Broad transcriptomic dysregulation across the cerebral cortex in ASD
— **Source link** ↗

Jillian R. Haney, Brie Wamsley, George T. Chen, Sepideh Parhami ...+19 more authors

Institutions: University of California, Los Angeles, Yale University, University of Wisconsin-Madison, University of California, Berkeley

Topics: Cerebral cortex, Visual cortex and Sensory system

Related papers:

- Global changes in patterning, splicing and primate specific lncRNAs in autism brain

- Elevated gene expression of most microglial markers, and reduced expression of most pyramidal neuron and interneuron markers, in postmortem autism cortex

- Association between resting-state functional brain connectivity and gene expression is altered in autism spectrum disorder

- Genome-wide changes in lncRNA, alternative splicing, and cortical patterning in autism

- Human brain organoids reveal accelerated development of cortical neuron classes as a shared feature of autism risk genes

Share this paper: 🟦 🐦 in ✉

View more about this paper here: https://typeset.io/papers/broad-transcriptomic-dysregulation-across-the-cerebral-247k322t51

1 **Title Page**

2

3 **Title:**

4

5 Broad transcriptomic dysregulation across the cerebral cortex in ASD.

6

7 **Authors**:

8

9 Jillian R. Haney[1,2,3], Brie Wamsley[1,2,3,4], George T. Chen[2], Sepideh Parhami[1,2,3], Prashant S. Emani[8], Nathan

10 Chang[8], Gil D. Hoftman[1,3], Diego de Alba[3], Gaurav Kale[3], Gokul Ramaswami[1,2], Christopher L. Hartl[1,2], Ting

11 Jin[9], Daifeng Wang[9], Jing Ou[2], Ye Emily Wu[1,2], Neelroop N. Parikshak[2,4], Vivek Swarup[2,6], T. Grant Belgard[10],

12 Mark Gerstein[8], Bogdan Pasaniuc[4,7], Michael J. Gandal[1,2,3,4,*] & Daniel H. Geschwind[1,2,3,4,*]

13

14 **Affiliations**:

15

16 1. Program in Neurobehavioral Genetics, Semel Institute, David Geffen School of Medicine, University of
17   California, Los Angeles, Los Angeles, CA 90095, USA.
18 2. Department of Neurology, Center for Autism Research and Treatment, Semel Institute, David Geffen School of
19   Medicine, University of California, Los Angeles, 695 Charles E. Young Drive South, Los Angeles, CA 90095,
20   USA.
21 3. Department of Psychiatry, Semel Institute, David Geffen School of Medicine, University of California, Los
22   Angeles, 695 Charles E. Young Drive South, Los Angeles, CA 90095, USA.
23 4. Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los
24   Angeles, CA 90095, USA.
25 5. Department of Neurology, School of Medicine, University of California, San Francisco, CA 94143, USA.
26 6. Institute for Memory Impairments and Neurological Disorders, University of California, Irvine, CA 92697, USA.
27 7. Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California,
28   Los Angeles, 90024, USA.
29 8. Computational Biology & Bioinformatics Program, Yale University, New Haven, CT 06520, USA.
30 9. Waisman Center and Department of Biostatistics and Medical Informatics, University of Wisconsin – Madison,
31   Madison, WI, 53705, USA.
32 10. The Bioinformatics CRO, 203 Muirfield Cove West, Niceville, FL 32578, USA.
33 * Correspondence to: mgandal@mednet.ucla.edu and dhg@mednet.ucla.edu

1 # Abstract

2 Classically, psychiatric disorders have been considered to lack defining pathology, but recent work has
3 demonstrated consistent disruption at the molecular level, characterized by transcriptomic and epigenetic
4 alterations.[1–3] In ASD, upregulation of microglial, astrocyte, and immune signaling genes, downregulation of
5 specific synaptic genes, and attenuation of regional gene expression differences are observed.[1,2,4–6] However,
6 whether these changes are limited to the cortical association areas profiled is unknown. Here, we perform RNA-
7 sequencing (RNA-seq) on 725 brain samples spanning 11 distinct cortical areas in 112 ASD cases and
8 neurotypical controls. We identify substantially more genes and isoforms that differentiate ASD from controls
9 than previously observed. These alterations are pervasive and cortex-wide, but vary in magnitude across
10 regions, roughly showing an anterior to posterior gradient, with the strongest signal in visual cortex, followed by
11 parietal cortex and the temporal lobe. We find a notable enrichment of ASD genetic risk variants among cortex-
12 wide downregulated synaptic plasticity genes and upregulated protein folding gene isoforms. Finally, using
13 snRNA-seq, we determine that regional variation in the magnitude of transcriptomic dysregulation reflects
14 changes in cellular proportion and cell-type-specific gene expression, particularly impacting L3/4 excitatory
15 neurons. These results highlight widespread, genetically-driven neuronal dysfunction as a major component of
16 ASD pathology in the cerebral cortex, extending beyond association cortices to involve primary sensory regions.

17

1 **Main Text**

2 *Transcriptomic changes across the cerebral cortex in ASD*
3
4 Similar to other neuropsychiatric disorders, the risk for autism spectrum disorder (ASD) involves
5 substantial genetic liability, which is profoundly complex and heterogeneous.[7,8] Despite this causal
6 heterogeneity, molecular profiling studies consistently show common patterns of shared transcriptomic and
7 epigenetic dysregulation in the majority of ASD cases in frontal and temporal association cortex.[1–3,5] Whether
8 this represents focal, regional, or more generalized dysfunction is not known. To address this question cortex-
9 wide, we conducted strand-specific RNA-sequencing (RNA-seq) to identify gene and isoform (transcriptomic)
10 changes in 725 samples across 11 brain regions spanning all four cortical lobules (frontal, parietal, temporal,
11 and occipital), from 49 subjects with idiopathic ASD and 54 matched neurotypical controls (**Fig. 1a**, **Methods**,
12 **Supplementary Table 1**, and **Extended Data Fig. 1-3**). Previous work using gene expression microarrays and
13 RNA-seq identified gene co-expression modules representing specific pathways differentially expressed in ASD
14 frontal and temporal cortices.[4,5] The number of samples profiled here is more than five times greater than these
15 prior studies, so we first used this multi-region RNA-seq resource to replicate and extend these previous findings.
16 We observed widespread dysregulation across all 11 cortical regions that replicated the previously identified
17 patterns of dysregulation in temporal and frontal cortex (**Fig. 1b**, **Methods**, **Supplementary Table 2**, **Extended**
18 **Data Fig. 3**). However, the magnitude of effect varied across regions, with the primary visual cortex (V1;
19 Brodmann Area (BA) 17) exhibiting the greatest degree of dysregulation, followed by parietal cortex (BA7) and
20 posterior superior temporal gyrus (BA 41/42/22) both in terms of fold changes and the number of genes
21 differentially expressed (**Fig. 1b**, **Extended Data Fig. 3**). To show that this was not due to regional variation in
22 sample sizes, we performed permutation testing, which indicated that this increased signal was not biased by
23 regional sample size differences (**Methods**, **Supplementary Table 2**).
24 Given that qualitatively similar transcriptomic changes were observed across regions (**Fig. 1b**), we next
25 combined all regions to increase our statistical power to detect previously unrecognized differentially expressed
26 (DE) genes and isoforms. We used a linear mixed model framework to control for individual effects and identify
27 changes across all 11 regions examined as well as within individual regions, separately (**Fig. 1c**, **Methods**,
28 **Supplementary Table 3**, **Extended Data Fig 4**). We found 4,223 genes and 9,474 isoforms (FDR < 0.05) DE
29 across all cortical regions, a notable increase compared to previous analyses (**Fig. 1c**, **Extended Data Fig. 3**).
30 We again observed the greatest signal in BA17, and 59% of DE genes in BA17 alone overlapped with what was
31 observed globally (**Supplementary Table 2**, **Extended Data Fig. 4**). Additionally, DE gene effect sizes in BA17
32 and BA7 were the highest in magnitude, more than other regions assessed (**Fig. 1d**, **Methods**). In comparing
33 DE genes and isoforms across all regions, we found both conserved and distinct dysregulation (**Extended Data**
34 **Fig. 4**, **Supplementary Table 2**). Notably, as previously observed in frontal and temporal cortex[1] we observed
35 that DE isoforms exhibited greater effect size changes in ASD than their matched genes (**Fig. 1e**,
36 **Supplementary Table 2**, **Extended Data Fig. 4**).
37 We next evaluated differential gene and isoform expression in an additional 83 pan-cortical samples from
38 9 subjects with dup15q syndrome, a rare genetic disorder with high penetrance for ASD, which previously was
39 shown to strongly parallel changes in idiopathic ASD in frontal and temporal cortex, but with greater magnitude
40 of effect.[5] We replicated these previous results broadly across the cortical regions examined, finding substantial
41 overlap in transcriptomic changes between dup15q and idiopathic ASD and with dup15q exhibiting a greater
42 magnitude of dysregulation overall (**Fig. 1f**, **Supplementary Table 2**, **Extended Data Fig. 4**). BA17 also
43 exhibited the greatest number of DE genes in dup15q (**Extended Data Fig. 4**). These results demonstrate that
44 the molecular pathology shared by this genetic form of ASD and idiopathic ASD is widespread across distinct
45 regions of the cortex, and that some commonalities in regional variance of effect exist, both impacting sensory
46 in addition to association cortex.
47

