Archival Report

The Neurogenetics of Functional Connectivity Alterations in Autism: Insights From Subtyping in 657 Individuals

Javier Rasero, Antonio Jimenez-Marin, Ibai Diez, Roberto Toro, Mazahir T. Hasan, and Jesus M. Cortes

ABSTRACT

BACKGROUND: There is little consensus and controversial evidence on anatomical alterations in the brains of people with autism spectrum disorder (ASD), due in part to the large heterogeneity present in ASD, which in turn is a major drawback for developing therapies. One strategy to characterize this heterogeneity in ASD is to cluster large-scale functional brain connectivity profiles.

METHODS: A subtyping approach based on consensus clustering of functional brain connectivity patterns was applied to a population of 657 autistic individuals with quality-assured neuroimaging data. We then used high-resolution gene transcriptomic data to characterize the molecular mechanism behind each subtype by performing enrichment analysis of the set of genes showing a high spatial similarity with the profiles of functional connectivity alterations between each subtype and a group of typically developing control participants.

RESULTS: Two major stable subtypes were found: subtype 1 exhibited hypoconnectivity (less average connectivity than typically developing control participants) and subtype 2, hyperconnectivity. The 2 subtypes did not differ in structural imaging metrics in any of the analyzed regions (68 cortical and 14 subcortical) or in any of the behavioral scores (including IQ, Autism Diagnostic Interview, and Autism Diagnostic Observation Schedule). Finally, only subtype 2, comprising about 43% of ASD participants, led to significant enrichments after multiple testing corrections. Notably, the dominant enrichment corresponded to excitation/inhibition imbalance, a leading well-known primary mechanism in the pathophysiology of ASD.

CONCLUSIONS: Our results support a link between excitation/inhibition imbalance and functional connectivity alterations, but only in one ASD subtype, overall characterized by brain hyperconnectivity and major alterations in somatomotor and default mode networks.

https://doi.org/10.1016/j.biopsych.2023.04.014

Autism encompasses multiple manifestations, from impaired social communication and language to restricted or repetitive behavior patterns, interests, and activities (1–3). Due to the vast heterogeneity in behavior, and as recommended in DSM-5, this condition is referred to as autism spectrum disorder (ASD), in which the term "spectrum" emphasizes the variation in the type and severity of manifestations (4). ASD is thought to result from complex interactions during development between genetic, cellular, circuit, epigenetic, and environmental factors (5–9). Several researchers have suggested that an excitation/inhibition (E/I) imbalance during development (10,11) may be an essential mechanism, yet specific factors driving the condition are not well understood. Therapeutic interventions aiming to restore the E/I balance in ASD are a major challenge (12).

Concerning neurobiology, heterogeneity in brain morphology (13) and brain networks has been found, e.g., in the frontal, default mode, and salience networks (14–19), as well as in the social network (20)—encompassing the primary motor cortex, fusiform gyrus, amygdala, cerebellum, insula, somatosensory cortex, and anterior cingulate cortex (14,21,22). ASD is also heterogeneous in relation to network characteristics; less segregation and greater efficiency (23,24), and the opposite as well (25) or a combination of both (26,27), have been shown. Furthermore, ASD neuroanatomical correlates are not static but undergo changes throughout development (28–30), and the same seems to occur behaviorally in social functioning and communication (31). Altogether, accumulated evidence has shown high heterogeneity within ASD in the participation of functional brain networks and behavioral manifestations and in the longitudinal trajectories at the individual level.

Moreover, with neuroimaging studies, recent work has shown additional sources of heterogeneity due to variations in diagnostic and inclusion criteria and differences in the processing neuroimaging pipeline (32,33). ASD is also a polygenic, highly heterogeneous condition, with 1010 genes associated with ASD as of July 8, 2022, according to the Simons Foundation Autism Research Initiative (SFARI) gene

human database [see also (34)]. Of those, 213 have a relevance score of 1, meaning that they have maximum published pathophysiological evidence to ASD. This high genetic complexity is another manifestation of the heterogeneity of this condition on several scales. Previous work has assessed the associations between transcriptomics and brain morphology (35), showing that genes that are downregulated and enriched for synaptic transmission in individuals with autism were associated with variations in cortical thickness.

Novel strategies for ASD subtyping are needed to overcome such multiscale heterogeneity, which is the most significant challenge in the development of effective therapies. Some studies have addressed the heterogeneity in ASD to better stratify this condition (36–39). Previous work performed clustering, pooling together ASD and typically developing control (TDC) groups (37), and found 2 groups of individuals showing hyperconnected or hypoconnected patterns (each group containing both ASD and TDC participants). Stratification yields reduced interindividual differences and, therefore, could complement—and even alleviate—the need for large sample sizes in autism-based biomarker discovery (40). Here, and following previous work (36,41,42), we looked at large-scale brain connectivity patterns common within groups of individuals to deploy subtyping in ASD. In particular, we applied consensus clustering strategies to multivariate connectivity patterns of brain regions (43,44) for associating connectivity-based ASD subtypes with their neurogenetic profile. Following previous work (45–52), we hypothesized different biological characterization underlying the neurodevelopmental and maturation brain connectivity profile for each subtype, unknown for this condition. For this, we used the Allen Human Brain Atlas (AHBA) of whole-brain transcriptional data (53) and performed subtyping on 657 individuals with ASD from the Autism Brain Imaging Data Exchange (ABIDE) repository (54), all of them having passed a very strict quality assurance criterion of elimination of participants by head movement during image acquisition, thus correcting a well-known spurious excess of functional connectivity driven by head movements, which is even more pronounced in the autistic condition. Moreover, to overcome interscanner variability in the functional connectivity values across different institutions, we applied rigorous harmonization strategies to transform heterogeneous data into equivalents (55–58).

METHODS AND MATERIALS

Participants

A total of 2156 participants from the ABIDE-I (54) and ABIDE-II (59) repositories were initially considered in this study, of which 1026 were individuals with ASD and 1130 were TDC participants. These data were collected across 35 different scanning cohorts. For each participant, both anatomical and functional magnetic resonance imaging (MRI) data were used. Acquisition parameters for each scanning site are found at http://fcon_1 000.projects.nitrc.org/indi/abide/. Additionally, we extracted several composite scores from the Autism Diagnostic Observation Schedule-Generic, Autism Diagnostic Interview-Revised, Vineland Adaptive Behavior Scales, Social Responsiveness Scale, Social Communication Questionnaire, and raw score of the Autism Quotient, and the verbal, performance, and Full Scale IQ scores to address cognitive performance and

disorder severity. After data quality assurance (see the Supplement), the final number of included participants was 1541 (884 TDC and 657 ASD).

