Table of contents

Supplemental Data

Supplemental Figures

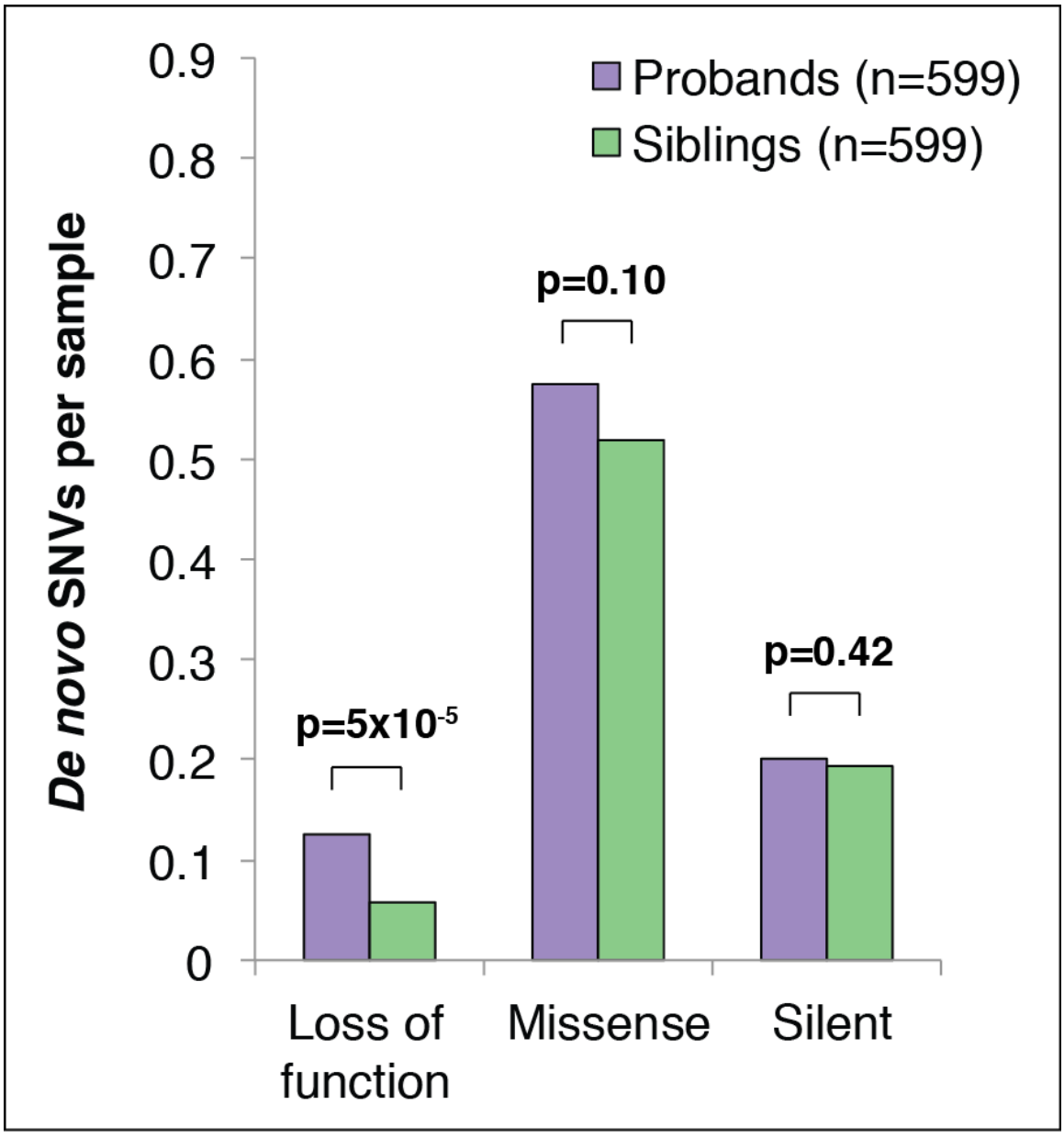

Figure S1. *De novo* burden analysis

The number of *de novo* mutations per sample are shown in 599 probands (purple) and 599 matched unaffected siblings (green) from the Simons Simplex Collection (SSC).

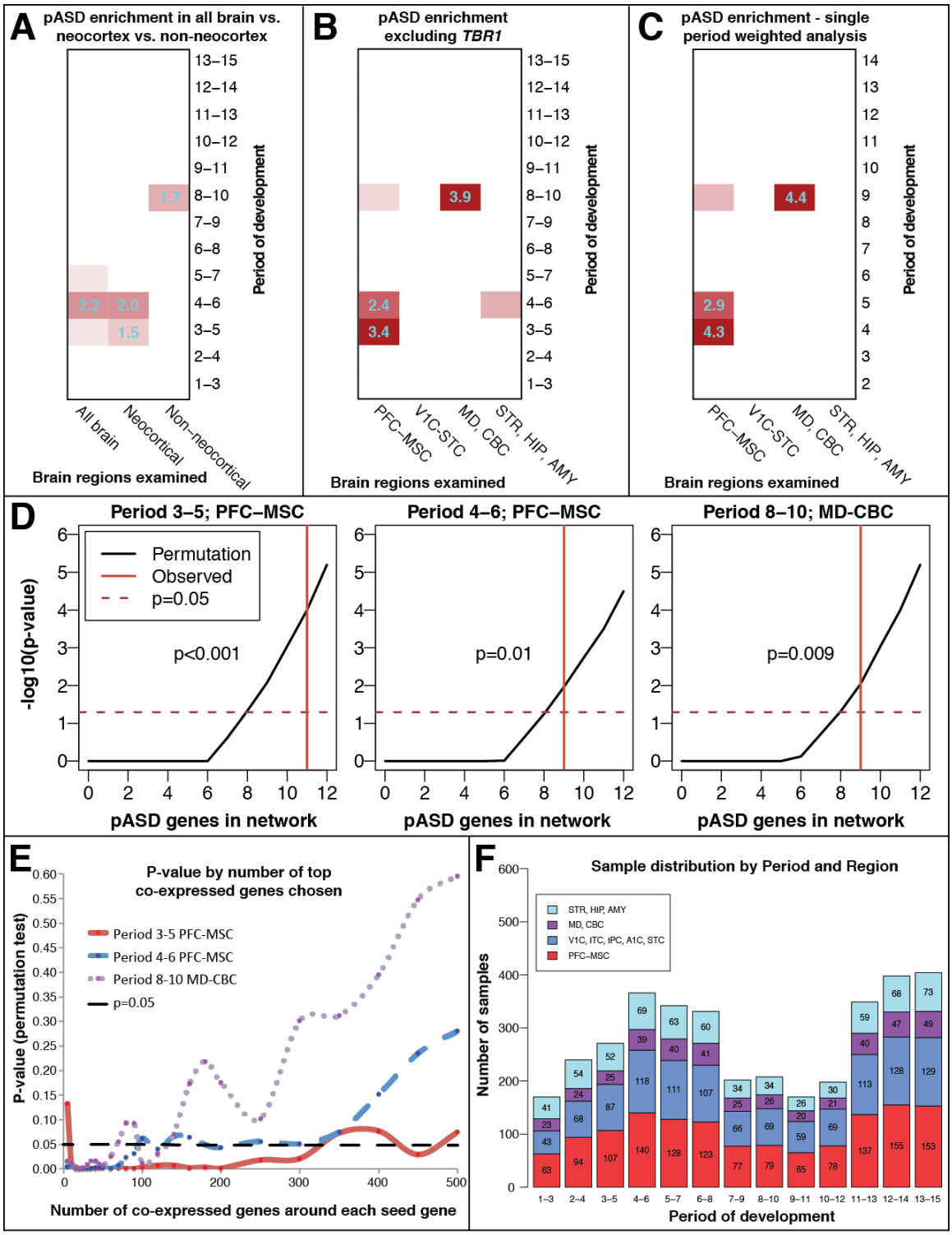

Figure S2. Additional spatio-temporal network analyses and permutation tests.

