Year : 2021

The impact of Copy Number Variants on brain morphometry

Modenato Claudia

LREN - Département de Neurosciences Cliniques

The impact of Copy Number Variants on brain morphometry

Thèse de doctorat en Neurosciences

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Claudia Modenato

Master en Neurosciences de l'Université de Genève

Jury

Prof. Jean-Pierre Hornung, Président
Prof. Bogdan Draganski, Directeur de thèse
Prof. Sébastien Jacquemont, Co-directeur de thèse (Centre de recherche CHU Sainte-Justine et
Université de Montréal, Canada)
Prof. Richard Delorme, Expert
Prof. Sven Cichon, Expert

Thèse n° 316

Lausanne 2021

Programme doctoral interuniversitaire en Neurosciences
des Universités de Lausanne et Genève

LREN - Département de Neurosciences Cliniques

The impact of Copy Number Variants on brain morphometry

Thèse de doctorat en Neurosciences

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Claudia Modenato

Master en Neurosciences de l'Université de Genève

Jury

Prof. Jean-Pierre Hornung, Président
Prof. Bogdan Draganski, Directeur de thèse
Prof. Sébastien Jacquemont, Co-directeur de thèse (Centre de recherche CHU Sainte-Justine et
Université de Montréal, Canada)
Prof. Richard Delorme, Expert
Prof. Sven Cichon, Expert

Thèse n° 316

Lausanne 2021

*Programme doctoral interuniversitaire en Neurosciences
des Universités de Lausanne et Genève*

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof.	Jean-Pierre Hornung
Directeur·trice de thèse	Monsieur	Prof.	Bogdan Draganski
Co-directeur·trice de thèse	Monsieur	Prof.	Sébastien Jacquemont
Expert·e·s	Monsieur	Prof.	Richard Delorme
	Monsieur	Prof.	Sven Cichon

le Conseil de Faculté autorise l'impression de la thèse de

Madame Claudia Modenato

Master en Neuroscience, Université de Genève, Suisse

intitulée

The impact of Copy Number Variants
on brain morphometry

Lausanne, le 26 novembre 2021

pour Le Doyen
de la Faculté de Biologie et de Médecine

Prof.   Jean-Pierre Hornung

Acknowledgments

I would like to express my deepest appreciation to all those who provided me the possibility to complete this thesis. Foremost, I would like to express my sincere gratitude to my thesis director Prof. Bogdan Draganski and co-director Sébastien Jacquemont for their support, motivation, enthusiasm, suggestions and encouragement. Under their guidance I've learned so much about science and research over the years, and I would like to thank them for this opportunity.

I would like to thank the SNSF for granting me with a doc.mobility and giving me the opportunity of spending one year abroad during my PhD to visit Sébastien Jacquemont lab in Montréal and Carrie Bearden lab in Los Angeles. A special thank goes also to Pr. Carrie Bearden and Paul Thompson for welcoming me in their teams during the 6 months I spent in Los Angeles.

Besides my supervisor I would like to thank all the fellow lab members who accompanied me during my journey. I want to thank the fellows of the Laboratoire de Recherche en Neuroimagerie (LREN) for their complicity and support during the first three years of my thesis, it was the best support I could have wished for and I miss it ever since. I want to thank all the fellows from Sébastien's, Carrie's and Paul's labs for welcoming me and helping me during my year abroad.

My sincere thanks also go to the professors Jean-Pierre Hornung, Richard Delorme and Sven Cichon who agreed to evaluate my thesis as jury president and external experts.

Finally I would like to thank my family, my friends and my partner for their endless love and support.

Abstract

Copy number variants (CNVs) are amongst the most frequent high-risk genetic factors implicated in neuropsychiatric disorders. Multiple rare CNVs confer risk for overlapping spectrums of neurodevelopmental symptoms and psychiatric conditions, including autism and schizophrenia. To date neuroimaging studies have typically been carried out one mutation at a time, showing that CNVs have large effects on global and regional brain morphometry. However it is unknown whether CNVs associated with similar risks for the same psychiatric condition also yield similar brain alterations.

The aim of this work was to better understand the effect of CNVs on brain structure. We characterized the volume, thickness, and surface of cortical and subcortical brain structures using the largest cross CNV T1-weighted MRI data to date (CNVs=720, controls=782).

First, we characterized distinct CNV-associated alteration profiles as well as shared latent gene-morphology dimensions across 8 CNVs on the cortex and whole brain morphometry measures. Second, we characterized the effects of 15 CNVs and 4 Polygenic Risk Scores on subcortical structures and compared the convergence with 6 idiopathic conditions.

We found global mirror effects for 6 genomic loci. The whole brain and cortical analysis showed mainly distinct brain patterns across CNVs, however, one-third of variance was explained by the first principal component (8 CNVs). The top regions contributing to latent dimension included the cingulate gyrus, insula, and supplementary motor cortex. Analysis of subcortical brain regions showed much larger effect sizes for CNVs (15 CNVs) than for polygenic risk scores and idiopathic conditions. CNV-associated subcortical brain alterations were correlated with their effect on cognition and on disease risk. Subcortical effect sizes normalized for gene intolerance decreased non-linearly from small oligogenic to larger multigenic CNVs.

Compared with CNVs effects on brain structure, the ones of idiopathic conditions and PRS were very weak and difficult to capture. We therefore conclude that CNVs are a powerful tool to investigate the effects of combinations of genes on brain structure and how these effects impact cognition and disease risk. Nevertheless, studies including many more variants are needed in order to understand brain mechanisms underlying psychiatric disorders.

Resumé

Les variations en nombre de copies (CNV) sont parmi les facteurs génétiques à haut risque, les plus fréquemment impliqués dans les troubles neuropsychiatriques. Plusieurs CNVs rares confèrent un risque pour différents symptômes neurodéveloppementaux et pour des troubles psychiatriques, y compris l'autisme et la schizophrénie. À ce jour, les études de neuroimagerie ont généralement été réalisées une mutation à la fois, montrant que les CNVs ont des effets importants sur la morphométrie cérébrale au niveau globale et régionale. Cependant, on ne sait pas si les CNVs qui sont associées à un risque pour les mêmes maladies psychiatriques, entraînent également des altérations cérébrales similaires.

Le but de ce travail était de mieux comprendre l'effet des CNV sur la structure du cerveau. Nous avons caractérisé le volume, l'épaisseur et la surface des structures cérébrales corticales et sous-corticales en utilisant la plus grande base de données d'IRM (T1-weighted) incluant plusieurs CNVs (CNV = 720, contrôles = 782).

Tout d'abord, nous avons caractérisé les profils d'altération spécifiques associés aux CNV, ainsi que des dimensions latentes partagées entre les 8 CNV sur le cortex et des mesures de morphométrie du cerveau entier (gene-morphometry dimensions). Deuxièmement, nous avons caractérisé les effets de 15 CNV et 4 scores de risque polygénique sur les structures sous-corticales et nous en avons comparé la convergence avec 6 conditions idiopathiques.

Nous avons trouvé des effets miroirs globaux pour 6 loci génomiques. L'ensemble du cerveau et l'analyse corticale ont montré principalement des schémas cérébraux distincts à travers les CNV, cependant, un tiers de la variance a été expliqué par la première composante principale (8 CNV). Les régions supérieures contribuant à la dimension latente comprenaient le gyrus cingulaire, l'insula et le cortex moteur supplémentaire. L'analyse des régions cérébrales sous-corticales a montré des tailles d'effet beaucoup plus importantes pour les CNV (15 CNV) que pour les scores de risque polygénique ou les maladies idiopathiques. Les altérations cérébrales

sous-corticales associées au CNV étaient corrélées à leur effet sur la cognition et sur le risque de maladie. Les tailles d'effet sous-corticales normalisées pour l'intolérance génique diminuent de manière non linéaire à partir des petits CNVs oligogéniques jusqu'aux plus grandes CNVs multigéniques.

Comparés aux effets des CNV sur la structure du cerveau, ceux des maladies idiopathiques et du PRS étaient très faibles et difficiles à capturer. Nous concluons donc que les CNV sont un outil puissant pour étudier les effets des gènes sur la structure du cerveau et de comment ces effets ont un impact sur la cognition et le risque de maladie. Néanmoins, des études comprenant de nombreuses autres variantes sont nécessaires afin de comprendre les mécanismes cérébraux sous-jacents aux troubles psychiatriques.

Table of Contents

Abbreviations and definitions table

Additive genetic effects	The combined effect of several genomic variants on a quantitative trait equals the sum of their individual effects. An alternative is multiplicative interaction effects (epistasis). For complex traits such as cognition and behavior, the genetic contribution to phenotypic variance has mainly been attributed to additive effects. Under the assumption of an additive model, the variance of a trait is expected to be the same in a group of CNV carriers and in the general population (e.g. IQ variance is the same in 16p11.2 deletion carriers and in unselected populations (2)). This is in line with the additional familial and genetic factors that have previously been associated with the variance of cognitive and behavioral phenotypes in CNV carriers (2–6).
ASD	Autism spectrum disorder
Breakpoints (BP), chromosomal	A specific site of breakage, usually associated with a recurrent chromosomal abnormality. For some CNVs, several low copy repeats (LCRs) in the region allow for multiple such BPs.
CNV	Copy number variant
Cortical thickness	Measures of local thickness of the cerebral cortex (or subcortical structures) can be obtained from widely used (and histologically validated) MRI processing techniques that measure the distance between the pial surface and gray/white matter boundary beneath the cortical ribbon.
de novo	A genomic variation that occurs spontaneously in the offspring and thus is not inherited from the parents.
Gene dosage effect	Effects related to a change in the number of genomic copies (deletions or duplications).
Gene dose response	The effect of altering the amount of genetic material in a region/the magnitude of the response of an organism to changes in gene presence.
Genetic heterogeneity	The same or similar phenotypes caused by different genetic mechanisms.
Genetics-first approach	A strategy used in epidemiological studies to associate specific genotypes (such as a specific CNV) with apparent clinical phenotypes of a complex disease or trait. Also called "genotype-first."
GM	Grey matter
Idiopathic	Any disease or condition for which the cause is unknown.
Mirror gene dosage response	Monotonic relationship between the number of genomic copies at a given locus and a quantitative trait. Negative/inverse mirror dose-response describes the negative correlation between the number of genomic copies at a genomic locus and a given trait, (e.g. larger volume in deletion carriers, smaller volume in duplication carriers). Likewise, positive mirror dose-response means smaller volume in deletion carriers, larger volume in duplication carriers.
MRI	magnetic resonance imaging
Noncarrier	In the context of CNVs, this is usually defined as an individual who does not carry the particular CNV being studied.
NPD	Neurodevelopmental disorder
Pathogenic CNV	A CNV with large effect size on neurodevelopmental traits. Such variants are also

	referred to as clinically significant (1).
Penetrance	The proportion of individuals with a genomic variant who present symptoms. The penetrance reflects the effect size of a genomic variant under the assumption of additive effects. CNVs with high effect sizes show high penetrance (e.g. DS decreases IQ by 3.5 SD), because symptoms will still be observed irrespective of genetic and environmental backgrounds. On the other hand, small effect sizes (low penetrance) CNVs (e.g. 15q11.2 deletion) will often be asymptomatic unless genetic and environmental background are conducive.
Pleiotropy	The phenomenon whereby one allele (or a pair of alleles) influences multiple, independent phenotypes.
Polygenic trait	A phenotype that is influenced by multiple genetic variants at different genomic sites.
Rare CNV	Typically defined as a CNV with <1% frequency in the population.
Reciprocal CNVs	Deletions and duplications that occur at the same locus, usually flanked by LCRs.
Recurrent CNVs	CNVs that occur as spontaneous de novo events at the same sites in the genome repeatedly in unrelated individuals due to the presence of flanking low copy repeats, or LCRs) (Hastings, Lupski, Rosenberg, & Ira, 2009). In other words, they occur de novo in the first individual, and hence are not observed in the CNV carrier's parents but are potentially inherited in subsequent generations.
SCZ	schizophrenia
Single nucleotide polymorphism (SNP)	The substitution of a single base (A, T, C, or G) for another base at a specific genetic location that occurs in at least 1% of the population. A SNP may or may not have functional consequences on gene expression.
Surface area	As with cortical thickness, semi-automated MRI processing techniques can be used to estimate the surface area of the cortical mantle and have been widely used to study brain development in healthy and clinical populations. MRI-derived measures of surface area are thought to be driven by the number of cortical columns and are likely under differential genetic control than cortical thickness.
TIV	Total intracranial volume
WM	White matter

Introduction

1. Bottom up versus top down approaches

1.1. Neuroscience research in psychiatric disorders - the search for biomarkers

The concept of psychiatric disorders has seen a remarkable evolution since his appearance more than a century ago. This evolution shaped the classification of mental disorders into clinically defined categories or psychiatric diagnosis. Psychiatric and later neuroscience research, was therefore defined as the study of patients characterized by the same diagnosis. However, research based on nosological approaches has so far failed in understanding what goes on in the brain of people with psychiatric disorders and in understanding the biological foundations of mental illnesses. Psychiatric diagnoses are based on behavioral defined criteria and therefore do not necessarily group patients with the same underlying biological conditions. In fact 1) patients classified under the same diagnosis can present different sets of symptoms, 2) overlap of symptoms is observed across diagnosis, 3) patients can transition from one diagnosis to another during their lifetime. These limitations likely induce heterogeneity in patient groups and could explain why so far studies attempting to identify brain biomarkers for psychiatric disorders (using structural MRI or other imaging modalities) have reported very small effect sizes.

More recently research has shifted from the nosological approach ("top-down") to an etiological or pathophysiological approach ("bottom-up"). In bottom-up approaches subjects are grouped based on biological characteristics instead of clinical diagnosis, to reduce biological heterogeneity within the patient group. By reducing biological heterogeneity,

bottom-up studies are expected to find much larger effect-sizes compared to top-down approaches. Another initiative, that proposes an alternative approach to top down methodologies in research studies, is the Research Domain Criteria (RDoC) Framework. Developed by the National Institute for Mental Health, the RDoC introduces a dimensional perspective for psychiatric research based on underlying neurobiological and behavioral mechanisms [1] (Figure 1).
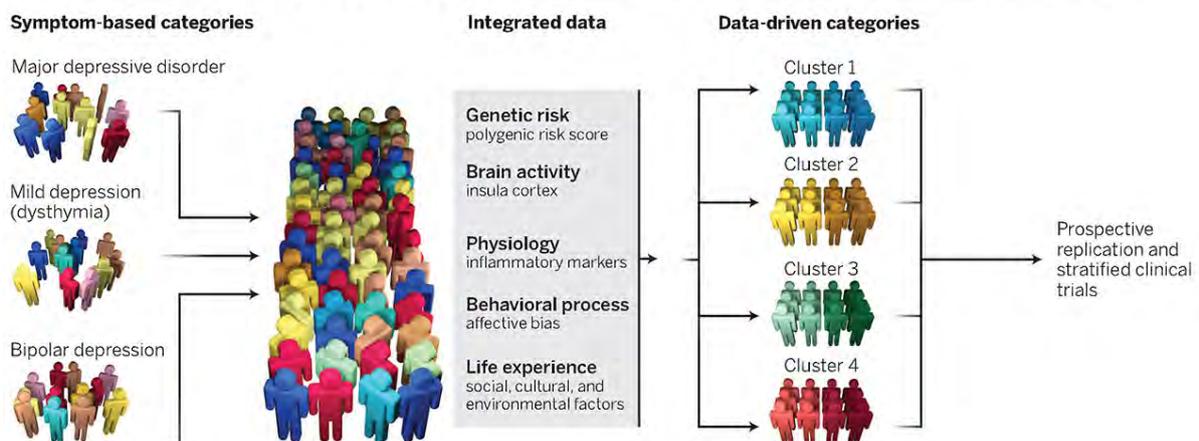
Figure 1: *from Insel and Cuthbert (2015). Illustration: V. Altounian and C. Smith/Science*

1.2. Neuroimaging: a tool to investigate brain mechanisms

Magnetic resonance imaging (MRI) is a non-invasive technique widely used in clinical and research environments to image anatomical and physiological features within the human body. As the name suggests, MRI exploits the properties of nuclear magnetic resonance of protons (hydrogen atoms), which are present in large amounts in water molecules inside

tissues and fluids of the human body. MRI provides an in vivo estimation of the structural organization in the human brain. Among the many methods used to investigate MRI images is voxel-based morphometry (VBM) [2] (a voxel is a cubic element of brain volume, usually of the size of 1 mm^3). The method evaluates macroscopic gray matter asymmetries with high regional specificity by using a regional smoothing kernel to quantify and compare between subjects the concentration of grey matter across voxels in the brain. Alternatively surface-based methods (SBM) are used to investigate cortical properties such as cortical thickness and surface area [3,4]. The difference between SBM and VBM approaches is that the analysis is performed in a 3-D voxel grid for VBM, while in SBM is performed at the nodes of a three-dimensional polygonal mesh [4].

Hardware and computational methods in neuroimaging have been rapidly developing over the past three decades, and this method remains a promising tool to investigate the brain and to define disease biomarkers that distinguish patients from controls, and also patients with different diagnoses. However so far neuroimaging has shown limited clinical utility and has not been linked to improvement of psychiatric diagnosis. The shifting of paradigm within psychiatric research towards bottom-up approaches holds hope for the identification of robust neuroimaging biomarkers.

1.3. Genetic first approaches

Bottom-up approaches in clinical neuroscience include genetic first approaches, where subjects are stratified based on the presence of a genetic variant.

The rationale for studying a genetically homogeneous group comes from the underlying hypothesis that psychiatric diseases are the endpoint of multiple biological pathways which are caused by different genetic predispositions and environmental factors [5]. The fact that the

same brain abnormalities are identified in patients with a given psychiatric diagnosis, as well as in individuals at high-risk for developing a given psychiatric disorder (i.e. family members), confirms the hypothesis that brain abnormalities are associated with genetic risk for the disease and not with disease symptoms [5]. Bottom up approaches have identified brain alterations regardless of the symptomatology of the participant, suggesting that alterations precede symptom development. This implies that if genetics or neuroscience can detect the risk for a specific disease before the appearance of symptoms, preventive intervention or monitoring could be foreseen [6].

Effect-sizes on brain alterations reported in studies using top-down approaches in psychiatric disorders are very small and therefore they are not suitable biomarkers (i.e cortical alterations in SCZ and ASD have an absolute Cohen's d $<=0.20$ [7,8]). While on the contrary genetic-first approaches have reported much larger effect-sizes on brain alterations (i.e. absolute Cohen's d ranging from 0.5 to 1.3 in 16p11.2 proximal and 22q11.2 [9–12]) and hold the promise of identifying brain mechanisms associated with specific biological risk [13–15]. Overall, the study of brain structural alterations associated with specific genetic conditions represent a powerful tool to study the relationship between genes, brain mechanisms and behavior.

2. Copy number variants

2.1. Definition of copy number variants

For genetic-first approaches, copy number variants (CNVs) have gained more and more interest since many of them occur recurrently in unrelated individuals allowing to establish cohorts of individuals with the exact same genomic variant irrespective of their clinical symptomatology. These cohorts are particularly well suited for the investigation of brain

mechanisms underlying the association between genetic risk and neuropsychiatric disorders [16,17].

Copy number variants (CNVs) are structural variations in the number of copies of large DNA segments that can range from 1000 base pairs, to several megabases and up to an entire chromosome (the case of monosomies and trisomies). CNVs can include coding and noncoding DNA regions and many have been associated with a wide range of diseases, yet many others are benign and constitute normal genetic variation within the population [18–20].

Discovered in 1959, Trisomy 21 was the first example of altered gene dosage across an entire chromosome. The discovery was followed in the 1980s-90s by discoveries of CNVs associated with clinically defined syndromes such as the 7q11.23 (Williams-Beuren syndrome [21]) and the 22q11.2 deletions (velocardiofacial syndrome [22]). The rate of discovery accelerated by several orders of magnitude around 2010 with the advent of high-throughput chromosomal microarrays, which allowed for genome-wide interrogation of CNVs. Approximately 15%, 9%, and 2-8% of individuals referred to the clinic for motor delay and intellectual disabilities [23], autistic spectrum disorder (ASD) [24], and schizophrenia (SCZ) [25,26], respectively, carry a pathogenic CNV.

There are two major classes of CNVs: 1) recurrent CNVs, which are deletions or duplications of chromosomal intervals defined by specific breakpoints (BP) leading to identical CNVs containing the same genes in non-related individuals [20] 2) non-recurrent CNVs occurring at random positions in the genome, thus individually extremely rare or even unique. This work focuses on recurrent CNVs. Pathogenicity of CNVs is linked to several factors such as their inheritance (de novo are usually more pathogenic than inherited), their size (larger CNVs can encompass more genes and eventually more gene-dosage sensitive genes), the type (deletions are more pathogenic than duplications) and gene content (CNVs may include a few genes but

some of the more deleterious ones may encompass dozens of genes including genes previously associated with severe neurodevelopmental disorders) [20].

2.2. Copy number variants and risk for psychiatric disorders

CNV carriers are at elevated risk for malformations, neurodevelopmental disorders [27], and other medical conditions [28–30]. In addition it has been recently shown that CNVs have a pleiotropic impact on major disease categories including psychiatric diseases [31].

In comparison to the highest odd ratios observed for individual Single Nucleotide Polymorphisms (SNP) in ASD, SCZ, attention deficit hyperactivity disorder (ADHD), or major depression disorder (OR=~1.05-1.25) [31–34], deleterious recurrent CNVs are associated with substantially higher risk (OR often >10) [34]. Ultra rare non-recurrent CNVs are distributed across the genome and can not be studied individually. Nevertheless, burden analyses have shown that, as a group, they are overrepresented in SCZ [34] and ASD, and that they decrease intelligence irrespectively of a neuropsychiatric disorder diagnosis [36]. Models estimate that the vast majority of 1 megabase deletions or duplications containing coding elements increase ASD risk [36] and that, when deleted, approximately 10000 genes negatively affect intelligence [37]. Part of this work is focused on the most frequent large effect-size genetic risk-factors for psychiatric conditions: eight CNVs at four genomic loci: 1q21.1, 16p11.2, 22q11.2 and 15q11.2, which affect dosage of 12, 29, 60 and 4 genes respectively [38–40].

2.2.1. 1q21.1

Deletions and duplication at the 1q21.1 BP4-BP5 locus (146.6-147.5 Mb, hg19) affect the dosage of on average 12 genes and are among the most frequently reported in neurodevelopmental disorders [41]. The estimated population frequency for 1q21.1 deletions is of 1/3500, while for duplications is of 1/2300 [41]. Deletions at the 1q21.1 locus are associated with primarily with SCZ (OR=6) and then with ASD (OR=1.56), while duplications are associated primarily with ASD (OR=8) and then with SCZ (OR=3) [34,38,42–44].

2.2.2. 16p11.2

Deletions and duplication at the 16p11.2 BP4-BP5 locus (29.6–30.3 Mb, hg19) affect the dosage of on average 29 genes and are also often reported in a wide range of neurodevelopmental conditions. 16p11.2 deletions and duplications have a prevalence of 1 over 2000 individuals each [45]. Similar to 1q21.1 carriers, in 16p11.2 carriers the risk for autism and schizophrenia in deletions and duplications follows an opposite pattern, although the pattern is inverted compared to 1q21.1. In fact 16p11.2 deletion carriers are associated with ASD (OR=9.5) while duplication carriers are associated with ASD and SCZ (OR=11.8 and 12) [35,47–50]. Moreover, 16p11.2 deletions and duplications are associated with a mirror effect on body mass index [50].

2.2.3. 22q11.2

Deletions and duplication at the 22q11.2 locus (1.5 to 3 Mb) affect the dosage of on average 60 genes. 22q11.2 deletions are also known as the DiGeorge or Velocardiofacial syndrome, and was the first CNV known for its association with SCZ. Among the CNVs included in this research, 22q11.2 deletions are associated with the highest risk for SCZ (OR=92) and ASD

(OR=32.4). While 22q11.2 duplications have an increased risk for ASD (OR=3.3) but interestingly a decreased risk for SCZ (OR=0.15) [35,49].

2.2.4. 15q11.2

Deletions and duplication at the 15q11.2 locus affect the dosage of on average 4 genes and have a relatively frequent prevalence of approximately 0.36% in the general population [41]. 15q11.2 deletions are associated with SCZ (OR=1.3), while 15q11.2 duplications have been associated with ASD (OR=1.8) [35,36]. No detectable gene dosage effect on brain volume has been identified for the 15q11.2 locus.

3. The effect of copy number variants on brain anatomy

Many CNVs have been shown to have an effect on global brain metrics. Large effects have been observed for deletions and duplications at the 16p11. 2, 22q11.2 and 1q21.1 loci and duplication at the 1q21.1 locus, while smaller effects have been observed for duplications at the 16p11.2 locus. For the 22q11.2 locus deletion carriers are associated with smaller brain size compared to controls [51]. Deletions and duplications at the 1q21.1 and at the 16p11.2 loci have reciprocal opposing effects on head size: smaller in deletion carriers and larger in duplication carriers and vice versa [42,45,52,53].

An increasing number of brain morphometry studies in CNV carriers have reported robust and large effects on brain structures that partially overlap with brain alterations previously observed in idiopathic psychiatric disorders [13,54,55]. Robust effects on total and regional brain volumes, cortical thickness (CT) and surface area (SA), have been reported in 22q11.2 [52,57,58], 16p11.2 BP4-5 [53,58,59], and 15q11.2 CNVs [61–64]. Opposing effects on global and-or regional

18

brain volumes between deletions and duplications were observed for 16p11.2 [53], 22q11.2 [51], 1q21.1 [52] and 15q11.2 [61] loci (hereafter referred to as "mirror effects").

Finally, most of the effects are observed irrespectively of psychiatric diagnoses and of symptoms [57], suggesting that the final clinical outcome may result from the effect of CNVs and from additional factors.

Only a few neuroimaging studies have simultaneously investigated brain alterations across multiple genomic variants. An investigation of 49 unaffected carriers of SZ-associated CNVs across 5 genomic loci in the UK biobank showed smaller volumes of the thalamus, hippocampus, and nucleus accumbens [64]. Functional connectivity similarities have also been demonstrated between 16p11.2 and 22q11.2 deletions as well as with idiopathic ASD and SZ [65]. A recent study on 6 different CNVs suggests that patterns of neuroimaging alterations are relatively distinct between them [66].

Otherwise most CNVs neuroimaging studies were conducted one mutation at a time, thus providing a complex catalogue of brain anatomy patterns linked to different genomic loci.

4.   Polygenic risk score - new tool in genetic-first studies

Recently Polygenic Risk Score (PRS) has been developed as a new genetic measure to investigate the risk for specific traits and diseases. PRS is the sum of an individual's risk alleles, weighted by risk allele effect sizes derived from genome-wide association study (GWAS) data [67]. While CNVs typically range from oligogenic (2-10 genes) to multigenic variants (30 to 58 genes), PRS integrates thousands of SNPs with very small individual effect sizes.

CNVs associated brain alterations have been reported consistently at the global and regional level [14], while the effect of PRS on brain structure has been investigated in a few studies which reported inconsistent results [69–71].

In our second paper we compare the effects sizes of PRS, CNVs and idiopathic conditions on brain anatomy in order to have a better understanding of the relationship between polygenicity and brain alterations.

5. Aims of the thesis

The aim of this thesis was to investigate and better understand the effect of CNVs on brain structure.

Knowledge gap: CNV neuroimaging studies are an emerging field. While effects on MRI derived measures are robust, all of the previously published CNV neuroimaging studies were conducted one CNV at a time. Therefore, it has been difficult to identify general principles or potential shared effects of neuropsychiatric CNVs on brain anatomy.

Specific aims (Figure 2):

Aim 1: We investigated the effect on structural brain imaging in the largest multi-site neuroimaging dataset of CNV carriers. We aimed at characterizing the amount of overlap of neuroanatomical alterations across eight CNVs at four genomic loci.

> 1a) First, we wanted to characterize the brain morphometry alterations associated with 8 CNVs using three morphometric measures: volume, thickness and surface area covering the whole brain and cortex.
>
> 1b) Second, we wanted to quantify the amount of distinct and shared variation associated with the morphometric alterations of eight CNVs.

1c) Finally, we wanted to use a data driven approach to jointly analyse genetic and morphometric brain data and to identify latent 'gene-morphometry dimensions'

Aim 2: We investigated the relationship between polygenicity and subcortical brain alterations or mechanisms involved in psychiatric conditions. To do so we looked at subcortical brain effect sizes across oligo- and multigenic CNVs (12 neuropsychiatric CNV, 3 non-psychiatric CNVs) as well as across polygenic risk scores (4 PRS scores) and idiopathic conditions (6 idiopathic conditions).

2a) Compare subcortical brain effect sizes across oligo- and multigenic CNVs as well as across polygenic risk scores and idiopathic conditions.

2b) Characterize potential subcortical brain alteration overlap between genetic risk factors and idiopathic conditions
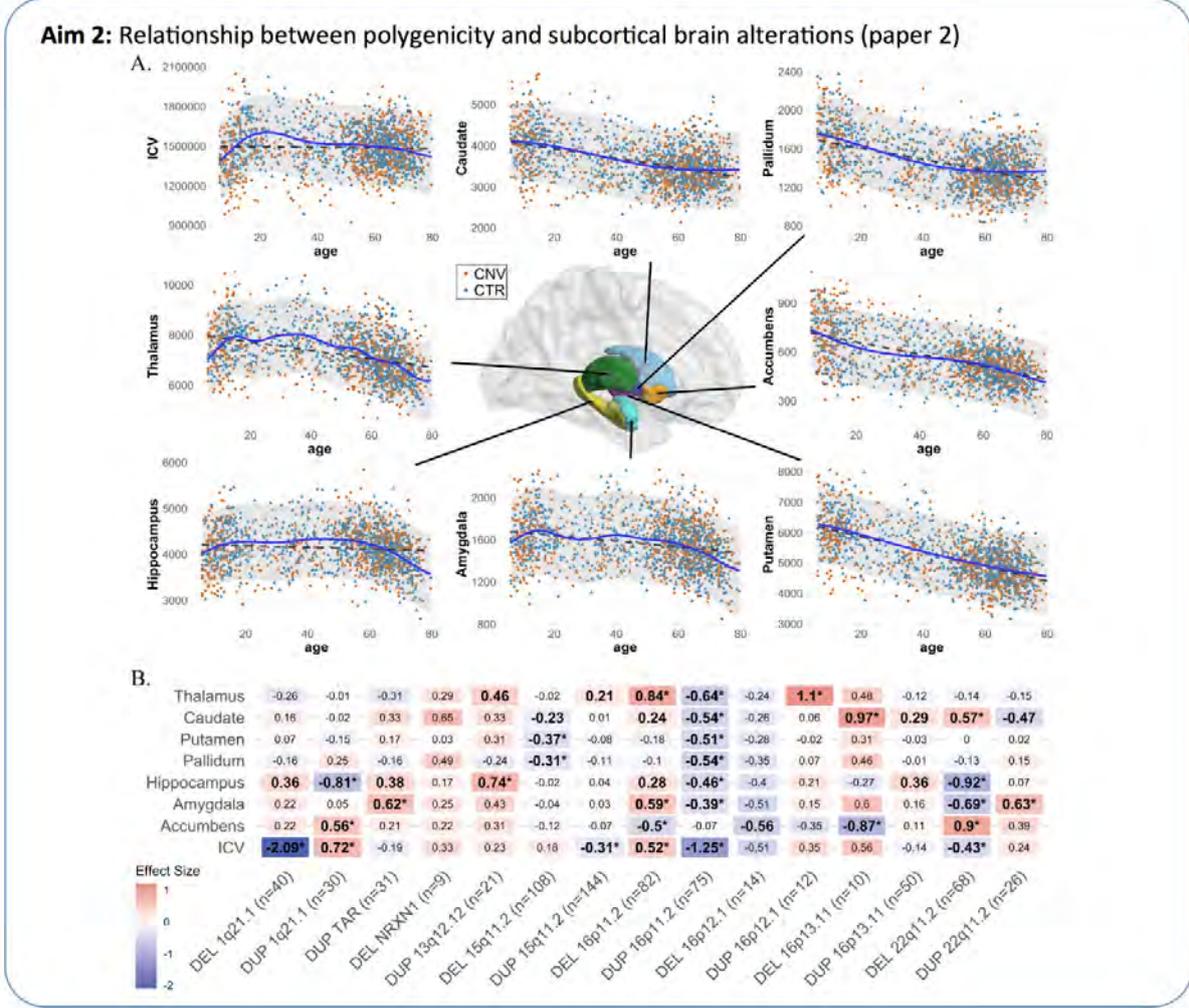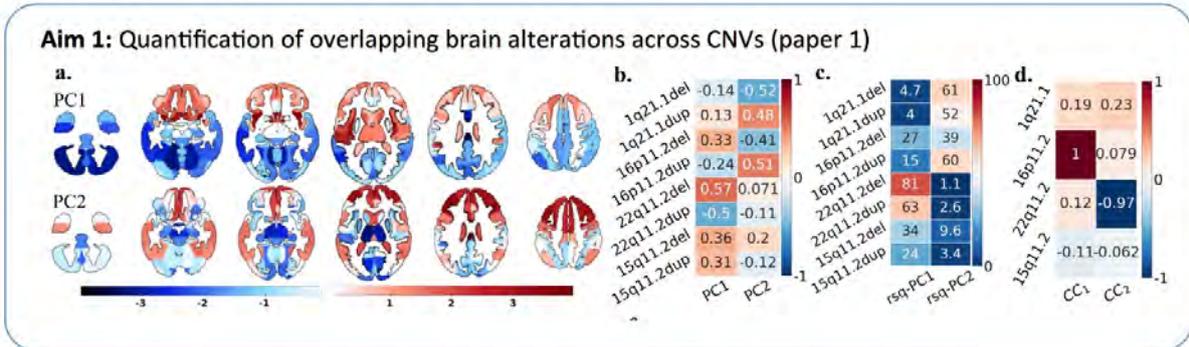
Figure 2: *Visual representation of thesis aims with snapchat of each resulting paper.*

Results

6. Effects of eight neuropsychiatric copy number variants on human brain structure

<u>Contribution:</u> designed the study, acquired and analysed imaging data and drafted the manuscript.

Modenato Claudia1* MSc, Kuldeep Kumar2* PhD, Moreau Clara2 PhD, Martin-Brevet Sandra1 PhD, Huguet Guillaume2 PhD, Schramm Catherine2 PhD, Jean-Louis Martineau2 MSc, Martin Charles-Olivier2 PhD, Younis Nadine2 BSc, Tamer Petra2 BSc, Douard Elise2 MSc, Thébault-Dagher Fanny2 BSc, Côté V.2 BSc, Charlebois A.R.2 MSc, Deguire F.2 MSc, Maillard Anne M.3 PhD, Rodriguez-Herreros Borja3 PhD, Pain Aurèlie3 MSc, Richetin Sonia3 MSc, 16p11.2 European Consortium, Simons Searchlight Consortium, Melie-Garcia Lester4 PhD, Kushan Leila MSc5, Silva Ana I.6,7 PhD, van den Bree Marianne B.M.7,8,9 PhD, Linden David E.J.6,7,9 MD, Owen Michael J.7,8 MD, PhD, Hall Jeremy7,8,9 MD, PhD, Lippé Sarah2 PhD, Chakravarty Mallar10 PhD, Bzdok Danilo11,12 MD, PhD, Bearden Carrie E. PhD5, Draganski Bogdan MD**1, 13 and Jacquemont Sébastien MD**2

* shared 1st authorship

** shared senior authorship

Abstract

Many Copy Number Variants (CNVs) confer risk for the same range of neurodevelopmental symptoms and psychiatric conditions including autism and schizophrenia. Yet, to date neuroimaging studies have typically been carried out one mutation at a time, showing that CNVs have large effects on brain anatomy. Here, we aimed to characterize and quantify the distinct brain morphometry effects and latent dimensions across 8 neuropsychiatric CNVs.

We analyzed T1-weighted MRI data from clinically and non-clinically ascertained CNV carriers (deletion/duplication) at the 1q21.1 (n=39/28), 16p11.2 (n=87/78), 22q11.2 (n=75/30), and 15q11.2 (n=72/76) loci as well as 1296 non-carriers (controls). Case control contrasts of all examined genomic loci demonstrated effects on brain anatomy, with deletions and duplications showing mirror effects at the global and regional levels. Although CNVs mainly showed distinct brain patterns, principal component analysis (PCA) loaded subsets of CNVs on two latent brain dimensions, which explained 32 and 29% of the variance of the 8 Cohen's d maps. The cingulate gyrus, insula, supplementary motor cortex, and cerebellum were identified by PCA and multiview pattern learning as top regions contributing to latent dimension shared across subsets of CNVs.

The large proportion of distinct CNV effects on brain morphology may explain the small neuroimaging effect sizes reported in polygenic psychiatric conditions. Nevertheless, latent gene brain morphology dimensions will help subgroup the rapidly expanding landscape of neuropsychiatric variants and dissect the heterogeneity of idiopathic conditions.

7. Effects of rare CNVs and common variants on subcortical structure

Contribution: designed the study, acquired and analysed imaging data and drafted the manuscript.

Abstract

Rare Copy number variants (CNVs) and Polygenic risk scores are well known genetic risks for neurodevelopmental psychiatric disorders (NPDs) such as autism and schizophrenia. CNVs can either include a few genes (oligo-) or several dozen genes (multigenic), whereas polygenic risk scores (PRS) are the sum of thousands of individual common variants with very small effects on risk for psychiatric diseases.

CNVs associated brain alterations have been reported consistently at the global and regional level, while the effect of PRS on brain structure has been investigated in a few studies which reported inconsistent results. Additionally most neuroimaging studies have extensively examined regional volumes, and only few studies have investigated whether CNVs differentially affect subregions or subfields of subcortical structures. Effects of CNVs on cognition and risk for disease can accurately be predicted using the sum of intolerant genes encompassed in the CNV, but whether this applies to structural brain alterations has not yet been studied.

Aim : In this study, we aimed to 1) compare effect sizes of oligo, multigenic and polygenic risk on subcortical structures; 2) investigate whether there are subcortical structures vulnerable to a broad spectrum of rare and common genetic risk.

Methods: To do so we characterized volume, thickness and surface of subcortical brain structures using T1-weighted MRI data across oligo- and multigenic CNVs (12

neuropsychiatric CNV, 3 non-psychiatric CNVs -1502 carriers and controls-) as well as across polygenic risk scores (4 PRS scores in 35000 individuals of the UKBB) and idiopathic conditions (6 idiopathic conditions).

Results: Most CNV showed significant alterations in morphometric measures of at least one subcortical structure. We observed significant correlations between CNVs effect sizes on SubCortical morphometry and on cognition. ES normalized for intolerance decreased dramatically from small oligogenic to larger multigenic CNVs. Moreover, ES of CNVs were up to an order of magnitude larger than those observed for psychiatric PRS and 2-fold larger than idiopathic conditions. PCA identified the same latent dimension -defined by opposing effects on basal ganglia and limbic structures- across CNVs and across idiopathic conditions.

Conclusion: CNVs broadly affect structural morphometry of subcortical structures. Moreover these effects were much higher than those observed for PRS and idiopathic conditions, and were correlated with CNVs effects on cognition and disease risk. CNVs prove to be a powerful tool to investigate how gene combinations affect neuroanatomy. Although many studies including many more variants are needed in order to dissect the heterogeneity of brain mechanisms underlying NPDs.

Introduction

Neurodevelopmental psychiatric disorders (NPDs), such as Autism (ASD) and Schizophrenia (SCZ), have complex genetic architecture involving common and rare variants such as Copy Number Variants (CNVs). CNVs are genomic deletions and duplications larger than 1000 base pairs. CNVs previously associated with NPDs typically range from oligogenic (eg. 2-10 genes) to multigenic variants (eg. 30 to 58 genes), Therefore they offer the opportunity to study the impact of defined sets of genes on neurobiological pathways contributing to NPDs.

Polygenic Risk Scores (PRS) have recently emerged as a method to study the aggregate effects of thousands of common risk variants for specific traits and diseases. A PRS is the sum of thousands of risk alleles, weighted by risk allele effect sizes derived from genome-wide association study (GWAS) data [68]. The latter risks are similar to some of the smaller oligogenic CNVs such as 15q11.2 (OR SZ= 1.5) or 16p13.11 (OR SZ= 1.5) [35].

CNVs' associated brain alterations have been reported consistently at the global and regional level [14], while the effect of PRS on brain structure has been investigated in a few studies which reported inconsistent results [69–71].

Subcortical regions have a critical role in cognitive, affective, and social functions in humans, and the structural and functional abnormalities of these regions have been associated with various psychiatric disorders. While most neuroimaging studies have extensively examined regional volumes, few studies have investigated whether CNVs differentially affect subregions or subfields of subcortical structures [13]. High-resolution shape analysis has been used to map fine grained subcortical alterations in schizophrenia, autism, bipolar disorder, offering insights into differential impact on subcompartments or subfields with known structural and functional connectivity patterns. Moreover shape measurements seem to capture additional effects which are either predictive of disease status, or related to genetic variants, and which are not detectable from volumetric measurements [72]. Large neuroimaging cohorts have shown that major psychiatric conditions, including schizophrenia (SCZ [73]), major depressive disorder (MDD [74]), bipolar disorder (BD [75]), obsessive-compulsive disorder (OCD [76]), autism spectrum disorder (ASD [8]) and attention deficit hyperactivity disorder (ADHD [77]) are associated with alterations of subcortical structures predominantly affecting hippocampus and the amygdala. Studies have also shown some level of correlation between brain structural alterations across these conditions [78]. However, idiopathic conditions are associated with small effect sizes on brain morphometry

which are discordant with the severity of these conditions, and this issue has been interpreted as a consequence of genetic and mechanistic heterogeneity underlying these highly polygenic conditions [13,14].

There is a growing body of literature on the subcortical alterations associated with genetic risk for NPDs. The effect of CNVs on subcortical structures has been characterized for several loci including 1q21.1, 16p11.2 BP4-5, 16p11.2 BP1-3, 15q11.2, 15q11-q13 and 22q11.2 showing that subcortical structures are broadly affected without a clear pattern shared across variants [14]. Our previous work on structural brain alterations in 8 CNVs demonstrated distinct CNV-associated alteration profiles as well as shared latent gene-morphology dimensions relevant to subsets of CNVs [79]. We hypothesized that the large proportion of distinct CNV effects on brain morphology may explain the small neuroimaging effect sizes reported in polygenic psychiatric conditions.

To date, effect size of common variants and CNVs on cognition and risk for disease have been estimated using additive models, where the combined effect of several genomic variants on a quantitative trait equals the sum of their individual effect. We previously demonstrated that the effects of CNVs on cognition and risk for disease can be accurately predicted using the sum of genes -weighted by their intolerance to haploinsufficiency- encompassed in the CNV. In addition, previous studies have demonstrated that PRS SCZ and ASD are associated with an increased risk for their corresponding condition with ORs ranging from 1.7 to 32 respectively.

Knowledge gap:

Studies have not investigated simultaneously the effect of several multigenic and polygenic psychiatric risk factors on subcortical brain alterations.

Overarching Aim :

28

Investigate the relationship between polygenicity and subcortical brain alterations or mechanisms involved in psychiatric conditions.

Specific aims:

1. Compare subcortical brain effect sizes across oligo- and multigenic CNVs as well as across polygenic risk scores and idiopathic conditions.

2. Characterize potential subcortical brain alteration overlap between genetic risk factors and idiopathic conditions

Hypothesis :

Polygenic risk and oligo- multi-genic CNVs with similar effect size on cognition and risk for disease should result in structural brain alterations of similar effect sizes.

Relevance : Investigating this hypothesis is critical to shed light on the heterogeneity of brain mechanisms involved in NPDs.

Methods

Participants: Deletions and duplications carriers' neuroimaging data included in the study were selected on the following breakpoints: 1q21.1 (Class I, 146.4-147.5MB & II, 145.3-147.5MB), 15q11.2 (BP1-2, 22.8-23.0MB), 16p11.2 (BP4-5, 29.6-30.2MB), 16p12.1 (chr16 21.95-22.43MB), 16p13.11 (chr16 15.51-16.29MB), and 22q11.2 (BPA-D, 19.04-21.5MB), together with NRXN1 deletion (chr2 50.14-51.26MB), TAR (chr1 145.39-145.81MB) and 13q12.12 (chr13 23.56-24.88 MB) duplications, as well as control individuals not carrying any CNVs at these loci (Table 1, and supplementary materials). Signed consents were obtained from all participants or legal representatives prior to the investigation. Data from clinically ascertained CNV carriers and controls were pooled from 5 different cohorts: Cardiff

University (UK), 16p11.2 European Consortium (Lausanne, Switzerland), University of Montreal (Canada), UCLA (Los Angeles, USA) and the Variation in individuals Project (SVIP, USA), and have been previously published in [79]. Data for non-clinically ascertained individuals were identified in the UK Biobank [80] and in part previously published in [79].

