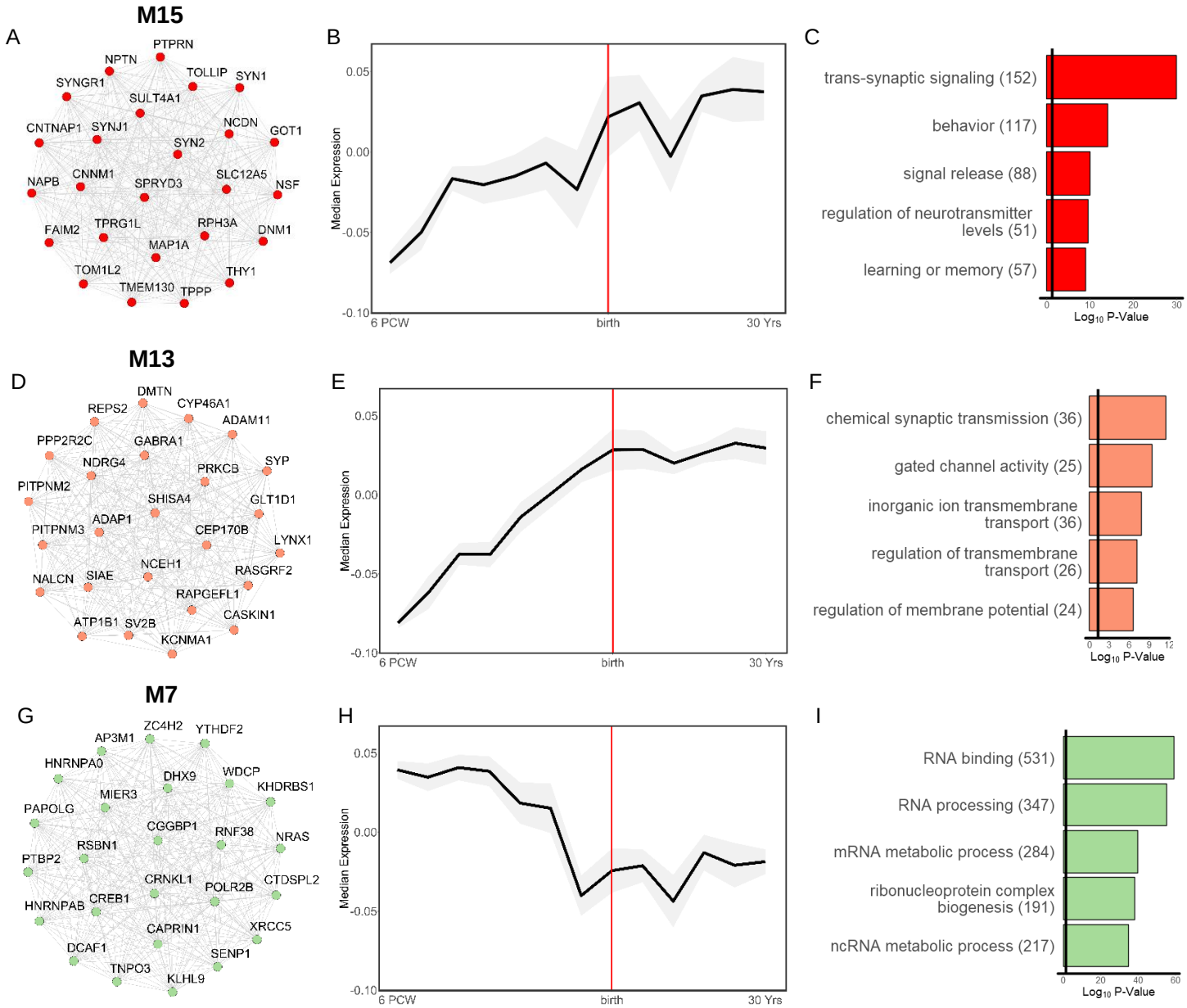
The 4 additional modules significantly associated with ASD were M1, M4, M6, and M12. Thus, rare variants, broadly, and LoF and protein-altering DNMs in ASD were consistently over-represented in M1, with common variants also nominally enriched. M1 peaked in expression during early-mid fetal development and was enriched for genes showing fetal-specific expression (Fig. S2A). M1 was involved in transcription, chromatin modification, and gene expression, and its hub genes included the transcription factors *SOX11* and *SOX4*, which are critical for promoting neuronal differentiation(47; Fig. 2J-L). Further, M1 contained numerous genes encoding subunits of the BAF chromatin remodeling complexes, which program gene expression changes that mediate embryonic stem cell differentiation into neurons(48). Thus, M1 appears to be particularly important for regulating fetal gene expression patterns that direct neuronal differentiation. LoF DNMs and SFARI-defined ASD genes were additionally over-represented in M4, with common and rare variants also nominally enriched. M4 was enriched for neuronal markers, peaked in expression during mid-fetal development, and was involved in neuronal differentiation and neuron projection development (Fig. 2M-O). Consistent

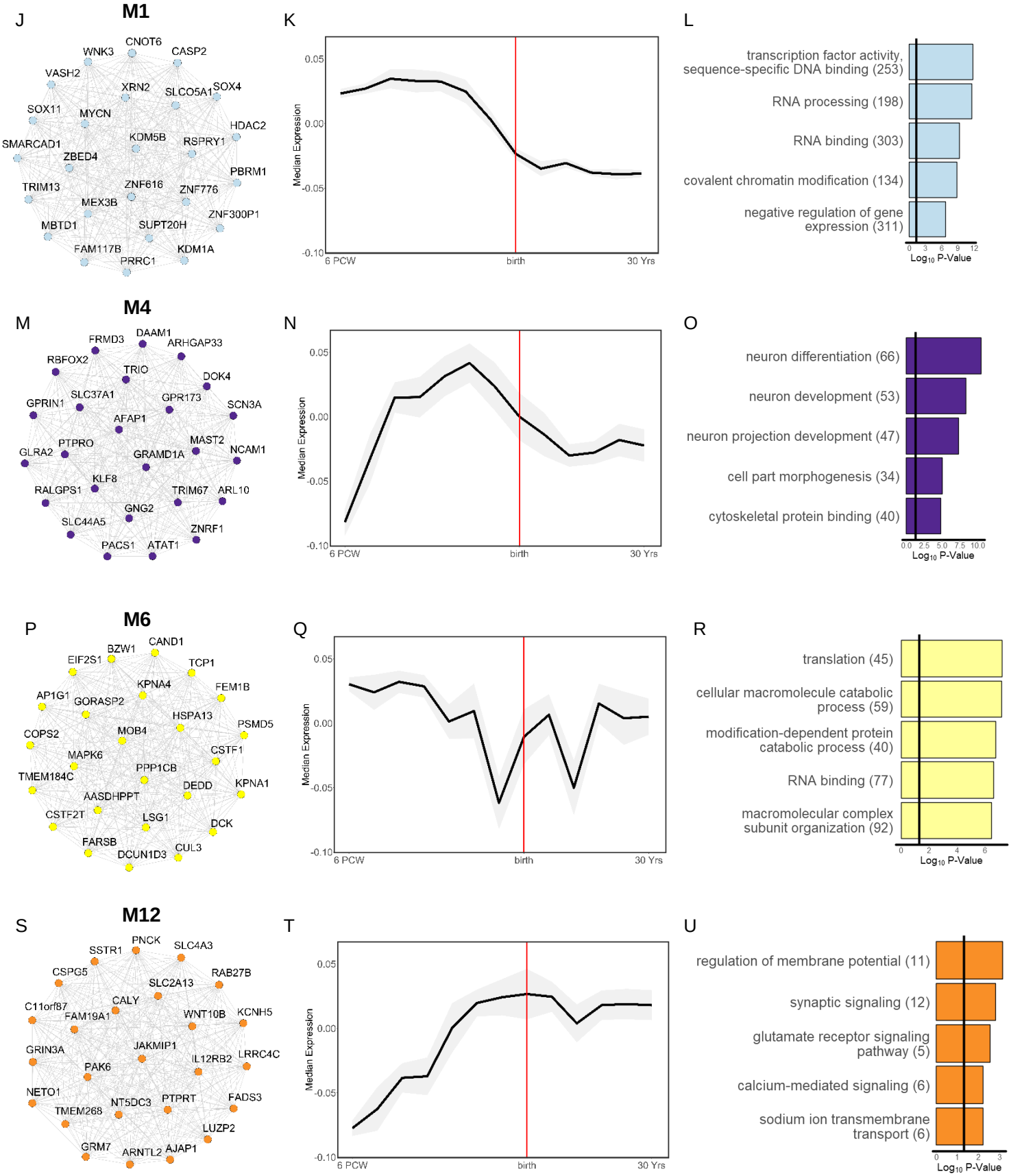
with this, M4 was enriched for genes with fetal-specific expression (Fig. S2C), and its hub genes included DOK4, ARGHAP33, and TRIO, which are involved in axon and dendrite morphogenesis(49–51). LoF DNMs in ASD were also over-represented in M6, with common variants nominally enriched. M6 showed maximal expression during fetal development but was not enriched for genes expressed in any specific period of development (Fig. S2E), and was involved in translation and protein catabolism (Fig. 2P-R). Finally, LoF DNMs in ASD were also over-represented in M12, with protein-altering DNMs, more broadly, nominally enriched. M12, a neuronal module that is closely related to M13 (Fig S1A,B), showed an increasing trajectory of expression until largely plateauing in late fetal development, and was involved in regulating the membrane potential (Fig. 2S-U). Module enrichment for gene-sets implicated in prior genetic studies of SCZ or ASD were consistent with GO analyses (Fig. S1C).

As expected, LoF and protein-altering DNMs in control subjects were not over-represented in any SCZ- or ASD-associated modules, and no modules were significantly enriched for synonymous DNMs in ASD, SCZ, or control subjects ($p>0.05$).

Thus, polygenic risk for SCZ and ASD appear to both coalesce on neurodevelopmental networks involved in postnatal neuronal and synaptic function and pan-developmental regulation of gene expression; however, risk variants for ASD additionally converge on networks involved in early neuronal differentiation and synapse formation.

