An interaction-based model for neuropsychiatric features of copy-number variants

Matthew Jensen[1, 2] and Santhosh Girirajan[1, 2, 3]

[1]Department of Biochemistry and Molecular Biology, [2]Bioinformatics and Genomics Program, and [3]Department of Anthropology, Pennsylvania State University, University Park, PA 16802

Correspondence to:

Santhosh Girirajan, MBBS, PhD
205A Life Sciences Building
The Pennsylvania State University
University Park, PA 16802
E-mail: sxg47@psu.edu
Phone: 814-865-0674

ABSTRACT

Variably expressive copy-number variants (CNVs) are characterized by extensive phenotypic heterogeneity of neuropsychiatric phenotypes. Approaches to identify single causative genes for these phenotypes within each CNV have not been successful. Here, we posit using multiple lines of evidence, including pathogenicity metrics, functional assays of model organisms, and gene expression data, that multiple genes within each CNV region are likely responsible for the observed phenotypes. We propose that candidate genes within each region likely interact with each other through shared pathways to modulate the individual gene phenotypes, emphasizing the genetic complexity of CNV-associated neuropsychiatric features.

A case for a multi-genic model of CNV pathogenicity

Since the advent of large-scale sequencing studies, the number of genes associated with neurodevelopmental disorders such as autism, intellectual disability, and schizophrenia has increased dramatically. For example, nearly 200 genes have been identified with recurrent *de novo* mutations in both individuals with autism and intellectual disability (1–8). In fact, complex human disease phenotypes can be influenced by variation in both a small number of core genes with large effect size and a large number of modifier genes with small effect size, accounting for the large number of candidate neurodevelopmental genes (9,10). The application of a multi-genic model for disease pathogenicity has not been fully expanded to cover copy-number variants (CNVs), or large duplications and deletions in the genome. The prevailing notion of single causative genes for CNV disorders is due to the paradigm of gene discoveries for CNVs associated with genetic syndromes in individuals with specific constellations of clinical features, such as Smith-Magenis syndrome (SMS). Although some variability in phenotypic expression has been documented, these disorders usually occur *de novo* and are characterized by high penetrance for the observed phenotypes (11,12) (**Figure 1**). In these cases, individuals manifesting the characteristic features of the syndrome but with either atypical breakpoints or mutations in individual genes within the CNV region were used to identify causative genes for the major phenotypes (13–15). These causative genes, such as *RAI1* for SMS, were then confirmed by recapitulating conserved phenotypes of the deletion using functional evaluations in animal models (16,17).

In contrast, another category of CNVs has been identified in individuals with neurodevelopmental disorders, including duplications and deletions at proximal 16p11.2, 3q29, distal 16p11.2, and 1q21.1 (18–21). Although these CNVs are enriched in affected individuals compared to population controls, they are primarily characterized by variable expressivity of clinical features (12,22–26) (**Figure 1B**). For example, the 16p11.2 deletion has been implicated in 1% of individuals with idiopathic autism (18,27), but only 25% of individuals with the deletion exhibit an autism phenotype (28–31), while others may manifest intellectual disability, obesity, or epilepsy at varying degrees of penetrance (28,32,33). In fact, certain CNVs, such as the 16p12.1 deletion and the 15q11.2 deletion, have a high frequency of carriers who only manifest mild neuropsychiatric features, in contrast to more severely affected individuals who also carry other rare variants in the genetic background (12,22,23,26,34,35). As such, many

variably expressive CNVs have a higher frequency of inherited compared to *de novo* occurrence (12) (**Figure 1A**).

Based on the success of gene discovery in CNVs with syndromic features, such as SMS, several studies have attempted to identify the causative genes in variably expressive CNVs (36–52). Several individual genes within variably expressive CNV regions have been associated with specific congenital or structural features of these disorders, including *TBX6* for scoliosis in 16p11.2 deletion (53), *TBX1* for cardiac phenotypes in 22q11.2 deletion (38,54), *GJA8* for cataracts and *GJA5* for heart defects in 1q21.1 deletion (55,56), and *MYH11* for aortic aneurysms in 16p13.11 duplication (57,58). However, approaches to identify single causative genes for the more prominent neuropsychiatric features of these CNVs have not been successful (59). Here, we show several lines of evidence from gene pathogenicity metrics, animal model studies, and gene expression data that support the involvement of multiple genes towards the neuropsychiatric features of variably expressive CNVs.

First, genome-wide metrics of pathogenicity, including those that measure haploinsufficiency (HI score, gene essentiality, GHIS and EpiScore) (60–63) and resistance to variation (RVIS, pLI and maximum CCR scores) (64–66), provide evidence for several candidate genes within CNV regions for developmental disorders **(Figure 2)**. For example, 45 out of 152 genes (30%) within 12 variably expressive CNV regions are intolerant to variation with RVIS metrics in the top 20th genome-wide percentile, similar to that of known neurodevelopmental genes such as *CHD8*, *NRXN1* and *SCN2A*, as well as genes responsible for major features of syndromic CNVs, such as *RAI1* and *NSD1* (**Figure 2A**). These top-ranked genes include *TAOK2*, *MVP*, *ALDOA* and *DOC2A* on chromosome 16p11.2, *BCL9* and *GJA5* on chromosome 1q21.1, and *ATXN2L*, *ATP2A1* and *SH2B1* on distal 16p11.2. Similarly, 32/165 genes (19%) are considered intolerant to loss-of-function mutations based on pLI scores (>0.9), and 36/160 genes (23%) have haploinsufficiency scores in the highest 20th percentile of the entire genome (**Figure 2A**). Further, the top 10% of all genes identified by a gene interaction-based machine learning classifier to be associated with autism included eight genes within 16p11.2 and four genes within 22q11.2 (67).

Second, several recent studies using animal and cellular models have demonstrated the critical involvement of several genes within CNVs towards neurological, cellular and developmental functions (36,37,46,47,51,52) (**Figure 2B**). For example, Blaker-Lee *et al*.

screened 22 homologs of 16p11.2 genes in zebrafish morpholino knockdown models, and identified 20 homologs that contributed to morphological defects and abnormal behavior (37). Iyer *et al.* also screened homologs of 16p11.2 genes in *Drosophila melaogaster* using RNAi knockdown, and found that 10 out of 14 homologs contributed to global developmental defects as well as specific neuronal and cellular defects in the developing fly eye (46). Further, mouse models for 15 genes within the 16p11.2 region have been generated to test for defects in development and neuronal behavior (45,48–50,68–80). For example, *Taok2*$^{-/-}$ mice have increased brain size, behavioral defects, and impaired synapse development (50), *Kcdt13*$^{+/-}$ mice show defects in hippocampal synaptic transmission and decreased dendritic complexity (45), *Mapk3*$^{+/-}$ mice show behavior anomalies, abnormal synapse function and reduced cell proliferation during development (68,69), and *Mvp*$^{+/-}$ mice show decreased plasticity and synaptic defects in ocular neurons (48) (**Figure 2B**). Importantly, these models of individual genes do not fully recapitulate the phenotypes observed in models of the entire CNV (81–85). For example, the decreased body weight, abnormal brain morphology and coordination defects observed in 16p11.2 deletion mouse models have not been observed in any individual gene knockdown models (81–84) (**Figure 2B**). Similarly, *Otud7a*$^{+/-}$ mouse models have low body weight, reduced vocalization, abnormal dendritic spine morphology, and seizures, but the 15q13.3 deletion mice also show learning and memory defects in addition to the above features (43,44,86). Further, mouse models for *Chrna7*$^{+/-}$, another candidate gene on chromosome 15q13.3, only show subtle behavioral phenotypes (87). These data suggest that haploinsufficiency of *CHRNA7* or *OTUD7A* alone is not sufficient to account for the pathogenicity of the entire CNV. Overall, a catalog of functional data from mouse (88), zebrafish (89), and fruit fly studies (90) indicates that 80% (131/163) of homologs for genes within CNV regions present lethality, behavioral, developmental, or neuronal phenotypes when disrupted. These data suggest that disruption of multiple genes within each CNV region can affect important developmental or neuronal functions that could contribute to the phenotypes of the entire CNV.