Loci	Chr (hg19) start-stop	Type	nGenes (Gene)	n tot/clin	Age (SD)	Sex (M/F)	Cohorts	Diagnosis (tot/clin)	IQ loss	OR ASD/SCZ
1q21.1	chr1 146.53-147.39	Del	7 CHDIL	40/28	38 (21)	22/18	BC-Cardiff-SVIP-UKBB	12/8	15	3.2/6.4
		Dup		30/17	47 (19)	18/12		5/5	25	5.3/2.9
TAR	chr1 145.39-145.81	Dup	15 RBM8	31/0	60 (8)	14/17	UKBB	-	2.4	-
NRXN1	chr2 50.14-51.26	Del	1 NRXN1	9/0	63 (5)	5/4	UKBB	-	9	7.9/4.7
13q12.12	chr13 23.56-24.88	Dup	5 SPATA13	21/0	62 (8)	11/10	UKBB	-	0.6	-
15q11.2	chr15 22.81-23.09	Del	4 CYF1P1	108/0	65 (7)	59/49	UKBB	2/0	5.7	1.3/1.9
		Dup		144/0	64 (7)	77/67		6/0	0.9	1.8/1
16p11.2	chr16 29.6530.20	Del	27 KCTD13	82/78	19 (15)	37/45	BC-Cardiff-SVIP-UKBB	49/49	26	14.3/1.1
		Dup		75/68	34 (17)	32/43		30/30	11	10.5/11.7
16p12.1	chr16 21.95-22.43	Del	7 PDZD9	14/0	60 (7)	4/9	UKBB	-	-	-/2.72
		Dup		12/0	66 (5)	9/3		-	-	-
16p13.11	chr16 15.51-16.29	Del	6 MYH11	10/0	63 (8)	4/6	UKBB	-	7.3	2.5/2.2
		Dup		50/0	66 (6)	26/24		-	8.7	1.5/2
22q11.2	chr22 19.04-21.47	Del	49 AIFM3	68/68	14 (6)	33/35	BC-Cardiff-UCLA-UKBB	41/41	28.8	32.3/23
		Dup		26/19	29 (23)	11/15		11/10	8.3	2/0.2
Controls				782/317	48 (21)	387/395	BC-Cardiff-SVIP-UCLA-UKBB	24/24	-	-

Table 1: Demographics

Legend: CNV carriers and controls from the clinically ascertained group come from 5 different cohorts (Supplementary Table 1), while non-clinically ascertained participants were

identified in the UK Biobank with 3 identical scanning sites. Other diagnosis included: language disorder, major depressive disorder, posttraumatic stress disorder (PTSD), unspecified disruptive and impulse-control and conduct disorder, social anxiety disorder, social phobia disorder, speech sound disorder, moderate intellectual disability, specific learning disorder, gambling disorder, bipolar disorder, conduct disorder, attention deficit / hyperactivity disorder ADHD, Substance abuse disorder, global developmental delay, motor disorder, obsessive-compulsive disorder, sleep disorder, tourettes disorder, mood disorder, eating disorders, transient tic disorder, trichotillomania, pervasive developmental disorder NOS, specific phobia, body dysmorphic disorder, mathematics disorder, dysthymic disorder.

MRI image acquisition and preprocessing: Data sample included T1-weighted (T1w) images at 0.8 - 1 mm isotropic resolution across all sites. MRI parameters are detailed in Supplemental Material.

Subcortical segmentation and shape analysis: FreeSurfer 5.3.0 was used to segment all scans into seven bilateral subcortical regions of interest: nucleus accumbens, amygdala, caudate, hippocampus, putamen, pallidum, and thalamus. The ENIGMA subcortical shape analysis pipeline was then applied to derive two measures of shape morphometry for each subcortical region: 1) the radial distance, which is the distance from each vertex to the medial curve of each region (referred to as thickness); 2) the logarithm of Jacobian determinant (LogJacs), which correspond to the surface dilation ratio between the subject structure and the template (referred to as surface area) [13]. Intracranial volume (ICV) was used as a global metric for total brain volume or head size in all our analyses. For the subcortical volumes, the primary analysis averaged the left and right hemisphere volumes for the seven structures, while a secondary analysis consisted of bilateral subcortical regions. For subcortical shape measures we analyzed vertices across all the 14 subcortical structures (bilateral).

Quality control: visual quality inspection was performed by the same rater for all volumes and shape models using the ENIGMA standardized quality control protocol.

Normative modeling: Changes in brain measures over age (in controls) were modeled using Gaussian processes [81], and compared with linear models (Figure 3, Supplement Figure 1). Similar to [82] we observed approximately linear effects for Accumbens, Caudate, Palladium, and Putamen, where the volumes peak during the first decade and decile over the years. Non-linear effects for age, flattened-inverted U-shape, were observed for ICV, Amygdala, Hippocampus, and Thalamus [82]. In subsequent analysis we used Gaussian Processes Regression (fitting model on controls and using age, sex, site, and ICV as covariates) to obtain W-scores (GPR based Z-scores w.r.t mean and standard deviation modelled in controls, Supplement Figure 2).

Statistical analysis: Linear regression models were used to compute case-control differences (Cohen's d) for each CNV using the Gaussian Processes Regression based W-scores. This approach was used for CNVs across ICV, Subcortical-volumes, and Subcortical-Shape analysis, as well as PRS and traits. FDR procedure was applied within CNVs, NPDs, PRS, and traits respectively. For subcortical shape analysis, we applied FDR procedure across all structures within a CNV. The significance was set at FDR-corrected p values<0.05.

Effect sizes: To compare across CNVs, NPDs, PRS, and traits, we used Cohen's d as effect sizes. For CNVs, Cohen's d was computed based on case-control linear regression. Idiopathic condition effect sizes were the Cohen's d values reported in the latest ENIGMA studies. For PRS and traits, continuous variables, we computed Cohen's d using top and bottom deciles as

32

cases and controls respectively. All effect sizes were computed after regressing for age, sex, site, and ICV. ICV was not used as a covariate for ICV.

Quantifying shared variance across CNVs: Principal Component Analysis (PCA) quantified shared variance across CNVs, using CNVs as input-variables and subcortical volumes (or vertex-wise measures) as observations (z-scored Cohen's d contrasts adjusted for ICV and nuisance variables; FactoMineR package in R). Similar approach was used for Idiopathic conditions, and other PCA analysis.

Results

1. Effects of CNVs on ICV, Subcortical volumes and shape

6 CNVs out of the 15 CNVs had significant effects on ICV. 7 CNVs decreased ICV and 8 increased ICV. Opposing effects were observed for deletions and duplication at the same locus (Figure 3).

11 out of 15 CNVs had significant effects on the volume of at least one subcortical structure that survived FDR correction (15 CNVs by 8 volumes). The largest effects were observed for 22q11.2 deletions followed by 16p13.11, 16p11.2, 1q21.1 deletions and 1q21.1, 16p11.2 duplications (Figure 3B). Every structure was affected by at least 2 CNVs and hippocampus, amygdala and accumbens were affected by 4 CNVs.
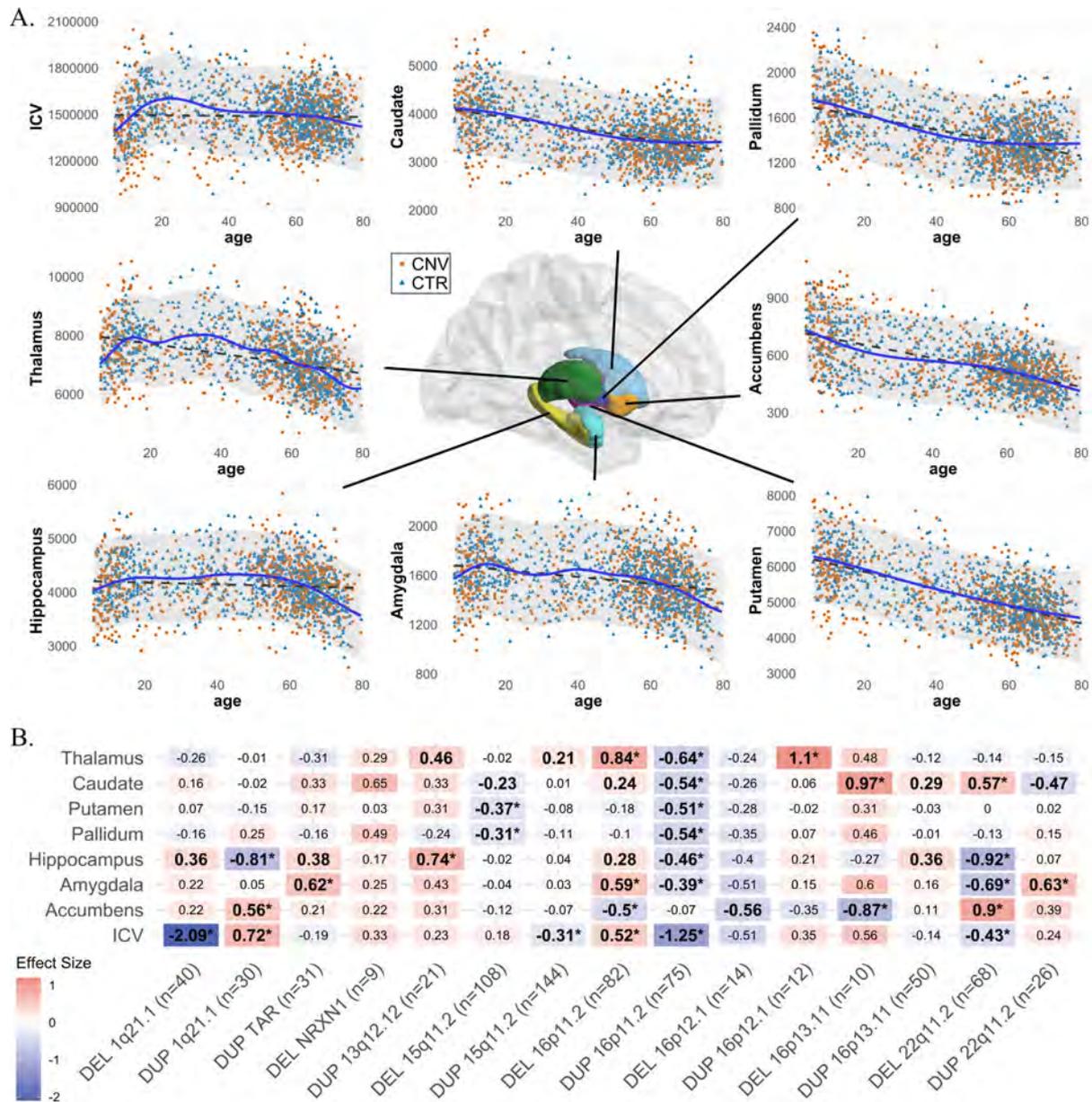
Figure 3: Normative age modeling and SubCortical volume effect sizes.

Legend: A) Scatterplots showing the distribution of ICV and SubCortical volumes with age, along with Gaussian processes modelling (solid line) and linear model (dotted line). B) Cohen's d values for subcortical structures and ICV for CNVs. Case-control differences are calculated (lm in R) using W-scores obtained from Gaussian processes regression (GPR, with age, sex, site, and/or ICV as covariates). Significant effect sizes with nominal **p-value <0.05** are in bold, and **FDR p-value <0.05** are shown with star (*), FDR correction is applied across 15 CNVs x 8 volumes. Darker color represents higher magnitudes. Sample sizes used for each analysis (for ICV) are reported in parentheses along with x-axis labels. DEL:

deletions; DUP: duplications. Detailed effect sizes, SE, p-values are reported in Supplement Figure 2.

To provide a more refined subcortical analysis we investigated subcortical shape differences using thickness and surface area (Log-Jacobian) measures. Shape analysis detected significant group differences across all CNVs (Figure 4, Supplement Figure 5), with both higher and lower thickness and surface relative to the control groups (Figure 4 and Supplement Figure 5). The number of significant vertices varied across CNVs and structures for both thickness and surface, with 22q11.2 deletion having the highest and 16p12.1 deletion the lowest number of significant vertices (Supplement Table 3-4). Across CNVs the largest number of significant vertices were, on average, observed for thickness in Caudate, and for surface in Thalamus and Hippocampus (Supplement Table 3-4). The top decile of the absolute Cohen's d values for thickness and surface were concordant (CCC= 0.6 and 0.68 respectively) with those reported for volume but were on average higher (Supplement Figure 8, and Supplement Table 1-2). Sensitivity analysis testing the effect of psychiatric diagnoses, pooling data across cohorts, and bilateral subcortical structures demonstrated that results were robust (Supplementary Figure 6). In addition, we replicated previously published [13] deletion 22q11.2 effect sizes (n=430 deletion carriers) using a subset of data (n=68 deletion carriers; r=0.93, p=2e-3; Supplementary Figure 6).
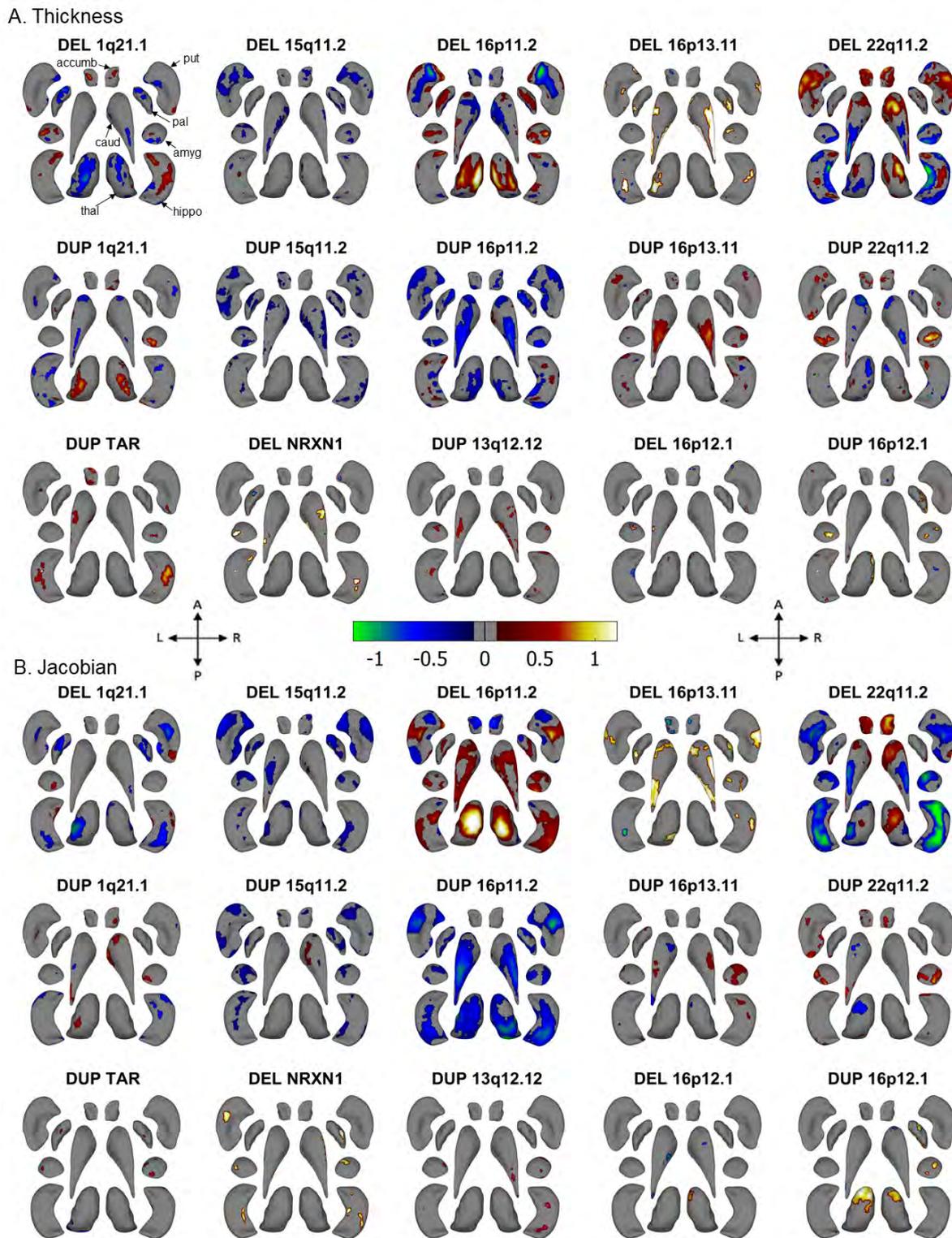
Figure 4: Cohen's d maps for SubCortical Shape analysis.

Legend: Effect sizes for Subcortical shape analysis of thickness (panel A); and Jacobian (panel B) measures. Cohen's d values are shown for 15 CNVs (Dorsal view) after applying FDR correction (<0.05) across all 14 vertices for 15 CNVs (within each measure/panel). Thickness represents local radial distance, and Jacobian represents local surface area

dilation/contraction. Blue/green colors indicate negative coefficients, or regions of lower thickness measures in the CNV group compared with the controls. Red/yellow colors indicate positive coefficients, or regions of greater thickness values in the CNV group compared with the controls. Gray regions indicate areas of no significant difference after correction for multiple comparisons. Each vertex was adjusted for sex, site, age, and intra-cranial volume (ICV) using Gaussian Processes Regression before running case-control analysis. DEL: deletion; DUP: duplication; **accumb***: Accumbens;* **amyg***: Amygdala;* **caud***: Caudate;* **hippo***: Hippocampus;* **pal***: Palladium;* **put***: Putamen;* **thal***: Thalamus. Ventral views are shown in Supplement Figure 5.*

2. CNV effect sizes for subcortical volume/shape, cognition and risk for disease

We observed a correlation between the effect size of CNVs on SubCortical volumes/thickness/surface and their previously reported effect size on IQ [83,84] (r=0.66-0.75, p<0.03, Figure 5A) and risk for either ASD [35,48,49] or SCZ [16,85] r=0.72-0.88, p<0.03, Figure 5A). On the contrary, correlations with ICV effect sizes were not significant (Figure 5A).
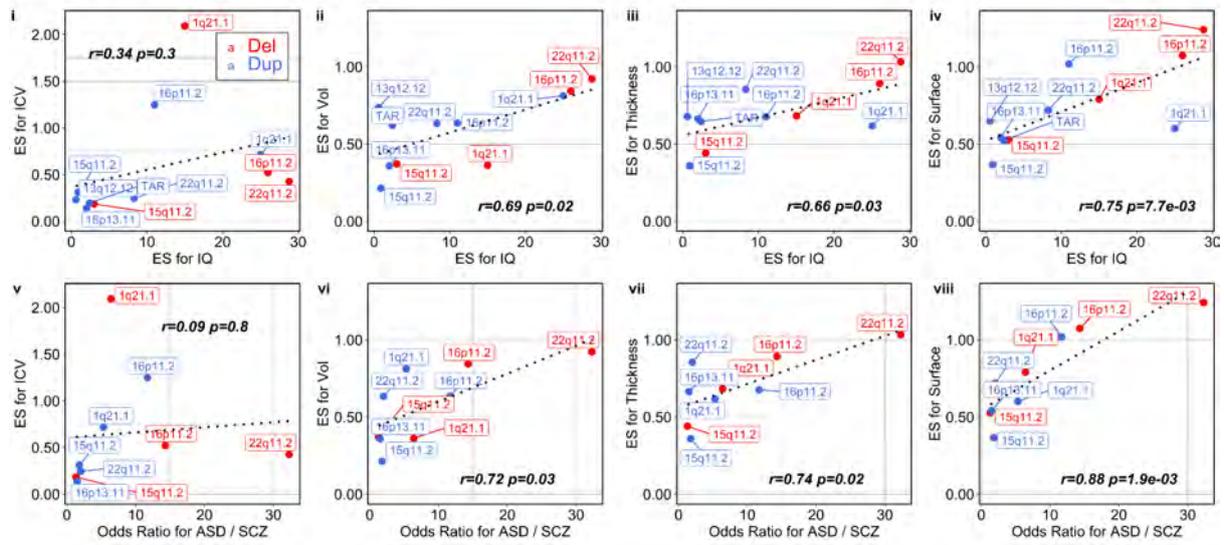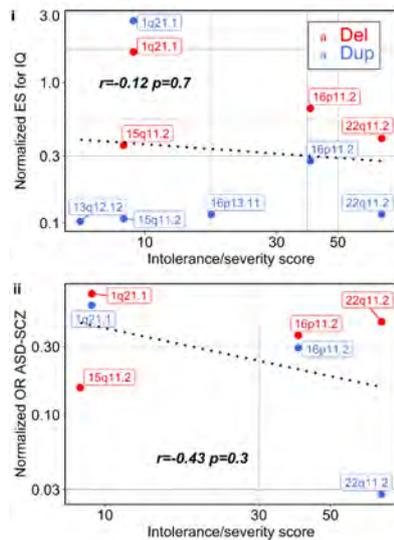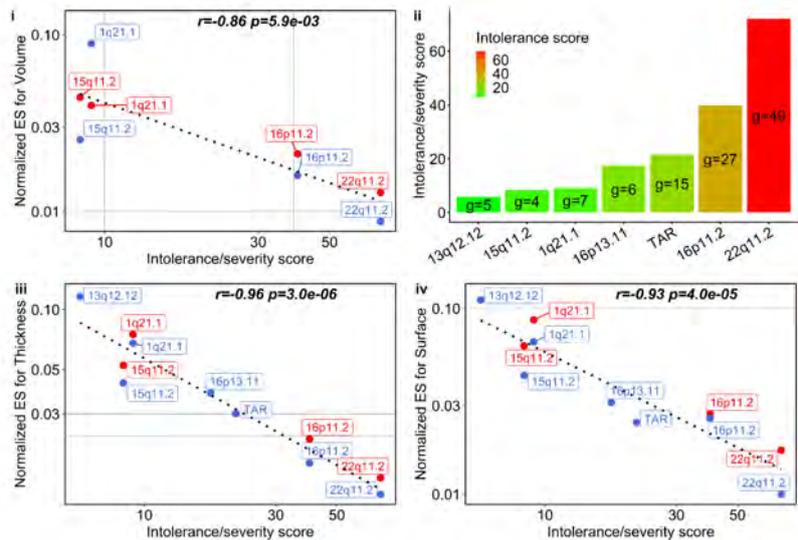
Figure 5: SubCortical Effect sizes and Polygenicity

Legend: Scatter plots showing association between Subcortical effect sizes (ES) and cognition / disease-risk (panel A); normalized effect sizes (ES divided by intolerance/severity score) and intolerance/severity score (sum of 1/LEOUF) for CNVs (panel B-C).

Panel A: Effect sizes for ICV, Subcortical volumes, and Subcortical shapes (thickness and surface) versus previously published effects of CNVs on IQ (i-iv); and risk for ASD/SCZ (v-viii; x-axis: maximum of ASD and SCZ Odds Ratio). Panel B: Effect sizes normalized by intolerance/severity score (sum of 1/LOEUF) versus effect sizes on cognition / disease-risk. Panel C: Effect sizes normalized by intolerance score versus intolerance score for subcortical effect sizes (i,iii,iv); and (ii) barplot showing intolerance score (sum of 1/LOEUF) for CNVs at different Loci (bars are color-coded based

on sum of 1/LOEUF, and number of genes (g) are added as labels). Panel B-C are in log10 scale. Deletions, and duplications are colored in red, and blue respectively. Linear fitted lines are added to each plot using geom_smooth and lm method. Pearson correlations and p-values are reported in each plot. Del: deletion; Dup: duplication; ES: Effect Size; Surface: logarithm of Jacobian determinant; 1/LOEUF: inverse of loss-of-function observed/expected upper bound fraction; g: number of genes within CNV; ASD: autism spectrum disorder; SCZ: schizophrenia; OR: odds ratio.

3. Normalized effect sizes

Previous studies demonstrated that the effect-size of CNVs on IQ increases linearly with the number of encompassed genes. Specifically, the number of genes (weighted by their sensitivity to gene dosage) can predict the effect of any CNV on IQ with 78% accuracy [83,84]. We tested if the same observation was true for subcortical structures. Specifically, does the effect size of CNVs on subcortical measures increase linearly with the number of intolerant genes encompassed in CNVs? We therefore computed for each CNV, an effect size "normalized by intolerance" (normalized ES= effect size divided by *intolerance score* of a CNV). Sum of 1/LOEUF (inverse of loss-of-function observed/expected upper bound fraction) was used as an intolerance score (Figure 5Cii). Sensitivity analysis using pLI, and number of genes are reported in Supplement Table 5 and Supplement Figure 9.

The normalized ES dramatically decreased from small oligogenic to large multigenic CNVs (volume: r=-0.6 (p=5e-2); thick: r=-0.74 (p=9e-3); surface: r=-0.74 (p=9.7e-3); Supplement Table 5) and this effect was stronger on a log-log scale (volume: r=-0.86 (p=6e-4); thick: r=-0.96 (p=3e-6); surface: r=-0.93 (p=4e-5); Figure 5 and Supplement Table 5). In other words, large multigenic CNVs have strikingly smaller effects on SubCortical volumes than expected based on the number of intolerant genes they encompass.

Anova of a linear vs non-linear (polynomial of degree 3, Supplement Table 5) model showed that this decrease in normalized ES was non-linear for volume (F=11.2, DF=2, p=7e-3), thickness (F=26.7, DF=2, p=5e-4), and surface (F=10, DF=2, p=9e-3).

4. PRS and idiopathic conditions.

If CNVs with increasing numbers of genes result in a decrease of normalized ES for subcortical structures, this may have extreme consequences for PRS which aggregates the effects of thousands of variants. We therefore investigated the effect sizes of PRS-ASD, SZ, MDD and IQ on subcortical volumes by comparing top and bottom decile (Supplemental Figure 7).

Overall, effect size associated with these 3 PRS scores were extremely low ranging from 0.02 to 0.08 for subcortical structures. This was approximately 6 to 10-fold smaller (Figure 3B, Supplement Figure 7) than those observed for CNVs with similar risk for ASD, SZ or effect on IQ (e.g. deletion 15q11.2, and duplication 1q21.1 and 22q11.2).

We also compared effects of CNVs to those of polygenic idiopathic conditions. Cohen's d values for idiopathic SZ, OCD, MDD, BP, ASD and ADHD obtained from previous Enigma studies were 2 to-6-fold smaller. Schizophrenia showed the largest effect size (Supplement Table 1).

5. Latent dimensions of shared variance across CNVs and NPDs

We performed a multivariate analysis (PCA) to understand potentially shared and distinct effects of CNVs and idiopathic conditions on subcortical volumes, thickness, and surface (Supplement Figure 10-11). The two principal components explained 68%, 35%, and 42% of variance respectively, showing a decrease in commonalities when moving from globals to higher granularity (Supplement Figure 11). Similar multivariate analysis for thickness and

40

surface showed that the two principal components explained a total of 35% and 42% of variance respectively (Supplement Figure 10-11). While the CNV groupings differed from the ones obtained from volume, and between thickness and surface, one cluster including deletion 16p11.2, and duplication 15q11.2, was preserved across volume and surface (Supplement Figure 11).

In addition, we ran a PCA to understand potentially shared and distinct effects of CNVs and idiopathic conditions on subcortical volumes. The two principal components explained 45 and 28% of the variance of Cohen's d profiles. Dimension 1 of idiopathic conditions and CNVs shows positive and negative loadings for the basal ganglia (Pallidum, Putamen) and limbic system (Thalamus, Hippocampus, Amygdala) respectively (Figure 6C). The 2nd PC dimension is characterized by the accumbens and thalamus loading on both extremes. K-means clustering using PC1 and PC2 identified 4 clusters, with cluster 4 (violet color) corresponding to adult IPCs, cluster 1 corresponding to ASD, and ADHD, and cluster 2 and 3 to CNVs. These groupings were reflected in the effect size correlation matrix (Figure 6A). To test if a specific group was driving the PCA, we performed 2 seperate PCAs on idiopathic conditions, and CNVs. The latter identified latent dimensions (PC's) that were highly correlated with each other (r= -0.93 (p=2.5e-3) between PC1 $_{IPCs}$ and PC1$_{CNVs}$; r= -0.83 (p=0.02) between PC2 $_{IPC}$ and PC2$_{CNV}$; Supplement Figure 12) as well as PC's from CNV+IPC (Supplement Figure 12).
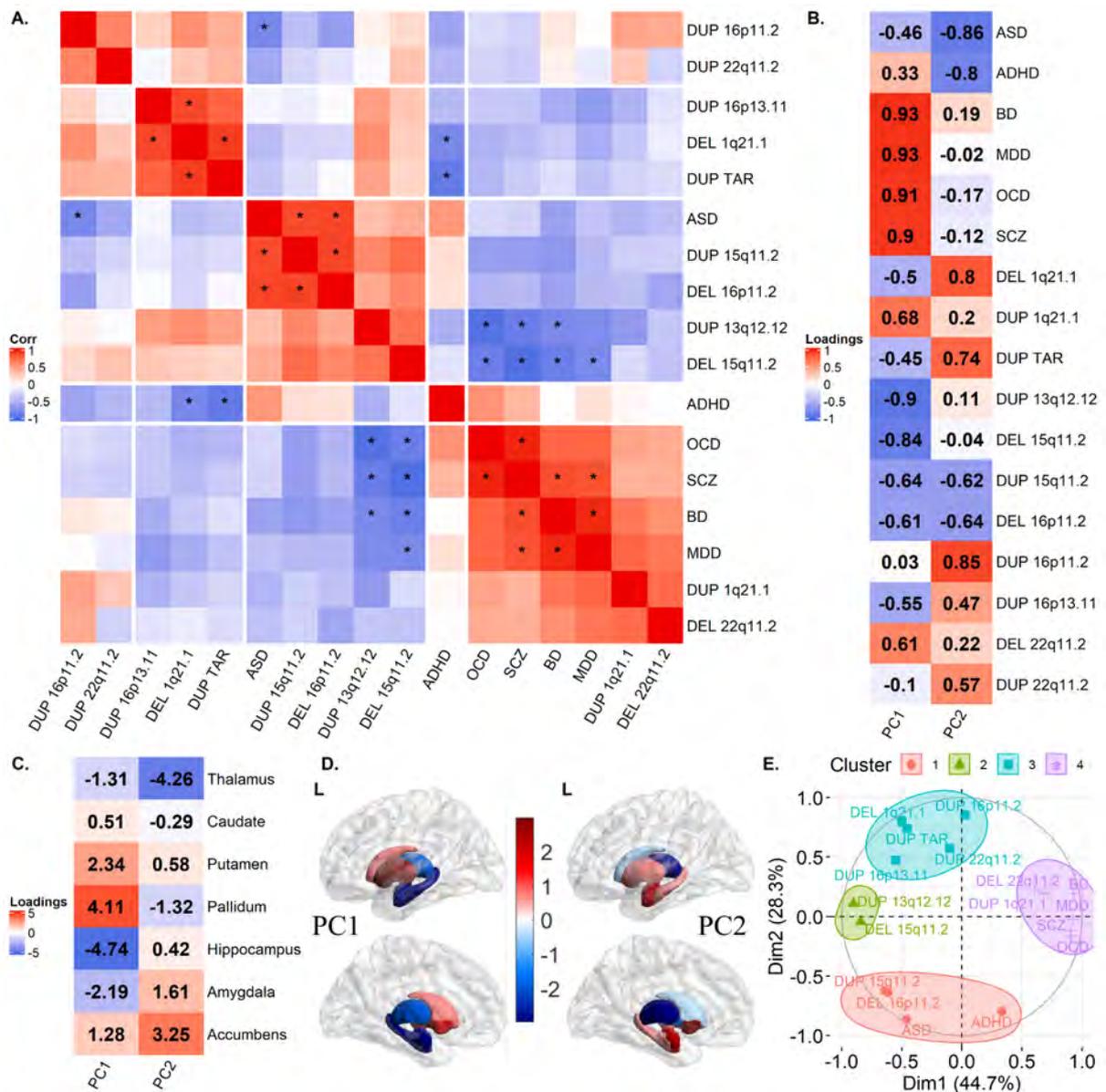
Figure 6: Principal component analysis.

Legend: A) Correlations between effect-size profiles of CNVs and Idiopathic conditions. * represent p-value <0.05, obtained using a parametric test (cor.test function in R). Hierarchical clustering (ward method) identified 4 clusters. B-E) Principal Component Analysis across SubCortical volumes of all rare and common genetic risk as well as IPCs. B) Variable loadings on PC1 and 2; C) Subcortical structure loadings; D) PC1 and PC2 loadings shown on SubCortical structures. The colorbar values above -3 and 3 are set to -3 for ease of visualization. (E) Correlation circle showing genetic risk and IPCs in PC1 and PC2 space. K-means clustering in the PC space (using euclidean distance) identified 4 clusters. DEL: Deletion; DUP: Duplication; L: left hemisphere; IPC: Idiopathic condition; ASD: autism spectrum disorder; ADHD: attention deficit hyperactivity disorder; BD: bipolar

Discussion

Here, in the largest cross-CNV neuroimaging study to date, we characterized and compared the effects of 12 NPD CNVs, 3 non-NPD CNVs, 4 psychiatric PRS and 6 psychiatric idiopathic conditions. Our results show that 8/12 NPD-CNVs, and 2/3 non-NPD CNVs impact subcortical volumes, and at least 4/14 subcortical structures have significant vertices across all CNVs for surface and thickness. The effect size of CNVs on subcortical structures was correlated to their effect size on cognition and risk for ASD-SZ. That is, larger and gene rich CNVs have larger impacts on brain and behavior. However, we also show that ES normalized for intolerance dramatically decreases from small oligogenic to larger multigenic CNVs. This was consistent with the fact that ES of CNVs were up to an order of magnitude larger than those observed for psychiatric PRS and 2 to-6-fold larger than those of idiopathic conditions. We identified the same latent dimension -defined by opposing loadings on basal ganglia and limbic structures- across CNVs and across idiopathic conditions explaining 44.7% of the Cohen's d map's variance across all CNVs and idiopathic conditions.

1. Effects of CNVs on subcortical structures and behavior.

CNVs showed broad significant effects on all subcortical measures: volume, thickness and surface area. In particular 11 out of 15 CNVs had a significant effect on at least one subcortical volume, and all CNVs showed significant thickness and surface subcortical alterations.

The moderate correlation between CNV effects on subcortical brain measures and their effect on behavior suggest that the relationship between MRI derived measures and behavior may be

43

much higher than those recorded in general population cohorts (eg. r=0.13 reported by Marek et al.).

2. Distinct and shared effects of CNVs

Over half of the mean CNV effects on subcortical structures are distinct. This is consistent with a recent study showing relative specificity of association between CNVs and brain alteration [67]. Moreover these results are very similar to the ones obtained in our previous work where we investigated overlapping brain alterations across CNVs at the level of, and study on whole-brain volume, cortical thickness and surface area [79]. On the other hand the presence of shared effects across CNV subgroups suggests that similar brain mechanisms might underlie brain alterations across subsets of CNVs.

3. Polygenicity and effects of CNVs on Subcortical structures

Polygenicity had an impact on Subcortical structures: As CNVs increased in size and number of deleted or duplicated genes, their normalized effect size on Subcortical structures (effect size / intolerance score) decreased. Suggesting that the extent of alterations on SubCortical structures was not an additive effect of individual genes encompassed in CNVs, otherwise we would have observed a constant association (close to zero correlation).

4. Effect sizes across CNVs, IPC, PRS

We show that CNV effect sizes on subcortical structures were 2- to 10-fold larger than the ones observed for idiopathic conditions and PRS respectively.

The extremely small effects of PRS on subcortical structures are consistent with previous smaller studies using cortical surface and thickness measures. A small number of studies have tested PGRS-SCZ in relation to structural brain imaging phenotypes with inconsistent results

44

finding either no associations between subcortical or cortical regional volume as well as white matter microstructure and PGRS-SCZ [69,85] or weak associations [86–88].

These results have implications for our conceptualization of polygenic psychiatric conditions. Indeed, studies estimate that 70 to 100% of any 1-MB window in the human genome encompasses variants (including CNVs) contributing to increased risk for schizophrenia and autism.

5. Limitations

Multiple sites included in the study may have introduced noise, but previous studies have shown that site effects do not influence the neuroanatomical patterns associated with CNVs at the 16p11.2, 22q11.2, and 15q11.2 loci [12,19,23]. While shared variation could have been influenced by clinical ascertainment or psychiatric diagnoses, our sensitivity analyses showed that this is not the case. Effect of medication on CNVs brain alterations could not be investigated in the current study as medication information was not available for the whole dataset. We were underpowered to properly investigate potential sex related effects of CNVs on subcortical structures. Of note, previous neuroimaging studies of large 22q11.2 and 16p11.2 samples were unable to identify any sex related effects [54,90].

Small sample size was available for 16p13.11, TAR, NRXN1, 13q12.12, 16q11.1 deletions and duplications, larger samples would improve the accuracy of the effect sizes. Systematic analysis through two computational neuroanatomy frameworks (volume-based and shape-based) shows that CNV effects could not be attributed to the processing pipeline. Gaussian Processes Regression -modelling for covariates using all controls- was used to model non-linear effects of age and other covariates, however, larger sample sizes with better coverage of age range would be required for more accurate modelling. Extending our approach to the rapidly expanding number of rare genomic variants associated with psychiatric disorders is

required to draw a robust conclusion on the distinct and shared effects of CNVs on brain structure.

Conclusions

NPD CNVs affect subcortical structures with ES that are correlated to their effect size for cognition and risk for disease. However, our results suggest that the ability of structural MRI methods to detect differences is inversely correlated to the complexity of genetic risk. Oligogenic variants are the easiest to investigate while PRS provides only a faint signal. How combinations of genes included in CNVs or variants in PRS affect neuroanatomy and risk for disease remains a critical question in neuroscience and psychiatry. Answering this question will require studies on a broader spectrum of risk variants.

Data and materials availability

UK Biobank data was downloaded under the application 40980, and can be accessed via their standard data access procedure (see http://www.ukbiobank.ac.uk/register-apply). UK Biobank CNVs were called using the pipeline developed in Jacquemont Lab, as described in https://github.com/labjacquemont/MIND-GENESPARALLELCNV. The final CNV calls are available for download from the UK Biobank returned datasets (Return ID: 3104, https://biobank.ndph.ox.ac.uk/ukb/dset.cgi?id=3104).

Discussion

1.   Global and regional effects

1.1.   Global effects

Overall CNVs have a substantial effect on global brain metrics but the directionality of their effects varies from one locus to another, as shown by our meta-analysis on several CNVs and aneuploidies [14] (Figure 7). For CNVs such as 16p11.2, 22q11.2 and 1q21.1 mirror dose responses on brain volume are observed, while for other CNVs such as 15q11.2 the sample and effect sizes are too small to provide definitive answers. Very similar effects of CNVs are observed on brain volume and total SA while mean CT seems to be differently affected. These findings are supported by recent GWAS studies which show that different neurodevelopmental mechanisms can affect cortical SA expansion and CT increase [91]. As also hypothesized by the radial unit hypothesis [92,93] different mechanisms underlie cortical development of SA and CT. In fact cortical SA development is primarily driven by the number of radial columns perpendicular to the pial surface while CT development is determined by the horizontal layers in the cortical columns.

Findings of our first study also show that effect sizes on global brain metrics are not correlated with the number of genes encompassed in the CNV, or with symptom severity. We would have expected that CNVs encompassing numerous genes would have had a stronger impact on global metrics and would have shown larger effect sizes, while CNVs encompassing fewer genes would have shown small effects, but this was not the case. In fact 1q21.1 deletions and duplications encompass only 16 genes but show among the largest effect sizes on global metrics. We could hypothesize that those large CNVs do not necessarily

encompass genes related to brain volume. Or that they encompass both, genes with a positive and genes with a negative effect on brain volume, and that the effect is therefore neutralized. Gene dosage effects on brain volume have been interpreted as evidence of abnormal neurogenesis [94], supposedly involving genes implicated in the control of cell size and proliferation [95–97].

A

CNV	Studies with TBV	Summary effect size	95% CI
Down syndrome	7	-1.65	[-2.34, -0.96]
XXX	6	-1.4	[-1.87, -0.93]
XXY	9	-0.76	[-1.19, -0.33]
XYY	5	0.13	[-0.33, 0.59]
XXYY	6	-0.76	[-1.26, -0.26]
TS	8	-0.18	[-0.7, 0.35]
7q11.23	8	-1.63	[-2.34, -0.92]
1q21.1 distal	1/1	-2.12/1.03	[-2.53, -1.71]/[0.56, 1.5]
15q11.2	1/1	0.09/-0.25	[-0.15, 0.33]/[-0.48, -0.02]
16p11.2 distal*	1/1	0.26/-0.9	[-0.32, 0.84]/[-1.45, -0.35]
16p11.2 proximal	3/3	0.87/-0.93	[0.36, 1.36]/[-1.52, -0.37]
22q11.2	9/1	-0.92/0.29	[-1.42, -0.36]/[-0.14, 0.72]
15q11-q13 (AS ■/PWS ●)	2/4	-1.76/-0.99	[-2.56, -0.95]/[-1.69, -0.2]
SHANK3	1	-0.62	[-1.42, 0.18]
ASD*	van Rooij et al, 2018	0.13	[ ]
SCZ*	van Erp et al, 2018	-0.12	[-0.24, 0.00]

Deletions / Duplications

B

CNV	Studies with SA	Summary effect size	95% CI
Down syndrome	1	-1.67	[-2.21, -1.13]
XXX	1	-1.22	[-1.67, -0.77]
XXY	1	-0.53	[-0.87, -0.19]
XYY	1	0.16	[-0.28, 0.6]
XXYY	1	-0.53	[-1.02, -0.04]
TS	1	0.19	[-0.3, 0.68]
7q11.23	1	-2.62	[-3.18, -2.06]
1q21.1 distal	2/2	-1.72/0.68	[-2.11, -1.33]/[0.24, 1.12]
15q11.2	2/2	-0.3/-0.1	[-0.49, -0.11]/[-0.27, 0.07]
16p11.2 distal	1/1	-0.14/-0.19	[-0.9, 0.61]/[-0.53, 0.16]
16p11.2 proximal	3/3	0.6/-0.96	[0.47, 0.73]/[-2.04, 0.12]
22q11.2	5/2	-1.18/0.29	[-1.41, -0.95]/[-0.21, 0.79]
15q11-q13 (AS/PWS)	-	-	[]/[]
SHANK3	-	-	[]
ASD	Boedhoe et al, 2020	0.01	[-0.06, 0.09]
SCZ	van Erp et al, 2018	-0.25	[-0.33, -0.17]

Deletions / Duplications

C

CNV	Studies with CT	Summary effect size	95% CI
Down syndrome	3	0.16	[-1.22, 1.54]
XXX	1	-0.46	[-0.89, -0.03]
XXY	1	-0.19	[-0.52, 0.14]
XYY	1	0.31	[-0.13, 0.75]
XXYY	1	-0.11	[-0.6, 0.38]
TS	1	0.29	[-0.2, 0.78]
7q11.23	1	0.8	[0.37, 1.23]
1q21.1 distal	2/2	0.37/0.08	[-0.01, 0.75]/[-0.36, 0.52]
15q11.2	2/2	0.38/-0.24	[0.17, 0.56]/[-0.41, -0.07]
16p11.2 distal	1/1	0.13/-0.27	[-0.58, 0.83]/[-0.75, 0.21]
16p11.2 proximal	3/3	-0.4/-0.56	[-1.28, 0.48]/[-0.98, -0.14]
22q11.2	5/2	0.66/-0.15	[0.49, 0.83]/[-1.38, 1.08]
15q11-q13 (PWS)	1	0.16	[]/[-0.56, 0.88]
SHANK3	-	-	[]
ASD	Bedford et al, 2019	0.22	[0.12, 0.33]
SCZ	van Erp et al, 2018	-0.52	[-0.63, -0.41]

Deletions / Duplications

49

Figure 7: From Modenato et al. (2021). Forest-plots with the summary estimates from the meta-analyses of the effects of 20 CNVs –as well as idiopathic ASD and SCZ– on three global metrics: total brain volume (TBV,a), total surface area (Total-SA, b) and mean cortical thickness (Mean-CT, c).

Legend: Summary estimates of the effect size of each CNV were derived from meta-analyses including all neuroimaging studies reporting TBV, Total-SA and Mean-CT data, respectively (see Table 1 and Supplementary Figure 1). Data for idiopathic ASD and SCZ were obtained from the biggest large-scale study to date (32-34). 95% CI provided by the meta-analyses are presented as solid error bars, whereas CNVs with only one available study show the 95% CI from that particular study as dotted error bars, together with bold italic font in the 95% CI column. The size of the square is proportional to the number of studies reporting data. Filled squares correspond to statistically significant effect sizes. CNV deletions are depicted in blue, duplications in red and data on idiopathic psychiatric conditions in black. * indicates intracranial volume (ICV) instead of TBV. Abbreviations: TS: Turner Syndrome; AS: Angelman Syndrome; PWS: Prader-Willis Syndrome; ASD: Autism Spectrum Disorder; SCZ: Schizophrenia.

1.2.   Cortical effects

Regarding cortical brain morphometry adjusted for global differences, no obvious pattern is recurrently observed across CNVs. Although some regions appear to be more frequently affected such as: insula, cingulate, dorsolateral prefrontal cortex, inferior frontal gyrus, orbitofrontal cortex, supplementary motor cortex, postcentral gyrus, superior parietal area, fusiform gyrus, and superior temporal area [14].

By comparing brain morphometry across CNVs we observed a dissociation between global and regional effects sizes. 1q21.1 deletion and duplication clearly show the divergence between a very large effect on global metrics and small effects on regional alterations on adjusted for global effects. A dissociation is also observed in the directionality of global and regional effects: deletions are associated with decreased cingulate and supplementary motor

cortex volumes irrespectively of the directionality of their effects on global measures. Therefore we can hypothesise that CNVs effects on global and regional morphometry are driven by different mechanisms, as also suggested by a recent study [98]. These findings highlight the importance of adjusting for global effect when looking at regional alteration, or else regional effect sizes may be inflated or canceled off.

1.3. Subcortical effects

The effect of CNVs on subcortical structures has been characterized for several loci including 1q21.1, 16p11.2 BP4-5, 16p11.2 BP1-3, 15q11.2, 15q11-q13 and 22q11.2 showing that subcortical structures are broadly affected without a clear pattern shared across variants [14].

In our second study we show that subcortical volumes were significantly different from controls in 11 out of the 15 CNVs included in the investigation. Mirror gene dosage effects were observed in adjusted subcortical volumes at the 16p11.2, 1q21.1, and 22q11.2 genomic loci. The effect size of CNVs on subcortical structures was correlated to their effect size on cognition and risk for ASD-SZ. That is, larger and gene rich CNVs have larger impacts on brain and behavior. However, we also show that the effect size normalized for intolerance dramatically decreases from small oligogenic to larger multigenic CNVs. This was consistent with the fact that ES of CNVs were up to an order of magnitude larger than those observed for psychiatric PRS, and 2-fold larger than those of idiopathic conditions. Finally we identified the same latent dimension -defined by opposing loadings on basal ganglia and limbic structures- across CNVs and across idiopathic conditions explaining 44.7% of the Cohen's d map's variance across all CNVs and idiopathic conditions.