Figure 2. Left: top 25 hub genes, Middle: developmental expression trajectory (mean ± standard deviation in grey shading), and Right: top 5 significantly enriched biological process and molecular function gene ontology (GO) terms for key SCZ-associated (M15, M13, M7) and ASD-associated (M15, M13, M7, M1, M4, M6, M12) BrainSpan modules. Hub genes and GO terms are colored by module. Edges between hub genes denote positive correlations between genes.




J
M1
CNOT6
WNK3
CASP2
VASH2
XRN2
SLCO5A1
SOX4
SOX11
MYCN
HDAC2
SMARCAD1
KDM5B
RSPRY1
ZBED4
PBRM1
TRIM13
ZNF616
ZNF776
MEX3B
ZNF300P1
MBTD1
SUPT20H
FAM117B
KDM1A
PRRC1
K
Median Expression
0.05
0.00
-0.05
-0.10
6 PCW
birth
30 Yrs
L
transcription factor activity, sequence-specific DNA binding (253)
RNA processing (198)
RNA binding (303)
covalent chromatin modification (134)
negative regulation of gene expression (311)
0 3 6 9 12
Log10 P-Value
M
M4
FRMD3
DAAM1
ARHGAP33
RBFOX2
TRIO
DOK4
SLC37A1
GPR173
GPRIN1
SCN3A
AFAP1
PTPRO
MAST2
GLRA2
GRAMD1A
NCAM1
KLF8
RALGPS1
TRIM67
ARL10
GNG2
SLC44A5
ZNRF1
PACS1
ATAT1
N
Median Expression
0.05
0.00
-0.05
-0.10
6 PCW
birth
30 Yrs
O
neuron differentiation (66)
neuron development (53)
neuron projection development (47)
cell part morphogenesis (34)
cytoskeletal protein binding (40)
0.0 2.5 5.0 7.5 10.0
Log10 P-Value
P
M6
CAND1
BZW1
EIF2S1
TCP1
KPNA4
FEM1B
AP1G1
GORASP2
HSPA13
PSMD5
COPS2
MOB4
CSTF1
MAPK6
PPP1CB
TMEM184C
DEDD
KPNA1
AASDHPPT
CSTF2T
LSG1
DCK
FARSB
CUL3
DCUN1D3
Q
Median Expression
0.05
0.00
-0.05
-0.10
6 PCW
birth
30 Yrs
R
translation (45)
cellular macromolecule catabolic process (59)
modification-dependent protein catabolic process (40)
RNA binding (77)
macromolecular complex subunit organization (92)
0 2 4 6
Log10 P-Value
S
M12
PNCK
SSTR1
SLC4A3
CSPG5
RAB27B
SLC2A13
CALY
C11orf87
FAM19A1
WNT10B
KCNH5
GRIN3A
JAKMIP1
IL12RB2
LRRC4C
PAK6
NETO1
NT5DC3
PTPRT
FADS3
TMEM268
LUZP2
GRM7
ARNTL2
AJAP1
T
Median Expression
0.05
0.00
-0.05
-0.10
6 PCW
birth
30 Yrs
U
regulation of membrane potential (11)
synaptic signaling (12)
glutamate receptor signaling pathway (5)
calcium-mediated signaling (6)
sodium ion transmembrane transport (6)
0 1 2 3
Log10 P-Value

The 22q$_{BD\text{-}PPI}$ Network Loads Highly onto SCZ and ASD-Associated Modules

Forty-six protein-coding genes are contained in the ~2.57 Mb 22q11.2 locus; relative to all equal-sized regions of the genome, this was in the 93.58 percentile. Thirty-eight 22q11.2 genes were indexed in BrainSpan (Table 2). There were nominally more 22q11.2 genes in M14 than expected given the number of genes in each module (i.e., 1 gene; SEPT5; p=0.02; Fig. 1), but this did not survive FDR correction. Similarly, although 23 22q11.2 genes were in SCZ- and/or ASD-associated modules, this was not significantly more than expected, p>0.05. Relative to all equal-sized regions of the genome (~2.57 Mb), 22q11.2 gene content was above the 90th percentile for 6 BrainSpan modules, including 4 modules associated with SCZ or ASD (M1, M6, M7, M15; Table S4). However, the genomic-region rank percentiles of the 22q11.2 locus for SCZ and ASD-associated modules, overall, were not significantly higher than for non-disease associated modules, p>0.05.

We next sought to characterize the 22q11.2 brain development PPI network (22q$_{BD\text{-}PPI}$ network), given that perturbation of protein networks is a common characteristic of mutations associated with many diseases(39). The 22q11.2 locus formed a large BD-PPI network, consisting of 239 genes from 38 seed genes. Given the number of genes in each BrainSpan module, the 22q$_{BD\text{-}PPI}$ network was significantly over-represented in multiple SCZ- and/or ASD-associated modules (M6, M7, M15), and nominally over-represented in M13 (Fig. 1). The 22q$_{BD\text{-}PPI}$ network was also significantly over-represented in M2 and M3, and nominally over-represented in M14. However, some BrainSpan modules consist of genes that are more highly connected than other modules (i.e., have higher mean BD-PPI partners per gene, ANOVA $p<2.0\text{x}10^{-16}$; Table S5), and highly connected genes are more likely to be found in the BD-PPI network of any random set of genes. We therefore asked 2 additional questions to better understand the 22q$_{BD\text{-}PPI}$ network relative to the rest of the genome: 1) how highly does the 22q$_{BD\text{-}PPI}$ network load onto disease-associated modules relative to all other regions of the genome (i.e., normalized by genomic region size); and 2) how highly connected within protein networks are genes in the 22q11.2 locus to genes in disease-associated modules, relative to random gene-sets (i.e., normalized by number of seed genes)?

Table 2. Module membership of 38 genes in the 22q11.2 locus indexed in BrainSpan. Genes are annotated by kME score (i.e., correlation with parent module eigengene) and pLI score (i.e., loss of function intolerance). ^Module enriched for neuronal cell-type markers.

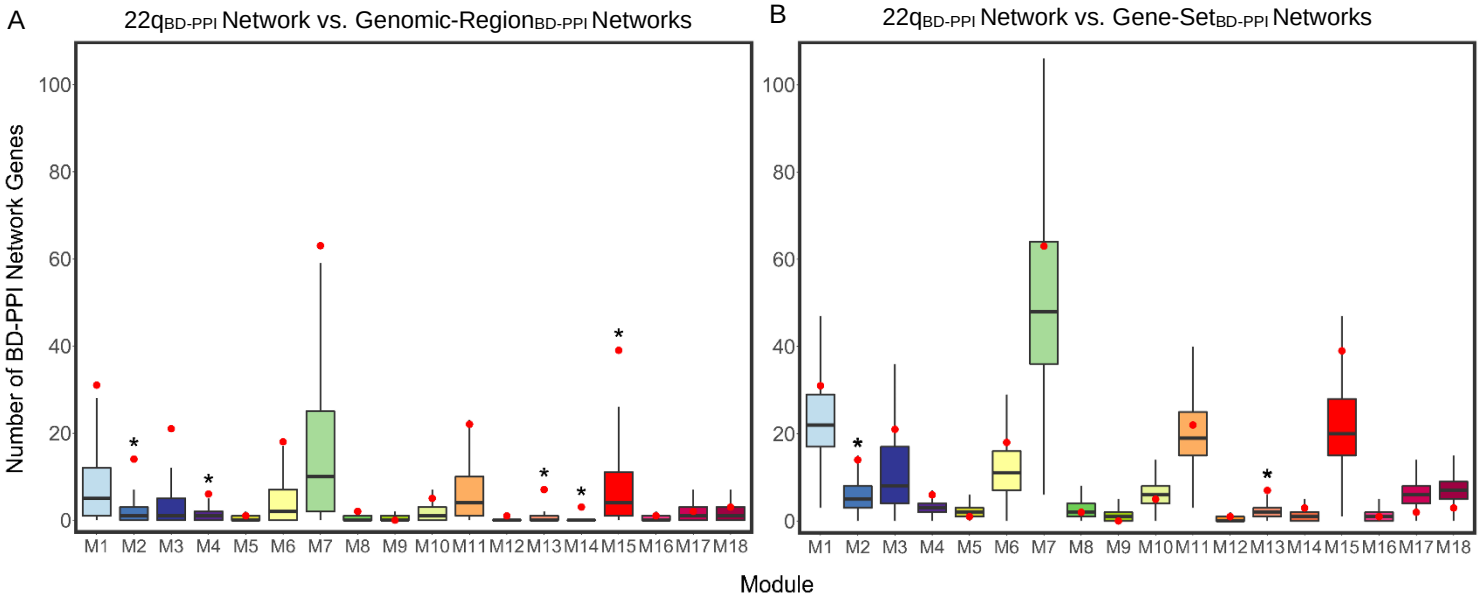
Module	Gene	Definition	kME Score	pLI Score
M1	ZNF74	zinc finger protein 74	0.843	0.0004
	DGCR8	DGCR8, microprocessor complex subunit	0.757	**0.9999**
	SLC25A1	solute carrier family 25 member 1	0.754	0.6335
	ARVCF	armadillo repeat gene deleted in velocardiofacial syndrome	0.669	0.0000
	CLTCL1	clathrin heavy chain like 1	0.607	0.0000
	KLHL22	kelch like family member 22	0.562	0.7026
	TMEM191A	transmembrane protein 191A	0.182	NA
	SCARF2	scavenger receptor class F member 2	0.009	NA
M2	CRKL	v-crk avian sarcoma virus CT10 oncogene homolog-like	0.511	0.1635
M3	CDC45	cell division cycle 45	0.775	0.0691
M6	MRPL40	mitochondrial ribosomal protein L40	0.779	0.0201
	ESS2	DiGeorge syndrome critical region gene 14	0.752	0.0000
M7	UFD1	ubiquitin recognition factor in ER associated degradation 1	0.797	**0.9974**
	SNAP29	synaptosome associated protein 29	0.749	0.1517
	HIRA	histone cell cycle regulator	0.635	**1.0000**
	RTL10	retrotransposon Gag like 10	0.547	0.0000
	TRMT2A	tRNA methyltransferase 2 homolog A	0.348	0.0000
	THAP7	THAP domain containing 7	0.208	0.0293
	DGCR2	DiGeorge syndrome critical region gene 2	0.165	0.0011
M10	AIFM3	apoptosis-inducing factor, mitochondrion-associated, 3	0.941	0.0000
	P2RX6	purinergic receptor P2X, ligand-gated ion channel, 6	0.799	0.0000
M11	PRODH	proline dehydrogenase (oxidase) 1	0.855	0.0000
	ZDHHC8	zinc finger, DHHC-type containing 8	0.755	**0.9890**
	COMT	catechol-O-methyltransferase	0.632	0.0008
	TXNRD2	thioredoxin reductase 2	0.590	0.0000
	SERPIND1	serpin peptidase inhibitor, clade D (heparin cofactor), member 1	0.416	0.0000
	C22orf39	chromosome 22 open reading frame 39	0.243	0.0000
	DGCR6L	DiGeorge syndrome critical region gene 6-like	0.172	0.1286
	GSC2	goosecoid homeobox 2	0.146	0.0372
M14^	SEPT5	septin 5	0.798	**0.9146**
M15^	PI4KA	phosphatidylinositol 4-kinase alpha	0.881	0.0008
	TANGO2	transport and golgi organization 2 homolog	0.820	0.0000
	SLC7A4	solute carrier family 7 member 4	0.722	0.0001
	RTN4R	reticulon 4 receptor	0.573	0.5264
	MED15	mediator complex subunit 15	0.381	**0.9610**
	LZTR1	leucine zipper like transcription regulator 1	0.247	0.0000
M18	CLDN5	claudin 5	0.434	0.7405
	TSSK2	testis-specific serine kinase 2	0.382	0.0042