Functional Connectivity Matrices

After neuroimaging preprocessing using state-of-the-art methodology (see the Supplement), FreeSurfer version 5.3.0 was used for brain segmentation and cortical parcellation. A total of 82 regions were generated from the Desikan-Killiany atlas, with 68 cortical regions (34 in each hemisphere) and 14 subcortical regions segmented from FreeSurfer (left/right thalamus, caudate, putamen, pallidum, hippocampus, amygdala, and accumbens). Different brain parcellations were also considered for the analyses (see the Supplement). For each participant, the parcellations were projected to the individual functional data and the mean functional time series of each region was obtained. Finally, one connectivity matrix for each participant was built by Fisher's Z-transformation of the Pearson correlation coefficients between the region pairs of the time series.

Data Harmonization

To harmonize our multi-institution functional connectivity data, and before performing subtyping, we used an in-house implementation of Combat (https://pypi.org/project/pycombat), adjusting these multi-institution batch effects by linear mixed modeling and the use of empirical Bayes methods (56). We also included in this model the diagnosis label (TDC or ASD) as a biological variable of interest, ensuring that group-level connectivity differences were preserved after harmonization. See the Supplement for details.

ASD Subtyping via Consensus Clustering

Consensus clustering was applied to brain connectivity matrices (43,44). Because connectivity matrices may contain effects of no interest (e.g., age), prior to subtyping, we regressed out age, sex, and motion from each connectivity entry of the participants with ASD. This regression-out step was only applied at this subtyping stage. In subsequent analyses, the harmonized connectivity matrices were used, and the effect of these variables was controlled for by using them as covariates. The stability of each subtype and the 95% CIs of the estimated maximum modularity were assessed by bootstrapping (60).

Statistical Differences in Brain Morphology and Behavior Between ASD Subtypes

We applied multiple linear regression to assess statistical differences between ASD subtypes in regionwise volume and thickness from FreeSurfer (https://surfer.nmr.mgh.harvard.edu), while controlling for age, sex, and total intracranial volume and a one-way analysis of variance for differences in behavior. Multiple testing was corrected by controlling the false discovery rate (FDR).

Association Between Subtypes and Transcriptomics

We computed the association between functional connectivity alterations represented by pseudo-R^2 maps (larger R^2 scores

corresponding to greater alterations) (see the Supplement) and brain transcriptomics maps using spatial autoregressive models, well known to reduce the correlation bias produced by the similar transcriptomic expression in proximal brain regions (49). This analysis was implemented by means of the maximum-likelihood estimator routine (ML_Lag) from the Python Spatial Analysis Library (pysal) (61). As a result, for each gene we obtained one t-statistic and one p value, which allowed us to assess the association with the pseudo-R^2 maps while accounting for possible spatial autocorrelations. Among the significantly associated genes, we identified as relevant those genes included in the SFARI database (https://gene.sfari.org/) with a gene score equal to 1. Several post hoc analyses tested the robustness of the significant association between the ASD subtypes and transcriptomics and whether our enrichment findings were specific to the ASD condition (see the Supplement).

Gene Set Enrichment Analysis and Protein Interaction Analysis

We only considered for the analyses such genes with FDR-corrected p (p_{FDR}) value < .05 in each subtype. After that, we performed a gene set enrichment analysis using WebGestalt (62) (http://www.webgestalt.org/), introducing as the input the list of the corrected genes and the t-statistic from the association analysis. We computed the gene set enrichment analysis for gene ontology (GO) biological process (63) and Reactome pathways (64) and only considered enriched categories p_{FDR} value < .05. We further applied an ensemble-based enrichment analysis, similar to the one developed in (65), to evaluate whether significant enrichment annotations were affected by inflation or false positive bias (65). First, we generated 10,000 surrogate brain maps with the same spatial autocorrelation as the original pseudo-R^2 maps using BrainSMASH tool (66). For each of the surrogate maps, we computed the association with brain transcriptome maps using spatial autoregressive models and used those genes with a p_{FDR} value < .05 for computing enrichment analysis. For each of the annotations of interest, we generated a distribution of the likelihood of that annotation being significant by these random maps. These distributions were used to compute a p value to evaluate false positive bias. For the protein interaction analysis, we used the tool STRING version 11.5 (67) to generate a physical protein-protein interaction network for each subtype, with experiments and databases as interaction sources. These networks were then analyzed using Cytoscape version 3.9.0.

RESULTS

We obtained harmonized functional connectivity matrices from 657 ASD and 884 TDC participants following the method represented in Figure S1. For subtyping, we first removed any effect from age, sex, and head motion in the brain connectivity matrices of the ASD group and then applied a consensus clustering. We thus found 2 main subtypes[1]: the first with 348

---

[1]For subtyping, we used connectivity matrices only from participants with ASD, although the association with brain transcriptomics was performed for functional connectivity alterations encoded in the pseudo-R^2 maps.

participants (52.97% of all participants with ASD) and the second with 284 participants (43.23%). In addition to these 2 subtypes, which were at the highest order in a hierarchy that broke down into smaller subtypes (Figure S2), we also found 2 more residual subtypes of only 23 participants (3.5%) and 2 participants (0.3%), respectively. Our clustering solution, which exhibited modularity statistically different from 0 (0.181, 95% CI, 0.169–0.194), provided for the 2 major subtypes a stability score of 0.953 and 0.819, respectively, suggesting high consistency and recovery after resampling (mean stability > 0.75) [see (60)]. In contrast, the residual subtypes were mostly not replicable during bootstrapping (mean stability < 0.5); as a result, they were ignored for further analysis. Furthermore, the robustness of the subtyping solution was assessed by 2 different strategies, namely multiresolution hierarchical clustering and cross-validation (see the Supplement). As expected, given that their effects were removed prior to subtyping, none of the resulting subtypes were differentiated by age (absolute Cohen's $|d|$ = 0.04, t test, p = .58), sex (Cramer's V = 0.02, χ^2 test, p = .55), or head motion (absolute Cohen's $|d|$ = 0.04, t test, p = .64).

With respect to cognitive and behavioral performance, the 2 subtypes were highly similar to each other, because among 10 different scores compared, only 2 of them gave uncorrected statistical differences (Autism Diagnostic Observation Schedule total p = .03, Social Responsiveness Scale total p = .05), which became nonsignificant after correcting for multiple comparisons (Autism Diagnostic Observation Schedule total p_{FDR} = .27, Social Responsiveness Scale total p_{FDR} = .27). For further details on the subtype comparisons, see Table 1. Furthermore, no significant structural differences between subtypes 1 and 2 were found in region volume or thickness. Therefore, all the following analyses are based on differences in functional connectivity that each ASD subtype has in relation to TDC.