A) Networks built from all brain regions, neocortical regions, and non-neocortical regions. To achieve spatio-temporal resolution, co-expression networks were formed from subsets of the expression data based on developmental stage (in 3-period

windows) and brain region (in windows based on transcriptional similarity of the 12 (periods 1-2) or 16 regions (periods 3-15)). Each of the networks was tested for enrichment of 122 pASD genes. This heatmap shows the negative $\log_{10}$(p-value) of this enrichment (hypergeometric test) for each network with developmental stages on the Y-axis and brain regions on the X-axis. Networks that are not significant are in white; nominally significant networks are in light red; networks that are significant after correction for multiple comparisons are in red and have the negative $\log_{10}$(p-value) overlaid with cyan text. The presence of a strong signal in period 3-5 and 4-6 in all brain and in the neocortex indicated clustering of brain regions by transcriptional similarity during fetal development (periods 3-7) would increase resolution of our network analysis by separating these regions meaningfully during this time. As well, gene expression changes are most dynamic during fetal development (Kang et al., 2011). B) The analysis in main text Figure 2 was repeated without *TBR1* as an hcASD gene. Excluding *TBR1* produces essentially identical results in terms of temporal and spatial specificity. C) The analysis in main text Figure 2 was repeated using a weighted approach that places the greatest emphasis on individual periods while including data from the prior and subsequent interval. Specifically, for each network, the co-expression data for the period is given a weight of 1 and the periods immediately before and after are given a weight of 0.5. This method shows essentially identical results in terms of temporal and spatial specificity as the three period sliding window analysis in Figure 2. D) Permutation test to estimate the significance of observing enrichment of multiple pASD genes in a co-expression network built around nine hcASD seeds by permuting the pASD genes. The black line shows the estimated p-value as the number of pASD genes in the co-expression network increases (corrected for multiple comparisons); the vertical red line shows the number of pASD genes observed for each specific network. The dashed purple horizontal line represents the point on the y-axis at which p=0.05. E) The co-expression networks were built from the top 20 co-expressed genes for each seed gene. To assess whether the results were robust to altering the choice of 20 genes a permutation test was performed in which this number was varied between 5 and 500. The red, blue and purple lines show how the p-value varied for the period 3-5 PFC-MSC, period 4-6 PFC-MSC, and period 8-10 MD-CBC respectively. The horizontal dashed black line represents the threshold of p=0.05. All three networks were robust (below p=0.05) to variation between 10 and 60 top co-expressed genes with the period 3-5 PFC-SMC network remaining significant up to 300 genes. F) The number of tissue samples available in the BrainSpan exon array dataset for each spatio-temporal window is plotted by developmental window (X axis) and group of brain regions (colors). The total BrainSpan dataset in Kang et al. (2011) comprises 1,340 tissue samples from 57 clinically unremarkable brain samples; however, the total number of samples listed in this figure is greater because the developmental windows are overlapping.

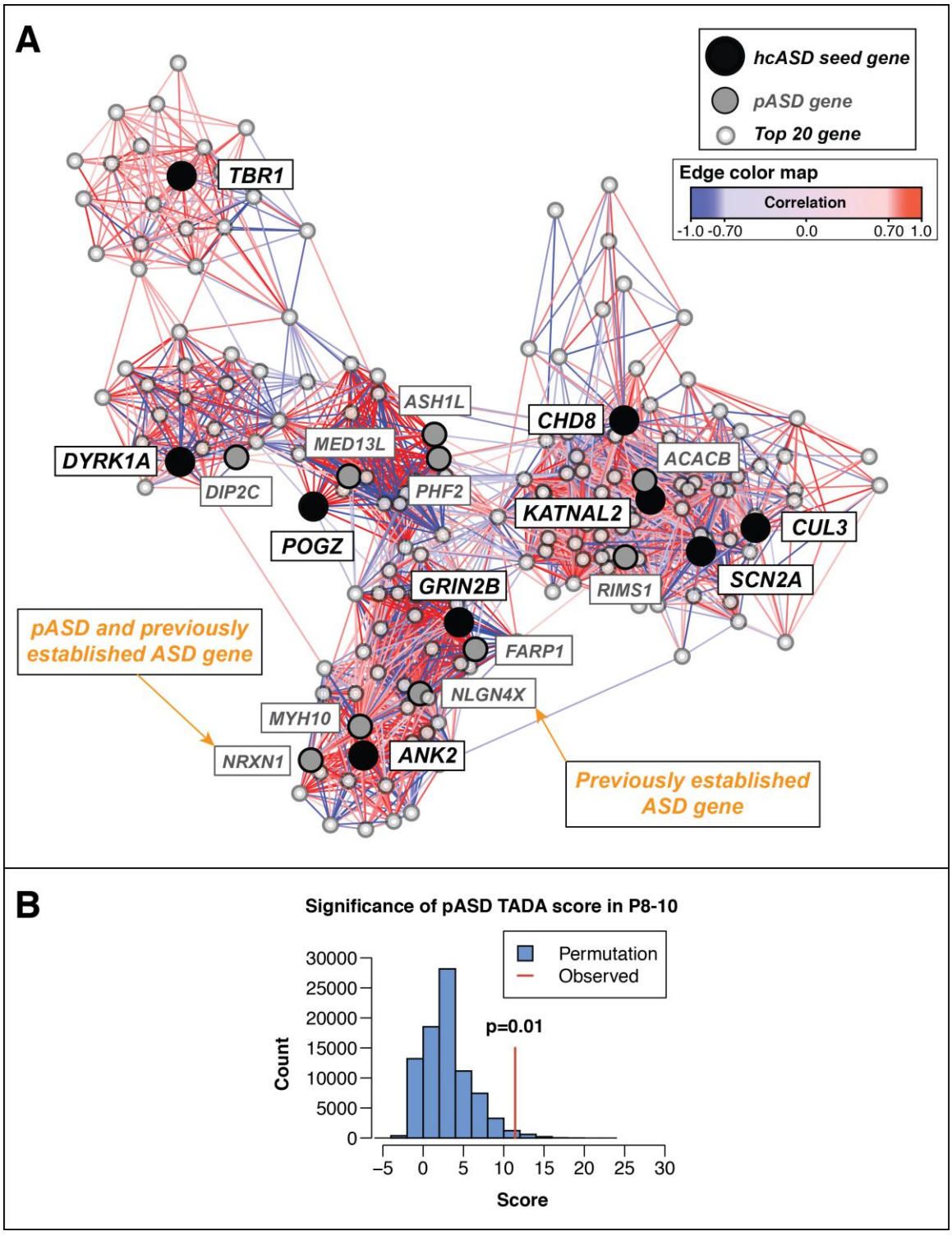

Figure S3. The period 8-10 MD-CBC network.