Relative to all regions of the genome (i.e., all Genomic-Region$_{BD\text{-}PPI}$ networks), the 22q$_{BD\text{-}PPI}$ network was above the 90th percentile for 10 of 18 modules, including 6 of 7 SCZ- or ASD-associated modules (Fig. 3A, Table S4). Indeed, the percentile ranks of the 22q$_{BD\text{-}PPI}$ network were significantly higher for SCZ- and/or ASD-associated modules than for non-SCZ- or ASD-associated modules, p=0.03, highlighting that relative to the rest of the genome, the 22q$_{BD\text{-}PPI}$ network loads disproportionately onto SCZ- and ASD-associated modules.

Relative to the connectivity of random sets of 38 seed genes (i.e., the Gene-Set$_{BD\text{-}PPI}$ networks), the 22q$_{BD\text{-}PPI}$ network was above the 75th percentile for 5 SCZ- and/or ASD-associated modules, including remaining in the ~96th and ~93rd percentiles of the SCZ- and ASD-associated M13 and M15 neuronal modules, respectively (Figure 3B; Table S4). The percentile ranks of the 22q$_{BD\text{-}PPI}$ network relative to the Gene-Set$_{BD\text{-}PPI}$ networks were similarly higher for SCZ- and/or ASD-associated modules than for non-SCZ- or ASD-associated modules, p=0.03.

Together, this suggests that 22q11.2 CNVs confer risk for SCZ and ASD by not only spanning a genomic region that is rich in protein-coding genes, but by spanning genes that form protein networks in the developing brain that are highly - and disproportionately - connected to broader networks associated with SCZ and ASD pathogenesis.

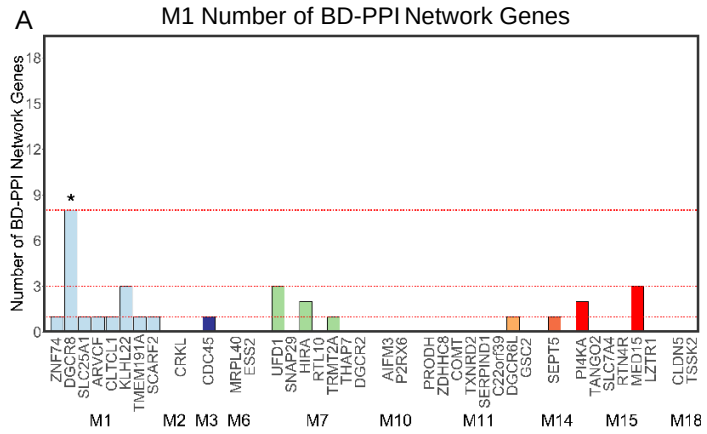
Figure 3. Distribution of number of BD-PPI network genes per module across all (A) genome-wide ~2.57 Mb genomic regions (i.e., 12,149 Genomic-Region$_{BD\text{-}PPI}$ networks) and (B) random 38 seed gene gene-sets (i.e., 10,000 Gene-Set$_{BD\text{-}PPI}$ networks) derived to characterize the 22q11.2 region relative to all other regions of the genome, and the connectivity of 22q11.2 genes within brain development protein networks relative to all BrainSpan genes, respectively. Boxplots mark the median genes per module, with lower and upper hinges corresponding to the 25th and 75th quartiles, respectively, and the whiskers extending to 1.5x interquartile range. The number of 22q$_{BD\text{-}PPI}$ network genes for each module and gene signature are marked in red. *22q$_{BD\text{-}PPI}$ network above the 95th percentile.

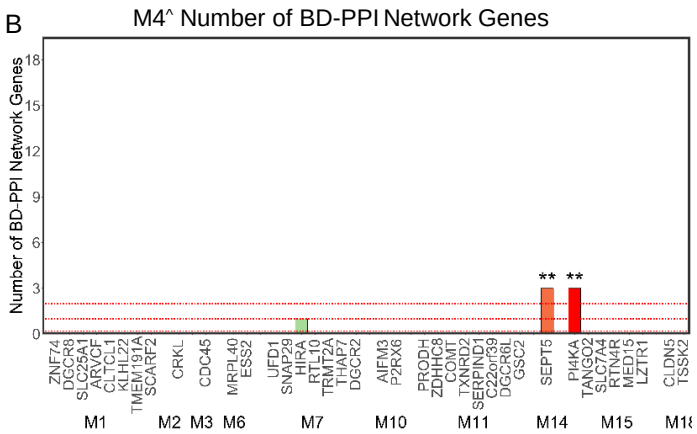


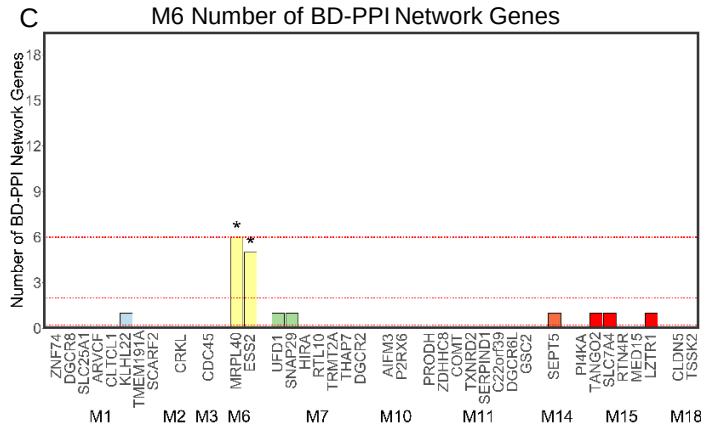
SEPT5, *PI4KA*, and *SNAP29* are Candidate Drivers of Synaptic Pathology, and *DGCR8* and *HIRA* are Candidate Drivers of Dysregulated Gene Expression

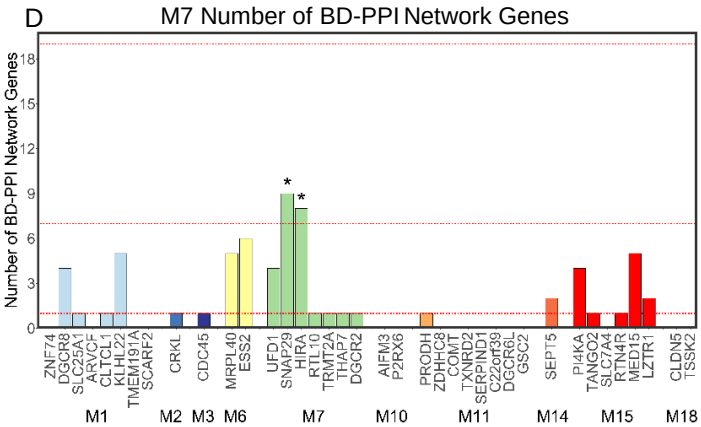
To identify specific genes driving the loading of the $22q_{BD\text{-}PPI}$ network onto SCZ- and ASD-associated modules, we next examined the size of each individual 22q11.2 gene's BD-PPI network for each module, relative to all BrainSpan-indexed genes. *SEPT5* and *PI4KA* were above the 99th percentile for BD-PPI network genes in 3 neuronal modules associated with SCZ and ASD (M4, M13, M15; Fig. 4); *PI4KA* was additionally above the 95th percentile for the ASD-associated module, M12. *SNAP29* was also above the 95th percentile for network genes in M13 and M15, as well as M7. Indeed, the individual *SEPT5*, *PI4KA*, and *SNAP29* networks were each enriched for cell parts and/or biological processes related to synaptic function (Fig. 5A-F). The $HIRA_{BD\text{-}PPI}$ network was additionally above the 95th percentile for genes in M7 and was enriched for genes involved in chromatin organization and gene expression (Fig. 5G,H), consistent with HIRA's role as a histone chaperone protein. Additionally, the $DGCR8_{BD\text{-}PPI}$ network was near the 99th percentile of BrainSpan genes for M1 and was enriched for RNA binding and processing, consistent with *DGCR8*'s known role in microRNA (miRNA) biogenesis (Fig. 5I,J). Finally, the *MRPL40* and *ESS2* BD-PPI networks were above the 95th percentile for M6; the *ESS2* network was not enriched for any GO terms, while *MRPL40*'s network was enriched for ribosomal components and mitochondrial elongation (Fig. 5K-M).