To assess the differences between groups in the overall connectivity per participant, defined here as the average positive correlation of the harmonized connectivity matrix (negative correlations were excluded due to the lack of consensus about their origin), we ran multiple linear regression while controlling for age, sex, and full IQ (Figure 1). Subtype 1 showed significant hypoconnectivity to TDC (β = −0.08, t_{1227} = −14.86, p < .01). The opposite was true for subtype 2 (β = 0.04, t_{1163} = 6.91, p < .01), thus corresponding to hyperconnectivity. Moreover, the difference in (absolute) β coefficients provided for subtype 1 higher values than for subtype 2, indicating a larger separability in connectivity to TDC. Other metrics for defining overall connectivity per participant led to similar conclusions (see the Supplement).

Next, we assessed the differences in connectivity patterns between each ASD subtype and the TDC group, measured by regionwise normalized pseudo-R^2 brain maps, resulting from multivariate distance matrix regression (Figure 2; Supplement). The spatial similarity between these maps was very low (r_{80} = 0.09, permutation-based p = .67, after using 5000 surrogates that preserved spatial autocorrelation), indicating that each subtype exhibited a distinct neurobiological profile of brain-wide connectivity, as expected since the subtypes were obtained by clustering the functional connectivity profiles. Specifically, for subtype 1, higher differences as compared

Table 1. Behavioral Characterization of ASD Subtypes

Measurement ($n_{subtype\ 1}/n_{subtype\ 2}$)	Subtype 1, Mean (95% CI)	Subtype 2, Mean (95% CI)	p Value	FDR-Corrected p Value
Full Scale IQ (321/266)	105.23 (103.54–106.99)	106.81 (104.96–108.61)	.25	.49
Verbal IQ (273/241)	104.43 (102.28–106.51)	106.78 (104.44–108.98)	.13	.45
Performance IQ (277/245)	105.0 (103.18–106.96)	105.4 (103.27–107.56)	.80	.97
ADI Total (228/190)	35.21 (34.07–36.33)	34.07 (32.59–35.51)	.21	.49
ADOS Total (221/179)	11.86 (11.37–12.36)	11.01 (10.46–11.58)	.03	.27
Vineland Sum Scores (53/49)	248.98 (237.57–260.68)	254.84 (240.00–268.96)	.54	.77
Vineland ABC Standard (53/49)	79.43 (76.3–82.7)	79.16 (75.53–82.61)	.91	.97
SRS Raw Total (213/166)	94.45 (90.46–98.35)	88.49 (83.64–93.14)	.05	.27
SCQ Total (87/71)	18.78 (17.22–20.38)	17.68 (16.14–19.23)	.32	.54
AQ Total (22/21)	32.05 (29.27–34.55)	32.14 (27.81–36.19)	.97	.97

p Values are from a one-way analysis of variance test to assess any statistical difference. For ABIDE data, we calculated the ADI total as the sum of the Abnormalities in Reciprocal Social Interaction and the Abnormalities in Communication (Verbal Subjects Only) scores, and the ADOS total as the sum of the Communication and the Social Interaction scores.

ABIDE, Autism Brain Imaging Data Exchange; ABC, Adaptive Behavior Composite; ADI-R, Autism Diagnostic Interview-Revised; ADOS, Autism Diagnostic Observation Schedule; AQ, Autism Spectrum Quotient; ASD, autism spectrum disorder; FDR, false discovery rate; SCQ, Social Communication Questionnaire; SRS, Social Responsiveness Scale.

with TDC were found in the superior temporal gyrus, posterior cingulate cortex, and the insula, covering the functional networks of default mode and salience. For subtype 2, higher differences existed in the thalamus, similar to previous work (68), putamen, and precentral gyrus. Thus, alterations affecting the default mode network were common to both subtypes, but one (subtype 1) also showed specific disruptions involving the

Figure 1. Two major stable ASD subtypes, one with hypoconnectivity and the other with hyperconnectivity. Histogram and box plots of the individual average connectivity values (measured as Fisher's Z) for the TDC group (blue), the population of all participants with ASD without subtyping (brown), and the 2 ASD subtypes (pink and orange). Two other subtypes that were not stable after permutation testing are not depicted here and have been ignored for further analysis. The median value of the TDC group is marked as the connectivity baseline by a dashed vertical line. Values above the baseline correspond to overall hyperconnectivity and those below the baseline to overall hypoconnectivity. Subtype 1 is dominated by hypoconnectivity and subtype 2 by hyperconnectivity. Additionally, for subtypes 1 and 2, we introduce 2 colors for participants to show within-group connectivity differences: blue for hypoconnectivity, and red for hyperconnectivity. ASD, autism spectrum disorder; TDC, typically developing control.

salience network and the other (subtype 2) in the somatomotor network.

For the biological characterization of each subtype, we set out to identify which genes had an expression across brain regions significantly associated ($p_{FDR} < .05$) with the differences in connectivity measured by the normalized R^2 brain maps (Figure 2, histograms), whereby larger R^2 values correspond to larger functional connectivity alterations. For subtype 1, a total of 195 negative-associated (NEG) genes and 364 positive-associated (POS) genes existed. Significant NEG genes, also present in the SFARI gene human database with a relevance score of 1, were *GFAP, CHD7, SKI, SHANK3, ANK3,* and *CACNA1E,* while POS genes were *ASXL3, MAP1A, STXBP1, DPYSL2, KNCB1, SCN8A, RIMS1,* and *CDKL5.* Similarly, for subtype 2, we found 142 NEG genes, of which *GRIA2, RFX3, SHANK2, GRIN2B, DLG4, LRRC4C, ARX,* and *GABRB3* were also present in the SFARI list, and 180 POS genes, including *MAGEL2* and *IQSEC2.* We next applied gene enrichment to the list of significant genes within each subtype, finding no significant enrichment for subtype 1, the type with brain hypoconnectivity. However, for subtype 2, the enrichment of the NEG genes included GO biological processes and Reactome pathways related to glutamate signaling (affecting both AMPA and NMDA receptors) and synapse organization in relation to the E/I imbalance occurring during the development of brain circuits (Figure 3A). We also assessed which NEG genes participated in each biological process and pathway (Figure 3B), finding that genes *DLG4, GRIN2B, GRIA2,* and *SHANK2* were participating in most of them; and, the gene *DLG4* plays a role in all of them. Additionally, the *DLG4* gene was the one with the highest degree in the protein interaction network.