Third, patterns of gene expression in humans and model organisms have identified multiple genes within each CNV region that are co-expressed in the developing brain along with known neurodevelopmental genes. For example, Maynard and colleagues examined expression patterns of 22q11.2 gene homologs in the developing mouse brain, and found that 27 out of 32

genes were expressed in the embryonic forebrain, with six genes expressed in neuronal tissues related to schizophrenia (39). In fact, a genome-wide weighted gene correlation network analysis (WGCNA) (91) from different brain tissues during development (92) shows several large modules of genes with similar expression patterns (**Figure 3**). For example, the five largest modules are each enriched ($p<0.05$ with Benjamini-Hochberg correction) for biological functions related to neurodevelopment, including protein modification and transport in module 1 (M1), nervous system development in M2, and cell communication and signal transduction in M5. Importantly, each of these modules contains multiple genes from the same CNV region, including 3q29 genes *PAK2*, *NCBP2*, and *BDH1* in M1, 1q21.1 genes *BCL9*, *CHD1L* and *FMO5* in M2, and 16p11.2 genes *MVP* and *QPRT* in M5. Therefore, it is clear that multiple genes in the same CNV region are co-expressed with each other in the developing brain and could share similar functions or regulatory patterns.

Dissecting the genetic complexity of CNV pathogenicity

Several scenarios could explain how the haploinsufficiency of multiple genes can predict the variable phenotypes associated with the entire CNV (**Figure 4A**). The simplest such model is an additive model, where disruption of individual genes within a CNV may only impart a mild phenotype on their own, but additively contribute to more severe features (93) (**Figure 4A**). However, an additive model may not always explain the phenotypic features manifested by CNVs containing multiple candidate genes that could lead to severe defects or lethality on their own. For example, heterozygous *Tbx1*$^{+/-}$ (within the 22q11.2 region) and *Mapk1*$^{+/-}$ (within the distal 22q11.2 region) mice both lead to perinatal or neonatal lethality (94–96). In humans, 14% (24/172) of CNV genes are under evolutionary constraint in control populations (pLI score >0.9 or maximum CCR score >99th percentile) and have no reported disease-associated variants (97–99), suggesting that these genes could be under strong purifying selection (66). Further, 18% (22/125) of CNV genes show evolutionary constraint for loss-of-function mutations (pLI>0.9) but not for copy-number changes within a control population (100). We therefore hypothesize that the pathogenicity of variably expressive CNVs can also be explained by complex interactions among the constituent genes within shared biological pathways. These interactions can enhance or suppress the phenotypes caused by disruption of individual genes. Under this model, the haploinsufficiency of certain genes can be modulated by haploinsufficiency of other

153 interacting genes in the same region that may or may not lead to phenotypes on their own
154 (**Figure 4A**). Further, variants in the genetic background that map within these shared pathways
155 can simultaneously modulate the effects of multiple genes, ultimately defining the phenotypic
156 trajectory in CNV carriers (**Figure 4A**). For example, Pizzo *et al.* found that the burden of rare
157 deleterious mutations within genes in the genetic background correlated with variability of IQ
158 scores and head circumference among 16p11.2 deletion carriers (35). The potential for complex
159 interactions within a CNV region depends on the functional convergence of the constituent
160 genes. For instance, both *KCTD13* and *TAOK2* within 16p11.2 participate in the RhoA signaling
161 pathway (45,50) and therefore are more likely to interact with each other than genes located in
162 different biological pathways. In fact, it has been shown that genes within pathogenic CNVs are
163 more similar in function compared to genes within benign CNVs, suggesting that variably
164 expressive CNVs are likely to contain interactions between functionally relevant genes (101).
165 Further, Noh and colleagues found an over-representation of interactions among genes within
166 autism-associated CNVs, and these interactions were enriched for synaptic transmission and
167 regulatory signaling pathways (102). Because of this, therapeutic targets for pathways shared
168 among CNV genes could be explored as potential treatments for CNV disorders.
169 The possibility of additive, suppressor and enhancer interactions between pairs of genes
170 underlies the potential for highly complex models of CNV pathogenicity. For instance, within a
171 CNV region spanning three genes, seven combinations of gene knockdown experiments
172 (haploinsufficiency of A, B, C, AB, BC, AC, and ABC) can be tested for the presence or absence
173 of a specific phenotype (**Figure 4B**). This set of knockdown experiments can yield 128 possible
174 experimental outcomes that can be used to further deduce 64 possible sets of pairwise
175 interactions for AB, BC, and AC (no interaction, additive, suppression, or enhancement for each
176 interaction) (**Figure 4B**). These possible combinations of interactions exponentially increase for
177 larger CNVs with more genes, and the complexity further increases if quantitative phenotypes
178 are used to determine the magnitude of interactions between genes or when interactions with
179 variants in the genetic background are taken into account. However, testing even a small number
180 of these interactions would still uncover the nature of the relationships among genes within a
181 CNV region and potentially a common pathway shared by those genes. For example, Grice and
182 colleagues used *D. melanogaster* RNAi models to identify six synergistic interactions out of 41
183 tested pairwise interactions between genes within *de novo* CNVs from autism patients, including

partial 3q29 and 22q11.2 deletions (103). Iyer *et al*. also used fly models to identify 24 additive, enhancer and suppressor interactions out of 52 tested pairwise interactions among homologs of 16p11.2 genes (46), providing further evidence for complex interactions within CNV regions. Further, these interaction models for CNV pathogenicity can be tested in cellular models of the entire CNV. For example, a more severe phenotype observed by restoring dosage of a candidate gene would suggest that disruption of this gene potentially suppresses the effects of other genes within the CNV.

Complex genetic interactions in the context of genome sequencing

In recent years, exome and whole-genome sequencing analysis has proven invaluable in identifying candidate genes for neurodevelopmental disorders (104). However, sequencing studies would not be able to capture the genetic complexity of a multi-genic CNV region. For example, genes that cause severe phenotypes or lethality on their own and are modulated by haploinsufficiency of other interacting genes within a CNV are less likely to have an enrichment of mutations in sequencing studies. Further, because of the strong phenotypic heterogeneity of these CNVs, it is not possible to determine whether the phenotypes of any individual candidate gene fully recapitulate the variable phenotypes of the entire CNV region. Candidate genes within CNVs identified through genome sequencing studies, such as *TAOK2* on chromosome 16p11.2 (50) or *CHRNA7* on chromosome 15q13.3 (105), do not preclude the possibility of other candidate genes in the same region. Because of this, a thorough systems-based approach for each gene within a CNV and its interactions is necessary to identify candidate genes responsible for the neuropsychiatric features of each region (106).