2.   Distinct and shared effects

To date, neuroimaging studies have typically been carried out one mutation at a time. An important question raised by the CNV literature is whether they lead to similar behavioral and cognitive phenotypes via numerous or a limited number of brain mechanisms. We investigated CNVs effects on brain morphometry in order to understand whether CNVs at risk for psychiatric conditions shared similar anatomical brain alterations or whether each CNV had a unique brain signature.

In the first cross CNV article we found that one third of the CNV effects on brain morphometry were shared across subgroups of the 8 CNVs, while the remaining two thirds of effects were distinct. In the second study, which included 15 CNVs and was focused on subcortical brain alterations, we found that almost half of the CNVs effects on subcortical volumes were shared across CNVs subgroups. The presence of distinct CNV effects is consistent with a recent study that has reported specific effects when looking at the association between brain patterns of gene expression and patterns of cortical anatomy changes across 6 CNVs and chromosomal aneuploidies [67]. The large proportion of distinct CNV-neuroimaging effects suggests that a broad diversity of brain mechanisms increase risk for autism and schizophrenia. In fact we did not find a single dimension explaining CNV effects, meaning that there is no single morphometric profile common to all CNVs. On the other hand, the presence of a small proportion of shared effects across subsets of CNVs, suggest that similar mechanisms might underlie brain alterations within CNVs subgroups. Nevertheless, gene-morphology dimensions alone, can not explain how subgroups of CNVs are associated with a similar range of symptoms and disorders [35,37,49,99,100].

2.1. Other factors affecting brain morphometry

Effects of environmental factors such as early-life adversities and stressful experiences can increase risk for psychiatric disorder and additionally affect corresponding structural abnormalities as shown by several studies on a range of psychiatric disorders [101–109].

The presentation of neurodevelopmental disorders is heavily affected by sex with a general excess of males observed in subjects diagnosed with ASD, intellectual disabilities, speech and language disorders and ADHD [110]. Overall, females are less likely to be referred to the clinic compared to males [111]. Despite this bias, the largest studies on single CNVs (i.e. 22q11.2 and 16p11.2 studies) did not report any interaction between genetic status and sex [54,57], suggesting that the bias observed in the neurodevelopmental disorder clinic might not be related to differential effects of genetic risk at the neuroanatomical level.

Age is another factor that can affect CNVs neuroimaging findings due to the fact that cohorts of individuals with rare genetic variants often have a broad age range and rarely include sufficient number of non-carrier controls to accurately model age effects. In our second study we show that new methods to model age such as Gaussian Processes are more accurate and can improve age modelling for statistical analysis of imaging data compared to linear models.

Nevertheless the wide age range in these cohorts offer a unique opportunity to study the effects of the same molecular mechanisms across broad neurodevelopmental periods. Data suggests that alterations associated with the 22q11.2 and 16p11.2 loci on SA and regional brian volumes appear early on and remain stable through adolescence and young adulthood [112–114], with the exception of subcortical volumes in 22q11.2 deletion carriers [115]. In contrast, studies on CT suggest that normal age-related thinning may be disrupted in some CNVs. Accelerated cortical thinning visible around 40 years of age in comparison to controls was observed for 16p11.2 duplication carriers [60], while an overall thicker cortex appears to be

associated with accelerated cortical thinning in prefrontal and posterior regions in 22q11.2 deletion carriers [114,116,117].

The effects of ascertainment, comorbidities and medication are concerns that are regularly raised in CNV studies. Regarding ascertainment, we have shown that neuroanatomical findings are similar in clinically and non-clinically ascertained individuals and that they remain unchanged after excluding CNV carriers with diagnoses. In one other study Cárdenas-de-la-Parra et al. show that neuroanatomical alterations are present before the onset of symptomatology in 16p11.2 carriers [112]. Finally we also show that CNVs have much larger effect sizes than those observed in psychiatric conditions. We could not investigate the effect of medication in the current research. Nevertheless studies on 22q11.2 and 16p11.2 s have reported no significant effects of medication on neuroimaging patterns[54,118] (Martin-Brevet 2018, Rogdaki 2020).

2.2. Brain hubs

CNVs alter cortical brain regions such as insula, cingulate, fusiform gyrus, and hippocampus, that are also showing alterations across schizophrenia, bipolar disorders, major depression, and obsessive compulsive disorders [106,119]. In particular cingulate, insula, fusiform gyrus and cerebellum were altered across deletions and duplications at the 22q11.2, 16p11.2, 1q21.1 and 15q11.2 loci. These brain regions are vulnerable hubs most likely affected during brain development. In fact abnormal brain hub connectivity has been found in several neurodevelopmental idiopathic conditions as well as in neonates at risk for these conditions [120]. Overall cross disorders and cross CNVs brain alterations commonalities highlight the pleiotropy and polygenic effects underlying psychiatric disorders [121,122].

54

The identification of sensitive brain hubs associated with CNVs leads to the more specific question of how gene expression affects brain morphometry. The development of brain atlases integrating anatomic and genetic information like the Allen Human Brain Atlas (http://human.brain-map.org/) represent a pivotal step to allow such investigations. Currently only a few studies have investigated the correlation between CNVs brain alterations and cortical gene expression. The study by Seidlitz *et al.* showed that spatial patterns of CNVs brain alterations are organized by normative expression gradients of disease-relevant genes in the human brain and that this could be linked to cell-type-dependent patterning of gene expression [67]. The second study by Moreau *et al.* showed a significant association between functional connectivity signatures of the 22q11.2 and 16p11.2 genomic loci, and the expression patterns of the 37 and 24 genes encompassed in the two CNVs respectively [66]. More investigations on broader ranges of genetic variants are required to have deeper insight into the underlying mechanisms that coordinate the mapping of genetic risks onto brain alterations.

3. Polygenicity and brain alterations

In our second paper we investigated whether the additive effects of genes encompassed in CNVs, observed for cognition and risk for disease, also apply to structural brain alterations. We observed that polygenicity had an impact on subcortical structures, in fact as the number of genes encompassed in the CNVs increased, their normalized effect size on subcortical structures (effect size / nGenes) decreased. Suggesting that the extent of alterations on subcortical structures is not related to the additive effect of individual genes encompassed in the CNVs. We also showed that compared to idiopathic conditions and PRS, CNVs provided

a more powerful signal on neuroimaging alterations. These findings suggest that CNVs are a powerful paradigm to investigate how combinations of genes affect the brain structure.

4. Limitations

This research was performed on the largest multi-site CNV cohort, this implies that noise could have been introduced by the inclusion of multiple sites. However, previous work has shown that neuroanatomical alterations associated with CNVs at the 16p11.2, 22q11.2, and 15q11.2 loci are not influenced by different scanning sites [54,57,64].

Although we were able to investigate the effect of ascertainment and psychiatric diagnosis and to determine that it does not affect neuroimaging findings, we could not investigate the effect of medication due to the lack of information for this dataset.

Additionally we did not investigate the interaction of the genetic effect with sex as we were underpowered for CNVs at the 1q21.1 and 15q11.2 loci.

Neuroimaging phenotypes have typically only been investigated in the most frequent recurrent CNVs, which represents a minuscule fraction of the diverse landscape of deleterious CNVs diagnosed in individuals referred for neurodevelopmental disorders. To extend our approach to more rare genetic variants is necessary in order to be able to investigate the association of shared and distinct neuroanatomical alterations and behavioral symptoms.

5. Conclusion

This work provides many insights into the effects of CNVs on brain structure, from the dissociation of these effects on global and regional brian alterations, to the amount of shared and distinct effects across several CNVs. Finally we also showed a correlation with effects on cognition and disease risk.

These findings show that both shared and distinct brain mechanisms across genetic variants contribute to the risk for psychiatric disorders and that compared with CNVs effects on brain structure, the ones of idiopathic conditions and PRS were very weak and difficult to capture. We therefore conclude that CNVs are a powerful tool to investigate the effects of combinations of genes on brain structure and how these effects impact cognition and disease risk. The access to multi-site and to large scale studies has provided us with the opportunity to simultaneously analyse and compare several genomic variants. Nevertheless in order to draw robust conclusions and to better understand the relationship between gene function and brain morphometry our approach needs to be extended to the rapidly expanding number of rare genomic variants associated with psychiatric disorders.

6. Future Perspectives of CNV studies

Although greatly advanced in recent years, neuroimaging for rare genomics variants still requires several improvements. For the most studied variants (i.e 16p11.2 and 22q11.2) effects are well known, but robustness of findings for the most recently studied variants still needs to be demonstrated. The lack of power, due to the relatively small sample size for some of the variants, is still a major issue which hinders our ability to stratify, test for interactions and conduct genome-wide analyses. Access to large cohorts (UKBB [123], ABCD [124]) and large-scale initiatives (ENIGMA-CNV) are very promising advancements that allow to increase the

power by merging smaller datasets. Unfortunately the limitation of such initiatives is that individuals with significant psychopathology and carrying large-effect variants are significantly underrepresented [38]. Therefore it is important to pursue in parallel the recruitment of individuals selected on the basis of a broad spectrum of cognitive and behavioral symptoms, and of deleterious genomic variants from the genetic clinic. Such a dataset, focused on developmental psychiatric/genetic-first individuals, would provide a 20 to 100-fold enrichment in deleterious variants in comparison to unselected populations [37,38].

References

1. Cuthbert, B. N. The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry* **13**, 28–35 (2014).

2. Ashburner, J. & Friston, K. J. Voxel-based morphometry--the methods. *NeuroImage* **11**, 805–821 (2000).

3. Fischl, B., Sereno, M. I. & Dale, A. M. Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *NeuroImage* **9**, 195–207 (1999).

4. Lerch, J. P. & Evans, A. C. Cortical thickness analysis examined through power analysis and a population simulation. *NeuroImage* **24**, 163–173 (2005).

5. Linden, D. E. J. The Challenges and Promise of Neuroimaging in Psychiatry. *Neuron* **73**, 8–22 (2012).

6. Insel, T. R. Rethinking Mental Illness. *JAMA* **303**, 1970 (2010).

7. van Erp, T. G. M. van *et al.* Cortical Brain Abnormalities in 4474 Individuals With Schizophrenia and 5098 Control Subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. *Biol. Psychiatry* **84**, 644–654 (2018).

8. van Rooij, D. *et al.* Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. *Am. J. Psychiatry* **175**, 359–369 (2018).

9. Modenato, C. *et al.* Neuropsychiatric copy number variants exert shared effects on human brain structure. *medRxiv* 2020.04.15.20056531 (2020) doi:10.1101/2020.04.15.20056531.

10. Lin, A. *et al.* Mapping 22q11.2 Gene Dosage Effects on Brain Morphometry. *J. Neurosci.* **37**, 6183–6199 (2017).

11. van der Meer, D. *et al.* Understanding the genetic determinants of the brain with

MOSTest. *Nat. Commun.* **11**, 3512 (2020).

12. Sønderby, I. E. *et al.* 1q21.1 distal copy number variants are associated with cerebral and cognitive alterations in humans. *Transl. Psychiatry* **11**, 1–16 (2021).

13. Ching, C. R. K. *et al.* Mapping Subcortical Brain Alterations in 22q11.2 Deletion Syndrome: Effects of Deletion Size and Convergence With Idiopathic Neuropsychiatric Illness. *Am. J. Psychiatry* appi.ajp.2019.19060583 (2020) doi:10.1176/appi.ajp.2019.19060583.

14. Modenato, C. *et al.* Lessons learnt from neuroimaging studies of Copy Number Variants, a systematic review. *Biol. Psychiatry* **0**, (2021).

15. Moreau, C. A., Ching, C. R., Kumar, K., Jacquemont, S. & Bearden, C. E. Structural and functional brain alterations revealed by neuroimaging in CNV carriers. *Curr. Opin. Genet. Dev.* **68**, 88–98 (2021).

16. Kirov, G., Rees, E. & Walters, J. What a psychiatrist needs to know about copy number variants. *BJPsych Adv.* **21**, 157–163 (2015).

17. Stessman, H. A., Bernier, R. & Eichler, E. E. A Genotype-First Approach to Defining the Subtypes of a Complex Disease. *Cell* **156**, 872–877 (2014).

18. Reardon, P. K. *et al.* An Allometric Analysis of Sex and Sex Chromosome Dosage Effects on Subcortical Anatomy in Humans. *J. Neurosci.* **36**, 2438–2448 (2016).

19. Sebat, J. *et al.* Large-scale copy number polymorphism in the human genome. *Science* **305**, 525–528 (2004).

20. Watson, C. T., Marques-Bonet, T., Sharp, A. J. & Mefford, H. C. The genetics of microdeletion and microduplication syndromes: an update. *Annu. Rev. Genomics Hum. Genet.* **15**, 215–244 (2014).

21. Pérez Jurado, L. A., Peoples, R., Kaplan, P., Hamel, B. C. & Francke, U. Molecular definition of the chromosome 7 deletion in Williams syndrome and parent-of-origin

effects on growth. *Am. J. Hum. Genet.* **59**, 781–792 (1996).

22. Shprintzen, R. J., Goldberg, R. B., Young, D. & Wolford, L. The velo-cardio-facial syndrome: a clinical and genetic analysis. *Pediatrics* **67**, 167–172 (1981).

23. Fan, Y.-S. *et al.* Detection of pathogenic gene copy number variations in patients with mental retardation by genomewide oligonucleotide array comparative genomic hybridization. *Hum. Mutat.* **28**, 1124–1132 (2007).

24. Munnich, A. *et al.* Impact of on-site clinical genetics consultations on diagnostic rate in children and young adults with autism spectrum disorder. *Mol. Autism* **10**, 33 (2019).

25. Rees, E. *et al.* Analysis of copy number variations at 15 schizophrenia-associated loci. *Br. J. Psychiatry* **204**, 108–114 (2014).

26. Costain, G. *et al.* Pathogenic rare copy number variants in community-based schizophrenia suggest a potential role for clinical microarrays. *Hum. Mol. Genet.* **22**, 4485–4501 (2013).

27. Kirov, G. *et al.* The Penetrance of Copy Number Variations for Schizophrenia and Developmental Delay. *Biol. Psychiatry* **75**, 378–385 (2014).

28. Macé, A. *et al.* CNV-association meta-analysis in 191,161 European adults reveals new loci associated with anthropometric traits. *Nat. Commun.* **8**, 744 (2017).

29. Owen, D. *et al.* Effects of pathogenic CNVs on physical traits in participants of the UK Biobank. *BMC Genomics* **19**, 867 (2018).

30. Crawford, K. *et al.* Medical consequences of pathogenic CNVs in adults: analysis of the UK Biobank. *J. Med. Genet.* **56**, 131–138 (2019).

31. Li, Y. R. *et al.* Rare copy number variants in over 100,000 European ancestry subjects reveal multiple disease associations. *Nat. Commun.* **11**, 255 (2020).

32. ADHD Working Group of the Psychiatric Genomics Consortium (PGC) *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder.

Nat. Genet. **51**, 63–75 (2019).

33. Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).

34. eQTLGen *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).

35. Marshall, C. R. *et al.* Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat. Genet.* **49**, 27–35 (2017).

36. Stefansson, H. *et al.* CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* **505**, 361–366 (2014).

37. Douard, E. *et al.* Effect Sizes of Deletions and Duplications on Autism Risk Across the Genome. *Am. J. Psychiatry* appi.ajp.2020.19080834 (2020) doi:10.1176/appi.ajp.2020.19080834.

38. Huguet, G. *et al.* Estimating the effect-size of gene dosage on cognitive ability across the coding genome. *bioRxiv* 2020.04.03.024554 (2020) doi:10.1101/2020.04.03.024554.

39. Mefford, H. C. *et al.* Recurrent Rearrangements of Chromosome 1q21.1 and Variable Pediatric Phenotypes. *N. Engl. J. Med.* **359**, 1685–1699 (2008).

40. Golzio, C. *et al.* KCTD13 is a major driver of mirrored neuroanatomical phenotypes associated with the 16p11.2 CNV. *Nature* **485**, 363–367 (2012).

41. Jonas, R. K., Montojo, C. A. & Bearden, C. E. The 22q11.2 Deletion Syndrome as a Window into Complex Neuropsychiatric Disorders Over the Lifespan. *Biol. Psychiatry* **75**, 351–360 (2014).

42. Kendall, K. M. *et al.* Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects. *Biol. Psychiatry* **82**, 103–110 (2017).

43. Bernier, R. *et al.* Clinical phenotype of the recurrent 1q21.1 copy-number variant. *Genet. Med.* **18**, 341–349 (2016).

44. Brunetti-Pierri, N. *et al.* Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat. Genet.* **40**, 1466–1471 (2008).

45. Dolcetti, A. *et al.* 1q21.1 Microduplication expression in adults. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **15**, (2013).

46. D'Angelo, D. *et al.* Defining the Effect of the 16p11.2 Duplication on Cognition, Behavior, and Medical Comorbidities. *JAMA Psychiatry* **73**, 20–30 (2016).

47. McCarthy, S. E. *et al.* Microduplications of 16p11.2 are associated with schizophrenia. *Nat. Genet.* **41**, 1223–1227 (2009).

48. Moreno-De-Luca, A. *et al.* The role of parental cognitive, behavioral, and motor profiles in clinical variability in individuals with chromosome 16p11.2 deletions. *JAMA Psychiatry* **72**, 119–126 (2015).

49. Sanders, S. J. *et al.* Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* **87**, 1215–1233 (2015).

50. Zufferey, F. *et al.* A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. *J. Med. Genet.* **49**, 660–668 (2012).

51. Jacquemont, S. *et al.* Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature* **478**, 97–102 (2011).

52. Lin, A. *et al.* Mapping 22q11.2 Gene Dosage Effects on Brain Morphometry. *J. Neurosci.* **37**, 6183–6199 (2017).

53. Sønderby, I. E. *et al.* 1q21.1 distal copy number variants are associated with cerebral and cognitive alterations in humans. *Transl. Psychiatry* **11**, 1–16 (2021).

54. Martin-Brevet, S. *et al.* Quantifying the Effects of 16p11.2 Copy Number Variants on

Brain Structure: A Multisite Genetic-First Study. *Biol. Psychiatry* **84**, 253–264 (2018).

55. Sønderby, I. E. *et al.* Dose response of the 16p11.2 distal copy number variant on intracranial volume and basal ganglia. *Mol. Psychiatry* **25**, 584–602 (2020).

56. Moreau, C. A. *et al.* Mutations associated with neuropsychiatric conditions delineate functional brain connectivity dimensions contributing to autism and schizophrenia. *Nat. Commun.* **11**, 5272 (2020).

57. Sun, D. *et al.* Large-scale mapping of cortical alterations in 22q11.2 deletion syndrome: Convergence with idiopathic psychosis and effects of deletion size. *Mol. Psychiatry* 1–13 (2018) doi:10.1038/s41380-018-0078-5.

58. Niarchou, M. *et al.* Psychiatric disorders in children with 16p11.2 deletion and duplication. *Transl. Psychiatry* **9**, 1–8 (2019).

59. Maillard, A. M. *et al.* The 16p11.2 locus modulates brain structures common to autism, schizophrenia and obesity. *Mol. Psychiatry* **20**, 140–147 (2015).

60. Qureshi, A. Y. *et al.* Opposing Brain Differences in 16p11.2 Deletion and Duplication Carriers. *J. Neurosci.* **34**, 11199–11211 (2014).

61. Stefansson, H. *et al.* CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* **505**, 361–366 (2014).

62. Silva, A. I. *et al.* Reciprocal White Matter Changes Associated With Copy Number Variation at 15q11.2 BP1-BP2: A Diffusion Tensor Imaging Study. *Biol. Psychiatry* **85**, 563–572 (2019).

63. Ulfarsson, M. O. *et al.* 15q11.2 CNV affects cognitive, structural and functional correlates of dyslexia and dyscalculia. *Transl. Psychiatry* **7**, e1109 (2017).

64. van der Meer, D. van der *et al.* Association of Copy Number Variation of the 15q11.2 BP1-BP2 Region With Cortical and Subcortical Morphology and Cognition. *JAMA Psychiatry* 1–11 (2019) doi:10.1001/jamapsychiatry.2019.3779.

64

65. Warland, A., Kendall, K. M., Rees, E., Kirov, G. & Caseras, X. Schizophrenia-associated genomic copy number variants and subcortical brain volumes in the UK Biobank. *Mol. Psychiatry* **25**, 854–862 (2020).

66. Moreau, C. A. *et al.* Mutations associated with neuropsychiatric conditions delineate functional brain connectivity dimensions contributing to autism and schizophrenia. *Nat. Commun.* **11**, 5272 (2020).

67. Seidlitz, J. *et al.* Transcriptomic and cellular decoding of regional brain vulnerability to neurogenetic disorders. *Nat. Commun.* **11**, 3358 (2020).

68. Purcell, S. M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).

69. Van der Auwera, S. *et al.* No association between polygenic risk for schizophrenia and brain volume in the general population. *Biol. Psychiatry* **78**, e41-42 (2015).

70. Neilson, E. *et al.* Effects of environmental risks and polygenic loading for schizophrenia on cortical thickness. *Schizophr. Res.* **184**, 128–136 (2017).

71. Ranlund, S. *et al.* Associations between polygenic risk scores for four psychiatric illnesses and brain structure using multivariate pattern recognition. *NeuroImage Clin.* **20**, 1026–1036 (2018).

72. Ge, T. *et al.* Multidimensional heritability analysis of neuroanatomical shape. *Nat Commun* **7**, 13291 (2016).

73. van Erp, T. G. M. *et al.* Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Mol Psychiatry* **21**, 585 (2016).

74. Schmaal, L. *et al.* Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. *Mol Psychiatry* **21**, 806–812 (2016).

75. Hibar, D. P. *et al.* Subcortical volumetric abnormalities in bipolar disorder. *Mol Psychiatry* **21**, 1710–1716 (2016).

76. Boedhoe, P. S. W. *et al.* Distinct Subcortical Volume Alterations in Pediatric and Adult OCD: A Worldwide Meta- and Mega-Analysis. *Am J Psychiatry* **174**, 60–69 (2017).

77. Hoogman, M. *et al.* Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *Lancet Psychiatry* **4**, 310–319 (2017).

78. Radonjić, N. V. *et al.* Structural brain imaging studies offer clues about the effects of the shared genetic etiology among neuropsychiatric disorders. *Mol Psychiatry* 1–10 (2021) doi:10.1038/s41380-020-01002-z.

79. Modenato, C. *et al.* Effects of eight neuropsychiatric copy number variants on human brain structure. *Transl. Psychiatry* **11**, 1–10 (2021).

80. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).

81. Marquand, A. F. *et al.* Conceptualizing mental disorders as deviations from normative functioning. *Mol Psychiatry* **24**, 1415–1424 (2019).

82. Dima, D., Modabbernia, A., Papachristou, E., & others. Subcortical volumes across the lifespan: Data from 18,605 healthy individuals aged 3–90 years. *Hum Brain Mapp*.

83. Huguet, G. *et al.* Measuring and Estimating the Effect Sizes of Copy Number Variants on General Intelligence in Community-Based Samples. *JAMA Psychiatry* **75**, 447–457 (2018).

84. Huguet, G. *et al.* Genome-wide analysis of gene dosage in 24,092 individuals estimates that 10,000 genes modulate cognitive ability. *Mol. Psychiatry* 1–14 (2021) doi:10.1038/s41380-020-00985-z.

85. Malhotra, D. & Sebat, J. CNVs: Harbingers of a Rare Variant Revolution in Psychiatric

Genetics. *Cell* **148**, 1223–1241 (2012).

86. Reus, L. M. *et al.* Association of polygenic risk for major psychiatric illness with subcortical volumes and white matter integrity in UK Biobank. *Scientific Reports* vol. 7 (2017).

87. Neilson, E. *et al.* Impact of Polygenic Risk for Schizophrenia on Cortical Structure in UK Biobank. *Biol. Psychiatry* **86**, 536–544 (2019).

88. Alnæs, D. *et al.* Brain Heterogeneity in Schizophrenia and Its Association With Polygenic Risk. *JAMA Psychiatry* **76**, 739–748 (2019).

89. Alemany, S. *et al.* Common Polygenic Variations for Psychiatric Disorders and Cognition in Relation to Brain Morphology in the General Pediatric Population. *J Am Acad Child Adolesc Psychiatry* **58**, 600–607 (2019).

90. Sun, D. *et al.* Large-scale mapping of cortical alterations in 22q11.2 deletion syndrome: Convergence with idiopathic psychosis and effects of deletion size. *Mol. Psychiatry* 1–13 (2018) doi:10.1038/s41380-018-0078-5.

91. Grasby, K. L. *et al.* The genetic architecture of the human cerebral cortex. *Science* **367**, (2020).

92. Rakic, P. The radial edifice of cortical architecture: From neuronal silhouettes to genetic engineering. *Brain Res. Rev.* **55**, 204–219 (2007).

93. Chen, Z. *et al.* Extensive brain structural network abnormality in first-episode treatment-naive patients with schizophrenia: morphometrical and covariation study. *Psychol. Med.* 1–13 (2014) doi:10.1017/S003329171300319X.

94. Poluch, S. & Juliano, S. L. Fine-tuning of neurogenesis is essential for the evolutionary expansion of the cerebral cortex. *Cereb. Cortex N. Y. N 1991* **25**, 346–364 (2015).

95. Deshpande, A. *et al.* Cellular Phenotypes in Human iPSC-Derived Neurons from a Genetic Model of Autism Spectrum Disorder. *Cell Rep.* **21**, 2678–2687 (2017).

96. Suzuki, I. K. Molecular drivers of human cerebral cortical evolution. *Neurosci. Res.* **151**, 1–14 (2020).

97. Fiddes, I. T. *et al.* Human-Specific NOTCH2NL Genes Affect Notch Signaling and Cortical Neurogenesis. *Cell* **173**, 1356-1369.e22 (2018).

98. van der Meer, D. van der *et al.* Association of Copy Number Variation of the 15q11.2 BP1-BP2 Region With Cortical and Subcortical Morphology and Cognition. *JAMA Psychiatry* 1–11 (2019) doi:10.1001/jamapsychiatry.2019.3779.

99. Huguet, G. *et al.* Measuring and Estimating the Effect Sizes of Copy Number Variants on General Intelligence in Community-Based Samples. *JAMA Psychiatry* **75**, 447–457 (2018).

100. Chawner, S. J. R. A. *et al.* Genotype–phenotype associations in children with copy number variants associated with high neuropsychiatric risk in the UK (IMAGINE-ID): a case-control cohort study. *Lancet Psychiatry* **6**, 493–505 (2019).

101. Copeland, W. E. *et al.* Association of Childhood Trauma Exposure With Adult Psychiatric Disorders and Functional Outcomes. *JAMA Netw. Open* **1**, e184493 (2018).

102. Scott, K. M., McLaughlin, K. A., Smith, D. A. R. & Ellis, P. M. Childhood maltreatment and DSM-IV adult mental disorders: comparison of prospective and retrospective findings. *Br. J. Psychiatry J. Ment. Sci.* **200**, 469–475 (2012).

103. Kessler, R. C. *et al.* Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *Br. J. Psychiatry J. Ment. Sci.* **197**, 378–385 (2010).

104. Arseneault, L. *et al.* Childhood trauma and children's emerging psychotic symptoms: A genetically sensitive longitudinal cohort study. *Am. J. Psychiatry* **168**, 65–72 (2011).

105. Rao, U. *et al.* Hippocampal changes associated with early-life adversity and vulnerability to depression. *Biol. Psychiatry* **67**, 357–364 (2010).

106. Opel, N. *et al.* Mediation of the influence of childhood maltreatment on depression

relapse by cortical structure: a 2-year longitudinal observational study. *Lancet Psychiatry* **6**, 318–326 (2019).

107. Busso, D. S. *et al.* Child Abuse, Neural Structure, and Adolescent Psychopathology: A Longitudinal Study. *J. Am. Acad. Child Adolesc. Psychiatry* **56**, 321-328.e1 (2017).

108. Van Dam, N. T., Rando, K., Potenza, M. N., Tuit, K. & Sinha, R. Childhood maltreatment, altered limbic neurobiology, and substance use relapse severity via trauma-specific reductions in limbic gray matter volume. *JAMA Psychiatry* **71**, 917–925 (2014).

109. Gorka, A. X., Hanson, J. L., Radtke, S. R. & Hariri, A. R. Reduced hippocampal and medial prefrontal gray matter mediate the association between reported childhood maltreatment and trait anxiety in adulthood and predict sensitivity to future life stress. *Biol. Mood Anxiety Disord.* **4**, 12 (2014).

110. Boyle, C. A. *et al.* Trends in the prevalence of developmental disabilities in US children, 1997-2008. *Pediatrics* **127**, 1034–1042 (2011).

111. Jacquemont, S. *et al.* Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature* **478**, 97–102 (2011).

112. Cárdenas-de-la-Parra, A. *et al.* Developmental trajectories of neuroanatomical alterations associated with the 16p11.2 Copy Number Variations. *NeuroImage* **203**, 116155 (2019).

113. Jalbrzikowski, M. *et al.* Categorical versus dimensional approaches to autism-associated intermediate phenotypes in 22q11.2 microdeletion syndrome. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2**, 53–65 (2017).

114. Schaer, M. *et al.* Deviant trajectories of cortical maturation in 22q11.2 deletion syndrome (22q11DS): a cross-sectional and longitudinal study. *Schizophr. Res.* **115**, 182–190 (2009).

115. Ching, C. R. K. *et al.* Mapping Subcortical Brain Alterations in 22q11.2 Deletion

Syndrome: Effects of Deletion Size and Convergence With Idiopathic Neuropsychiatric Illness. *Am. J. Psychiatry* appi.ajp.2019.19060583 (2020) doi:10.1176/appi.ajp.2019.19060583.

116.   Bearden, C. E. *et al.* Alterations in midline cortical thickness and gyrification patterns mapped in children with 22q11.2 deletions. *Cereb. Cortex N. Y. N 1991* **19**, 115–126 (2009).

117.   Jalbrzikowski, M. *et al.* Structural abnormalities in cortical volume, thickness, and surface area in 22q11.2 microdeletion syndrome: Relationship with psychotic symptoms. *NeuroImage Clin.* **3**, 405–415 (2013).

118.   Rogdaki, M. *et al.* Magnitude and heterogeneity of brain structural abnormalities in 22q11.2 deletion syndrome: a meta-analysis. *Mol. Psychiatry* **25**, 1704–1717 (2020).

119.   Goodkind, M. *et al.* Identification of a common neurobiological substrate for mental illness. *JAMA Psychiatry* **72**, 305–315 (2015).

120.   Oldham, S. & Fornito, A. The development of brain network hubs. *Dev. Cogn. Neurosci.* **36**, 100607 (2018).

121.   Ripke, S. *et al.* Biological Insights From 108 Schizophrenia-Associated Genetic Loci. *Nature* **511**, 421–427 (2014).

122.   Lee, S. H. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45**, 984–994 (2013).

123.   Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Med.* **12**, e1001779 (2015).

124.   NIH's Adolescent Brain Cognitive Development (ABCD) Study. *Alcohol Res. Curr. Rev.* **39**, 97 (2018).

Article study 1

Effects of eight neuropsychiatric copy number variants on human brain structure

Modenato Claudia[1]* MSc, Kuldeep Kumar[2]* PhD, Moreau Clara[2] PhD, Martin-Brevet Sandra[1] PhD, Huguet Guillaume[2] PhD, Schramm Catherine[2] PhD, Jean-Louis Martineau[2] MSc, Martin Charles-Olivier[2] PhD, Younis Nadine[2] BSc, Tamer Petra[2] BSc, Douard Elise[2] MSc, Thébault-Dagher Fanny[2] BSc, Côté V.[2] BSc, Charlebois A.R.[2] MSc, Deguire F.[2] MSc, Maillard Anne M.[3] PhD, Rodriguez-Herreros Borja[3] PhD, Pain Aurèlie[3] MSc, Richetin Sonia[3] MSc, 16p11.2 European Consortium, Simons Searchlight Consortium, Melie-Garcia Lester[4] PhD, Kushan Leila MSc[5], Silva Ana I.[6,7] PhD, van den Bree Marianne B.M.[7,8,9] PhD, Linden David E.J.[6,7,9] MD, Owen Michael J.[7,8] MD, PhD, Hall Jeremy[7,8,9] MD, PhD, Lippé Sarah[2] PhD, Chakravarty Mallar[10] PhD, Bzdok Danilo[11,12] MD, PhD, Bearden Carrie E. PhD[5], Draganski Bogdan MD**[1, 13] and Jacquemont Sébastien MD**[2]

[1] LREN - Department of clinical neurosciences, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Switzerland;

[2] Centre de recherche CHU Sainte-Justine and University of Montréal, Canada;

[3] Service des Troubles du Spectre de l'Autisme et apparentés, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Switzerland;

[4] Applied Signal Processing Group (ASPG), Swiss Federal Institute Lausanne (EPFL), Lausanne, Switzerland;

[5] Semel Institute for Neuroscience and Human Behavior, Departments of Psychiatry and Biobehavioral Sciences and Psychology, UCLA, Los Angeles, USA;

[6] School for Mental Health and Neuroscience, Maastricht University, Netherlands;

[7] MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, United Kingdom;

[8] Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, United Kingdom;

[9] Neuroscience and Mental Health Research Institute, Cardiff University, Cardiff, United Kingdom;

[10] Douglas Research Centre, McGill University, Montréal, Canada;

[11] Department of Biomedical Engineering, McConnell Brain Imaging Centre; Montreal Neurological Institute, McGill University, Montréal, Canada;

[12] Mila - Quebec Artificial Intelligence Institute, Montréal, Canada

[13] Neurology Department, Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Corresponding author: Sébastien Jacquemont (Affiliation: Centre de recherche CHU Sainte-Justine and University of Montréal, Canada; Postal address: Centre hospitalier universitaire Sainte-Justine, 3175, Chemin de la Côte Sainte-Catherine, H3T 1C5, Montréal, Canada; Email: sebastien.jacquemont@umontreal.ca; Phone: + 1 514 345-4931)

Running title: Effects of 8 neuropsychiatric copy number variants

Abstract

Many Copy Number Variants (CNVs) confer risk for the same range of neurodevelopmental symptoms and psychiatric conditions including autism and schizophrenia. Yet, to date neuroimaging studies have typically been carried out one mutation at a time, showing that CNVs have large effects on brain anatomy. Here, we aimed to characterize and quantify the distinct brain morphometry effects and latent dimensions across 8 neuropsychiatric CNVs.

We analyzed T1-weighted MRI data from clinically and non-clinically ascertained CNV carriers (deletion/duplication) at the 1q21.1 (n=39/28), 16p11.2 (n=87/78), 22q11.2 (n=75/30), and 15q11.2 (n=72/76) loci as well as 1296 non-carriers (controls). Case control contrasts of all examined genomic loci demonstrated effects on brain anatomy, with deletions and duplications showing mirror effects at the global and regional levels. Although CNVs mainly showed distinct brain patterns, principal component analysis (PCA) loaded subsets of CNVs on two latent brain dimensions, which explained 32 and 29% of the variance of the 8 Cohen's d maps. The cingulate gyrus, insula, supplementary motor cortex, and cerebellum were identified by PCA and multiview pattern learning as top regions contributing to latent dimension shared across subsets of CNVs.

The large proportion of distinct CNV effects on brain morphology may explain the small neuroimaging effect sizes reported in polygenic psychiatric conditions. Nevertheless, latent gene brain morphology dimensions will help subgroup the rapidly expanding landscape of neuropsychiatric variants and dissect the heterogeneity of idiopathic conditions.

Introduction

Genomic copy number variants (CNVs) are deletions or duplications of DNA segments of more than 1000 base pairs. Rare CNVs with large effects have been associated with a range of often overlapping developmental psychiatric phenotypes and conditions, including autism spectrum disorder (ASD) and schizophrenia (SZ) [1–4]. A looming question in psychiatric genetics pertains to the underlying basis of polygenicity: How do different variants lead to risk for the same psychiatric condition?

Some of the most frequent risk factors for neuropsychiatric disorders identified in pediatric clinics include CNVs at the 22q11.2, 16p11.2, 1q21.1, and 15q11.2 genomic loci [5,6]. They affect the dosage of 60, 29, 12 and 4 genes, respectively [7–9]. The largest increases in risk for SZ have been documented for the 22q11.2 deletion (30 to 40-fold) followed by 16p11.2 duplication (10-fold), 1q21.1 deletion and 15q11.2 deletion (1.5-fold) [2]. ASD risk is highest for 16p11.2 deletions and duplications (10-fold) followed by 1q21.1 duplications and 22q11.2 duplications (3 to 4-fold) [1,2,10–13]. The nature and specificity of CNV effects on cognitive and behavioral traits is an area of intense investigation. All CNVs studied to date affect cognition to varying degrees and a broad range of cognitive functions [14,15]. A recent study found that the range of affected traits was broadly similar for 13 CNVs at 8 loci and specific genotypes accounted for a low proportion of phenotypic variance [3]. These variants are therefore opportunities to investigate brain phenotypes conferring high-risk for mental illness.

Neuroimaging studies have only been performed for a few CNVs. Robust effects on total and regional brain volumes, cortical thickness (CT) and surface area (SA), have been reported in 22q11.2 [12,13,16], 16p11.2 BP4-5 [17–19], and 15q11.2 CNVs [20–23]. Opposing effects on global and-

or regional brain volumes between deletions and duplications were observed for 16p11.2 [19], 22q11.2 [16], 1q21.1 [24] and 15q11.2 [20] loci (hereafter referred to as "mirror effects").

Neuroanatomical alterations associated with 16p11.2 and 22q11.2 show overlap with those observed in idiopathic ASD and SZ [17–19,21,25]. Finally, most of the effects are observed irrespective of psychiatric diagnoses and symptoms [12], suggesting that the final clinical outcome may result from the effect of CNVs and additional factors.

Neuroimaging studies across genomic variants are scarce. An investigation of 49 unaffected carriers of SZ-associated CNVs across 5 genomic loci in the UK biobank showed smaller volumes of the thalamus, hippocampus, and nucleus accumbens [26]. Functional connectivity similarities have also been demonstrated between 16p11.2 and 22q11.2 deletions as well as with idiopathic ASD and SZ [27]. Alternatively, a recent study suggests a relatively distinct association between neuroimaging alterations and 6 different CNVs [28].

In this study, we aimed to characterize shared and distinct neuroanatomical alterations associated with 8 CNVs at 4 genomic loci. We analyzed high-resolution structural brain scans from the largest multi-site dataset of CNV carriers (n=484, of which 87 have not yet been published) and controls (n=1296) to date. Different approaches were implemented, from simple case-control contrasts to one-view and multi-view multivariate pattern learning [29,30]. First, we compared brain morphometry features associated with each deletion and duplication using univariate linear models. Second, we quantified the shared variation of brain morphometry associated with 8 CNVs using Principal Component Analysis (PCA). To complement this single-view approach, a multi-view pattern-learning algorithm was carried out for the joint analysis of genetic and morphometry brain data, to identify latent 'gene-morphometry dimensions' (Canonical Correlation Analysis, CCA). Primary analyses were

performed using VBM for consistency with previous studies [19]. In addition, we carried out the same multivariate analyses using freesurfer derived cortical surface area and thickness to ensure that shared variation was not limited to one neuroimaging modality or analytical pipeline.

Methods

Participants: Deletions and duplications carriers' neuroimaging data included in the study were selected on the following breakpoints (hg 19): 16p11.2 (BP4-5, 29.6-30.2MB), 1q21.1 (Class I, 146.4-147.5MB & II, 145.3-147.5MB), 22q11.2 (BPA-D, 18.8-21.7MB) and 15q11.2 (BP1-2, 22.8-23.0MB), together with control individuals not carrying any CNVs at these loci (Table 1, Supplementary Table 1 and supplementary materials). Signed consents were obtained from all participants or legal representatives prior to the investigation. Of note, data of 87 CNV carriers have never been published. Clinically ascertained CNV carriers were recruited as either probands referred for genetic testing, or as relatives. Controls were either non-carriers within the same families or individuals from the general population. We pooled data from 5 cohorts. CNVs from non-clinical populations were identified in the UK Biobank [31,32].

MRI data: Details for methods and analyses are provided in supplementary material and Supplementary Methods 1 to 8. Data sample included T1-weighted (T1w) images at 0.8 - 1 mm isotropic resolution across all sites. Population description is available in Table 1 and Supplementary Table 1.

Data quality check: All data included in the analysis were quality checked by the same researcher (CM). A total of 107 structural brain scans from carriers and controls were excluded from further analysis based on visual inspection that identified significant artifacts compromising accurate tissue classification and boundary detection (supplementary materials).

MRI data processing: Data for Voxel-Based Morphometry were preprocessed and analysed with SPM12 (http://www. fil.ion.ucl.ac.uk/spm/software/spm12/) [33–35] running under

7

MATLAB R2018b (https://www.mathworks.com/products/new_products/release2018b.html). For surface-based feature extraction, we used FreeSurfer 5.3.0 (http://surfer.nmr.mgh.harvard.edu [36,37]). Quality control was performed using standardized ENIGMA quality control procedures (http://enigma.ini.usc.edu/protocols/imaging-protocols/).

Statistical analysis for global brain measures: Global brain aggregate measures (Total Intracranial Volume (TIV), total Gray Matter volume (GM), total surface area (SA), and mean Cortical Thickness (CT)) were adjusted for age, age^2, and sex as fixed effects and scanning site as random factor. Non-clinically ascertained subjects from the UKBB are on average 30 years older than the clinically ascertained subjects. Because of this age difference we used age matched control groups for univariate analysis. Global measure z-scores for each CNV for clinically and non-clinically ascertained CNVs were calculated using 331 and 965 controls, respectively. All statistical analyses were performed in R, version 3.4.4 (https://www.r-project.org/), or in MatlabR2018b.

Voxel-based measures and statistical analyses: We performed whole-brain voxel-based analysis testing for voxel-wise volume differences within the mass-univariate analysis framework implemented in SPM (Supplementary Method 4). Cohen's d (i.e. effect size) [38] maps were obtained by converting SPM T-maps using the CAT12 toolbox for SPM (http://www.neuro.uni-jena.de/cat/).

Surface-based measures and statistical analyses: In parallel to VBM, we used surface-based GLM-based analysis to test differences in CT and SA (SurfStat toolbox [39]).

Neuromorphometrics and Desikan parcellations: Parcellation into regions of interest (ROIs) was performed using neuromorphometric atlas (http://www.neuromorphometrics.com/) for

8

grey matter volume (130 ROIs excluding white matter ROIs), and using Desikan parcellation [37] for FreeSurfer derived CT and SA (68 ROIs).

Comparison of ranked Cohen's d maps across CNVs: To adjust for the unequal power to detect alterations across different CNV groups, which have different sample and effect sizes, we ranked the Cohen's d values of all voxels (/vertices) for each statistical maps (CNV versus controls contrast). We then tested for spatial overlap between maps across CNVs after thresholding the tails of the distribution at the 15th & 85th quantiles. Dice index was calculated using publicly available Matlab scripts (https://github.com/rordenlab/spmScripts).

Null hypothesis testing using spin permutations and label shuffling: We used spin permutation and label shuffling [40,41] to calculate empirical p-values for 1) the deletion and duplication convergence pattern and 2) the correlation/dice-index between two maps.

Quantifying shared variation across CNVs using principal components (PC): PCs were derived to quantify shared morphometry variation across CNVs. We used Cohen's d values of 130 neuroanatomical GM regions (neuromorphometrics atlas) of 8 CNVs as input-variables (z-scored Cohen's d contrasts adjusted for total GM and nuisance variables; FactoMineR package in R). The variance explained (coefficient of determination, R-squared) for each CNV-associated Cohen's d map by PCs was obtained by running a linear model (lm) in R; with PC1 and PC2 as independent explanatory variables and the CNV Cohen's d map as a dependent variable.

Jointly modeling of gene-morphology dimensions using canonical correlation analysis (CCA): We re-purposed CCA to simultaneously model the shared and distinct impact of the CNVs in causing distributed alterations in brain morphometry (130 grey matter regions) [29,30]. This principled doubly-multivariate approach, widely used in neuroimaging studies [29,30], was

9

performed to identify modes of coherent co-variation that jointly characterize how CNVs and patterns of regional volumes systematically co-occur across subjects. Henceforth, we refer to the ensuing modes of co-variation as 'CCA dimensions' or 'gene-morphology dimensions'.

Results

1. CNV effects on global brain morphometry

Deletions and duplications of each genomic loci showed opposing effects on one or more global metrics: total intracranial volume (TIV), total grey matter volume (GM) total surface area (SA) or mean cortical thickness (CT) (Figure 1, Supplementary Table 2). The directionality of global effects differed across loci (Figure 1a-c.). Effects on GM and SA were less pronounced once adjusted for TIV (Supplementary Figure 1).

2. Overlapping deletion effects on regional morphometry

Whole-brain VBM analyses contrasting each deletion and duplication group with controls showed mostly distinct brain patterns across CNVs (Figure 2a, c, e, Supplementary Table 3). To investigate potential overlap across the 4 genomic regions, we ranked Cohen's d maps and overlapped voxels with similar rankings. Using a threshold for voxels with Cohen's d <15th and >85th percentiles separately (Figure 3c, e, g, i), we observed significant overlap between deletions (p-value$_{SHUFFLE}$<10e-4, Figure 3a). Volumes of the middle and anterior cingulate extending to the supplementary motor cortex and of the cerebellum were decreased in all deletions while volume was increased in the thalamus (Figure 3a).

