



Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism

Rafael Romero-Garcia¹ · Varun Warriar² · Edward T. Bullmore^{1,3,4} · Simon Baron-Cohen^{2,5} · Richard A. I. Bethlehem²

Received: 24 October 2017 / Revised: 27 November 2017 / Accepted: 15 December 2017 / Published online: 26 February 2018
© The Author(s) 2018. This article is published with open access

Abstract

Differences in cortical morphology—in particular, cortical volume, thickness and surface area—have been reported in individuals with autism. However, it is unclear what aspects of genetic and transcriptomic variation are associated with these differences. Here we investigate the genetic correlates of global cortical thickness differences (Δ CT) in children with autism. We used Partial Least Squares Regression (PLSR) on structural MRI data from 548 children (166 with autism, 295 neurotypical children and 87 children with ADHD) and cortical gene expression data from the Allen Institute for Brain Science to identify genetic correlates of Δ CT in autism. We identify that these genes are enriched for synaptic transmission pathways and explain significant variation in Δ CT. These genes are also significantly enriched for genes dysregulated in the autism post-mortem cortex (Odd Ratio (OR) = 1.11, $P_{\text{corrected}} = 10^{-14}$), driven entirely by downregulated genes (OR = 1.87, $P_{\text{corrected}} = 10^{-15}$). We validated the enrichment for downregulated genes in two independent data sets: Validation 1 (OR = 1.44, $P_{\text{corrected}} = 0.004$) and Validation 2 (OR = 1.30; $P_{\text{corrected}} = 0.001$). We conclude that transcriptionally downregulated genes implicated in autism are robustly associated with global changes in cortical thickness variability in children with autism.

Rafael Romero-Garcia and Varun Warriar contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41380-018-0023-7>) contains supplementary material, which is available to authorized users.

✉ Richard A. I. Bethlehem
rb643@medschl.cam.ac.uk

¹ Department of Psychiatry, University of Cambridge, Cambridge CB2 0SZ, UK

² Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge CB2 0SZ, UK

³ Cambridgeshire and Peterborough National Health Service Foundation Trust, Cambridge CB21 5EF, UK

⁴ ImmunoPsychiatry, GlaxoSmithKline Research and Development, Stevenage SG1 2NY, UK

⁵ CLASS Clinic, Cambridgeshire and Peterborough NHS Foundation Trust (CPFT), Cambridgeshire, UK

Introduction

Autism Spectrum Conditions (henceforth ‘autism’) are characterized by difficulties in social communication alongside unusually narrow interests and restrictive, repetitive behaviours, a resistance to unexpected change and sensory hypersensitivity [1]. In addition to behavioural and clinical differences, differences in cortical morphology between individuals with autism compared to typical controls have been reported [2–5]. While heterogeneous, recent studies have reported increased cortical volumes in the first years of life with autism compared to controls, with accelerated decline or arrest in growth in adolescents [3, 4]. Changes in cortical volume may be attributed to changes in cortical thickness (CT), changes in surface area or both [3]. In support of this, studies have separately identified differences in both surface area [6] and CT [7] in children with autism. For example, Smith et al. [7] show that the developmentally accelerated gain in grey matter volume in autism is largely driven by the lack of typical age-related CT decrease (<http://ajp.psychiatryonline.org/doi/abs/10.1176/aip.ajp.2017.17010100?code=ajp-site>). Furthermore, earlier studies identified differential trajectories in CT

Table 1 Descriptive statistics for all four datasets

	Discovery		Validation 1		Validation 2		ADHD	
	Autism	Controls ^a	Autism	Controls	Autism	Controls	ADHD	Controls ^a
<i>n</i>	62 (0 F)	87 (0 F)	48 (8 F)	54 (27 F)	56 (15 F)	154 (56 F)	69 (0 F)	87 (0 F)
Age	10.07 (±1.11)	10.04 (±1.13)	10.98 (±1.53)	10.43 (±1.71)	10.32 (±1.51)	10.34 (±1.20)	9.99 (±1.17)	10.04 (±1.13)
FIQ	108.86 (±16.94)	110.98 (±10.39)	118.68 (±15.01)	122.04 (±13.27)	103.42 (±15.99)	114.4 (±10.55)	107.95 (±14.18)	110.98 (±10.39)

The Discovery cohort was obtained from ABIDE-I. The validation cohorts were obtained from the ABIDE-II (Validation 1: Georgetown University, Validation 2: Kennedy Krieger Institute). The *n*-row denotes the number of subjects with the number of females (F) provided in parenthesis, FIQ denotes the full-scale IQ, with standard deviations in parenthesis below. Further details on the Discovery and ADHD datasets are described elsewhere [34]

^aIndicates that the same controls were used for both the Autism Discovery and the ADHD datasets.

development in autism [8] as well as CT differences in autism in specific brain regions [9, 10]. In addition, Hardan et al. [11] found areas of increased CT in children with autism, predominantly in temporal and parietal lobules. In contrast however, Hadjikhani et al. [12] report a pattern of cortical thinning in adults with autism, mainly within the mirror neuron system. Both these studies point towards significant heterogeneity within findings that is at least partially related to differences in age. Despite significant heterogeneity in cortical morphology across autism imaging studies [13], recent studies have also indicated alterations in areas associated with higher cognition (e.g. language, social perception and self-referential processing) [7, 14]. This has been supported by observed differences in cortical minicolumns in association areas in individuals with autism [15].

It is unclear what contributes to these differences in cortical morphology in individuals with autism. Genetic factors play a major role in the development of brain networks and volumes in typically developing individuals [16–18]. For instance, twin heritability of CT measures suggest modest to high heritability for most regions of the brain [19]. In parallel, the contribution of genetic factors for autism has been estimated between 50–90% [20–22]. Different classes of genetic variation have been associated with risk for autism. Several recent studies have identified a significant contribution of rare, de novo putative loss-of-function mutations for autism [23–27]. In addition, common genetic variants, cumulatively, account for approximately half of the total variance in risk for autism [20]. Studies have also identified genes dysregulated in the autism post-mortem cortex [28–30], enriched in processes such as synaptic transmission and astrocyte and microglial genes. These dysregulated genes may either represent causal mechanisms for risk or compensatory mechanisms as a result of upstream biological and cellular changes. Genes

dysregulated in the autism post-mortem cortex are also enriched in specific gene co-expression modules identified using both adult [28–30] and fetal [31] cortical post-mortem samples.

Despite considerable progress in understanding neuroanatomical and genetic risk for autism, several questions remain. Mechanistically, it is likely that genetic risk variants alter neuroanatomical structural and functional properties, contributing to behavioural and clinical phenotypes. Given the heterogeneity in autism imaging findings [13], it is pertinent to ask how genetic risk for autism is associated with variability in cortical morphology observed in individuals with autism. Thus, the goal of the present study was to identify molecular correlates of disease-related neuroanatomy irrespective of regional specific neuroanatomical differences that may not replicate well across studies [13]. Here, focusing on CT, we ask 3 specific questions: (Q1) Which genes and biological pathways are associated with in CT variability (Δ CT) in children with autism? (Q2) What is the spatial expression profile of genes associated with Δ CT? and (Q3) Are these genes enriched for three different classes of risk factors associated with autism: rare, de novo variants, common genetic variants and/or dysregulated genes in the post-mortem cortex? We address these questions by combining analysis of Δ CT in autism, as measured with MRI, with gene expression post-mortem data provided by the Allen Institute for Brain Science (AIBS; [32, 33]).