In summary, genomic and functional data have implicated multiple genes in variably expressive CNV regions towards neuropsychiatric phenotypes, suggesting that single causative genes are not responsible for the heterogeneous features of these CNVs. Here, we propose a complex interaction-based model for these CNVs, where candidate genes within each region interact with each other to influence the variable clinical outcome. The CNV phenotype is therefore distinct from the phenotype manifested by any individual gene, or in some cases, the additive effects of all genes in the region. This multi-genic model of CNVs agrees with a broader complex genetic view of neurodevelopmental disorders, where hundreds of genes with varying effect sizes and complex interactions influence developmental features (10). Further studies on

the role of individual genes in CNV regions towards neurodevelopment, especially those that identify key interactions between genes, will be useful in uncovering the cellular pathways and mechanisms responsible for the observed neuropsychiatric features.

Acknowledgements

The authors thank Lucilla Pizzo, Vijay Kumar and Maitreya Das for their helpful discussions and comments on the manuscript.

Contributions

M.J. and S.G. conceptualized and wrote the Viewpoint article.

Funding

This work was supported by NIH R01-GM121907, SFARI Pilot Grant (#399894) and resources from the Huck Institutes of the Life Sciences to SG, and NIH T32-GM102057 to MJ. The funders had no role in the preparation of this manuscript.

Competing interests

The authors declare that they have no financial or non-financial competing interests.

FIGURE LEGENDS

Figure 1. Phenotypic profiles of syndromic and variably expressive CNVs. **(A)** Table listing variably expressive (top) and syndromic (bottom) CNV regions is shown. The colored boxes indicate frequency of *de novo* versus inherited CNV cases for deletions (del) and duplications (dup) previously identified in a cohort of 2,312 children with developmental disorders (12). The twelve variably-expressive CNV regions highlighted in bold were selected for the analysis described in the manuscript. **(B)** Table listing average frequencies of neurodevelopmental phenotypes for select variably-expressive and syndromic CNVs, curated from GeneReviews reports on individual CNVs (107), is shown. White boxes represent no available data from GeneReviews, but do not necessarily indicate a lack of association between the CNV and the phenotype (for example, 1q21.1 deletion and schizophrenia). Data for this figure are available in the **Supporting Information** file.

Figure 2. (A) Percentile-rank scores compared to the whole genome for intolerance to variation (RVIS, pLI and maximum CCR) and haploinsufficiency (HI, essentiality, GHIS and EpiScore) metrics for genes within select variably expressive CNV regions (60–66). Lower percentile scores indicate a gene is more likely to be haploinsufficient or intolerant to variation. Grey boxes indicate metrics were not available for a particular gene. **(B)** Developmental phenotypes in animal models for homologs of individual genes within the 16p11.2 region, as catalogued from animal model databases (MGI, ZFIN and FlyBase). Black boxes indicate presence of phenotype, white boxes indicate absence of phenotype, and grey boxes indicate no homolog is present for a particular gene in a model organism. The phenotypes observed in 16p11.2 deletion and duplication mice are distinct from those observed in the individual gene models (81–85). Data for this figure, including gene metrics and animal phenotypes for other CNV genes not shown in this figure, are available in the **Supporting Information** file. (Abbreviations: RVIS—Residual Variance to Intolerance Score; pLI—Probability of Loss-of-function Intolerance; CCR—Constrained Coding Regions; HI—Haploinsufficiency score, GHIS—Genome-wide haploinsufficiency score; MGI—Mouse Genome Informatics; ZFIN—Zebrafish Information Network)

Figure 3. Modules of co-expressed genes derived from WGCNA analysis of BrainSpan Atlas RNA-Seq data (Gencode v10) (92) across 524 tissues and timepoints the developing brain. Networks of interactions among genes within three select top WGCNA modules (M1, M2 and M5) were obtained from the BioGrid interaction database (108) and visualized with Cytoscape (109). Genes within variably expressive CNV regions are highlighted as colored nodes in each network. Bar graphs show enrichment ($p<0.05$ with Benjamini-Hochsberg correction, represented by red dotted line) of genes within each module for Gene Ontology (GO) Biological Process terms, calculated using PantherDB (110). Data for this figure are available in the **Supporting Information** file.

Figure 4. Models for genetic interactions within CNV regions. **(A)** Several models of interactions among CNV genes are shown. These models include (i) a single-gene model where one gene is sufficient to account for the phenotype; additive models where the phenotype is due to the additive effects of multiple CNV genes that (ii) may or (iii) may not account for phenotypes on their own; and (iv) a complex interaction model where additive, enhancer and suppressor interactions between genes in the CNV region modulate the phenotype, including when additive effects could lead to lethality. The size of the circles in the plot indicates the relative contribution of each gene to the overall neurodevelopmental phenotype. Thick circles indicate genes that contribute to the observed phenotypes on their own, while connector lines indicate the nature of interaction between pairs of genes. Connected modifier genes (M) can further modulate these interactions to ultimately define the phenotypic trajectory in individuals carrying the CNV. **(B)** For a hypothetical CNV region with three genes, there are seven combinations of gene knockdowns (A, B, C, AB, BC and ABC) that can be tested for the presence or absence of a specific phenotype. These knockdown experiments can yield 128 potential outcomes for each phenotype tested, with each individual set of outcomes corresponding to one of 64 combinations of pairwise gene interactions (additive, enhancer, suppressor or no interaction). One possible outcome highlighted in orange shows presence of a particular phenotype for knockdowns of single genes A and B and two-hit knockdowns AB and BC. The single-gene knockdowns indicate that only genes A and B contribute to the phenotype, and that the phenotype of pairwise knockdown AB is due to the additive effects of the two genes.

While the phenotype is observed for BC, the phenotype is not observed for AC and ABC, suggesting that gene C suppresses the phenotype of gene A.

REFERENCES

1. O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science. 2012 Dec 21;338(6114):1619–22.
2. Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature. 2012 Apr 4;485(7397):237–41.
3. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, et al. The contribution of de novo coding mutations to autism spectrum disorder. Nature. 2014 Nov 13;515(7526):216–21.
4. Turner TN, Hormozdiari F, Duyzend MH, McClymont SA, Hook PW, Iossifov I, et al. Genome Sequencing of Autism-Affected Families Reveals Disruption of Putative Noncoding Regulatory DNA. Am J Hum Genet. 2016;98(1):58–74.
5. Yuen RKC, Merico D, Bookman M, Howe JL, Thiruvahindrapuram B, Patel R V, et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat Neurosci. 2017 Apr 6;20(4):602–11.
6. Gilissen C, Hehir-Kwa JY, Thung DT, Van De Vorst M, Van Bon BWM, Willemsen MH, et al. Genome sequencing identifies major causes of severe intellectual disability. Nature. 2014 Jul 17;511(7509):344–7.
7. Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. Nature. 2017 Feb 25;542(7642):433–8.
8. Kosmicki JA, Samocha KE, Howrigan DP, Sanders SJ, Slowikowski K, Lek M, et al. Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples. Nat Genet. 2017 Apr 13;49(4):504–10.
9. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell. 2017 Jun;169(7):1177–86.
10. Jensen M, Girirajan S. Mapping a shared genetic basis for neurodevelopmental disorders. Genome Med. 2017 Dec 14;9(1):109.
11. Girirajan S, Eichler EE. Phenotypic variability and genetic susceptibility to genomic disorders. Hum Mol Genet. 2010 Oct 15;19(R2):R176-87.
12. Girirajan S, Rosenfeld JA, Coe BP, Parikh S, Friedman N, Goldstein A, et al. Phenotypic

Heterogeneity of Genomic Disorders and Rare Copy-Number Variants. N Engl J Med. 2012 Oct 4;367(14):1321–31.