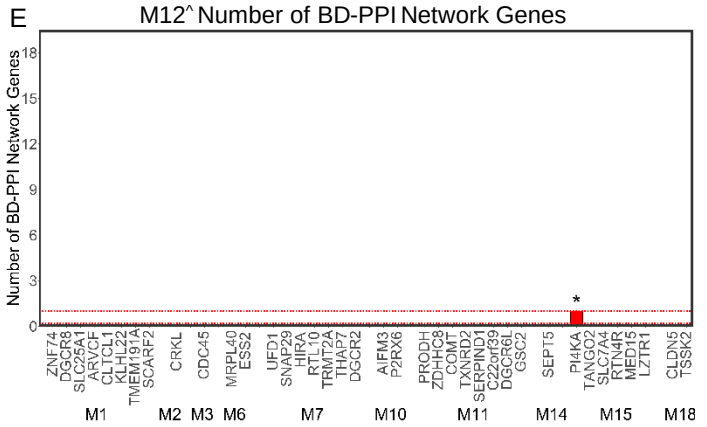
Figure 4. Number of BD-PPI network genes per individual 22q11.2 gene in the (A) M1, (B) M4, (C) M6, (D) M7, (E) M12, (F) M13, and (G) M15 SCZ and/or ASD-associated modules. 22q11.2 genes are shown grouped by module membership. The top, second, and third red dotted lines in each plot mark the number of BD-PPI network genes nearest to the 99th, 95th, and 75th percentiles for each module. Genes above the 95th or 99th percentile for each module are denoted by * or **, respectively. ^Module enriched for neuronal cell-type markers.

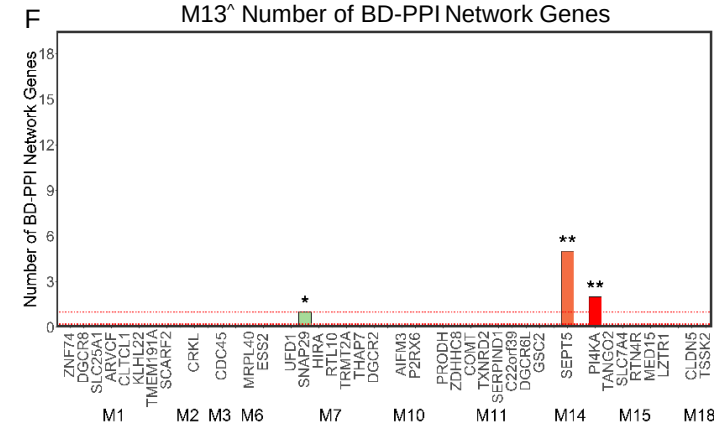

A
M1 Number of BD-PPI Network Genes
Number of BD-PPI Network Genes
0
3
6
9
12
15
18
*
ZNF74
DGCR8
SLC25A1
ARVCF
CLTCL1
KLHL22
TMEM191A
SCARF2
CRKL
CDC45
MRPL40
ESS2
UFD1
SNAP29
HIRA
RTL10
TRMT2A
THAP7
DGCR2
AIFM3
P2RX6
PRODH
ZDHHC8
COMT
TXNRD2
SERPIND1
C22orf39
DGCR6L
GSC2
SEPT5
PI4KA
TANGO2
SLC7A4
RTN4R
MED15
LZTR1
CLDN5
TSSK2
M1
M2
M3
M6
M7
M10
M11
M14
M15
M18


B
M4^ Number of BD-PPI Network Genes
Number of BD-PPI Network Genes
0
3
6
9
12
15
18
**
**
ZNF74
DGCR8
SLC25A1
ARVCF
CLTCL1
KLHL22
TMEM191A
SCARF2
CRKL
CDC45
MRPL40
ESS2
UFD1
SNAP29
HIRA
RTL10
TRMT2A
THAP7
DGCR2
AIFM3
P2RX6
PRODH
ZDHHC8
COMT
TXNRD2
SERPIND1
C22orf39
DGCR6L
GSC2
SEPT5
PI4KA
TANGO2
SLC7A4
RTN4R
MED15
LZTR1
CLDN5
TSSK2
M1
M2
M3
M6
M7
M10
M11
M14
M15
M18


C
M6 Number of BD-PPI Network Genes
Number of BD-PPI Network Genes
0
3
6
9
12
15
18
*
*
ZNF74
DGCR8
SLC25A1
ARVCF
CLTCL1
KLHL22
TMEM191A
SCARF2
CRKL
CDC45
MRPL40
ESS2
UFD1
SNAP29
HIRA
RTL10
TRMT2A
THAP7
DGCR2
AIFM3
P2RX6
PRODH
ZDHHC8
COMT
TXNRD2
SERPIND1
C22orf39
DGCR6L
GSC2
SEPT5
PI4KA
TANGO2
SLC7A4
RTN4R
MED15
LZTR1
CLDN5
TSSK2
M1
M2
M3
M6
M7
M10
M11
M14
M15
M18


D
M7 Number of BD-PPI Network Genes
Number of BD-PPI Network Genes
0
3
6
9
12
15
18
*
*
ZNF74
DGCR8
SLC25A1
ARVCF
CLTCL1
KLHL22
TMEM191A
SCARF2
CRKL
CDC45
MRPL40
ESS2
UFD1
SNAP29
HIRA
RTL10
TRMT2A
THAP7
DGCR2
AIFM3
P2RX6
PRODH
ZDHHC8
COMT
TXNRD2
SERPIND1
C22orf39
DGCR6L
GSC2
SEPT5
PI4KA
TANGO2
SLC7A4
RTN4R
MED15
LZTR1
CLDN5
TSSK2
M1
M2
M3
M6
M7
M10
M11
M14
M15
M18


E
M12^ Number of BD-PPI Network Genes
Number of BD-PPI Network Genes
0
3
6
9
12
15
18
*
ZNF74
DGCR8
SLC25A1
ARVCF
CLTCL1
KLHL22
TMEM191A
SCARF2
CRKL
CDC45
MRPL40
ESS2
UFD1
SNAP29
HIRA
RTL10
TRMT2A
THAP7
DGCR2
AIFM3
P2RX6
PRODH
ZDHHC8
COMT
TXNRD2
SERPIND1
C22orf39
DGCR6L
GSC2
SEPT5
PI4KA
TANGO2
SLC7A4
RTN4R
MED15
LZTR1
CLDN5
TSSK2
M1
M2
M3
M6
M7
M10
M11
M14
M15
M18


F
M13^ Number of BD-PPI Network Genes
Number of BD-PPI Network Genes
0
3
6
9
12
15
18
*
**
**
ZNF74
DGCR8
SLC25A1
ARVCF
CLTCL1
KLHL22
TMEM191A
SCARF2
CRKL
CDC45
MRPL40
ESS2
UFD1
SNAP29
HIRA
RTL10
TRMT2A
THAP7
DGCR2
AIFM3
P2RX6
PRODH
ZDHHC8
COMT
TXNRD2
SERPIND1
C22orf39
DGCR6L
GSC2
SEPT5
PI4KA
TANGO2
SLC7A4
RTN4R
MED15
LZTR1
CLDN5
TSSK2
M1
M2
M3
M6
M7
M10
M11
M14
M15
M18