A) The period 8-10 MD-CBC co-expression network. Gene co-expression analysis identified the 20 genes best correlated with each hcASD gene. The hcASD seed genes are in black; pASD genes identified within the network are in gray; and top 20 co-

expressed genes that are not pASD genes are in white. The lines (edges) link genes with co-expression correlation >0.7 and the shade represents the strength of the correlation; positive correlations are in in red; negative correlations are in blue. Previously established ASD genes *NRXN1* and *NLGN4X* (also a pASD gene) are highlighted with orange arrows. B) The TADA score uses inherited and mutation probability for each gene to assess which pASD genes are likely to be true ASD genes (He et al., 2013). This histogram shows the results of a permutation test (100,000 iterations) assessing the combined TADA score in the period 8-10 network; the observed score (11.4) is shown by the vertical red line and corresponds to a corrected p-value of 0.01.

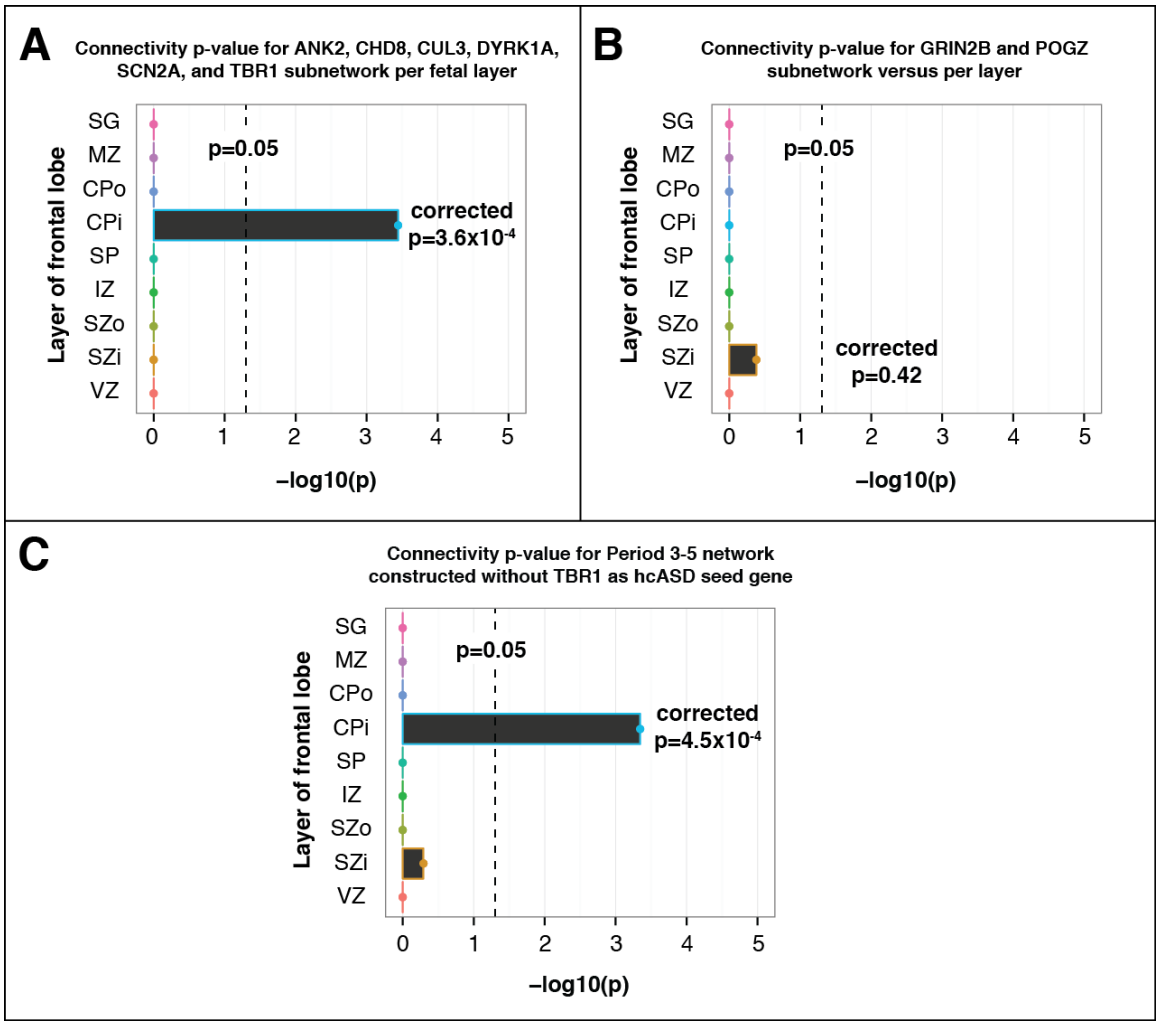

Figure S4. Additional layer-specific analyses

A) The connectivity of two subnetworks within the period 3-5 network were assessed in a separate prenatal transcriptome (BrainSpan, www.brainspan.org). This dataset is comprised of microarray-based gene expression profiles of laser micro-dissected human mid-fetal (period 4-6) brains. Significance of the observed extent of connectivity (sum of correlations along network connections) for each layer was assessed by permutation

test. A) The *ANK2*, *CHD8*, *CUL3*, *DYRK1A*, *SCN2A*, and *TBR1* subnetwork shows greatest connectivity in the CPi region (inner cortical plate) corresponding to neocortical layers 5-6 (corrected $p=3.6x10^{-4}$). B) The *GRIN2B* and *POGZ* subnetwork shows greatest connectivity in the VZ region (corrected p=0.42). B) The connectivity of the period 3-5 network built excluding *TBR1* (see Supplemental Figure S2) was assessed as in 'A'. This network shows greatest connectivity in the CPi region (corrected $p=4.5x10^{-4}$), and therefore, excluding *TBR1* produces essentially identical results with respect to laminar specificity as the three period sliding window analysis in main text Figure 4. Abbreviations: SG: subpial granular zone; MZ: marginal zone; CPo: outer cortical plate; CPi: inner cortical plate; SZ: subplate zone; IZ: intermediate zone; SZo: outer subventricular zone; SZi: inner subventricular zone; VZ: ventricular zone.

Supplemental Tables

Table S1. *De novo* mutations (attached xslx file)

This spreadsheet details the genes and *de novo* mutations that were used to build and interrogate the co-expression network. There are several sheets within the excel spreadsheet: **hcASD**, the high-confidence ASD genes with multiple *de novo* loss of function (LoF) mutations that were used as seeds in the co-expression network; **pASD**, the probable ASD genes with single *de novo* LoF mutations that were used to assess enrichment in the co-expression network (only the subset with expression data from Kang et al. were used in the network analysis – see column m); **Burden_LoF**, the *de novo* LoF mutations in 599 matched probands and siblings used to estimate *de novo* burden; **56_families**, the 56 additional families from the Simons Simplex Collection (SSC) that were analyzed with exome sequencing.

Table S2. Cross-validation experiment.

This table summarizes the results from the cross-validation experiment. In each of the 200 iterations performed, 1 hcASD and 10% (12) of the pASD genes were removed randomly, the co-expression analysis was conducted in each of the 52 spatio-temporal points, and the networks with the greatest number of pASD genes were recorded. 100% of the time, the network with the most pASD genes corresponded to one of the three networks identified with the full set of data.