13. Slager RE, Newton TL, Vlangos CN, Finucane B, Elsea SH. Mutations in RAI1 associated with Smith–Magenis syndrome. Nat Genet. 2003 Apr 24;33(4):466–8.

14. Bi W, Saifi GM, Shaw CJ, Walz K, Fonseca P, Wilson M, et al. Mutations of RAI1, a PHD-containing protein, in nondeletion patients with Smith-Magenis syndrome. Hum Genet. 2004 Nov 30;115(6):515–24.

15. Girirajan S, Vlangos CN, Szomju BB, Edelman E, Trevors CD, Dupuis L, et al. Genotype-phenotype correlation in Smith-Magenis syndrome: Evidence that multiple genes in 17p11.2 contribute to the clinical spectrum. Genet Med. 2006 Jul;8(7):417–27.

16. Bi W, Ohyama T, Nakamura H, Yan J, Visvanathan J, Justice MJ, et al. Inactivation of Rai1 in mice recapitulates phenotypes observed in chromosome engineered mouse models for Smith-Magenis syndrome. Hum Mol Genet. 2005 Apr 15;14(8):983–95.

17. Walz K, Caratini-Rivera S, Bi W, Fonseca P, Mansouri DL, Lynch J, et al. Modeling del(17)(p11.2p11.2) and dup(17)(p11.2p11.2) contiguous gene syndromes by chromosome engineering in mice: phenotypic consequences of gene dosage imbalance. Mol Cell Biol. 2003 May;23(10):3646–55.

18. Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, et al. Association between Microdeletion and Microduplication at 16p11.2 and Autism. N Engl J Med. 2008 Feb 14;358(7):667–75.

19. Willatt L, Cox J, Barber J, Cabanas ED, Collins A, Donnai D, et al. 3q29 microdeletion syndrome: clinical and molecular characterization of a new syndrome. Am J Hum Genet. 2005 Jul;77(1):154–60.

20. Bachmann-Gagescu R, Mefford HC, Cowan C, Glew GM, Hing A V, Wallace S, et al. Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. Genet Med. 2010 Oct 30;12(10):641–7.

21. Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, Sahoo T, et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. Nat Genet. 2008 Dec 23;40(12):1466–71.

22. Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, et al. A copy number

variation morbidity map of developmental delay. Nat Genet. 2011 Sep 14;43(9):838–46.

23. Girirajan S, Rosenfeld JA, Cooper GM, Antonacci F, Siswara P, Itsara A, et al. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. Nat Genet. 2010 Mar 14;42(3):203–9.

24. Rosenfeld JA, Coe BP, Eichler EE, Cuckle H, Shaffer LG. Estimates of penetrance for recurrent pathogenic copy-number variations. Genet Med. 2013 Jun 20;15(6):478–81.

25. Kirov G, Rees E, Walters JTR, Escott-Price V, Georgieva L, Richards AL, et al. The Penetrance of Copy Number Variations for Schizophrenia and Developmental Delay. Biol Psychiatry. 2014 Mar 1;75(5):378–85.

26. Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. Nature. 2014 Jan 18;505(7483):361–6.

27. Kumar RA, Karamohamed S, Sudi J, Conrad DF, Brune C, Badner JA, et al. Recurrent 16p11.2 microdeletions in autism. Hum Mol Genet. 2008 Nov 7;17(4):628–38.

28. Shinawi M, Liu P, Kang SHL, Shen J, Belmont JW, Scott DA, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. J Med Genet. 2010 May 1;47(5):332–41.

29. Zufferey F, Sherr EH, Beckmann ND, Hanson E, Maillard AM, Hippolyte L, et al. A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. J Med Genet. 2012 Oct;49(10):660–8.

30. Hanson E, Bernier R, Porche K, Jackson FI, Goin-Kochel RP, Snyder LG, et al. The Cognitive and Behavioral Phenotype of the 16p11.2 Deletion in a Clinically Ascertained Population. Biol Psychiatry. 2015 May 1;77(9):785–93.

31. Moreno-De-Luca A, Evans DW, Boomer KB, Hanson E, Bernier R, Goin-Kochel RP, et al. The role of parental cognitive, behavioral, and motor profiles in clinical variability in individuals with chromosome 16p11.2 deletions. JAMA Psychiatry. 2015 Feb 1;72(2):119–26.

32. Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, Blaszczyk K, et al. Large, rare chromosomal deletions associated with severe early-onset obesity. Nature. 2010 Feb 6;463(7281):666–70.

389 33. Walters RG, Jacquemont S, Valsesia A, de Smith AJ, Martinet D, Andersson J, et al. A
390 new highly penetrant form of obesity due to deletions on chromosome 16p11.2. Nature.
391 2010 Feb 4;463(7281):671–5.
392 34. Chaste P, Sanders SJ, Mohan KN, Klei L, Song Y, Murtha MT, et al. Modest impact on
393 risk for autism spectrum disorder of rare copy number variants at 15q11.2, Specifically
394 Breakpoints 1 to 2. Autism Res. 2014 Jun;7(3):355–62.
395 35. Pizzo L, Jensen M, Polyak A, Rosenfeld JA, Mannik K, Krishnan A, et al. Rare variants in
396 the genetic background modulate cognitive and developmental phenotypes in individuals
397 carrying disease-associated variants. Genet Med. 2018 Sep 7;1–10.
398 36. Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S, et al. KCTD13 is
399 a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number
400 variant. Nature. 2012 May 16;485(7398):363–7.
401 37. Blaker-Lee A, Gupta S, McCammon JM, De Rienzo G, Sive H. Zebrafish homologs of
402 genes within 16p11.2, a genomic region associated with brain disorders, are active during
403 brain development, and include two deletion dosage sensor genes. Dis Model Mech. 2012
404 Nov 1;5(6):834–51.
405 38. Lindsay EA, Vitelli F, Su H, Morishima M, Huynh T, Pramparo T, et al. Tbx1
406 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice.
407 Nature. 2001 Mar 1;410(6824):97–101.
408 39. Maynard TM, Haskell GT, Peters AZ, Sikich L, Lieberman JA, LaMantia A-S. A
409 comprehensive analysis of 22q11 gene expression in the developing and adult brain. Proc
410 Natl Acad Sci. 2003 Nov 25;100(24):14433–8.
411 40. Prasad SE, Howley S, Murphy KC. Candidate genes and the behavioral phenotype in
412 22q11.2 deletion syndrome. Dev Disabil Res Rev. 2008;14(1):26–34.
413 41. Gillentine MA, Schaaf CP. The human clinical phenotypes of altered CHRNA7 copy
414 number. Biochem Pharmacol. 2015 Oct 15;97(4):352–62.
415 42. Gillentine MA, Yin J, Bajic A, Zhang P, Cummock S, Kim JJ, et al. Functional
416 Consequences of CHRNA7 Copy-Number Alterations in Induced Pluripotent Stem Cells
417 and Neural Progenitor Cells. Am J Hum Genet. 2017 Dec 7;101(6):874–87.
418 43. Uddin M, Unda BK, Kwan V, Holzapfel NT, White SH, Chalil L, et al. OTUD7A
419 Regulates Neurodevelopmental Phenotypes in the 15q13.3 Microdeletion Syndrome. Am

J Hum Genet. 2018 Feb 1;102(2):278–95.