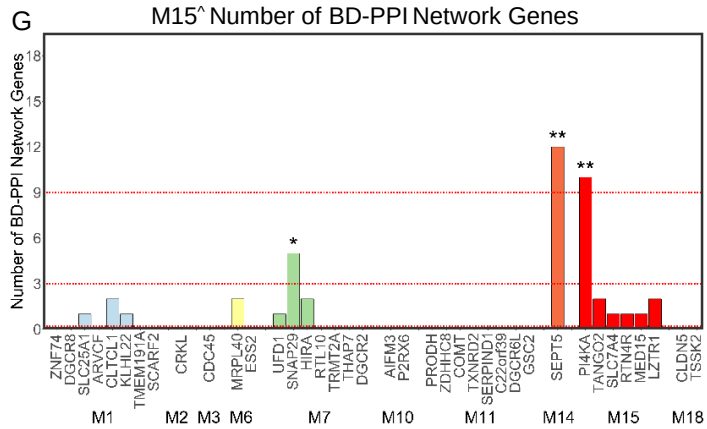
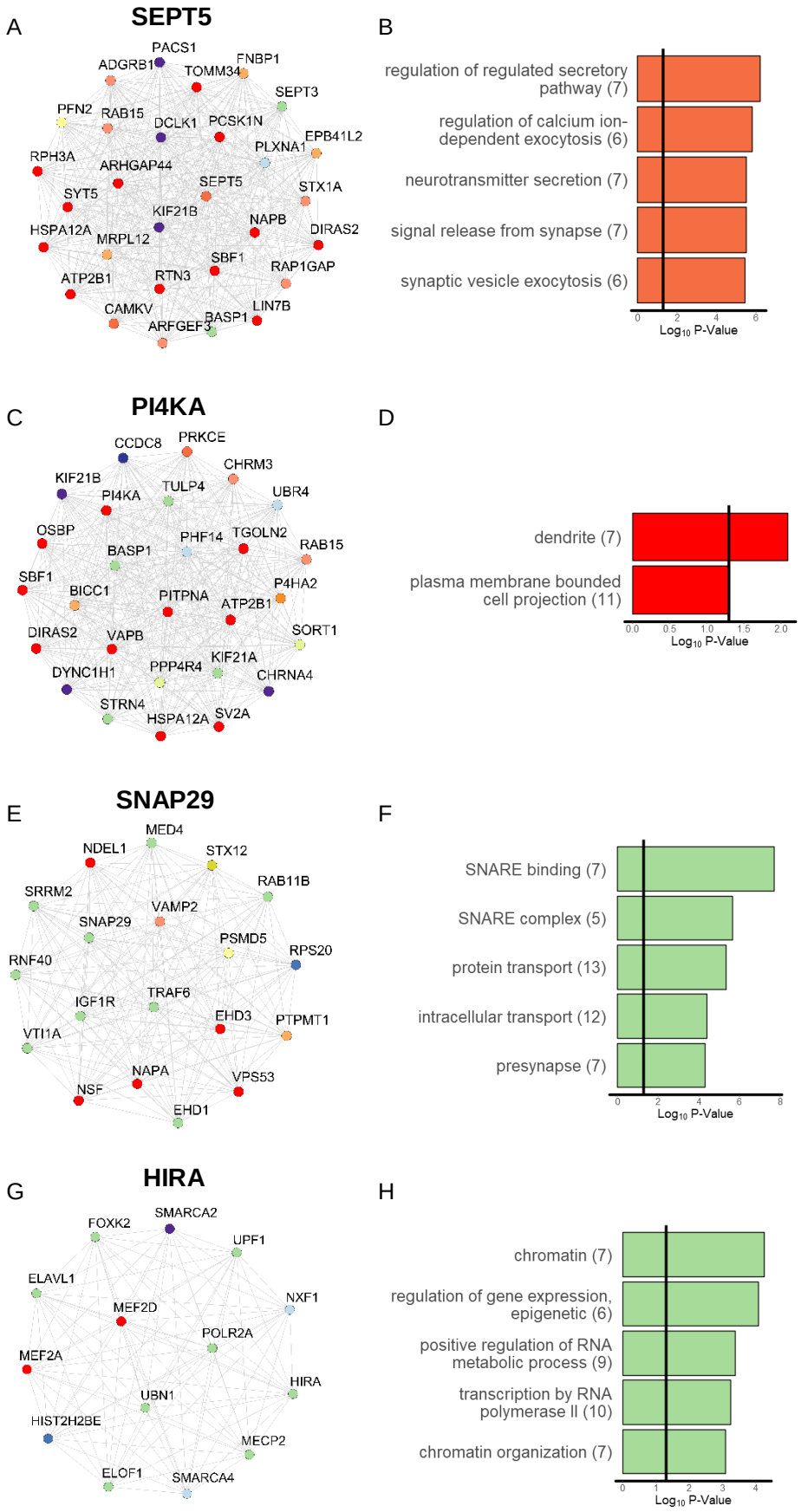
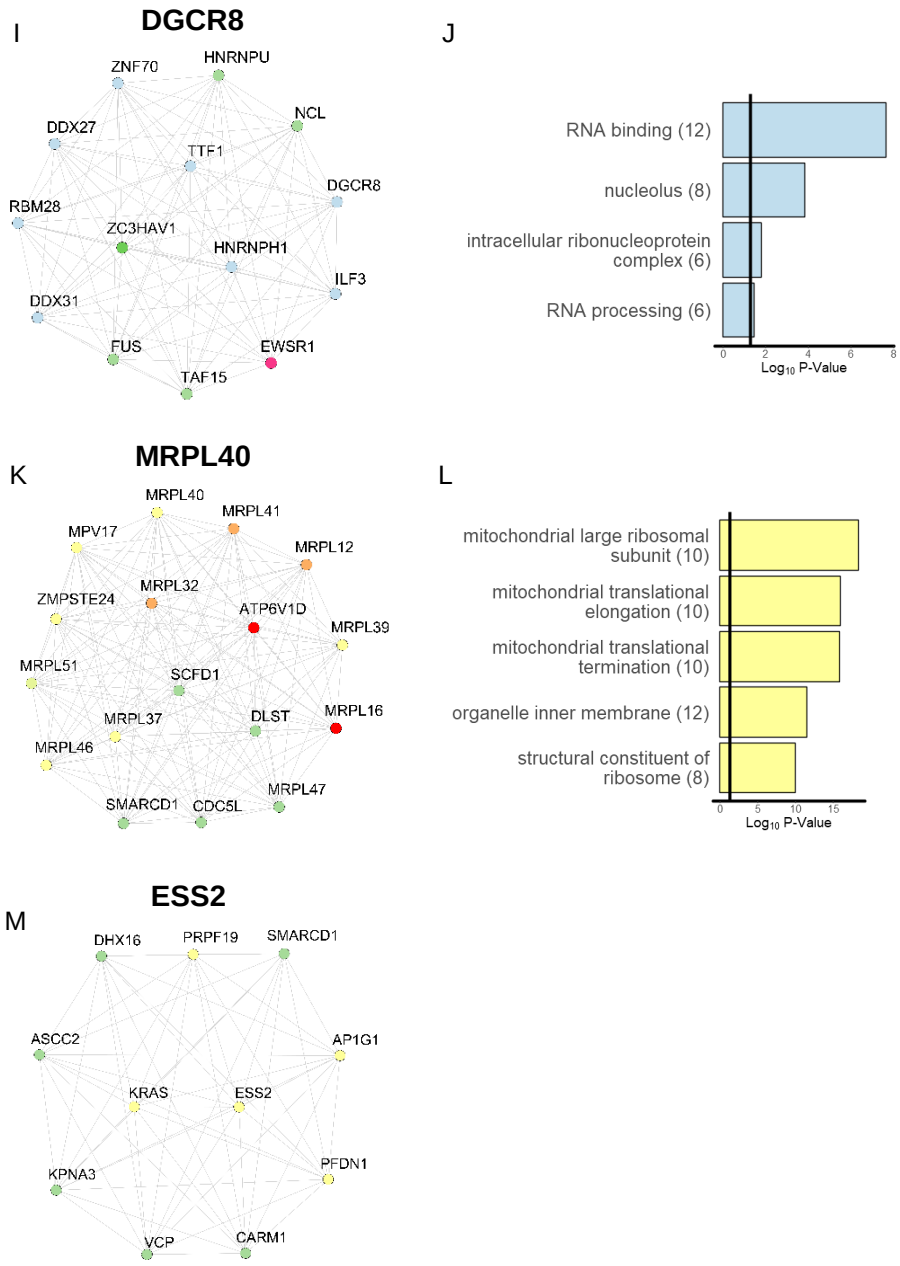

G
M15^ Number of BD-PPI Network Genes
Number of BD-PPI Network Genes
0
3
6
9
12
15
18
*
**
**
ZNF74
DGCR8
SLC25A1
ARVCF
CLTCL1
KLHL22
TMEM191A
SCARF2
CRKL
CDC45
MRPL40
ESS2
UFD1
SNAP29
HIRA
RTL10
TRMT2A
THAP7
DGCR2
AIFM3
P2RX6
PRODH
ZDHHC8
COMT
TXNRD2
SERPIND1
C22orf39
DGCR6L
GSC2
SEPT5
PI4KA
TANGO2
SLC7A4
RTN4R
MED15
LZTR1
CLDN5
TSSK2
M1
M2
M3
M6
M7
M10
M11
M14
M15
M18

Figure 5. Left: BD-PPI network genes and Right: significantly enriched GO terms (up to top 5 terms, where applicable) for: (A,B) SEPT5, (C,D) PI4KA, (E,F) SNAP29, (G,H) HIRA, (I,J) DGCR8, (K,L) MRPL40, and (M) ESS2. Gene nodes in each network are colored by primary BrainSpan module assignment.





The 22q$_{BD\text{-}PPI}$ Network Loads Highly onto SCZ and ASD Pathophysiologic Signatures

Post-mortem brain tissue from SCZ and ASD patients is usually ascertained from adults, limiting the detectability of differential gene expression (DGE) profiles that may exist during early brain development. Nevertheless, DGE profiles provide important pathophysiological signatures of disease. All modules enriched for genetic risk for SCZ were significantly enriched for down-regulated genes in SCZ cortex (M7,M13,M15; Fig. 6A). Similarly, 6 of 7 modules enriched for genetic risk for ASD were nominally or significantly enriched for down-regulated genes in ASD cortex (M4,M6,M7,M12,M13,M15). M1 was consistently enriched for ASD risk variants, but not for down-regulated genes in ASD cortex. However, M1 was highly enriched for genes expressed specifically during fetal development, suggesting that this likely reflects a floor effect in postnatal expression of these genes. Down-regulated genes in ASD and SCZ cortex

therefore appear to at least partially reflect direct effects of genetic risk, and we focused on $22q_{BD\text{-}PPI}$ network overlap with down-regulated genes for subsequent analyses.

Given the size of the $22q_{BD\text{-}PPI}$ network and each DGE gene-set, the $22q_{BD\text{-}PPI}$ network was significantly enriched for down-regulated genes in both SCZ (~11% of the $22q_{BD\text{-}PPI}$ network; p=0.006) and ASD (~14% of the $22q_{BD\text{-}PPI}$ network; p=2.0x10^{-4}). Additionally, compared to all regions of the genome (i.e., Genomic-$Region_{BD\text{-}PPI}$ networks), the $22q_{BD\text{-}PPI}$ network was in the ~93rd and ~96th percentiles for overlap with down-regulated genes in SCZ and ASD, respectively (Fig. 6B; Table S4). Compared to the connectivity of random sets of 38 seed genes (i.e., Gene-$Set_{BD\text{-}PPI}$ networks), the $22q_{BD\text{-}PPI}$ network was in the ~77th and ~92nd percentiles for down-regulated SCZ and ASD genes, respectively (Fig. 6C; Table S4). Notably, $22q_{BD\text{-}PPI}$ network genes overlapping down-regulated genes in SCZ and ASD were both enriched for neuronal and synaptic gene-sets (Fig. 6D,E), indicating that the $22q_{BD\text{-}PPI}$ network overlaps key pathophysiological signatures in SCZ and ASD via neuronal and synaptic features.

***SEPT5* and *PI4KA* are Candidate Drivers of Synaptic Pathology Shared in SCZ and ASD**

Among the 22q11.2 genes, *SEPT5* was above the 99th percentile and *PI4KA* was near the 99th percentile of BrainSpan genes for BD-PPI network overlap with down-regulated genes in SCZ and ASD (Fig. 7). Indeed, *SEPT5* accounted for ~50% and ~26% of the overall $22q_{BD\text{-}PPI}$ network overlapping down-regulated genes in SCZ and ASD cortex, respectively, and *PI4KA* accounted for ~37% and ~18% of the $22q_{BD\text{-}PPI}$ network overlap with down-regulated genes in SCZ and ASD, respectively. *SNAP29* and *KLHL22* were also above the 95th percentile for BD-PPI network overlap with down-regulated genes in ASD, and accounted for ~15% and ~12% of $22q_{BD\text{-}PPI}$ network overlap with down-regulated genes in ASD, respectively. Thus, *SEPT5* and *PI4KA* were further highlighted as potential key drivers of synaptic pathology in 22q11.2 CNVs relevant to SCZ and ASD risk.