Sensitivity analyses tested the effect of ascertainment and control groups: 1) We recomputed the deletion convergence map using 1q21.1 deletion carriers from UK Biobank instead of those clinically ascertained (Table 1). The new deletion convergence map was similar to the initial one presented above with a dice index of 39.4% (p-value$_{SPIN}$< 10e-4); 2) We excluded all subjects with autism, schizophrenia, or other psychiatric diagnoses. Again, this did not change the overlap (Supplementary Figure 2); 3) We tested the effects of the control group by

11

recomputing contrasts only using controls from the same site (instead of the initial ANOVA pooling all controls together and controlling for site). This again did not alter the convergence maps (Supplementary Figure 3). Finally, we performed the same analysis using Freesurfer derived SA and CT measures. We also identified spatial overlaps but regions identified were different especially for CT (Supplementary Table 4 & Figure 4). Overlap maps are provided in Supplementary Figures 5-8 & Tables 5-6.

3. Overlapping duplication effects on regional morphometry

Contrasts computed for duplications (Figure 2b, d, f) showed smaller effect sizes compared to deletions. The same analysis using Cohen's d values <15th and >85th percentiles (Figure 3d, f, h, j) demonstrated spatial overlap across all 4 duplications (p-value$_{SHUFFLE}$<10e-4, Figure 3b). The resulting pattern was mainly distinct from the one observed in deletions and was characterized by smaller volumes in anterior insula and frontal operculum, and larger volumes in the middle cingulate gyrus and supplementary motor cortex compared to controls.

Sensitivity analysis testing the effect of clinical ascertainment, psychiatric diagnoses, control groups and volume versus Freesurfer derived measures demonstrated that results were robust (Supplementary Figures 2-8).

The deletion/duplication ratio of Cohen's d distributions ranged from 1.24 to 2-fold across the 4 genomic loci (F-test, p<10e^{-16}, Figure 3c-j, Supplementary Table 7). Similar effect-size ratios were also observed for SA alterations (Supplementary Table 7), except for the 15q11.2 locus.

We tested opposing (mirror) effects on VBM contrast maps between deletion and duplications. The strongest anticorrelation of Cohen's d values was observed for 16p11.2 (p-value$_{SPIN}$<10e-4) followed by 15q11.2 (p-value$_{SPIN}$<10e-4), 1q21.1 (p-value$_{SPIN}$<0.033) and
12

22q11.2 (p-value$_{SPIN}$<0.038) (Supplementary Figure 9 and Tables 8-10). Mirror effects were observed in clinically and non-clinically ascertained CNV carriers, as well as for SA at all 4 genomic loci but not for CT (Supplementary Tables 8-10). Hence mirror effects were observed in global metrics and, independently, in regional alterations.

4. Quantifying distinct and shared effects on brain morphometry associated with 8 CNVs

We performed a multivariate PCA based on Cohen's d profiles obtained from contrasts between the 8 CNV groups and controls (using 130 neuromorphometric regional volumes, Supplementary Table 11). The first 2 components explained 31.8 and 28.7 % of the variance of Cohen's d maps, respectively. The third component dropped to 13.8% and was therefore not investigated further.

Deletions and duplications at each genomic loci showed opposite loading on PC1 or PC2 (Figure 4c). Regions with the highest loadings on PC1 and PC2 were also those identified in the convergence maps presented above: in particular the middle cingulate gyrus and the supplementary motor cortex. Anterior and posterior insula, cerebellum, fusiform gyrus and thalamus were also top regions altered across subsets of CNVs (Figure 4a-b and Supplementary Table 12). The variance explained by both components for each CNV's Cohen's d map ranged from 27% to 82% (Figure 4d). Finally, we performed the same analysis using Freesurfer-derived SA and CT measures which also provided latent dimensions with comparable variance explained, opposing loadings for deletions and duplications of each genomic loci (Supplementary Figure 10). However CNV loadings differ across brain morphometry metrics.

5. Gene - morphology dimensions across 8 CNVs

As a next step, we performed a multi-view pattern-learning analysis, jointly analyzing the genetic and morphometry brain data. This doubly-multivariate method allowed testing whether shared dimensions could be identified in a data driven approach, without performing any individual contrast. We interrogated 2 hypotheses: 1) CNVs show levels of shared brain effects at the morphometry level and 2) deletions and duplications show opposing effects. We investigated the same 130 regional volumes in 484 carriers of CNVs at 4 genomic loci. To test hypothesis 2), deletions and duplications were coded as opposing gene dosage. CCA confirmed both hypotheses by identifying 2 significant 'gene-morphometry dimensions' (r=0.84, 0.79, p-value<0.05, Figure 4e-f). Regional brain contributions to canonical dimension 1 and 2 were well correlated with those of PC2 and 1 respectively (r=0.83, r=-0.81).

Top ranking brain regions contributing to either of the 2 CCA dimensions of morphological variation included supplementary motor cortex, posterior and anterior insula, middle cingulate gyrus, calcarine cortex, cuneus and accumbens (Supplementary Figure 11 and Supplementary Table 13). 16p11.2 and 22q11.2 preferentially contributed to dimension 1 and 2 respectively, and 1q21.1 loaded similarly on both dimensions. 15q11.2 CNVs showed the smallest loadings on both dimensions (Figure 4e).

Sensitivity analyses are detailed in supplementary material (Supplementary Figures 12-16 and Tables 14-15).

Discussion

Here, in the largest cross-CNV neuroimaging study to date, we tested potentially shared effects of 8 neuropsychiatric CNVs on brain morphometry. CNVs showed a combination of distinct and shared profiles of brain alterations, as demonstrated by the spatial overlap of Cohen's d maps across deletions and duplications. A multivariate approach (PCA) quantified distinct and shared alterations across subsets of CNVs and identified 2 latent dimensions explaining 31.8 and 28.7 % of Cohen's d map's variance. A second multivariate approach (CCA), jointly analyzing genetic and morphometry data, confirmed the latent CNV-brain dimensions identified by PCA. Genomic loci contributed to the latent CCA dimensions in proportion to their effect-sizes. Even for small effect size deletions at the 1q21.1 and 15q11.2 loci, the PCA components explained between 43 and 65% of their Cohen's d profile. All three approaches -spatial overlap, CCA and PCA- identified a similar set of regions altered by CNVs including cingulate gyrus and supplementary motor cortex.

Distinct and shared effects of CNVs

Our results show that two-thirds of the average CNV effects on brain morphometry are distinct. This is consistent with a recent study showing relative specificity of association between brain patterns of gene expression and patterns of cortical anatomy changes across 6 CNVs and chromosomal aneuploidies [28]. One third of the effects on brain morphometry are shared as demonstrated by latent gene-morphology dimensions identified across subsets of CNVs. There is no single dimension explaining CNV effects. Instead, subsets of CNVs load on either dimension, which may suggest similar brain mechanisms within subgroups of CNV. Yet CNVs within subgroups were not characterized by the same risk for ASD or SCZ.

These results have implications for our conceptualization of polygenic psychiatric conditions. Indeed, studies estimate that 70 to 100% of any 1-MB window in the human genome encompasses variants (including CNVs) contributing to increased risk for schizophrenia and autism [4,42]. Gene-morphology dimensions alone, can't explain the fact that subgroups of CNVs are associated with a similar range of behavioral symptoms [43], and psychiatric disorders [1,2,4,44]. In fact, the large proportion of distinct CNV-neuroimaging effects suggests that a broad diversity of brain mechanisms increase risk for autism and schizophrenia. Extreme examples include CNVs associated with opposing loadings on the same latent gene-morphology dimension while increasing risk for the same psychiatric condition (ie. 16p11.2 deletions, duplications and autism). The presence of such genomic variants in studies of ASD and SZ may explain heterogeneity and small neuroimaging effect sizes [45,46]. Why opposing effects on the same latent brain dimension increase risk for the same psychiatric condition is an unsolved question. Further observations on a broad variety of genomic variants are required to address this question.

Brain hubs vulnerable to altered gene dosage

Insula, cingulate, fusiform gyrus, and hippocampus are regions showing alterations across schizophrenia, bipolar disorders, major depression, and obsessive compulsive disorders [45,47]. The cingulate, insula, and fusiform gyrus were also among regions markedly altered across 8 CNVs. CNVs have either negative or positive effects on these brain regions, however, the number of CNVs included in this study did not allow us to associate the directionality of these effects with phenotypic traits. Alterations of the cingulate cortex have been associated with genetic and environmental risk for schizophrenia [48]. The supplementary motor cortex has been shown to play a critical role in 16p11.2, 22q11.2 CNVs as well as autism and schizophrenia

by functional connectivity studies, but not by cross-diagnostic neuroimaging structural studies [49,50]. Several cerebellar regions (vermis lobule VIII-X and cerebellar cortex) are highly sensitive to CNVs, which may be due to the cerebellum's protracted development [51]. The cerebellum has either been excluded or not reported by cross-disorder structural neuroimaging studies, but volume alterations have been associated with autism and schizophrenia separately [52,53]. Multiple genetic mouse models of autism, as well as Down Syndrome, also show abnormal cerebellar development [54]. The same level of spatial overlap was observed for SA and CT but implicated mostly distinct sets of brain regions. This is in line with the distinct genetic contributions previously demonstrated for these cortical metrics [55].

Dissociation between global and regional effects

Results suggest that global and local effects may be mechanistically unrelated. 1q21.1 deletions and duplications highlight the contrast between very large effects on global measures, with small regional effects once adjusted for total GM. Dissociation is also observed between the directionalities of global and regional effects: all deletions are associated with a smaller cingulate and supplementary motor cortex volume irrespective of their effect on TIV and GM. Animal studies have proposed mechanisms for global [8,56], but not regional effects of CNVs.

Limitations

Multiple sites included in the study may have introduced noise, but previous studies have shown that site effects do not influence the neuroanatomical patterns associated with CNVs at the 16p11.2, 22q11.2, and 15q11.2 loci [12,19,23]. While shared variation could have been influenced by clinical ascertainment or psychiatric diagnoses, our sensitivity analyses showed that this is not the case. Effect of medication on CNVs brain alterations could not be

investigated in the current study as medication information was not available for the whole dataset. We were underpowered to properly investigate potential sex related effects of 1q21.1 and 15q11.2 on brain morphometry. Of note, previous neuroimaging studies of large 22q11.2 and 16p11.2 samples were unable to identify any sex related effects [19,25].

15q11.2 deletions and duplications have small effect sizes and larger samples would improve the accuracy of the brain morphometry signature. Systematic analysis through the two most widespread computational neuroanatomy frameworks (voxel-based and surface-based) shows that effects could not be attributed to the processing pipeline. Extending our approach to the rapidly expanding number of rare genomic variants associated with psychiatric disorders is required to draw a robust conclusion on the distinct and shared effects of CNVs on brain structure.

Conclusions

The simultaneous analyses and comparisons of several genomic variants demonstrate distinct CNV-associated alteration profiles as well as shared latent gene-morphology dimensions relevant to subsets of CNVs. Large proportions of distinct effects may provide some answers to the small neuroimaging effect sizes reported in idiopathic psychiatric conditions. The mechanisms underlying the identified latent dimensions remain unknown and pathway convergence may occur early on at the transcriptome and protein level, or at later stages (ie. brain architecture or behavior). The hotly debated omnigenic model postulates that convergence may occur at early stages due to highly interconnected cell regulatory networks [57]. These approaches may help subgroup genomic variants based on their morphometry signature and dissect the heterogeneity of psychiatric conditions.

18

Acknowledgments

sites, as well as the Simons Searchlight Consortium. We appreciate obtaining access to imaging and phenotypic data on SFARI Base. Approved researchers can obtain the Simons Searchlight population dataset described in this study by applying at https://base.sfari.org. We are grateful to all families who participated in the 16p11.2 European Consortium.

Disclosures

MvdB reports grants from Takeda Pharmaceuticals, outside the submitted work. All other authors reported no biomedical financial interests or potential conflicts of interest.

Author contributions

C.Mod., K.K., B.D. and S.J. designed the study, analysed imaging data and drafted the manuscript.

Analyses: C.Mod. and K.K. did all the preprocessing and analysis of neuroimaging data, DB provided scripts and mentored the CCA analysis. C.Mor., C.E.B. and D.B. contributed in result interpretation and in the editing of the manuscript.

Data collection: C.Mod., A.M., A.P., S.R. and S.M-B. recruited and scanned participants in the 16p11.2 European Consortium. S.L., C.O.M., N.Y., P.T., E.D., F. T-D, V.C., A.R.C., F.D. recruited and scanned participants in the Brain Canada cohort. L.K. collected and provided the data for the UCLA cohort. D.E.J.L., M.J.O., M.B.M. V.d.B., J.H. and A.I.S., provided the data for the Cardiff cohort.

All authors provided feedback on the manuscript.

Article Informations

Each cohort and corresponding study received approval from their local Institutional review board and this study was approved by the Institutional review board of the CHU Ste Justine
20

Annapurna Poduri, PhD, Nicholas J. Pojman, BS, Ken Porche, MS, Monica B. Proud, MD, Saba Qasmieh, BA, Melissa B. Ramocki, MD, PhD, Beau Reilly, PhD, Timothy P. L. Roberts, PhD, Dennis Shaw, MD, Tuhin Sinha, PhD, Bethanny Smith-Packard, MS, CGC, Anne Snow Gallagher, PhD, Vivek Swarnakar, PhD, Tony Thieu, BA, MS, Christina Triantafallou, PhD, Roger Vaughan, PhD, Mari Wakahiro, MSW, Arianne Wallace, PhD, Tracey Ward, BS, Julia Wenegrat, MA, and Anne Wolken, BS. European 16p11.2 Consortium principal investigator Sébastien Jacquemont. Members of the European 16p11.2 Consortium include the following: Addor Marie-Claude, Service de génétique médicale, Centre Hospitalier Universitaire Vaudois, Lausanne University, Switzerland; Andrieux Joris, Institut de Génétique Médicale, CHRU de Lille, Hopital Jeanne de Flandre, France; Arveiler Benoît, Service de génétique médicale, CHU de Bordeaux- GH Pellegrin, France; Baujat Geneviève, Service de Génétique Médicale, CHU Paris - Hôpital Necker-Enfants Malades, France; Sloan-Béna Frédérique, Service de médecine génétique, Hôpitaux Universitaires de Genève - HUG, Switzerland; Belfiore Marco, Service de génétique médicale, Centre Hospitalier Universitaire Vaudois, Lausanne University, Switzerland; Bonneau Dominique, Service de génétique médicale, CHU d'Angers, France; Bouquillon Sonia, Institut de Génétique Médicale, Hopital Jeanne de Flandre, Lille, France; Boute Odile, Hôpital Jeanne de Flandre, CHRU de Lille, Lille, France; Brusco Alfredo, Genetica Medica, Dipartimento di Scienze Mediche, Università di Torino, Italy; Busa Tiffany, Département de génétique médicale, CHU de Marseille, Hôpital de la Timone, France; Caberg Jean- Hubert, Centre de génétique humaine, CHU de Liège, Belgique; Campion Dominique, Service de psychiatrie, Centre hospitalier de Rouvray, Sotteville lès Rouen, France; Colombert Vanessa, Service de génétique médicale, Centre Hospitalier Bretagne Atlantique CH Chubert- Vannes, France; Cordier Marie-Pierre, Service de génétique clinique, CHU de Lyon, Hospices Civils de Lyon,

France; David Albert, Service de Génétique Médicale, CHU de Nantes, Hôtel Dieu, France; Debray François-Guillaume, Service de Génétique Humaine, CHU Sart Tilman - Liège, Belgique; Delrue Marie-Ange, Service de génétique médicale, CHU de Bordeaux, Hôpital Pellegrin, France; Doco-Fenzy Martine, Service de Génétique et Biologie de la Reproduction, CHU de Reims, Hôpital Maison Blanche, France; Dunkhase- Heinl Ulrike, Department of Pediatrics, Aabenraa Hospital, Sonderjylland, Denmark; Edery Patrick, Service de génétique clinique, CHU de Lyon, Hospices Civils de Lyon, France; Fagerberg Christina, Department of Clinical Genetics, Odense University hospital, Denmark; Faivre Laurence, Centre de génétique, Hôpital d'Enfants, CHU Dijon Bourgogne - Hôpital François Mitterrand, France; Forzano Francesca, Ambulatorio di Genetica Medica, Ospedali Galliera di Genova, Italy and Clinical Genetics Department, 7th Floor Borough Wing, Guy's Hospital, Guy's & St Thomas' NHS Foundation Trust, Great Maze Pond, London SE1 9RT, UK; Genevieve David, Département de Génétique Médicale, Maladies Rares et Médecine Personnalisée, service de génétique clinique, Université Montpellier, Unité Inserm U1183, CHU Montpellier, Montpellier, France; Gérard Marion, Service de Génétique, CHU de Caen, Hôpital Clémenceau, France; Giachino Daniela, Genetica Medica, Dipartimento di Scienze Cliniche e Biologiche, Università di Torino, Italy; Guichet Agnès, Service de génétique, CHU d'Angers, France; Guillin Olivier, Service de psychiatrie, Centre hospitalier du Rouvray, Sotteville lès Rouen, France; Héron Delphine, Service de Génétique clinique, CHU Paris-GH La Pitié Salpêtrière-Charles Foix - Hôpital Pitié Salpêtrière, France; Isidor Bertrand, Service de Génétique Médicale, CHU de Nantes, Hôtel Dieu, France; Jacquette Aurélia, Service de Génétique clinique, CHU Paris-GH La Pitié Salpêtrière-Charles Foix - Hôpital Pitié-Salpêtrière, France; Jaillard Sylvie, Service de Génétique Moléculaire et Génomique – Pôle biologie, CHU de Rennes, Hôpital Pontchaillou, France; Journel Hubert, Service de génétique

médicale, Centre Hospitalier Bretagne Atlantique CH Chubert- Vannes, France; Keren Boris, Centre de Génétique Moléculaire et Chromosomique, CHU Paris-GH La Pitié Salpêtrière-Charles Foix - Hôpital Pitié-Salpêtrière, France; Lacombe Didier, Service de génétique médicale, CHU de Bordeaux-GH Pellegrin, France; Lebon Sébastien, Pediatric Neurology Unit, Department of Pediatrics, Lausanne University Hospital, Lausanne, Switzerland; Le Caignec Cédric, Service de Génétique Médicale - Institut de Biologie, CHU de Nantes, France; Lemaître Marie-Pierre, Service de Neuropédiatrie, Centre Hospitalier Régional Universitaire de Lille, France; Lespinasse James, Service génétique médicale et oncogénétique, Hotel Dieu, Chambéry, France; Mathieu-Dramart Michèle, Service de Génétique Clinique, CHU Amiens Picardie, France; Mercier Sandra, Service de Génétique Médicale, CHU de Nantes, Hôtel Dieu, France; Mignot Cyril, Service de Génétique clinique, CHU Paris-GH La Pitié Salpêtrière-Charles Foix - Hôpital Pitié-Salpêtrière, France; Missirian Chantal, Département de génétique médicale, CHU de Marseille, Hôpital de la Timone, France; Petit Florence, Service de génétique clinique Guy Fontaine, Hôpital Jeanne de Flandre, CHRU de Lille, France; Pilekær Sørensen Kristina, Department of Clinical Genetics, Odense University Hospital, Denmark; Pinson Lucile, Département de Génétique Médicale, Maladies Rares et Médecine Personnalisée, service de génétique clinique, Université Montpellier, Unité Inserm U1183, CHU Montpellier, Montpellier, France; Plessis Ghislaine, Service de Génétique, CHU de Caen, Hôpital Clémenceau, France; Prieur Fabienne, Service de génétique clinique, CHU de Saint-Etienne - Hôpital Nord, France; Raymond Alexandre, Center for Integrative Genomics, Lausanne University, Switzerland; Rooryck-Thambo Caroline, Laboratoire de génétique moléculaire, CHU de Bordeaux-GH Pellegrin, France; Rossi Massimiliano, Service de génétique clinique, CHU de Lyon, Hospices Civils de Lyon, France; Sanlaville Damien, Laboratoire de Cytogénétique Constitutionnelle, CHU de Lyon,

Hospices Civils de Lyon, France; Schlott Kristiansen Britta, Department of Clinical Genetics, Odense University Hospital, Denmark; Schluth-Bolard Caroline, Laboratoire de Cytogénétique Constitutionnelle, CHU de Lyon, Hospices Civils de Lyon, France; Till Marianne, Service de génétique clinique, CHU de Lyon, Hospices Civils de Lyon, France; Van Haelst Mieke, Department of Genetics, University Medical Center Utrecht, Holland; Van Maldergem Lionel, Centre de Génétique humaine, CHRU de Besançon - Hôpital Saint-Jacques, France.

References

1. Sanders, S. J. *et al.* Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* **87**, 1215–1233 (2015).

2. Marshall, C. R. *et al.* Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet* **49**, 27–35 (2017).

3. Chawner, S. J. R. A. *et al.* A Genetics-First Approach to Dissecting the Heterogeneity of Autism: Phenotypic Comparison of Autism Risk Copy Number Variants. *AJP* **178**, 77–86 (2021).

4. Douard, E. *et al.* Effect Sizes of Deletions and Duplications on Autism Risk Across the Genome. *AJP* **178**, 87–98 (2020).

5. Moreno-De-Luca, D. *et al.* Using large clinical data sets to infer pathogenicity for rare copy number variants in autism cohorts. *Mol Psychiatry* **18**, 1090–1095 (2013).

6. Crawford, K. *et al.* Medical consequences of pathogenic CNVs in adults: analysis of the UK Biobank. *Journal of Medical Genetics* **56**, 131–138 (2019).

7. Mefford, H. C. *et al.* Recurrent Rearrangements of Chromosome 1q21.1 and Variable Pediatric Phenotypes. *N Engl J Med* **359**, 1685–1699 (2008).

8. Golzio, C. *et al.* KCTD13 is a major driver of mirrored neuroanatomical phenotypes associated with the 16p11.2 CNV. *Nature* **485**, 363–367 (2012).

9. Jonas, R. K., Montojo, C. A. & Bearden, C. E. The 22q11.2 Deletion Syndrome as a Window into Complex Neuropsychiatric Disorders Over the Lifespan. *Biological Psychiatry* **75**, 351–360 (2014).

10. Zufferey, F. *et al.* A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. *Journal of Medical Genetics* **49**, 660–668 (2012).

11. D'Angelo, D. *et al.* Defining the Effect of the 16p11.2 Duplication on Cognition,

Behavior, and Medical Comorbidities. *JAMA Psychiatry* **73**, 20–30 (2016).

12. Sun, D. *et al.* Large-scale mapping of cortical alterations in 22q11.2 deletion syndrome: Convergence with idiopathic psychosis and effects of deletion size. *Mol Psychiatry* 1–13 (2018) doi:10.1038/s41380-018-0078-5.

13. Niarchou, M. *et al.* Psychiatric disorders in children with 16p11.2 deletion and duplication. *Transl Psychiatry* **9**, 1–8 (2019).

14. Moberg, P. J. *et al.* Neurocognitive Functioning in Patients with 22q11.2 Deletion Syndrome: A Meta-Analytic Review. *Behav Genet* **48**, 259–270 (2018).

15. Gur, R. E. *et al.* Neurocognitive development in 22q11.2 deletion syndrome: comparison with youth having developmental delay and medical comorbidities. *Molecular Psychiatry* **19**, 1205–1211 (2014).

16. Lin, A. *et al.* Mapping 22q11.2 Gene Dosage Effects on Brain Morphometry. *J. Neurosci.* **37**, 6183–6199 (2017).

17. Maillard, A. M. *et al.* The 16p11.2 locus modulates brain structures common to autism, schizophrenia and obesity. *Mol Psychiatry* **20**, 140–147 (2015).

18. Qureshi, A. Y. *et al.* Opposing Brain Differences in 16p11.2 Deletion and Duplication Carriers. *J. Neurosci.* **34**, 11199–11211 (2014).

19. Martin-Brevet, S. *et al.* Quantifying the Effects of 16p11.2 Copy Number Variants on Brain Structure: A Multisite Genetic-First Study. *Biological Psychiatry* **84**, 253–264 (2018).

20. Stefansson, H. *et al.* CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* **505**, 361–366 (2014).

21. Silva, A. I. *et al.* Reciprocal White Matter Changes Associated With Copy Number Variation at 15q11.2 BP1-BP2: A Diffusion Tensor Imaging Study. *Biological Psychiatry*

85, 563–572 (2019).

22. Ulfarsson, M. O. *et al.* 15q11.2 CNV affects cognitive, structural and functional correlates of dyslexia and dyscalculia. *Transl Psychiatry* **7**, e1109 (2017).

23. van der Meer, D. van der *et al.* Association of Copy Number Variation of the 15q11.2 BP1-BP2 Region With Cortical and Subcortical Morphology and Cognition. *JAMA Psychiatry* 1–11 (2019) doi:10.1001/jamapsychiatry.2019.3779.

24. Sønderby, I. E. *et al.* 1q21.1 distal copy number variants are associated with cerebral and cognitive alterations in humans. *Translational Psychiatry* **11**, 1–16 (2021).

25. Sun, D. *et al.* Large-scale mapping of cortical alterations in 22q11.2 deletion syndrome: Convergence with idiopathic psychosis and effects of deletion size. *Mol Psychiatry* 1–13 (2018) doi:10.1038/s41380-018-0078-5.

26. Warland, A., Kendall, K. M., Rees, E., Kirov, G. & Caseras, X. Schizophrenia-associated genomic copy number variants and subcortical brain volumes in the UK Biobank. *Molecular Psychiatry* **25**, 854–862 (2020).

27. Moreau, C. A. *et al.* Mutations associated with neuropsychiatric conditions delineate functional brain connectivity dimensions contributing to autism and schizophrenia. *Nature Communications* **11**, 5272 (2020).

28. Seidlitz, J. *et al.* Transcriptomic and cellular decoding of regional brain vulnerability to neurogenetic disorders. *Nature Communications* **11**, 3358 (2020).

29. Smith, S. M. *et al.* A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat Neurosci* **18**, 1565–1567 (2015).

30. Wang, H.-T. *et al.* Finding the needle in high-dimensional haystack: A tutorial on canonical correlation analysis. *NeuroImage* (2020).

31. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective

28

epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).

32. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine* **12**, e1001779 (2015).

33. Ashburner, J. & Friston, K. J. Unified segmentation. *NeuroImage* **26**, 839–851 (2005).

34. Lorio, S. *et al.* New tissue priors for improved automated classification of subcortical brain structures on MRI. *Neuroimage* **130**, 157–166 (2016).

35. Ashburner, J. A fast diffeomorphic image registration algorithm. *NeuroImage* **38**, 95–113 (2007).

36. Fischl, B., Sereno, M. I. & Dale, A. M. Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *NeuroImage* **9**, 195–207 (1999).

37. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006).

38. Cohen, J. *Statistical power analysis for the behavioral sciences*. (Psychology Press, 1988).

39. Worsley, K. J. *et al.* SurfStat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. *Neuroimage* (2009) doi:10.1016/S1053-8119(09)70882-1.

40. Alexander-Bloch, A. *et al.* On testing for spatial correspondence between maps of human brain structure and function. *Neuroimage* **178**, 540–551 (2018).

41. Reardon, P. K. *et al.* Normative brain size variation and brain shape diversity in humans. *Science* **360**, 1222–1227 (2018).

42. Loh, P.-R. *et al.* Contrasting genetic architectures of schizophrenia and other complex

diseases using fast variance-components analysis. *Nat Genet* **47**, 1385–1392 (2015).

43. Huguet, G. *et al.* Measuring and Estimating the Effect Sizes of Copy Number Variants on General Intelligence in Community-Based Samples. *JAMA Psychiatry* **75**, 447–457 (2018).

44. Chawner, S. J. R. A. *et al.* Genotype–phenotype associations in children with copy number variants associated with high neuropsychiatric risk in the UK (IMAGINE-ID): a case-control cohort study. *The Lancet Psychiatry* **6**, 493–505 (2019).

45. Opel, N. *et al.* Cross-Disorder Analysis of Brain Structural Abnormalities in Six Major Psychiatric Disorders: A Secondary Analysis of Mega- and Meta-analytical Findings From the ENIGMA Consortium. *Biological Psychiatry* (2020) doi:10.1016/j.biopsych.2020.04.027.

46. Bedford, S. A. *et al.* Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and cortical morphometry in autism spectrum disorder. *Mol Psychiatry* **25**, 614–628 (2020).

47. Goodkind, M. *et al.* Identification of a common neurobiological substrate for mental illness. *JAMA Psychiatry* **72**, 305–315 (2015).

48. Tost, H., Champagne, F. A. & Meyer-Lindenberg, A. Environmental influence in the brain, human welfare and mental health. *Nat Neurosci* **18**, 1421–1431 (2015).

49. Moreau, C. *et al.* Neuropsychiatric mutations delineate functional brain connectivity dimensions contributing to autism and schizophrenia. *bioRxiv* 862615 (2019) doi:10.1101/862615.

50. Kebets, V. *et al.* Somatosensory-Motor Dysconnectivity Spans Multiple Transdiagnostic Dimensions of Psychopathology. *Biological Psychiatry* **86**, 779–791 (2019).

51. Sathyanesan, A. *et al.* Emerging connections between cerebellar development, behavior,

and complex brain disorders. *Nat Rev Neurosci* **20**, 298–313 (2019).

52. Moberget, T. *et al.* Cerebellar volume and cerebellocerebral structural covariance in schizophrenia: a multisite mega-analysis of 983 patients and 1349 healthy controls. *Molecular Psychiatry* **23**, 1512–1520 (2018).

53. Traut, N. *et al.* Cerebellar Volume in Autism: Literature Meta-analysis and Analysis of the Autism Brain Imaging Data Exchange Cohort. *Biol. Psychiatry* **83**, 579–588 (2018).

54. Ellegood, J. *et al.* Clustering autism: using neuroanatomical differences in 26 mouse models to gain insight into the heterogeneity. *Mol. Psychiatry* **20**, 118–125 (2015).

55. Grasby, K. L. *et al.* The genetic architecture of the human cerebral cortex. *Science* **367**, (2020).

56. Richter, M. *et al.* Altered TAOK2 activity causes autism-related neurodevelopmental and cognitive abnormalities through RhoA signaling. *Mol Psychiatry* **24**, 1329–1350 (2019).

57. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An expanded view of complex traits: from polygenic to omnigenic. *Cell* **169**, 1177–1186 (2017).

Figures legends

Figure 1: 1q21.1, 16p11.2, 22q11.2 and 15q11.2 exert rich effects on global brain measures.

Legend: Total intracranial volume (a), total surface area (b), total grey matter volume (c) and mean cortical thickness (d) for clinically and non-clinically ascertained CNVs. Z-scores for clinically and non-clinically ascertained CNVs were calculated using 331 and 965 controls respectively, adjusting for age, age^2, sex and site as a random factor. Y axis values are z-scores. X axis are CNV groups. Significant difference between CNV group and corresponding control group is indicated with a star. Horizontal bars with stars show significant differences between deletions and duplications within the same locus. TIV: total intracranial volume, SA: surface area, GM: grey matter, CT: cortical thickness.

Figure 2: Cohen's d maps of VBM regional brain differences in deletion and duplication carriers at the 1q21.1, 16p11.2 and 22q11.2 loci compared to controls.

Legend: Regional brain differences adjusted for total grey matter volume. Left and right columns show results for deletions (a, c, e) and duplication (b, d, f) carriers respectively. Color maps show the significant effects of each CNV, thresholded at q< 0.05 FWE. Color scale represents positive and negative Cohen's d effect sizes were estimated. Linear model were adjusted for sex, linear and quadratic expansion of age and total grey matter volume. 15q11.2 was not displayed because only a few voxels survived family-wise error (FWE) correction. Corresponding maps for Surface Area and Cortical Thickness are reported in Supplementary Figures 4-5.

Figure 3: Spatial overlap across deletions and duplications at 4 genomic loci.

Legend: Spatial overlap across clinically and non-clinically ascertained deletions (a) and duplications (b) at 4 genomic loci shown separately for <15th and >85th percentile of Cohen's d values. Overlap of all four deletions (a) or all four duplications (b) is shown in blue. Overlaps of any combination of three deletions (a) or any combination of three duplications (b) are shown in red. Top ranking Cohen's d values used in (a, b) are presented on the density plots for all eight deletions and duplications: 1q21.1 (c, d), 16p11.2 (e, f), 22q11.2 (g, h), 15q11.2 (i, j). The x axes values of the 8 density plots are Cohen's d. Corresponding maps for surface area and cortical thickness are reported in Supplementary Figures 6-7.

Figure 4: Principal Component Analysis and Canonical Correlation Analysis of brain alterations due to 8 CNVs.

Legend: (a) PCA dimension 1 and 2 regional relevances projected on axial brain slices. The darker the red or blue color, the stronger the positive or negative association with the PCA dimensions. PCA was run on z-scored Cohen's d values, with the 8 CNVs as variables and 130 neuroanatomical GM regions as observations. GM region volumes were adjusted for total grey matter, age, age^2, sex and site. The first 2 components explained respectively 31.77 and 28.66 % of the variance. (b) Loading of 8 CNVs on the 2 PCA dimensions. Values are PC loading magnitudes and represent the contribution of a CNV to the PC. (c) Variance explained (coefficient of determination, R-squared) of each CNV Cohen's d profile by PC1 and PC2. Values and color scale represent the "percent of variance". (d) Loadings of the first and second CCA dimension on 4 CNV genomic loci. Shows contribution of a CNV loci to the canonical dimension. (e) Loading of Neuromorphometrics Regions of Interests (ROIs) on the 2 PCA dimensions. ROIs are averaged across the left and right hemisphere for visualization.

The font size is correlated to the region's contribution to dimensions. ROI names are color coded as being part of the deletion (red), duplication (blue) and both deletion and duplication (magenta) convergence patterns. (f) Scatterplot showing the participant/specific expressions of each of the 484 carriers of 8 different CNVs along 2 dominant gene-morphometry Canonical Correlation (CC) dimensions established using 130 neuroanatomical GM regions of CNV carriers. GM region volumes were adjusted for total grey matter, age, age^2, sex and site. The empty and full symbols represent deletions and duplication respectively. The grey hexagonal bin plot represents the frequency of controls (n=1296). Controls were not used to calculate the CCA and were projected post hoc on the 2 dimensions using CCA prediction. CCA ROI loadings are reported in Supplementary Figure 10. Results for Surface Area and Cortical Thickness are reported in Supplementary Figures 9 (PCA), 14-15 (CCA).

CLINICAL ASCERTAINMENT								
CNV loci	**Copy number**	**Age mean (SD)**	**Male/ Female**	**TIV mean (SD)**	**FSIQ mean (SD)**	**ASD**	**SCZ**	**Other diagnosis**
1q21.1	Deletions N=29	29(18)	11/18	1.22(0.14)	90.85 (21.75) N=26	1	-	7
	Duplication N=19	34(17)	10/9	1.57(0.11)	95.56 (23.19) N=18	1	-	4
16p11.2	Deletions N=83	17(12)	47/36	1.54(0.17)	82.17 (14.99) N=64	13	-	36
	Duplication N=73	31(14.9)	41/32	1.33(0.17)	85.47 (19.48) N=63	10	1	19
22q11.2	Deletions N=74	16(8.6)	35/39	1.30(0.15)	77.42 (13.51) N=48	9	2	32
	Duplication N=22	20(14.2)	15/7	1.47(0.16)	97.83 (20.34) N=12	2	-	8
Controls N=331		26(14.6)	189/142	1.46(0.15)	106.73 (15.03) N=224	1	-	23
NON-CLINICAL ASCERTAINMENT								
CNV loci	**Copy number**	**Age mean (SD)**	**Male/ Female**	**TIV mean (SD)**	**UKB FI mean (SD)**	**ASD**	**SCZ**	**Other diagnosis**
1q21.1	Deletions N=10	59.1(6.7)	6/4	1.35(0.12)	-0.8 (0.5) N=9	-	1*	3
	Duplication N=9	60.6(7)	2/7	1.55(0.14)	0.2 (1.3) N=9	-	-	-
15q11.2	Deletions N=72	63.4(7.6)	31/41	1.54(0.15)	-0.3 (0.9) N=63	-	-	2
	Duplication N=76	62.9 (7.3)	36/40	1.49(0.15)	0 (1.1) N=71	-	-	6
16p11.2	Deletion N=4	65.6 (3.2)	3/1	1.56(0.13)	0.8 (0.5) N=2	-	-	-
	Duplication N=4	69.3 (2.1)	1/3	1.29(0.11)	-1.6 (0.2) N=4	-	-	-
22q11.2	Deletion N=1	69.8(-)	1/-	1.44(-)	-	-	-	-
	Duplication N=8	62(9.5)	4/4	1.55(0.17)	-0.2 (1.1) N=8	-	-	1
Controls N=965		62.1(7.4)	358/607	1.51(0.14)	0 (1) N=866	-	2*	65

Table 1: Demographics

Legend: CNV: Copy Number Variant, SD: Standard deviation, TIV: total intracranial volume, FSIQ: Full scale IQ, UKB FI: UK Biobank fluid intelligence, ASD: Autism Spectrum Disorders, SCZ: Schizophrenia (including * ICD10 code F25.9 Schizoaffective disorder, unspecified). CNV carriers and controls from the clinically ascertained group come from 5 different cohorts (Supplementary Table 1), while non-clinically ascertained participants were identified in the UK Biobank. 16p11.2 and 22q11.2 from the UKBB were not included in the VBM and SBM due to small sample size. Other diagnosis included: language disorder, major depressive disorder, posttraumatic stress disorder (PTSD), unspecified disruptive and impulse-control and conduct disorder, social anxiety disorder, social phobia disorder, speech sound disorder, moderate intellectual disability, specific learning disorder, gambling disorder, bipolar disorder, conduct disorder, attention deficit / hyperactivity disorder ADHD, Substance abuse disorder, global developmental delay, motor disorder, obsessive-compulsive disorder, sleep disorder, Tourette's disorder, mood disorder, eating disorders, transient tic disorder, trichotillomania, pervasive developmental disorder NOS, specific phobia, body dysmorphic disorder, mathematics disorder, dysthymic disorder.

Spatial overlap Cohen's d tails

a.

■ Overlap of any 3 CNV deletions   ■ Overlap of 4 CNV deletions

<15th   >85th

b.

■ Overlap of any 3 CNV duplications   ■ Overlap of 4 CNV duplications

<15th   >85th

Distribution of Cohen's d

deletions   duplications

15th 85th   15th 85th

c.   d.   1q21.1

e.   f.   16p11.2

g.   h.   22q11.2

i.   j.   15q11.2 UKBB

Supplementary Material, Methods and Results: Effects of eight neuropsychiatric copy number variants on human brain structure

INDEX

MRI data acquisition and quality check

Participants

Clinically ascertained CNV carriers were recruited as either probands referred for genetic testing, or as relatives. Controls were either non-carriers within the same families or individuals from the general population. We pooled data from 5 different cohorts: Cardiff University (UK), 16p11.2 European Consortium (Lausanne, Switzerland), University of Montreal (Canada), UCLA (Los Angeles, USA) and the Variation in individuals Project (SVIP, USA). A subset of the participants with 16p11.2 and 22q11.2 CNVs were included in prior publications [1–4] (Supplementary Table 1). CNVs from non-clinical populations were identified in the UK Biobank [5,6]. PennCNV and QuantiSNP were used, with standard quality control metrics, to identify CNVs [7,8].

16p11.2 European Consortium

MRI data of the EU participants were acquired on two 3T whole-body scanners. 14 carriers of a 16p11.2 deletion and 17 duplication carriers, together with 59 controls (21 familial and 38 unrelated controls) were examined on a Magnetom TIM Trio (Siemens Healthcare, Erlangen, Germany), using a 12-channel RF receive head coil and RF body transmit coil. The remaining 16p11.2 (13 deletions, 6 duplications), 1q11.2 (9 deletions, 7 duplications) carriers and controls (n=38) in the european cohort, were scanned on a Magnetom Prisma Syngo (Siemens Healthcare, Erlangen, Germany) using a 64-channel RF receive head coil and RF body transmit coil. T1-weighted (T1w) anatomical images acquired with the TIM Trio scanner used a Multi-Echo Magnetization Prepared RApid Gradient Echo sequence (ME-MPRAGE: 176 slices; 256×256 matrix; echo time (TE): TE1 = 1.64 ms, TE2 = 3.5 ms, TE3 = 5.36 ms, TE4 = 7.22 ms; repetition time (TR): 2530 ms; flip angle 7°). On the Prisma Syngo scanner, T1w images were acquired using a single-echo MPRAGE sequence (176 slices; 256×256 matrix; TE = 2.39 ms; TR = 2000 ms; flip angle 9°).

Simons Searchlight Consortium

Data were acquired using multi and single-echo sequences. 176 participants (38 del/ 34 dup 16p11.2 carriers, 2 dup 1q21.1 carriers and 102 familial controls) underwent the research MRI protocol at two imaging core sites on matched 3T Magnetom TIM Trio MRI scanners (Siemens Healthcare, Erlangen, Germany), using the vendor-supplied 32-channel phased-array radio-frequency head coils. 68 participants were scanned at University of California sites (UC) and 108 at the Children Hospital of Philadelphia (CHOP). Structural MRI data included multi-echo T1w ME-MPRAGE using the following parameters: 176 slices, 256×256 matrix, TR = 2530 ms, TI = 1200 ms, TE = 1.64 ms, and flip angle 7°. Clinical MRI images (single-echo) obtained at the phenotyping core sites were also analyzed. The remaining 79 subjects (19 del/ 13 dup 16p11.2 carriers, 12 del/8 dup 1q21.1 carriers and 27 familial controls) were scanned at University of Washington Medical Center, Baylor University Medical Center and Boston Children's Hospital on two matched 3T Philips Achieva (Philips Healthcare, United States of America) and one unmatched Magnetom TIM Trio scanner (Siemens Healthcare, Erlangen, Germany), respectively. T1w images were acquired using a single-echo MPRAGE sequence and the following parameters: 160 slices; 256×256 matrix; TE = 2.98 ms; TR = 2300 ms; flip angle 9°. All multi-echo images were averaged following a Root-Mean Square (RMS) averaging method.

Brain Canada

MRI scans for the Brain Canada cohort have been performed at the Montreal Neurological Institute with the same 3T scanner: Magnetom Prisma Syngo (Siemens Healthcare, Erlangen, Germany). Data included 16p11.2 (3 deletions, 3 duplications), 1q11.2 (5 deletions, 1 duplication), 22q11.2 (1 duplication) carriers and controls (n=26) T1w images were acquired using using MPRAGE sequences, scanning protocol description is detailed on this website: http://www.bic.mni.mcgill.ca/users/jlewis/BrainCanada/MCIN/.

4

UCLA

Imaging data of 22q11.2 CNV carriers and typically developing (TD) controls were acquired at the University of California, Los Angeles (UCLA). Patients were ascertained from the UCLA or Children's Hospital, Los Angeles Pediatric Genetics, Allergy/Immunology and/or Craniofacial Clinics. We excluded 11 individuals from the analysis due to insufficient quality of the imaging data (cf. Supplementary Methods, quality control). The final 22q11.2 sample includes 144 individuals (71 deletions, 19 duplications and 54 controls). Demographically comparable TD comparison subjects were recruited from the same communities as patients via web-based advertisements and by posting flyers and brochures at local schools, pediatric clinics, and other community sites. Exclusion criteria for all study participants included significant neurological or medical conditions (unrelated to 22q11.2 mutation) that might affect brain structure, history of head injury with loss of consciousness, insufficient fluency in English, and/or substance or alcohol abuse or dependence within the past 6 months. The UCLA Institutional Review Board approved all study procedures and informed consent documents. Scanning was conducted on an identical 3 tesla Siemens Trio MRI scanner with a 12-channel head coil at the University of California at Los Angeles Brain Mapping Center or at the Center for Cognitive Neuroscience.

Cardiff

Imaging acquisition in Cardiff was performed on a 3 T General Electric HDx MRI system (GE Medical Systems, Milwaukee, WI) using an eight-channel receive-only head RF coil. T1-weighted structural images were acquired with a 3D fast spoiled gradient echo (FSPGR) sequence (TR = 7.8 ms, TE = 3.0 ms, voxel size = 1 mm³ isomorphic). Data included 1 16p11.2 deletion, 1q11.2 (3 deletions, 1 duplication), 22q11.2 (3 deletions, 2 duplications) carriers and 15 controls.

MRI quality control

All MRI T1w nifti images were visually inspected by the same rater (CM) for head coverage, ghosting and susceptibility artifacts. Images were also screened after segmentation to ensure good

5

tissue classification accuracy. From the clinically ascertained dataset 55 subjects were excluded for insufficient image quality or artifacts while from the non-clinically ascertained dataset 52 subjects were excluded following the same criteria. Quality assurance protocol for Freesurfer based cortical reconstructions led to exclusion of an additional 34 scans. Numbers reported in Table 1 and Supplementary Table 1 are after exclusion.

List of abbreviations

CNV: Copy Number Variant, ASD: Autism Spectrum Disorders, SCZ: Schizophrenia, VBM: Voxel Based Morphometry, SBM: Surface Based Morphometry, TIV: Total Intracranial Volume, GM: Grey Matter, WM: White Matter, ICV: Intracranial Volume, CT: Cortical Thickness, SA: Surface Area, IQ: Intelligence Quotient, CCA: Canonical Correlation Analysis, PCA: Principal Component Analysis, pLI: Probability of being loss-of-function intolerant.

Supplementary Method 1: MRI data processing, Voxel-Based Morphometry

We used the probabilistic tissue classification within SPM12 "unified segmentation" framework [9] and enhanced tissue priors for optimal delineation of subcortical structures [10]. The obtained grey and white matter maps were screened for tissue misclassification by the same researcher (CM). Individuals' gray matter (GM) maps were subsequently spatially registered to the standard Montreal Neurological Institute (MNI) space using DARTELs diffeomorphic registration [11] followed by voxel-based scaling using the Jacobian determinants of the deformation field (i.e., "modulation"). Finally, GM maps were smoothed with a Gaussian of 8mm full-width-at-half-maximum. Parcellation into regions of interest (ROIs) was performed using maximum probability tissue labels (http://www.neuromorphometrics.com/) build using data from the OASIS project (http://www.oasis-brains.org). ROIs' volumes obtained for each subject were used for PCA and CCA, with the exception of white matter ROIs that were excluded.