Methods

Overview

We first assessed differences in CT (Δ CT) across 308 cortical regions in individuals with autism by extracting CT estimates for 62 children with autism (cases) and 87

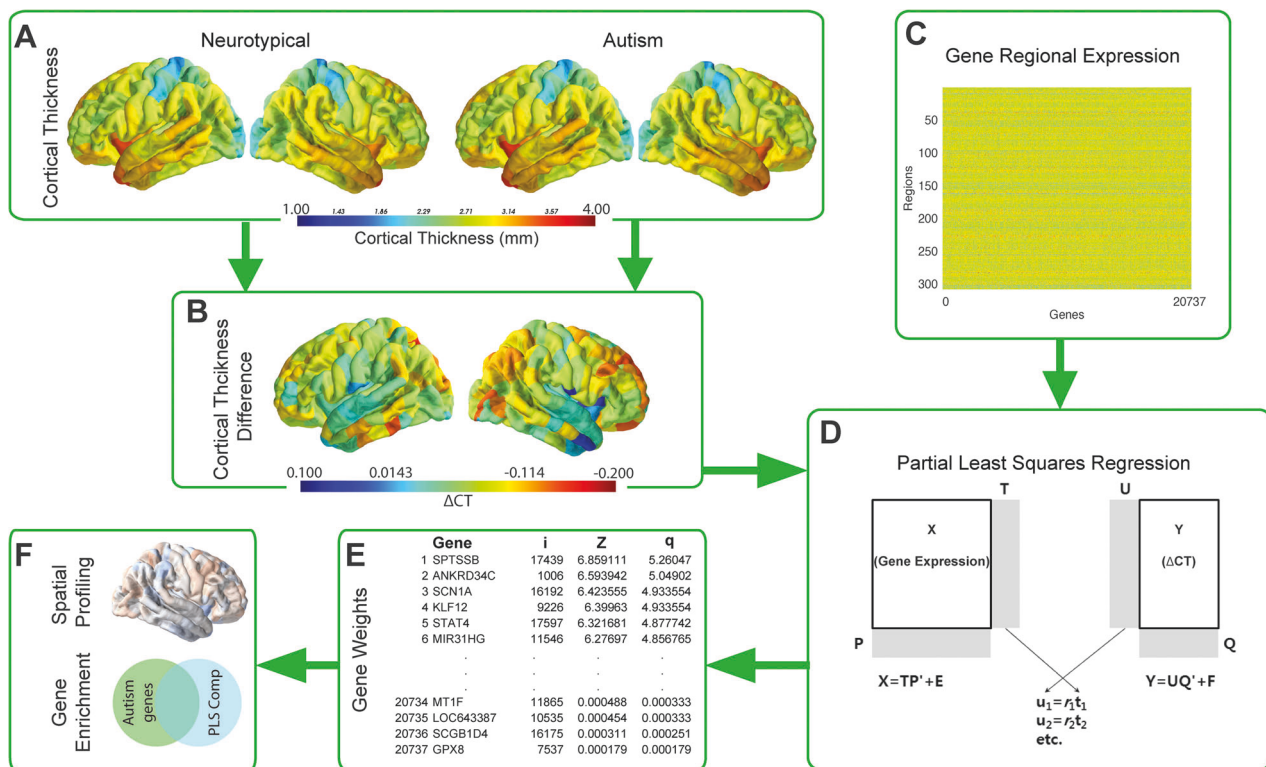


Fig. 1 Schematic overview of the methodology used to identify gene contribution. Mean cortical thickness was extracted for both the autism and the neurotypical groups across 308 cortical nodes (a). A difference score in cortical thickness (Δ CT; autism—neurotypical) was calculated between these two groups (b). In parallel the median AIBS gene expression profiles for 20,737 genes were calculated across the same 308 cortical nodes used in the imaging analysis (c). Both these streams were included in a bootstrapped PLSR analysis that used the gene expression profiles as predictors and the Δ CT as response variable (d). The PLSR assigns weights to each gene in terms of its contribution to

the overall model in each component. Bootstrapped standard errors were derived and the gene weights were Z-transformed and corrected for multiple comparison using a FDR inverse quantile transform correction to account for winners curse (e; i = gene index number, z = z-score for that gene's association and q = FDR corrected z-score). Genes that were significant after FDR correction (z -score > 1.96) were analysed in terms of their spatial expression as well as tested for enrichment against three classes of risk for autism: dysregulated autism genes in the postmortem cortex, genes harbouring rare de novo variants and common genetic variants in autism (f)

matched typically developing individuals (controls) from the ABIDE-I (Table 1; Discovery dataset). Using median gene expression of 20,737 genes from six post-mortem cortical brain samples [32], we conduct a partial least squares regression (PLSR), a data reduction and regression technique, to identify significant genes and enriched pathways that contribute to Δ CT (Q1). We next quantified the expression of the same significant genes in terms of their spatial profile by comparing them across the different brain regions and Von Economo classes [45], which provides a way of assessing the hypothesis that there would be a global differentiation between higher order cognitive processing and more primary sensory processing (Q2). We tested any significant genes for enrichment for classes of genetic and transcriptomic variation associated with autism (Q3): (1) Genes dysregulated in the autism post-mortem cortex; (2) Adult cortical gene co-expression modules associated with dysregulated genes in the autism post-mortem cortex; (3) Fetal cortical gene co-expression modules associated with dysregulated genes in the autism post-mortem cortex; (4)

Genes enriched for rare, de novo loss of function mutations in autism; and (5) Common genetic variants associated with autism. To assess the replicability of the findings, we validate the results using two independent datasets from ABIDE-II (Table 1). In parallel, we also used a second list of genes dysregulated in autism identified using a partially overlapping cortical gene expression data set of autism and control post-mortem brain samples to validate the enrichment analysis across all datasets. To assess specificity of our results, we furthermore sought to answer these questions in a matched MRI dataset of children with ADHD, another childhood psychiatric condition. A schematic overview of the study protocol is provided in Fig. 1.

Discovery dataset

Neuroimaging, gene expression and PLS regression

Discovery imaging data used in this study are described in detail in the supplementary materials and elsewhere [34]. In

short, structural T1-weighted MPRAGE images were obtained from the ABIDE I database (http://fcon_1000.projects.nitrc.org/indi/abide/), selecting participants in the age range from 9 to 11, all males. All subjects were matched on age, and IQ between groups (see Table 1; Discovery Data) (see ref. [34] and Supplementary Materials for details on matching and scanner site). CT estimates were extracted using freesurfer and visually inspected for quality of segmentation by two independent researchers. Only when there was consensus between researchers were images included. Next, images were parcellated into 308 cortical regions and mean CT for these regions was extracted. In addition, scanner site was regressed out from CT estimates and the residuals were added to the group mean to allow for easier interpretation. The final sample consisted of 62 children with autism (cases) and 87 neurotypical individuals (controls).