1
Broad attenuation of transcriptomic regional identity

3

4 　　We previously observed an attenuation of typical gene expression differences between two regions,
5 frontal and temporal lobe in ASD,[4,5] which we refer to here as an "Attenuation of Transcriptomic Regional Identity"
6 (ARI). To assess whether this was a broader phenomenon, we systematically contrasted all unique pairs of 11
7 cortical regions (55 comparisons in all) using a conservative statistical approach to account for differences in
8 sample size across regions, while correcting stringently for multiple comparisons (**Fig. 2a**, **Methods**). We further
9 validated the identified transcriptomic regional identity patterns in our control samples with those from an external
10 data source, the Allen Brain Atlas[9] (**Supplementary Table 4**, **Extended Data Fig. 5**, **Methods**). Ten pairs of
11 regions exhibited significantly greater ARI patterns in ASD compared to controls, with an additional 31 out of the
12 55 pairs of regions exhibiting a trend towards attenuation in ASD (**Fig. 2b**, **Supplementary Table 4**, **Extended**
13 **Data Fig. 5**, **Methods**). These results provide evidence in support of widespread ARI across the cerebral cortex
14 in ASD for the first time, across both gene and isoform levels (**Extended Data Fig. 5**). Additionally, we observed
15 a regional anterior - posterior gradient, with nine of the ten region pairs exhibiting significant ARI in ASD
16 containing either BA17 or BA39-40 (**Fig. 2c-d**). Notably, BA17 was also one of the regions with the largest case-
17 control differences in gene expression. To determine how gene expression changes were dispersed across
18 regions in these pairs, we used a conservative filtering process to identify individual genes exhibiting ARI
19 (**Methods**, **Supplementary Table 4**). Although these genes were widely dysregulated, the posterior regions
20 BA17 and BA39-40 exhibited the greatest changes (**Fig. 2c-d**, **Extended Data Fig. 6**). ARI genes were also
21 comparably disrupted in the dup15q samples (**Extended Data Fig. 6**), suggesting that transcriptomic regional
22 identity attenuation in the cerebral cortex is shared across heterogenous forms of ASD.
23 　　To identify the biological processes contributing to ARI gene dysregulation in ASD, we grouped together
24 all of the ARI genes that were either downregulated (1,881 genes) or upregulated (1,695 genes) with a
25 pronounced posterior effect in ASD (**Methods**). The downregulated set of ARI genes showed broad enrichment
26 for neuronal cell-type-specific markers and RNA processing pathways, and contained many transcription factors
27 (**Fig. 2c**, **Supplementary Table 4**). The upregulated ARI genes also contained many transcription factors and
28 were enriched for oligodendrocyte progenitor cell (OPC) and astrocyte cell-type markers along with metabolic
29 and development pathways. ARI gene dysregulation was further characterized by subsequent co-expression
30 network analysis, which further refined the topology and pathways involved.

31

Refining disrupted gene co-expression networks in ASD

33

34 　　We next used weighted gene correlation network analysis (WGCNA)[10] across all samples to partition
35 genes into co-expression modules capturing potentially shared biological functions or regulation (**Methods**). We
36 identified a total of 35 gene modules, of which 9 were downregulated and 15 were upregulated in ASD
37 (**Supplementary Table 5-6**, **Extended Data Fig. 7**). We further generated networks using isoform-level
38 quantifications, identifying 61 isoform modules. Of these, 39 were distinct from the gene modules, with 5
39 downregulated and 9 upregulated in ASD (**Supplementary Table 5-6**, **Extended Data Fig. 8**). In total, 38 gene
40 and isoform modules were dysregulated in at least one region in ASD. Most of these fell into two broad groups
41 - either dysregulated (1) cortex-wide with comparable magnitude across regions, or (2) with significantly variable
42 magnitude across regions. Again, dup15q effects were similar to ASD effects, but were greater in magnitude
43 (**Extended Data Fig. 7-8**, **Supplementary Table 6**).

44 Cortex-wide dysregulation observed for ASD risk genes

45 　　Eighteen gene and isoform modules exhibited a consistent pattern of dysregulation in ASD across all
46 cortical regions assessed (linear mixed model, FDR < 0.05; **Fig. 3a**, **Extended Data Fig. 7-8**, **Supplementary**
47 **Table 6**). These include GeneM9, an upregulated neuronal module with a significant enrichment for non-coding

genes; GeneM32, a strongly upregulated reactive astrocyte module with the greatest overall magnitude of dysregulation; and GeneM24, a downregulated module enriched for endothelial and pericyte marker genes which are involved in blood-brain-barrier functions (**Fig. 3b, Extended Data Fig. 7, Supplementary Table 6**). These modules replicate previous findings of neuronal upregulation, astrocyte reactivity, and BBB disruption in ASD,[1,4–6] but extend these findings by demonstrating that these processes are widespread across the cerebral cortex.

Two modules - GeneM5 and IsoM37 - demonstrated cortex-wide dysregulation along with significant enrichment for ASD-associated common genetic variation (**Fig. 3b-d**).[11] GeneM5 is down-regulated in ASD, contains many neuronal genes involved in synaptic plasticity, and significantly overlaps with the synaptic module CTX.M16 previously identified by Parikshak et al.[5] (**Fig. 3a**, **Fig. 3d**, **Supplementary Table 5-6**). In addiiton to common genetic variation, GeneM5 is also significantly enriched for genes containing rare *de novo* protein disrupting mutations associated with ASD, including the high-confidence risk genes *GRIN2A*, *MYO5A*, and *BTRC*[12] (**Supplementary Table 5-6**, **Methods**). This demonstrates another point of convergence of rare and common risk variants on shared biological processes in ASD.[39] GeneM5 is enriched in cortical lower layer 4-6 excitatory neuron cell-type markers (**Extended Data Fig. 7**),[13] identifying them as a point of convergence for rare and common genetic risk in ASD. Finally, IsoM37 is enriched for ASD common genetic risk variants (but not rare mutations), is upregulated in ASD, and contains genes involved in protein folding (**Fig. 3a**, **Fig. 3c**, **Supplementary Table 6**). To our knowledge, this is the first report of an upregulated ASD transcriptomic signature that is associated with known ASD risk variants.

Magnitude of effect parallels anterior-posterior gradients

In addition to observing profound cortex-wide dysregulation in ASD, we found 13 modules that exhibited their most pronounced ASD effect in BA17, as measured against a permuted distribution containing all regions (**Extended Data Fig. 7**, **Supplementary Table 6**, **Methods**). Of these, 12 showed significant enrichment for ARI genes (half up-regulated and half down-regulated in ASD) and all 13 had anterior - posterior gradients of expression in neurotypical samples, indicating that these modules contribute to transcriptomic regional identities that are observed in neurotypical controls, but attenuated in ASD (**Extended Data Fig. 7**, **Supplementary Table 6**). Six of these modules were more highly expressed in posterior regions in neurotypical subjects and were observed to be downregulated in ASD across the cortex (**Fig. 4a**, **Extended Data Fig. 7**, **Supplementary Table 6**). These include GeneM23, an oligodendrocyte-specific module consisting of genes important for organelle regulation and intracellular restructuring; GeneM14, a neuronal module that contains genes involved in neurite morphogenesis and is also strongly downregulated in BA41-42-22; and GeneM3, a neuronal module enriched for energy generation and neuronal processes that are highly energy dependent, such as vesicle transport (**Fig. 4b-c**, **Supplementary Table 5-6**). GeneM3 is also significantly enriched for cell-type markers specific to layer 4-5 excitatory neurons (**Extended Data Fig. 7**).[13] The next four modules were more highly expressed in anterior regions in neurotypical subjects and exhibited cortex-wide upregulation in ASD (**Fig. 4a**, **Extended Data Fig. 7**, **Supplementary Table 6**). These include GeneM8, a microglial module containing genes involved in immune signaling and phagocytosis; and GeneM7, an immune response module containing genes such as NF-kB and interferon response pathways (**Fig. 4b**, **Supplementary Table 5-6**). Although neuronal and oligodendrocyte downregulation along with immune and microglia upregulation have been previously reported in ASD,[1,4–6] these findings indicate that this dysregulation is widespread across the cerebral cortex, with increased magnitude in posterior regions, a pattern most pronounced in BA17.

The last three modules were observed to be significantly upregulated in ASD only in BA17. One of these, GeneM4, is an inhibitory neuron module containing many genes important for various intracellular signaling and maturation processes, such as *SCN9A* (**Fig. 4a-c**, **Supplementary Table 5-6**). Additionally, GeneM4 is significantly enriched for lincRNAs and for previously reported gene modules associated with upregulated pathways related to development[5] and signaling[1,6] in ASD, although we observe this effect in BA17 for the first time (**Extended Data Fig. 7**). We also identified four other modules exhibiting strong region-specific dysregulation in regions other than BA17 (**Extended Data Fig. 7**). For example, the module GeneM34, which

1 contains genes involved in cellular stress response regulatory processes, is upregulated with the greatest
2 magnitude in BA4-6 and shows no significant effect in BA17 (**Extended Data Fig. 7**, **Supplementary Table 5-**
3 **6**). None of the gene modules with regionally-variable magnitudes of ASD effect were significantly enriched for
4 known ASD genetic risk variants.

5 Cell-type changes mirror regional variation

6 We finally sought to determine what might be driving the observed changes in magnitude of ASD effect
7 across regions. It is well established that BA17 is the most neuronally dense region in the human brain, with a
8 notable expansion in the thickness of L3/4, compared with other cortical regions.[18] Likewise, there is an anterior-
9 posterior gradient in neuronal density observed in mice and primates.[14–17] As such, we posited that regional
10 variation in cell density could be contributing to regional differences in magnitude of ASD effect. Regional
11 neuronal density across multiple brain regions has not been quantitatively studied in the human brain, but such
12 gradients have been established across some regions in non-human primates.[15,16] Therefore, we compared the
13 region-specific ASD effect size changes in our gene modules to regional neuronal nuclei density measured in
14 primates[15] for 6 matched regions across species. We observed a significant association between neuronal
15 density and the effect sizes for several modules dysregulated in ASD (seven with FDR < 0.05, and an additional
16 eight with FDR < 0.1, **Extended Data Fig. 9**, **Supplementary Table 7**). Further, L3/4 thickness was also
17 associated with the region-specific ASD effect sizes in dysregulated modules (**Supplementary Table 7**).