Supplementary Method 2: MRI data processing, Surface-Based Morphometry

ENIGMA quality control procedures (http://enigma.ini.usc.edu/protocols/imaging-protocols/) led to the exclusion of additional 34 scans. Applying FreeSurfer's recon-all pipeline, estimates of local cortical thickness (CT) and surface area (SA) were calculated at each vertex. Statistical analyses were conducted on each vertex (Freesurfer fsaverage, 327,684 vertices), after spatially smoothing with a Gaussian kernel of 10-mm full width at half maximum (FWHM). In addition, FreeSurfer provided estimates of global brain measures of mean cortical thickness, total surface area per hemisphere, and regional measures based on the Desikan Freesurfer atlas [12].

Supplementary Method 3: Statistical analysis for global brain measures

Within the two cohort ascertainments, we used ANOVA design to compare group means of global brain measures. P-value correction for multiple comparisons was performed with the Tukey honest significant differences test. Wilcoxon rank sum and signed rank tests were used to compare distributions between deletions and duplications of each CNV.

Supplementary Method 4: Voxel-based measures and statistical analyses

GM maps were entered as dependent variables, group - as independent variable (variance between groups was set as unequal) and age, age^2, sex, scanning site, total grey matter volume - as fixed effects. After model estimation, post hoc contrasts were calculated to compare each CNV group to the corresponding control group. VBM results on grey matter are significant at $p < 0.05$, FWE corrected [2-4]. We tested for significant effects in a conjunction analysis across the pre-defined contrasts.

Supplementary Method 5: Surface-based measures and statistical analyses

In parallel to VBM, we used surface-based GLM-based analysis to test differences in CT and SA (SurfStat toolbox [13]). Each GLM used the surface feature as the dependent variable, the groups - as independent variable, which were adjusted for age, age^2, sex, site, and Total-SA/Mean-CT. Post-hoc contrasts compared each CNV group against controls, including estimation of Cohen's d effect size estimates from t-values [14]. False Discovery Rate (FDR), with p-value at 0.05 was applied to control for false positive errors due to multiple comparisons.

Supplementary Method 6: Spin permutation testing

The spin permutation test provides a null hypothesis quantifying the probability of observing by chance a dice index or correlation value, while controlling for spatial auto-correlations inherent in neuroimaging data. This method has been established previously in [15,16]. To perform the spin permutation test, we applied 10,000 random surface-based rotations of XYZ to a given map, generating random maps while preserving the relative spatial organization of the vertices. We then generate a null distribution of dice indices or correlation values for deletion and duplication convergence patterns or deletion-duplication anti-correlation respectively. An empirical p-value (p-value$_{SPIN}$) was then obtained as the proportion of 10,000 null dice-indices or correlations that had a higher value than the observed value. P-value$_{SPIN}$ < 0.05 was used as a statistical significance threshold [15,16].

Supplementary Method 7: Label shuffling

We additionally tested the overlap significance by performing permutation of control and CNV labels, generating empirical null distributions and calculating dice index distribution with regard to the convergence pattern. We performed permutation for control and CNV labels (7 groups, clinically ascertained) 10,000 times, and calculated the same overlap using the 15th and 85th percentile of

8

Cohen's d values (the top 15% of positive and negative effect sizes) between the three random groups. For 15q11.2 CNVs (non-clinically ascertained), we shuffled control and CNV labels 10,000 times. We then calculated the same overlap using the 15th and 85th percentile of Cohen's d values (the top 15% of positive and negative effect sizes) with 3-deletion convergence pattern (cf. above). To generate empirical null distributions to obtain p-values (p-value$_{SHUFFLE}$) for significance testing, we calculated dice index distribution w.r.t the convergence pattern. Similarly, for a new map, we calculated the dice index with respect to the convergence pattern and 10,000 random patterns generated above. An empirical p-value (p-value$_{SHUFFLE}$) was then obtained based on the proportion of 10,000 random map based dice-indices that had a higher value than the observed value. p-value$_{SHUFFLE}$ < 0.05 was used as a statistical significance threshold [15,16].

Supplementary Method 8: Multi-view pattern-learning analysis

Canonical correlation analysis (CCA) allows a single integrated multivariate analysis to simultaneously co-analyze brain morphometry measures along with the Genomic loci information (CNV status). This aims to identify symmetric linear relations between the two sets of variables. In our study, the first variable set (X) represents the CNV status of a subject, based on encoding with -1 (deletion) 0 (normal diploid) or +1 (duplication). All regional volume measures were normalized by de-meaning to 0 and re-scaling to 1. That is, the set of CNV carrier status was jointly encoded, capturing genomic loci. The second variable set (Y) was encoded as a matrix of 130 regional grey matter volumes (68 Desikan ROIs for SA, and CT) adjusted for total grey matter volume (Total SA, and Mean CT for SA, and CT). The brain volume measures in variable set Y were fed into CCA after a confound-removal procedure including age, age^2, sex, site, and total grey matter analogous to previous studies [17–19].

Our CCA strategy involved finding pairs of canonical vectors u and v that maximize the correlation in the embedding space (canonical variates) between a linear combination of the set of CNV indicators (X) and a linear combination of regional volumes (Y). In other words, canonical vectors project high dimensional variable sets onto a new low-rank space defined by the linear combination of original

variables, called canonical variates. The Canonical Correlation, calculated as Pearson's correlation (r) between X and Y canonical variates, can be seen as a metric of joint information reduction and serves as a performance measure for CCA. In our study, CCA identifies modes of coherent co-variation that jointly characterize how CNVs and patterns of regional volumes systematically co-occur across subjects. We refer to these modes of co-variation as 'CCA dimensions' or 'gene-morphology dimensions'.

The statistical significance of the derived dominant modes of gene-morphology co-variation was determined by null hypothesis testing based on a non-parametric permutation procedure [17]. Relying on minimal modeling assumptions [20], an empirical null distribution representing the absence of correlation between CNVs and brain morphometry. 1,000 gene-volume correlations were computed after permuting the original data such that the CNV status and set of region volume expressions were unrelated across individuals. In 1,000 permutation iterations, the CNV labels were held constant, while the regional volumes were subject to participant-wise random shuffling. The CCA analysis was rerun for each permutation iteration (the maximum possible correlation coefficient). An original CCA dimension was declared statistically significant if the associated canonical correlation values exceeded the 99% percentile of the null distribution (i.e., $p < 0.05$).

CLINICAL ASCERTAINMENT

CNV loci	Copy number	Cohort	Age mean(SD)	Age mean(SD)	Male/Female	Male/Female	TIV mean(SD)	TIV mean(SD)	FSIQ mean(SD)	FSIQ mean(SD)
1q21.1	Deletions N=29	EU N=9	29.43 (18.40)	18.91 (12.91)	11/18	1/8	1.22 (0.14)	1.15 (0.09)	90.85 (21.75) N=26	81.67 (21.57) N=9
		VIP N=12		35.11 (20.66)		6/6		1.27 (0.15)		99.33 (23.96) N=12
		BC N=5		28.25 (18.46)		1/4		1.17 (0.17)		87 (4.76) N=5
		Cardiff N=3		40.18 (13.15)		3/0		1.36 (0.02)		(-)
	Duplications N=19	EU N=7	34.29 (17.19)	37.32 (19.58)	10/9	5/2	1.57 (0.11)	1.58 (0.07)	95.56 (23.19) N=18	96.57 (11.59) N=7
		VIP N=10		34.25 (15.84)		5/5		1.54 (0.13)		93.80 (30.17) N=10
		BC N=1		8.34 (-)		0/1		1.57 (-)		106 (-) N=1
		Cardiff N=1		39.45 (-)		0/1		1.76 (-)		(-)
16p11.2	Deletions N=83	EU N=26	17.14 (11.97)	21.23 (13.51)	47/36	15/11	1.54 (0.17)	1.44 (0.14)	82.17 (14.99) N=64	74.38 (14.61) N=13
		VIP N=53		13.82 (9.47)		30/23		1.58 (0.17)		83.98 (14.36) N=48
		BC N=3		31.64 (10.26)		1/2		1.46 (0.09)		87 (21) N=3
		Cardiff N=1		43.04 (-)		1/0		2.01 (-)		(-)
	Duplications N=73	EU N=23	31.01(14.91)	32.55 (13.20)	41/32	13/10	1.33 (0.17)	1.36 (0.16)	85 (19.70) N=63	72.71 (16.30) N=17
		VIP N=47		30.52 (15.30)		25/22		1.30 (0.15)		90.44 (18.49) N=43
		BC N=3		26.89 (25.29)		3/0		1.49 (0.42)		86.67 (23.50) N=3
22q11.2	Deletions N=74	UCLA N=71	16.35 (8.56)	15.66 (7.24)	35/39	34/37	1.30 (0.15)	1.30 (0.15)	77.42 (13.51) N=48	77.41 (13.51) N=50
		Cardiff N=3		32.56 (20.60)		1/2		1.37 (0.11)		(-)
	Duplications N=22	UCLA N=19	19.66 (14.24)	17.32 (12.51)	15/7	13/6	1.47 (0.16)	1.45 (0.17)	97.83 (20.34) N=12	99.18 (20.76) N=11
		BC N=1		13.67 (-)		1/0		1.50 (-)		83 (-) N=1
		Cardiff N=2		44.94 (4.77)		1/1		1.55 (0.16)		(-)
	Controls N=331	EU N=97	25.91 (14.57)	30.33 (12.95)	189/142	62/35	1.46 (0.15)	1.49 (0.14)	106.73 (15.03) N=224	102.15 (12.98) N=59
		VIP N=129		24.03 (14.56)		71/58		1.44 (0.14)		108.88 (11.82) N=90
		UCLA N=54		13.37 (4.96)		30/24		1.39 (0.13)		112.43 (20.95) N=44
		BC N=36		33.56 (14.90)		19/17		1.52 (0.17)		101.11 (13.18) N=31
		Cardiff N=15		40.18 (11.12)		7/8		1.54 (0.15)		(-)

NON-CLINICAL ASCERTAINMENT

CNV loci	Copy number	Cohort	Age mean(SD)	Male/Female	TIV mean(SD)	UKB FI mean(SD)
1q21.1	Deletions N=10	UKBB	59.11 (6.71)	6/4	1.35 (0.12)	-0.8 (0.5) N=9
	Duplications N=9		60.56 (7)	2/7	1.55 (0.14)	0.2 (1.3) N=9
15q11.2	Deletions N=72		63.38 (7.60)	31/41	1.54 (0.15)	-0.3 (0.9) N=63
	Duplications N=76		62.92 (7.31)	36/40	1.49 (0.15)	0 (1.1) N=71
16p11.2	Deletions N=4		65.6 (3.2)	3/1	1.56 (0.13)	0.8 (0.5) N=2
	Duplications N=4		69.3 (2.1)	1/3	1.29 (0.11)	-1.6 (0.2) N=4
22q11.2	Deletions N=1		69.8 (-)	31/41	1.44 (-)	-
	Duplications N=8		62 (9.5)	4/4	1.55 (0.17)	-0.2 (1.1) N=8
	Controls N=965		62.13 (7.40)	358/607	1.51 (0.14)	0 (1) N=866

Supplementary Table 1 : Demographics

Legend: EU: 16p11.2 European Consortium, VIP: Simons Searchlight Consortium, BC: Brain Canada, CNV: Copy Number Variant, SD: Standard deviation, TIV: total intracranial volume, FSIQ: Full-scale IQ, UKB FI: UK Biobank fluid intelligence, ASD: Autism Spectrum Disorders, SCZ: schizophrenia (including * ICD10 code F25.9 Schizoaffective disorder, unspecified). CNV carriers and controls from the clinically ascertained group come from 5 different cohorts, while non-clinically ascertained participants were identified in the UK Biobank. UK Biobank fluid intelligence scores (UKB field:20016) were adjusted for age, age^2, sex, site, and then z-scored. 16p11.2 and 22q11.2

from the UKBB were not included in the VBM and SBM due to small sample size. Data for 16p11.2 and 22q11.2 deletions and duplications from the VIP, EU and UCLA datasets were previously published in Martin-Brevet *et al.* 2018 and Lin *et al.* 2017 [1,4].

		total Surface Area		mean Cortical Thickness	
		Z-score mean	Z-score sd	Z-score mean	Z-score sd
clinical ascertainment	1q21.1 deletion	-1.96	0.65	0.37	0.94
	1q21.1 duplication	1.16	1.01	0.13	1.76
	16p11.2 deletion	0.52	1.24	-0.26	1.07
	16p11.2 duplication	-1.17	1.00	-0.51	0.89
	22q11.2 deletion	-0.97	0.98	0.41	1.16
	22q11.2 duplication	0.44	0.94	-0.81	0.99
	controls	0.00	1.00	0.00	1.00
non-clinical ascertainment	1q21.1 deletion	-2.06	0.56	0.58	0.89
	1q21.1 duplication	0.83	0.99	0.10	1.18
	15q11.2 deletion	-0.13	0.99	0.38	0.97
	15q11.2 duplication	-0.17	1.09	-0.29	0.92
	16p11.2 deletion	-0.70	0.24	0.87	1.18
	16p11.2 duplication	1.24	1.04	-1.47	0.97
	22q11.2 deletion	-0.58	NA	0.86	NA
	22q11.2 duplication	0.04	0.63	0.04	1.10
	controls	0.00	1.00	0.00	1.00
		Total Brain Volume		total Grey Matter	
		Z-score mean	Z-score sd	Z-score mean	Z-score sd
clinical ascertainment	1q21.1 deletion	-1.72	0.82	-1.85	0.75
	1q21.1 duplication	0.89	0.83	1.02	1.06
	16p11.2 deletion	0.61	1.23	0.40	1.24

	16p11.2 duplication	-1.10	1.12	-1.12	1.09
	22q11.2 deletion	-0.60	0.95	-0.75	0.98
	22q11.2 duplication	0.36	1.15	0.40	1.08
	controls	0.00	1.00	0.00	1.00
non-clinical ascertainment	1q21.1 deletion	-1.91	0.97	-2.03	0.56
	1q21.1 duplication	0.67	1.28	0.70	1.31
	15q11.2 deletion	-1.64	0.93	0.09	1.08
	15q11.2 duplication	0.17	1.22	-0.35	1.12
	16p11.2 deletion	-0.11	0.71	-0.17	0.83
	16p11.2 duplication	-1.78	0.59	-1.31	0.59
	22q11.2 deletion	0.17	NA	-0.07	NA
	22q11.2 duplication	-0.32	0.88	0.35	0.69
	controls	0.00	1.00	0.00	1.00

Supplementary Table 2: Mean and standard deviation for each CNV on global metrics.

Legend: Mean and standard deviations of global metrics z-scores for each CNV group. These data are represented in Figure 1

	Median Cohen's d 5th percentile	Median Cohen's d 95th percentile	% of voxels surviving FWE	% of voxels with negative effects surviving FWE	% of voxels with positive effects surviving FWE
1q21.1 deletion	-0.29	0.30	1.49%	0.64%	0.84%
1q21.1 duplication	-0.23	0.22	0.04%	0.01%	0.03%
16p11.2 deletion	-0.48	0.65	17.13%	8.03%	9.09%
16p11.2 duplication	-0.37	0.25	2.84%	2.66%	0.17%
22q11.2 deletion	-0.45	0.59	14.81%	7.28%	7.52%
22q11.2 duplication	-0.23	0.27	0.32%	0.06%	0.26%
1q21.1 deletion UKBB	-0.14	0.17	0.31%	-	0.31%
1q21.1 duplication UKBB	-0.14	0.15	0.04%	-	0.04%
15q11.2 deletion UKBB	-0.13	0.14	0.02%	0.02%	-
15q11.2 duplication UKBB	-0.12	0.14	0.002%	-	0.002%

Supplementary Table 3: CNV effect sizes on local grey matter volume changes

Legend: First and second column report medians of 5th and 95th percentiles of Cohen's d distributions. Columns three to five report voxels surviving FWE error correction.

	Median Cohen's d 5th percentile	Median Cohen's d 95th percentile	% of vertices surviving FDR (q<0.05)	% of vertices with negative effects surviving FDR (q<0.05)	% of vertices with positive effects surviving FDR (q<0.05)
1q21.1 deletion	-0.21	0.21	0.31%	-	0.31%
1q21.1 duplication	-0.18	0.20	0.06%	-	0.06%
16p11.2 deletion	-0.51	0.63	42.45%	19.97%	22.48%
16p11.2 duplication	-0.35	0.30	13.91%	8.81%	5.10%
22q11.2 deletion	-0.44	0.49	34.07%	14.60%	19.47%
22q11.2 duplication	-0.18	0.21	1.31%	0.31%	1.00%
1q21.1 deletion UKBB	-0.11	0.11	-	-	-
1q21.1 duplication UKBB	-0.12	0.15	0.10%	0.03%	0.07%
15q11.2 deletion UKBB	-0.12	0.13	0.07%	-	0.07%
15q11.2 duplication UKBB	-0.13	0.14	-	-	-

Supplementary Table 4: CNV effect sizes on local Surface Area changes

Legend: First and second column report medians of 5th and 95th percentiles of Cohen's d distributions. Columns three to five report vertices surviving FDR correction (q<0.05).

	Median Cohen's d 5th percentile	Median Cohen's d 95th percentile	% of vertices surviving FDR (q<0.05)	% of vertices with negative effects surviving FDR (q<0.05)	% of vertices with positive effects surviving FDR (q<0.05)
1q21.1 deletion	-0.21	0.24	2.44%	0.55	1.89%
1q21.1 duplication	-0.20	0.18	-	-	-
16p11.2 deletion	-0.31	0.35	18.31%	8.67%	9.64%
16p11.2 duplication	-0.27	0.26	6.88%	4.02%	2.86%
22q11.2 deletion	-0.44	0.39	27.96%	18.52%	9.44%
22q11.2 duplication	-0.22	0.20	0.71%	0.46%	0.25%
1q21.1 deletion UKBB	-0.12	0.14	-	-	-
1q21.1 duplication UKBB	-0.23	0.08	19.23%	19.23%	-
15q11.2 deletion UKBB	-0.12	0.14	-	-	-
15q11.2 duplication UKBB	-0.11	0.12	-	-	-

Supplementary Table 5: CNV effect sizes on local Cortical Thickness changes

Legend: First and second column report medians of 5th and 95th percentiles of Cohen's d distributions. Columns three to five report vertices surviving FDR correction (q<0.05).

	Ratio del/dup (GM)	F test del/dup (GM)	Degrees of freedom df1-df2 (GM)	p-value F test (GM)	Ratio del/dup (SA)	F test del/dup (SA)	Degrees of freedom df1-df2 (SA)	p-value F test (SA)
1q21.1	1.297	1.682	416407-416408	0	1.087	1.174	327683-327683	0
16p11.2	1.846	3.485	416408-416408	0	1.772	3.141	327683-327683	0
22q11.2	2.019	4.0748	416408-416407	0	2.267	5.139	327683-327683	0
ukb 1q21.1	1.013	1.025	385045-385046	0	0.795	0.631	327683-327683	0
ubk 15q11.2	1.071	1.148	385045-385045	0	0.944	0.892	327683-327683	0

Supplementary Table 6: Within loci comparison of Cohen's d variances between deletions and duplications

Legend: Standard deviation ratio and two-sample F test for equal variances on Cohen's D distributions of GM volume and SA. Second and sixth columns report standard deviations ratios between deletion and duplication Cohen's d distributions. Column 3-5 and 7-9 report two-sample F test statistics for GM volume and SA for each loci.

	1q21.1	16p11.2	22q11.2	ukb1q21.1	ukb15q11.2
1q21.1	-0.187	-0.058	0.258*	0.213*	0.150
16p11.2	0.141	-0.626*	-0.141	0.035	-0.025
22q11.2	0.035	0.208	-0.232*	-0.006	-0.061
ukb1q21.1	0.535*	0.064	0.058	-0.384*	0.083
ukb15q11.2	-0.138	0.020	0.173	-0.175	-0.184

Supplementary Table 7: Brain-wide Cohen's d map Correlations for VBM

Legend: Pearson correlation between Cohen's d maps (brain wide) for Deletion-Duplication (uiagonal, purple cells), Deletion-Deletion (lower triangle, red cells), and Duplication-Duplication (upper triangle, blue cells) for clinical and non-clinical CNVs. ukb15q11.2 and ukb1q21.1 correspond to UKBB CNV carriers. * Bonferoni correction, p-value$_{SPIN}$ <0.002 (=0.05/25).

	1q21.1	16p11.2	22q11.2	ukb1q21.1	ukb15q11.2
1q21.1	-0.147	0.232*	-0.030	0.318*	0.021
16p11.2	0.136	-0.551*	-0.006	0.083	0.025
22q11.2	0.111	0.133	-0.139	0.035	0.050
ukb1q21.1	0.332*	0.118	-0.143	-0.275*	0.041
ukb15q11.2	0.058	0.054	0.178	-0.035	-0.231*

Supplementary Table 8: Cortex-wide Cohen's d map Correlations for SA

Legend: Pearson correlation between Cohen's d maps (coretx wide) for Deletion-Duplication (diagonal, purple cells), Deletion-Deletion (lower triangle, red cells), and Duplication-Duplication (upper triangle, blue cells) for clinical and non-clinical CNVs. ukb15q11.2 and ukb1q21.1 correspond to UKBB CNV carriers. * Bonferoni correction, p-value$_{SPIN}$ <0.002 (=0.05/25).

	1q21.1	16p11.2	22q11.2	ukb1q21.1	ukb15q11.2
1q21.1	-0.074	0.083	0.165	0.059	0.041
16p11.2	0.132	-0.169	-0.094	-0.068	0.066
22q11.2	-0.037	0.032	0.012	0.198	-0.044
ukb1q21.1	0.381*	0.123	-0.130	-0.103	-0.009
ukb15q11.2	0.032	0.074	-0.016	0.115	-0.092

Supplementary Table 9: Cortex-wide Cohen's d map Correlations for CT

Legend: Pearson correlation between Cohen's d maps (coretx wide) for Deletion-Duplication (diagonal, purple cells), Deletion-Deletion (lower triangle, red cells), and Duplication-Duplication (upper triangle, blue cells) for clinical and non-clinical CNVs. ukb15q11.2 and ukb1q21.1 correspond to UKBB CNV carriers. * Bonferoni correction, p-value$_{SPIN}$ <0.002 (=0.05/25).

Neuromorphometric brain regions	1q21.1 del	1q21.1 dup	16p11.2 del	16p11.2 dup	22q11.2 del	22q11.2 dup	15q11.2 del	15q11.2 dup
X3rd.Ventricle	0.01	0.06	0.14	0.05	0.26	0.12	0.35	-0.17
X4th.Ventricle	0.47	-0.58	-0.07	0.00	0.02	-0.06	0.10	-0.10
R.Accumbens.Area	0.15	0.53	-0.50	0.14	0.97	-0.87	-0.24	0.08
L.Accumbens.Area	0.01	0.27	-0.39	0.09	0.95	-0.99	-0.11	0.07
R.Amygdala	0.26	0.20	0.13	-0.03	0.16	0.28	-0.23	0.09
L.Amygdala	0.40	-0.19	0.04	-0.10	0.11	0.09	-0.28	0.13
Brain.Stem	0.12	-0.32	-0.02	-0.18	-0.46	0.68	-0.27	0.20
R.Caudate	0.27	-0.12	-0.20	-0.22	0.63	-0.60	0.14	-0.09
L.Caudate	0.29	-0.32	-0.14	-0.35	0.55	-0.65	0.01	-0.06
R.Cerebellum.Exterior	0.02	-0.34	-0.48	0.17	-0.47	0.73	-0.20	-0.15
L.Cerebellum.Exterior	0.08	-0.34	-0.46	0.11	-0.50	0.85	-0.20	-0.18
CSF	-0.14	0.15	0.08	0.09	0.48	-0.14	0.34	-0.23
R.Hippocampus	0.56	-0.48	-0.08	0.02	0.01	0.33	-0.36	-0.04
L.Hippocampus	0.58	-0.65	-0.15	0.13	0.08	-0.04	-0.36	-0.01
R.Inf.Lat.Vent	-0.22	-0.06	0.12	0.18	0.34	-0.22	0.06	-0.14
L.Inf.Lat.Vent	-0.09	-0.01	0.16	0.19	0.31	-0.24	0.02	-0.12
R.Lateral.Ventricle	-0.07	0.49	-0.05	0.43	0.44	-0.03	0.28	-0.17
L.Lateral.Ventricle	-0.04	0.52	-0.03	0.40	0.38	0.05	0.28	-0.21
R.Pallidum	0.18	-0.61	-0.05	-0.31	0.12	-0.19	-0.17	0.17
L.Pallidum	0.03	-0.70	-0.05	-0.39	0.13	-0.29	-0.16	0.14
R.Putamen	0.45	0.01	-0.10	-0.29	0.02	-0.24	-0.24	-0.02
L.Putamen	0.42	0.02	-0.16	-0.28	0.05	-0.22	-0.33	0.00
R.Thalamus.Proper	0.64	-0.44	0.17	-0.56	0.18	-0.01	0.01	0.19
L.Thalamus.Proper	0.60	-0.54	0.20	-0.62	0.22	-0.10	-0.06	0.21
R.Ventral.DC	0.42	-0.30	0.27	-0.24	0.00	0.40	-0.31	0.33
L.Ventral.DC	0.38	-0.41	0.26	-0.33	0.01	0.32	-0.29	0.31
Optic.Chiasm	-0.06	0.24	-0.52	0.20	-0.03	0.45	0.00	-0.23
Cerebellar.Vermal.Lobules.I.V	0.02	-0.44	-0.19	-0.17	-0.06	0.39	-0.01	-0.27
Cerebellar.Vermal.Lobules.VI.VII	-0.01	-0.25	-0.37	-0.13	-0.31	0.44	-0.14	-0.12
Cerebellar.Vermal.Lobules.VIII.X	0.44	-0.43	0.00	-0.17	-0.39	0.76	-0.13	-0.03
L.Basal.Forebrain	0.09	0.70	0.30	-0.04	0.55	-0.47	-0.12	0.15
R.Basal.Forebrain	0.13	0.65	0.29	0.08	0.61	-0.38	-0.14	0.29
R.ACgG.anterior.cingulate.gyrus	-0.49	0.41	-0.15	-0.20	-0.22	0.40	-0.10	0.05

L.ACgG.anterior.cingulate.gyrus	-0.85	0.72	-0.20	0.03	-0.15	0.41	-0.02	0.07
R.AIns.anterior.insula	-0.21	0.16	0.75	-0.50	0.65	-0.57	0.50	-0.14
L.AIns.anterior.insula	-0.15	0.08	0.66	-0.51	0.72	-0.60	0.35	-0.07
R.AOrG.anterior.orbital.gyrus	-0.35	0.48	-0.09	-0.14	0.07	-0.09	-0.08	0.18
L.AOrG.anterior.orbital.gyrus	-0.50	0.06	0.01	-0.06	0.07	0.09	0.03	0.18
R.AnG.angular.gyrus	-0.11	0.36	-0.48	0.25	-0.29	0.17	-0.06	-0.12
L.AnG.angular.gyrus	0.11	-0.06	-0.65	0.32	-0.28	0.26	-0.10	-0.18
R.Calc.calcarine.cortex	0.02	0.21	0.85	-0.59	-0.46	0.58	-0.22	-0.06
L.Calc.calcarine.cortex	0.08	0.11	0.83	-0.56	-0.50	0.46	-0.28	-0.01
R.CO.central.operculum	-0.30	0.25	0.03	0.02	0.29	-0.25	0.13	0.06
L.CO.central.operculum	-0.27	0.18	0.31	-0.16	0.45	-0.15	0.17	0.06
R.Cun.cuneus	0.06	0.14	0.20	-0.19	-0.59	0.50	-0.20	-0.20
L.Cun.cuneus	-0.06	0.22	0.29	-0.43	-0.69	0.74	-0.19	-0.24
R.Ent.entorhinal.area	0.17	0.47	0.08	0.11	-0.06	0.50	-0.11	0.07
L.Ent.entorhinal.area	0.17	0.31	-0.03	0.06	-0.16	0.56	-0.23	-0.02
R.FO.frontal.operculum	-0.10	-0.11	0.40	-0.42	0.26	-0.18	0.05	-0.10
L.FO.frontal.operculum	-0.04	-0.24	0.36	-0.45	0.24	-0.26	-0.07	0.02
R.FRP.frontal.pole	-0.41	0.80	-0.45	0.07	0.27	-0.07	0.21	-0.01
L.FRP.frontal.pole	-0.40	0.52	-0.27	0.12	0.29	-0.11	0.24	-0.03
R.FuG.fusiform.gyrus	-0.26	0.24	-0.41	0.40	-0.28	0.66	-0.03	-0.22
L.FuG.fusiform.gyrus	-0.22	0.23	-0.47	0.30	-0.40	0.65	-0.16	-0.23
R.GRe.gyrus.rectus	-0.64	0.70	-0.51	-0.04	0.25	-0.47	0.11	0.08
L.GRe.gyrus.rectus	-0.61	0.49	-0.27	-0.10	0.40	-0.50	0.00	0.00
R.IOG.inferior.occipital.gyrus	-0.27	0.28	0.31	-0.09	-0.35	0.59	0.05	-0.10
L.IOG.inferior.occipital.gyrus	-0.04	0.49	0.06	-0.20	-0.10	0.30	0.09	-0.27
R.ITG.inferior.temporal.gyrus	-0.38	-0.17	-0.28	0.43	-0.31	0.63	0.04	-0.22
L.ITG.inferior.temporal.gyrus	-0.36	-0.08	-0.25	0.36	-0.32	0.63	-0.04	-0.37
R.LiG.lingual.gyrus	0.35	0.04	0.45	-0.43	-0.43	0.76	-0.09	-0.31
L.LiG.lingual.gyrus	0.25	-0.04	0.29	-0.47	-0.50	0.64	-0.14	-0.26
R.LOrG.lateral.orbital.gyrus	-0.37	0.10	0.14	-0.26	0.07	0.07	-0.07	0.07
L.LOrG.lateral.orbital.gyrus	-0.49	-0.27	0.17	-0.13	0.06	0.14	-0.14	0.25
R.MCgG.middle.cingulate.gyrus	-0.55	0.79	-0.74	0.11	-0.55	0.80	-0.35	0.05
L.MCgG.middle.cingulate.gyrus	-0.91	0.88	-0.81	0.21	-0.62	0.77	-0.34	0.00
R.MFC.medial.frontal.cortex	-0.70	0.69	-0.16	0.11	0.21	0.11	0.18	0.14

22

L.MFC.medial.frontal.cortex	-0.69	0.76	-0.03	0.08	0.27	0.17	0.20	0.12
R.MFG.middle.frontal.gyrus	-0.90	0.64	-0.55	0.32	0.14	-0.17	-0.05	0.06
L.MFG.middle.frontal.gyrus	-1.00	0.68	-0.53	0.32	0.21	-0.35	-0.09	0.00
R.MOG.middle.occipital.gyrus	-0.04	0.14	-0.03	-0.04	-0.22	0.43	0.07	-0.05
L.MOG.middle.occipital.gyrus	0.14	0.37	-0.36	-0.05	-0.15	0.20	-0.11	-0.13
R.MOrG.medial.orbital.gyrus	-0.36	0.52	0.09	-0.37	0.21	-0.20	-0.01	0.13
L.MOrG.medial.orbital.gyrus	-0.26	0.37	0.13	-0.35	0.25	-0.40	0.02	0.11
R.MPoG.postcentral.gyrus.medial.segment	0.13	0.18	-0.23	-0.04	-0.24	0.21	-0.29	0.03
L.MPoG.postcentral.gyrus.medial.segment	-0.06	0.41	-0.20	0.12	-0.18	0.21	-0.10	0.08
R.MPrG.precentral.gyrus.medial.segment	-0.37	0.36	-0.40	0.14	-0.41	0.30	-0.24	0.07
L.MPrG.precentral.gyrus.medial.segment	-0.35	0.61	-0.45	0.05	-0.48	0.47	-0.18	0.04
R.MSFG.superior.frontal.gyrus.medial.segment	-0.75	0.78	-0.38	0.27	-0.03	-0.02	0.11	0.11
L.MSFG.superior.frontal.gyrus.medial.segment	-0.72	0.69	-0.37	0.15	-0.02	-0.01	0.10	0.09
R.MTG.middle.temporal.gyrus	-0.26	0.04	-0.57	0.42	-0.28	0.42	0.03	-0.16
L.MTG.middle.temporal.gyrus	-0.19	-0.20	-0.61	0.50	-0.32	0.49	0.02	-0.32
R.OCP.occipital.pole	0.34	-0.10	0.57	-0.48	-0.32	0.42	-0.02	-0.03
L.OCP.occipital.pole	0.49	0.06	0.37	-0.35	-0.27	0.39	-0.06	-0.05
R.OFuG.occipital.fusiform.gyrus	0.03	0.06	0.09	-0.18	-0.33	0.68	-0.08	-0.31
L.OFuG.occipital.fusiform.gyrus	0.13	0.02	0.11	-0.22	-0.16	0.26	-0.10	-0.28
R.OpIFG.opercular.part.of.the.inferior.frontal.gyrus	-0.37	0.33	-0.11	0.37	0.21	0.13	0.02	0.11
L.OpIFG.opercular.part.of.the.inferior.frontal.gyrus	-0.41	0.21	-0.21	0.25	0.20	-0.35	-0.04	0.21
R.OrIFG.orbital.part.of.the.inferior.frontal.gyrus	-0.22	-0.34	0.35	-0.43	0.15	-0.12	0.03	0.05
L.OrIFG.orbital.part.of.the.inferior.frontal.gyrus	-0.31	-0.31	0.31	-0.41	0.18	-0.19	-0.10	0.14
R.PCgG.posterior.cingulate.gyrus	0.42	0.22	0.03	-0.04	-0.29	0.70	0.09	-0.13
L.PCgG.posterior.cingulate.gyrus	0.32	-0.08	-0.02	-0.08	-0.36	0.59	-0.04	0.01
R.PCu.precuneus	0.14	0.24	0.04	0.08	-0.41	0.40	0.05	-0.18
L.PCu.precuneus	0.15	0.14	0.09	-0.12	-0.61	0.55	-0.03	-0.04
R.PHG.parahippocampal.gyrus	0.33	0.50	0.41	-0.02	0.14	0.44	-0.01	-0.09
L.PHG.parahippocampal.gyrus	0.22	0.13	0.16	-0.03	-0.02	0.26	-0.24	-0.12
R.PIns.posterior.insula	-0.15	0.31	0.74	-0.31	0.59	-0.44	0.27	0.07
L.PIns.posterior.insula	-0.04	0.15	0.85	-0.39	0.69	-0.37	0.39	0.05
R.PO.parietal.operculum	0.00	0.23	0.66	-0.22	0.12	0.16	0.05	0.15
L.PO.parietal.operculum	-0.14	0.04	1.07	-0.43	0.48	-0.04	0.00	0.23
R.PoG.postcentral.gyrus	0.47	-0.25	-0.41	0.08	-0.19	0.09	-0.05	-0.05

L.PoG.postcentral.gyrus	0.30	-0.39	-0.53	0.11	0.04	-0.40	-0.14	-0.07
R.POrG.posterior.orbital.gyrus	0.08	-0.11	0.69	-0.42	0.25	0.26	0.05	0.14
L.POrG.posterior.orbital.gyrus	-0.04	0.13	0.60	-0.41	0.22	-0.02	-0.04	0.22
R.PP.planum.polare	-0.14	0.20	0.39	-0.07	0.11	0.12	0.22	-0.11
L.PP.planum.polare	-0.19	-0.10	0.55	-0.19	-0.01	0.43	0.12	0.07
R.PrG.precentral.gyrus	-0.44	-0.28	-0.66	0.43	0.09	-0.21	-0.17	-0.16
L.PrG.precentral.gyrus	-0.33	-0.04	-0.58	0.29	0.34	-0.33	-0.03	-0.05
R.PT.planum.temporale	-0.02	0.26	0.62	-0.15	0.06	0.34	0.17	0.02
L.PT.planum.temporale	-0.26	-0.14	1.28	-0.43	0.52	0.02	0.11	0.14
R.SCA.subcallosal.area	-0.57	0.70	0.23	-0.10	0.57	-0.26	0.27	0.00
L.SCA.subcallosal.area	-0.73	0.68	0.28	-0.03	0.67	-0.53	0.23	-0.02
R.SFG.superior.frontal.gyrus	-0.73	0.64	-0.45	0.20	0.39	-0.39	0.15	-0.25
L.SFG.superior.frontal.gyrus	-0.72	0.46	-0.26	0.05	0.35	-0.36	0.26	-0.18
R.SMC.supplementary.motor.cortex	-0.70	0.43	-0.67	0.39	-0.37	0.26	-0.02	0.01
L.SMC.supplementary.motor.cortex	-0.65	0.65	-0.58	0.28	-0.42	0.34	0.00	0.06
R.SMG.supramarginal.gyrus	-0.05	-0.23	-0.40	0.16	-0.04	-0.05	-0.22	-0.12
L.SMG.supramarginal.gyrus	0.00	0.12	-0.31	0.00	0.09	-0.17	-0.27	-0.02
R.SOG.superior.occipital.gyrus	0.11	-0.21	0.15	-0.31	-0.27	0.24	-0.09	-0.25
L.SOG.superior.occipital.gyrus	0.24	-0.07	-0.02	-0.21	-0.25	0.07	-0.18	-0.08
R.SPL.superior.parietal.lobule	0.09	-0.20	-0.15	-0.11	-0.08	-0.10	-0.09	0.00
L.SPL.superior.parietal.lobule	0.19	-0.35	-0.19	0.03	-0.25	0.17	-0.02	-0.02
R.STG.superior.temporal.gyrus	0.02	0.36	-0.63	0.23	-0.14	0.37	0.00	-0.16
L.STG.superior.temporal.gyrus	-0.12	0.27	-0.67	0.29	-0.18	0.50	-0.01	-0.16
R.TMP.temporal.pole	0.12	-0.29	-0.22	0.12	-0.30	0.34	0.10	-0.10
L.TMP.temporal.pole	-0.15	-0.19	-0.06	0.06	-0.37	0.37	0.01	-0.13
R.TrIFG.triangular.part.of.the.inferior.frontal.gyrus	-0.41	0.38	0.15	-0.15	0.10	0.18	0.06	0.21
L.TrIFG.triangular.part.of.the.inferior.frontal.gyrus	-0.50	0.35	0.14	-0.26	0.04	-0.14	-0.14	0.29
R.TTG.transverse.temporal.gyrus	-0.05	0.20	0.85	-0.22	0.09	0.43	0.16	0.08
L.TTG.transverse.temporal.gyrus	-0.27	-0.01	1.37	-0.47	0.53	0.16	0.13	0.15

Supplementary Table 10: Cohen's d values for neuromorphometric brain regions.

Legend: Cohen's d profiles obtained from contrasts between the 8 CNV groups and controls (adjusted for age, age^2, sex, site, and total grey matter). Whole brain was parcellated with the neuromorphometric atlas (Supplementary Method 1) and 130 grey matter regions were used to create Cohen's d profiles.

	ROI	PC1		ROI	PC2
1.	Left_Cerebellum_Exterior	-3.186		Left_Thalamus_Proper	-3.268
2.	Right_Cerebellum_Exterior	-2.992		Right_Thalamus_Proper	-2.988
3.	Left_middle_cingulate_gyrus	-2.808		Left_Ventral_DC	-2.800
4.	Left_fusiform_gyrus	-2.666		Right_Ventral_DC	-2.616
5.	Right_middle_cingulate_gyrus	-2.664		Left_calcarine_cortex	-2.563
6.	Left_middle_temporal_gyrus	-2.608		Right_occipital_pole	-2.426
7.	Left_inferior_temporal_gyrus	-2.473		Right_calcarine_cortex	-2.308
8.	Right_fusiform_gyrus	-2.215		Cerebellar_Vermal_Lobules_VIII-X	-2.108
9.	Left_cuneus	-2.152		Right_Pallidum	-2.031
10.	Left_angular_gyrus	-2.120		Left_Pallidum	-2.030
11.	Left_Accumbens_Area	2.487		Right_frontal_pole	2.392
12.	Right_subcallosal_area	2.617		Left_superior_frontal_gyrus	2.443
13.	Left_parietal_operculum	2.646		Left_superior_frontal_gyrus_medial_segme	2.459
14.	Left_planum_temporale	2.838		Left_supplementary_motor_cortex	2.487
15.	Left_transverse_temporal_gyrus	2.876		Right_supplementary_motor_cortex	2.626
16.	Left_subcallosal_area	3.022		Left_middle_cingulate_gyrus	2.734
17.	Right_posterio_.insula	3.315		Right_superior_frontal_gyrus_medial_segm	2.862
18.	Left_anterior_insula	3.654		Right_middle_frontal_gyrus	3.073
19.	Left_posterior_insula	3.666		Right_superior_frontal_gyrus	3.099
20.	Right_anterior_insula	3.784		Left_middle_frontal_gyrus	3.302

Supplementary Table 11: Principal Component Analysis ROIs loadings (volume)

Legend: Top 20 loadings (10 negative, 10 positive) of Neuromorphometric Regions of Interests (ROIs) on the 2 PCA dimensions (the first 2 components explained respectively 31.77 and 28.66 % of the variance). Loadings values are listed in increasing order, top 10 negative followed by top 10 positive.

	ROI	|CC1|		ROI	|CC2|
1.	Left_transverse_temporal_gyrus	0.54424372		Right_Accumbens_Area	0.716341907
2.	Left_planum_temporale	0.516375486		Left_Accumbens_Area	0.693450348
3.	Left_parietal_operculum	0.478605816		Left_subcallosal_area	0.598630484
4.	Left_calcarine_cortex	0.434979442		Left_cuneus	0.587180378
5.	Right_supplementary_motor_cortex	0.423161744		Left_precuneus	0.542742079
6.	Right_calcarine_cortex	0.417872614		Brain_Stem	0.538584491
7.	Left_posterior_insula	0.41658809		Right_superior_frontal_gyrus	0.530594854
8.	Left_middle_cingulate_gyrus	0.404137535		Cerebellar_Vermal_Lobules_VIII-X	0.530148089
9.	Left_middle_temporal_gyrus	0.380330333		Left_anterior_insula	0.527867312
10.	Right_anterior_insula	0.378170585		Right_lingual_gyrus	0.526385942
11.	Right_posterior_orbital_gyrus	0.375392577		Left_lingual_gyrus	0.52497161
12.	Left_anterior_insula	0.37348146		Right_cuneus	0.513604455
13.	Right_occipital_pole	0.367616101		Right_Caudate	0.502454679
14.	Left_supplementary_motor_cortex	0.35555794		Right_anterior_insula	0.49874064
15.	Right_precentral_gyrus	0.354152555		Left_calcarine_cortex	0.486781547
16.	Right_posterior_insula	0.352835632		Left_superior_frontal_gyrus	0.481515906
17.	Right_middle_temporal_gyrus	0.351133642		Left_Cerebellum_Exterior	0.477760835
18.	Right_middle_cingulate_gyrus	0.343736614		Left_gyrus_rectus	0.468197071
19.	Right_middle_frontal_gyrus	0.341059273		Right_subcallosal_area	0.465928639
20.	Left_Thalamus_Proper	0.340754975		Right_calcarine_cortex	0.463409682

Supplementary Table 12: Canonical Correlation Analysis ROIs loadings (volume)

Legend: Top 20 loadings of Neuromorphometric Regions of Interests (ROIs) on the 2 main CCA dimensions (r=0.84, 0.79; statistically significant at p-value<0.05). Loadings are in absolute values and listed in decreasing order.

	ROI	|CC1|		ROI	|CC2|
1.	L_superiortemporal	0.6587446385		R_supramarginal	0.4556194543
2.	R_superiortemporal	0.4716305815		R_postcentral	0.4078724885
3.	R_pericalcarine	0.3752495588		L_caudalanteriorcingulate	0.3529766412
4.	L_bankssts	0.3625881379		L_postcentral	0.3515379435
5.	L_transversetemporal	0.3385850519		R_pericalcarine	0.3243996464
6.	L_insula	0.3145652172		L_inferiortemporal	0.3037588837
7.	R_caudalanteriorcingulate	0.3088200827		R_inferiortemporal	0.2836174352
8.	L_parahippocampal	0.3074088432		L_parstriangularis	0.2467471024
9.	L_caudalanteriorcingulate	0.2822321503		R_superiortemporal	0.2382066534
10.	L_lingual	0.2806560547		L_fusiform	0.2357128834
11.	L_pericalcarine	0.2775258853		L_lingual	0.2303667725
12.	L_inferiorparietal	0.249919579		R_fusiform	0.2224770313
13.	R_insula	0.243533976		L_pericalcarine	0.2217702203
14.	L_inferiortemporal	0.2406929256		R_lateraloccipital	0.2163886563
15.	R_lingual	0.2355838232		L_cuneus	0.2143733739
16.	R_posteriorcingulate	0.2275119787		R_superiorparietal	0.2064224327
17.	L_fusiform	0.2234845007		L_supramarginal	0.2042270199
18.	R_bankssts	0.2160471626		L_lateralorbitofrontal	0.200380036
19.	L_superiorparietal	0.213339904		L_lateraloccipital	0.1888223609
20.	R_cuneus	0.2122297907		R_rostralanteriorcingulate	0.1811606572

Supplementary Table 13: Canonical Correlation Analysis ROIs loadings (CT)

Legend: Top 20 loadings of Desikan Regions of Interests (ROIs) on the 2 main CCA dimensions (r=0.69, 0.65; statistically significant at p-value<0.05). Loadings are in absolute values and listed in decreasing order.