We used the transcriptomic dataset of the adult human brain created by the Allen Institute for Brain Science (<http://human.brain-map.org>) [32, 33]. The anatomical structure associated to each tissue sample was determined using the MRI data provided by the AIBS for each donor. The high-resolution parcellation with 308 cortical regions, employed in the neuroimaging dataset, was warped from the anatomical space of the average subject provided by FreeSurfer (fsaverage) into the surface reconstruction of each AIBS donor brain. After pre-processing regional gene expression values were represented as a $308 \times 20,737$ matrix that contained the whole-genome expression data for the 308 MRI regions of interest [35]. Code used to determine regional gene expression levels is available at (http://github.com/RafaelRomeroGarcia/geneExpression_Repository) and data used can be downloaded from Cambridge Data Repository [36]. More details on tissue sample handling, processing, batch correction and consistency of gene expression data across donors are provided in the Supplementary Materials. Cortical surface representations were plotted using BrainsForPublication v0.2.1 (<https://doi.org/10.5281/zenodo.1069156>).

We used PLSR to identify which genes were significantly associated with Δ CT. After obtaining PLS weights for each gene, these were z-transformed (based on standard errors obtained from bootstrapping) and FDR-adjusted using a FDR inverse quantile transformation correction to account for winners curse bias [37]. Only genes that passed FDR correction of $p < 0.05$ were included in enrichment analysis. We used significant genes with both negative and positive weights in our analysis. As our dependent variable, Δ CT, had both positive and negative values, weight signs were not informative about directionality in the analysis. A detailed description of the PLSR regression and the detailed rationale behind choosing the unsigned weights is provided in the Supplementary Materials.

Genetic modules and enrichment analyses

We used Enrichr (<http://amp.pharm.mssm.edu/Enrichr>) [38, 39] to test for enrichment of significant PLSR genes for each component against Gene Ontology Biological Processes and report significant results after Benjamini–Hochberg FDR correction ($q < 0.05$). Cell-type specific enrichment was conducted for five broad classes of cells: neurons, astrocytes, oligodendrocytes, microglia and vascular cells [40]. We defined cell-type specific genes as the top 500 genes with higher expression in the cell-type compared to the remaining five genes. As these classes of genes are largely distinct with minimal overlap, we used a Bonferroni correction to correct for cell-type specific enrichment.

We also investigated the enrichment in different classes of risk genes for autism using logistic regression (more detail on each class of genes can be found in the supplementary materials):

1. Transcriptionally dysregulated genes ($n = 1143$, 584 upregulated and 558 downregulated in the autism cortex) were identified from Parikshak et al. [28].
2. Adult gene co-expression modules [28].
3. Fetal gene co-expression modules [31].
4. Genes enriched for rare, de novo, putative loss of function variants (rare, de novo genes, $n = 65$) were identified from Sanders et al. [23].
5. Common genetic variants associated with autism were downloaded from the latest data freeze from the Psychiatric Genomics Consortium (5305 cases and 5305 pseudocontrols). Gene based P -values and Z scores were obtained using MAGMA for each gene [41].

Enrichment analyses for the different classes of autism risk genes were done using logistic regression after accounting for gene length as a covariate. Enrichments are reported as significant if they had a Benjamini-Hochberg FDR adjusted P -value < 0.05 [42] and if they have an enrichment odds ratio (OR) > 1 . The supplementary material provides further details about the gene sets and the methods used.

Validation and specificity

We conducted extensive validation of our initial results against two independent datasets and checked for specificity of an autism effect against a matched ADHD dataset. There are significant phenotypic and genetic correlations between the two conditions, and we had access to MRI data from children with ADHD [34], making this a suitable dataset for testing specificity. Details on all these three datasets are provided in the supplementary materials.

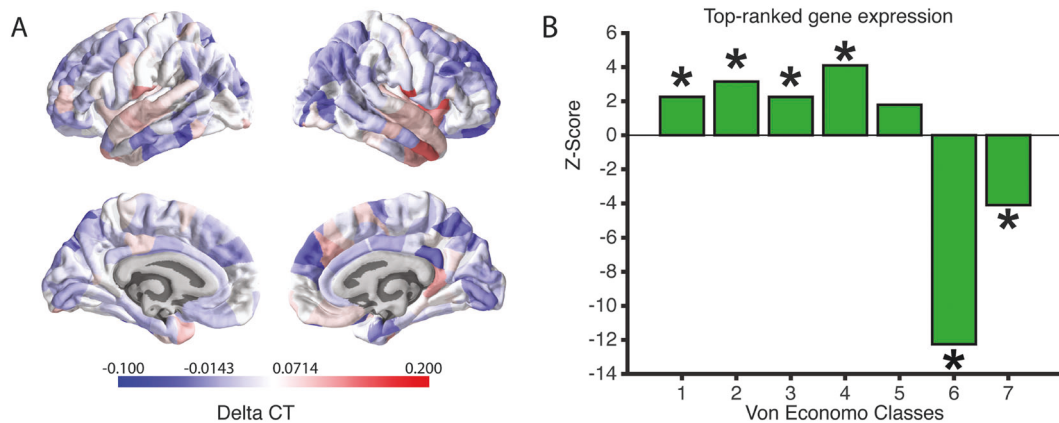


Fig. 2 Expression and Von Economo classification for PLSR1. The heatmap in **a** shows the Δ CT distribution across all 308 cortical regions. The barplot in **b** shows the z-scores of the mean distribution across the different Von Economo Classes (Class 1: granular cortex, primary motor cortex. Class 2: association cortex. Class 3: association cortex. Class 4: dysgranular cortex, secondary sensory cortex. Class 5:

agranular cortex, primary sensory cortex. Class 6: limbic regions, allocortex. Class 7: insular cortex.). All significant over- or under-expression classes are marked with an asterisk. To determine significance, we used permutation testing and a false discovery rate corrected p-value < 0.025 to fully account for two-tailed testing

Results

Autism discovery MRI dataset

PLSR analyses and characterization

Cross-validation using an initial 35 component analyses identified that a 13-component model had the best fit (Supplementary Table S1). Note that the number of components chosen for the model does not affect the individual component composition. Consequently, PLSR was run using a 13-component model. Four components (Components 1, 3, 4 and 6) explained more than 10% of the variance (Supplementary Figure S1). However, variance in Δ CT explained by PLS components was higher than expected by chance only for the first component ($P = 0.009$, 10,000 permutations) but not for the remaining components ($P = 0.303$, $P = 0.693$ and $P = 0.394$, for components 3, 4 and 6, respectively). Thus, only PLSR1 was used for subsequent analyses and we only included genes that passed FDR correction ($q < 0.05$). Only the GO term “Synaptic Transmission” in component 1 (PLSR1) survived FDR correction for multiple comparisons ($P_{\text{corrected}} = 0.00006$). PLSR1 was also significantly enriched for 11 pathways (Table S2) in the Kyoto Encyclopedia of Genes and Genomes (KEGG). There was a significant positive correlation between Δ CT and the scores of PLSR1 ($r = 0.32$; $P = 4.15 \times 10^{-9}$).