Figure 6. (A) BrainSpan module enrichment for significantly up- and down-regulated genes in SCZ (red) and ASD (blue) post-mortem cortex. Log p values are indicated by color for all enrichment odds ratios >1.0. Enrichment values for p<0.05 are annotated ($^{+}$p<0.05, and *FDR corrected p<0.05 for number of modules), with numerical values denoting odds ratios for enriched modules. ^Module enriched for neuronal cell-type markers. Distribution of number of genes overlapping down-regulated genes in SCZ and ASD cortex for all: A) genome-wide ~2.57 Mb genomic regions (i.e., Genomic-$Region_{BD\text{-}PPI}$ networks) and (B) random 38 seed gene gene-sets (i.e., Gene-$Set_{BD\text{-}PPI}$ networks) derived to characterize the 22q11.2 region relative to all other regions of the genome, and the connectivity of 22q11.2 genes relative to all BrainSpan genes, respectively. Boxplots mark the median genes per module, with lower and upper hinges corresponding to the 25th and 75th quartiles, respectively, and the whiskers extending to 1.5x interquartile range. The number of $22q_{BD\text{-}PPI}$ network

genes for each gene signature are marked in red. *22q$_{BD\text{-}PPI}$ network above the 95th percentile. Significantly enriched gene ontology (GO) terms for 22q$_{BD\text{-}PPI}$ network genes overlapping down-regulated genes in (D) SCZ and (E) ASD cortex.

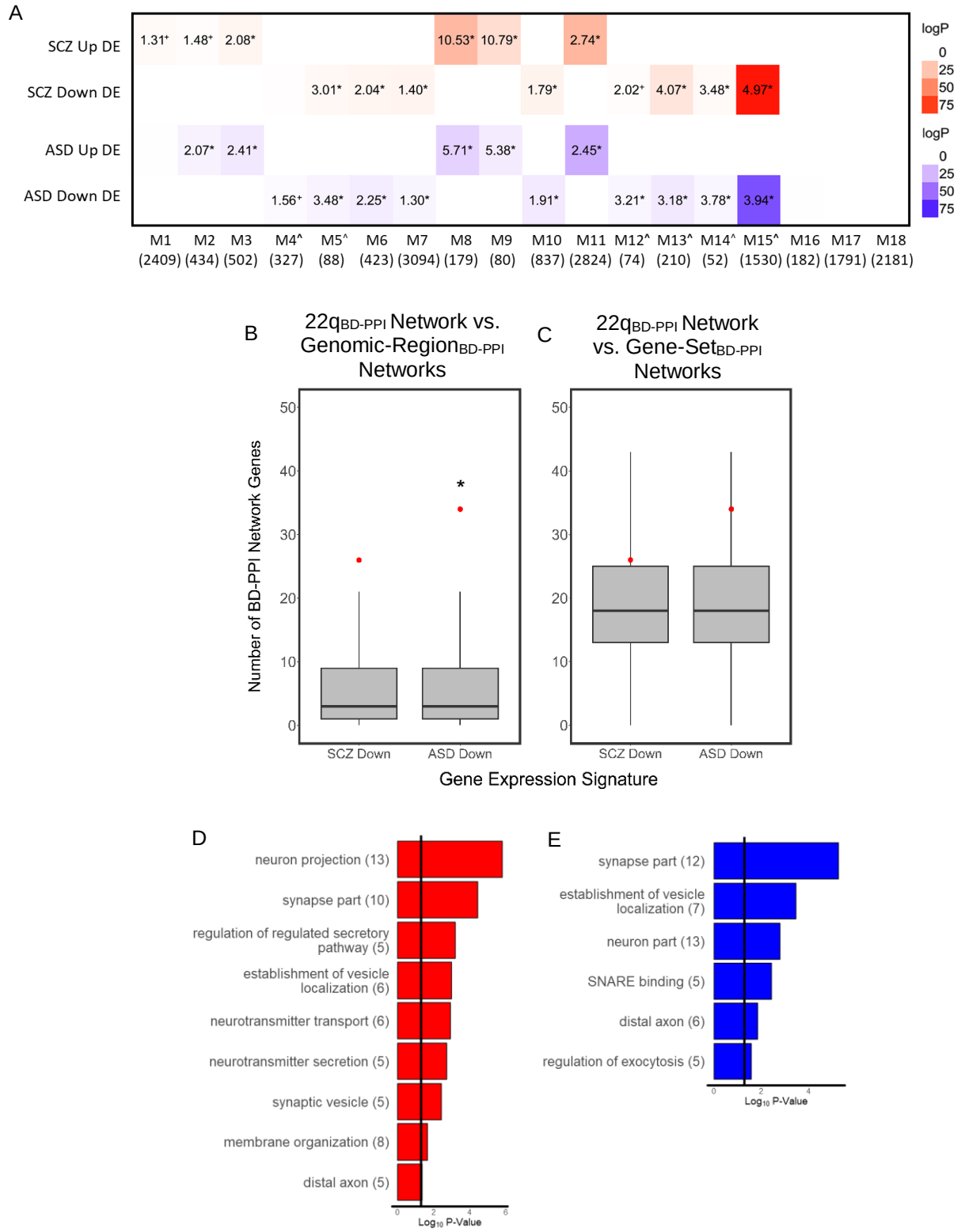
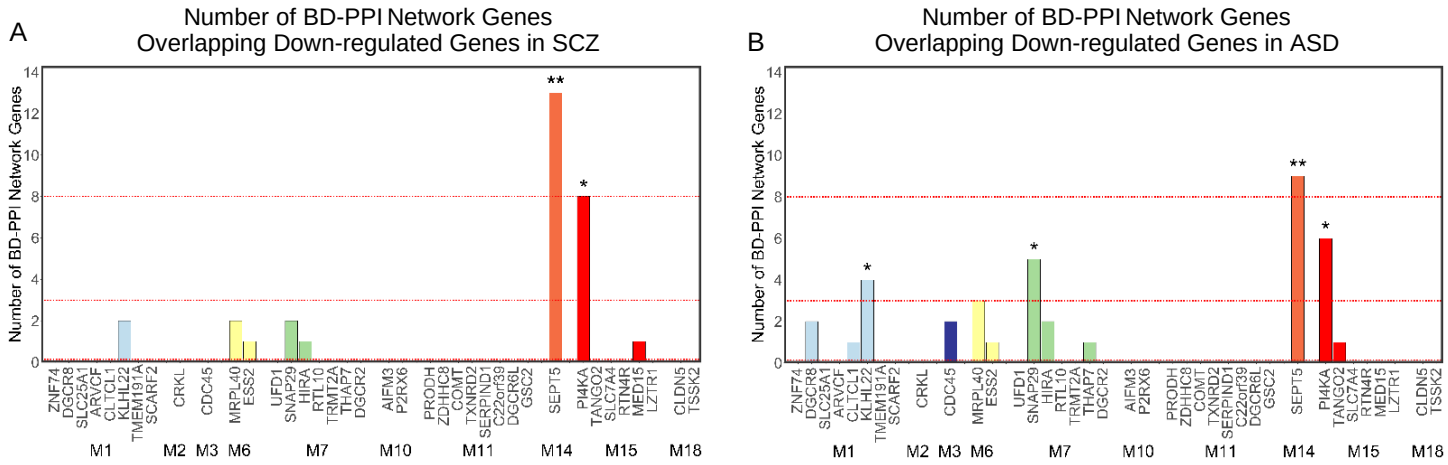


Figure 7. Number of BD-PPI network genes per individual 22q11.2 gene overlapping down-regulated genes in (A) SCZ and (B) ASD cortex. 22q11.2 genes are shown grouped by module membership. The top, second, and third red dotted lines on each plot mark the number of genes nearest to the 99th, 95th, and 75th percentiles for all BrainSpan genes' BD-PPI networks, when above 0. Genes above the 95th or 99th percentile for each module are denoted by * or **, respectively.



Exploratory Analyses Supporting the Current Framework for Understanding 22q11.2 CNV Risk for SCZ and ASD

Exploratory analyses of broader CNVs and genes associated with SCZ and ASD confirmed that high BD-PPI network loading onto SCZ- and ASD-associated BrainSpan modules and down-regulated genes in SCZ or ASD cortex is a common characteristic of disease-associated CNVs and genes (see Supplementary Information). Similarly, using lower co-expression thresholds (i.e., more inclusive thresholds; Spearman's $\rho>0.50$ and $\rho>0.60$) to define the BD-PPI networks of each gene highlighted similar 22q11.2 genes within the 22q11.2 locus as likely drivers of SCZ and ASD risk. Further, a machine learning approach to classify SCZ and ASD associated genes based on BD-PPI network rankings across BrainSpan modules and DGE signatures prioritized similar 22q11.2 genes as likely disease-associated. Finally, study bias could not account for our prioritization of 22q11.2 genes likely to drive risk for SCZ and ASD. See Supplementary Information for details.

Discussion

Recent genetic discoveries in SCZ and ASD have brought renewed focus to understanding the mechanisms through which genetic variation confers risk for these disorders at both the polygenic and individual variant level. Here, using a developmental functional genomics approach, we identified overlapping and distinct neurodevelopmental processes associated with risk for SCZ and ASD, and developed a framework leveraging this disease-informed lens to elucidate mechanisms underlying disease risk in highly penetrant multi-gene CNVs. Our analyses confirmed the capacity

of 22q11.2 CNVs to disrupt disease-relevant neurodevelopmental networks and highlighted both well-studied (e.g., *DGCR8* and *HIRA*) and relatively understudied genes (e.g., *SEPT5*, *PI4KA*, and *SNAP29*) as candidate 22q11.2 genes likely to drive risk for SCZ and/or ASD.