We found a significant enrichment with biological processes related to E/I imbalance for subtype 2 but not for subtype 1. To test whether these findings suggested that the functional connectivity in subtype 1 was different from that in previous studies of ASD, we calculated the similarity of the connectivity profiles of our 2 subtypes with typical connectivity alterations in ASD, represented by brain maps in (69) and calculated from 4 different ASD databases of resting functional MRI data. We
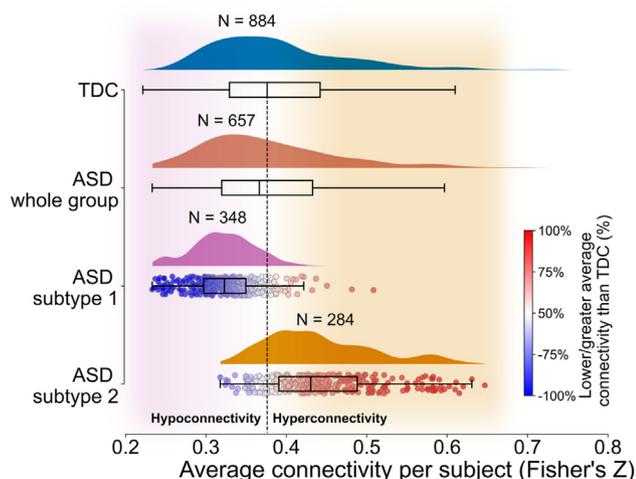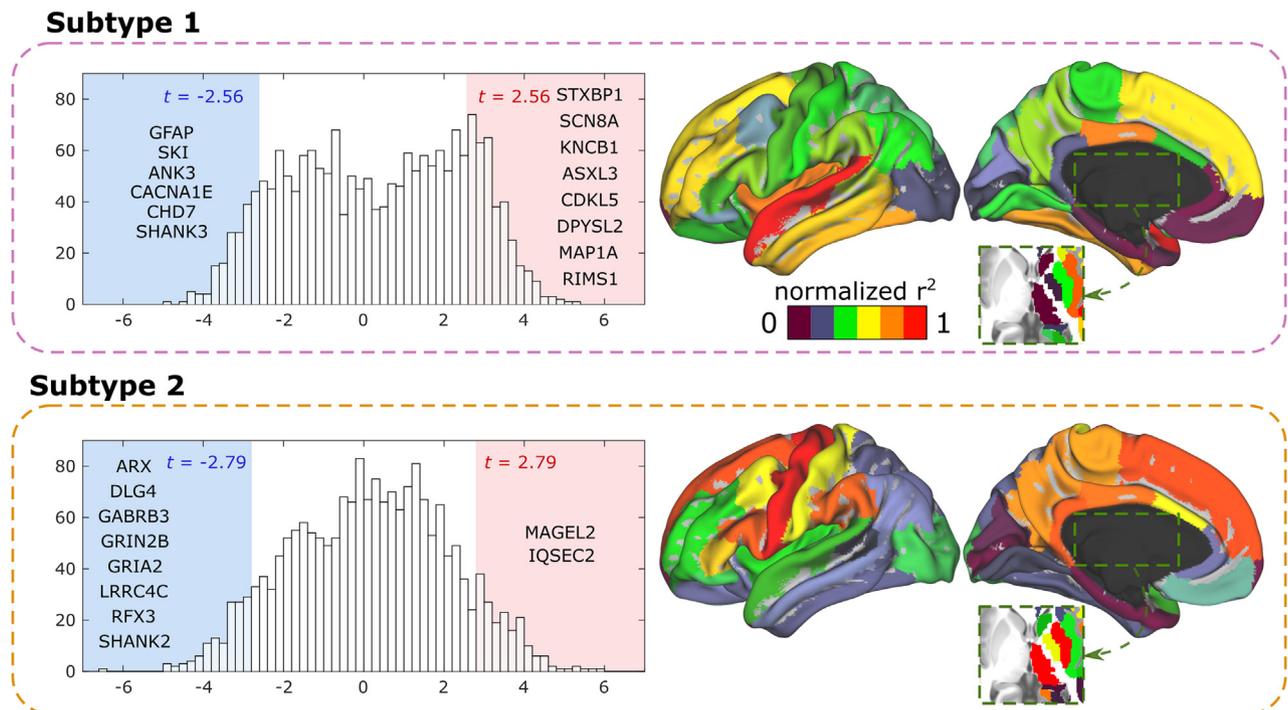
Figure 2. Association between transcriptomics and connectivity patterns for each autism spectrum disorder subtype. For subtypes 1 and 2, we calculated the pseudo-R^2 map, considering the differences in the connectivity pattern that each subtype has from typically developing control participants. (Right) Brain maps of normalized pseudo-R^2. (Left) Histograms of association values between pseudo-R^2 and gene transcription activity (different values correspond to association with different genes). This procedure was repeated using the pseudo-R^2 map for each subtype. The tail of the negative genes (false discovery rate–corrected $p < .05$ and t-statistic < 0) is marked with a blue rectangle and the tail of the positive genes (false discovery rate–corrected $p < .05$ and t-statistic > 0) with a red one for both subtypes. Significance limits (t) are also shown. For each distribution tail, we also show the relevant genes present in the SFARI autism spectrum disorder genes with a score = 1.

calculated for each subtype the average spatial similarity across the 4 existing brain maps of connectivity alterations in (69). For subtype 1 the average similarity was not significant ($r_{80} = 0.19$, $p = .45$), but for subtype 2 it was significant ($r_{80} = 0.46$, $p = .02$), indicating that subtype 2 more closely resembled the typical connectivity alterations reported in ASD, which in fact is the subtype for which we found significant enrichment toward E/I imbalance.

We also compared the transcriptomic-connectivity results from the generalized Louvain algorithm to those found by multiresolution clustering (see the Supplement). As a measure of similarity between the 2 solutions, Dice index values of the solutions were 0.99 and 0.89, respectively, which indicated a high level of reproducibility of the gene expression association with brain alterations between the 2 clustering strategies.

Additionally, we studied the effects of considering a different brain partition on the results of the transcription-connectivity association (see the Supplement). Using the functionally defined Schaefer brain partition with 100 different regions, the association results obtained from the Desikan-Killiany atlas as compared to those from the Schaefer partition had very low similarity for subtype 1 ($r_{1880} = -0.11$, $p < .001$), and slightly higher results were found for subtype 2 ($r_{1880} = 0.40$, $p < .001$). By adding the same subcortical regions to the Schaefer partition (see the Supplement), the gene association became very similar for the 2 brain partitions and for

the 2 subtypes (subtype 1, $r_{1880} = 0.87$, $p < .001$; subtype 2, $r_{1880} = 0.92$, $p < .001$), suggesting a strong contribution of the subcortical alterations to the robustness of our association results. A description of the relevant genes related to subtype 2 is shown in Table S2. These results were obtained by using left-hemisphere transcription sites[2], but the results were also preserved when we repeated the analysis for the 2 brain hemispheres (Figure S3).