44. Yin J, Chen W, Chao ES, Soriano S, Wang L, Wang W, et al. Otud7a Knockout Mice Recapitulate Many Neurological Features of 15q13.3 Microdeletion Syndrome. Am J Hum Genet. 2018 Feb 1;102(2):296–308.

45. Escamilla CO, Filonova I, Walker AK, Xuan ZX, Holehonnur R, Espinosa F, et al. Kctd13 deletion reduces synaptic transmission via increased RhoA. Nature. 2017 Nov 9;551(7679):227–31.

46. Iyer J, Singh MD, Jensen M, Patel P, Pizzo L, Huber E, et al. Pervasive genetic interactions modulate neurodevelopmental defects of the autism-associated 16p11.2 deletion in Drosophila melanogaster. Nat Commun. 2018 Dec 29;9(1):2548.

47. McCammon JM, Blaker-Lee A, Chen X, Sive H. The 16p11.2 homologs fam57ba and doc2a generate certain brain and body phenotypes. Hum Mol Genet. 2017 Oct 1;26(19):3699–712.

48. Ip JPK, Nagakura I, Petravicz J, Li K, Wiemer EAC, Sur M. Major vault protein, a candidate gene in 16p11.2 microdeletion syndrome, is required for the homeostatic regulation of visual cortical plasticity. J Neurosci. 2018 Apr 18;38(16):2034–17.

49. Yadav S, Oses-Prieto JA, Peters CJ, Zhou J, Pleasure SJ, Burlingame AL, et al. TAOK2 Kinase Mediates PSD95 Stability and Dendritic Spine Maturation through Septin7 Phosphorylation. Neuron. 2017 Jan 18;93(2):379–93.

50. Richter M, Murtaza N, Scharrenberg R, White SH, Johanns O, Walker S, et al. Altered TAOK2 activity causes autism-related neurodevelopmental and cognitive abnormalities through RhoA signaling. Mol Psychiatry. 2018 Feb 21;1–22.

51. Park SM, Littleton JT, Park HR, Lee JH. Drosophila Homolog of Human KIF22 at the Autism-Linked 16p11 . 2 Loci Influences Synaptic Connectivity at Larval Neuromuscular Junctions. Exp Neurobiol. 2016 Feb;25(1):33–9.

52. Loviglio MN, Arbogast T, Jønch AE, Collins SC, Popadin K, Bonnet CS, et al. The Immune Signaling Adaptor LAT Contributes to the Neuroanatomical Phenotype of 16p11.2 BP2-BP3 CNVs. Am J Hum Genet. 2017 Oct 5;101(4):564–77.

53. Wu N, Ming X, Xiao J, Wu Z, Chen X, Shinawi M, et al. TBX6 Null Variants and a Common Hypomorphic Allele in Congenital Scoliosis. N Engl J Med. 2015 Jan 22;372(4):341–50.

54. Gong W, Gottlieb S, Collins J, Blescia A, Dietz H, Goldmuntz E, et al. Mutation analysis of TBX1 in non-deleted patients with features of DGS/VCFS or isolated cardiovascular defects. J Med Genet. 2001 Dec;38(12):E45.

55. Ponnam SPG, Ramesha K, Tejwani S, Ramamurthy B, Kannabiran C. Mutation of the gap junction protein alpha 8 (GJA8) gene causes autosomal recessive cataract. J Med Genet. 2007 Jan 19;44(7):e85–e85.

56. Soemedi R, Topf A, Wilson IJ, Darlay R, Rahman T, Glen E, et al. Phenotype-specific effect of chromosome 1q21.1 rearrangements and GJA5 duplications in 2436 congenital heart disease patients and 6760 controls. Hum Mol Genet. 2012 Apr 1;21(7):1513–20.

57. Pannu H, Tran-Fadulu V, Papke CL, Scherer S, Liu Y, Presley C, et al. MYH11 mutations result in a distinct vascular pathology driven by insulin-like growth factor 1 and angiotensin II. Hum Mol Genet. 2007 Oct 15;16(20):2453–62.

58. Kuang SQ, Guo DC, Prakash SK, McDonald MLN, Johnson RJ, Wang M, et al. Recurrent chromosome 16p13.1 duplications are a risk factor for aortic dissections. PLoS Genet. 2011 Jun 16;7(6):e1002118.

59. Girirajan S, Campbell CD, Eichler EE. Human Copy Number Variation and Complex Genetic Disease. Annu Rev Genet. 2011 Dec 15;45(1):203–26.

60. Huang N, Lee I, Marcotte EM, Hurles ME. Characterising and predicting haploinsufficiency in the human genome. PLoS Genet. 2010 Oct 14;6(10):1–11.

61. Steinberg J, Honti F, Meader S, Webber C. Haploinsufficiency predictions without study bias. Nucleic Acids Res. 2015 Sep 3;43(15):e101–e101.

62. Khurana E, Fu Y, Chen J, Gerstein M. Interpretation of Genomic Variants Using a Unified Biological Network Approach. PLoS Comput Biol. 2013 Mar 7;9(3):e1002886.

63. Han X, Chen S, Flynn E, Wu S, Wintner D, Shen Y. Distinct epigenomic patterns are associated with haploinsufficiency and predict risk genes of developmental disorders. Nat Commun. 2018 May 30;9(1):2138.

64. Lek M, Karczewski KJ, Minikel E V., Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016 Aug 18;536(7616):285–91.

65. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic Intolerance to Functional Variation and the Interpretation of Personal Genomes. PLoS Genet. 2013 Aug

22;9(8):e1003709.

66. Havrilla JM, Pedersen BS, Layer RM, Quinlan AR. A map of constrained coding regions in the human genome. bioRxiv. Cold Spring Harbor Laboratory; 2017 Nov 22;220814.

67. Krishnan A, Zhang R, Yao V, Theesfeld CL, Wong AK, Tadych A, et al. Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder. Nat Neurosci. 2016 Nov 1;19(11):1454–62.

68. Mazzucchelli C, Vantaggiato C, Ciamei A, Fasano S, Pakhotin P, Krezel W, et al. Knockout of ERK1 MAP kinase enhances synaptic plasticity in the striatum and facilitates striatal-mediated learning and memory. Neuron. 2002 May 30;34(5):807–20.

69. Voisin L, Saba-El-Leil MK, Julien C, Fremin C, Meloche S. Genetic Demonstration of a Redundant Role of Extracellular Signal-Regulated Kinase 1 (ERK1) and ERK2 Mitogen-Activated Protein Kinases in Promoting Fibroblast Proliferation. Mol Cell Biol. 2010 Jun 15;30(12):2918–32.

70. Haller M, Au J, O'Neill M, Lamb DJ. 16p11.2 transcription factor MAZ is a dosage-sensitive regulator of genitourinary development. Proc Natl Acad Sci. 2018 Feb 20;115(8):201716092.

71. Ohsugi M, Adachi K, Horai R, Kakuta S, Sudo K, Kotaki H, et al. Kid-Mediated Chromosome Compaction Ensures Proper Nuclear Envelope Formation. Cell. 2008 Mar 7;132(5):771–82.