	ROI	|CC1|		ROI	|CC2|
1.	L_bankssts	0.6521173442		L_cuneus	0.4721872379
2.	L_lateraloccipital	0.6217516421		R_precentral	0.4287453486
3.	L_insula	0.5772778171		L_caudalmiddlefrontal	0.3853658259
4.	R_bankssts	0.5730431727		L_lingual	0.3836281323
5.	R_lateraloccipital	0.5575884583		L_precentral	0.3658962124
6.	R_insula	0.5274745577		R_lingual	0.3579919796
7.	R_lingual	0.4881908917		R_superiorfrontal	0.3560307613
8.	L_pericalcarine	0.4619077081		L_lateralorbitofrontal	0.3541644729
9.	L_lateralorbitofrontal	0.443217953		L_insula	0.3306590802
10.	L_entorhinal	0.4227964373		R_insula	0.3119737927
11.	R_pericalcarine	0.4101667004		L_precuneus	0.3104848234
12.	L_lingual	0.4071596452		R_cuneus	0.3006340465
13.	R_parsorbitalis	0.4015238848		R_precuneus	0.2940527986
14.	L_postcentral	0.3973332167		L_superiorparietal	0.2856777808
15.	L_caudalmiddlefrontal	0.3914227421		R_lateralorbitofrontal	0.2854749482
16.	L_middletemporal	0.3882833195		R_superiorparietal	0.2735339712
17.	R_middletemporal	0.3879412723		L_superiorfrontal	0.271298681
18.	R_caudalmiddlefrontal	0.3821714147		L_pericalcarine	0.2670122335
19.	L_inferiortemporal	0.379691104		R_parsopercularis	0.2515318479
20.	R_superiorfrontal	0.3789751734		L_postcentral	0.2398170347

Supplementary Table 14: Canonical Correlation Analysis ROIs loadings (SA)

Legend: Top 20 loadings of Desikan Regions of Interests (ROIs) on the 2 main CCA dimensions (r=0.79, 0.67; statistically significant at p-value<0.05). Loadings are in absolute values and listed in decreasing order.

Desikan-Killiany Atlas		Tail1 (<15%)			Tail2 (>85%)		
ROIname	#vertices	SA-CT	SA-VBMfs	CT-VBMfs	SA-CT	SA-VBMfs	CT-VBMfs
bankssts	4333	0	0	0	0	0	0
caudalanteriorcingulate	3047	7	61	28	0	0	0
caudalmiddlefrontal	7230	0	0	0	0	0	0
cuneus	3268	0	0	0	0	0	0
entorhinal	2004	0	0	0	0	0	0
fusiform	9375	0	0	0	0	0	0
inferiorparietal	17547	0	0	0	0	0	0
inferiortemporal	8613	0	0	0	0	0	0
isthmuscingulate	4919	0	0	0	0	52	0
lateraloccipital	12342	0	0	0	0	0	0
lateralorbitofrontal	8542	0	0	0	0	69	0
lingual	8099	0	0	0	0	0	0
medialorbitofrontal	5454	0	0	0	0	0	0
middletemporal	9509	0	0	0	0	0	0
parahippocampal	3580	0	0	91	0	0	0
paracentral	7125	0	0	0	0	0	0
parsopercularis	5591	0	0	0	0	5	0
parsorbitalis	1902	0	0	0	0	0	0
parstriangularis	4426	0	0	0	0	0	0
pericalcarine	3735	0	0	0	0	0	0
postcentral	18657	0	0	0	0	0	0
posteriorcingulate	6260	0	0	0	0	0	0
precentral	21445	0	0	0	0	0	0
precuneus	15283	0	0	0	0	14	0
rostralanteriorcingulate	2401	0	0	0	0	0	0
rostralmiddlefrontal	15107	0	0	0	0	0	0
superiorfrontal	24057	0	0	0	0	0	0
superiorparietal	20678	0	0	0	0	0	0
superiortemporal	14139	0	0	0	0	0	0
supramarginal	16750	0	0	0	0	0	0
frontalpole	641	0	0	0	0	0	0
temporalpole	1656	0	0	0	0	0	0

transversetemporal	1845	0	0	0	0	0	0
insula	10319	0	0	0	0	123	0

Supplementary Table 15: Intersection of overlap maps across SA-CT-VBM

Legend: Table showing the number of vertices (sum of left and right hemispheres) intersecting between pairs of SA-CT-VBM overlap maps (Deletions of 1q21.1, 15q11.2, 16p11.2, and 22q11.2) for both tails <15% (left side) and >85% (right side), mapped to Desikan Regions of Interests (ROIs). See Supplementary Figure 16 for brain maps. VBM overlap maps are projected onto fsaverage using mri_vol2surf function in Freesurfer, and are restricted to cortex for comparison. Cells are highlighted for ease of visualization. Only 7 vertices intersect across all three - SA-CT-VBMfs overlap maps - for tail 1 (<15%). No intersection between pairs of SA-CT-VBM overlap maps is observed for duplications and hence not included. SA: Surface Area; CT: Cortical Thickness; VBM: Voxel Based Morphometry; #vertices: number of vertices for a ROI in fsaverage (left and right hemisphere sum in the above table).
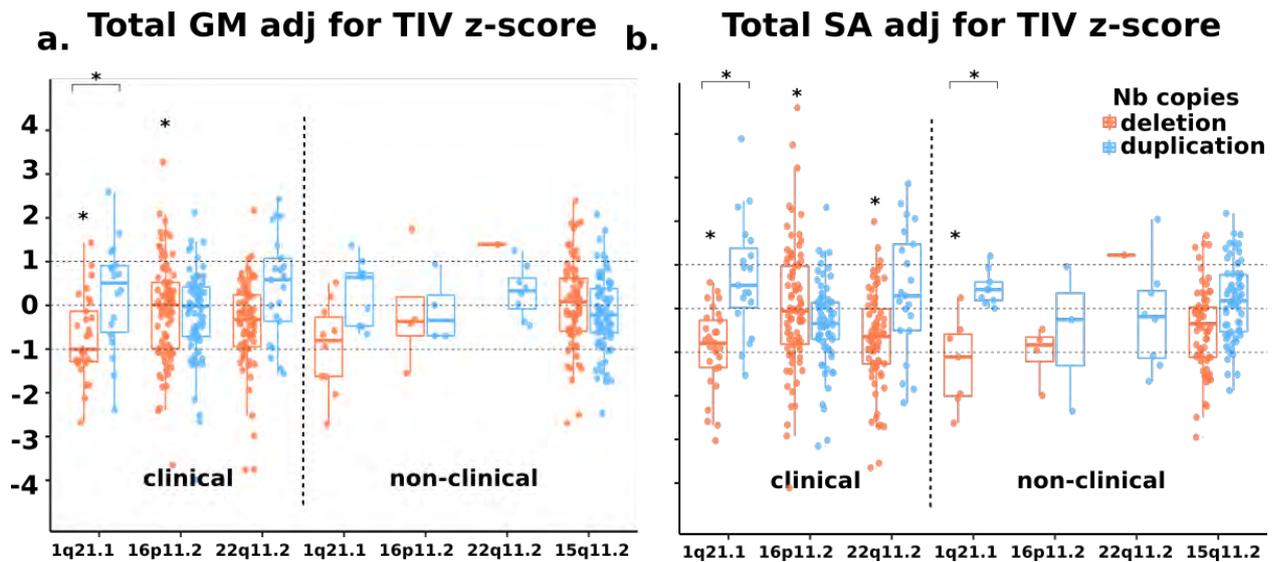
Supplementary Figure 1: The effect of 1q21.1, 16p11.2, 22q11.2 and 15q11.2 on total GM and total SA adjusted for TIV

Legend: Total grey matter volume (a) and total surface area (b) for clinically and non-clinically ascertained CNVs. Z-scores for clinically and non-clinically ascertained CNVs were calculated using 331 and 965 controls respectively, adjusting for age, age^2, sex, TIV and site as a random factor. Y axis values are z scores. X axis are CNV groups. Significant difference between CNV group and corresponding control group is indicated with a star. Horizontal bars with stars show significant differences between deletions and duplications within the same locus.
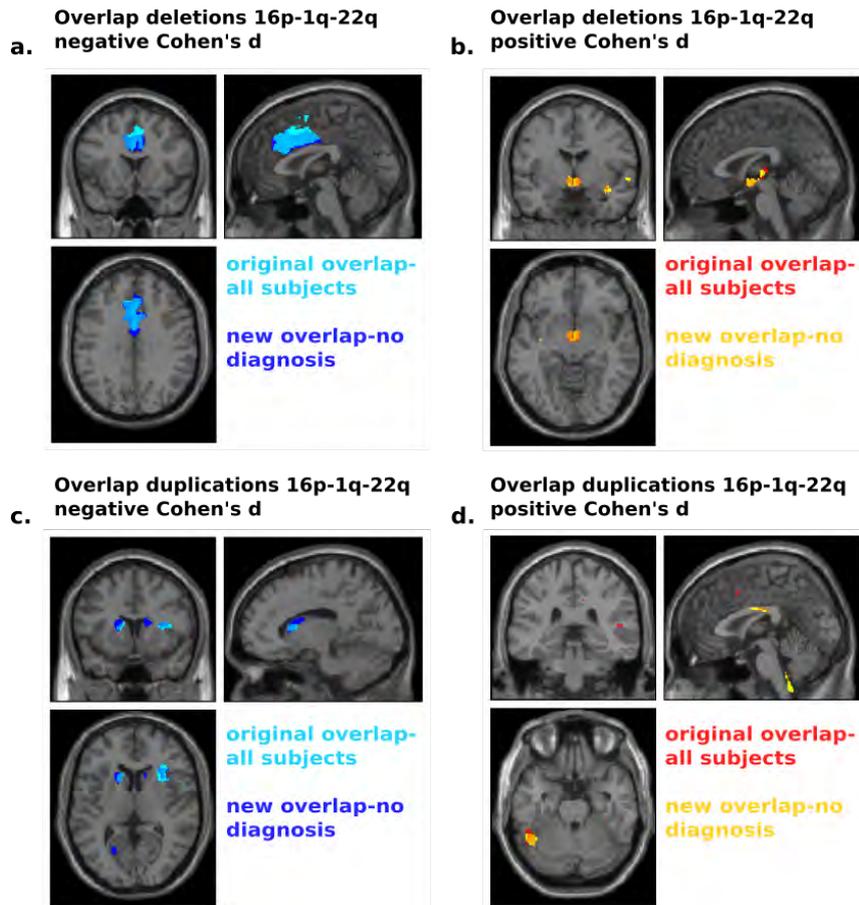
Supplementary Figure 2: Robustness of spatial overlap after removing subjects with psychiatric diagnosis

Legend: Comparison of spatial overlap between the grey matter volume effects of 3 clinically ascertained CNVs 1q21.1, 16p11.2 and 22q11.2 obtained from contrast controls including and excluding subjects with diagnosis. Overlap comparison of (a) negative effects and (b) positive effects for deletions. Overlap comparison of (c) negative effects and (d) positive effects for duplications. Original overlaps calculated from contrasts including all subjects are shown in cyan (a,c) and red (b,d). New overlaps calculated from contrast excluding subjects with a psychiatric diagnosis are shown in blue (a,c) and yellow (b,d).
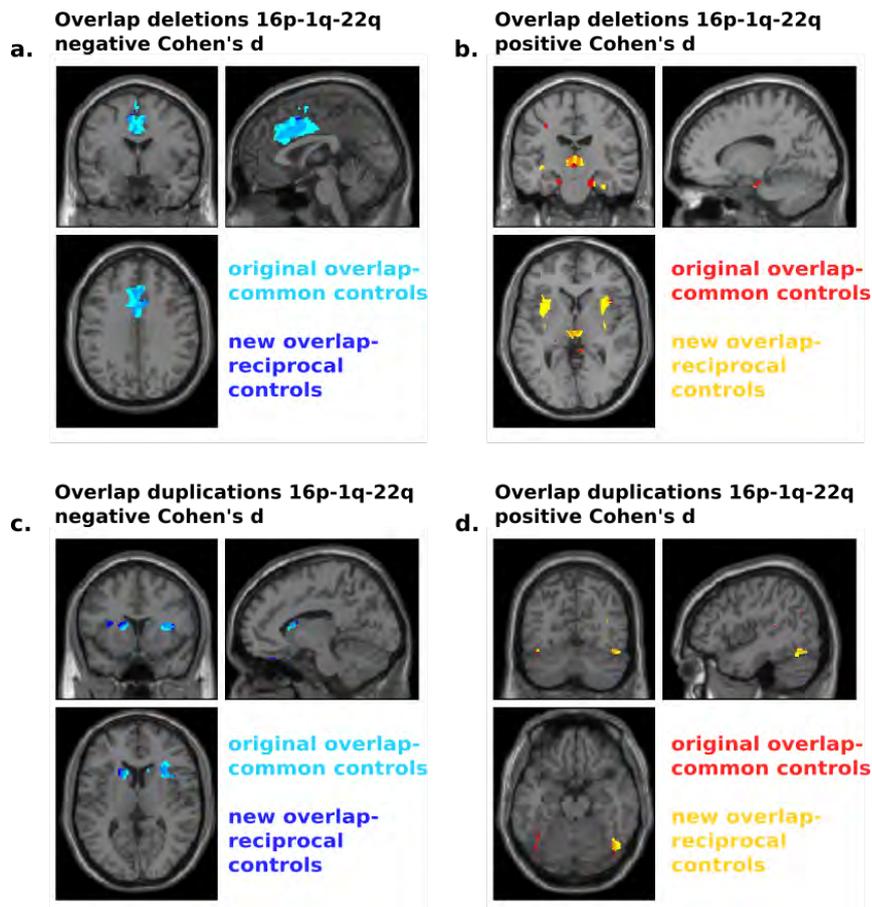
Supplementary Figure 3: Robustness of spatial overlap after matching controls by site

Legend: Comparison of spatial overlap between the grey matter volume effects of 3 clinically ascertained CNVs 1q21.1, 16p11.2 and 22q11.2 obtained from contrast with merged or contrast with separated reciprocal controls. Overlap comparison of (a) negative effects and (b) positive effects for deletions. Overlap comparison of (c) negative effects and (d) positive effects for duplications. Original overlaps calculated from contrasts with the merged group of controls are shown in cyan (a,c) and red (b,d). New overlaps calculated from contrast with reciprocal control groups for each CNV are shown in blue (a,c) and yellow (b,d).
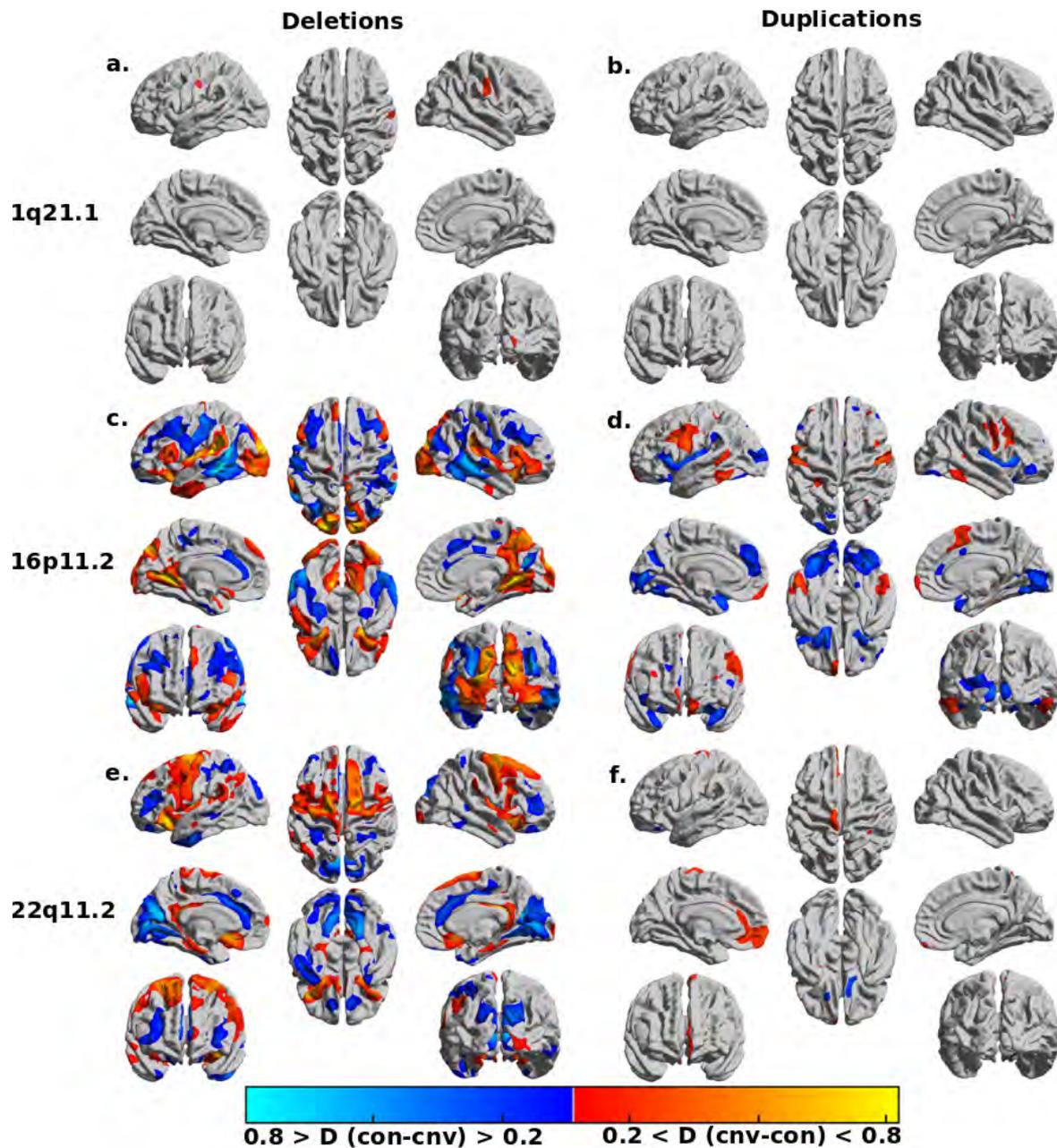
Supplementary Figure 4: FDR corrected Cohen's d maps of Surface Area

Legend: Vertex-wise brain differences in deletion and duplication carriers at the 1q21.1, 16p11.2 and 22q11.2. loci. 15q11.2 was not displayed because few vertices survived FDR correction. The left column shows regional brain differences for deletion carriers (a, c, e) while right columns show regional brain differences for duplication carriers (b, d, f). Cohen's d effect sizes were estimated in

SurfStat using vertex-wise SA estimate maps obtained from Freesurfer. Linear model was adjusted for sex, linear and quadratic expansion of age and total surface area. The results are significant at the FDR threshold q < 0.05, and shown within a range of 0.2 and 0.8 for comparison across CNVs.
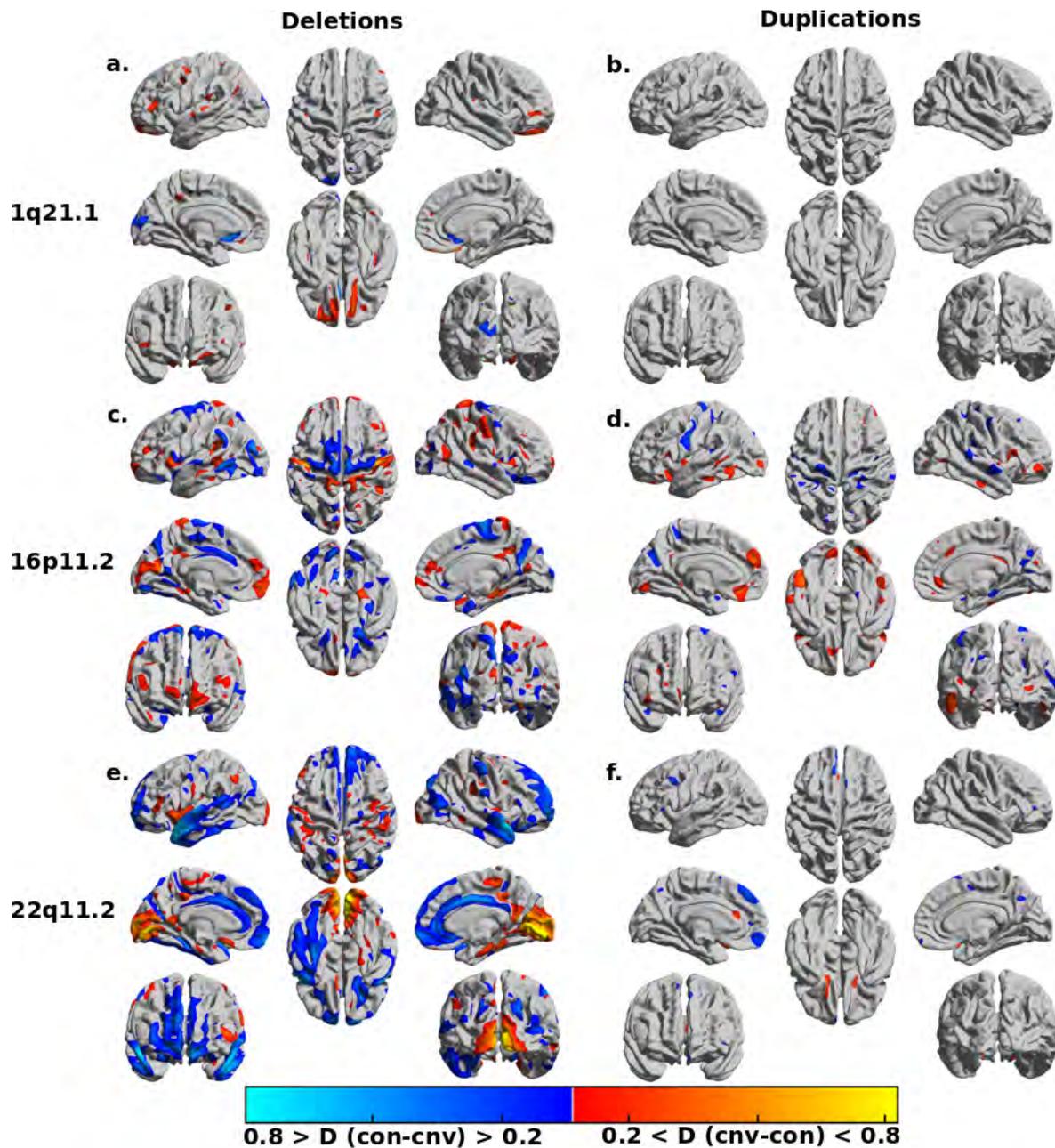
Supplementary Figure 5: FDR corrected Cohen's d maps of Cortical

Thickness

Legend: Vertex-wise brain differences in deletion and duplication carriers at the 1q21.1, 16p11.2 and

22q11.2. loci. 15q11.2 was not displayed because few vertices survived FDR correction. Left column

shows vertex-wise brain differences for deletion carriers (a, c, e) while right columns show vertex-

wise brain differences for duplication carriers (b, d, f). Cohen's d effect sizes were estimated in

SurfStat using vertex-wise CT estimate maps obtained from Freesurfer. Linear model was adjusted for sex, linear and quadratic expansion of age and mean cortical thickness. The results are significant at the FDR threshold $q < 0.05$ and shown within a range of 0.2 and 0.8 for comparison across CNVs.
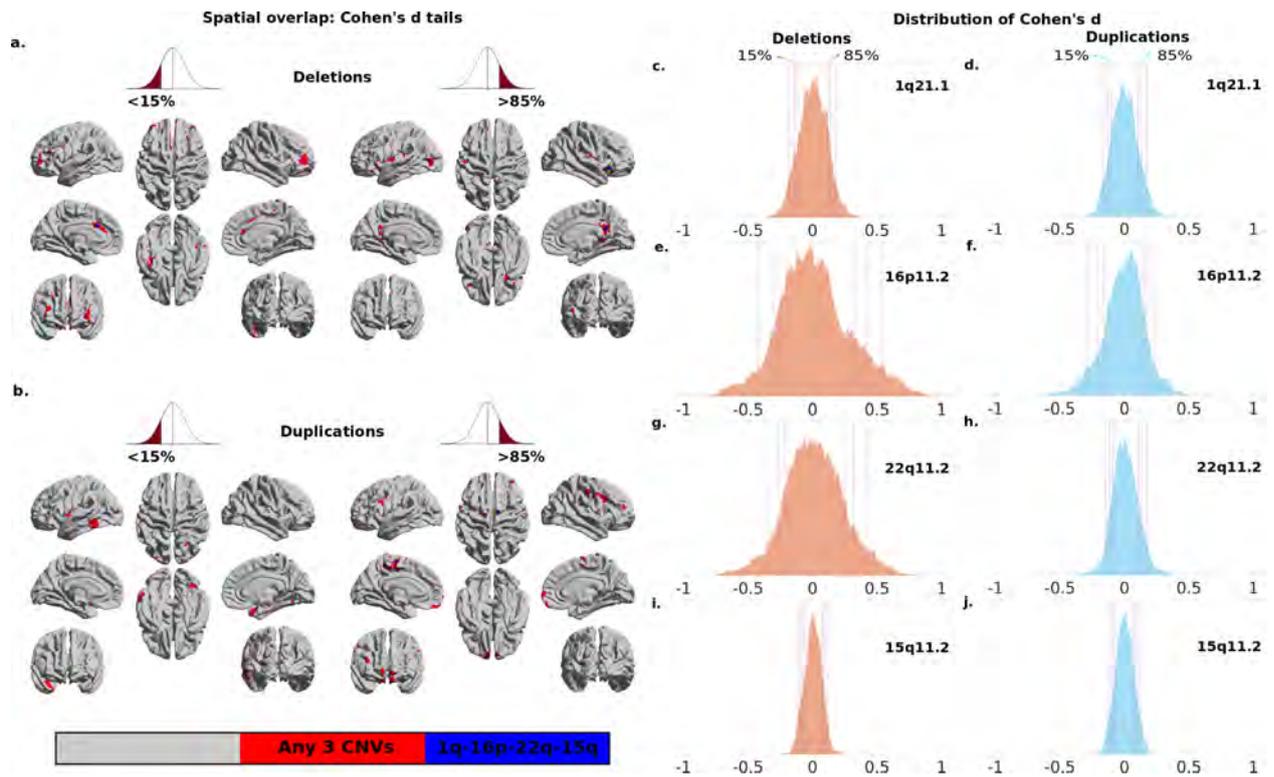
Supplementary Figure 6: Spatial overlap for surface area (SA) across deletions and duplications

Legend: Spatial overlap across clinically and non-clinically ascertained deletions (a) and duplications (b) at 4 genomic loci shown separately for <15th and >85th percentile of Cohen's d values. Overlap of all four deletions (a) or all four duplications (b) is shown in blue. Overlaps of any combination of three deletions (a) or any combination of three duplications (b) are shown in red. Top ranking Cohen's d values used in (a, b) are presented on the density plots for all eight deletions and duplications: 1q21.1 (c, d), 16p11.2 (e, f), 22q11.2 (g, h), and 15q11.2 (i, j). The X axes values of the 8 density plots are Cohen's d.
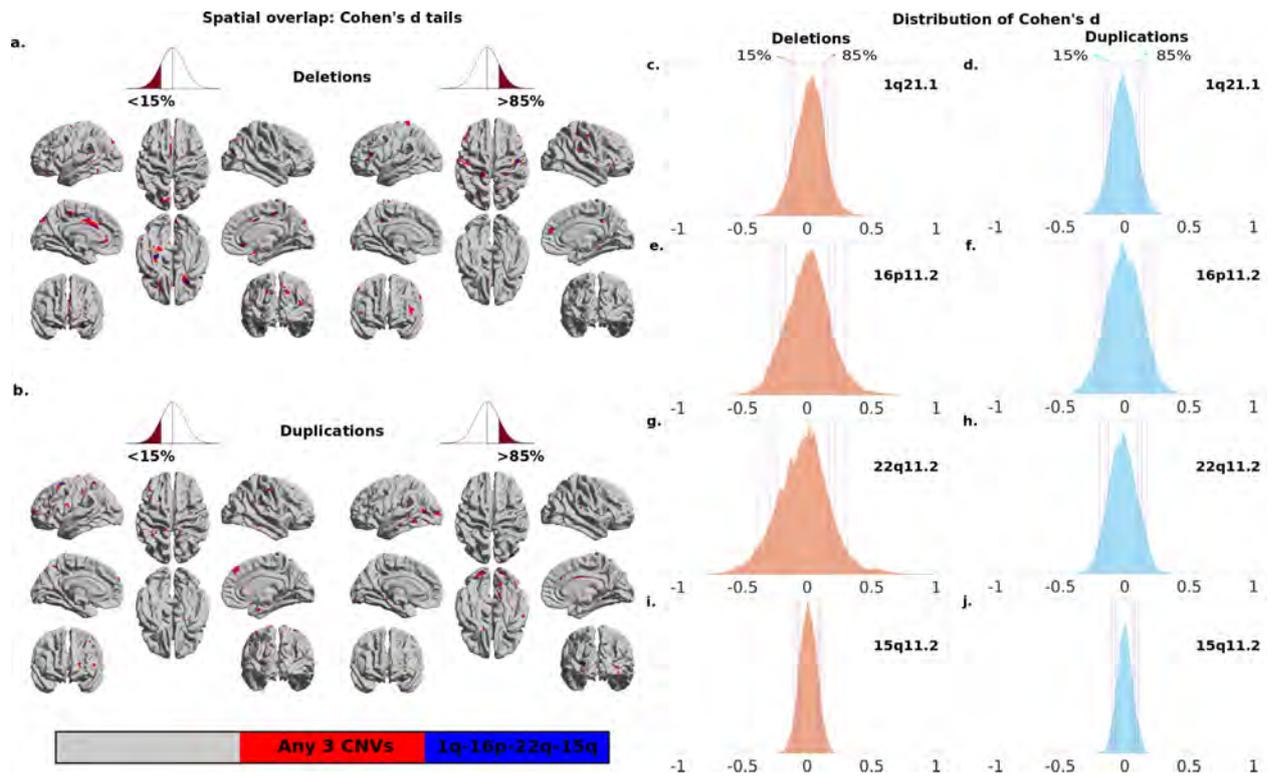
Supplementary Figure 7: Spatial overlap for mean cortical thickness (CT) across deletions

Legend: Spatial overlap across clinically and non-clinically ascertained deletions (a) and duplications (b) at 4 genomic loci shown separately for <15th and >85th percentile of Cohen's d values. Overlap of all four deletions (a) or all four duplications (b) is shown in blue. Overlaps of any combination of three deletions (a) or any combination of three duplications (b) are shown in red. Top ranking Cohen's d values used in (a, b) are presented on the density plots for all eight deletions and duplications: 1q21.1 (c, d), 16p11.2 (e, f), 22q11.2 (g, h), and 15q11.2 (i, j). The X axes values of the 8 density plots are Cohen's d.
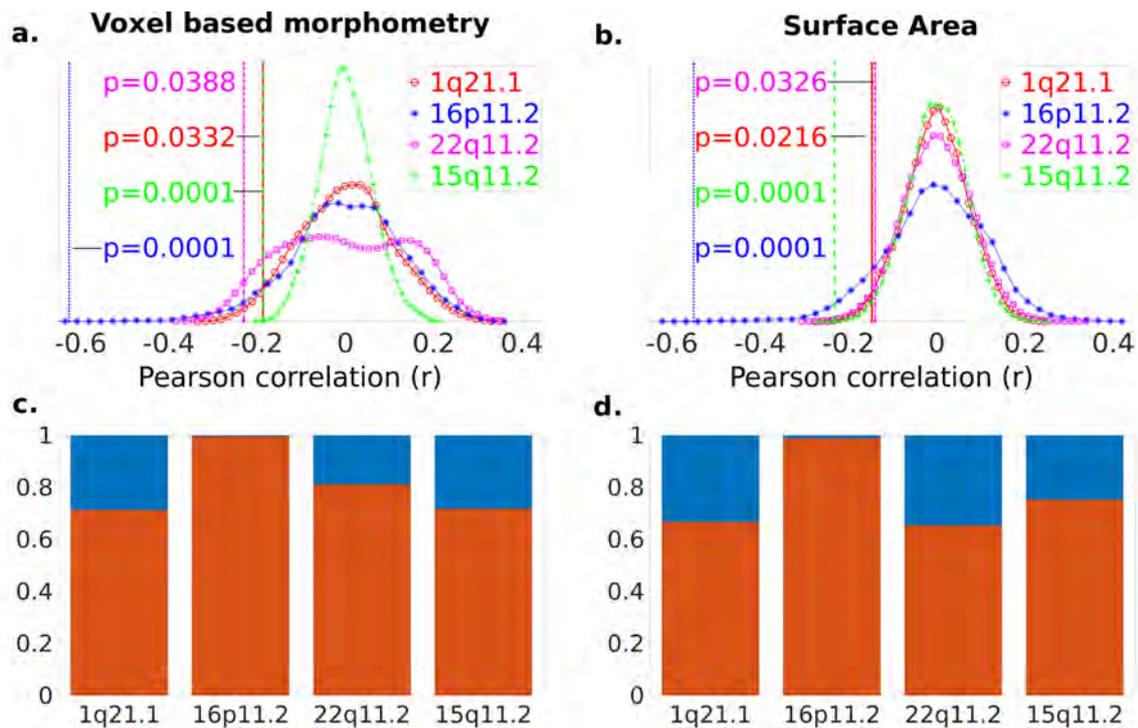
Supplementary Figure 8: Cortex-wide mirror effects between deletions and duplications

Legend: Pearson correlations between Cohen's d values for deletions and duplications. Voxel-Based Morphometry (a) and Surface Area (b) are adjusted for total gray matter and total surface area respectively. The 4 vertical lines represent the correlation (Pearson r) between deletions and duplications at each locus: 1q21.2 (red), 15q11.2 (green), 16p11.2 (blue), and 22q11.2 (magenta), with the corresponding empirical p-values (uncorrected) shown next to them in same color code. The 4 density plots represent the distribution of Pearson Correlations obtained by performing 10,000 spin permutations of duplication maps while keeping deletion maps fixed. Negative correlations between deletions and duplication are observed across loci and are significantly different (Bonferroni) from the null distributions for 16p11.2 and 15q11.2 (p-values are uncorrected). X axis = Pearson r coefficients, y axis = the surface under the curve is 100% of the distribution. (c-d) Mirror effects between deletions and duplications at both tails of the distribution c) The red bar is the proportion of voxels that are in the top 85th percentile for deletions and the lower 15th percentile for duplications and vice versa. The

blue bar represents voxels that are either in the top 85th or lower 15th percentile for both deletions and duplications. d) The same bar plots are presented for surface area. (All Correlation values are reported in supplemental eTable 8-10).

a. Scatter plot of brain regions in PC1 (x-axis, −6 to 6) vs PC2 (y-axis, −3 to 4) space.

b. Heatmap of PC1 and PC2 loadings:

	PC1	PC2
1q21.1del	−0.47	0.095
1q21.1dup	0.46	−0.036
16p11.2del	−0.27	0.56
16p11.2dup	0.36	−0.43
22q11.2del	0.43	0.42
22q11.2dup	−0.38	−0.31
15q11.2del	0.21	0.31
15q11.2dup	0.038	0.36

c. Heatmap of rsq-PC1 and rsq-PC2:

	rsq-PC1	rsq-PC2
1q21.1del	54	1.9
1q21.1dup	52	0.28
16p11.2del	19	68
16p11.2dup	33	39
22q11.2del	46	37
22q11.2dup	35	20
15q11.2del	11	20
15q11.2dup	0.37	27

d. PC1 and PC2 brain surface maps.

e. Scatter plot of brain regions in PC1 (x-axis, −4 to 4) vs PC2 (y-axis, −3 to 3) space.

f. Heatmap of PC1 and PC2 loadings:

	PC1	PC2
1q21.1del	−0.23	0.58
1q21.1dup	0.14	−0.43
16p11.2del	0.39	−0.3
16p11.2dup	−0.11	0.4
22q11.2del	0.58	0.28
22q11.2dup	−0.6	−0.17
15q11.2del	0.25	0.35
15q11.2dup	−0.028	−0.047

g. Heatmap of rsq-PC1 and rsq-PC2:

	rsq-PC1	rsq-PC2
1q21.1del	11	56
1q21.1dup	4.3	31
16p11.2del	32	15
16p11.2dup	2.4	27
22q11.2del	70	13
22q11.2dup	76	5
15q11.2del	13	20
15q11.2dup	0.16	0.37

h. PC1 and PC2 brain surface maps.

43

Supplementary Figure 9: Principal Component Analysis of brain alterations associated with 8 CNVs for SA and CT

Legend: Top half (a, b, c, d) corresponds to SA, and bottom half (e, f, g, h) to CT. (a,e) Loading of ROIs on the 2 PCA dimensions; the font size is correlated to the region's contribution to dimensions. ROI names are color coded as being part of the deletion (red), duplication (blue) and both deletion and duplication (magenta) convergence maps. PCA was computed on z-scored Cohen's d values, with the 8 CNVs as variables and 68 Desikan ROIs (FreeSurfer derived) as observations. Cohen's d for SA/CT regions were obtained using linear models adjusted for Total-SA/Mean-CT, age, age2, sex and site. The first 2 components explained respectively 31.25 and 26.75 % of the variance for SA; and 26.13 and 20.79 % of the variance for CT. (b,f) Loading of 8 CNVs on the 2 PCA dimensions. Values are PC loading magnitudes and represent the contribution of a CNV to the PC. (c,g) Variance explained (rsq) of each CNV Cohen's d profile by PC1 and PC2. Values and color scale represent the "percent of variance". (d,h) PC1 and PC2 projected on the brain. The darker the red or blue color, the stronger the positive or negative loadings for PC1 and PC2.
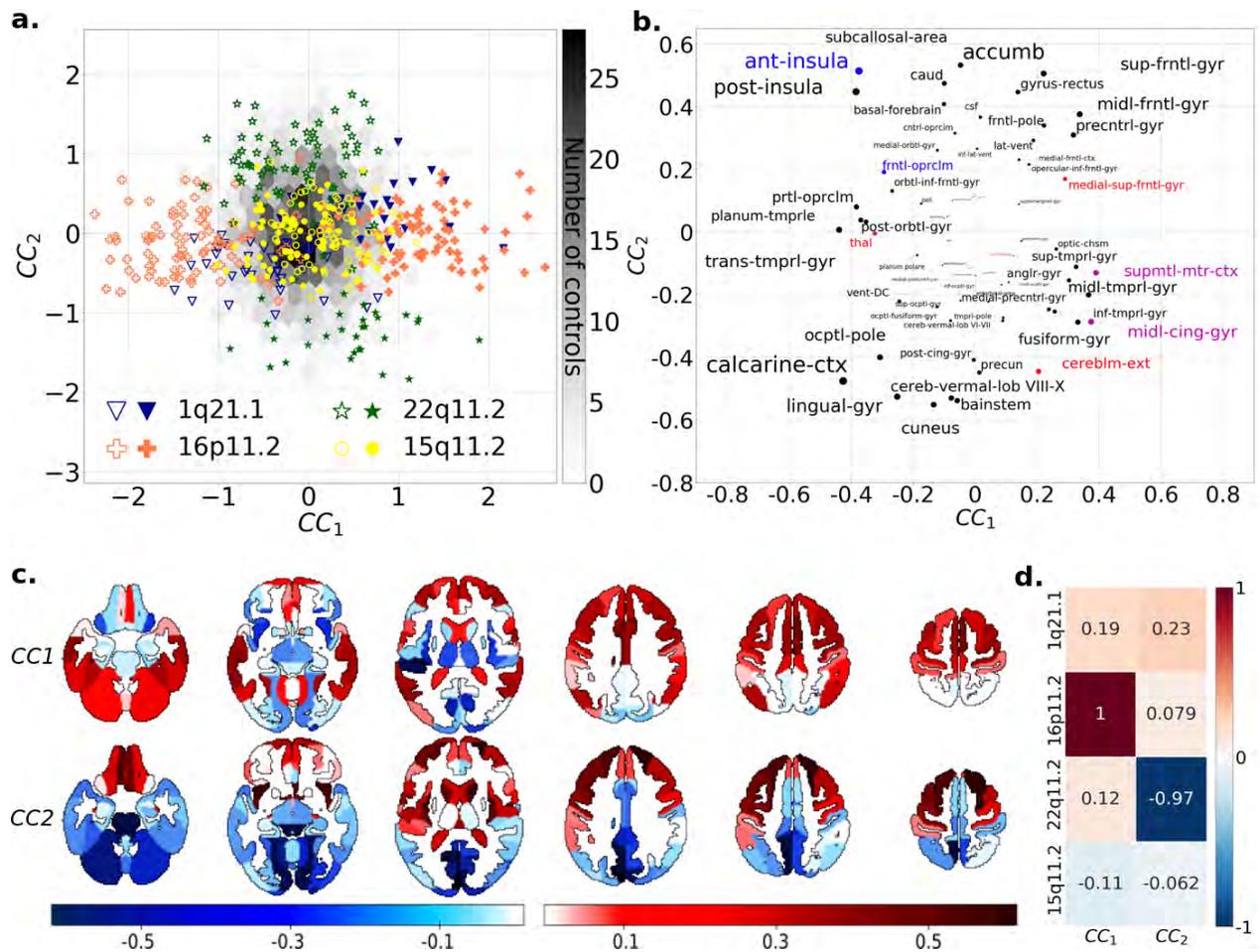
Supplementary Figure 10: Co-analysis of shared brain alterations due to 8 CNVs (4 Genomic Loci)

Legend: Legend: (a) Scatterplot showing the position of each of the 484 carriers of 8 different CNVs along 2 dominant brain-gene Canonical Correlation (CC) dimensions established using 130 neuroanatomical GM regions of CNV carriers. GM region volumes were obtained using neuromorphometric and were adjusted for total grey matter, age, age^2, sex and site. The empty and full symbols represent deletions and duplication respectively. The grey hexagonal bin plot represents the frequency of controls (n=1296). Controls were not used to calculate the CCA and were projected post hoc on the 2 dimensions using CCA prediction. X and Y axis values: z-scores of regional volumes. (b) Loading of Neuromorphometric Regions of Interests (ROIs) on the 2 CC dimensions.

The font size is correlated to the region's contribution to dimensions. ROI names are color coded as being part of the deletion (red), duplication (blue) and both deletion and duplication (magenta) convergence patterns. (c) CCA dimension 1 and 2 regional relevances projected on axial brain slices. The darker the red or blue color, the stronger the positive or negative association with the CCA dimensions. (d) Loading of the first and second CCA dimension on 4 CNV genomic loci. Values are CCA loading magnitudes and represent the contribution of a CNV loci to the canonical dimension.
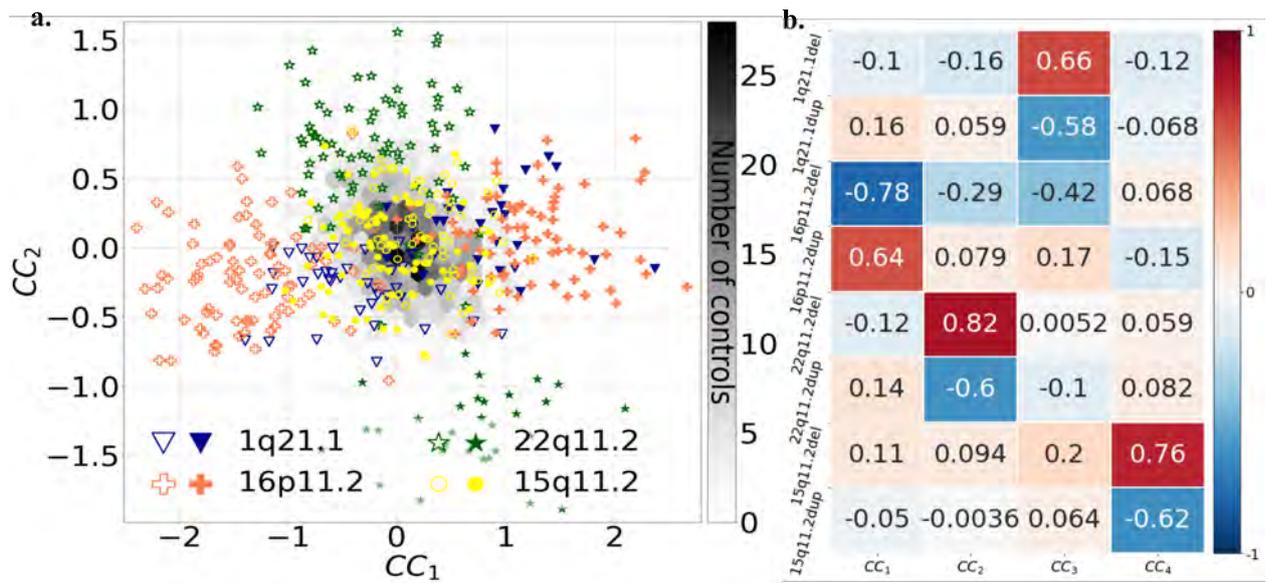
Supplementary Figure 11: CCA analysis of 8 CNVs (Deletion and Duplication as independent groups)

Legend: (a) Scatterplot showing the position of each of the 484 carriers of 8 different CNVs along 2 Canonical Correlation (CC) dimensions established using 130 neuroanatomical GM regions of CNV carriers. GM region volumes were obtained using neuromorphometric and were adjusted for total grey matter, age, age^2, sex and site. The empty and full symbols represent deletions and duplication respectively. The grey hexagonal bin plot represents the frequency of controls (n=1296). Controls were not used to compute the CCA and were projected post hoc on the 2 dimensions using CCA prediction. (b) Loading of the first and second CCA dimension on 8 CNVs. Values are CCA loading magnitudes and represent the contribution of a CNV to the canonical dimension.