Finally, it is important to note that no significant enrichment was found for subtypes lower in the dendrogram level corresponding to the 2 subtypes described above (Figure S2). The significant enrichment did not exist after repeating the same analysis using the entire ASD group, indicating the need for subtyping first in the entire population to reveal our findings. To prove that our gene enrichment findings were specific to the ASD condition (see the Supplement), we repeated the same procedure using only the TDC population in 2 matched subgroups of TDC participants, one used for subtyping and the other for estimating the pseudo-R^2 maps. As a result, no gene survived FDR correction in any subtype, thus indicating that the E/I imbalance found in the hyperconnected autistic

---

[2]We focused on the left hemisphere, as all donors provided sampling sites of genes in this hemisphere, and only 2 of the 6 donors from the AHBA dataset were sampled in both left and right hemispheres.
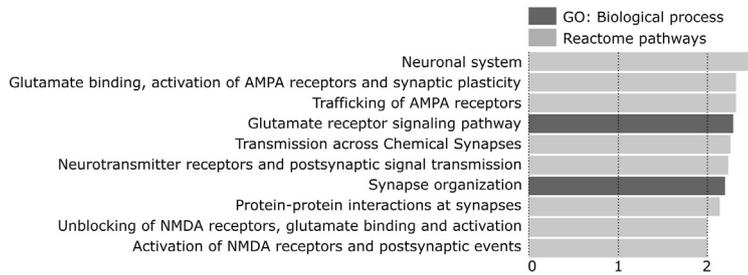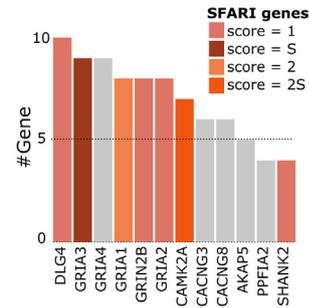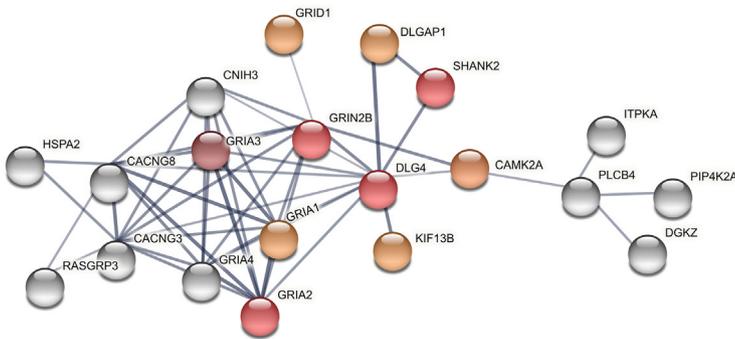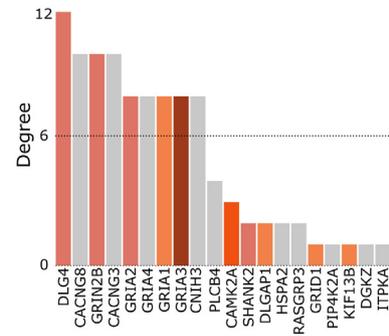
Figure 3. Excitation/inhibition imbalance enrichment for only one class of participants with autism spectrum disorder (subtype 2). **(A)** GSEA characterization of the false discovery rate–significant genes in subtype 2, including the GO biological processes (dark gray) and Reactome pathways (light gray) enrichments. We further tested whether the enrichment findings were affected by their reporting rate in the literature, and for all cases reported here, we obtained false discovery rate–corrected $p < .05$. **(B)** Participation count for each gene in the processes shown in **(A)** ranging from 4 to 10 (corresponding to a participation in all processes that only occurred for *DLG4*). **(C)** Protein-protein interaction physical network from the list of false discovery rate–significant genes. For ease of visualization, only subnetworks with a minimum of 10 genes are depicted. **(D)** Node degree of the genes participating in the network shown in **(C)**. *DLG4* is the gene with the highest degree. **(B–D)** Bars corresponding to genes with SFARI score = 1 are colored in red; SFARI score = 1S in dark red; SFARI score = 2 in orange; and SFARI score = 2S in dark orange, and the same color code was used in **(C)** and **(D)** for network nodes. GO, gene ontology; GSEA, gene set enrichment analysis.

subtype is specific to the autistic condition. Likewise, although the subtyping performed in both ASD and TDC groups resulted in solutions with similar overall connectivity separation, resulting in 2 sets of hypo- and hyperconnected brains, a multivariate distance matrix regression analysis applied to the functional correlation patterns showed that the hyperconnected subtype found in ASD was statistically different from that in TDC ($p < .001$). The hypoconnected subtypes in ASD and TDC were also different from each other ($p < .001$). This might explain why no similar findings in the enrichment were found for the hyperconnected TDC subgroup. In summation, the significant association between E/I imbalance and altered functional connectivity was observed when subtyping in ASD, and only in the ASD group characterized by overall hyperconnectivity, demonstrating the specificity of the reported enrichment.

DISCUSSION

Two significant subtypes result from functional connectivity–based subtyping in a cohort of 657 individuals with ASD. The two are indistinguishable by behavioral scores, and also by morphometric comparisons based on structural neuroimaging,

in agreement with recent results ([40](#)). Compared with the TDC group, the first subtype is characterized by hypoconnectivity, with major implications in the superior temporal gyrus, posterior cingulate cortex, and insula, showing connectivity alterations in the default mode and salience networks with no significant gene enrichment after correcting for multiple comparisons. The second subtype, representing 43% of participants with autism, is characterized by hyperconnectivity, with major implications in the thalamus, putamen, and precentral gyrus and showing network alterations in somatomotor and default mode networks. In a recent analysis linking genomics and resting functional connectivity in 32,726 individuals with psychiatric conditions, significant ASD contributions were shown in the thalamic and somatomotor networks ([68](#)), consistent with our results for subtype 2. Only subtype 2 had a significant gene enrichment toward glutamate signaling (affecting both AMPA and NMDA receptors), consistent with the E/I imbalance that occurs during brain development and one of the most accepted hypotheses in the pathophysiology of autism ([70](#)). Indeed, it is thought that in the development of ASD, there is an increase in the ratio between excitation and inhibition, leading to hyperexcitability of cortical circuits ([10](#)). It is also possible that differential E/I alteration of selective brain

circuits might result in an unaltered E/I ratio at the network level (11). Our work maps patterns of functional connectivity alterations with genes that are involved in E/I balance. While it is true that perturbation in these genes in animal models strongly affects E/I imbalance in brain networks (71), the participant data that we analyzed in this study do not directly address the E/I imbalance, and this is a limitation of our methodology. It is also important to emphasize that the E/I enrichment found in our study is specific to the ASD condition and, as such, does not occur in the TDC group. Moreover, the connectivity profile in the entire autistic population, i.e., if no subtyping is performed, does not have significant enrichment, indicating the need for subtyping first to find the connection with E/I imbalance in one subtype of individuals with ASD.