72. Michetti C, Castroflorio E, Marchionni I, Forte N, Sterlini B, Binda F, et al. The PRRT2 knockout mouse recapitulates the neurological diseases associated with PRRT2 mutations. Neurobiol Dis. 2017 Mar;99:66–83.

73. Kumar A, Lualdi M, Loncarek J, Cho YW, Lee JE, Ge K, et al. Loss of function of mouse Pax-Interacting Protein 1-associated glutamate rich protein 1a (Pagr1a) leads to reduced Bmp2 expression and defects in chorion and amnion development. Dev Dyn. 2014 Jul;243(7):937–47.

74. Lee C-M, He CH, Nour AM, Zhou Y, Ma B, Park JW, et al. IL-13Rα2 uses TMEM219 in chitinase 3-like-1-induced signalling and effector responses. Nat Commun. 2016 Sep 15;7:12752.

75. Sakaguchi G, Manabe T, Kobayashi K, Orita S, Sasaki T, Naito A, et al. Doc2α is an activity-dependent modulator of excitatory synaptic transmission. Eur J Neurosci. 1999

513 Dec;11(12):4262–8.
514 76. Shui J-W, Hu MC-T, Tan T-H. Conditional Knockout Mice Reveal an Essential Role of
515 Protein Phosphatase 4 in Thymocyte Development and Pre-T-Cell Receptor Signaling.
516 Mol Cell Biol. 2007 Jan 1;27(1):79–91.
517 77. Toyo-oka K, Mori D, Yano Y, Shiota M, Iwao H, Goto H, et al. Protein phosphatase 4
518 catalytic subunit regulates Cdk1 activity and microtubule organization via NDEL1
519 dephosphorylation. J Cell Biol. 2008 Mar 24;180(6):1133–47.
520 78. Chapman DL, Papaioannou VE. Three neural tubes in mouse embryos with mutations in
521 the T-box gene Tbx6. Nature. 1998 Feb 12;391(6668):695–7.
522 79. Takemoto T, Uchikawa M, Yoshida M, Bell DM, Lovell-Badge R, Papaioannou VE, et al.
523 Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells.
524 Nature. 2011 Feb 17;470(7334):394–8.
525 80. Pagès G, Guérin S, Grall D, Bonino F, Smith A, Anjuere F, et al. Defective thymocyte
526 maturation in p44 MAP kinase (Erk 1) knockout mice. Science. 1999 Nov
527 12;286(5443):1374–7.
528 81. Pucilowska J, Vithayathil J, Tavares EJ, Kelly C, Karlo JC, Landreth GE. The 16p11.2
529 Deletion Mouse Model of Autism Exhibits Altered Cortical Progenitor Proliferation and
530 Brain Cytoarchitecture Linked to the ERK MAPK Pathway. J Neurosci. 2015 Feb
531 18;35(7):3190–200.
532 82. Arbogast T, Ouagazzal A-M, Chevalier C, Kopanitsa M, Afinowi N, Migliavacca E, et al.
533 Reciprocal Effects on Neurocognitive and Metabolic Phenotypes in Mouse Models of
534 16p11.2 Deletion and Duplication Syndromes. PLOS Genet. 2016 Feb
535 12;12(2):e1005709.
536 83. Portmann T, Yang M, Mao R, Panagiotakos G, Ellegood J, Dolen G, et al. Behavioral
537 abnormalities and circuit defects in the basal ganglia of a mouse model of 16p11.2
538 deletion syndrome. Cell Rep. 2014 May 22;7(4):1077–92.
539 84. Horev G, Ellegood J, Lerch JP, Son Y-EE, Muthuswamy L, Vogel H, et al. Dosage-
540 dependent phenotypes in models of 16p11.2 lesions found in autism. Proc Natl Acad Sci.
541 2011 Oct 11;108(41):17076–81.
542 85. Blizinsky KD, Diaz-Castro B, Forrest MP, Schürmann B, Bach AP, Martin-de-Saavedra
543 MD, et al. Reversal of dendritic phenotypes in 16p11.2 microduplication mouse model

544 neurons by pharmacological targeting of a network hub. Proc Natl Acad Sci. 2016 Jul
545 26;113(30):8520–5.
546 86. Forsingdal A, Fejgin K, Nielsen V, Werge T, Nielsen J. 15q13.3 homozygous knockout
547 mouse model display epilepsy-, autism- and schizophrenia-related phenotypes. Transl
548 Psychiatry. 2016 Jul 26;6(7):e860.
549 87. Yin J, Chen W, Yang H, Xue M, Schaaf CP. Chrna7 deficient mice manifest no consistent
550 neuropsychiatric and behavioral phenotypes. Sci Rep. 2017 Dec 3;7(1):39941.
551 88. Smith CL, Blake JA, Kadin JA, Richardson JE, Bult CJ, Mouse Genome Database Group.
552 Mouse Genome Database (MGD)-2018: knowledgebase for the laboratory mouse. Nucleic
553 Acids Res. 2018 Jan 4;46(D1):D836–42.
554 89. Howe DG, Bradford YM, Conlin T, Eagle AE, Fashena D, Frazer K, et al. ZFIN, the
555 Zebrafish Model Organism Database: increased support for mutants and transgenics.
556 Nucleic Acids Res. 2012 Oct 15;41(D1):D854–60.
557 90. Gramates LS, Marygold SJ, Dos Santos G, Urbano JM, Antonazzo G, Matthews BB, et al.
558 FlyBase at 25: Looking to the future. Nucleic Acids Res. 2017 Jan 4;45(D1):D663–71.
559 91. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network
560 analysis. BMC Bioinformatics. 2008 Dec 29;9(1):559.
561 92. Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, Szafer A, et al. Transcriptional
562 landscape of the prenatal human brain. Nature. 2014 Apr 2;508(7495):199–206.
563 93. Golzio C, Katsanis N. Genetic architecture of reciprocal CNVs. Curr Opin Genet Dev.
564 2013 Jun;23(3):240–8.
565 94. Hu T, Yamagishi H, Maeda J, McAnally J, Yamagishi C, Srivastava D. Tbx1 regulates
566 fibroblast growth factors in the anterior heart field through a reinforcing autoregulatory
567 loop involving forkhead transcription factors. Development. 2004 Nov 1;131(21):5491–
568 502.
569 95. Vitelli F, Zhang Z, Huynh T, Sobotka A, Mupo A, Baldini A. Fgf8 expression in the Tbx1
570 domain causes skeletal abnormalities and modifies the aortic arch but not the outflow tract
571 phenotype of Tbx1 mutants. Dev Biol. 2006 Jul 15;295(2):559–70.
572 96. Satoh Y, Endo S, Ikeda T, Yamada K, Ito M, Kuroki M, et al. Extracellular signal-
573 regulated kinase 2 (ERK2) knockdown mice show deficits in long-term memory; ERK2
574 has a specific function in learning and memory. J Neurosci. 2007 Oct 3;27(40):10765–76.