47
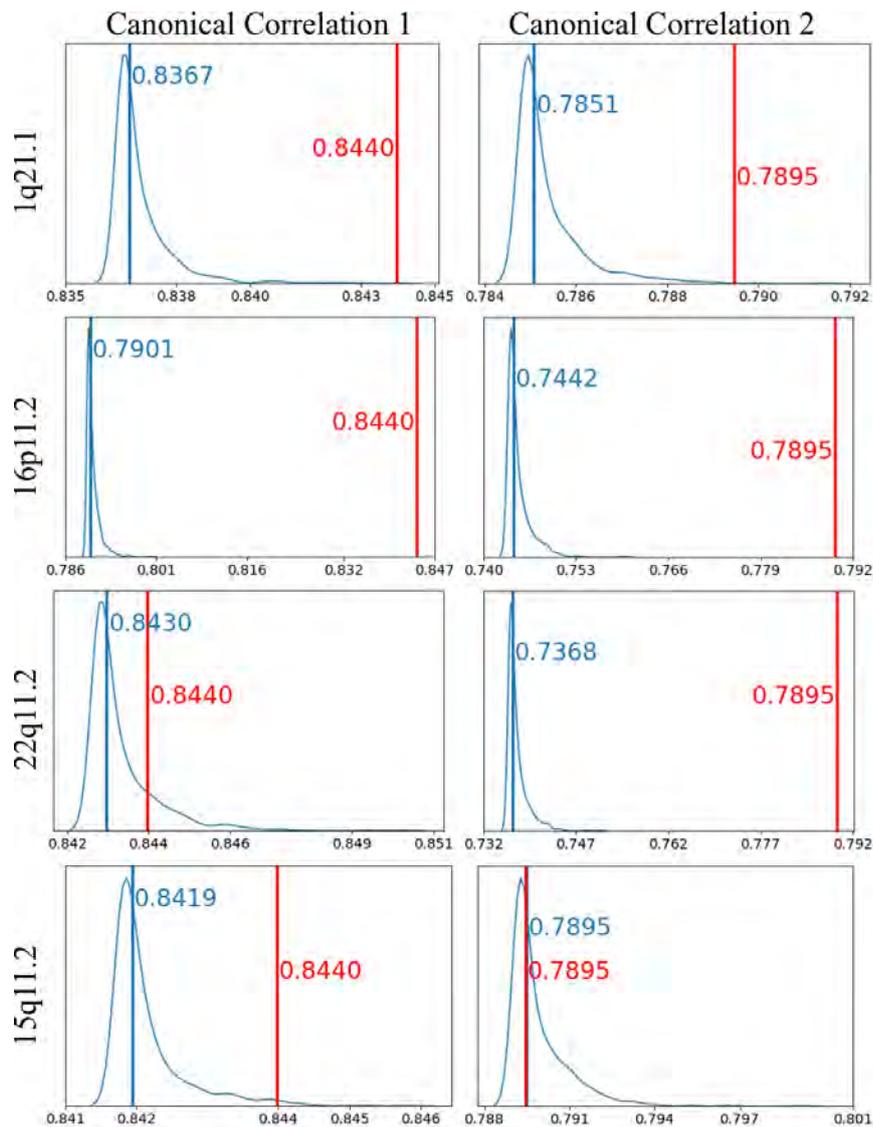
Supplementary Figure 12: Impact of each genomic loci on Canonical Correlations

Legend: Density plots of canonical correlations (CC1: left column; CC2: right column) obtained by shuffling labels 1000 times for each genomic loci separately. The red vertical line is the original Canonical Correlation values including all genomic loci. red X-axis: Canonical Correlation. Y-axis: Frequency. Blue vertical line: median of null Canonical Correlations. We consistently observe that the original value is significantly different from the distribution of the randomly permuted values.
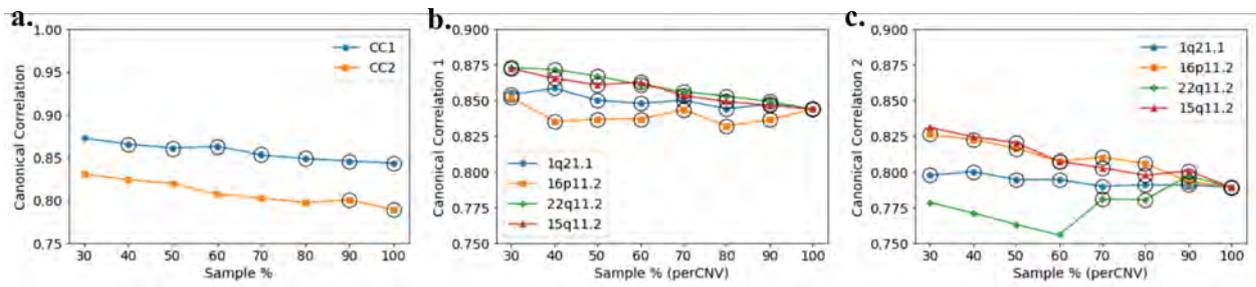
Supplementary Figure 13: Impact of sample size on CCA analysis

Legend: (a) Plot showing the impact of sample size on Canonical Correlation 1 and 2, for all CNV carriers (4 CNV loci). X-axis: Sample %; Y-axis: Canonical Correlation. Significant CCs are encircled in black. (b-c) Plots showing the impact of sample size of single CNV loci on Canonical Correlation 1 (b), and Canonical Correlation 2 (c).
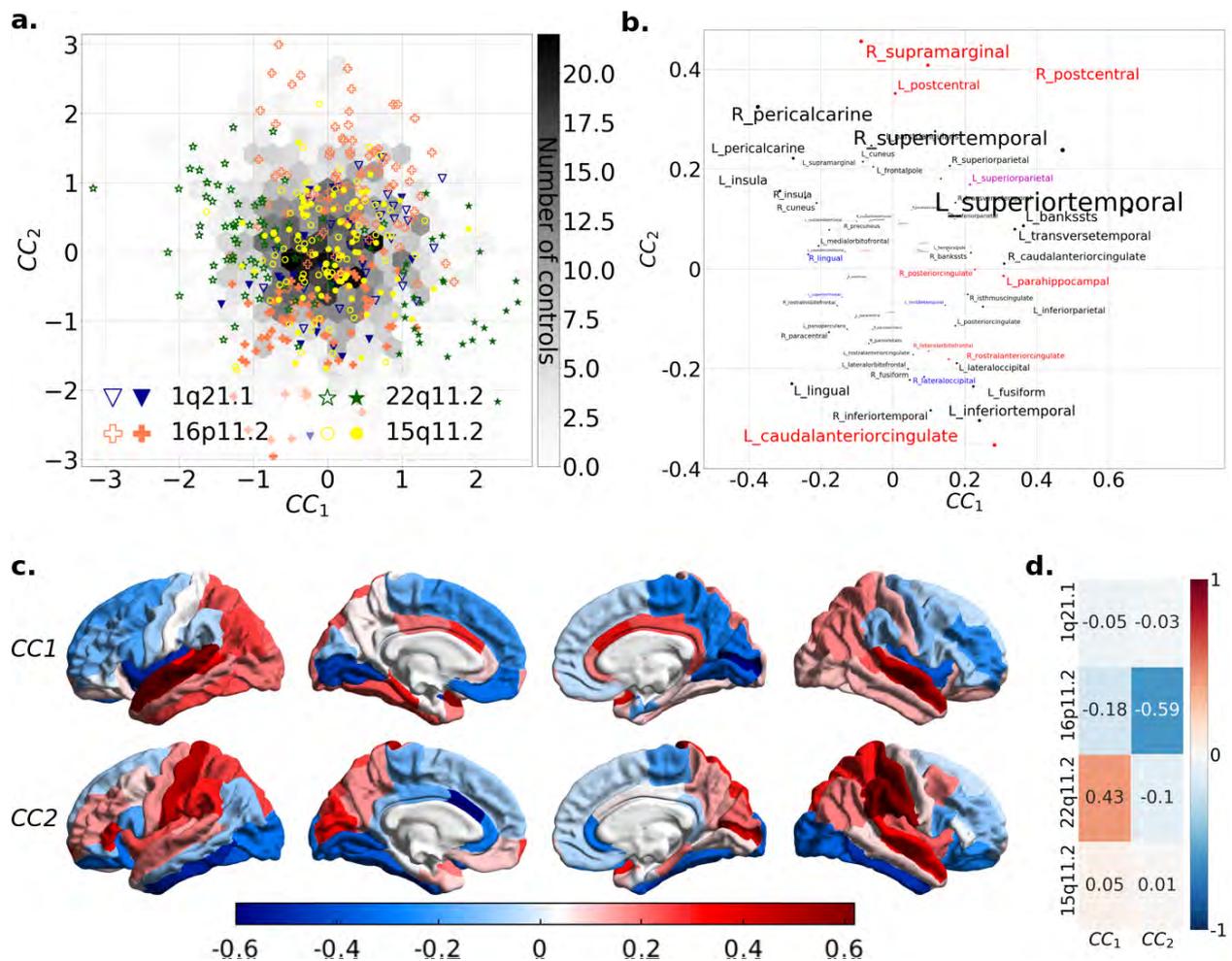
Supplementary Figure 14: CCA analysis across 4 Genomic Loci, Cortical Thickness

Legend: (a) Scatterplot showing the position of each of the 468 carriers of 8 different CNVs along 2 Canonical Correlation (CC) dimensions established using 68 Freesurfer regions of CNV carriers. Cortical Thickness (CT) was calculated using Freesurfer recon-all and was adjusted for mean CT, age, age^2, sex, and site. The empty and full symbols represent deletions and duplication respectively. The grey hexagonal bin plot represents the frequency of controls (n=1252). Controls were not used to compute the CCA and were projected post hoc on the 2 dimensions using CCA prediction. X and Y axis values: z-scores of regional volumes. (b) Loading of Freesurfer Regions of Interests (ROIs) on the 2 CC dimensions. The font size is correlated to the region's contribution to CC dimensions. ROI names are color-coded as being part of the deletion (red), duplication (blue) and both deletion and

50

duplication (magenta) convergence patterns. (c) CCA dimensions 1 and 2 projected on the brain. Brain regions most strongly associated with CCA dimensions 1 and 2. The darker the red or blue color, the stronger the positive or negative association with the CCA dimensions. (d) Loading of the first and second CCA dimensions on 4 CNV genomic loci. Values are CCA loading magnitudes and represent the contribution of a CNV locus to the canonical dimension. CCA was calculated using 449 CNV carriers without controls and identified 2 dimensions (r=0.69, 0.65 statistically significant at p-value<0.05). Top 20 ROI loadings (absolute values) for CC1 and CC2 are listed in eTable 7.
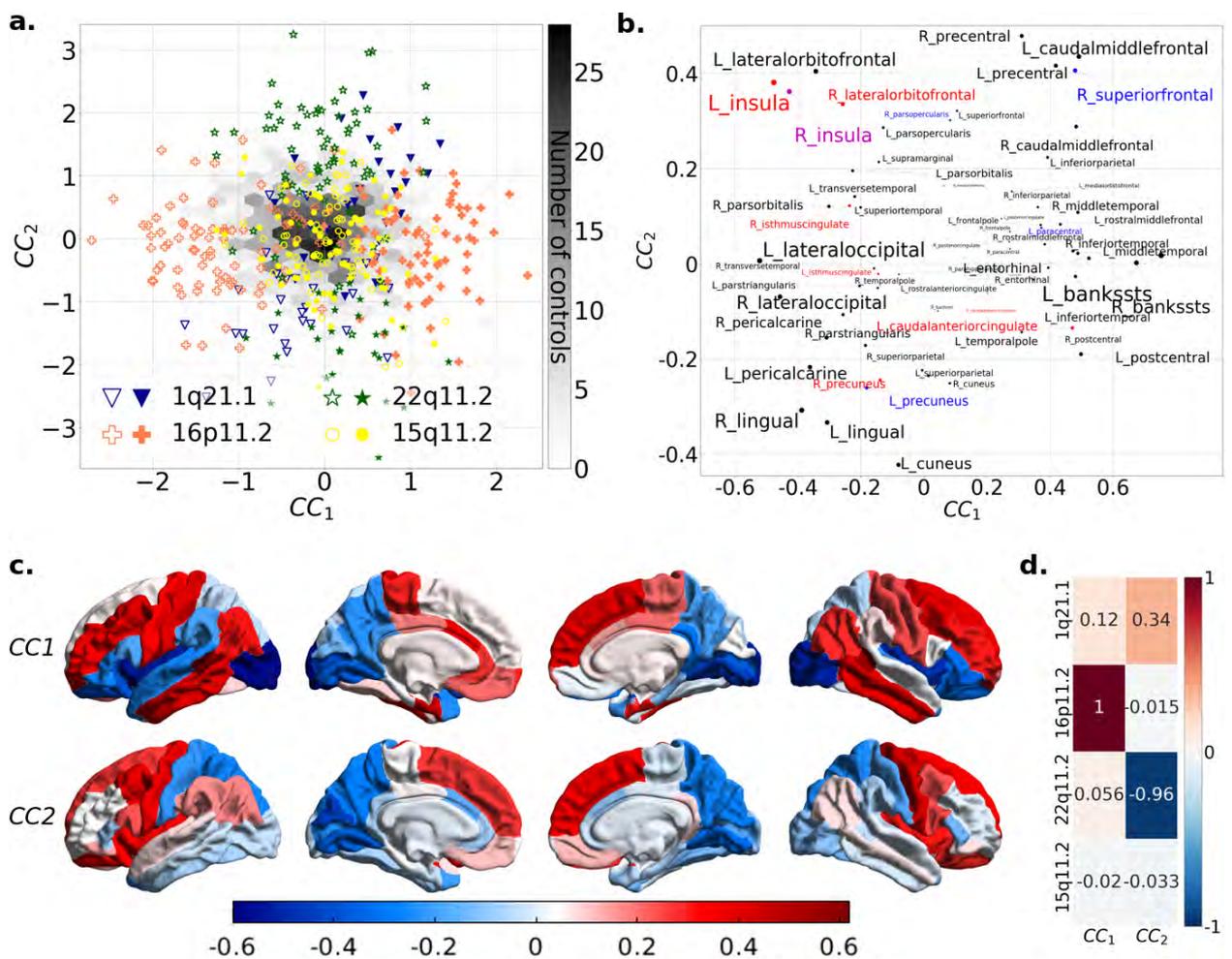
Supplementary Figure 15: CCA analysis across 4 Genomic Loci, Surface Area

Legend: (a) Scatterplot showing the position of each of the 468 carriers of 8 different CNVs along 2 Canonical Correlation (CC) dimensions established using 68 Freesurfer regions of CNV carriers. Surface Area (SA) was calculated using Freesurfer recon-all and were adjusted for mean total SA, age, age^2, sex and site. The empty and full symbols represent deletions and duplication respectively. The grey hexagonal bin plot represents the frequency of controls (n=1252). Controls were not used to compute the CCA and were projected post hoc on the 2 dimensions using CCA prediction. X and Y axis values: z-scores of regional volumes. (b) Loading of Freesurfer Regions of Interests (ROIs) on the 2 CC dimensions. The font size is correlated to the region's contribution to CC dimensions. ROI names are color coded as being part of the deletion (red), duplication (blue) and both deletion and duplication (magenta) convergence patterns. (c) CCA dimensions 1 and 2 projected on the brain. Brain regions most strongly associated with CCA dimensions 1 and 2. The darker the red or blue color, the stronger the positive or negative association with the CCA dimensions. (d) Loading of the first and second CCA dimensions on 4 CNV genomic loci. Values are CCA loading magnitudes and represent the contribution of a CNV loci to the canonical dimension. CCA was calculated using 449 CNV carriers without controls and identified 2 dimensions (r=0.79, 0.67 statistically significant at p-value<0.05). Top 20 ROI loadings (absolute values) for CC1 and CC2 are listed in eTable 8.

Supplementary Figure 16: Intersection of deletion and duplication overlap maps of 4 CNVs across SA-CT-VBM

Legend: Deletion (left) and duplication (right) overlap maps for 1q21.1, 15q11.2, 16p11.2, and 22q11.2 and intersection between pairs of SA-CT-VBM overlap maps for both tails <15% (top row) and >85% (bottom row). For comparison, VBM overlap maps are projected onto fsaverage using mri_vol2surf function in Freesurfer. See Supplementary Table 15 for a summary of the number of intersecting vertices within Desikan ROIs between pairs of SA-CT-VBM. More intersection is observed for SA-VBM (purple) as opposed to CT-VBM or SA-CT, mainly in insula, isthmus cingulate, and lateral orbitofrontal. Only 7 vertices intersect (in caudal anterior cingulate) across all three - SA-CT-VBM deletion overlap maps - for tail 1 (<15%). No intersection between pairs of SA-CT-VBM overlap maps is observed for duplications. SA: Surface Area; CT: Cortical Thickness; VBM: Voxel Based Morphometry;

References

1. Lin, A. *et al.* Mapping 22q11.2 Gene Dosage Effects on Brain Morphometry. *J. Neurosci.* **37**, 6183–6199 (2017).

2. Maillard, A. M. *et al.* The 16p11.2 locus modulates brain structures common to autism, schizophrenia and obesity. *Mol. Psychiatry* **20**, 140–147 (2015).

3. Qureshi, A. Y. *et al.* Opposing Brain Differences in 16p11.2 Deletion and Duplication Carriers. *J. Neurosci.* **34**, 11199–11211 (2014).

4. Martin-Brevet, S. *et al.* Quantifying the Effects of 16p11.2 Copy Number Variants on Brain Structure: A Multisite Genetic-First Study. *Biol. Psychiatry* **84**, 253–264 (2018).

5. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).

6. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Med.* **12**, e1001779 (2015).

7. Huguet, G. *et al.* Measuring and Estimating the Effect Sizes of Copy Number Variants on General Intelligence in Community-Based Samples. *JAMA Psychiatry* **75**, 447–457 (2018).

8. Douard, E. *et al.* Effect Sizes of Deletions and Duplications on Autism Risk Across the Genome. *Am. J. Psychiatry* **178**, 87–98 (2020).

9. Ashburner, J. & Friston, K. J. Unified segmentation. *NeuroImage* **26**, 839–851 (2005).

10. Lorio, S. *et al.* New tissue priors for improved automated classification of subcortical brain structures on MRI. *Neuroimage* **130**, 157–166 (2016).

11. Ashburner, J. A fast diffeomorphic image registration algorithm. *NeuroImage* **38**, 95–113 (2007).

12. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006).

13. Worsley, K. J. *et al.* SurfStat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. *Neuroimage* (2009) doi:10.1016/S1053-8119(09)70882-1.

14.   Cohen, J. *Statistical power analysis for the behavioral sciences*. (Psychology Press, 1988).

15.   Alexander-Bloch, A. *et al.* On testing for spatial correspondence between maps of human brain structure and function. *NeuroImage* **178**, 540–551 (2018).

16.   Reardon, P. K. *et al.* Normative brain size variation and brain shape diversity in humans. *Science* **360**, 1222–1227 (2018).

17.   Smith, S. M. *et al.* A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat. Neurosci.* **18**, 1565–1567 (2015).

18.   Wang, H.-T. *et al.* Finding the needle in high-dimensional haystack: A tutorial on canonical correlation analysis. *NeuroImage* (2020).

19.   Kernbach, J. M. *et al.* Subspecialization within default mode nodes characterized in 10,000 UK Biobank participants. *Proc. Natl. Acad. Sci.* **115**, 12295–12300 (2018).

20.   Bzdok, D. & Yeo, B. T. T. Inference in the age of big data: Future perspectives on neuroscience. *NeuroImage* **155**, 549–564 (2017).

Supplement material study 2: Effects of rare CNVs and common variants on subcortical structure

INDEX

MRI data acquisition and quality check

Participants

Clinically ascertained CNV carriers were recruited as either probands referred for genetic testing, or as relatives. Controls were either non-carriers within the same families or individuals from the general population. We pooled data from 5 different cohorts: Cardiff University (UK), 16p11.2 European Consortium (Lausanne, Switzerland), University of Montreal (Canada), UCLA (Los Angeles, USA) and the Variation in individuals Project (SVIP, USA). A subset of the participants with 16p11.2 and 22q11.2 CNVs were included in prior publications [1–4] (Supplementary Table 1). CNVs from non-clinical populations were identified in the UK Biobank [5,6]. PennCNV and QuantiSNP were used, with standard quality control metrics, to identify CNVs [7,8].

16p11.2 European Consortium

MRI data of the EU participants were acquired on two 3T whole-body scanners. 14 carriers of a 16p11.2 deletion and 17 duplication carriers, together with 59 controls (21 familial and 38 unrelated controls) were examined on a Magnetom TIM Trio (Siemens Healthcare, Erlangen, Germany), using a 12-channel RF receive head coil and RF body transmit coil. The remaining 16p11.2 (13 deletions, 6 duplications), 1q11.2 (9 deletions, 7 duplications) carriers and controls (n=38) in the european cohort, were scanned on a Magnetom Prisma Syngo (Siemens Healthcare, Erlangen, Germany) using a 64-channel RF receive head coil and RF body transmit coil. T1-weighted (T1w) anatomical images acquired with the TIM Trio scanner used a Multi-Echo Magnetization Prepared RApid Gradient Echo sequence (ME-MPRAGE: 176 slices; 256×256 matrix; echo time (TE): TE1 = 1.64 ms, TE2 = 3.5 ms, TE3 = 5.36 ms, TE4 = 7.22 ms; repetition time (TR): 2530 ms; flip angle 7°). On the Prisma Syngo scanner, T1w images were acquired using a single-echo MPRAGE sequence (176 slices; 256×256 matrix; TE = 2.39 ms; TR = 2000 ms; flip angle 9°).

1

Simons Searchlight Consortium

Data were acquired using multi and single-echo sequences. 176 participants (38 del/ 34 dup 16p11.2 carriers, 2 dup 1q21.1 carriers and 102 familial controls) underwent the research MRI protocol at two imaging core sites on matched 3T Magnetom TIM Trio MRI scanners (Siemens Healthcare, Erlangen, Germany), using the vendor-supplied 32-channel phased-array radio-frequency head coils. 68 participants were scanned at University of California sites (UC) and 108 at the Children Hospital of Philadelphia (CHOP). Structural MRI data included multi-echo T1w ME-MPRAGE using the following parameters: 176 slices, 256×256 matrix, TR = 2530 ms, TI = 1200 ms, TE = 1.64 ms, and flip angle 7°. Clinical MRI images (single-echo) obtained at the phenotyping core sites were also analyzed. The remaining 79 subjects (19 del/ 13 dup 16p11.2 carriers, 12 del/8 dup 1q21.1 carriers and 27 familial controls) were scanned at University of Washington Medical Center, Baylor University Medical Center and Boston Children's Hospital on two matched 3T Philips Achieva (Philips Healthcare, United States of America) and one unmatched Magnetom TIM Trio scanner (Siemens Healthcare, Erlangen, Germany), respectively. T1w images were acquired using a single-echo MPRAGE sequence and the following parameters: 160 slices; 256×256 matrix; TE = 2.98 ms; TR = 2300 ms; flip angle 9°. All multi-echo images were averaged following a Root-Mean Square (RMS) averaging method.

Brain Canada

MRI scans for the Brain Canada cohort have been performed at the Montreal Neurological Institute with the same 3T scanner: Magnetom Prisma Syngo (Siemens Healthcare, Erlangen, Germany). Data included 16p11.2 (3 deletions, 3 duplications), 1q11.2 (5 deletions, 1 duplication), 22q11.2 (1 duplication) carriers and controls (n=26) T1w images were acquired using using MPRAGE sequences, scanning protocol description is detailed on this website: http://www.bic.mni.mcgill.ca/users/jlewis/BrainCanada/MCIN/.

UCLA

Imaging data of 22q11.2 CNV carriers and typically developing (TD) controls were acquired at the University of California, Los Angeles (UCLA). Patients were ascertained from UCLA or Children's Hospital, Los Angeles Pediatric Genetics, Allergy/Immunology and/or Craniofacial Clinics. We excluded 11 individuals from the analysis due to insufficient quality of the imaging data (cf. Supplementary Methods, quality control). The final 22q11.2 sample includes 144 individuals (71 deletions, 19 duplications and 54 controls). Demographically comparable TD comparison subjects were recruited from the same communities as patients via web-based advertisements and by posting flyers and brochures at local schools, pediatric clinics, and other community sites. Exclusion criteria for all study participants included significant neurological or medical conditions (unrelated to 22q11.2 mutation) that might affect brain structure, history of head injury with loss of consciousness, insufficient fluency in English, and/or substance or alcohol abuse or dependence within the past 6 months. The UCLA Institutional Review Board approved all study procedures and informed consent documents. Scanning was conducted on an identical 3 tesla Siemens Trio MRI scanner with a 12-channel head coil at the University of California at Los Angeles Brain Mapping Center or at the Center for Cognitive Neuroscience.

Cardiff

Imaging acquisition in Cardiff was performed on a 3 T General Electric HDx MRI system (GE Medical Systems, Milwaukee, WI) using an eight-channel receive-only head RF coil. T1-weighted structural images were acquired with a 3D fast spoiled gradient echo (FSPGR) sequence (TR = 7.8 ms, TE = 3.0 ms, voxel size = 1 mm³ isomorphic). Data included 1 16p11.2 deletion, 1q11.2 (3 deletions, 1 duplication), 22q11.2 (3 deletions, 2 duplications) carriers and 15 controls.

MRI quality control

All MRI T1w nifti images were visually inspected by the same rater (CM) for head coverage, ghosting and susceptibility artifacts. Images were also screened after segmentation to ensure good

3

tissue classification accuracy. From the clinically ascertained dataset 55 subjects were excluded for insufficient image quality or artifacts while from the non-clinically ascertained dataset 52 subjects were excluded following the same criteria. Quality assurance protocol for Freesurfer based cortical reconstructions led to exclusion of an additional 34 scans. Numbers reported in Table 1 and Supplementary Table 1 are after exclusion.

List of abbreviations

CNV: Copy Number Variant, ASD: Autism Spectrum Disorders, SCZ: Schizophrenia, VBM: Voxel Based Morphometry, SBM: Surface Based Morphometry, TIV: Total Intracranial Volume, GM: Grey Matter, WM: White Matter, ICV: Intracranial Volume, CT: Cortical Thickness, SA: Surface Area, IQ: Intelligence Quotient, CCA: Canonical Correlation Analysis, PCA: Principal Component Analysis, pLI: Probability of being loss-of-function intolerant.

Supplementary Tables

Loci	Type	nGenes	Phenotypes/CNV-scores				Effect Sizes			
			IQ_loss	OR ASD/SCZ	pLIsum	1/LOEUF	ICV	SubCortVol	Thickness	Jacobian
1q21.1	Del	7	15	6.4	1.7	9.098	2.09	0.36	0.68	0.79
	Dup	7	25	5.3	1.7	9.098	0.72	0.81	0.62	0.6
16p11.2	Del	27	26	14.3	6.56	39.76	0.52	0.84	0.89	1.07
	Dup	27	11	11.7	6.56	39.76	1.25	0.64	0.68	1.02
22q11.2	Del	49	28.8	32.3	10.5	72.04	0.43	0.92	1.03	1.24
	Dup	42	8.3	2	10.5	72.04	0.24	0.63	0.85	0.72
15q11.2	Del	4	3	1.3	1.48	8.39	0.18	0.37	0.44	0.53
	Dup	4	0.9	1.8	1.48	8.39	0.31	0.21	0.36	0.37
16p12.1	Del	7	-	-	0.57	8.41	0.51	0.56	0.7	0.66
	Dup	7	-	-	0.57	8.41	0.35	1.1	0.96	1.14
16p13.11	Del	6	-	-	2.28	17.405	0.56	0.97	1.76	1.17
	Dup	6	2	1.5	2.28	17.405	0.14	0.36	0.66	0.54
13q12.12	Dup	5	0.6	-	0.07	5.84	0.23	0.74	0.68	0.65
NRXN1	Del	1	-	-	1	1.56	0.33	0.65	0.94	0.82
TAR	Dup	15	2.4	-	3.41	21.47	0.19	0.62	0.65	0.52
ASD	PRS	-	-	2.4	-	-	0.03	0.06	-	-
SCZ		-	-	2.6	-	-	0.09	0.05	-	-
IQ		-	5	-	-	-	0.26	0.09	-	-
FluidIntel	Traits	-	-	-	-	-	0.47	0.15	-	-
Neuroticism		-	-	-	-	-	0.09	0.05	-	-
ASD	ENIGMA NPDs	-	-	-	-	-	0.13	0.13	-	-
ADHD		-	-	-	-	-	0.1	0.19	-	-
BD		-	-	-	-	-	0	0.23	-	-
MDD		-	-	-	-	-	0.03	0.14	-	-
OCD		-	-	-	-	-	0.01	0.16	-	-
SCZ		-	-	-	-	-	0.12	0.46	-	-

Supplement Table 1: Summary of effect sizes.

Legend: Summary effect sizes for IQ-loss/disease-risk, SubCortical volumes, Thickness, LogJacobian, and ICV. OR: Odds ratio; ICV: Intracranial volume; Del: deletion; Dup: Duplication; Thick: Thickness (RAD);

Comparison		Welch Two Sample t-test					Pearson		Concordance Correlation Coefficient (CCC)			
ES1	ES2	MeanES1	MeanES2	DF	t-stat	Pval	Corr	CorrPval	CCC_rho	95% CI [low	high]	Bias Factor
ICV	SubCortVol	0.537	0.652	20.438	-0.781	0.444	-0.127	0.652	-0.097	-0.47	0.306	0.761
ICV	Thick	0.537	0.793	23.688	-1.637	0.115	-0.003	0.991	-0.002	-0.393	0.389	0.759
ICV	LogJacs	0.537	0.789	21.432	-1.681	0.107	0.234	0.402	0.162	-0.21	0.493	0.693
SubCortVol	Thick	0.652	0.793	26.411	-1.328	0.195	0.696	0.004	0.599	0.224	0.82	0.861
SubCortVol	LogJacs	0.652	0.789	27.811	-1.423	0.166	0.78	0.001	0.679	0.34	0.862	0.871
Thick	LogJacs	0.789	0.793	27.247	-0.036	0.971	0.751	0.001	0.741	0.401	0.901	0.986

Supplement Table 2: Compare effect sizes across CNVs: ICV vs SubCortVol vs Thick vs LogJacs.

Legend: Welch t-test comparing the effect sizes across SubCortical measures and ICV reported in Supplement Table 2.

ROIs	1q21.1		16p11.2		22q11.2		15q11.2		16p12.1		16p13.11		13q12.12	NRXN1	TAR
	Del	Dup	Del	Dup	Del	Dup	Del	Dup	Del	Dup	Del	Dup	Dup	Del	Dup
Lthal	521	231	1074	578	1006	190	41	50	5	413	168	42	25	1	17
Lcaud	33	246	906	934	1094	346	286	498	47	5	508	525	92	64	176
Lput	83	59	1203	966	1136	163	571	530	3	8	83	65	3	6	16
Lpal	214	30	455	209	137	13	95	264	0	14	31	85	26	14	3
Lhippo	326	234	298	305	1193	16	79	67	31	22	88	122	46	21	60
Lamyg	135	0	316	284	483	228	173	302	17	61	1	88	21	49	25
Laccumb	97	1	303	181	662	131	10	169	2	0	28	28	5	3	263
Rthal	489	175	838	578	815	132	51	63	32	217	51	58	0	4	50
Rcaud	102	74	629	643	1433	143	128	634	0	23	890	420	236	147	7
Rput	243	70	823	735	1112	53	456	261	31	11	119	68	47	9	54
Rpal	199	1	458	377	257	23	345	485	4	122	134	11	0	50	0
Rhippo	608	71	504	669	1545	215	112	177	9	12	53	196	36	27	142
Ramyg	168	130	756	602	742	393	139	164	0	47	93	351	71	28	86
Raccumb	37	98	101	148	595	66	93	8	26	3	41	0	0	6	1

Supplement Table 3: FDR-corrected vertices count for Thickness.

Legend: Summary of nVertices surviving FDR correction (<0.05, across all vertices of 15 CNVs) for SubCortical Shape analysis Thickness measure.

ROIs	1q21.1		16p11.2		22q11.2		15q11.2		16p12.1		16p13.11		13q12.12	NRXN1	TAR
	Del	Dup	Del	Dup	Del	Dup	Del	Dup	Del	Dup	Del	Dup	Dup	Del	Dup
Lthal	591	93	1389	1703	1261	189	350	30	0	690	220	3	118	0	71
Lcaud	26	118	1495	1412	742	195	287	97	83	51	643	167	18	21	50
Lput	58	13	1101	1289	897	208	1009	709	0	17	76	49	10	32	0
Lpal	311	4	203	367	72	62	266	281	0	0	54	104	0	0	6
Lhippo	320	343	552	1117	1522	59	253	382	1	5	84	265	166	32	2
Lamyg	155	1	232	252	564	151	410	445	0	1	0	120	44	11	33
Laccumb	8	20	518	1	785	126	26	74	0	0	129	15	0	0	32
Rthal	367	3	1262	1647	970	32	137	4	86	185	28	5	0	24	38
Rcaud	15	321	1184	670	1339	14	240	346	12	12	553	70	26	32	34
Rput	327	13	743	1153	792	29	1022	405	0	66	249	36	0	1	52
Rpal	263	2	455	584	181	1	496	438	0	221	172	11	0	40	8
Rhippo	591	225	1220	1600	1920	63	297	170	0	0	43	220	100	48	40
Ramyg	76	55	755	607	812	364	293	185	0	48	115	418	157	32	99
Raccumb	0	24	339	48	850	62	144	30	7	0	53	4	3	0	8

Supplement Table 4: FDR-corrected vertices count for surface (Jacobian).

Legend: Summary of nVertices surviving FDR correction (<0.05,across all vertices of 15 CNVs) for SubCortical Shape analysis of surface (Log-Jacobian) measure.

Experiment		Pearson Correlation (linear scale)				Pearson Correlation (log10 scale)				Model comparison: linear vs non-linear (poly 3)				
Measure 1	Measure 2	Corr	95% CI [low, high]		Pval	Corr	95% CI [low, high]		Pval	RSS lm	RSS nlm	anova-DF	anova-F	anova-Pval
IQ loss	ES ICV	0.34	-0.33	0.78	0.31	0.63	0.05	0.89	0.04	1054.14	301.78	2	8.73	0.013
	ES ScortVol	0.69	0.16	0.91	0.02	0.54	-0.09	0.86	0.09	619.32	356.51	2	2.58	0.145
	ES Thick	0.66	0.1	0.9	0.03	0.61	0.02	0.89	0.04	674	605.42	2	0.4	0.687
	ES LogJacs	0.75	0.28	0.93	0.008	0.74	0.26	0.93	0.009	518.42	501.94	2	0.11	0.893
OR ASD/SCZ	ES ICV	0.09	-0.61	0.71	0.83	0.6	-0.11	0.9	0.09	806.94	580.57	2	0.97	0.439
	ES ScortVol	0.72	0.1	0.94	0.03	0.72	0.11	0.94	0.03	395.92	168.96	2	3.36	0.119
	ES Thick	0.74	0.14	0.94	0.02	0.68	0.04	0.93	0.04	371.01	126.17	2	4.85	0.067
	ES LogJacs	0.88	0.51	0.97	0.002	0.88	0.52	0.97	0.002	186.79	17.65	2	23.95	0.003
pLIsum	ES ICV	-0.06	-0.64	0.56	0.86	0.19	-0.46	0.71	0.58	138.67	119.94	2	0.55	0.602
	ES ScortVol	0.52	-0.12	0.85	0.1	0.19	-0.46	0.71	0.58	101.78	95.78	2	0.22	0.809
	ES Thick	0.78	0.34	0.94	0.005	0.42	-0.24	0.81	0.2	54.73	38.07	2	1.53	0.281
	ES LogJacs	0.71	0.19	0.92	0.01	0.45	-0.2	0.83	0.16	68.92	66.61	2	0.12	0.888
1/LOEUF sum	ES ICV	-0.12	-0.67	0.52	0.73	0.04	-0.57	0.62	0.91	6196.8	5554.91	2	0.4	0.682
	ES ScortVol	0.53	-0.11	0.86	0.1	0.5	-0.15	0.84	0.12	4540.31	4249.36	2	0.24	0.793
	ES Thick	0.8	0.38	0.95	0.003	0.72	0.22	0.92	0.01	2275.54	1564.28	2	1.59	0.269
	ES LogJacs	0.69	0.15	0.91	0.02	0.67	0.11	0.9	0.02	3298.54	3110.91	2	0.21	0.815
Normalized effect sizes (ES/nGenes OR ES/sum_LOEUF_inv)														
nGenes	ES ICV / nGenes	-0.42	-0.81	0.24	0.2	-0.69	-0.91	-0.14	0.02	2184.27	1086.78	2	3.53	0.087
	ES ScortVol / nGenes	-0.71	-0.92	-0.19	0.01	-0.88	-0.97	-0.6	0.0003	1323.06	158.6	2	25.7	0.0006
	ES Thick / nGenes	-0.87	-0.97	-0.58	0.0004	-0.97	-0.99	-0.89	7.00E-07	624	147.74	2	11.28	0.006
	ES LogJacs / nGenes	-0.85	-0.96	-0.52	0.0009	-0.94	-0.99	-0.8	1.00E-05	727.48	349.17	2	3.79	0.077
	IQloss / nGenes	-0.25	-0.74	0.41	0.47	-0.07	-0.64	0.55	0.84	2491.7	2205.11	2	0.45	0.652
	OR ASD-SCZ / nGenes	-0.2	-0.76	0.54	0.61	-0.28	-0.8	0.47	0.46	2368.8	1182.16	2	2.51	0.176
1/LOEUF sum	ES ICV / nGenes	-0.46	-0.83	0.19	0.16	-0.77	-0.94	-0.31	0.006	4961.12	2231.61	2	4.28	0.061
	ES ScortVol / nGenes	-0.73	-0.92	-0.23	0.01	-0.93	-0.98	-0.75	3.00E-05	2941.82	359.49	2	25.14	0.0006
	ES Thick / nGenes	-0.85	-0.96	-0.51	0.0009	-0.95	-0.99	-0.82	6.00E-06	1736.17	334.78	2	14.65	0.003
	ES LogJacs / nGenes	-0.85	-0.96	-0.5	0.001	-0.94	-0.99	-0.79	1.00E-05	1778.85	703.59	2	5.35	0.039
	IQ loss / nGenes	-0.3	-0.76	0.37	0.38	-0.13	-0.68	0.51	0.7	5729.98	5098.88	2	0.43	0.665
	OR ASD-SCZ / nGenes	-0.33	-0.81	0.43	0.39	-0.43	-0.85	0.33	0.25	5075.76	2501.49	2	2.57	0.171
1/LOEUF sum	ES ICV / LOEUF_inv	-0.42	-0.81	0.24	0.2	-0.72	-0.92	-0.21	0.01	5187.4	3010.24	2	2.53	0.149
	ES ScortVol / LOEUF_inv	-0.6	-0.88	0.01	0.05	-0.86	-0.96	-0.55	6.00E-04	4048.27	962.24	2	11.22	0.007
	ES Thick / LOEUF_inv	-0.74	-0.93	-0.26	0.009	-0.96	-0.99	-0.85	3.00E-06	2828.02	328.13	2	26.66	0.0005
	ES LogJacs / LOEUF_inv	-0.74	-0.93	-0.25	0.0097	-0.93	-0.98	-0.74	4.00E-05	2873.43	744.69	2	10	0.009
	IQ loss / LOEUF_inv	-0.28	-0.76	0.38	0.4	-0.12	-0.67	0.51	0.71	5777.86	4833.53	2	0.68	0.535
	OR ASD-SCZ / LOEUF_inv	-0.24	-0.78	0.51	0.54	-0.32	-0.81	0.44	0.4	5369.77	2917.52	2	2.1	0.218

Supplement Table 5: Linear vs log10 scale comparison of effect sizes.

Legend: Effect sizes vs cognition / disease-risk / nGenes / LOEUF_inv, comparing linear vs log-scale correlations (left and middle), and linear versus non-linear model fit using anova (right). Linear scale: both variables in linear scale; Log10 scale x-y: both variables in log10 scale;. ASD: autism spectrum disorder; SCZ: schizophrenia; PRS: Polygenic Risk Score; ES: Effect Size; LogJacs: logarithm of Jacobian determinant; Thick: Thickness; nGenes: number of genes within CNV. lm: linear model; nlm: non-linear model; RSS: residual sum of squares; DF: degrees of freedom; 95% CI: 95 % confidence intervals; LOEUF: loss-of-function observed/expected upper bound fraction; pLI: probability of being loss-of-function intolerant.

Correlation values are color-coded with blue-to-red gradient corresponding to negative and positive values. P-values are color coded so that smallest to largest p-values are on a red-yellow-green gradient.

Supplementary Figures
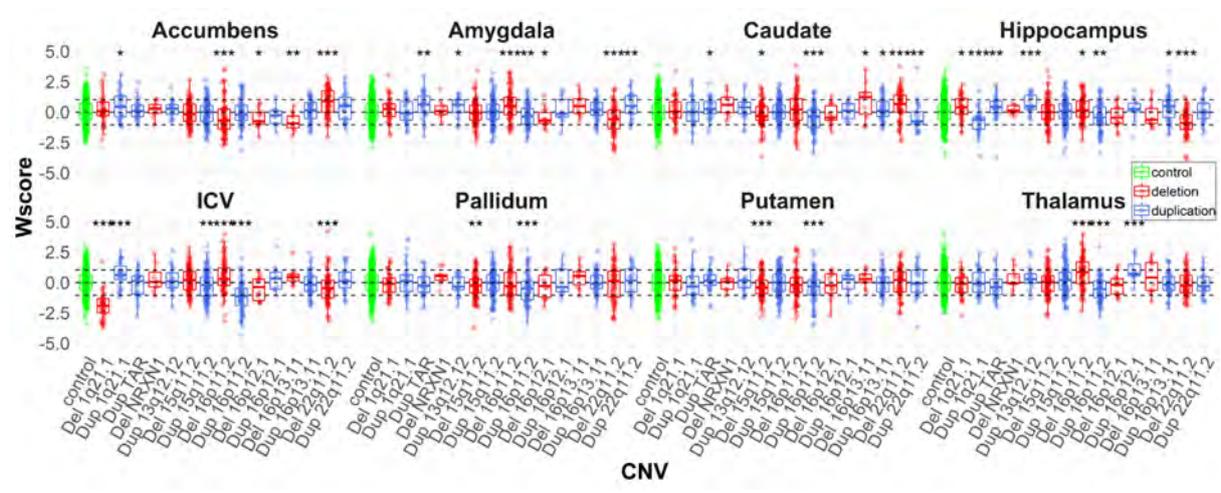
Supplement Figure 1: W-score obtained using Gaussian Processes modelling.

Legend: Boxplots showing the distribution of ICV and SubCortical W-scores (Z Score using mean and sigma of controls based on Gaussian Processes modeling) for CNVs. Gaussian Processess are modelled using age, sex, site, and ICV (except for ICV) as covariates.
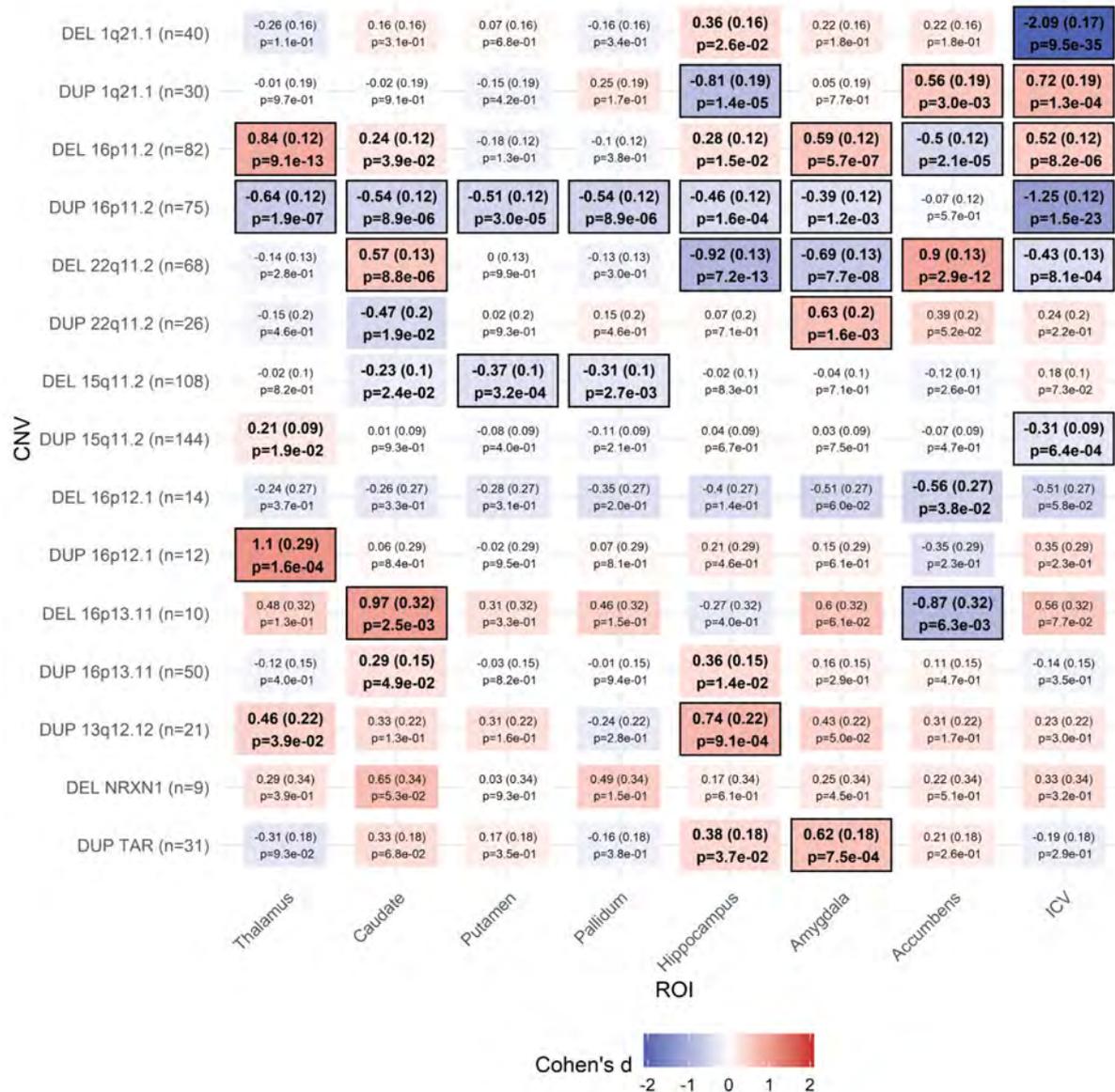
Supplement Figure 2: Detailed effect sizes, SE, and p-values for SubCortical volumes and ICV.