18 These observations motivated us to perform single-nucleus RNA sequencing (snRNA-seq) in a small
19 cohort of individuals to help evaluate how distinct neural cell-types could be contributing to the regional variance
20 in ASD transcriptomic dysregulation identified with bulk RNA-seq (**Fig. 4d**, **Extended Data Fig. 9**,
21 **Supplementary Table 7, Methods**). We sequenced over 150,000 nuclei from ASD and control samples across
22 frontal and occipital cortices with matching bulk RNA-seq. From these data, we identified 35 distinct cell clusters
23 and 4,953 cell-type-specific DE genes in ASD subjects in the frontal and occipital cortex. The vast majority of
24 these were DE in excitatory neurons in both regions, and exhibited larger effects overall in the occipital lobe (**Fig.**
25 **4f**). While statistical power limited our ability to detect significant cell-type proportion differences between regions
26 or diagnoses (**Methods**), we do observe that excitatory neurons are increased in proportion by ~5% in BA17
27 across both control and ASD subjects compared to frontal regions (**Extended Data Fig. 9**), corresponding with
28 the primate neuronal density measurements. To predict how cell-type proportions may vary across our entire
29 bulk RNA-seq dataset, next we utilized cell-type markers from our snRNA-seq to perform cell-type deconvolution
30 in all samples (**Methods**). We identified 11 significant cell subtype proportion changes present across six
31 different regions in ASD, characterized by neuronal decreases and astrocyte and microglia increases (**Fig. 4e**,
32 **Extended Data Fig. 9**, **Supplementary Table 7**). We also found many anterior-posterior cell-type proportion
33 gradients in control subjects that are attenuated in ASD (**Extended Data Fig. 9**, **Supplementary Table 7**),
34 mirroring patterns observed with our bulk RNA-seq transcriptomic regional identity analysis.

35 When directly comparing cell-type-specific DE and deconvolved proportional changes in ASD with
36 regional variability in the larger bulk transcriptome sample, we observed a convergent signal within excitatory
37 neurons – in particular, those in L3/4 (ExNeuron4; **Fig. 4e-f**, **Extended Data Fig. 9**). Recapitulating the known
38 increase in thickness of L3/4 in BA17 compared with other cortical regions[18], we observed a significant increase
39 in the estimated proportion of the ExNeuron4_L3/4 cluster in posterior regions, peaking in BA17 (**Fig. 4e;**
40 **Extended Data Fig. 9g).** This regional pattern was significantly attenuated in ASD, with a ~2-fold median
41 reduction in estimated Ex4 neuronal proportion in BA17 compared with controls. This cell cluster, with marker
42 genes *RORB*, *PCP4*, *CUX2*, *PHACTR2*, and *EYA4*, also exhibited substantially greater cell-type-specific DE in
43 snRNA-seq profiling from BA17 compared with frontal cortex (90 vs 0 DE genes, respectively; **Extended Data**
44 **Fig. 9d**). Similarly, BA17 shows a substantially greater upregulation of inhibitory neuron genes in the single cell
45 data, consistent with the observed greater up-regulation of GeneM4 (inhibitory neuron) in BA17 (**Fig. 4f**).
46 Substantial changes in gene expression are also evident in other cell subtypes (**Extended Data Fig. 9,**
47 **Supplementary Table 7**), such as microglia_2, which shows a strong and specific increase in DE genes in ASD

BA17 compared to frontal regions. These observed intracellular/cell-type changes in neuronal and microglial gene expression are further supported by another snRNA-seq dataset containing a small ASD cohort, which assessed a single region.[19] Here, through performing multi-region snRNA-seq and cell-type deconvolution, we show that predicted cell-type proportions as well as cell-type-specific gene expression profiles are impacted across the ASD cerebral cortex. Importantly, we see increased cell-type-specific transcriptomic dysregulation and lowered neuronal proportions with a notable convergence within L3/4 excitatory neurons in ASD BA17, a region where neuronal proportions are neurotypically abundant. These changes likely contribute to the pronounced ASD effect we observe with bulk RNA-seq in this region.

Discussion

Overall, the findings presented here substantially expand our understanding of ASD pathology beyond the previously established 'downregulated neuron' and 'upregulated glia/immune' functional categories observed in frontal and temporal lobe. We identify gene and isoform expression changes in ASD that extend across the cerebral cortex, many neural cell-types, and specific biological processes (**Extended Data Fig. 10**), including primary sensory areas in addition to association areas.[1,4–6] We find that the recently observed reactive astrocyte upregulation and blood-brain barrier membrane transport downregulation[1] is extended cortex-wide in ASD. It is interesting to speculate that the substantial changes observed in area 17, a primary sensory region, may be related to the widespread observation of sensory hypersensitivity or processing abnormalities in ASD.[40] Nevertheless, this region shows the most profound changes in gene expression, a clear demonstration that these alterations are not specific to higher association areas.

Furthermore, we find that other dysregulated pathways observed before in ASD - particularly upregulated immune response and reactive microglia genes, along with downregulated neurite morphogenesis and neuronal energy pathway genes - are not only impacted cortex-wide in ASD, but impacted in a regional gradient that reflects fundamental elements of cortical cytoarchitecture, such as neuronal density. It is also notable that the magnitude of region-level differences in ASD parallels regional variance in attenuation of transcriptomic identity (ARI gene dysregulation), suggesting that they reflect related processes. That the gradient of region-specific changes between ASD and controls coincides with both neuronal proportion differences and cell-type-specific transcriptomic dysregulation further suggests that the interplay of cytoarchitecture and cell-type gene expression, rather than a single one of these features, influences our ability to observe transcriptomic changes in bulk tissue. Given the connection between regional cytoarchitecture, local circuits and long-range brain connectivity,[20,21] parsimony suggests that in addition to developmental patterning contributions,[5,22] the diminution of transcriptomic regional identity reflects changes in local neuronal circuit dysfunction and deficits in synaptic plasticity and homeostasis that are widely propagated.[20] This is supported by our observation that the gene co-expression module representing synaptic plasticity genes is downregulated cortex-wide and is significantly enriched with common and rare ASD genetic risk variants, further emphasizing that synaptic plasticity is a convergent pathway in ASD. Given this result, along with our observations of profound neuronal dysregulation present throughout the ASD cortex, future work should determine which specific aspects of synaptic plasticity may contribute to causal mechanisms in the disorder across specific brain regions and developmental timepoints.

Several additional factors should guide the interpretation of these results. The samples utilized in this work were obtained from heterogeneous postmortem cortical tissue, meaning that the results reported here are broadly applicable to the postnatal ASD cortex across both sexes and a span of ages from two to 68 years old, and they should be interpreted in this context. Rigorous methodology was utilized at every step to account for biological and technical variability, ensuring that the results reported here are conservative and widely applicable. Additionally, bulk tissue RNA-seq, in contrast to single cell and nucleus RNA-seq, does not have the cellular resolution to assess dissection variability across cortical regions and cell-type specificity of transcriptomic changes. We addressed this by performing snRNA-seq, which significantly enhanced our understanding of regional variation in ASD transcriptomic dysregulation. However, snRNA-seq also has its own limitations. While

1   snRNA-seq can profile tens of thousands of cells, snRNA-seq experiments typically have fewer unique samples
2   than bulk RNA-seq experiments, and the comparability of snRNA-seq cell-type proportions to true sample cell-
3   type proportions is currently unclear.[23] It is also challenging to estimate isoform quantifications with single cell
4   RNA-seq approaches, whereas this remains a strength of bulk tissue RNA-seq.[24] Leveraging this, we
5   subsequently identified an upregulated isoform-specific co-expression module enriched with ASD GWAS
6   variants, implicating increased protein folding dysfunction for the first time as a putative pathway contributing to
7   ASD causal mechanisms. Interestingly, upregulated proteostasis is also implicated in Down's Syndrome,[25,26]
8   supporting that protein folding machinery may be an affected biological process in multiple neurodevelopmental
9   disorders. The utilization of methods that have greater cellular resolution is necessary for the improved and
10   continued mapping of the results presented here to specific cortical cell-types. As we seek to gain a complete
11   understanding of ASD neural pathology, future approaches which integrate different sources of biological data -
12   including this cortex-wide transcriptomic resource - to determine how ASD risk genes are acting in the brain will
13   be essential.
14

Methods

Sample Acquisition and Preparation for RNA-seq

Postmortem cortical brain samples were acquired from the Harvard Brain Bank as part of the Autism BrainNet project (formerly the Autism Tissue Project, ATP) and the University of Maryland Brain Banks (UMDB). A total of 842 samples from subjects with ASD, dup15q syndrome, and non-psychiatric controls (112 unique subjects) across 11 cortical regions encompassing all major cortical lobes – frontal: BA4/6, BA9, BA44/45, BA24; temporal: BA38, BA41/42/22, BA20/37; parietal: BA3/1/2/5, BA7, BA39/40; and occipital, BA17 - were acquired. These included 253 samples previously published in Parikshak et al., Nature 2016[5] from BA9 and BA41/42/22 and/or Gandal et al., Science 2018b[1,5] from BA9, BA4/6, and BA41/42/22. An ASD diagnosis was confirmed by the Autism Diagnostic Interview-Revised (ADIR) in 30 of the subjects. In the remaining 19 subjects, diagnosis was supported by clinical history. Frozen brain samples were stored at -80 deg C. To extract RNA from these samples, first approximately 50-100mg of tissue were dissected from the cortical regions of interest on dry ice in a dehydrated dissection chamber to reduce degradation effects from sample thawing or humidity. Then, RNA was isolated from each sample using the miRNeasy kit with no modifications (Qiagen). For each RNA sample, RNA quality was quantified using the RNA Integrity Number (RIN) on an Agilent Bioanalyzer.