Network(s) with highest pASD enrichment	Occurrence	Percent (%)
Period 3-5 MFC-MSC	106	53
Period 4-6 PFC-MSC	0	0
Period 8-10 MD-CBC	42	21
Period 3-5 and Period 4-6 PFC-MSC	4	2
Period 3-5 PFC-MSC and Period 8-10 MD-CBC network	47	23.5
Period 4-6 PFC-MSC and Period 8-10 MD-CBC	0	0
Period 3-5 and Period 4-6 PFC-MSC, and Period 8-10 MD-CBC	1	0.5

Table S3. Genes in networks (attached xslx file)

This spreadsheet shows the genes that make up the co-expression networks in the main manuscript. There are three sheets corresponding to the three spatiotemporal networks showing enrichment for pASD genes: 1) **Periods3-5_PFC-MSC**; 2) **Periods4-6_PFC-MSC**; 3) **Periods8-10_MD-CBC**. Each sheet has two lists of genes: 1) A list of all the genes in the network and whether each gene was a seed (hcASD) or top 20 co-expressed gene (pASD or top20); and 2) a list of network connections (each with pairwise correlation value). The **Periods3-5_PFC-MSC** sheet has one additional column identifying genes as in the *ANK2*, *CHD8*, *DYRK1A*, *SCN2A*, and *TBR1* cluster or the *GRIN2B* and *POGZ* cluster.

Table S4. Genes differentially expressed in the *TBR1* knockout

This spreadsheet shows the genes determined via RNA-seq to be differentially expressed in the TBR1 knockout (KO) mouse versus its wildtype littermate. There are two sheets: 1) **Downregulated_Genes** lists genes downregulated in the KO, and 2) **Upregulated_Genes** lists genes upregulated in the KO. The corresponding human orthologues are also listed (overlap with spatiotemporal networks was assessed based on the corresponding human orthologue).

Table S5. Layer specific genes during mouse development (attached xslx file)

This spreadsheet shows the genes determined via RNA-seq to be exclusively differentially expressed in deep layers (DL; correspond to future layers L5 and L6), layer 4 (L4; corresponds to future layer L4), or superficial layers (SL; correspond to future layers L2 and L3) of mouse neocortex during the course of brain development. Layer specificity was assessed at postnatal day (P) 4, P6, P8, P10, and P180. Within each time point, the three layers (DL, L4, and SL) were compared pairwise and the upregulated layer is indicated in Column L (prefixed by the relevant timepoint; i.e. L4 at P4 is abbreviated P4L4). To identify layer markers for a particular timepoint and layer, filter by column L and take the unique set of genes (each gene will appear twice due to the two pairwise comparisons).

The following criteria determined differential expression:

1) Upregulated with log2(fold change) ≥ 1 versus other two layers
2) P-value and q-value both ≤ 0.05
3) FPKM ≥ 10 in at least one sample

These genes were used to assess enrichment by layer of the period 3-5 and 4-6 PFC-MSC networks. The corresponding human orthologues are also listed (overlap with spatiotemporal networks was assessed based on the corresponding human orthologue).

Table S6. Cell-type specific marker genes

This table summarizes cell-type specific marker genes from Kang et al. (2011). Enrichment of these markers within the period 3-5 and period 4-6 PFC-MSC networks were assessed by permutation test.

Cell type	Marker genes
Astrocyte	*GFAP, S100B, ALDOC*
Cortical GABA interneuron	*CALB2, CALB1, NOS1, PVALB, CCK, VIP, DLX1, DLX2, NKX2-1, ASCL1, GAD1, GAD2*
Cortical glutamatergic neuron	*RELN, CUX1, UNC5D, RORB, BCL11B, ETV1, FEZF2, OTX1, FOXP2, NTSR1, SOX5, SSTR2, TBR1, TLE4, ZFPM2, CTGF, UNC5C*
L5 & L6 Cortical glutamatergic neuron	*ZFPM2, NTSR1, TLE4, FOXP2, TBR1, SOX5, SSTR2, FEZF2, BCL11B, OTX1, ETV1*
L1-L4 Cortical glutamatergic neuron	*CUX1, UNC5D, RORB, CUX2, SATB2, WFS1, RELN*
Microglia	*CFH, FCER1G, TNIP2*
Oligodendrocyte	*CNP, CSPG4, OLIG1, OLIG2, PDGFRA*

Extended Experimental Procedures

Exome data and mutations

Published *de novo* mutations

Published *de novo* variants detected by exome sequencing in 943 ASD probands were obtained from the supplemental data of four publications: 225 from the Simons Simplex Collection (SSC) (Sanders et al., 2012), 343 from the SSC (Iossifov et al., 2012), 200 from the SSC (O'Roak et al., 2012b), and 175 from the ARRA consortium (Neale et al., 2012).

In addition *de novo* variants detected by exome sequencing in 543 matched siblings of ASD probands were obtained from the supplemental data of two publications: 200 from the SSC (Sanders et al., 2012) and 343 from the SSC (Iossifov et al., 2012).

To maximize our detection of *de novo* loss of function mutations we included two further papers: 1) genome-wide sequencing of 44 probands with ASD (Kong et al., 2012) in which two *de novo* loss of function (LoF) mutations were described in the main text, however the sample IDs were not stated; 2) targeted sequencing of genes with known *de novo* variants (O'Roak et al., 2012a). To keep the analysis fair only mutations in samples already sequenced using exome technology were included. These two publications added two further 2-hit hcASD genes, which were used as seeds: *CUL3* (Kong et al., 2012) and *TBR1* (O'Roak et al., 2012a). All *de novo* variants from all studies were re-annotated according to the Consensus CDS database (Harte et al., 2012; Pruitt et al., 2007).

The total samples size of published data is therefore 987 probands, 543 of which have *de novo* data for matched siblings.

Additional exome samples from 56 families

Whole-blood derived DNA for 56 families chosen at random from the Simons Simplex Collection was subjected to exon capture using the Nimblegen EZExomeV2.0 array and sequenced with 74bp paired-end reads on the Illumina HiSeq 2000. Sequence reads were aligned to hg19 using BWA (Li and Durbin, 2009); duplicate reads were removed with Samtools (Li et al., 2009). Single nucleotide variants were predicted using Samtools (Li et al., 2009) and *de novo* variants were identified using in-house scripts (Sanders et al., 2012). All *de novo* variants were confirmed using PCR and Sanger sequencing. Complete details of our methods for sequencing and analyzing *de novo* variants can be seen in our previous publication (Sanders et al., 2012). Details of the 56 additional SSC families are shown in Supplemental Table S1.

Detection of insertion-deletions (indels)

To maximize our detection of *de novo* loss of function variants we reassessed 225 trios from our previous publication (Sanders et al., 2012) to detect *de novo* frameshift indels. The aligned BAM files were analyzed using local realignment with dindel (Albers et al., 2011). Ten further *de novo* frameshift indels were detected and confirmed using PCR and Sanger sequencing.

Confirmation of *de novo* events

The rarity of *de novo* variants increases the potential for false positive variants in the children and false negatives in the parents. The vast majority of *de novo* LoF variants used as the input for the analyses presented in the main manuscript have been confirmed through PCR of the interval in the child and both parents, followed by Sanger sequencing to demonstrate that a variant is present and that it is truly *de novo*. Out of 144 *de novo* LoF variants in this analysis, 138 (96%) have been confirmed. Of the remaining six variants, five are *de novo* frameshift indels from the Iossifov paper (*MED13L*, *BCL11A*, *EFCAB5, TRIM17*, *TROVE2*) in which confirmation has not been attempted, however they do demonstrate a 96% (148/154) *de novo* confirmation rate for the *de novo* variants for which confirmation was attempted. The final *de novo* LoF that is not explicitly confirmed is *CUL3* from the Kong paper; they demonstrate a 99% (93/94) *de novo* confirmation rate, however it is not stated whether this *de novo* LoF mutation was included within this set.