Prior studies using transcriptomic data from the developing brain to understand how risk variants contribute to disease largely focused on rare variants in ASD. These studies suggested that ASD risk variants preferentially disrupt neuronal development and transcriptional and chromatin regulation during fetal development(15,16,34). One early study of DNMs in SCZ also implicated disruptions in neurogenesis and transcriptional regulation(25); however, this was a relatively small study. Here, by creating co-expression networks from brain samples spanning the full period of development and testing an expanded and updated set of risk variants for SCZ and ASD simultaneously, we confirmed prior enrichment patterns in ASD and identified co-expression modules with distinct enrichment patterns for ASD versus SCZ. We found robust enrichment in SCZ for a neuronal module that increased in expression across development and was involved broadly in postnatal synaptic transmission and refinement of neurotransmission; a neuronal module involved in regulating neuronal excitability that stabilized in expression around birth; and a module involved in regulating transcription and splicing across development. ASD risk variants showed overlapping enrichment for these modules, as well as robust enrichment for distinct modules involved in neuronal differentiation, and in regulating chromatin organization and gene transcription during fetal development. While our findings thus confirm prior findings implicating neuronal and synaptic genes in both SCZ and ASD, risk variants for ASD also appear to partially diverge from those for SCZ by loading highly onto gene networks that orchestrate neuronal differentiation and synapse formation during fetal development. It is tempting to suggest that the higher loading of ASD risk variants onto early neuronal differentiation networks may be a key mechanism underlying the earlier onset of ASD relative to SCZ; however, experimental validation is necessary to confirm this.

By integrating BrainSpan co-expression and PPI data, we identified the protein network likely to be strongly affected by 22q11.2 CNVs during brain development. One prior study focused on 16p11.2 CNVs integrated PPI and BrainSpan co-expression data and found that 'neuropsychiatric' CNV networks were enriched for varying developmental periods, with the 22q11.2 network enriched during fetal and postnatal development(40). Here, using a disease-informed framework, we focused on understanding how 22q11.2 CNVs confer risk for SCZ and ASD, more specifically. We found that 22q11.2 genes formed a large BD-PPI network that: 1) loaded highly onto multiple SCZ- and ASD-associated modules, relative to all other regions of the genome; 2) was more highly connected to genes in SCZ- and/or ASD-associated modules than to genes in non-disease associated modules; and 3) overlapped highly with genes that are down-regulated in SCZ and ASD cortex, with overlapping genes enriched for neuronal and synaptic genes. This suggests that 22q11.2 CNVs may confer risk for SCZ and ASD, at least partially, by perturbing synaptic networks that modulate

neuronal excitability and signaling. Consistent with this, mouse models of 22q11DS show altered neuronal excitability, dendritic spine stability, and synaptic activity(52,53).

A primary advantage of our approach was our ability to systematically survey 22q11.2 genes to prioritize genes likely to contribute to disease risk. Our analyses consistently highlighted *SEPT5* and *PI4KA* as candidate drivers of synaptic pathology relevant to SCZ and ASD, and *DGCR8* as a candidate driver of dysregulated gene expression that may be most relevant to ASD. *SNAP29* and *HIRA* were also prioritized as contributing to SCZ and ASD. Of note, both *SEPT5* and *SNAP29* have been found to negatively modulate neurotransmitter release. *SEPT5* encodes the filamentous protein septin 5, which is thought to serve as a molecular "brake" for neurotransmitter release by associating with synaptic vesicles to create a barrier that inhibits vesicle fusion with the membrane(55,56). Similarly, *SNAP29* encodes a SNARE protein that competes with α-SNAP for binding to SNARE complexes after membrane fusion, thus reducing SNARE protein recycling and synaptic vesicle availability(57). Conversely, altered *PI4KA* expression may contribute to SCZ and ASD risk by perturbing the proteomic composition of the plasma membrane (PM). *PI4KA* encodes the lipid kinase PI4KIIIα, which is involved in synthesizing phosphatidylinositol 4,5-biphosphate (PI(4,5)P_2) that recruits and regulates PM proteins and thereby governs diverse cellular processes including vesicular trafficking, ion transport, and actin cytoskeletal dynamics(58). *PI4KA* depletion was found to disrupt localization of PI(4,5)P_2-associated proteins, reduce cellular responses to external stimuli(59, 60), and impair synapse maintenance(61). Our prioritization of *PI4KA* in contributing to SCZ and ASD risk is consistent with prior associations of *PI4KA* polymorphisms in SCZ(62-64) and with growing evidence that *PI4KA* may play an important role in synaptic function.

DGCR8 is a core component of the microprocessor complex involved in processing miRNAs, which regulate gene expression at the protein level by binding target mRNAs to silence their translation(65). *HIRA* is a histone chaperone that influences gene expression by facilitating the deposition of the non-canonical histone variant, H3.3, into chromatin(66,67). Given these gene regulatory functions, *DGCR8* and *HIRA* have both been previously hypothesized to play an important role in 22q11.2 CNV phenotypes(67,68). Here, using our disease-informed functional genomic lens, we provide independent, unbiased evidence for the contribution of *DGCR8* and *HIRA* haploinsufficiency to risk for neurodevelopmental disorders.

Some caveats to the current study should be noted. First, in our analyses testing SCZ and ASD risk variants for module enrichment, some studies shared overlapping samples. As the genetic architectures of SCZ and ASD include many variants with small effects, collaborative analyses across studies are increasingly common to facilitate power. We therefore utilized independent datasets where possible, and focused on modules showing consistent disease association. Additionally, generating BD-PPI networks identifies protein networks putatively affected by 22q11.2 CNVs, but cannot tell us the functional output of increasing or reducing the expression of each gene. We therefore cannot determine how

deletions vs. duplications of these genes affect broader cascades within these networks. Similarly, while using each gene's BD-PPI network ranking is useful for prioritizing the most highly connected 22q11.2 genes within disease-associated modules, for genes whose mechanism of action involves regulating the expression of other genes, this cannot quantify the broader consequences of mutations in these genes. As high-throughput functional assays become increasingly feasible in the future, quantifying the transcriptomic and proteomic consequences of altered expression of each 22q11.2 gene will likely yield further insights into 22q11.2 CNV-mediated disease risk. Nevertheless, identifying convergent neurodevelopmental networks affected by SCZ and ASD genetic risk, and systematically prioritizing 22q11.2 genes that load highly onto these networks offers a powerful step towards understand the specific mechanisms through which 22q11.2 CNVs confer risk for SCZ and ASD.

Acknowledgements. The authors thank Gerhard Helleman for statistical advice and Elizabeth Ruzzo for feedback and advice on analyses. This work was supported by grants from the National Institutes of Mental Health (R01MH085953 to CEB, R01MH100900 to CEB, T32MH096682 to JKF, K08MH11877 to JKF).

Conflicts of Interest. The authors have no conflicts of interest to disclose.

References

1. Guna A, Butcher NJ, Bassett AS (2015): Comparative mapping of the 22q11.2 deletion region and the potential of simple model organisms. *J Neurodev Disord*. 7: 18.

2. Van L, Boot E, Bassett AS (2017): Update on the 22q11.2 deletion syndrome and its relevance to schizophrenia. *Curr Opin Psychiatry*. 30: 191–196.

3. Rees E, Kirov G, Sanders A, Walters JTR, Chambert KD, Shi J, *et al.* (2014): Evidence that duplications of 22q11.2 protect against schizophrenia. *Mol Psychiatry*. 19: 37–40.

4. Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, *et al.* (2017): Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet*. 49: 27–35.

5. Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, *et al.* (2011): Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*. 70: 863–885.

6. Gandal MJ, Leppa V, Won H, Parikshak NN, Geschwind DH (2016): The road to precision psychiatry: translating genetics into disease mechanisms. *Nat Neurosci*. 19: 1397–1407.

7. Pocklington AJ, Rees E, Walters JTR, Han J, Kavanagh DH, Chambert KD, *et al.* (2015): Novel Findings from CNVs Implicate Inhibitory and Excitatory Signaling Complexes in Schizophrenia. *Neuron*. 86: 1203–1214.

8. Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, *et al.* (2018): Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet*. . doi: 10.1038/s41588-018-0059-2.

9. Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, *et al.* (2014): De novo mutations in schizophrenia implicate synaptic networks. *Nature*. 506: 179–184.

10. Genovese G, Fromer M, Stahl EA, Ruderfer DM, Chambert K, Landén M, *et al.* (2016): Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. *Nat Neurosci*. 19: 1433–1441.

11. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014): Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 511: 421–427.

12. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, *et al.* (2014): The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. 515: 216–221.

13. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, *et al.* (2014): Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*. 515: 209–215.

14. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, *et al.* (2015): Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron*. 87: 1215–1233.

15. Parikshak NN, Luo R, Zhang A, Won H, Lowe JK, Chandran V, *et al.* (2013): Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*. 155: 1008–1021.

16. Willsey AJ, Sanders SJ, Li M, Dong S, Tebbenkamp AT, Muhle RA, *et al.* (2013): Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell*. 155: 997–1007.

17. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, *et al.* (2011): Spatio-temporal transcriptome of the human brain. *Nature*. 478: 483–489.

18. Geschwind DH, Konopka G (2009): Neuroscience in the era of functional genomics and systems biology. *Nature*. 461: 908–915.

19. Seyfried NT, Dammer EB, Swarup V, Nandakumar D, Duong DM, Yin L, *et al.* (2017): A Multi-network Approach Identifies Protein-Specific Co-expression in Asymptomatic and Symptomatic Alzheimer's Disease. *Cell Syst*. 4: 60–72.e4.

20. Reimand J, Kull M, Peterson H, Hansen J, Vilo J (2007): g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res*. 35: W193–200.

21. Dougherty JD, Schmidt EF, Nakajima M, Heintz N (2010): Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. *Nucleic Acids Res*. 38: 4218–4230.

22. Grove J, Ripke S, Als TD, Mattheisen M, Walters R, Won H, *et al.* (2017, November 25): Common risk variants identified in autism spectrum disorder. *bioRxiv*. .

23. Richards AL, Leonenko G, Walters JT, Kavanagh DH, Rees EG, Evans A, *et al.* (2016): Exome arrays capture polygenic rare variant contributions to schizophrenia. *Hum Mol Genet*. 25: 1001–1007.