RNA-seq and RNA Data Processing

Initial sequencing in BA9 and BA41/42/22 was performed in three batches as published by Parikshak et al., Nature 2016.[5] The remaining regions, along with additional BA9 and BA41/42/22 samples, were sequenced across three new batches. For all of these batches, strand-specific RNA-seq libraries were prepared. For the first two batches, the TruSeq Stranded Total RNA sample prep kit with RiboZero Gold (Illumina) was used to obtain rRNA-depleted libraries. The remaining batch was prepared with the TruSeq RNA Exome sample prep kit (formerly the TruSeq RNA Access sample prep kit; Illumina). All libraries were randomly pooled to multiplex 24 samples per lane using Illumina TruSeq barcodes. Each lane was sequenced five times on an Illumina HiSeq 2500 or 4000 instrument using high output mode with standard chemistry and protocols for 50, 69, or 100 bp paired-end reads (read length varied by batch) to achieve a target depth of 70 million reads.

After sequencing, the resulting sample FASTQ files from all batches (including the Parikshak et al.[5] samples) were subjected to the same processing pipeline. First, FASTQ files were assessed with FastQC[27] (v0.11.2) to verify that quality was sufficient for further processing. FASTQ files were then aligned to the human reference genome (GRCh37[28] Ensembl v75) with STAR[29] (v2.5.2b). Picard tools[30] (v2.5.0) was used with the resulting BAM files to collect various read quality measures, in addition to the quality measures collected by STAR. verifyBAMID[31] was also used with these BAM files along with known sample genotypes from Parikshak et al.[5] to validate that sample identity was correct for all BAM files. Additionally, the expression of XIST (a female-specific gene) was assessed to contribute to sample identity verification. Finally, RSEM[32] (v1.3.0) was used for quantification (Gencode[33] release 25lift37) to obtain expected read counts at the gene and isoform levels.

Expected gene and isoform read counts were then subjected to several processing steps in preparation for downstream analysis, mainly using R.[34] First, Counts Per Million (CPM) were obtained from counts for gene and isoform filtering purposes. Genes and isoforms were filtered such that genes/isoforms with a CPM > 0.1 in at least 30% of samples were retained. Genes/isoforms were also removed which had an effective length (measured by RSEM) of less than 15 bp. Isoforms were additionally filtered such that all isoforms corresponded with genes in the gene-level analysis. The counts for the remaining genes (24,836) and isoforms (99,819) passing these filters were normalized using the limma-trend approach in the limma[35] R package.

1 Briefly, the limma-trend approach obtains normalized expression data through taking the $\log_2$(CPM) of read
2 counts with an adjustment for sample read depth variance. An offset value calculated with CQN[36] accounting
3 for GC content bias and gene/isoform effective length bias in read quantification was also incorporated during
4 the normalization process. With this normalized expression data, sample outliers were identified in each
5 sequencing batch by cortical lobe (frontal, parietal, temporal, and occipital) group that had both (1) an absolute
6 z-score greater than 3 for any of the top 10 expression principal components (PCs) and (2) a sample
7 connectivity score less than -2. Sample connectivity was calculated using the fundamentalNetworkConcepts
8 function in the WGCNA[10] R package, with the signed adjacency matrix (soft power of 2) of the sample biweight
9 midcorrelation as input. This process identified 34 outliers, resulting in a final total of 808 samples
10 (341=Control, 384=ASD, 83=dup15q) which were carried forward for analysis.

11
12 **Evaluating Previous Co-Expression Modules and ASD DE Genes/Isoforms Cortex-wide**
13
14 Linear models for all subsequent analyses are described in the Supplementary Methods.
15
16 To determine how gene co-expression modules previously identified in Parikshak et al.[5,35] and Voineagu et al.[4]
17 were effected across distinct cortical regions, we first created a regressed gene expression dataset that only
18 contained the effects of biological covariates (subject, diagnosis, region, sequencing batch, sex, ancestry, age,
19 and age^2). This regressed dataset was created with the 'lmerTest'[37] package in R through subtracting the
20 effects of technical covariates from each gene, leaving only the random intercept, biological covariate effects,
21 and the residual. ASD-associated module eigengene region-specific ASD effects were identified using
22 contrasts (eg. Control_BA17 - ASD_BA17) with the limma[34] R package with this regressed expression dataset,
23 accounting for all biological covariates. Region-specific contrasts with a p-value < 0.05 were considered
24 significant (FDR-correction was unwarranted since only eight module eigengenes were examined).
25
26 To identify genes and isoforms dysregulated in ASD both within specific regions and cortex-wide, the limma[35]
27 R package was applied with the gene and isoform expression data using our full gene and isoform models
28 (both biological and technical covariates). The standard limma[35] workflow was implemented as recommended
29 for linear mixed models. Region-specific dysregulation was identified as described above for the Parikshak et
30 al.[5] and Voineagu et al.[4] modules. Whole cortex dysregulation was established through subtracting the sum of
31 the ASD region-specific effects from the sum of the Control region-specific effects. For both region-specific and
32 whole cortex effects, genes and isoforms with an FDR-corrected p-value < 0.05 were considered significantly
33 dysregulated. dup15q region-specific and whole cortex dysregulation was also established in this manner. The
34 fixed effects of sex, age, and age^2 were also acquired (shared in **Supplementary Table 3**) using the full gene
35 and isoform models.
36
37 The methodology used to evaluate region-specific ASD effects compared to whole cortex ASD effects is
38 described in the Supplementary Methods.
39
40 **Transcriptomic Regional Identity Analysis**
41
42 To identify differentially expressed genes and isoforms between all 55 pairs of cortical regions, a regressed
43 gene expression dataset containing only the random effect of subject and the fixed effects of diagnosis and
44 region (along with the model residual) was used. Regression was performed as described for evaluation of
45 previously identified co-expression modules. Significant attenuation of DE genes between each pair of regions
46 (a reduction in transcriptomic regional identity differences) in ASD was established through the following
47 process. (1) ASD and Control subjects containing each region in the regional pair were extracted for use in the
48 analysis. (2) Separately in ASD and Control subjects, the number of DE genes between regions was calculated

10

1 using the paired Wilcoxon signed-rank test. Genes with an FDR-corrected p-value < 0.05 were considered DE.
2 (3) The difference in the number of DE genes between regions for ASD v Control subjects was calculated (the
3 'true' difference). (4) A permuted distribution of the difference in DE genes between regions for ASD v Control
4 subjects was generated to test the 'true' difference. Each permutation (10,000 in total) randomly assigned
5 'ASD' and 'Control' status to subjects, but kept the number of ASD and Control subjects consistent with the
6 true number of ASD and Control subjects. (5) A two-tailed p-value was obtained from testing the 'true'
7 difference against the permuted distribution. If the regional comparison p-value < 0.05, with the number of DE
8 genes between regions in ASD less than that in Controls, then the regional comparison was considered
9 significantly attenuated in ASD. Otherwise, the regional comparison was considered over-patterned in ASD.
10 This procedure was repeated with isoform level regressed gene expression data (similarly, only containing the
11 random effect of subject and the fixed effects of diagnosis and region, along with the model residual) to identify
12 altered transcriptomic identities in ASD at the isoform-level.
13
14 The previously described permutation approach was designed to identify differences in transcriptomic regional
15 identity in ASD. Importantly, this method is not appropriate for assessing variance in expected numbers of DE
16 genes between regions across regional pairs and diagnoses, since the number of ASD and Control subjects
17 varied across regional pairs. To examine this, for each regional comparison we subset to 10 pairs of ASD and
18 Control subjects (10 was selected since every regional comparison had at least this many subjects). When
19 subsetting, subjects were removed such that the remaining subjects were closest in age to the median age of
20 the available samples for that regional comparison. A bootstrap approach was then used to calculate the
21 number of DE genes (p-value < 0.05) between regions separately in Control and ASD subjects through
22 sampling subjects with replacement (mean taken across 10,000 bootstraps). The same regressed expression
23 dataset used for the permutation approach was utilized for this bootstrap analysis. Any regional comparison in
24 which the number of DE genes between regions was less in ASD than in Control subjects was considered
25 trending towards attenuation in ASD.
26
27 To validate our bootstrapped estimates for the number of DE genes between pairs of regions in Controls, we
28 compared these estimates to those of the Allen Brain Atlas[9], which is the best publicly available work for
29 comparison. Allen Brain Atlas regions were matched to Brodmann regions (**Supplementary Table 4**) and
30 matching regional pairs were extracted for comparison with this work. When the Allen Brain Atlas had two or
31 more regional pairs matching one regional pair in this work, the mean was taken across the Allen Brain Atlas
32 regional pairs. A p-value for the association of the number of DE genes between regions in Controls obtained
33 in this work compared to the Allen Brain Atlas was calculated from a linear model (cortex-wide bootstrap mean
34 ~ allen brain atlas mean).
35
36 We applied a stringent filtering process to identify high-confidence attenuated regional identity (ARI) genes
37 from each significantly attenuated regional comparison identified with the permutation procedure described
38 above. First, for each of the attenuated regional comparisons, we extracted the genes which were identified as
39 DE between regions in subjects labeled as Controls in each of the 10,000 permutations. Then, we calculated
40 how many times each of the genes truly DE between pairs of regions in the Control subjects were present in
41 their respective permuted groups (ranging from a possible 0 to 10,000 occurrences). Those 'true' DE genes
42 which were present in less than 95% of their respective permutations were retained as ARI genes for each
43 attenuated regional comparison. For each set of ARI genes (ten total), each gene was matched to the region in
44 which it had higher expression in Control subjects. The paired Wilcoxon signed-rank p-values identified for
45 these genes in Controls (those subjects used for the permutation analysis) were also extracted and are shared
46 in **Supplementary Table 4**.
47

ARI gene groups (ARI downregulated genes, those highly expressed in BA17 and BA39-40 relative to other regions in Controls; ARI upregulated genes, those lowly expressed in BA17 and BA39-40 relative to other regions in Controls) were created through taking the union (without duplicates) across all ten identified ASD-attenuated regional comparisons, and sorting genes into the two groups based on gene expression profiles across regions. The details of this process are described in the Supplementary Methods, along with functional annotation procedures.