De novo mutation burden

The combination of published and new exome data gave us a total data set of 1,043 probands and 599 matched unaffected siblings. To estimate the significance of observing multiple *de novo* variants in the same gene it was necessary to estimate the odds ratio between matched probands and siblings. The results for 599 matched probands and siblings are shown in Supplemental figure S1. The *de novo* variants used for this analysis are shown in the tab marked 'Burden_LoF' in Supplemental Table S1. The odds ratio for *de novo* LoF mutations was 2.21 (95%CI: 1.45-3.36; $p=5x10^{-5}$, binomial exact test).

The difference between *de novo* LoF mutations in probands and siblings was highly significant with an odds ratio of 2.21 (95%CI: 1.45-3.36; $p=5x10^{-5}$, binomial exact test). Considering all the 1,043 ASD samples we obtained a list of 144 *de novo* LoF mutations, these are shown in the tabs marked 'hcASD' and 'pASD' in Supplemental table S1.

Significance of multiple *de novo* LoF mutations in the same gene

To estimate the significance of observing multiple *de novo* LoF mutations in the same gene for probands we need to estimate the chance of observing this outcome for the non-ASD associated *de novo* LoF mutations in siblings. In 1,043 ASD probands we observed 144 *de novo* LoF mutations. Based on the observed odds ratio of 2.21 we would expect 63 *de novo* LoF mutations in siblings. From this number we can estimate the probability of seeing 2 *de novo* LoF mutations in the same gene under the null hypothesis (a 2-hit gene).

P-value: Based on the birthday problem logic we can estimate the chance of observing 2 out of 63 LoF mutations in the same gene out of the 18,933 possible RefSeq genes as being 0.10 (this is the genome-wide p-value for a 2-hit gene and also the number of expected 2-hit genes). However genes differ in size and GC content, both of which are associated with more *de novo* mutations. A permutation test based on the 18,933 RefSeq genes was performed to allow size and GC content to be taken into account and the p-value was revised upwards to 0.1975.

Q-value: The p-value assess the chance of observing a 2-hit gene at least once in the experiment. The q-value, or false discovery rate (FDR), assess the chance of observing a two hits in a specific gene. Based on the permutation test we expect to observe 0.20 such genes, however we observe 9 genes with at least two *de novo* LoF mutations. For each gene with two *de novo* mutations is 45.57 times more likely to be a true ASD gene than a random 2-hit gene (9/0.1975). The q-value/FDR is therefore 0.022 (1/45.57) meaning that each 2-hit gene is an ASD gene with 97.8% confidence (1 - 0.022 * 100).

Applying these same methods to a 3-hit gene (i.e. 3 or more *de novo* LoF mutations in the same gene) gave a p-value of 0.0008 and a q-value of 0.0002 giving a confidence of 99.98% that such a gene is true ASD gene and exceeding the threshold for being genome-wide significant. Likewise a 1-hit gene has a q-value of 0.45 and a 55% confidence of being a true ASD gene.

Based on these metrics we designated a gene as being a high-confidence ASD gene (hcASD) if there were 2 or more *de novo* LoF mutations; 9 genes met or exceeded this threshold: one previously unreported gene with two *de novo* LoF mutations: *Ankyrin 2, neuronal* (*ANK2*); and confirms eight hcASD genes: four with three or more *de novo* LoF mutations – *Chromatin Helicase Domain 8 (CHD8); Dual Specificity Tyrosine Kinase 1A (DYRK1A); Glutamate receptor subunit 2B (GRIN2B);* and *Sodium channel, voltage-gated, type II, alpha subunit (SCN2A)*; and four with two *de novo* LoF mutations – *Cullin 3* (*CUL3); Katanin p60 subunit A-like 2 (KATNAL2); Pogo transposable element with ZNF domain (POGZ); and T-box brain gene 1* (*TBR1*). A gene was designated as a probable ASD gene (pASD) if there was a single *de novo* LoF mutation; 122 genes met this criterion. A complete list of these genes and the *de novo* LoF mutations are shown in the tabs marked 'hcASD' and 'pASD' in Supplemental Table S1.

Construction of spatio-temporal co-expression networks

Gene level expression data (Platform GPL5175; Affymetrix GeneChip Human Exon 1.0 ST Array) were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO accession number GSE25219) (Kang et al., 2011). Based on the quality control and quantile normalization by Kang et al. (2011) as well as their calculated gene-level expression values, expression data from the core probe set were used in co-expression analysis (exceptions: for hcASD gene *CHD8* and the pASD genes *FLG*, *FREM3*, *FRG2C*, *LMTK3*, *THSD7A*, *UBN2*, and *ZNF594,* data from the extended probe set were utilized). Genes with multiple transcript IDs (316 total) were removed from the analysis (exceptions: for *CHD8* and the pASD genes *FAM91A1*, *FREM3*, and *UBN2*, gene expression values from multiple transcript IDs were combined by averaging all data that mapped to the correct genomic locus – see below). Of note, only 116 of the 122 pASD genes had gene expression data (see Supplemental Table S3) and so all pASD enrichment analyses were based on these 116 genes.

Gene	Probe Set	Transcript IDs	Action
FAM91A1	Core	3114358, 3114365	Average both
CHD8	Extended	3528078, 3528085, 3555882	Average first two (3528078, 3528085)
FREM3	Extended	2787812, 2787824, 2787831	Average all
UBN2	Extended	3026891, 3075658	Use first only (3026891)

Co-expression analysis was conducted with an in-house R script. Briefly, within each post-mortem brain, gene expression values were determined per region: multiple tissue samples within the same region (i.e. different hemispheres or immature forms of mature regions) were averaged, resulting in a vector of expression values (expression profile) for each gene by brain region and brain sample. Within each spatio-temporal window, the vector of expression values was trimmed to relevant regions and brain samples only. Next, the absolute value of the Pearson correlation coefficient for the trimmed expression profile was calculated for each pairwise combination of genes. For each seed gene, the top 20 best correlated genes with an absolute correlation coefficient with the seed gene of ≥0.7 selected.

Temporal windows were determined solely by taking overlapping sets of three consecutive periods of development, starting with periods 1-3 and ending at periods 13-15. This provided adequate numbers of data points, as well as sufficiently dynamic data in each window, to assess correlations of changes in gene expression robustly while still preserving temporal resolution. To determine spatial windows, coherent subsets of brain regions were grouped based on transcriptional similarity. The presence of a strong signal in period 3-5 and 4-6 in all brain and in the neocortex (Supplemental Figure S2A) indicated clustering of brain regions by transcriptional similarity during fetal development would increase resolution of our network analysis by separating these regions meaningfully during this time. As well, gene expression changes are most dynamic during fetal development (Kang et al., 2011). Therefore, using expression data from fetal development (periods 3-7), within each brain region gene expression values were summarized by taking the median across all samples (expression values of probe sets targeting the same gene were averaged). Pairwise Spearman correlations between brain regions were then determined, and regions were hierarchically clustered in R using the

'hclust' function (1 – $corr^2$ as the distance, clustering using Ward's method). Distinct clusters of brain regions informed spatial window construction (see main text Figure 2A). Specifically, within the developing neocortex, the prefrontal cortex (PFC) areas (MFC, OFC, DFC and VFC) formed a distinct cluster, while the frontal primary motor cortex (M1C) clustered with its topographic neighbor, the primary somatosensory area (S1C). The clustering of M1C and S1C is at least in part driven by the fact that the two areas cannot be reliably distinguished until late mid-fetal development, due to the absence of the central sulcus. Together, the PFC, M1C, and S1C areas formed a separate larger cluster herein collectively referred to as the PFC-MSC (prefrontal cortex and primary motor-somatosensory cortex), which is clearly distinguishable from another large cluster comprised of the posterior perisylvian areas (IPC, STC, and A1C) and ITC, while the V1C was the most transcriptionally divergent among all analyzed areas.