Cell-type specific analysis identified a significant enrichment for neurons (OR = 1.1; $P_{\text{corrected}} = 3.19 \times 10^{-12}$), but no enrichment for genes enriched in astrocytes (OR = 1; $P_{\text{corrected}} = 1$), oligodendrocytes (OR = 0.99; $P_{\text{corrected}} = 1$), microglia (OR = 0.97; $P_{\text{corrected}} = 0.43$) or vascular cells (OR = 0.97; $P_{\text{corrected}} = 0.62$).

Topographical enrichment analyses

Previous studies reported an association between CT and cytoarchitectural cortical features [43] linked to specific abnormalities in laminar thickness of supragranular layers of the cortex of schizophrenia patients [44]. Here, we also conducted spatial characterization of PLSR1 across all 5 Von Economo classes [45] as well as an additional 2 subtype classes covering limbic regions and allocortex (class 6) and insular cortex (class 7) [17, 46]. We expected a potential differentiation between higher order cognitive processing and more primary sensory processing. The genes in PLSR1 were significantly over-expressed in secondary sensory and association cortices (VE classes 2, 3 and 4: all $P_{\text{corrected}} < 0.01$) compared to a null distribution. In limbic and insular regions, however these genes appeared to be under-expressed (VE classes 6 and 7: all $P_{\text{corrected}} < 0.01$). However, they also appear to be over-expressed in granular and primary motor cortices (VE Class 1). Figure 2 shows the results from the spatial characterization of the first component across all VE classes.

Gene enrichment analyses

We identified a significant enrichment for genes that are dysregulated in the autism post-mortem cortex (OR = 1.21; $P_{\text{corrected}} < 2.81 \times 10^{-15}$), driven entirely by genes downregulated in autism cortex (OR = 1.87; $P_{\text{corrected}} < 3.55 \times 10^{-16}$). In comparison, there was no enrichment for upregulated genes (OR = 1.01; $P_{\text{corrected}} = 0.49$). The downregulated genes have been previously reported to be enriched for several GO terms including synaptic transmission [28].

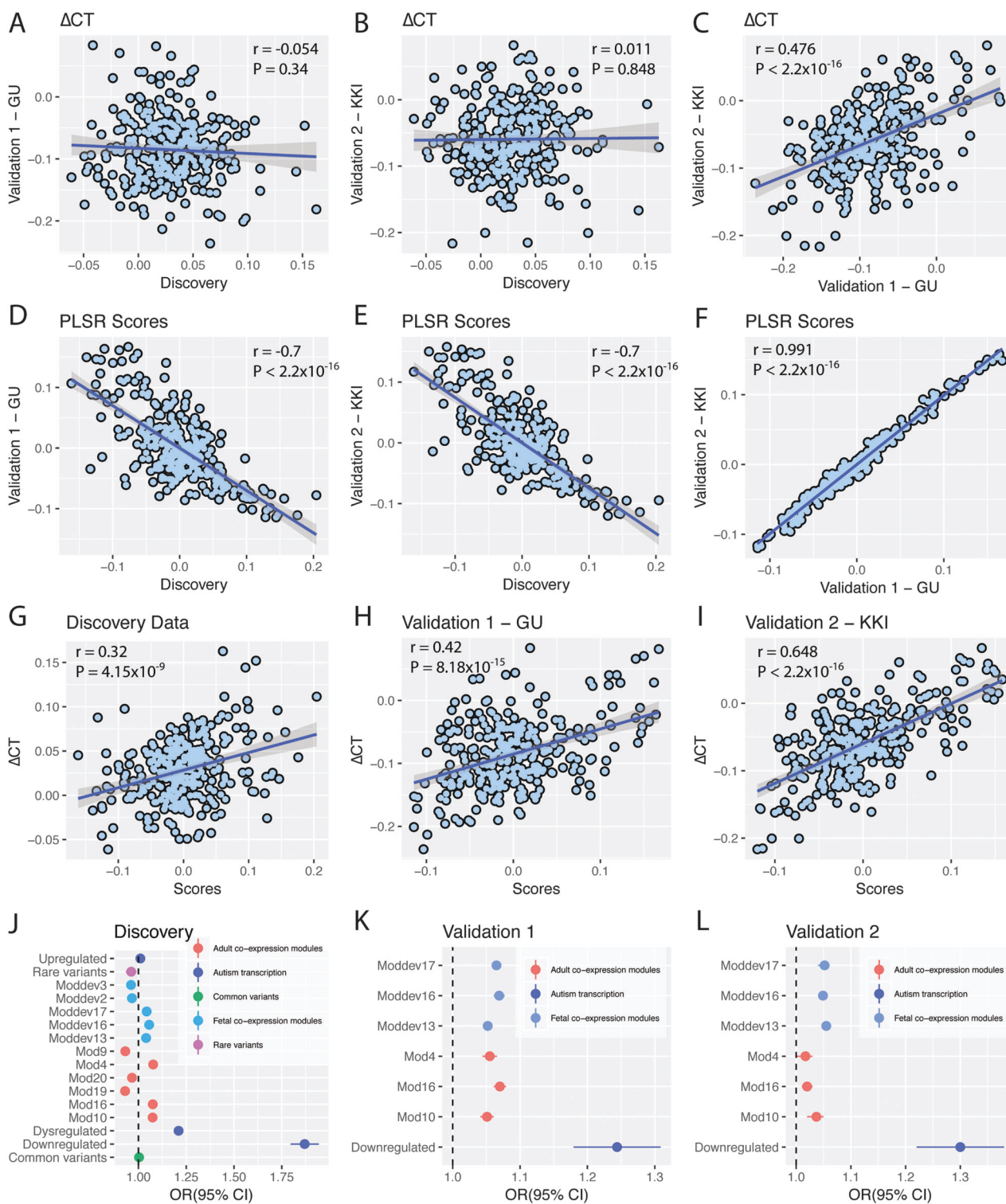


Fig. 3 Gene enrichment and dataset comparisons. **a–c** Show the correlation between Δ CT in the three datasets. **d–f** Show the correlation between the PLSR scores of all three datasets. **g–i** Show the correlation between Δ CT and the PLSR scores in all three datasets (indicating that increased scores are strongly correlated with increased Δ CT). **j** Shows

the odds ratios for the gene-enrichment analysis in the discovery dataset. All significantly enriched modules were replicated in the validation datasets (**k** and **l**) apart from module 4 of the adult co-expression modules. Pearson correlation coefficient and P -values of the correlations are provided in the top of the respective panels