Network-Based Functional Characterization

Standard workflows, as previously described in Parikshak et al.[5] and Gandal et al.,[1] were followed (with minor modifications) to identify gene and isoform co-expression modules using Weighted Gene Correlation Network Analysis (WGCNA).[10] Details regarding network formation, module identification, and module functional characterization are described in the Supplementary Methods.

snRNA-seq and Cell-type Deconvolution

Cell types were annotated based on expression of known marker genes visualized on the UMAP plot, violin plots, and by performing unbiased gene marker analysis. To gain insight into the regional enrichment or diagnostic enrichment of cell types, the relative proportion of the number of nuclei in each cell type was normalized to the total number of nuclei captured from each library. Average cell-type proportions and standard errors (across libraries) were scaled such that each Lobule x Diagnosis group sums to 100%, so that cell-type proportions in these groups could be fairly compared across all cell-types. To determine if any changes in cell-type proportion were statistically significant, we implemented scDC[38] to bootstrap proportion estimates for our samples (**Supplementary Table 7**). We employed a linear mixed model (random effect of subject) to determine if any changes in cell-type proportion were present across regions and diagnoses. None of the model covariates were statistically significant ($p > 0.05$ for all model covariates). However, we did find several significantly different predicted cell-type proportions in ASD with cell-type deconvolution analysis. We describe methods for cell-type deconvolution in detail in the Supplementary Methods. To identify genes differentially expressed in ASD compared to control in each cell type, the non-parametric Wilcoxon rank sum test was applied including gene detection rate and sequencing depth within the model. We compared frontal cortex ASD cells to frontal cortex control cells within each cluster and likewise for the occipital cortical cells. The bars in **Figure 4e** are the summation of all differentially expressed genes identified in each cell subtype for the broader cell-type (eg. all excitatory neuron subtype DE genes are summed to obtain the number of DE genes in the broad excitatory neuron cell class). Further details regarding the snRNA-seq analysis are included in the Supplementary Methods.

Data availability

All of the raw bulk RNA-seq data (FASTQ files) and processed (utilized for DE gene analysis, transcriptomic regional identity analysis, and WGCNA) bulk RNA-seq data that support the findings of this study will be deposited in a publicly accessible repository. snRNA-seq data will be made available by the corresponding authors upon reasonable request. All of the code, raw data, and processed data for the bulk RNA-seq analysis that support the findings of this study will also be made available in a publicly accessible GitHub repository, where readers may also access an R Shiny tool to visualize RNA-seq data across the biological covariates assessed in this study.

1 **Code availability**
2
3 All of the code, raw data, and processed data for the bulk RNA-seq analysis that support the findings of this
4 study will also be made available in a publicly accessible GitHub repository, where readers may also access an
5 R Shiny tool to visualize RNA-seq data across the biological covariates assessed in this study.
6
7 **Extended Data Figure Legends**
8
9 **Extended Data Figure 1 | Experiment Workflow and Sample Overview.** a. Overview of experiment workflow.
10 b. Summary of sample composition (biological data, brain bank source, and PMI).
11
12 **Extended Data Figure 2 | Quality Control Measures.** a. Sequencing batch parameters. b. Sequencing batches
13 (top) and sequencing statistics (bottom) by region and diagnosis. c. Top 15 expression PCs (gene and isoform,
14 with % of variance explained denoted) association with meta data (top) and sequencing statistics (bottom).
15
16 **Extended Data Figure 3 | Model Covariates and Previous Studies Across 11 Cortical Regions.** a. For the
17 covariates selected for the gene (left) and isoform (right) linear mixed models, % of expression variance
18 explained across all genes/isoforms. b-c. For the Voineagu et al. and Parikshak et al. studies, b. ASD associated
19 gene module ASD effect (standard error bars and cortical lobes indicated) and c. ASD log2 FC of DE genes
20 identified in these studies, compared to this dataset (Spearman's correlation rho, R, is plotted along with the
21 linear least squares regression best fit line).
22
23 **Extended Data Figure 4 | Transcriptomic Changes Across 11 Cortical Regions.** a. Overlap of Whole-Cortex
24 DE ASD genes and isoforms (blue) with other cortical region DE genes (no color). Regions with no third numeric
25 label on the right completely overlap with the Whole-Cortex DE genes. b. For the Whole-Cortex DE, overlap of
26 genes and isoforms. Regions not shown have no unique DE. c. log2(FC) (top) and standard error (SE, bottom)
27 of the Whole-Cortex ASD DE overlapping and distinct genes and isoforms. d. Overlap in DE ASD and dup15q
28 genes and isoforms. e. For regions with DE ASD genes (left) and isoforms (right), ASD log2(FC) v. dup15q
29 log2(FC) for specific regions (with principal components regression slope, S).
30
31 **Extended Data Figure 5 | Transcriptomic Regional Identity Attenuation in ASD.** a. Mean-centered
32 distribution of 10,000 permutations for the significantly attenuated regional comparisons in ASD. Red bar = true
33 difference in the number of DE genes between controls and ASD. b. Sample size for all regional comparisons.
34 c. Permutation p-values for all regional comparisons. d. For 10,000 regional comparison bootstraps, ratio of DE
35 genes in ASD compared to controls. e. Number of DE genes between pairs of regions in this study (mean across
36 bootstraps in controls, y-axis) compared to the Allen Brain Atlas (ref. 10, mean across matched regions, x-axis;
37 see Methods for matched regions). This Allen Brain Atlas dataset, with only 2 unique brains, is the best publicly
38 available dataset for comparison (linear least squares regression best fit line plotted).
39
40 **Extended Data Figure 6 | Additional ARI gene dysregulation.** a. First principal component (PC1) of posteriorly
41 downregulated (1,881, left) and upregulated (1,695, right) ARI genes identified in ASD, plotted in Controls and
42 dup15q (loess regression line plotted). b. For each significantly attenuated regional comparison, the identified
43 attenuated regional identity (ARI) genes. At center, number of ARI genes with greater neurotypical expression
44 in each pair of regions. On either side of the barplot, the PC1 of the genes with greater neurotypical anterior (left)
45 or posterior (right) expression is plotted across the pair of regions in Controls and ASD. The Wilcoxon signed-
46 rank test (unpaired) p-value is shown.
47
48 **Extended Data Figure 7 | Gene-level Co-Expression Network Analysis Module Associations.** Top:
49 average-linkage hierarchical clustering of module eigengene biweight midcorrelations. Significant FDR corrected
50 p-values are indicated (FDR < 0.05; for GWAS, FDR < 0.1). Any signed –log10(p) colors greater or less than 5/-
51 5 are set at a max/min of 5/-5 . For ASD, dup15q, and Age covariates, FDR p-value from the linear mixed model
52 testing the association of these covariates with module eigengenes is depicted. For the ASD and dup15q region-
53 specific comparisons, cortical lobule colors are indicated (Fig. 1a), and bold-italic FDR p-values indicate that
54 these regions are effected with significantly greater magnitude than the ASD whole-cortex (Methods). For gene

1   biotypes, both positive and negative enrichment is shown (Methods). Positive enrichment is shown for cell-type,
2   neuronal subtype (ref: Hodge et al, Nature 2019), ARI gene, GWAS, and rare variant enrichment (Methods).
3
4   **Extended Data Figure 8 | Isoform-level Co-Expression Network Analysis Module Associations.** Top:
5   average-linkage hierarchical clustering of module eigengene biweight midcorrelations. Significant FDR corrected
6   p-values are indicated (FDR < 0.05; for GWAS, FDR < 0.1). Any signed –log10(p) colors greater or less than 5/-
7   5 are set at a max/min of 5/-5 . For ASD, dup15q, and Age covariates, FDR p-value from the linear mixed model
8   testing the association of these covariates with module eigengenes is depicted. For the ASD and dup15q region-
9   specific comparisons, cortical lobule colors are indicated (Fig. 1a). For gene biotypes, both positive and negative
10 enrichment is shown (Methods). Positive enrichment is shown for cell-type, GWAS, and rare variant enrichments
11 (Methods).
12
13 **Extended Data Figure 9 | Neuronal Density Associations, snRNA-seq, and Cell-type Deconvolution.** a.
14 Macaque neuronal density v. module eigengene ASD effect for modules featured in Fig. 4c-d (linear least
15 squares regression). Both p-value and FDR corrected p-value are plotted. b. P-value histogram of all gene
16 modules' linear least squares regression with macaque region-specific neuronal density. c. UMAP plots of
17 snRNA-seq with cell sub-types (top) and brain regions (bottom) depicted. d. Number of genes differentially
18 expressed in ASD in each cell subtype. Upregulated genes are above 0 (red arrow) and downregulated genes
19 are below 0 (blue arrow). e. Average proportion of each broad cell-type in each diagnosis x cortical lobule group,
20 derived directly from the snRNA-seq data. f. Additional significant (Bonferroni corrected p-value < 0.05) cell-type
21 proportion differences in ASD from cell-type deconvolution. Region and cell-type are indicated in the title of each
22 plot. g. For two example cell-types, cell-type proportion attenuation in ASD across regions. ANOVA p-values
23 stratified by diagnosis are shown.
24
25 **Extended Data Figure 10 | Results Summary.** Overview of RNA-sequencing experiment and results. Region-
26 specific dysregulation scale in the top right corner and the leftmost portion of the bottom panel depict the region-
27 specific slopes compared to the whole cortex effect from Fig 1d. Median PC 1 of the ARI dysregulated genes is
28 plotted in the middle panel. In the right portion of the bottom panel, the median ME of GeneM4 (left) and GeneM3
29 (right) is depicted.
30
31

References

1. Gandal, M. J. *et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* **362**, (2018).

2. Wu, Y. E., Parikshak, N. N., Belgard, T. G. & Geschwind, D. H. Genome-wide, integrative analysis implicates microRNA dysregulation in autism spectrum disorder. *Nat. Neurosci.* **19**, 1463–1476 (2016).

3. Sun, W. *et al.* Histone Acetylome-wide Association Study of Autism Spectrum Disorder. *Cell* **167**, 1385–1397.e11 (2016).