Permutation tests

In the main manuscript we describe how spatiotemporal co-expression networks built around 9 hcASD seed genes (each with >97% confidence for being an ASD gene) can be assessed for enrichment of pASD genes (each with 54% confidence for being an ASD gene). The enrichment of pASD genes was observed in four specific brain regions and time periods: period 3-5 PFC-MSC, period 4-6 PFC-MSC, period 7-9 V1C-STC, and period 8-10 MD-CBC (Figure 2). The analysis was performed using a hypergeometric test (Figure 2), which makes the assumption that there is an equal chance of a pASD gene and a non-pASD gene being included in the network. However de novo mutations are more frequent in larger genes, effectively reducing the number of genes likely to contain a mutation and potentially leading to liberal estimates of p-values.

To obtain a measure of significance that better reflects the subtleties of co-expression data we used a permutation test based on selecting 9 pseudo-hcASD seed genes, based on the likelihood of observing 2-hit *de novo* LoF mutations by chance (taking gene size and GC content into account (Sanders et al., 2012)). Co-expression networks were built around each set of 9 pseudo-hcASD genes using the period 3-5 PFC-MSC, period 4-6 PFC-MSC, period 7-9 V1C-STC, or period 8-10 MD-CBC data. By assessing enrichment for pASD this analysis demonstrated a significant result for both mid-fetal modules as well as the period 8-10 MD-CBC network, but not for the period 7-9 V1C-STC network (100,000 iterations). Of note, for this and all downstream analyses that evaluate multiple hypotheses (e.g. multiple stages of development and/or regions of the brain), p-values are corrected for multiple comparisons by using the conservative Bonferroni correction within a family of tests.

Permuting pASD genes

The above permutation analysis tests the hypothesis that the identification of the 9 hcASD genes is critical to detection of the pattern of spatiotemporal pASD enrichment that we observed. To further confirm the robustness of our findings we considered the hypothesis that the identification of the 122 pASD genes was also critical to the detection of spatiotemporal enrichment. To test this we performed a permutation test in which the 9 hcASD genes remained constant while the 116 pASD genes with expression data were selected based on the likelihood of observing 1-hit *de novo* LoF mutations by chance (taking gene size and GC content into account and excluding the 9 seed genes from the list of possible genes (Sanders et al., 2012)). This permutation test was run to

100,000 iterations on each of the three spatiotemporal networks showing *de novo* LoF enrichment: period 3-5 PFC-SMC ($p<0.001$, corrected for 91 comparisons); period 4-6 PFC-SMC ($p=0.01$, corrected for 91 comparisons); period 8-10 CBC-MD ($p=0.009$, corrected for 91 comparisons); these results are shown in Supplemental Figure S2D. The 100,000 iterations were insufficient to detect a single iteration of 11 or 12 permuted pASD genes in the network therefore the results were extrapolated from the data from 0 to 10 permuted pASD genes to estimate the p-value.

Permuting number of co-expressed genes

Given the reliance on a bottom-up approach (i.e. building networks round a few high confidence genes as opposed to a top-down approach in which all genes are used to make a network which is then split into modules e.g. WGCNA) we considered how robust our results were to varying the number of co-expressed genes that were selected for each seed gene. In the main manuscript we used the 20 most highly correlated (Pearson's correlation) genes for each seed gene. To test this we used the same permutation test described in the main manuscript (permuting the 9 hcASD seed genes) but varied the number of most highly correlated genes selected for each seed between 5 and 500. For each number of most highly correlated genes 1,000 permutations were performed. The results show a great deal of stability with 10-300 genes exceeding a significance threshold of 0.05 for period 3-5 PFC-SMC and 5-90 genes exceeding the same threshold for period 4-6 PFC-SMC (Supplemental Figure S2E).

Permutation tests for enrichment of other genes

The 100,000 networks built around 9 pseudo-hcASD genes (see Permutation Tests above) for period 3-5 PFC-MSC, and the further 100,000 networks for period 4-6 PFC-MSC and for period 8-10 MD-CBC, were used as the basis for other permutation-based downstream analyses including TADA enrichment, connectivity, layer specific marker genes and cell-type specific marker genes.

Cross-validation experiment

To test further the robustness of our findings, we performed a cross-validation experiment with 200 iterations: for each iteration we randomly removed one hcASD seed gene and 10% (12) of the pASD genes, constructed all 52 spatio-temporal co-expression networks, and assessed the number of remaining pASD genes found within each network. We assessed the number of times that one of the top three networks (PFC-MSC in period 3-5 and 4-6, and MD-CBC in period 8-10; Figure 2B) was the most enriched for pASD genes. Consistent with our initial findings, we observed this 100% of the time (Supplemental Table S2), suggesting the observed spatio-temporal localization is robust to the set of input genes.

Single period weighted analysis

Finally, to ensure that the observed temporal and spatial localization was not an artifact of the choice of three period sliding windows ('windowed' analysis), we conducted a complementary single period 'weighted' analysis approach that places the greatest

emphasis on individual periods while including data from the prior and subsequent interval. Specifically, for each network, the co-expression data for the period is given a weight of 1 and the periods immediately before and after are given a weight of 0.5. Using this method, we observed essentially identical results in terms of temporal and spatial specificity (Supplemental Figure S2C).

Transmission and *De novo* Association Test (TADA)

Comparison of the burden of de novo LoF mutations detected by whole-exome sequencing (WES) in probands and their matched unaffected siblings shows that a gene with a single proband de novo LoF mutation has a 54% chance of being a true ASD risk gene. By considering further information about the gene's characteristics and other potential risk variants it is possible to refine this 54% estimate on a per-gene basis. He et al. (2013) developed the Transmission And *De novo* Association (TADA) test to integrate WES data from family and case-control studies alongside estimates of gene mutability based on gene size and GC content. TADA has been validated theoretically and by simulations of data on rare, large-effect mutations affecting risk for ASD. TADA p-values from He et al. were updated with the additional *de novo* and inherited variant data from the additional 56 quartets sequenced as part of this study. These updated p-values were utilized in the TADA permutation test.

Period 8-10 MD-CBC network

The period 8-10 MD-CBC network was significantly enriched for pASD genes (9/116 pASD genes, corrected $p=3.5x10^{-4}$ by hypergeometric test, corrected p=0.04 by permutation test), but there are relatively few samples underlying this result: 26 compared to 107 and 140 in the period 3-5 and period 4-6 networks, respectively. Thus, the remainder of the analyses in the main text focused on the networks from period 3-5 and 4-6 in the PFC-MSC cluster. However, we did assess several metrics of this network including, connectivity, and TADA enrichment by the same permutation test used in the main text.

Figure S3A summarizes the period 8-10 MD-CBC network. Visually, this network does not appear highly connected in comparison to the two mid-fetal networks. While significant enrichment of pASD is observed by hypergeometric test and the more stringent permutation test, the connectivity of this network is not significantly enriched (p=0.80) in contrast to the period 3-5 (p=0.03) and period 4-6 (p=0.02) networks. Finally, like the two mid-fetal networks, the set of pASD genes present in the network is enriched for pASD genes with a high probability of association with ASD based on TADA (p=0.01, Figure S3B). However, this result is two orders of magnitude less significant than the corresponding scores for the two mid-fetal networks (Figures 3C-D).