Our subtyping approach was based on patterns of functional connectivity alterations. There are 3 major reasons supporting our choice not to use structural features for our subtyping analysis. First, it would require a different clustering approach to the one adopted here, which is based on the consensus of connectivity patterns. Second, and based on recent data-driven results from an international autism imaging biomarker challenge (40) with more than 146 institutions submitting prediction algorithms, the 10 best-performing algorithms (with ASD prediction accuracies having area under the curve > 0.80) showed a dominant contribution of the functional modality, with a much higher discriminative power than the structural MRI data. Third, our main goal was to study the origin of functional connectivity–based heterogeneity in autism, and structural features (representing different brain aspects) give rise to a different kind of heterogeneity. As a result, the proper combination of these 2 diverse sources of heterogeneity would require a multimodal approach different from the one developed here.

Our approach is unique in several ways. First, our study is based on a large cohort of individuals with ASD (N = 657) from the ABIDE initiative, all of them having passed the rigorous criteria of motion removal, and it combines anatomical and functional neuroimaging data from 24 different institutions. Second, we used Combat, a rigorous data harmonization method to eliminate the variability between MRI scans across the 24 institutions, one of the largest sources of variability when combining imaging data from multiple institutions (72). Third, our analysis of brain connectivity was carried out on a large scale, in which each brain region is represented by its connectivity pattern across the entire brain. Therefore, we did not consider a priori any brain region as more dominant or relevant than the others. Fourth, we made use of a consensus clustering approach that we developed (43,44), and that has been successfully tested by others (73), to group participants in the same subtype if the connectivity profiles are similar across all the analyzed regions. Finally, we made use of the AHBA to describe the neurogenetic profiles of each subtype, which has been used before for morphometric information in ASD (35) but never for characterizing subtypes based on functional connectivity patterns of this condition.

Due to the heterogeneity and diversity reported in ASD genetics, the use of AHBA may shed new light, because it provides information on the transcriptome across the brain in unprecedented detail, accounting for 3702 sampling sites with transcription information on 20,500 genes as a specific signature for each anatomical region. Moreover, the use of AHBA is complementary to other techniques, such as genome-wide association studies (74), that simultaneously address genotype-phenotype associations from hundreds of thousands to millions of genetic variants in a data-driven manner. Indeed, genome-wide association studies have previously been used for ASD subtyping (75,76) using behavioral scores as traits and, therefore, the subtypes obtained were more closely related to symptom severity and not to functional connectivity.

Our enrichment results for subtype 2 show that *DLG4*, also known as *PSD95*, is a gene with major implications in the protein interaction network of subtype 2. *DLG4* mediates NMDA and AMPA receptor clustering and function; it affects glutamatergic transmission and has been shown to have an aberrant function in ASD (77–81). *DLG4* also influences the size and density of dendritic spines during brain development, having strong effects on synaptic connectivity and activity, e.g., reduced *DLG4* activity leads to increased dendritic spine numbers (82).

Some limitations should be noted. First, our transcriptomic analysis was based on AHBA, which is derived from healthy, and not from ASD, brain tissues. Therefore, the relations studied here between ASD-dependent connectivity patterns and healthy transcriptomics highlight large-scale organization aspects of the connectivity alterations to gene expression. Future studies should confirm our findings using gene expression data from a pathological cohort, which is not currently available. Second, the number of donors from the AHBA data is very limited (n = 6) and the sampling sites available do not cover the full brain. Third, our subtyping method found 2 subtypes of ASD participants who were hypoconnected and hyperconnected at the network level. The same classes of subtypes were found by subtyping the TDC group. However, when comparing first the hypoconnectivity subtypes between the ASD and TDC groups and then the hyperconnectivity subtypes, connectivity patterns were significantly different in both cases, which justifies the significant enrichment found for the hyperconnectivity ASD subtype, but not for TDC. Finally, our main neurogenetic finding in one ASD subtype, involving genes largely affecting the E/I imbalance, is based only on the statistical association between transcriptome activity and patterns of functional connectivity alterations. Future studies should explicitly test the causal link between E/I imbalance and functional connectivity alterations in ASD.

In summary, our novel approach, which includes data harmonization, multivariate distancing in large-scale functional connectivity patterns, and transcriptome brain maps, reveals strong enrichment for glutamate signaling (affecting both AMPA and NMDA receptors) and synapse organization in one subgroup of participants with ASD, reinforcing the hypothesis of an E/I imbalance occurring during brain development of individuals with ASD.

ARTICLE INFORMATION

From the Cognitive Axon Laboratory, Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania (JR); Computational Neuro-imaging Laboratory, Biocruces-Bizkaia Health Research Institute, Bar-akaldo, Spain (AJ-M, JMC); Biomedical Research Doctorate Program, University of the Basque Country, Leioa, Spain (AJ-M); Department of Radiology, Division of Nuclear Medicine and Molecular Imaging, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts (ID); Gordon Center for Medical Imaging, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts (ID); Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts (ID); Institut Pasteur, Université de Paris, Département de neuroscience, Paris, France (RT); Laboratory of Brain Circuits Therapeutics, Achucarro Basque Center for Neuroscience, Leioa, Spain (MTH); Ikerbas-que, The Basque Foundation for Science, Bilbao, Spain (MTH, JMC); and Department of Cell Biology and Histology, University of the Basque Country, Leioa, Spain (JMC).

JR and AJ-M contributed equally to this work as joint first authors.

Address correspondence to Javier Rasero, Ph.D., at jrasero.daparte@gmail.com.