575 97. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar:
576 improving access to variant interpretations and supporting evidence. Nucleic Acids Res.
577 2018 Jan 4;46(D1):D1062–7.
578 98. Turner TN, Yi Q, Krumm N, Huddleston J, Hoekzema K, F Stessman HA, et al. denovo-
579 db: a compendium of human de novo variants. Nucleic Acids Res. 2017 Jan
580 4;45(D1):D804–11.
581 99. Gonzalez-Mantilla AJ, Moreno-De-Luca A, Ledbetter DH, Martin CL. A cross-disorder
582 method to identify novel candidate genes for developmental brain disorders. JAMA
583 Psychiatry. 2016 Mar 1;73(3):275–83.
584 100. Ruderfer DM, Hamamsy T, Lek M, Karczewski KJ, Kavanagh D, Samocha KE, et al.
585 Patterns of genic intolerance of rare copy number variation in 59,898 human exomes. Nat
586 Genet. 2016 Oct 17;48(10):1107–11.
587 101. Andrews T, Honti F, Pfundt R, de Leeuw N, Hehir-Kwa J, Vulto-van Silfhout A, et al.
588 The clustering of functionally related genes contributes to CNV-mediated disease.
589 Genome Res. 2015 Jun;25(6):802–13.
590 102. Noh HJ, Ponting CP, Boulding HC, Meader S, Betancur C, Buxbaum JD, et al. Network
591 Topologies and Convergent Aetiologies Arising from Deletions and Duplications
592 Observed in Individuals with Autism. PLoS Genet. 2013 Jun 6;9(6):e1003523.
593 103. Grice SJ, Liu JL, Webber C. Synergistic Interactions between Drosophila Orthologues of
594 Genes Spanned by De Novo Human CNVs Support Multiple-Hit Models of Autism. PLoS
595 Genet. 2015 Mar 27;11(3):e1004998.
596 104. Mitchell KJ. The genetics of neurodevelopmental disease. Curr Opin Neurobiol. 2011
597 Feb;21(1):197–203.
598 105. Hoppman-Chaney N, Wain K, Seger P, Superneau D, Hodge J. Identification of single
599 gene deletions at 15q13.3: further evidence that *CHRNA7* causes the 15q13.3
600 microdeletion syndrome phenotype. Clin Genet. 2013 Apr;83(4):345–51.
601 106. Gokhale A, Freeman AA, Hartwig C, Bassell JL, Zlatic SA, Sapp C, et al. Systems
602 Analysis of the 22q11.2 Microdeletion Syndrome Converges on a Mitochondrial
603 Interactome Necessary for Synapse Function and Behavior. bioRxiv. Cold Spring Harbor
604 Laboratory; 2018 May 4;315143.
605 107. Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJ, Stephens K, et al.

606 GeneReviews®. GeneReviews®. University of Washington, Seattle; 2018.
607 108. Chatr-aryamontri A, Oughtred R, Boucher L, Rust J, Chang C, Kolas NK, et al. The
608 BioGRID interaction database: 2017 update. Nucleic Acids Res. 2017 Jan
609 4;45(D1):D369–79.
610 109. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A
611 Software Environment for Integrated Models of Biomolecular Interaction Networks.
612 Genome Res. 2003 Nov 1;13(11):2498–504.
613 110. Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis
614 with the PANTHER classification system. Nat Protoc. 2013 Jul 18;8(8):1551–66.
615

Figure 1

A List of pathogenic CNVs

CNV region	% De novo Del	% De novo Dup	CNV region	% De novo Del	% De novo Dup
1q21.1			2q23.1		
3q29			6p25		
10q23			**15q11.2**		
15q13.3			15q25.2		
16p13.11			**16p12.1**		
16p11.2			**16p11.2 distal**		
17p13.3			**17q12**		
17q23			19p13.12		
22q11.2			**22q11.2 distal**		
Wolf-Hirschhorn			Sotos/5q35		
Williams/7q11.23			Phelan-McDermid		
15q24			Rubinstein-Taybi		
Smith-Magenis/ Potocki-Lupski			Prader-Willi/ Angelmans		

Legend

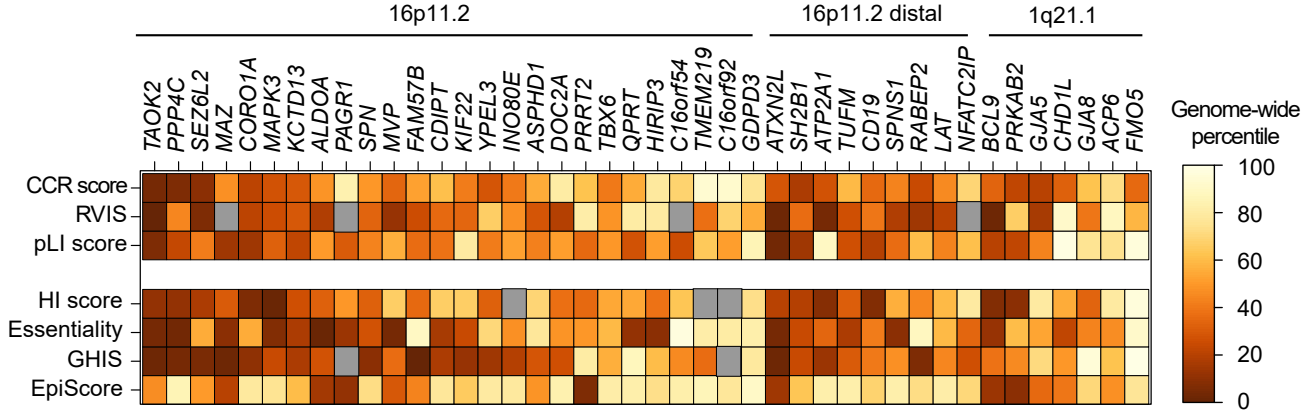
0-25%
25-50%
50-75%
75-100%

B Frequency of neuropsychiatric features

CNV disorder	ID/DD	Speech delay	Autism	Epilepsy	Micro/macrocephaly	ADHD	Schizophrenia	Behavioral features	Other psych. features
1q21 del									
3q29 del									
7q11.23 dup									
15q11.2q13.1 dup									
15q13.3 del									
16p11.2 del									
16p12.1 del									
17q12 del									
17q12 dup									
22q11.2 del									
Williams-Beuren									
Smith-Magenis									
Potocki-Lupski									
Sotos									

Figure 2

A



B

		Mouse phenotypes					Zebrafish phenotypes					*Drosophila* phenotypes				
		Behavioral	Embryonic	Growth/size	Lethality	Nervous system	Behavioral	Embryonic	Growth/size	Lethality	Nervous system	Behavior	Development	Lethality	Neuroanatomy	Neurophysiology
16p11.2 individual gene models	*SPN*															
	QPRT															
	C16orf54															
	KIF22															
	MAZ															
	PRRT2															
	C16orf53															
	MVP															
	CDIPT															
	SEZ6L2															
	ASPHD1															
	KCTD13															
	TMEM219															
	TAOK2															
	HIRIP3															
	INO80E															
	DOC2A															
	C16orf92															
	FAM57B															
	ALDOA															
	PPP4C															
	TBX6															
	YPEL3															
	GDPD3															
	MAPK3															
	CORO1A															
16p11.2 CNV models	Del(7Coro1a-Spn)1Dolm															
	Del(7Slx1b-Sept1)4Aam															
	Del(7Sult1a1-Spn)6Yah															
	Dp(7Slx1b-Sept1)5Aam															
	Dp(7Sult1a1-Spn)7Yah															