Table 2 Gene enrichment

Category	Dataset	OR	Upper CI (95%)	Lower CI (95%)	<i>P</i>	<i>P</i> _{corrected}
Autism transcription	Dysregulated	1.21	1.23	1.19	1.76E−15	2.81E−15
Autism transcription	Downregulated	1.87	1.94	1.8	2.00E−16	3.55E−16
Autism transcription	Upregulated	1.01	1.02	1	4.99E−01	4.99E−01
Adult co-expression modules	Mod4	1.08	1.08	1.07	2.00E−16	3.55E−16
Adult co-expression modules	Mod10	1.07	1.08	1.07	2.00E−16	3.55E−16
Adult co-expression modules	Mod16	1.08	1.08	1.07	2.00E−16	3.55E−16
Adult co-expression modules	Mod9	0.93	0.94	0.92	2.01E−14	2.92E−14
Adult co-expression modules	Mod19	0.93	0.94	0.92	2.00E−16	3.55E−16
Adult co-expression modules	Mod20	0.97	0.97	0.96	6.22E−05	7.66E−05
Common variants	Common variants	1	1.01	1	2.75E−01	2.93E−01
Rare variants	Rare variants	0.96	0.99	0.93	2.42E−01	2.76E−01
Fetal co-expression modules	Moddev2	0.97	0.97	0.96	1.28E−11	1.70E−11
Fetal co-expression modules	Moddev3	0.96	0.97	0.96	2.00E−16	3.55E−16
Fetal co-expression modules	Moddev13	1.04	1.04	1.04	2.00E−16	3.55E−16
Fetal co-expression modules	Moddev16	1.06	1.06	1.05	2.00E−16	3.55E−16
Fetal co-expression modules	Moddev17	1.04	1.05	1.04	2.00E−16	3.55E−16

Odds ratio scores, confidence intervals and significance of all major classes of gene enrichment investigated in the discovery dataset

Transcriptionally dysregulated genes can reflect several different underlying processes. To provide better resolution of the processes involved, we next investigated if this enrichment was associated with six adult co-expression modules associated with dysregulated autism genes [28]. Three of these were associated with genes downregulated in the autism postmortem cortex (M4, M10, M16), and three were enriched for genes upregulated in the autism post-mortem cortex (M9, M19 and M20) compared to controls. As we had identified a significant enrichment for down-regulated autism genes but not for the upregulated autism genes, we hypothesized that gene co-expression modules associated with downregulated genes would also be enriched for association with PLSR1 genes. Indeed, PLSR1 was enriched for all three downregulated modules but none of the 3 upregulated modules. See Fig. 3j, Table 2 and supplementary Table S6.

We also investigated if the significant genes in PLSR1 were enriched in specific cortical developmental modules [31]. The Mdev13, Mdev16 and the Mdev17 modules are enriched for transcriptionally dysregulated genes in autism postmortem frontal and temporal cortices [31]. The Mdev2 and the Mdev3 modules are enriched for rare variants identified in autism [31]. Again, we identified significant enrichment for three adult co-expression modules enriched for transcriptionally dysregulated genes. For the two modules associated with rare, de novo variants, we identified fewer PLSR1 genes than expected by chance. See Fig. 3j, Table 2 and Supplementary Table S6. We did not identify a significant enrichment for rare, de novo genes. We also did not identify a significant enrichment for common variants

using MAGMA to collapse SNP based *P*-values to gene based *P*-values (OR = 1.00; *P*_{corrected} = 0.29). Results of the gene enrichment analysis are provided in Fig. 3j, Table 2 and Supplementary Table S6.

Validation of initial findings

PLSR analyses and characterization

We validated all analyses using Δ CT from two independent cohorts (Table 1). There was no correlation in Δ CT between the discovery and the two validation datasets (Fig. 3a,b), which is in line with recent large scale assessments of autism neuroimaging studies [13]. This may be explained by factors such as heterogeneity due to scanner sites in the discovery dataset, age of onset of puberty and clinical conditions. There was a significant positive correlation in Δ CT between the two validation datasets ($r = 0.476$; $P < 2.2 \times 10^{-16}$). Heterogeneity in autism neuroimaging studies is well documented and complex [13, 47], but it should be emphasized that the present analysis focuses on the relation between whole-brain variation in Δ CT and whole-brain variation in gene expression, thus a lack of spatial overlap in Δ CT does not affect the Δ CT—Gene relation.

Again, only the first component (PLSR1-validation1 and PLSR1-validation2; see Fig. 4) ($P < 10^{-14}$, 10,000 permutations) (Supplementary Figure S2) explained a significant amount of the variance. There was a significant positive correlation between Δ CT and the gene expression scores in both validation datasets (Fig. 3h,i). Further, PLSR1 was enriched for the GO term ‘Synaptic transmission’

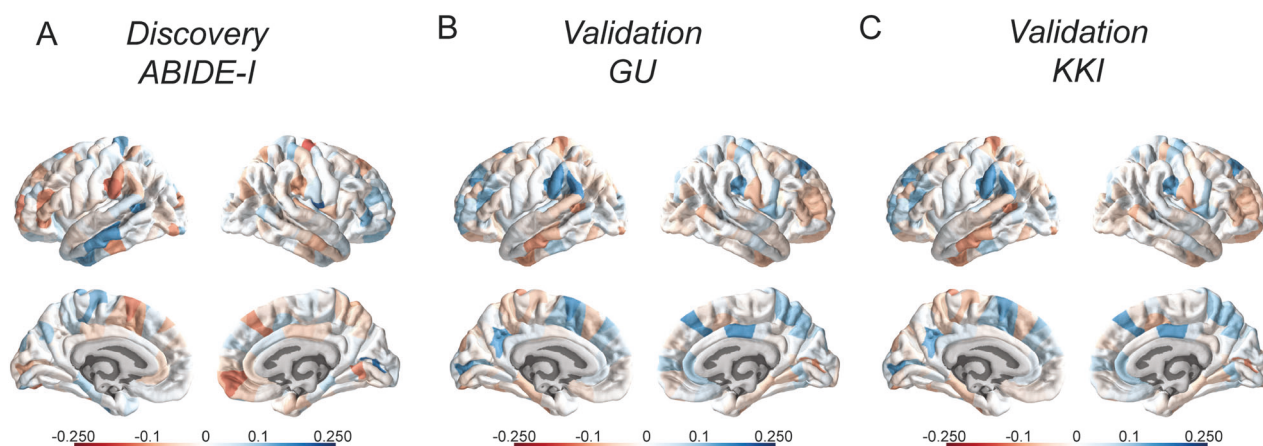


Fig. 4 PLSR1 scores for all autism datasets. **a–c** represent the PLSR1 scores for the three autism datasets across 308 cortical regions.

a Represents the discovery dataset, **b** represents Validation 1 and **c** represents Validation 2

($P_{\text{corrected}} = 1 \times 10^{-4}$ for both Validation 1 and Validation 2). In addition, for Validation 2, the PSLR component was also enriched for the GO term ‘Membrane depolarization’ ($P_{\text{corrected}} = 7 \times 10^{-4}$). KEGG pathway enrichment for both datasets are provided in Supplementary Tables S3 and S4. Further, we replicated the cell-type enrichment in genes expressed in neurons (OR = 1.13; $P = 2.01 \times 10^{-7}$ for Validation 1 and OR = 1.06; $P = 0.009$ for Validation 2).