4. Voineagu, I. *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474**, 380–384 (2011).

5. Parikshak, N. N. *et al.* Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature* **540**, 423–427 (2016).

6. Gandal, M. J. *et al.* Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).

7. Sullivan, P. F. & Geschwind, D. H. Defining the Genetic, Genomic, Cellular, and Diagnostic Architectures of Psychiatric Disorders. *Cell* **177**, 162–183 (2019).

8. de la Torre-Ubieta, L., Won, H., Stein, J. L. & Geschwind, D. H. Advancing the understanding of autism disease mechanisms through genetics. *Nat. Med.* **22**, 345–361 (2016).

9. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).

10. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

11. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).

12. Abrahams, B. S. *et al.* SFARI Gene 2.0: a community-driven knowledgebase for the autism

spectrum disorders (ASDs). *Mol. Autism* **4**, 36 (2013).

13. Hodge, R. D. *et al.* Conserved cell types with divergent features in human versus mouse cortex. *Nature* vol. 573 61–68 (2019).

14. Keller, D., Erö, C. & Markram, H. Cell Densities in the Mouse Brain: A Systematic Review. *Frontiers in Neuroanatomy* vol. 12 (2018).

15. Collins, C. E., Airey, D. C., Young, N. A., Leitch, D. B. & Kaas, J. H. Neuron densities vary across and within cortical areas in primates. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15927–15932 (2010).

16. Collins, C. E. *et al.* Cortical cell and neuron density estimates in one chimpanzee hemisphere. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 740–745 (2016).

17. Cadwell, C. R., Bhaduri, A., Mostajo-Radji, M. A., Keefe, M. G. & Nowakowski, T. J. Development and Arealization of the Cerebral Cortex. *Neuron* **103**, 980–1004 (2019).

18. Wagstyl, K. *et al.* BigBrain 3D atlas of cortical layers: Cortical and laminar thickness gradients diverge in sensory and motor cortices. *PLoS Biol.* **18**, e3000678 (2020).

19. Velmeshev, D. et al. Single-cell genomics identifies cell type-specific molecular changes in autism. Science 364, 685–689 (2019).

20. Ecker, C. *et al.* Intrinsic gray-matter connectivity of the brain in adults with autism spectrum disorder. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13222–13227 (2013).

21. Hilgetag, C. C., Beul, S. F., van Albada, S. J. & Goulas, A. An architectonic type principle integrates macroscopic cortico-cortical connections with intrinsic cortical circuits of the primate brain. *Netw Neurosci* **3**, 905–923 (2019).

22. Parikshak, N. N. *et al.* Integrative Functional Genomic Analyses Implicate Specific Molecular Pathways and Circuits in Autism. *Cell* vol. 155 1008–1021 (2013).

23. Patrick, E. *et al.* Deconvolving the contributions of cell-type heterogeneity on cortical gene expression. *PLoS Comput. Biol.* **16**, e1008120 (2020).

24. Arzalluz-Luque, Á. & Conesa, A. Single-cell RNAseq for the study of isoforms—how is that possible? *Genome Biology* vol. 19 (2018).

25. Zhu, P. J. *et al.* Activation of the ISR mediates the behavioral and neurophysiological abnormalities in Down syndrome. *Science* **366**, 843–849 (2019).

26. Halliday, M. & Mallucci, G. R. Translating translation in Down syndrome. *Science* vol. 366 797–798 (2019).

27. Andrews, S. *et al.* FastQC. (2010).

28. Yates, A. D. *et al.* Ensembl 2020. *Nucleic Acids Res.* (2019) doi:10.1093/nar/gkz966.

29. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

30. Picard Toolkit. *Broad Institute, GitHub repository*.

31. Jun, G. *et al.* Detecting and Estimating Contamination of Human DNA Samples in Sequencing and Array-Based Genotype Data. *The American Journal of Human Genetics* vol. 91 839–848 (2012).

32. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

33. Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Research* vol. 47 D766–D773 (2019).

34. Website. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

35. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

36. Hansen, K. D., Irizarry, R. A. & Wu, Z. Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics* vol. 13 204–216 (2012).

37. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* vol. 82 (2017).

38. Cao, Y. *et al.* scDC: single cell differential composition analysis. *BMC Bioinformatics* **20**, 721 (2019).

39. Walker, R. L. *et al.* Genetic Control of Expression and Splicing in Developing Human Brain

1   Informs Disease Mechanisms. *Cell* **179**, 750–771 (2019).

2   40. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-*

3   *5®)*. (American Psychiatric Pub, 2013).

1 **Acknowledgements**
2
3 Tissue, biological specimens or data used in this research were obtained from the Autism BrainNet (formerly
4 the Autism Tissue Program), which is sponsored by the Simons Foundation, and the University of Maryland
5 Brain and Tissue Bank, which is a component of the NIH NeuroBioBank. We are grateful to the patients and
6 families who participate in the tissue donation programs. Funding for this work was provided by grants to
7 D.H.G (NIMH R01MH110927, U01MH115746, P50-MH106438, and R01 MH-109912, R01 MH094714), grants
8 to MJG (SFARI Bridge to Independence Award, NIMH R01-MH121521, NIMH R01-MH123922, NICHD-P50-
9 HD103557), grants to JRH (Achievement Rewards for College Scientists Foundation Los Angeles Founder
10 Chapter, UCLA Neuroscience Interdepartmental Program). We also thank Janet Sinsheimer for discussion of
11 the transcriptomic regional identity analysis methodology.
12
13 **Author Contributions**
14
15 J.R.H contributed to the entire manuscript, including all analysis of the bulk RNA-seq data, supporting analysis
16 of the snRNA-seq data, writing the main text, creating all figures, and compiling all supplementary information.
17
18 B.W. completed the snRNA-seq experiment with matched samples, and also performed the main cell-type
19 proportion and DE gene analyses with the snRNA-seq data. B.W. also contributed to Figure 4.
20
21 G.T.C. completed an experiment with human *in vitro* neuronal cultures over-expressing ETV4, to determine if
22 this gene may be a transcriptional regulator of some of the attenuated regional identity genes identified in this
23 work. Results were inconclusive, and were ultimately not included in this manuscript.
24
25 S.P. completed all of the raw data processing of the bulk RNA-seq data (obtaining gene and isoform level
26 quantifications and RNA-seq quality metrics).
27
28 P.E. conducted the cell-type deconvolution analysis.
29
30 N.C. contributed to the cell-type deconvolution analysis.
31
32 G.D.H provided expert functional circuit and regional neuroanatomy input for the manuscript and assisted with
33 figure design.
34
35 D.A. carried out the protein-protein interaction enrichment analysis for gene and isoform modules.
36
37 G.K. assisted with making the GitHub repository and made the R Shiny data visualization application
38 accompanying this work.
39
40 G.R. provided assistance with the bulk RNA-seq analysis data interpretation and visualization.
41
42 C.L.H. created the RNA-seq computational pipeline used to process the raw bulk RNA-seq data and obtain
43 gene and isoform level gene expression quantification.
44
45 T.J. contributed to the cell-type deconvolution analysis.
46
47 D.W. contributed to the cell-type deconvolution analysis.
48

19

1   J.O. extracted RNA and prepared RNA libraries for the bulk RNA-seq.

3   Y.E.W. assisted with sample dissection for the bulk RNA-seq RNA extraction.

5   N.N.P. contributed to the initial planning of this experiment and assisted with sample dissections for the bulk
6   RNA-seq RNA extraction.

8   V.S. contributed to the initial planning of this experiment and assisted with sample dissections for the bulk
9   RNA-seq RNA extraction.

11   G.B. contributed to the initial planning of this experiment.

13   M.Ge. served in an advisory and project-planning role for the cell-type deconvolution analysis.

15   B.P. provided expert statistical advice for the bulk RNA-seq analysis, particularly for the permutation tests
16   assessing region-specific increases in magnitude of ASD gene expression effects.

18   M.J.G. provided project leadership, expert advice, substantial contributions to the entire manuscript, and
19   mentorship for J.R.H. along with funding for this work. M.J.G also contributed to sample dissections for the bulk
20   RNA-seq RNA extraction.

22   D.H.G. provided project leadership, expert advice, substantial contributions to the entire manuscript, and
23   mentorship for J.R.H. along with funding for this work. D.H.G also obtained all of the samples for RNA-seq.