Figure 3

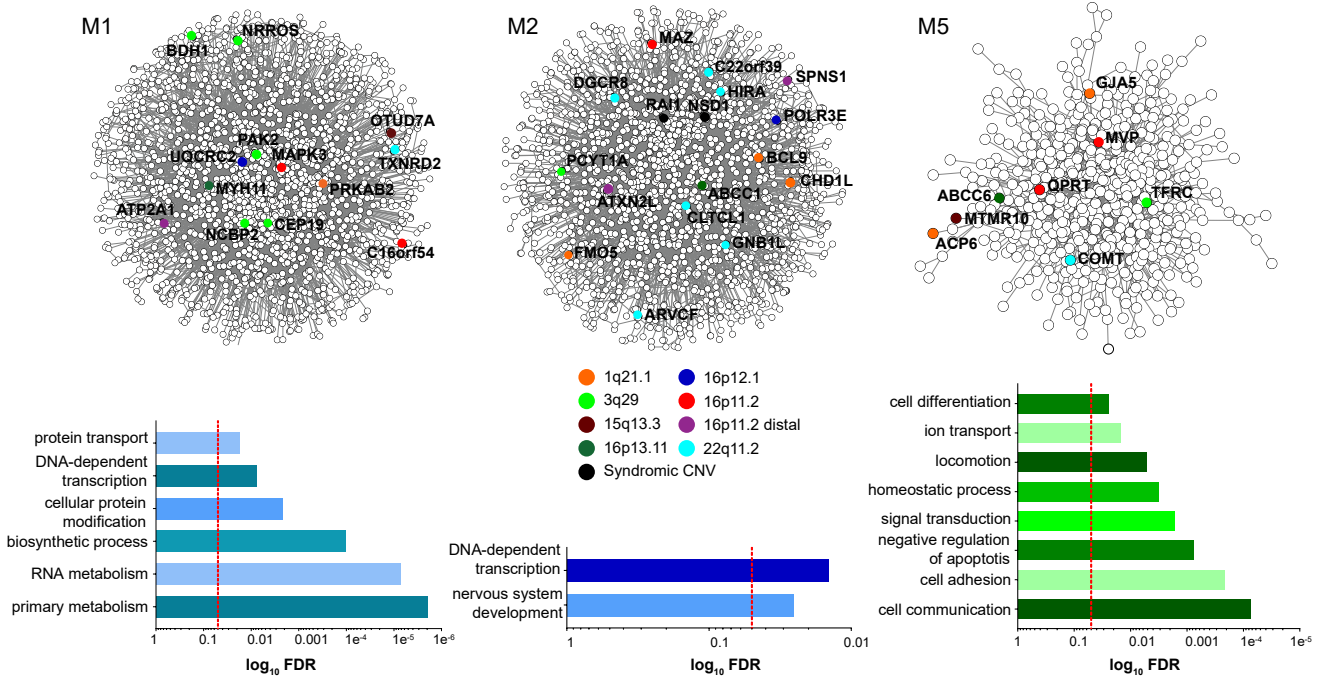
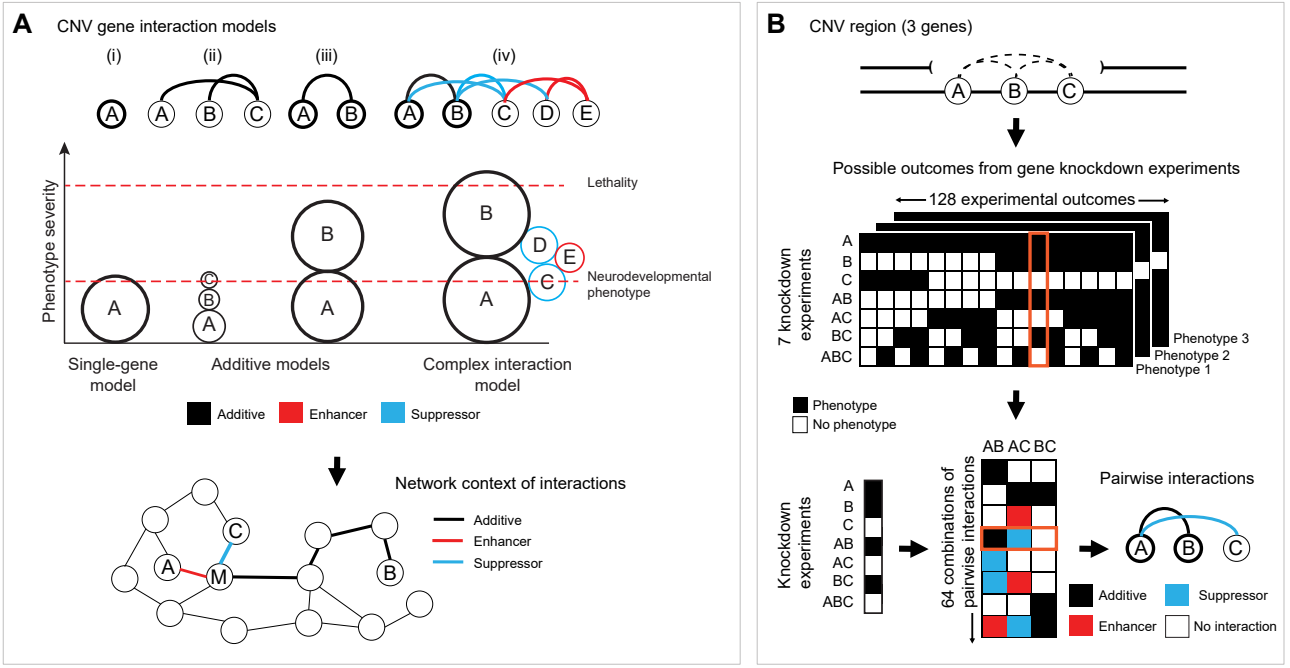

M1
BDH1
NRROS
GTUD7A
PAK2
UQCRC2
MAPK3
TXNRD2
MYH11
PRKAB2
ATP2A1
CEP19
NCBP2
C16orf54
M2
MAZ
C22orf39
SPNS1
DGCR8
HIRA
RAI1
NSD1
POLR3E
PCYT1A
BCL9
CHD1L
ATXN2L
ABCC1
CLTCL1
GNB1L
FMO5
ARVCF
M5
GJA5
MVP
QPRT
ABCC6
TFRC
MTMR10
ACP6
COMT
1q21.1
3q29
15q13.3
16p13.11
Syncromic CNV
16p12.1
16p11.2
16p11.2 distal
22q11.2
protein transport
DNA-dependent transcription
cellular protein modification
biosynthetic process
RNA metabolism
primary metabolism
log10 FDR
DNA-dependent transcription
nervous system development
log10 FDR
cell differentiation
ion transport
locomotion
homeostatic process
signal transduction
negative regulation of apoptotis
cell adhesion
cell communication
log10 FDR

Figure 4


A
CNV gene interaction models
(i)
(ii)
(iii)
(iv)
Phenotype severity
Lethality
Neurodevelopmental phenotype
Single-gene model
Additive models
Complex interaction model
Additive
Enhancer
Suppressor
Network context of interactions
Additive
Enhancer
Suppressor
B
CNV region (3 genes)
Possible outcomes from gene knockdown experiments
128 experimental outcomes
7 knockdown experiments
A
B
C
AB
AC
BC
ABC
Phenotype 3
Phenotype 2
Phenotype 1
Phenotype
No phenotype
Knockdown experiments
64 combinations of pairwise interactions
AB AC BC
Pairwise interactions
Additive
Suppressor
Enhancer
No interaction