Given the lack of consistent correlations in ΔCT between the original discovery dataset and both validation datasets we also added another validation dataset to confirm the PLS findings (Supplementary section 8, Figure S6). Results from this dataset were in line with the two original validation datasets and PLS weights were consistent across datasets.

Gene enrichment analyses

We replicated the significant enrichment of transcriptionally downregulated genes in the autism post mortem cortex (Validation 1: OR = 1.24; $P_{\text{corrected}} = 0.004$; Validation 2: OR = 1.3; $P_{\text{corrected}} = 0.001$), providing confidence in the robustness of our initial results (Fig. 3k,l and Supplementary Table S6). Mirroring the enrichment with the downregulated genes in the autism post-mortem cortex, we also identified enrichment for the three adult gene co-expression modules that are enriched for downregulated genes in Validation 1 (M4, M10 and M16) and two of the three (M10 and M16) adult gene co-expression modules for Validation 2 (Fig. 3k,l, Supplementary Table S6). We also replicated the enrichment for the three fetal gene co-expression modules (Mdev13, Mdev16 and Mdev17) in both validation datasets (Fig. 3k,l, Supplementary Table S6).

To understand the genes involved in this process, we focused on the top 200 genes (approximately 1%) based on their gene weights in PLS1 across all the three datasets

(Supplementary Dataset). This identified several genes that have been implicated in autism, synaptic processes and neural development. For example, one of the genes identified is *SCN1A*, which encodes a voltage-gated sodium channel, is one of the genes identified to be frequently mutated in autism [48–50]. *SLIT1* and *SLIT3* are important for regulating midline axon crossing in the developing forebrain [51]. Focusing on just the top 1% in the two validation datasets, given their high correlation, we identified several other genes that have been implicated in autism and synaptic transmission including *GABRA3*, *GABRA5* and *GABRB1*, all of which encode subunits of the GABA receptor. Mutations in *PTCHD1*, another gene identified in the top 1%, have been implicated in autism and intellectual disabilities and contributes to dysfunction of excitatory synapses [52]. *SYN2* and *SYT17* encode synapsin2 and synaptotagmin17 respectively, which are present in the pre-synaptic terminal and regulate neurotransmitter release.

Comparison with ADHD

PLSR analysis of ΔCT in ADHD data did not identify any components that significantly explained the variance in ΔCT (Supplementary Table S5). Thus, we did not consider the ADHD dataset for further analyses. Details of the number of components, the model fit and the variance explained are provided in the supplementary materials.

Discussion

Here we report the association of transcriptionally downregulated genes in the autism post-mortem cortex with global differences in CT in 166 children with autism and 295 neurotypical children. Using partial least squares regression on a discovery dataset of 62 cases and 87

controls, we identify one component (PLSR1) that explains a significant proportion of variance in Δ CT and is enriched for the GO term 'Synaptic Transmission' and for neuronal genes. This component was enriched for genes downregulated in the autism post-mortem cortex and validated using two independent datasets. We also find that PLSR1 genes are enriched for fetal and adult developmental cortical modules that have been previously reported to be enriched for transcriptionally dysregulated genes in the post-mortem autism cortex and for genes involved in synaptic transmission [28, 31]. We were unable to identify genes associated with Δ CT in ADHD, another childhood condition. Our study provides robust evidence linking disease-related variance in CT to synaptic genes and dysregulated genes in the autism post-mortem cortex, linking molecular and macroscopic pathology.

Validation using two independent autism MRI datasets suggests that the results are valid even using MRI data from different cohorts that had different scanner settings. The results were valid despite non-significant correlations in global Δ CT between the discovery and the two validation datasets and sex did not contribute to any of the observed differences between datasets (see Supplementary material). This suggests that the same sets of genes are associated with Δ CT regardless of sex. Studies have identified differences in cortical morphology between neurotypical males and females and between males and females with autism [2, 53]. Here we identified a high correlation between a males-only dataset and two males and females combined MRI datasets for the gene weights and gene scores in the first PLS component.

Changes in CT may be due to a host of factors such as changes in myelination, synaptic pruning and dendritic arborisation. Evidence from rare genetic variants [54, 55] and transcriptionally dysregulated genes in autism have highlighted a role for synaptic transmission in the aetiology of autism [28, 29]. Transcriptional dysregulation may reflect either a causative risk mechanism for autism, or a compensatory consequence of genetic, hormonal and environmental risk for autism. Here, we are unable to disentangle if transcriptionally dysregulated genes causally contribute to cortical morphology changes, or if they are both downstream of a common risk mechanism, or both. It is possible that both CT variability and transcriptional dysregulation are downstream processes of genetic variation implicated in autism, and, as such the enrichment for transcriptionally dysregulated genes need not be causative of cortical morphological changes.

We did not identify enrichments for rare, de novo loss of function genes or common variants implicated in autism. The lack of enrichment with rare, de novo loss of function genes may be due to both the relative low frequencies of such variants and small proportion of variance in liability

explained by rare de novo variants [20]. In contrast, the lack of enrichment with common variants may be explained by the lack of statistical power of the largest available autism GWAS dataset. Indeed, there is no enrichment for common genetic variants associated with autism in co-expression modules enriched for transcriptionally dysregulated genes in autism [28]. In contrast, common variants for schizophrenia are enriched in co-expression modules associated with dysregulated genes in schizophrenia [56]. It is likely that larger samples will better reveal the role of common genetic variants in cortical morphology differences in autism. While we do not know the genetic make-up of the cases and controls, our results likely represent common downstream convergence of upstream genetic perturbations.

Animal studies have shown that several candidate genes for autism risk are regulated by synaptic activity, leading to the hypothesis that dysregulation in synaptic homeostasis is a major risk for autism [55]. The effects of this can contribute to both neural signal processing, and to more morphological changes in neuroanatomy via processes such as activity dependent synaptic pruning and dendritic arborization. Post-mortem studies of the brains of children and adolescents with autism have identified deficits in synaptic pruning [57]. Investigating the specific role of synaptic genes in altering neural circuitry and cortical morphology will help elucidate the precise molecular mechanisms underlying CT differences seen in autism.