REFERENCES

1. Masi A, DeMayo MM, Glozier N, Guastella AJ (2017): An overview of autism spectrum disorder, heterogeneity and treatment options. Neurosci Bull 33:183–193.
2. Lai M-C, Lombardo MV, Baron-Cohen S (2014): Autism. Lancet 383:896–910.
3. Hodges H, Fealko C, Soares N (2020): Autism spectrum disorder: Definition, epidemiology, causes, and clinical evaluation. Transl Pediatr 9(suppl 1):S55–65.
4. Mehling MH, Tassé MJ (2016): Severity of autism spectrum disorders: Current conceptualization, and transition to DSM-5. J Autism Dev Disord 46:2000–2016.
5. Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J (2018): Autism spectrum disorder. Lancet 392:508–520.
6. Eapen V (2011): Genetic basis of autism: Is there a way forward? Curr Opin Psychiatry 24:226–236.
7. Bhandari R, Paliwal JK, Kuhad A (2020): Neuropsychopathology of autism spectrum disorder: Complex interplay of genetic, epigenetic, and environmental factors. Adv Neurobiol 24:97–141.
8. Hertz-Picciotto I, Schmidt RJ, Krakowiak P (2018): Understanding environmental contributions to autism: Causal concepts and the state of science. Autism Res 11:554–586.
9. Momoi T, Fujita E, Senoo H, Momoi M (2009): Genetic factors and epigenetic factors for autism: Endoplasmic reticulum stress and impaired synaptic function. Cell Biol Int 34:13–19.
10. Rubenstein JLR, Merzenich MM (2003): Model of autism: Increased ratio of excitation/inhibition in key neural systems. Genes Brain Behav 2:255–267.
11. Nelson SB, Valakh V (2015): Excitatory/inhibitory balance and circuit homeostasis in autism spectrum disorders. Neuron 87:684–698.
12. Sohal VS, Rubenstein JLR (2019): Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders. Mol Psychiatry 24:1248–1257.
13. Lefebvre A, Beggiato A, Bourgeron T, Toro R (2015): Neuroanatomical diversity of corpus callosum and brain volume in autism: Meta-analysis, analysis of the autism brain imaging data exchange project, and simulation. Biol Psychiatry 78:126–134.
14. Patriquin MA, DeRamus T, Libero LE, Laird A, Kana RK (2016): Neuroanatomical and neurofunctional markers of social cognition in autism spectrum disorder. Hum Brain Mapp 37:3957–3978.
15. Catani M, Dell'Acqua F, Budisavljevic S, Howells H, Thiebaut de Schotten M, Froudist-Walsh S, et al. (2016): Frontal networks in adults with autism spectrum disorder. Brain 139:616–630.
16. Uddin LQ, Supekar K, Lynch CJ, Khouzam A, Phillips J, Feinstein C, et al. (2013): Salience network–based classification and prediction of symptom severity in children with autism. JAMA Psychiatry 70:869–879.
17. Heinsfeld AS, Franco AR, Craddock RC, Buchweitz A, Meneguzzi F (2018): Identification of autism spectrum disorder using deep learning and the ABIDE dataset. Neuroimage Clin 17:16–23.
18. Anderson JS, Nielsen JA, Froehlich AL, DuBray MB, Druzgal TJ, Cariello AN, et al. (2011): Functional connectivity magnetic resonance imaging classification of autism. Brain 134:3742–3754.
19. Pua EPK, Ball G, Adamson C, Bowden S, Seal ML (2019): Quantifying individual differences in brain morphometry underlying symptom severity in autism spectrum disorders. Sci Rep 9:9898.
20. Kennedy DP, Adolphs R (2012): The social brain in psychiatric and neurological disorders. Trends Cogn Sci 16:559–572.
21. Thompson A, Murphy D, Dell'Acqua F, Ecker C, McAlonan G, Howells H, et al. (2017): Impaired communication between the motor and somatosensory homunculus is associated with poor manual dexterity in autism spectrum disorder. Biol Psychiatry 81:211–219.
22. Nebel MB, Joel SE, Muschelli J, Barber AD, Caffo BS, Pekar JJ, Mostofsky SH (2014): Disruption of functional organization within the primary motor cortex in children with autism. Hum Brain Mapp 35:567–580.
23. Rudie JD, Brown JA, Beck-Pancer D, Hernandez LM, Dennis EL, Thompson PM, et al. (2012): Altered functional and structural brain network organization in autism. Neuroimage Clin 2:79–94.
24. Itahashi T, Yamada T, Watanabe H, Nakamura M, Jimbo D, Shioda S, et al. (2014): Altered network topologies and hub organization in adults with autism: A resting-state fMRI study. PLoS One 9:e94115.
25. Alaerts K, Geerlings F, Herremans L, Swinnen SP, Verhoeven J, Sunaert S, Wenderoth N (2015): Functional organization of the action observation network in autism: A graph theory approach. PLoS One 10:e0137020.
26. Maximo JO, Kana RK (2019): Aberrant "deep connectivity" in autism: A cortico-subcortical functional connectivity magnetic resonance imaging study. Autism Res 12:384–400.
27. Yerys BE, Herrington JD, Satterthwaite TD, Guy L, Schultz RT, Bassett DS (2017): Globally weaker and topologically different: Resting-state connectivity in youth with autism. Mol Autism 8:39.
28. Nomi JS, Uddin LQ (2015): Developmental changes in large-scale network connectivity in autism. Neuroimage Clin 7:732–741.

29. Courchesne E, Carper R, Akshoomoff N (2003): Evidence of brain overgrowth in the first year of life in autism. JAMA 290:337–344.

30. He C, Chen H, Uddin LQ, Erramuzpe A, Bonifazi P, Guo X, et al. (2020): Structure-function connectomics reveals aberrant developmental trajectory occurring at preadolescence in the autistic brain. Cereb Cortex 30:5028–5037.

31. Fountain C, Winter AS, Bearman PS (2012): Six developmental trajectories characterize children with autism. Pediatrics 129:e1112–e1120.

32. Mueller S, Wang D, Fox MD, Yeo BTT, Sepulcre J, Sabuncu MR, et al. (2013): Individual variability in functional connectivity architecture of the human brain. Neuron 77:586–595.

33. Mottron L, Bzdok D (2020): Autism spectrum heterogeneity: Fact or artifact? Mol Psychiatry 25:3178–3185.

34. SFARI Gene (2023): Human gene module. Available at: https://gene.sfari.org/database/human-gene/. Accessed July 15, 2020.

35. Romero-Garcia R, Warrier V, Bullmore ET, Baron-Cohen S, Bethlehem RAI (2019): Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism. Mol Psychiatry 24:1053–1064.

36. Tang S, Sun N, Floris DL, Zhang X, Di Martino A, Yeo BTT (2020): Reconciling dimensional and categorical models of autism heterogeneity: A brain connectomics and behavioral study. Biol Psychiatry 87:1071–1082.

37. Easson AK, Fatima Z, McIntosh AR (2019): Functional connectivity-based subtypes of individuals with and without autism spectrum disorder. Netw Neurosci 3:344–362.

38. Urchs SGW, Tam A, Orban P, Moreau C, Benhajali Y, Nguyen HD, et al. (2022): Functional connectivity subtypes associate robustly with ASD diagnosis. eLife 11:e56257.

39. Ecker C, Murphy D (2014): Neuroimaging in autism—From basic science to translational research. Nat Rev Neurol 10:82–91.

40. Traut N, Heuer K, Lemaître G, Beggiato A, Germanaud D, Elmaleh M, et al. (2022): Insights from an autism imaging biomarker challenge: Promises and threats to biomarker discovery. Neuroimage 255:119171.

41. Hong SJ, Vogelstein JT, Gozzi A, Bernhardt BC, Yeo BTT, Milham MP, Di Martino A (2020): Toward neurosubtypes in autism. Biol Psychiatry 88:111–128.

42. Lombardo MV, Lai MC, Baron-Cohen S (2019): Big data approaches to decomposing heterogeneity across the autism spectrum. Mol Psychiatry 24:1435–1450.

43. Rasero J, Pellicoro M, Angelini L, Cortes JM, Marinazzo D, Stramaglia S (2017): Consensus clustering approach to group brain connectivity matrices. Netw Neurosci 1:242–253.