24. de Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015): MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol*. 11: e1004219.

25. Gulsuner S, Walsh T, Watts AC, Lee MK, Thornton AM, Casadei S, *et al.* (2013): Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell*. 154: 518–529.

26. Ambalavanan A, Girard SL, Ahn K, Zhou S, Dionne-Laporte A, Spiegelman D, *et al.* (2016): De novo variants in sporadic cases of childhood onset schizophrenia. *Eur J Hum Genet*. 24: 944–948.

27. Guipponi M, Santoni FA, Setola V, Gehrig C, Rotharmel M, Cuenca M, *et al.* (2014): Exome sequencing in 53 sporadic cases of schizophrenia identifies 18 putative candidate genes. *PLoS One*. 9: e112745.

28. Xu B, Ionita-Laza I, Roos JL, Boone B, Woodrick S, Sun Y, *et al.* (2012): De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet*. 44: 1365–1369.

29. Wang Q, Li M, Yang Z, Hu X, Wu H-M, Ni P, *et al.* (2015): Increased co-expression of genes harboring the damaging de novo mutations in Chinese schizophrenic patients during prenatal development. *Sci Rep*. 5: 18209.

30. McCarthy SE, Gillis J, Kramer M, Lihm J, Yoon S, Berstein Y, *et al.* (2014): De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol Psychiatry*. 19: 652–658.

31. Girard SL, Gauthier J, Noreau A, Xiong L, Zhou S, Jouan L, *et al.* (2011): Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat Genet*. 43: 860–863.

32. Takata A, Xu B, Ionita-Laza I, Roos JL, Gogos JA, Karayiorgou M (2014): Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. *Neuron*. 82: 773–780.

33. Kosmicki JA, Samocha KE, Howrigan DP, Sanders SJ, Slowikowski K, Lek M, *et al.* (2017): Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples. *Nat Genet*. 49: 504–510.

34. Takata A, Miyake N, Tsurusaki Y, Fukai R, Miyatake S, Koshimizu E, *et al.* (2018): Integrative Analyses of De Novo Mutations Provide Deeper Biological Insights into Autism Spectrum Disorder. *Cell Rep*. 22: 734–747.

35. Ware JS, Samocha KE, Homsy J, Daly MJ (2015): Interpreting de novo Variation in Human Disease Using denovolyzeR. *Curr Protoc Hum Genet*. 87: 7.25.1–15.

36. Simons Foundation Autism Research Initiative (n.d.): SFARI Gene. . Retrieved November 28, 2017, from https://gene.sfari.org/.

37. Shen L, Sinai M (2013): GeneOverlap: Test and visualize gene overlaps. *R package version*. 1.

38. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, *et al.* (2013): Software for computing and annotating genomic ranges. *PLoS Comput Biol*. 9: e1003118.

39. Sahni N, Yi S, Taipale M, Fuxman Bass JI, Coulombe-Huntington J, Yang F, *et al.* (2015): Widespread macromolecular interaction perturbations in human genetic disorders. *Cell*. 161: 647–660.

40. Lin GN, Corominas R, Lemmens I, Yang X, Tavernier J, Hill DE, *et al.* (2015): Spatiotemporal 16p11.2 protein network implicates cortical late mid-fetal brain development and KCTD13-Cul3-RhoA pathway in psychiatric diseases. *Neuron*. 85: 742–754.

41. Chatr-Aryamontri A, Breitkreutz B-J, Oughtred R, Boucher L, Heinicke S, Chen D, *et al.* (2015): The BioGRID interaction database: 2015 update. *Nucleic Acids Res*. 43: D470–8.

42. Li T, Wernersson R, Hansen RB, Horn H, Mercer J, Slodkowicz G, *et al.* (2017): A scored human protein-protein interaction network to catalyze genomic interpretation. *Nat Methods*. 14: 61–64.

43. Gandal MJ, Haney JR, Parikshak NN, Leppa V, Ramaswami G, Hartl C, *et al.* (2018): Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science*. 359: 693–697.

44. Li Z, Langhans SA (2015): Transcriptional regulators of Na,K-ATPase subunits. *Front Cell Dev Biol*. 3: 66.

45. Montgomery JR, Meredith AL (2012): Genetic activation of BK currents in vivo generates bidirectional effects on neuronal excitability. *Proc Natl Acad Sci U S A*. 109: 18997–19002.

46. Ortega-Martínez S (2015): A new perspective on the role of the CREB family of transcription factors in memory consolidation via adult hippocampal neurogenesis. *Front Mol Neurosci*. 8: 46.

47. Bergsland M, Werme M, Malewicz M, Perlmann T, Muhr J (2006): The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes Dev*. 20: 3475–3486.

48. Ronan JL, Wu W, Crabtree GR (2013): From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet*. 14: 347–359.

49. Uchida M, Enomoto A, Fukuda T, Kurokawa K, Maeda K, Kodama Y, *et al.* (2006): Dok-4 regulates GDNF-dependent neurite outgrowth through downstream activation of Rap1 and mitogen-activated protein kinase. *J Cell Sci*. 119: 3067–3077.

50. Schuster S, Rivalan M, Strauss U, Stoenica L, Trimbuch T, Rademacher N, *et al.* (2015): NOMA-GAP/ARHGAP33 regulates synapse development and autistic-like behavior in the mouse. *Mol Psychiatry*. 20: 1120–1131.

51. Peng Y-J, He W-Q, Tang J, Tao T, Chen C, Gao Y-Q, *et al.* (2010): Trio is a key guanine nucleotide exchange factor coordinating regulation of the migration and morphogenesis of granule cells in the developing cerebellum. *J Biol Chem*. 285: 24834–24844.

52. Fénelon K, Xu B, Lai CS, Mukai J, Markx S, Stark KL, *et al.* (2013): The pattern of cortical dysfunction in a mouse model of a schizophrenia-related microdeletion. *J Neurosci*. 33: 14825–14839.

53. Sun Z, Williams DJ, Xu B, Gogos JA (2018): Altered function and maturation of primary cortical neurons from a 22q11.2 deletion mouse model of schizophrenia. *Transl Psychiatry*. 8: 85.

54. Kaplan C, Steinmann M, Zapiorkowska NA, Ewers H (2017): Functional Redundancy of Septin Homologs in Dendritic Branching. *Front Cell Dev Biol*. 5: 11.

55. Beites CL, Campbell KA, Trimble WS (2005): The septin Sept5/CDCrel-1 competes with alpha-SNAP for binding to the SNARE complex. *Biochem J*. 385: 347–353.

56. Beites CL, Xie H, Bowser R, Trimble WS (1999): The septin CDCrel-1 binds syntaxin and inhibits exocytosis. *Nat Neurosci*. 2: 434–439.

57. Su Q, Mochida S, Tian JH, Mehta R, Sheng ZH (2001): SNAP-29: a general SNARE protein that inhibits SNARE disassembly and is implicated in synaptic transmission. *Proc Natl Acad Sci U S A*. 98: 14038–14043.

58. Falkenburger BH, Jensen JB, Dickson EJ, Suh B-C, Hille B (2010): Phosphoinositides: lipid regulators of membrane proteins. *J Physiol*. 588: 3179–3185.

59. Nakatsu F, Baskin JM, Chung J, Tanner LB, Shui G, Lee SY, *et al.* (2012): PtdIns4P synthesis by PI4KIIIα at the plasma membrane and its impact on plasma membrane identity. *J Cell Biol*. 199: 1003–1016.

60. Balakrishnan SS, Basu U, Shinde D, Thakur R, Jaiswal M, Raghu P (2018): Regulation of PI4P levels by PI4KIIIα during G-protein-coupled PLC signaling in Drosophila photoreceptors. *J Cell Sci*. 131. doi: 10.1242/jcs.217257.

61. Bulat V, Rast M, Pielage J (2014): Presynaptic CK2 promotes synapse organization and stability by targeting Ankyrin2. *J Cell Biol*. 204: 77–94.

62. Jungerius BJ, Hoogendoorn MLC, Bakker SC, Van't Slot R, Bardoel AF, Ophoff RA, *et al.* (2008): An association screen of myelin-related genes implicates the chromosome 22q11 PIK4CA gene in schizophrenia. *Mol Psychiatry*. 13: 1060–1068.

63. Kaur H, Jajodia A, Grover S, Baghel R, Jain S, Kukreti R (2014): Synergistic association of PI4KA and GRM3 genetic polymorphisms with poor antipsychotic response in south Indian schizophrenia patients with low severity of illness. *Am J Med Genet B Neuropsychiatr Genet*. 165B: 635–646.

64. Vorstman JAS, Chow EW, Ophoff RA, van Engeland H, Beemer FA, Kahn RS, *et al.* (2009): Association of the PIK4CA schizophrenia-susceptibility gene in adults with the 22q11.2 deletion syndrome. *Am J Med Genet B Neuropsychiatr Genet*. 150B: 430–433.

65. Rajman M, Schratt G (2017): MicroRNAs in neural development: from master regulators to fine-tuners. *Development*. 144: 2310–2322.

66. Chen W-Y, Shih H-T, Liu K-Y, Shih Z-S, Chen L-K, Tsai T-H, *et al.* (2015): Intellectual disability-associated dBRWD3 regulates gene expression through inhibition of HIRA/YEM-mediated chromatin deposition of histone H3.3. *EMBO Rep*. 16: 528–538.

67. Valenzuela N, Fan Q, Fa'ak F, Soibam B, Nagandla H, Liu Y, *et al.* (2016): Cardiomyocyte-specific conditional knockout of the histone chaperone HIRA in mice results in hypertrophy, sarcolemmal damage and focal replacement fibrosis. *Dis Model Mech*. 9: 335–345.

68. Fénelon K, Mukai J, Xu B, Hsu P-K, Drew LJ, Karayiorgou M, *et al.* (2011): Deficiency of Dgcr8, a gene disrupted by the 22q11.2 microdeletion, results in altered short-term plasticity in the prefrontal cortex. *Proc Natl Acad Sci U S A*. 108: 4447–4452.