Competing Interests Statement

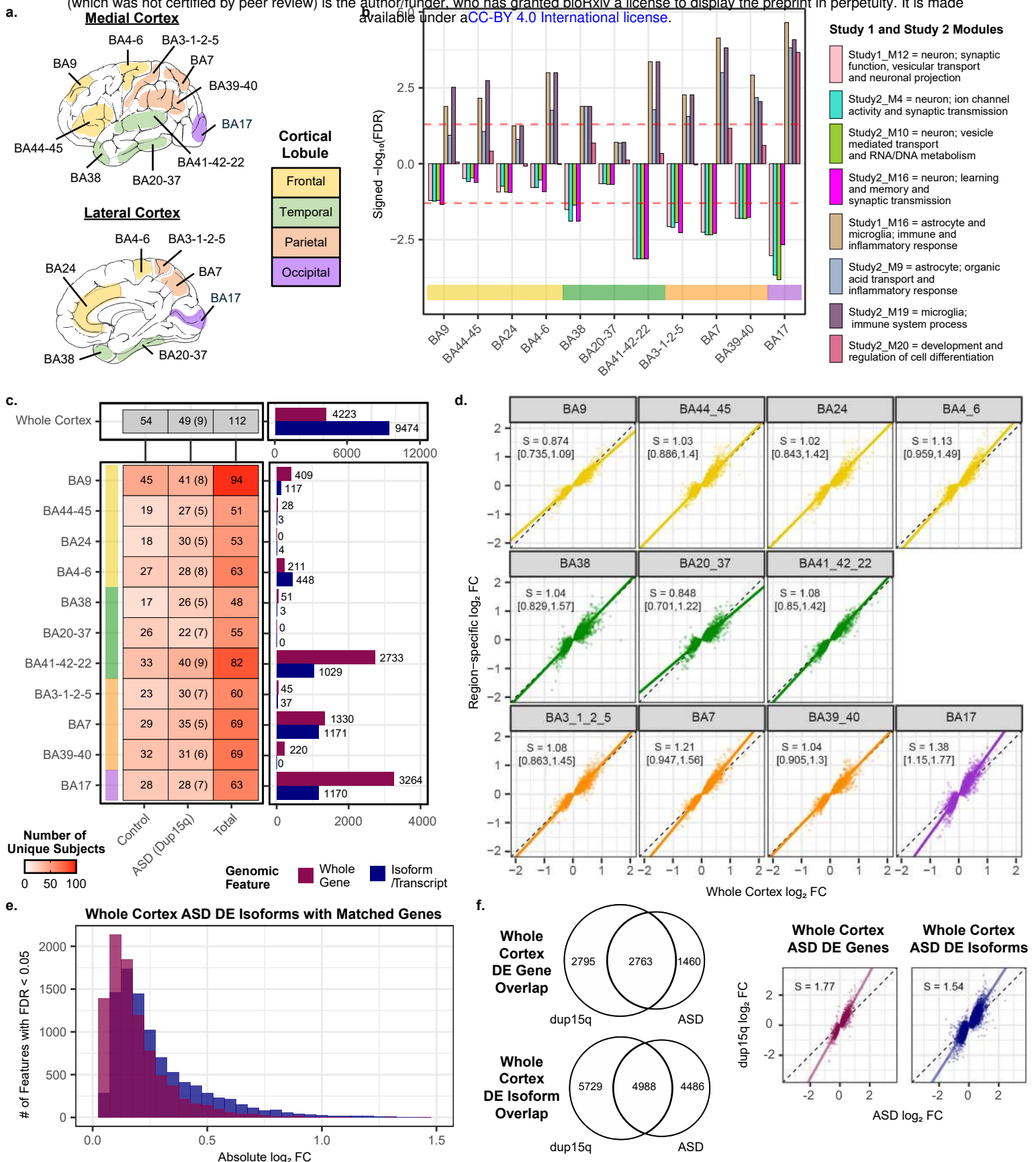The authors declare no competing interests.

Figures

20

Figure 1 | ASD Transcriptomic Differences Across 11 Cortical Regions. a. Human cortical Brodmann Areas (BA) with cortical lobules indicated. Cortical lobule colors are consistent throughout the figure. **b.** Dysregulation of previously identified co-expressed gene modules (Study 1: Voineagu et al., *Nature* 2011; Study 2: Parikshak et al., *Nature* 2016)[4-5] across cortical regions. Red dashed line marks FDR < 0.05. **c.** Unique subjects by region and diagnosis (left), with number of differentially expressed (DE; linear mixed model FDR < 0.05) genes and isoforms (right) across the whole cortex (top) or within individual regions (bottom). Gene and isoform colors are consistent throughout this figure. **d.** log₂ fold change (FC) of individual regions compared to the whole cortex log₂ FC for the 4,223 whole cortex DE genes. Slope (S, with 95% confidence interval in brackets) is calculated with principal components regression. **e.** For the whole cortex DE isoforms, a histogram of all isoforms DE across the whole cortex in ASD with their matched genes. **f.** Left: venn diagrams depicting the number of genes and isoforms DE across the whole cortex in dup15q samples compared to idiopathic ASD samples. Right: for the ASD whole cortex DE genes (left) and isoforms (right), the idiopathic ASD whole cortex log₂ FC compared to the dup15q whole cortex log₂ FC. Slope (S) is calculated with principal components regression.
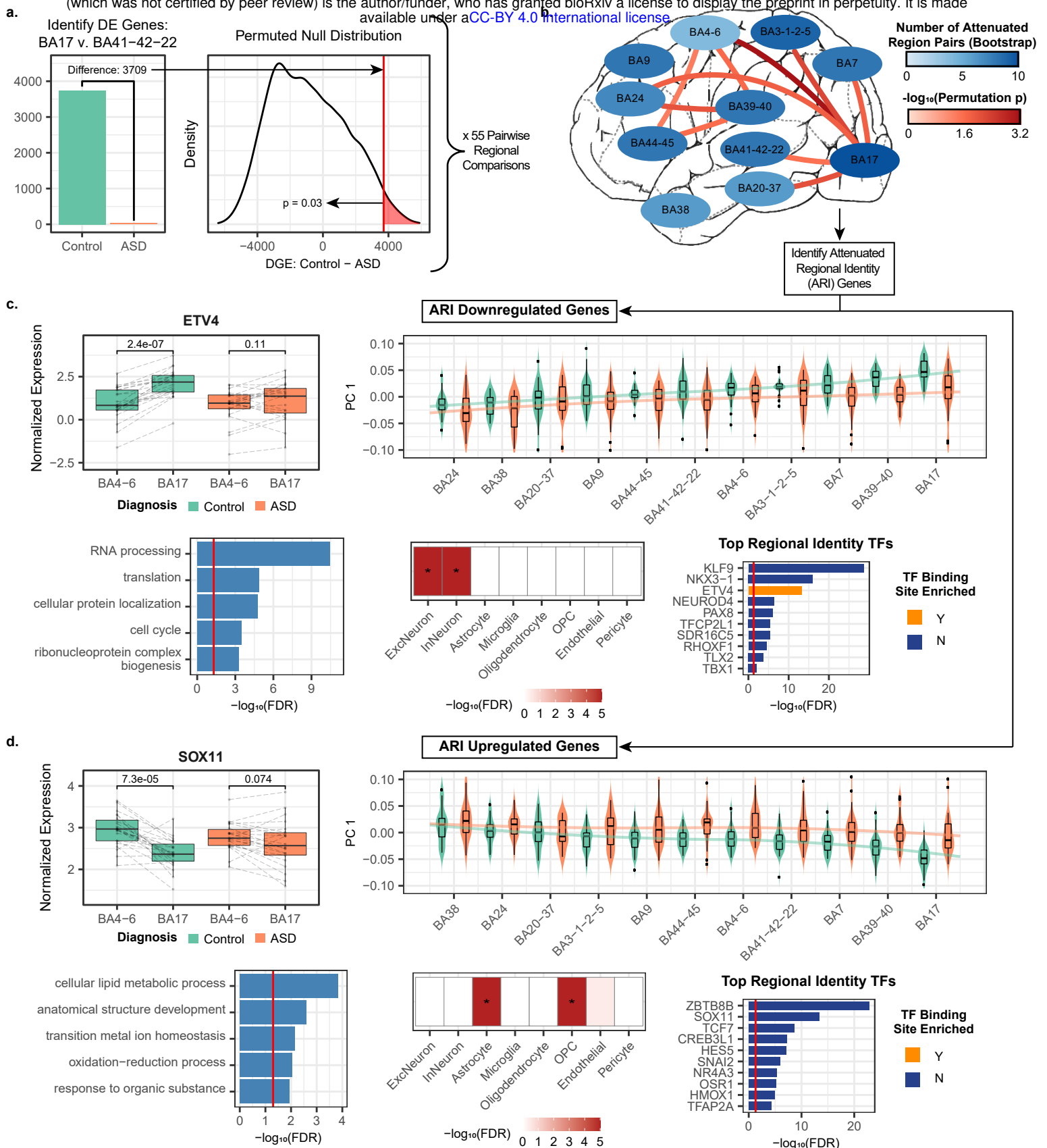
Figure 2 | Transcriptomic Regional Identity Attenuation in ASD. a. Methods overview for identifying statistically significant differences in transcriptomic regional identity in ASD. The regional comparison of BA17 v. BA41-42-22 is used here as an example. Number of DE genes between regions is calculated in controls and ASD samples (left). A permuted null distribution is then used to determine the significance of the difference in DE genes between controls and ASD samples (right). **b.** Regional comparisons with attenuation of transcriptomic regional identity in ASD with permutation p < 0.05 are connected with a bar (red scale), and a general trend towards cortex-wide attenuation is summarized by region color (blue scale; 0=no pairs exhibiting attenuation in ASD, 10=every pair exhibits attenuation in ASD). For region color, a regional pair is considered attenuated if it contains less DE in ASD compared to controls as measured with a bootstrap approach (**Methods**). Attenuated regional identity (ARI) genes are extracted from regional comparisons with permutation p < 0.05 (**Methods**). **c-d.** Overview of ARI downregulated (**c.**) and upregulated (**d.**) genes. Top left, select attenuated transcription factors in BA17 and BA4-6. Lines link paired samples from the same subject, and the paired Wilcoxon signed-rank test p-value is plotted above boxplots. Top right, PC 1 of ARI genes across all regions. Bottom left and bottom center, gene ontology and cell-type enrichment, respectively. Bottom right, top 10 attenuated transcription factors (TFs), where FDR is representative of how well these TFs distinguish BA17 and BA39-40 from the other nine cortical regions assessed here in controls (Methods). Enrichment for transcription factor binding sites is also depicted (Bonferroni-corrected p-value < 0.05 required for enrichment). Please see the Supplementary Methods for a description of all boxplots included in figures.
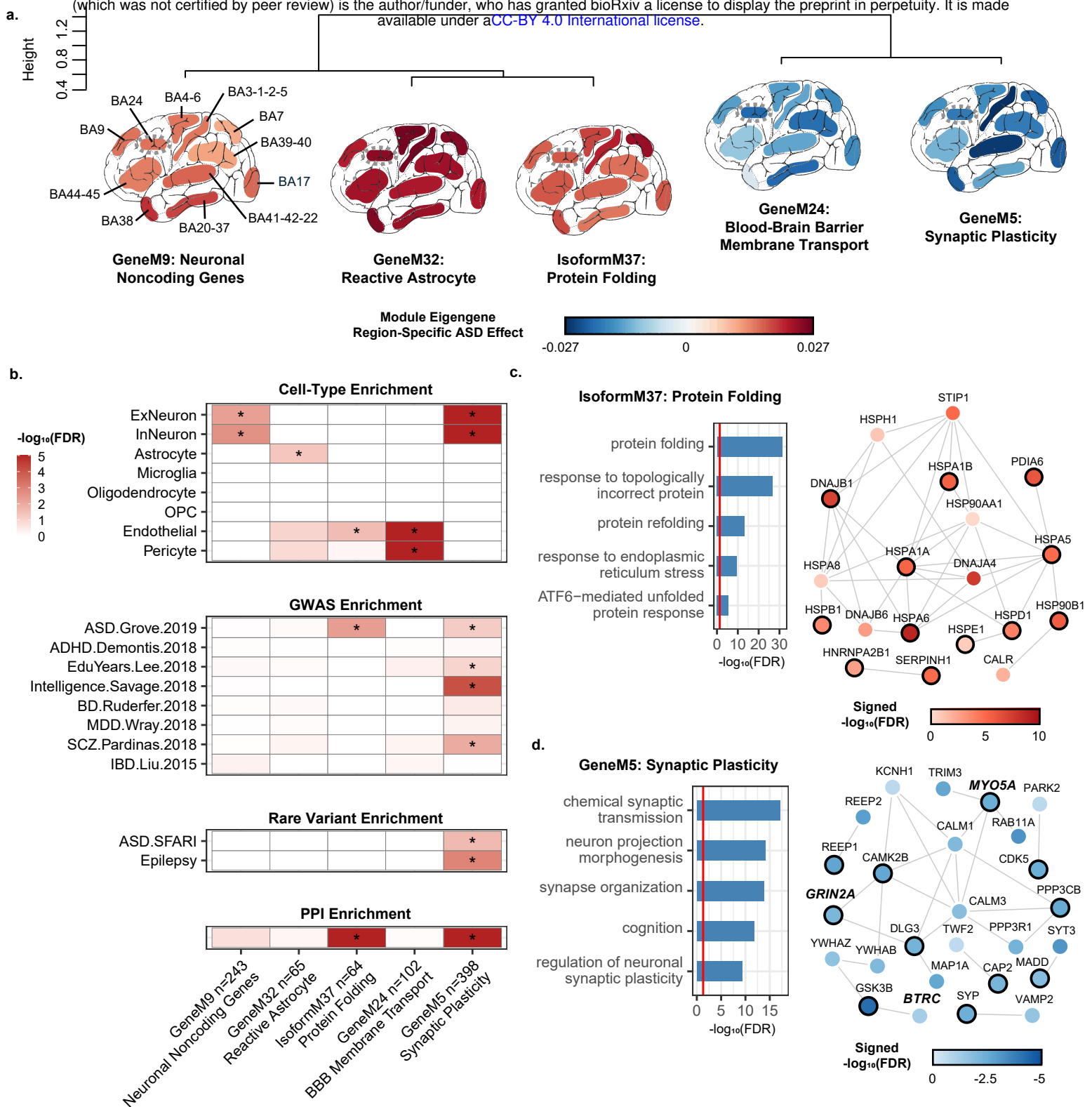
Figure 3 | Co-expression Network Analysis Characterizes Cortex-wide Dysregulation of ASD Risk Genes. a. Average linkage hierarchical clustering of the biweight midcorrelation of the top 5 most dysregulated gene and isoform co-expression module eigengenes (first principal component of the module) with regionally-consistent patterns of ASD dysregulation. The module eigengene ASD effect is indicated for each cortical region examined (**Methods**). **b.** -log10(FDR) for cell-type, GWAS, rare variant, and protein-protein interaction (PPI) enrichment for the modules depicted in **a**. '*' indicates a significant enrichment (FDR < 0.05 for cell-type, rare variant, and PPI enrichment, and FDR < 0.1 for GWAS enrichment). 'n' indicates the number of genes/isoforms in each module. **c-d.** For ASD GWAS enriched modules Isoform M37 (**c**) and GeneM5 (**d**), top gene ontology terms (left) and hub genes (module genes within the top 20 genes with the highest correlation with the module's eigengene) that participate in a protein-protein interaction (PPI) with any other module gene are depicted along with their PPI partners (right). Node color is the signed -log10(FDR) of the whole cortex ASD effect, edges denote direct PPIs, and hub genes are indicated with a black outline. SFARI database[14] gene names are in bold italic.