TBR1 KO mRNA-Seq and Data Analysis

Total RNA was isolated from freshly dissected P0 neocortices of $Tbr1^{-/-}$ and $Tbr1^{+/+}$ littermates using a Qiagen RNeasy Mini Kit. Libraries were prepared using an Illumina mRNA-Seq Sample Prep Kit. Amplified cDNA was size-selected at 250 bp and validated using the Agilent Bioanalyzer DNA 1000 system. The final product was subjected to cluster generation using an Illumina Standard Cluster Generation Kit v4. Libraries were

sequenced as single end 74mers using the Illumina Genome Analyzer pipeline, and image analysis (Firecrest module), base-calling (Bustard module), and primary sequence analysis (Gerald module) were performed. Reads were mapped to the mouse reference genome (release mm9) with ELAND software. Filtered reads that were uniquely mapped to exons of Ensembl gene model with up to two mismatches were used for the quantification of gene expression. To detect differentially expressed genes between Tbr1−/− and Tbr1+/+, we designed a three-step stringent filtering process. First, counts of mapped reads within genes were compared using Fisher's exact test (adjusted $P \leq 0.05$). Second, reads per kb of exon model per million reads ($RPKM = 10^9 C/NL$, where C is the number of mappable reads that fall onto the gene's exons, N is the total number of mappable reads in the lane, and L is the sum of the exons in base pairs) were used to quantify gene expression, and the fold change was set as another filtering scale ($|\log2(\text{fold change})| \geq 0.5$). In our series of relative experiments, the genes were finally identified as differentially expressed unless they were shared by all groups.

Animals

All experiments were carried out in accordance with a protocol approved by the Committee on Animal Research at Yale University. *Tbr1* mutant mice were a generous gift from John Rubenstein. The generation of *Tbr1*$^{-/-}$ allele was described previously (Bulfone et al., 1998).

Network analysis with human laminar-specific expression data

Region and layer specific gene expression data was obtained from the BRAINSPAN Prenatal LMD Microarray project. This dataset profiles gene expression in four brains spanning periods 4 to 6 of development (15-21 postconceptional weeks). In each brain, gene expression profiles were assessed for 347 finely laser microdissected tissue samples from subdivisions distributed across cortical and non-cortical regions. For more information on data generation, please see the corresponding technical white paper on the project website (http://www.brainspan.org/).

In order to test laminar location of the two mid-fetal prefrontal and primary motor/somato-sensory cortex (PFC-MSC) networks we assumed they were localized to the brain regions contained within the PFC-MSC area, and therefore, utilized expression data from the LMD microarray corresponding to these anatomical regions only (dorsolateral and ventrolateral prefrontal, orbital frontal, posterior frontal (motor), frontal pole, and primary somatosensory cortical regions). Within each of these regions samples from nine layers were assessed: subpial granular zone, marginal zone, outer cortical plate, inner cortical plate, subplate zone, intermediate zone, outer subventricular zone, inner subventricular zone, and ventricular zone (listed from cortical to ventricular surface).

To assess the relevance of the two networks to each of the nine layers, we recreated the original networks within each layer (based on expression data from samples collected within that layer only) and then assessed their connectivity. More specifically, we determined Pearson's correlation coefficients between network genes connected in the original mid-fetal networks, using expression data from each layer separately. By summing these correlations, we were able to estimate the overall connectivity of the networks layer-by-layer, and then to assess the significance of this connectedness using a permutation test with 100,000.

We calculated normed correlations in order to remove biases that have to do with global changes in connectivity:

Let *i* and *j* denote genes and *k* denote layer. Define r_{ijk} as the correlation between genes *i* and *j* in layer *k*, and define $\bar{r}_k$ as the average correlation, over all pairs of the genes measured in layer *k*. Then define a normed correlation as

$$c_{ijk} = \frac{r_{ijk}}{\bar{r}_k}$$

Connectivity of a network in layer *k* was then calculated by summing c_{ijk} for all pairs in the network.

For each iteration of the permutation test, the 2,952 and 3,389 gene-gene correlations from the original period 3-5 and 4-6 networks, respectively, were kept constant, however the normed correlation value between each of the pairs of genes within a network was chosen randomly from one of the 9 layers for each gene. In other words, we permuted the (*i,j*) connections over the set of *k* 100,000 times in order to get the null distribution of connectivity. This approach was chosen in order to simulate the set of possible sums of a network not enriched in a specific layer.

For subdivision and analysis of the two subnetworks in the period 3-5 PFC-MSC network, we visually separated the whole network into two subnetworks (*ANK2-CHD8-CUL3-DYRK1A-SCN2A-TBR1* and *POGZ-GRIN2B*; corresponding genes are in Table S3). We then performed the same connectivity analysis described above for both sub-networks but using the smaller pool of within subnetwork connections.

Enrichment analysis in mouse neocortical layers across development

Different layers of the mouse cerebral neocortex were microdissected from live tissue sections of the *Dcdc2a*-GFP transgenic reporter mouse (obtained from the GENSAT project (Schmidt et al., 2013)), which expresses GFP selectively in L4 pyramidal and stellate glutamatergic excitatory neurons. Tissue samples were collected from the primary somatosensory area (e.g., equivalent to human SC1 area) of mice brain at postnatal days 4 (P4), P6, P8, P10, P14, and adult (P180 or six months old mice). Based on a mathematical model to translate neurodevelopmental time across mammals (www.translatingtime.net; (Workman et al., 2013)), a brain growth event in the cortex of the mouse at P4 [post-conception (PC) day 22.5] translates to PC day 159, or approximately to the middle of period 6, in the human. Mouse P6 (PC 24.5), P8 (PC 26.5), P10 (PC 28.5), P14 (PC 32.5) and P180 (PC 198.5), translate to PC day 188 (early period 7), 219 (middle period 7), 253 (late period 7), 328 (early period 8), and young adult (period 13), respectively.

We assumed that each layer contains multiple but also unique cell types with identifiable and divergent gene expression profiles that allow us to follow the molecular dynamics of different neuronal and glial populations. For example, the DL are composed by the first born glutamatergic excitatory projection (pyramidal) neurons to arrive to the developing cortical plate, which project axons greatly to subcortical regions. On the other hand, UL are composed by the last born glutamatergic excitatory projection (pyramidal) neurons to

arrive to the developing cortical plate, which for the most part project intracortically (also known as cortico-cortical projections). Lastly, L4 is generally composed by two groups of glutamatergic excitatory neurons: smaller size pyramidal neurons, and spiny stellate cells (also known as granule cells), which specifically within the sensory cortices establish local synaptic connections with neighboring cells.

Details on the microdissection and RNA-seq will be published separately. Briefly, the microdissection of the superficial layers (SL) (i.e., L2/3 including marginal zone or L1, and pia), L4 and deep layers (DL) (i.e., L5/6 including transient subplate zone or adult white matter) was performed on 250 □m thick slices, while maintaining the section submerged in ice-cold ACSF inside a petri dish, using a scalpel and a dissecting needle, under a fluorescence stereoscope (Discovery V8, Zeiss). Total RNA was extracted from each sample (e.g., SL, L4, and DL). The cDNA libraries were prepared using the TruSeq mRNA Sample Prep Kit (Illumina) as per the manufacturer's instructions with some modifications. Briefly, polyA RNA was purified from 1 µg of total RNA using oligo(dT) beads. Quaint-IT RiboGreen RNA Assay Kit (Invitrogen) was used to quantitate purified mRNA with the NanoDrop 3300 (Thermo Scientific). The denatured libraries were diluted to 15 pM, followed by cluster generation on a single-end HiSeq flow cell (v1.5) using an Illumina cBOT, according to the manufacturer's instructions. The HiSeq flow cell was run for 75 cycles using a single-read recipe (v2 sequencing kit) according to the manufacturer's instructions.