Legend: Tile plot showing the Cohen's d (SE) and p-values for SubCortical volumes and ICV. Case-control differences are calculated (lm in R) using W-scores obtained from Gaussian processes regression (GPR, with age, sex, site, and ICV as covariates). Significant Cohen's d with **nominal p-value <0.05 are in bold**, and **FDR p-value <0.05 are shown with rectangles**. Darker color represents higher magnitudes. lh and rh denote the left and right hemisphere respectively. DEL: deletions; DUP: duplications.
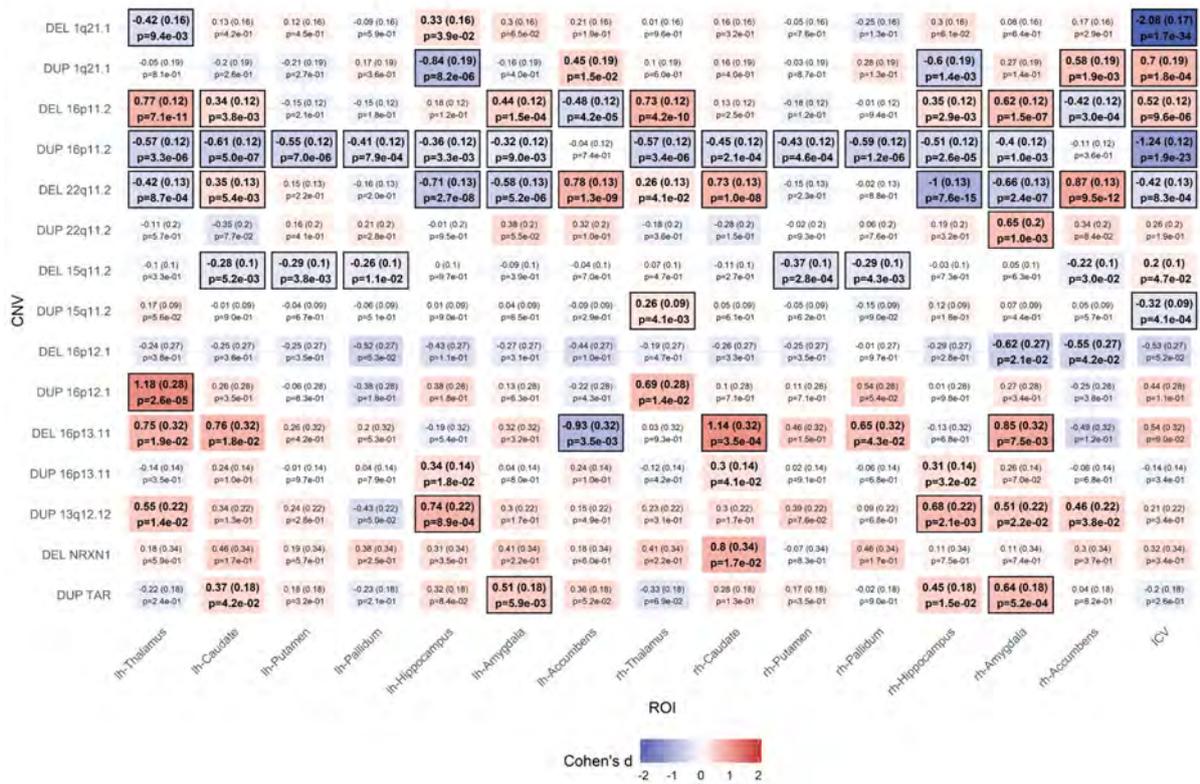
Supplement Figure 3: Sensitivity analysis- Bilateral effect sizes (Cohen's d) for SubCortical structures and ICV.

Cohen's d (SE) and p-values for subcortical structures for CNVs. Case-control differences are calculated (lm in R) using W-scores obtained from Gaussian processes regression (GPR, with age, sex, site, and ICV as covariates). Significant Cohen's d with **nominal p-value <0.05** are in bold, and **FDR p-value <0.05** are shown with rectangles. Darker color represents higher magnitudes. lh and rh denote the left and right hemisphere respectively. DEL: deletions; DUP: duplications.
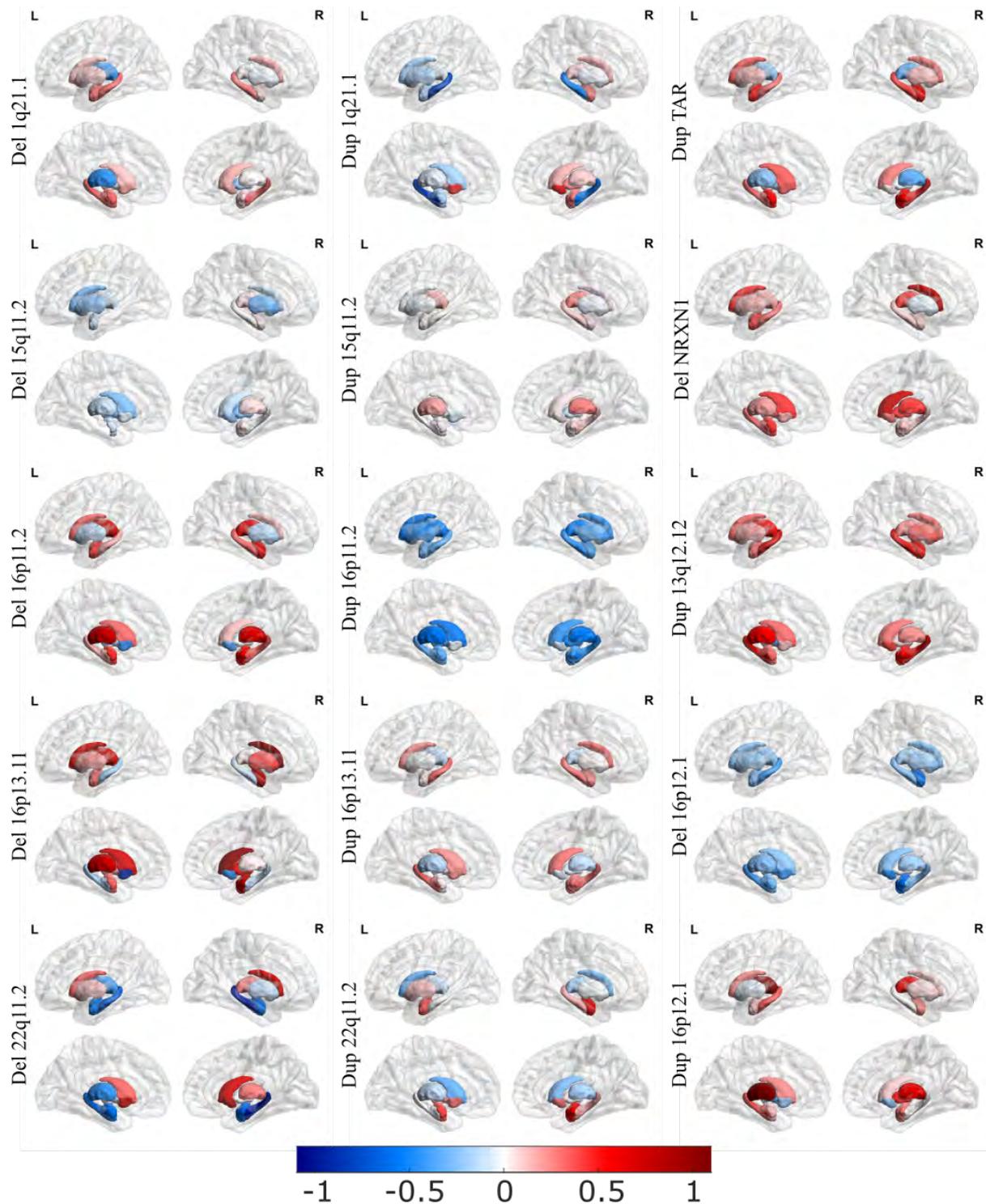
Supplement Figure 4: Sensitivity analysis- visualization of bilateral SubCortical volume effect sizes for CNVs.

Legend: Brain projections of Cohen's d values for subcortical structures (bilateral 14 ROIs). Case-control differences are calculated (lm in R) using W-scores obtained from Gaussian processes regression (GPR, with age, sex, site, and ICV as covariates). Effect sizes, standard

error, and p-values are reported in Supplement Figure 4. Darker color represents higher magnitudes. Plots generated using BrainNet Viewer. L and R denote the left and right hemisphere respectively. Del: deletions; Dup: duplications; ICV: Intracranial Volume.
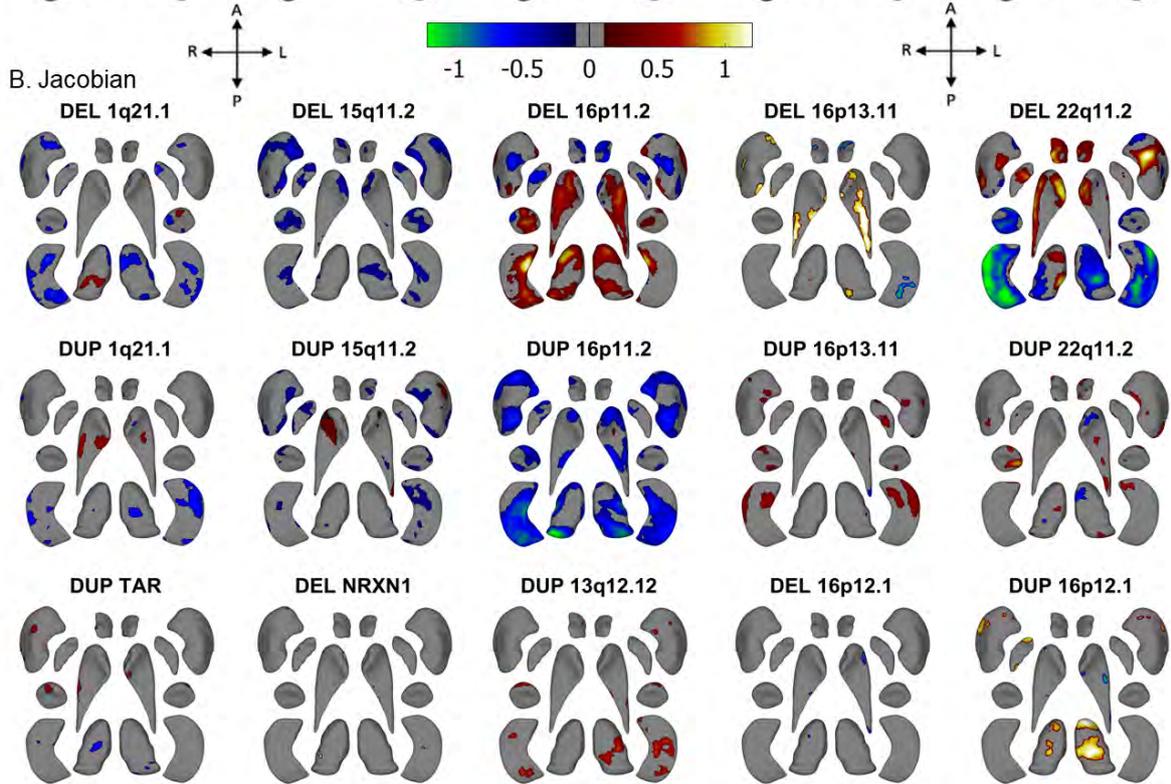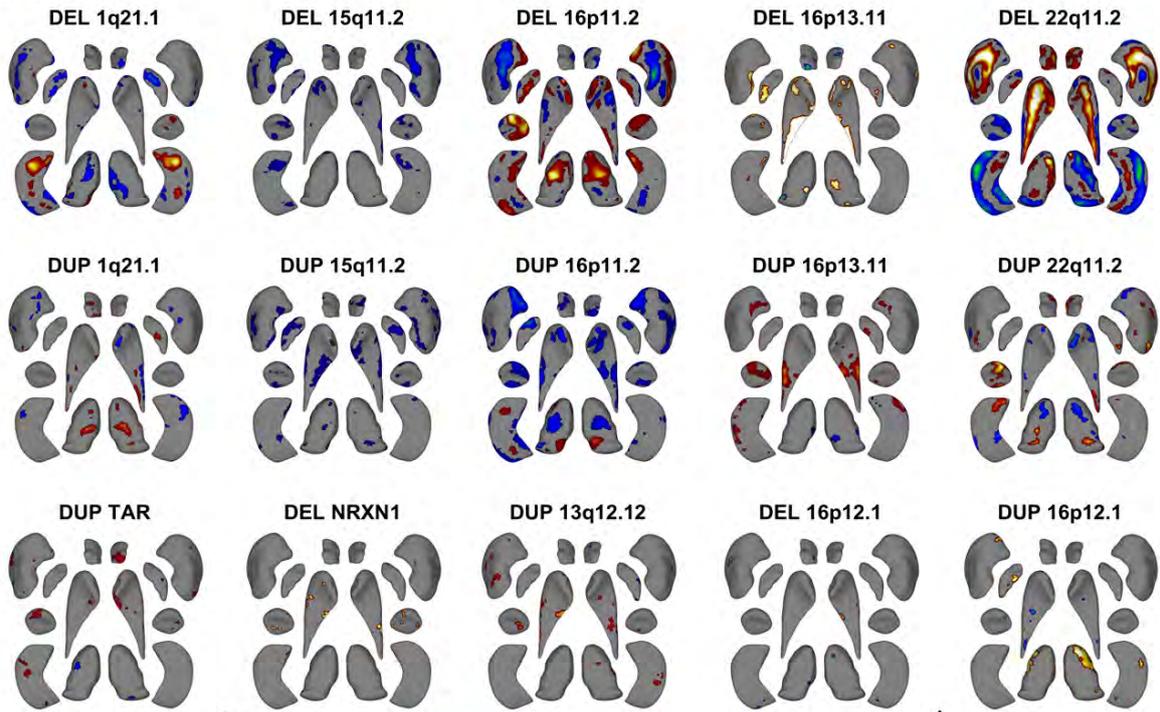
Supplement Figure 5: Cohen's d maps for SubCortical Shape analysis (Ventral view).

Legend: Effect sizes for Subcortical shape analysis of thickness (panel A); and Jacobian (panel B) measures. Cohen's d values are shown for 15 CNVs (Ventral view) after applying FDR correction (<0.05) across all vertices of 15 CNVs (within each measure/panel). Thickness represents local radial distance, and Jacobian represents local surface area dilation/contraction. Blue/green colors indicate negative coefficients, or regions of lower thickness measures in the CNV group compared with the controls. Red/yellow colors indicate positive coefficients, or regions of greater thickness values in the CNV group compared with the controls. Gray regions indicate areas of no significant difference after correction for multiple comparisons. Each vertex was adjusted for sex, site, age, and intra-cranial volume (ICV) using Gaussian Processes Regression before running case-control analysis. DEL: deletion; DUP: duplication.

Supplement Figure 6: Sensitivity analysis subcortical volume effect sizes.

Legend: Concordance plots between 22q11.2 deletion effect sizes (Cohen's d) for our main analysis (7ROIs) and different experiments. A) Concordance with literature: comparison with Cohen's d values (average of left and right hemisphere) from Ching et al. 2020; B) Impact of data pooling: comparison with effect sizes estimated by running analysis case-control linear model in separate cohorts; C) Impact of diagnosis: comparison with effect sizes estimated after removing 22q11.2 deletion carriers with a diagnosis (linear regression within 22q11.2 data from UCLA); D) Impact of bilateral ROIs: comparison between bilateral effect sizes estimated in our analysis (secondary) and those reported in Ching et a. 2020; E) Concordance between left and right hemisphere effect sizes in bilateral (secondary) analysis; F) Concordance between effect sizes estimated in our main analysis (average left and right hemisphere volumes before modelling or adjusting for covariates) versus left and right hemisphere averaged effect sizes (secondary analysis).

Perfect concordance line is shown in magenta. Linear fitted lines are shown with black dots.

15

A. Idiopathic conditions (ENIGMA Cohen's d) **B. PRS (top vs bottom decile)** **C. Traits**

	ASD (n=1522)	ADHD (n=1693)	BD (n=1710)	MDD (n=1728)	OCD (n=1493)	SCZ (n=2028)	PRS ASD (n=7161)	PRS MDD (n=7158)	PRS SCZ (n=7158)	PRS IQ (n=7158)	FluidIntel (n=8850)	Neuroticism (n=8313)
Thalamus	0	-0.03	-0.15*	-0.04	-0.08	-0.31*	-0.03	-0.01	-0.03	0	0.15*	-0.03
Caudate	-0.05	-0.11*	-0.09	-0.02	0.03	0.02	-0.01	0.03	0.02	0.03	0.03	0.05*
Putamen	-0.1*	-0.14*	0	0.01	0.02	0.08	-0.02	0.08*	0.02	-0.05	0.06*	0.05*
Pallidum	-0.08	0	-0.03	0	0.16*	0.21*	-0.03	0	-0.06	0.06*	0.11*	0.01
Hippocampus	-0.05	-0.11*	-0.23*	-0.14*	-0.14*	-0.46*	-0.05	-0.03	-0.05	0.03	0.15*	0.02
Amygdala	-0.08	-0.19*	-0.11*	-0.06	-0.11	-0.31*	-0.02	0.02	0.02	0.02	0.09*	0.03
Accumbens	-0.13*	-0.15*	-0.06	-0.02	-0.04	-0.25*	-0.05	0.02	-0.01	0.03	0.09*	0.01
ICV	0.13*	-0.1*	0	-0.03	0.01	-0.12	-0.02	-0.1*	-0.08*	0.26*	0.47*	-0.09*

Effect Size

1
0
-1
-2
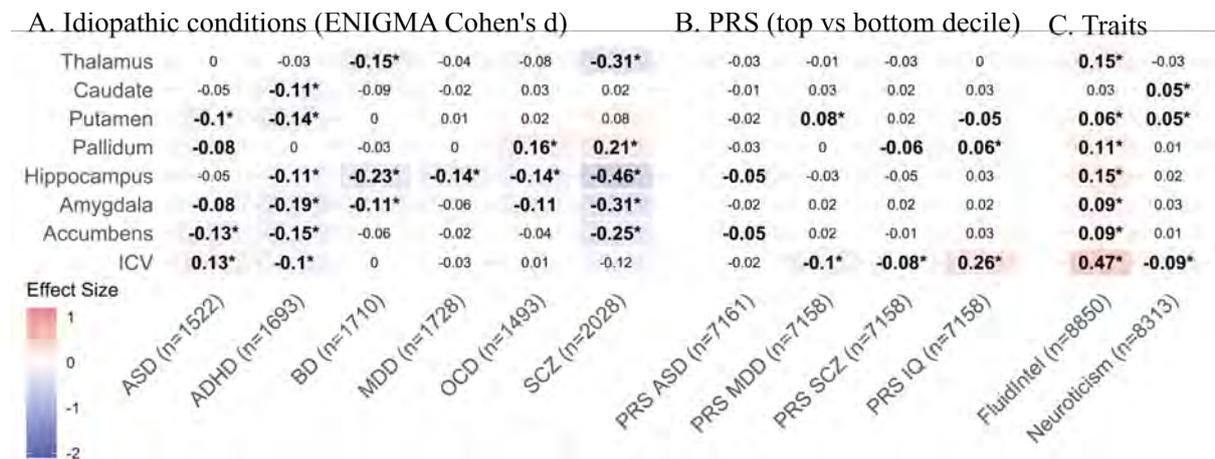
Supplement Figure 7: SubCortical volume alteration effect sizes for IPCs, PRS, and Traits.
Legend: A) Effect sizes for 6 Idiopathic conditions. Cohen's d values were previously published in ENIGMA-studies for: ASD [9], ADHD [10], BD [11], MDD [12], OCD [13], and SCZ [14].
B) Effect sizes for 4 PRS. Cohen's d values are estimated using top vs bottom decile case-control analysis (top and bottom deciles of PRS from UKBB data with a total sample size of 35000). C) Effect sizes for Traits (UKBB measures of Fluid-Intelligence (FI), and Neuroticism (NI)). For comparison purposes, Cohen's d values are estimated using top vs bottom decile case-control analysis (top and bottom deciles of traits from UKBB data with a total sample size of 40000). PRS-IQ uses the Fluid Intelligence measure from UKBB. Significant effect sizes with nominal **p-value <0.05** are in bold, and **FDR p-value <0.05** are shown with star (*), FDR correction is applied to each panel separately. Same color code (legend) is used across A-C as CNV's effect sizes (Figure 1B), for the ease of comparison. Darker color represents higher magnitudes. Sample sizes used for each analysis (for ICV) are reported in parentheses along with x-axis labels. DEL: deletions; DUP: duplications; PRS:

Polygenic Risk Score; IPC: Idiopathic condition; ASD: autism spectrum disorder; ADHD: attention deficit hyperactivity disorder; BD: bipolar disorder; MDD: major depressive disorder; OCD: obsessive-compulsive disorder; SCZ: schizophrenia; FluidIntel: Fluid Intelligence.

Supplement Figure 8: Concordance Correlation Coefficient between Subcortical volume, Thickness, and Surface effect sizes.

Legend: Concordance plots between effect sizes for Subcortical Volumes, Thickness, and surface (Log-Jacobian). Perfect concordance line is shown in magenta. Linear fitted lines are shown with black dots. CNV labels are shown for deletion and duplication. Del: deletion; Dup: duplication; ES: effect size; SubCortVol: subcortical volumes; Thick: Thickness; LogJacs: Log Jacobian; CI: confidence interval; CCC: concordance correlation coefficient (using DescTools package in R).

A. Normalized effect sizes (ES/nGenes) versus nGenes (log10)

B. Normalized effect sizes (ES/nGenes) versus pLI-sum (log10)
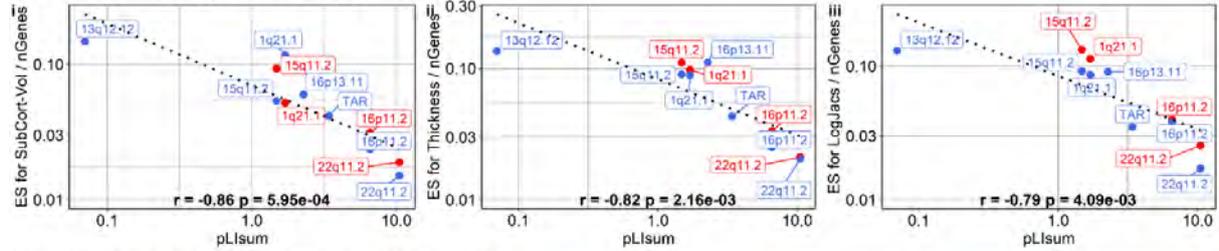
C. Normalized effect sizes (ES/nGenes) versus LOEUF_inv (log10)
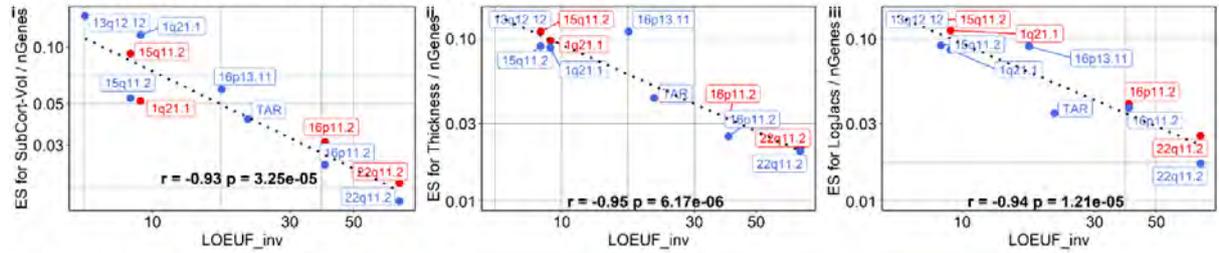
D. Normalized effect sizes (ES/nGenes) versus LOEUF_inv
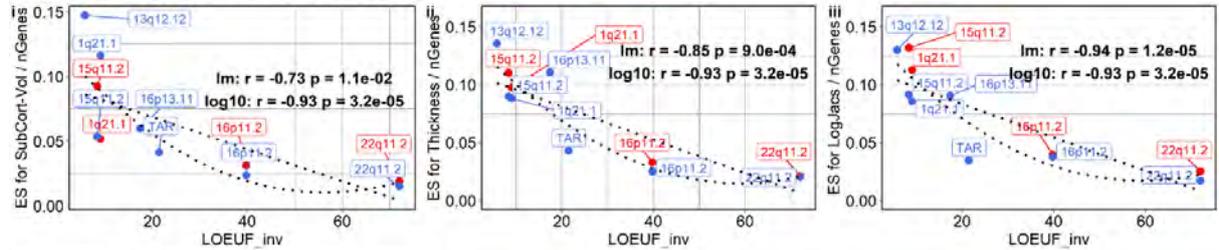
E. Normalized effect sizes (ES/nGenes) versus nGenes
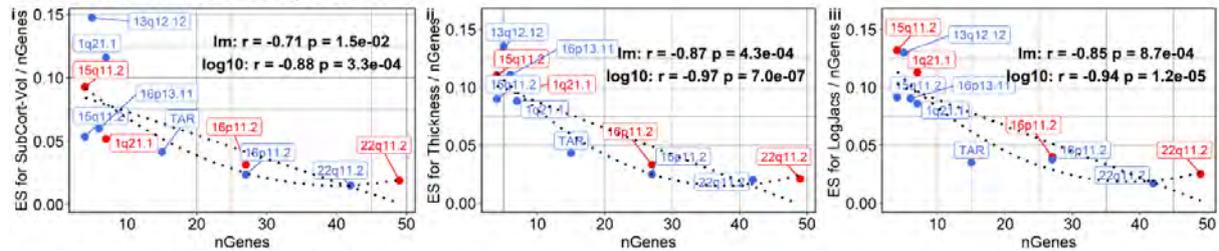
F. Normalized effect sizes (ES/LOEUFinv) versus LOEUF_inv

18
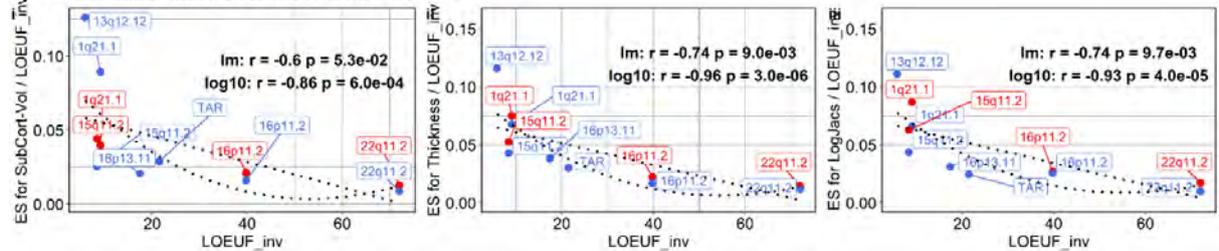
Supplement Figure 9: Sensitivity analysis for normalized effect sizes and Polygenicity

Legend: Normalized effect sizes (Effect sizes by nGenes) for Subcortical volumes, thickness, and LogJacs versus nGenes (Panel A), pLIsum (Panel B), sum of 1/LOEUF (Panel C). Panel A-C are in log10 scale, also the Pearson Correlation is computed after transforming values in log10 scale.

Comparing linear versus non-linear fit for normalized effect sizes (Effect sizes by nGenes or LOEUF_inv) for Subcortical volumes, thickness, and LogJacs versus sum of 1/LOEUF (Panel D and F) and nGenes (Panel E). Panel D-F are in linear scale. The Pearson correlation in linear scale (lm) and log10 scale (log10) are reported. Linear and nonlinear (LOESS function) curve fits are shown.

Deletions, and duplications are colored in red, and blue respectively. Del: deletion; Dup: duplication; ASD: autism spectrum disorder; SCZ: schizophrenia; PRS: Polygenic Risk Score; ES: Effect Size; LogJacs: logarithm of Jacobian determinant; LOEUF: loss-of-function observed/expected upper bound fraction; pLI: probability of being loss-of-function intolerant.

19
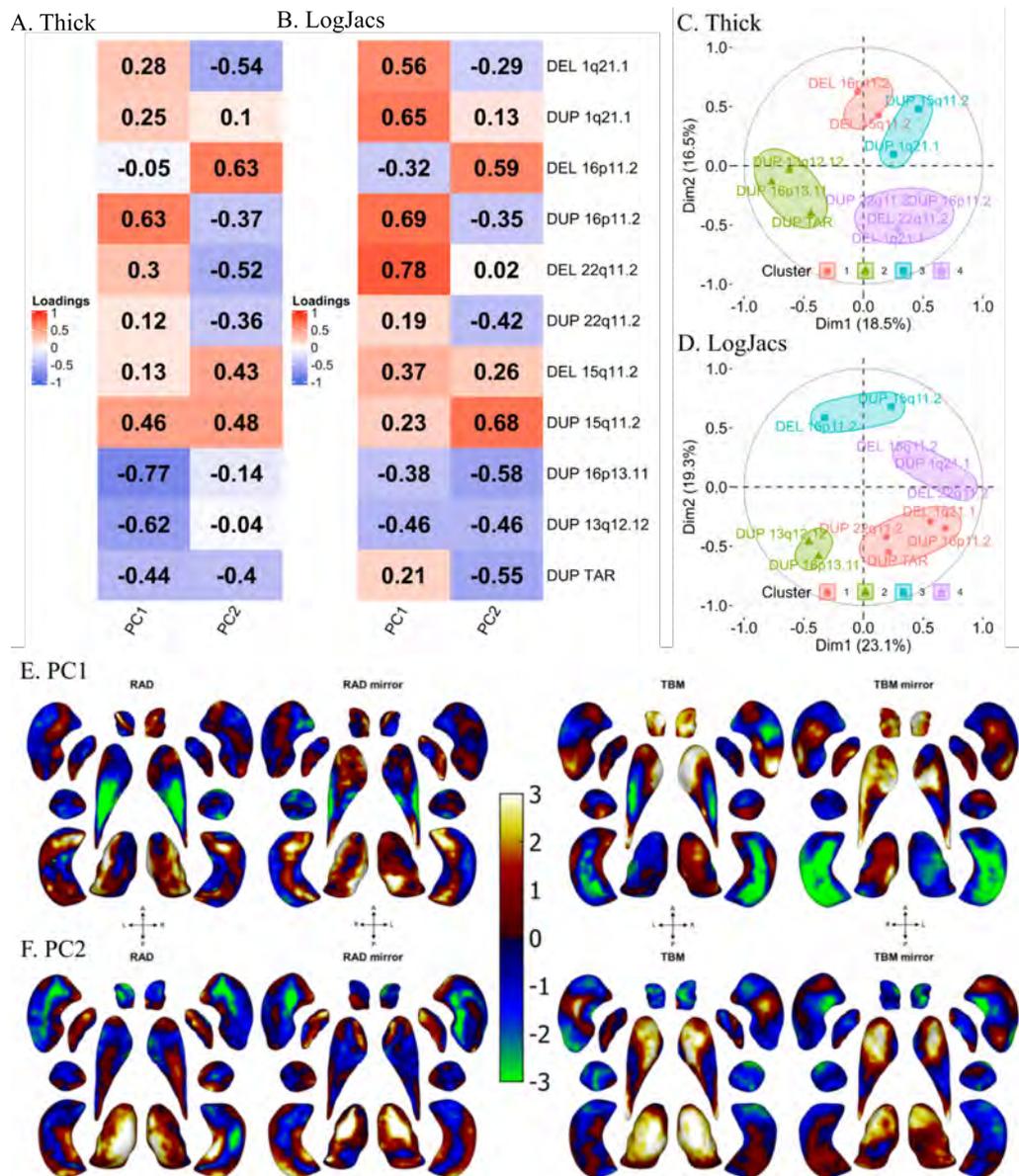
Supplement Figure 10: Principal Component Analysis across vertex-wise Cohen's d maps of CNVs.

Legend: Principal Component Analysis across vertex-wise Cohen's d maps of 11 CNVs for thickness (RAD) and Jacobian (TBM). Thickness represents local radial distance, and Jacobian represents local surface area dilation/contraction. Principal Component Analysis was run with CNVs (n>20) as variables and vertices (stacked across all SubCortical structures) as observations (Z-scored). Blue/green colors indicate negative Z-scores, or regions of lower Jacobian measure in the case group compared with the controls. Red/yellow

colors indicate positive coefficients, or regions of greater Jacobian measure in the case group

compared with the controls. Dorsal (TBM) and ventral (TBM mirror) views. Each vertex was

adjusted for sex, site, age, and intra-cranial volume (ICV) using Gaussian Processes

Regression.

A. CNV loadings for PC1 and PC2

i. Vol

PC1	PC2	
0.57	0.8	DEL 1q21.1
-0.76	0.19	DUP 1q21.1
0.57	0.7	DUP TAR
0.88	0.04	DUP 13q12.12
0.73	-0.09	DEL 15q11.2
0.55	-0.69	DUP 15q11.2
0.6	-0.7	DEL 16p11.2
-0.1	0.84	DUP 16p11.2
0.64	0.54	DUP 16p13.11
-0.66	0.26	DEL 22q11.2
0.06	0.49	DUP 22q11.2

Loadings: 1, 0.5, 0, -0.5, -1

ii. Thick

PC1	PC2	
0.28	-0.54	DEL 1q21.1
0.25	0.1	DUP 1q21.1
-0.05	0.63	DEL 16p11.2
0.63	-0.37	DUP 16p11.2
0.3	-0.52	DEL 22q11.2
0.12	-0.36	DUP 22q11.2
0.13	0.43	DEL 15q11.2
0.46	0.48	DUP 15q11.2
-0.77	-0.14	DUP 16p13.11
-0.62	-0.04	DUP 13q12.12
-0.44	-0.4	DUP TAR

Loadings: 1, 0.5, 0, -0.5, -1

iii. LogJacs

PC1	PC2	
0.56	-0.29	DEL 1q21.1
0.65	0.13	DUP 1q21.1
-0.32	0.59	DEL 16p11.2
0.69	-0.35	DUP 16p11.2
0.78	0.02	DEL 22q11.2
0.19	-0.42	DUP 22q11.2
0.37	0.26	DEL 15q11.2
0.23	0.68	DUP 15q11.2
-0.38	-0.58	DUP 16p13.11
-0.46	-0.46	DUP 13q12.12
0.21	-0.55	DUP TAR

Loadings: 1, 0.5, 0, -0.5, -1

B. Corrcircle in PC1-PC2 space with CNV clusters
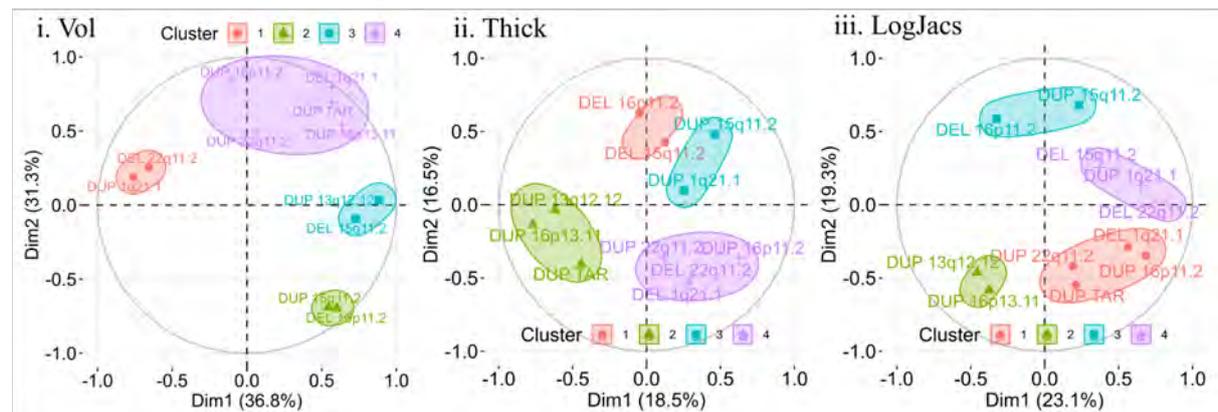
Supplement Figure 11: Comparison of CNV loadings and PC1-PC2 clusters across SubCortical Volume, Thickness, and Jacobian.

Legend: Panel A: CNV loadings for PCA analysis of Cohen's d maps of 11 CNVs for (i) volume (ii) thickness (RAD), and (iii) Jacobian (TBM). Thickness represents local radial distance, and Jacobian represents local surface area dilation/contraction. Principal Component Analysis was run with CNVs (n>20) as variables and volumes/vertices (stacked across all SubCortical structures) as observations (Z-scored).

Panel B: Correlation circle plot of CNV loadings in PC1-PC2 space with k-means clustering

used to detect groupings of CNVs. The percentage of variance explained by each principal

component dimension is also included in axis labels.

A. ROI loadings across PC's
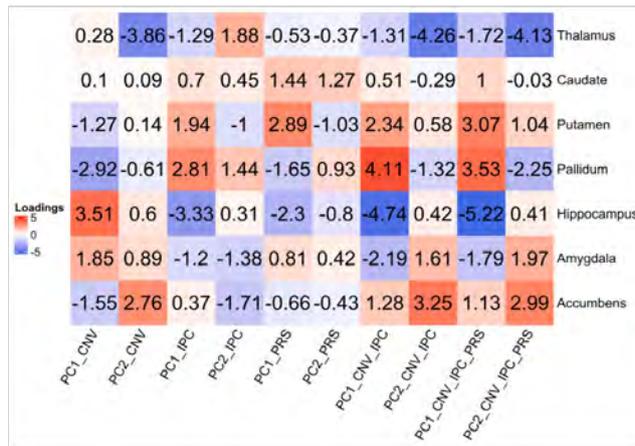
B. Correlation between ROI loadings
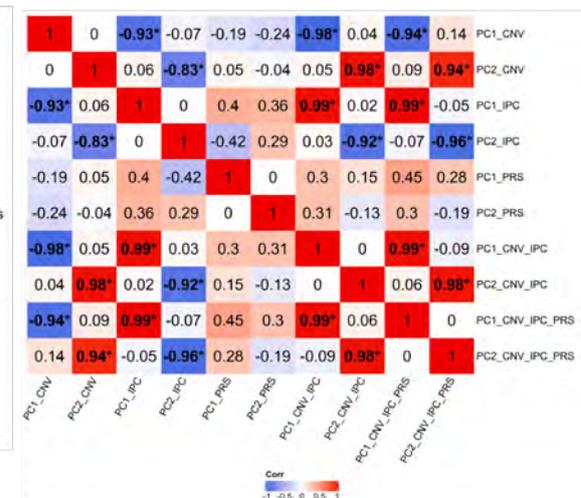
C. Brain projections for ROI loadings
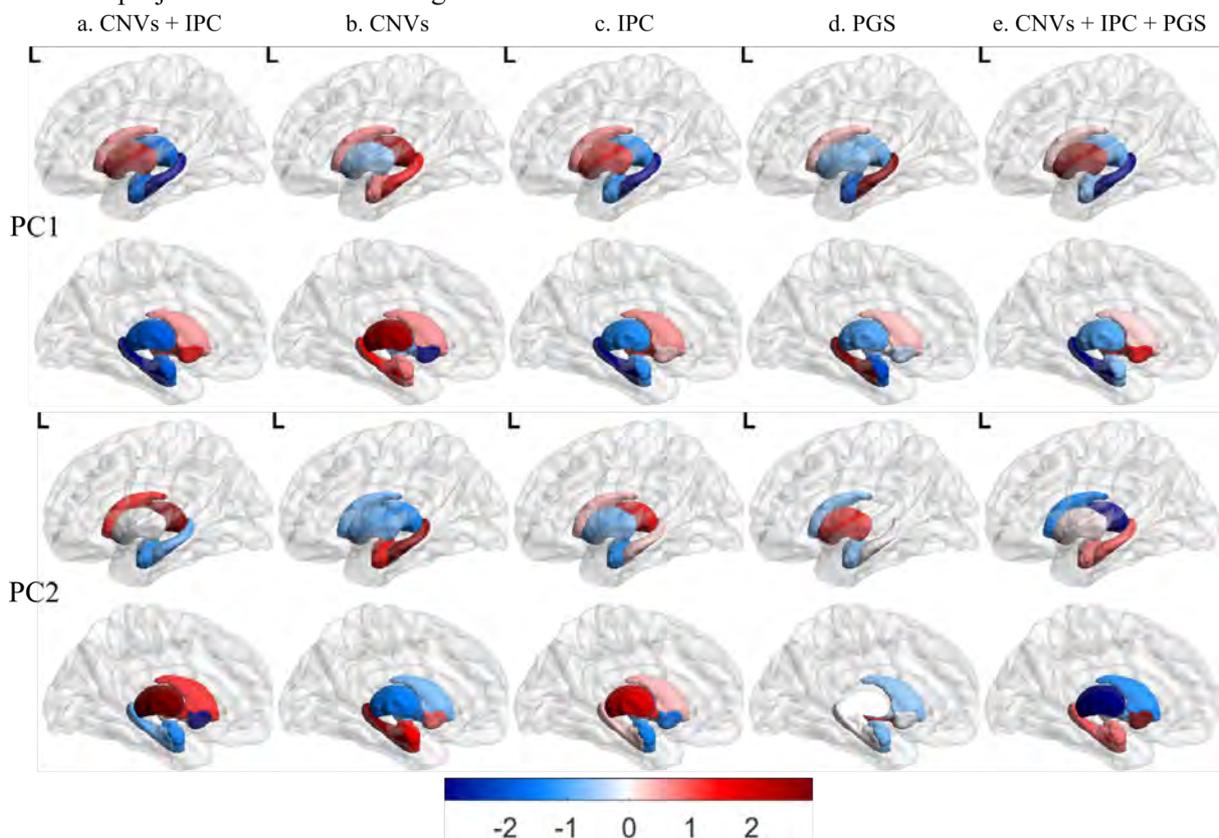
Supplement Figure 12: Sensitivity analysis- Principal component analysis across different sets of Cohen's D.

Compare PCA ROI loadings across 1) CNV only PCA; 2) IPC PCA; 3) CNV + IPC PCA; 4) PRS PCA; 5) CNV+IPC+PRS (Combined) PCA. Panel A: ROI loadings across different PCA are shown. Panel B: Pearson Correlation between PCA ROI loadings are reported. Panel

C: Brain projections for PCA ROI loadings. PC1 (top half) and PC2 (bottom half) shown on the same color scale for ease of visualization. The sign of PC loadings varies from one set to another, as such some PC loading will be flipped. PC_Comb refers to PC_CNV_IPC_PRS combined. PCA analyses were run using 4 PRS: ASD, MDD, SCZ, and IQ, 11 CNVs (n>20), and 6 Idiopathic conditions: ASD, ADHD, BD, MDD, OCD, and SCZ.

References

1. Lin, A. *et al.* Mapping 22q11.2 Gene Dosage Effects on Brain Morphometry. *J. Neurosci.* **37**, 6183–6199 (2017).

2. Maillard, A. M. *et al.* The 16p11.2 locus modulates brain structures common to autism, schizophrenia and obesity. *Mol. Psychiatry* **20**, 140–147 (2015).

3. Qureshi, A. Y. *et al.* Opposing Brain Differences in 16p11.2 Deletion and Duplication Carriers. *J. Neurosci.* **34**, 11199–11211 (2014).

4. Martin-Brevet, S. *et al.* Quantifying the Effects of 16p11.2 Copy Number Variants on Brain Structure: A Multisite Genetic-First Study. *Biol. Psychiatry* **84**, 253–264 (2018).

5. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).

6. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Med.* **12**, e1001779 (2015).

7. Huguet, G. *et al.* Measuring and Estimating the Effect Sizes of Copy Number Variants on General Intelligence in Community-Based Samples. *JAMA Psychiatry* **75**, 447–457 (2018).

8. Douard, E. *et al.* Effect Sizes of Deletions and Duplications on Autism Risk Across the Genome. *Am. J. Psychiatry* **178**, 87–98 (2020).

9. van Rooij, D. *et al.* Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. *Am. J. Psychiatry* **175**, 359–369 (2018).

10. Hoogman, M. *et al.* Subcortical brain volume differences in participants with attention

deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *Lancet Psychiatry* **4**, 310–319 (2017).

11. Hibar, D. P. *et al.* Subcortical volumetric abnormalities in bipolar disorder. *Mol Psychiatry* **21**, 1710–1716 (2016).

12. Schmaal, L. *et al.* Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. *Mol Psychiatry* **21**, 806–812 (2016).

13. Boedhoe, P. S. W. *et al.* Distinct Subcortical Volume Alterations in Pediatric and Adult OCD: A Worldwide Meta- and Mega-Analysis. *Am J Psychiatry* **174**, 60–69 (2017).

14. van Erp, T. G. M. *et al.* Cortical Brain Abnormalities in 4474 Individuals With Schizophrenia and 5098 Control Subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. *Biol. Psychiatry* **84**, 644–654 (2018).