There are some caveats that need to be taken into consideration while interpreting these results. Gene expression data was derived from only six post-mortem adult brain samples. Gene expression is known to vary with age [58, 59]. Unfortunately, we are restricted in using the adult gene expression data from the AIBS for several reasons. First, this is the most spatially detailed dataset of gene expression. Second, the availability of MNI coordinates in the adult gene expression datasets allows for mapping of gene expression in distinct brain regions to CT differences extracted from MRI scans. Third, gene expression changes with age are limited and restricted to specific brain regions. A recent study identified only 9 genes significantly altered globally across the 10 regions investigated in post-mortem tissue samples [60], largely driven by glial genes. Cell specific enrichment in our dataset implicated neuronal genes only. Fourth, as autism is a developmental condition, investigating differences in cortical morphology at an early age is important to limit the role of environmental factors that contribute to differences in cortical morphology later in life [8, 61]. Fifth, enrichment for gene expression modules associated with autism risk in the developing cortex provides further confidence that the genes identified here are relevant across the age-spectrum. We do acknowledge that investigating a paediatric specific gene-expression dataset will help further refine the analyses, once this data becomes available.

Lastly, the present study used CT in contrast to other morphological features such as cortical volume. It is known that grey matter volume relies on the relationship between two different morphometric parameters, CT and surface area that are unrelated genetically [62] and associated with different developmental trajectories [63]. The combination of at least two different sources of genetic and maturational influences on cortical volume would complicate meaningful analysis of associated genetic weights.

To our knowledge, this is the first study linking different genetic risk mechanisms in autism with changes in cortical morphology. In sum, we have shown that genes that are enriched for synaptic transmission and downregulated in individuals with autism are significantly associated with global changes in CT. We also show that these genes are generally overexpressed in association cortices. We validated the results in multiple independent datasets but not in a matched MRI dataset that included individuals with ADHD, showing both replicability as well as selectivity.

Acknowledgements This study was funded by grants from the Medical Research Council, UK, the Templeton World Charity Foundation, the Autism Research Trust and the Wellcome Trust. We thank Luke Kweku Abraham for help with interpretation of the PLSR gene weights. We thank Dr. Elijah Mak and František Váša for help with quality control of the MRI images. We thank Dr. Neelroop Parikhshak for help with regression-based enrichment analyses and Prof. Dan Geschwind for valuable comments on the manuscript. RRG was funded by the Neuroscience in Psychiatry Network, Wellcome Trust and the Guarantors of Brain. VW was funded by St. John's College, the Cambridge Commonwealth Trust. RAIB was funded by the Medical Research Council, Autism Research Trust, Pinsent Darwin Trust and Cambridge Trust. SBC is supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East of England at Cambridgeshire and Peterborough NHS Foundation Trust. ETB is employed half-time by the University of Cambridge and half-time by GlaxoSmithKline (GSK); he holds stock in GSK. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, GSK or the Department of Health.

Compliance with ethical standards

Conflict of interest ETB is employed half-time by GSK and holds stock in GSK. The remaining authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Press; 2013. .
2. Lai M-C, Lombardo MV, Suckling J, Ruigrok ANV, Chakrabarti B, Ecker C, et al. Biological sex affects the neurobiology of autism. *Brain*. 2013;136:2799–815.
3. Ecker C. The neuroanatomy of autism spectrum disorder: an overview of structural neuroimaging findings and their translatability to the clinical setting. *Autism*. 2016;21:18–28.
4. Mensen VT, Wierenga LM, van Dijk S, Rijks Y, Oranje B, Mandl RCWW, et al. Development of cortical thickness and surface area in autism spectrum disorder. *NeuroImage Clin*. 2016;13:215–22.
5. Ecker C, Ginestet CE, Feng Y, Johnston P, Lombardo MV, Lai M-C, et al. Brain surface anatomy in adults with autism. *JAMA Psychiatry*. 2013;70:59.
6. Hazlett HC, Gu H, Munsell BC, Kim SH, Styner M, Wolff JJ, et al. Early brain development in infants at high risk for autism spectrum disorder. *Nat Publ Gr*. 2017;542:348–51.
7. Smith E, Thurm A, Greenstein D, Farmer C, Swedo S, Giedd JN, et al. Cortical thickness change in autism during early childhood. *Hum Brain Mapp*. 2016;37:2616–29.
8. Zielinski BA, Prigge MBDD, Nielsen JA, Froehlich AL, Abildskov TJ, Anderson JS, et al. Longitudinal changes in cortical thickness in autism and typical development. *Brain*. 2014;137:1799–812.
9. Hyde KL, Samson F, Evans AC, Mottron L. Neuroanatomical differences in brain areas implicated in perceptual and other core features of autism revealed by cortical thickness analysis and voxel-based morphometry. *Hum Brain Mapp*. 2010;31:556–66.
10. Scheel C, Rotarska-Jagiela A, Schilbach L, Lehnhardt FG, Krug B, Vogeley K, et al. Imaging derived cortical thickness reduction in high-functioning autism: key regions and temporal slope. *Neuroimage*. 2011;58:391–400.
11. Hardan AY, Muddasani S, Vemulapalli M, Keshavan MS, Minshew NJ. An MRI study of increased cortical thickness in autism. *Am J Psychiatry*. 2006;163:1290–2.
12. Hadjikhani N, Joseph RM, Snyder J, Tager-Flusberg H. Anatomical differences in the mirror neuron system and social cognition network in autism. *Cereb Cortex*. 2006;16:1276–82.
13. Haar S, Berman S, Behrmann M, Dinstein I. Anatomical abnormalities in autism? *Cereb Cortex*. 2016;26:1440–52.
14. Yang DY-J, Beam D, Pelphrey KA, Abdullahi S, Jou RJ. Cortical morphological markers in children with autism: a structural magnetic resonance imaging study of thickness, area, volume, and gyrification. *Mol Autism*. 2016;7:11.
15. McKavanagh R, Buckley E, Chance SA. Wider mimocolumns in autism: a neural basis for altered processing? *Brain*. 2015;138:2034–45.
16. Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivieres S, Jahanshad N, et al. Common genetic variants influence human subcortical brain structures. *Nature*. 2015;520:224–9.
17. Whitaker KJ, Vértes PE, Romero-Garcia R, Váša F, Moutoussis M, Prabhu G, et al. Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proc Natl Acad Sci USA*. 2016;113:9105–10.
18. Elman JA, Panizzon MS, Hagler DJ, Fennema-Notestine C, Eyler LT, Gillespie NA, et al. Genetic and environmental influences on cortical mean diffusivity. *Neuroimage*. 2017;146:90–99.
19. Eyler LT, Chen C, Panizzon MS, Fennema-Notestine C, Neale MC, Jak A, et al. A comparison of heritability maps of cortical surface area and thickness and the influence of adjustment for whole brain measures: a magnetic resonance imaging twin study. *Twin Res Hum Genet*. 2012;15:304–14.

20. Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB, et al. Most genetic risk for autism resides with common variation. *Nat Genet.* 2014;46:881–5.
21. Bourgeron T. Current knowledge on the genetics of autism and propositions for future research. *C R Biol.* 2016;339:300–7.
22. Tick B, Bolton P, Happé F, Rutter M, Rijdsdijk F. Heritability of autism spectrum disorders: a meta-analysis of twin studies. *J Child Psychol Psychiatry.* 2016;57:585–95.
23. Sanders SJSJ HeX, Willsey AJJ, Ercan-Sencicek AGG, KEKE Samocha, Cicek AEE, et al. Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron.* 2015;87:1215–33.
24. Kosmicki JA, Samocha KE, Howrigan DP, Sanders SJ, Slowikowski K, Lek M, et al. Refining the role of de novo protein truncating variants in neurodevelopmental disorders using population reference samples. *bioRxiv.* 2016;49:52886.
25. Wang J, Barstein J, Ethridge LE, Mosconi MW, Takarae Y, Sweeney JA. Resting state EEG abnormalities in autism spectrum disorders. *J Neurodev Disord.* 2013;5:1–14.
26. Yuen RK, Merico D, Cao H, Pellecchia G, Alipanahi B, Thiruvahindrapuram B, et al. Genome-wide characteristics of de novo mutations in autism. *npj Genom Med.* 2016;1:16027.
27. Stessman HAF, Xiong B, Coe BP, Wang T, Hoekzema K, Fenckova M, et al. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat Genet.* 2017;49:515–26.
28. Parikshak NN, Swarup V, Belgard TG, Irimia M, Ramaswami G, Gandal MJ, et al. Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature.* 2016;540:423–7.
29. Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature.* 2011;474:380–4.
30. Gupta S, Ellis SE, Ashar FN, Moes A, Bader JS, Zhan J, et al. Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. *Nat Commun.* 2014;5:5748.
31. Parikshak NN, Luo R, Zhang A, Won H, Lowe JK, Chandran V, et al. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell.* 2013;155:1008–21.
32. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts A, Shen EH, Ng L, Miller JA, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature.* 2012;489:391–9.
33. Hawrylycz M, Miller JA, Menon V, Feng D, Dolbeare T, Guillozet-Bongaarts AL, et al. Canonical genetic signatures of the adult human brain. *Nat Neurosci.* 2015;18:1832–44.
34. Bethlehem RAI, Romero-Garcia R, Mak E, Bullmore ET, Baron-Cohen S. Structural covariance networks in children with autism or ADHD. *Cereb Cortex.* 2017;27:4267–76.
35. Romero-Garcia R, Whitaker KJ, Váša F, Seidlitz J, Shinn M, Fonagy P, et al. Structural covariance networks are coupled to expression of genes enriched in supragranular layers of the human cortex. *Neuroimage.* 2017;171:256–67.
36. Romero-Garcia R, Whitaker KJ, Vasa F, Seidlitz J, Shinn M, Fonagy P, et al. Data supporting NSPN publication 'Structural covariance networks are coupled to expression of genes enriched in supragranular layers of the human cortex.' Cambridge Data Repos. 2017. <https://doi.org/10.17863/CAM.11392>.
37. Bigdeli TB, Lee D, Webb BT, Riley BP, Vladimirov VI, Fanous AH, et al. A simple yet accurate correction for winner's curse can predict signals discovered in much larger genome scans. *Bioinformatics.* 2016;32:2598–603.
38. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics.* 2013;14:128.
39. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 2016;44:W90–7.
40. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keefe S, et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci.* 2014;34:11929–47.
41. de Leeuw CA, Mooij JM, Heskes T, Posthuma D, De Leeuw CA, Mooij JM, et al. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol.* 2015;11:1–19.
42. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B.* 1995;57:289–300.
43. Wagstyl K, Ronan L, Goodyer IM, Fletcher PC. Cortical thickness gradients in structural hierarchies. *Neuroimage.* 2015;111:241–50.
44. Wagstyl K, Ronan L, Whitaker KJ, Goodyer IM, Roberts N, Crow TJ, et al. Multiple markers of cortical morphology reveal evidence of supragranular thinning in schizophrenia. *Transl Psychiatry.* 2016;6:e780.
45. Von Economo C, Koskinas GN. Atlas of cytoarchitectonics of the adult human cerebral cortex. 1st ed. Basel, Switzerland: Karger; 2008.
46. Zilles K, Amunts K. Segregation and wiring in the brain. *Science.* 2012;335:1582–4.
47. Ecker C, Murphy D. Neuroimaging in autism—from basic science to translational research. *Nat Rev Neurol.* 2014;10:82–91.
48. Weiss LA, Escayg A, Kearney JA, Trudeau M, MacDonald BT, Mori M, et al. Sodium channels SCN1A, SCN2A and SCN3A in familial autism. *Mol Psychiatry.* 2003;8:186–94.
49. O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature.* 2012;485:246–50.
50. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Ercument Cicek A, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature.* 2014;515:209–15.
51. Plump AS, Erskine L, Sabatier C, Brose K, Epstein CJ, Goodman CS, et al. Slit1 and Slit2 cooperate to prevent premature midline crossing of retinal axons in the mouse visual system. *Neuron.* 2002;33:219–32.
52. Ung DC, Iacono G, Méziane H, Blanchard E, Papon M-A, Selten M et al. Ptchd1 deficiency induces excitatory synaptic and cognitive dysfunctions in mouse. *Mol Psychiatry.* 2017. <https://doi.org/10.1101/208223>.
53. Sowell ER, Peterson BS, Kan E, Woods RP, Yoshii J, Bansal R, et al. Sex differences in cortical thickness mapped in 176 healthy individuals between 7 and 87 years of age. *Cereb Cortex.* 2007;17:1550–60.
54. Bourgeron T. A synaptic trek to autism. *Curr Opin Neurobiol.* 2009;19:231–4.
55. Bourgeron T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat Rev Neurosci.* 2015;16:551–63.
56. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci.* 2016;19:1442–53.
57. Tang G, Gudsnek K, Kuo SH, Cotrina ML, Rosoklija G, Sosunov A, et al. Loss of mTOR-dependent macroautophagy causes autistic-like synaptic pruning deficits. *Neuron.* 2014;83:1131–43.
58. Glass D, Viñuela A, MNMN Davies, Ramasamy A, Parts L, Knowles D, et al. Gene expression changes with age in skin, adipose tissue, blood and brain. *Genome Biol.* 2013;14:R75.
59. Somel M, Guo S, Fu N, Yan Z, Hu HYH, Xu Y, et al. MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res.* 2010;20:1207–18.

60. Soreq L, Rose J, Soreq E, Hardy J, Trabzuni D, Cookson MR, et al. Major shifts in glial regional identity are a transcriptional hallmark of human brain aging. *Cell Rep.* 2017;18:557–70.
61. Lange N, Travers BG, Bigler ED, Prigge MBD, Froehlich AL, Nielsen JA, et al. Longitudinal volumetric brain changes in autism spectrum disorder ages 6-35 years. *Autism Res.* 2015;8:82–93.
62. Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M, et al. Distinct genetic influences on cortical surface area and cortical thickness. *Cereb Cortex.* 2009;19:2728–35.
63. Herting MM, Gautam P, Spielberg JM, Dahl RE, Sowell ER. A longitudinal study: changes in cortical thickness and surface area during pubertal maturation. *PLoS ONE.* 2015;10:e0119774.