Reads passing the default purify filtering of Illumina CASAVA pipeline (released version 1.7) were aligned to mouse reference genome (GRCm38/mm10) using Tophat 2.0.8 (Kim et al., 2013; Kim and Salzberg, 2011; Langmead et al., 2009; Trapnell et al., 2009). Full length reads were used (75 bp). The Cufflinks 2.1.1 program Cuffdiff (Trapnell et al., 2012) identified differentially expressed genes using the mm10 annotation from UCSC. The two mice at each time point were treated as biological replicates.

At each time point (P4, P6, P8, P10, P14, and P180) genes exclusively upregulated in each layer were identified by the following criteria: FPKM ≥10, and for both comparisons with the other two layers, a q-value (FDR) ≤ 0.05, and fold change ≥2. The stringent threshold of FPKM ≥10 was used to identify differentially expressed genes more likely to be important to layer identity.

Genes exclusively upregulated in each layer at each time point were termed marker genes for that layer and time. We then assessed enrichment of the set of markers for each layer and time point using the hypergeometric distribution function 'phyper' in R. At each time point, q = number of marker genes in network; m = total number of marker genes; n = the difference between the number of genes with status 'ok' and expression level in at least one layer ≥10 FPKM, and m; and k = the number of genes in the network.

In this and all downstream analyses, an in-house database of mouse and human orthologs converted mouse genes to human genes in order to assess enrichment in the networks built with human genes.

Cell-type specific markers

Cell-type specific marker genes were derived from a previously published and independent list derived from human expression data (Kang et al., 2011). Kang et al (2011) assessed the BrainSpan exon array developmental transcriptome and determined markers of five cell types: astrocytes, cortical GABA interneurons, cortical glutamatergic neurons, microglia, and oligodendrocytes (Table S6). After observation of significant enrichment of cortical glutamatergic neuron (CGPN) marker genes only, enrichment of marker gene specific to layer (L) 5/6 CGPNs and to L1-4 CGPNs (Table S6) was determined. The significance of the observed enrichment for a particular cell type was calculated by assessing the test set of 100,000 networks for the distribution of marker gene enrichment.

Immunostaining and *in situ* hybridization

Coronal sections from 17-20 PCW human frontal cortex were used to validate the cellular resolution of ASD gene expression. For immunohistochemistry, sections were pre-treated with 0.3% H_2O_2 followed by incubation in blocking buffer at room temperature and then incubation for 24-48 hours at 4°C in primary antibodies. Tissue was then incubated with biotin-labeled secondary antibodies, conjugated with Avidin-Biotin-Peroxidase Complex (Vector Laboratories), and visualized with DAB (Vector Laboratories) following the manufacturer's recommended protocol. Antibodies used were: rabbit anti-TBR1 (1:250, Abcam); rabbit anti-POGZ (1:250, Sigma-Aldrich); rabbit anti-CHD8 (1:500, Novus); and biotin-conjugated donkey anti-rabbit (1:200, Jackson ImmunoResearch). Developed sections were mounted on Superfrost Plus charged slides (Fisher), dehydrated through an ethanol series, cleared with xylenes, and preserved with Permount (Fisher) and glass coverslips. Images were taken using a digital scanner (Aperio ScanScope).

Immunofluorescent staining was performed similarly as aforementioned, without peroxide pre-treatment, and with longer incubations in blocking buffer, primary and secondary antibodies, and intermittent washes. Additional antibodies used were goat anti-FOXP2 (1:200, SantaCruz); mouse anti-SATB2 (1:200, GenWay); rat anti-CTIP2 (1:200, SantaCruz); and flurophore-conjugated secondary antibodies (AlexaFluor from Life Technologies). DAPI was included during incubation with secondary antibodies to counterstain nuclei. After staining, sections were mounted as before and preserved with Aqua Poly/Mount (Polysciences, Inc.) and glass coverslips.

For *in situ* hybridization, brain sections were first mounted on charged slides, followed by post-fixation in 4% paraformaldehyde-PBS and then a 0.1M triethanolamine with 0.25% acetic anhydride pre-treatment. Sections were then incubated overnight at 65°C with 500 ng/ml of digoxigenin (DIG)-labeled cRNA probes corresponding to human *SCN2A* nucleotides 1817-2954 (NM_001040143). Sections were washed and incubated overnight at 4°C with alkaline phosphatase-conjugated anti-DIG antibodies, followed by signal detection using NBT/BCIP chromogen (Roche) diluted in a PVA buffer. Tissue was again post-fixed in paraformaldehyde, then dehydrated and coverslipped as for the DAB-stained sections.

Supplemental References

Albers, C.A., Lunter, G., MacArthur, D.G., McVean, G., Ouwehand, W.H., and Durbin, R. (2011). Dindel: accurate indel calls from short-read data. Genome Res *21*, 961-973.

Bulfone, A., Wang, F., Hevner, R., Anderson, S., Cutforth, T., Chen, S., Meneses, J., Pedersen, R., Axel, R., and Rubenstein, J.L. (1998). An olfactory sensory map develops in the absence of normal projection neurons or GABAergic interneurons. Neuron *21*, 1273-1282.

Harte, R.A., Farrell, C.M., Loveland, J.E., Suner, M.-M., Wilming, L., Aken, B., Barrell, D., Frankish, A., Wallin, C., Searle, S.*, et al.* (2012). Tracking and coordinating an international curation effort for the CCDS Project. Database *2012*.

He, X., Sanders, S.J., Liu, L., De Rubeis, S., Lim, E.T., Sutcliffe, J.S., Schellenberg, G.D., Gibbs, R.A., Daly, M.J., Buxbaum, J.D.*, et al.* (2013). Integrated model of de novo and inherited genetic variants yields greater power to identify risk genes. PLoS Genet *9*, e1003671.

Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.-h., Narzisi, G., Leotta, A.*, et al.* (2012). De Novo Gene Disruptions in Children on the Autistic Spectrum. Neuron *74*, 285-299.

Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M.M., Pletikos, M., Meyer, K.A., Sedmak, G.*, et al.* (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483-489.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biology *14*, R36.

Kim, D., and Salzberg, S. (2011). TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biology *12*, R72.

Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., Wong, W.S.*, et al.* (2012). Rate of de novo mutations and the importance of father's age to disease risk. Nature *488*, 471-475.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, G.P.D.P. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., Lin, C.F., Stevens, C., Wang, L.S., Makarov, V.*, et al.* (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature *485*, 242-245.

O'Roak, B.J., Vives, L., Fu, W., Egertson, J.D., Stanaway, I.B., Phelps, I.G., Carvill, G., Kumar, A., Lee, C., Ankenman, K.*, et al.* (2012a). Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science *338*, 1619-1622.

O'Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P., Levy, R., Ko, A., Lee, C., Smith, J.D.*, et al.* (2012b). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature *485*, 246-250.

Pruitt, K.D., Tatusova, T., and Maglott, D.R. (2007). NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res *35*, D61-65.

Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L.*, et al.* (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature *485*, 237-241.

Schmidt, E.F., Kus, L., Gong, S., and Heintz, N. (2013). BAC transgenic mice and the GENSAT database of engineered mouse strains. Cold Spring Harb Protoc *2013*.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Workman, A.D., Charvet, C.J., Clancy, B., Darlington, R.B., and Finlay, B.L. (2013). Modeling Transformations of Neurodevelopmental Sequences across Mammalian Species. J Neurosci *33*, 7368-7383.