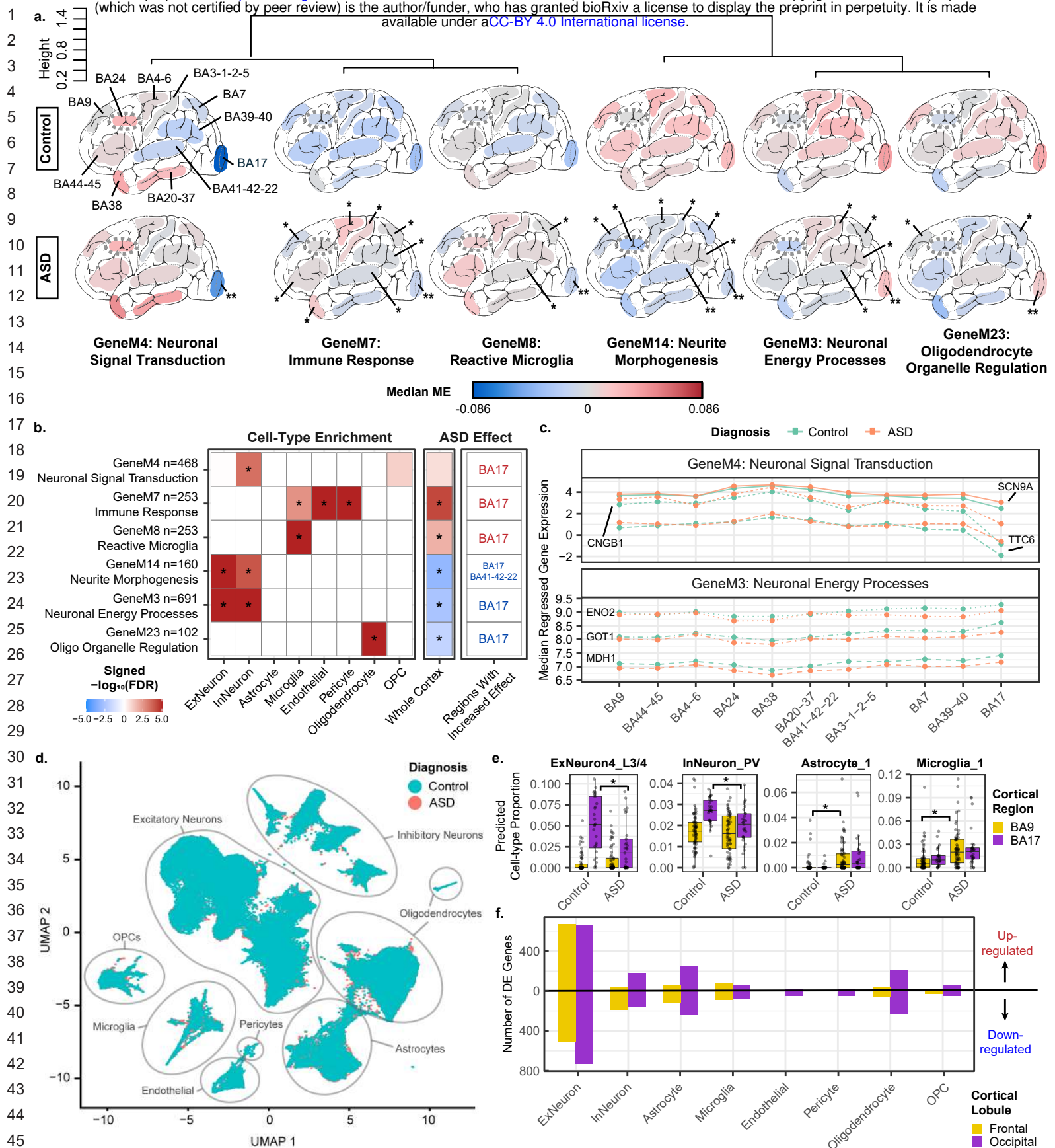**Figure 4 | Functional Characterization of Regionally-variable Transcriptomic Dysregulation in ASD. a.** Average linkage hierarchical clustering of the biweight midcorrelation of the top 6 most dysregulated gene co-expression module eigengenes (ME; first principal component of the module) with regionally-variable patterns of dysregulation. The median of the ME, stratified by diagnosis, is depicted for each cortical region examined. Significant region-specific dysregulation in ASD is marked with '*', and regions with a significantly increased magnitude of effect compared to the whole cortex effect is marked with '**' (**Methods**). **b.** Left and center, signed -log₁₀(FDR) for cell-type enrichment (left) and whole cortex ASD effect for the modules depicted in **a** (center). '*' indicates a significant enrichment (FDR < 0.05). Right, regions marked with '**' in **a** are listed. Red text indicates upregulation and blue text indicates downregulation in ASD. 'n' is the number of genes in each module. **c.** Median regressed gene expression for the top three hub genes for GeneM4 (top) and GeneM3 (bottom). **d.** UMAP plot for snRNA-seq, containing matched ASD and neurotypical control samples from both the frontal (BA9 and BA4_6; ASD: 4, Control: 2) and occipital (BA17; ASD:2, Control: 2) cortical lobules. **e.** Predicted neural cell-type proportions obtained from cell-type deconvolution with bulk RNA-seq samples, using snRNA-seq cell-type markers. Cell-types with significant proportion differences in ASD are shown, with significant ASD changes marked with '*'. **f.** DE genes identified with snRNA-seq. Total number of down-regulated (bottom) and up-regulated (top) genes in ASD samples compared to controls is shown (the sum of all cell subtype DE genes within each broad cell-type).