44. Rasero J, Diez I, Cortes JM, Marinazzo D, Stramaglia S (2019): Connectome sorting by consensus clustering increases separability in group neuroimaging studies. Netw Neurosci 3:325–343.

45. Diez I, Sepulcre J (2018): Neurogenetic profiles delineate large-scale connectivity dynamics of the human brain. Nat Commun 9:3876.

46. Sepulcre J, Grothe MJ, d'Oleire Uquillas F, Ortiz-Terán L, Diez I, Yang HS, et al. (2018): Neurogenetic contributions to amyloid beta and tau spreading in the human cortex. Nat Med 24:1910–1918.

47. Bueichekú E, Aznárez-Sanado M, Diez I, d'Oleire Uquillas F, Ortiz-Terán L, Qureshi AY, et al. (2020): Central neurogenetic signatures of the visuomotor integration system. Proc Natl Acad Sci USA 117:6836–6843.

48. Ritchie J, Pantazatos SP, French L (2018): Transcriptomic characterization of MRI contrast with focus on the T1-w/T2-w ratio in the cerebral cortex. Neuroimage 174:504–517.

49. Burt JB, Demirtaş M, Eckner WJ, Navejar NM, Ji JL, Martin WJ, et al. (2018): Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. Nat Neurosci 21:1251–1259.

50. Fornito A, Arnatkevičiūtė A, Fulcher BD (2019): Bridging the gap between connectome and transcriptome. Trends Cogn Sci 23:34–50.

51. Romero-Garcia R, Whitaker KJ, Váša F, Seidlitz J, Shinn M, Fonagy P, et al. (2018): Structural covariance networks are coupled to expression of genes enriched in supragranular layers of the human cortex. Neuroimage 171:256–267.

52. Jimenez-Marin A, Diez I, Labayru G, Sistiaga A, Caballero MC, Andres-Benito P, et al. (2021): Transcriptional signatures of synaptic vesicle genes define myotonic dystrophy type I neurodegeneration. Neuropathol Appl Neurobiol 47:1092–1108.

53. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, et al. (2012): An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 489:391–399.

54. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, et al. (2014): The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry 19:659–667.

55. Fortin JP, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, et al. (2018): Harmonization of cortical thickness measurements across scanners and sites. Neuroimage 167:104–120.

56. Johnson WE, Li C, Rabinovic A (2007): Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8:118–127.

57. Wachinger C, Rieckmann A, Pölsterl S (2021): Detect and correct bias in multi-site neuroimaging datasets. Med Image Anal 67:101879.

58. Yu M, Linn KA, Cook PA, Phillips ML, McInnis M, Fava M, et al. (2018): Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. Hum Brain Mapp 39:4213–4227.

59. Di Martino A, O'Connor D, Chen B, Alaerts K, Anderson JS, Assaf M, et al. (2017): Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. Sci Data 4:170010.

60. Hennig C (2007): Cluster-wise assessment of cluster stability. Comp Stat Data Anal 52:258–271.

61. Fischer MM, Getis A, editors. (2010). Handbook of Applied Spatial Analysis. Berlin, Heidelberg: Springer.

62. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B (2019): WebGestalt 2019: Gene set analysis toolkit with revamped UIs and APIs. Nucleic Acids Res 47:W199–205.

63. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. (2000): Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25–29.

64. Fabregat A, Sidiropoulos K, Viteri G, Forner O, Marin-Garcia P, Arnau V, et al. (2017): Reactome pathway analysis: A high-performance in-memory approach. BMC Bioinformatics 18:142.

65. Fulcher BD, Arnatkeviciute A, Fornito A (2021): Overcoming false-positive gene-category enrichment in the analysis of spatially resolved transcriptomic brain atlas data. Nat Commun 12:2669.

66. Burt JB, Helmer M, Shinn M, Anticevic A, Murray JD (2020): Generative modeling of brain maps with spatial autocorrelation. Neuroimage 220:117038.

67. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. (2021): The STRING database in 2021: Customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 49:D605–D612.

68. Moreau CA, Kumar K, Harvey A, Huguet G, Urchs SGW, Schultz LM, et al. (2023): Brain functional connectivity mirrors genetic pleiotropy in psychiatric conditions. Brain 146:1686–1696.

69. Holiga Š, Hipp JF, Chatham CH, Garces P, Spooren W, D'Ardhuy XL, et al. (2019): Patients with autism spectrum disorders display reproducible functional connectivity alterations. Sci Transl Med 11:eaat9223.

70. Cellot G, Cherubini E (2014): Reduced inhibitory gate in the barrel cortex of Neuroligin3R451C knock-in mice, an animal model of autism spectrum disorders. Physiol Rep 2:e12077.

71. Schmeisser MJ, Ey E, Wegener S, Bockmann J, Stempel AV, Kuebler A, et al. (2012): Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. Nature 486:256–260.

72. He Y, Byrge L, Kennedy DP (2020): Nonreplication of functional connectivity differences in autism spectrum disorder across multiple sites and denoising strategies. Hum Brain Mapp 41:1334–1350.

73. Kabbara A, Khalil M, O'Neill G, Dujardin K, El Traboulsi Y, Wendling F, Hassan M (2019): Detecting modular brain states in rest and task. Netw Neurosci 3:878–901.

74. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D (2019): Benefits and limitations of genome-wide association studies. Nat Rev Genet 20:467–484.

75. Asif M, Martiniano HFMC, Marques AR, Santos JX, Vilela J, Rasga C, et al. (2020): Identification of biological mechanisms underlying a multidimensional ASD phenotype using machine learning [no. 1]. Transl Psychiatry 10:43.

76. Yousaf A, Waltes R, Haslinger D, Klauck SM, Duketis E, Sachse M, et al. (2020): Quantitative genome-wide association study of six phenotypic subdomains identifies novel genome-wide significant variants in autism spectrum disorder. Transl Psychiatry 10:215.

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

78. Gilman SR, Iossifov I, Levy D, Ronemus M, Wigler M, Vitkup D (2011): Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. Neuron 70:898–907.

79. Tsai NP, Wilkerson JR, Guo W, Maksimova MA, DeMartino GN, Cowan CW, Huber KM (2012): Multiple autism-linked genes mediate synapse elimination via proteasomal degradation of a synaptic scaffold PSD-95. Cell 151:1581–1594.

80. Mariner R, Jackson AW, Levitas A, Hagerman RJ, Braden M, McBogg PM, et al. (1986): Autism, mental retardation, and chromosomal abnormalities. J Autism Dev Disord 16:425–440.

81. Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, et al. (1999): A genomic screen of autism: Evidence for a multilocus etiology. Am J Hum Genet 65:493–507.

82. Penzes P, Cahill ME, Jones KA, VanLeeuwen JE, Woolfrey KM (2011): Dendritic spine pathology in neuropsychiatric disorders. Nat Neurosci 14:285–293.