

Spatio-temporal transcriptome of the human brain

Hyo Jung Kang^{1*}, Yuka Imamura Kawasawa^{1*}, Feng Cheng^{1*}, Ying Zhu^{1*}, Xuming Xu^{1*}, Mingfeng Li^{1*}, André M. M. Sousa^{1,2}, Mihovil Pletikos^{1,3}, Kyle A. Meyer¹, Goran Sedmak^{1,3}, Tobias Guennel⁴, Yurae Shin¹, Matthew B. Johnson¹, Željka Krsnik¹, Simone Mayer^{1,5}, Sofia Fertuzinhos¹, Sheila Umlauf⁶, Steven N. Lisgo⁷, Alexander Vortmeyer⁸, Daniel R. Weinberger⁹, Shrikant Mane⁶, Thomas M. Hyde^{9,10}, Anita Huttner⁸, Mark Reimers⁴, Joel E. Kleinman⁹ & Nenad Sestan¹

Brain development and function depend on the precise regulation of gene expression. However, our understanding of the complexity and dynamics of the transcriptome of the human brain is incomplete. Here we report the generation and analysis of exon-level transcriptome and associated genotyping data, representing males and females of different ethnicities, from multiple brain regions and neocortical areas of developing and adult post-mortem human brains. We found that 86 per cent of the genes analysed were expressed, and that 90 per cent of these were differentially regulated at the whole-transcript or exon level across brain regions and/or time. The majority of these spatio-temporal differences were detected before birth, with subsequent increases in the similarity among regional transcriptomes. The transcriptome is organized into distinct co-expression networks, and shows sex-biased gene expression and exon usage. We also profiled trajectories of genes associated with neurobiological categories and diseases, and identified associations between single nucleotide polymorphisms and gene expression. This study provides a comprehensive data set on the human brain transcriptome and insights into the transcriptional foundations of human neurodevelopment.

Human neurodevelopment is a complex and precisely regulated process that occurs over a protracted period of time^{1–3}. Human-specific features of this process are likely to be important factors in the evolution of human specializations^{2–5}. However, in addition to giving us remarkable cognitive and motor abilities, the formation of molecularly distinct and intricate neural circuits may have also increased our susceptibility to certain psychiatric and neurological disorders^{4–9}. Furthermore, sex differences are important in brain development and function, and are a risk factor for conditions such as autism spectrum disorders (ASDs) and depression^{9–13}. Research and progress in all these areas could be enhanced by a comprehensive analysis of the spatio-temporal dynamics of gene expression and transcript variants in the human brain.

Previous transcriptome studies of the developing human brain have used relatively small numbers of samples and predominantly focused on only a few regions or developmental time points^{14–18}. In this Article, we explore the transcriptomes of 16 regions comprising the cerebellar cortex, mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus and 11 areas of the neocortex. The data set was generated from 1,340 tissue samples collected from 57 developing and adult post-mortem brains of clinically unremarkable donors representing males and females of multiple ethnicities.

Study design, data generation and quality control

To investigate the spatio-temporal dynamics of the human brain transcriptome, we created a 15-period system spanning the periods from embryonic development to late adulthood (Table 1 and Supplementary Information, section 2.1). We sampled transient prenatal

structures and immature and mature forms of 16 brain regions, including 11 neocortex (NCX) areas, from multiple specimens per period (Table 2; Supplementary Information, section 2.2; Supplementary Figs 1–3; and Supplementary Table 1). The 11 NCX areas are collectively referred to hereafter as the region NCX. We also genotyped donor DNA using an Illumina 2.5-million single nucleotide polymorphism (SNP) chip (Supplementary Fig. 4 and Supplementary Table 2). Only brains from clinically unremarkable donors with no signs of large-scale genomic abnormalities were included in the study ($N = 57$, including 39 with both hemispheres; age, 5.7 weeks post-conception

Table 1 | Periods of human development and adulthood as defined in this study

Period	Description	Age
1	Embryonic	4 PCW ≤ Age < 8 PCW
2	Early fetal	8 PCW ≤ Age < 10 PCW
3	Early fetal	10 PCW ≤ Age < 13 PCW
4	Early mid-fetal	13 PCW ≤ Age < 16 PCW
5	Early mid-fetal	16 PCW ≤ Age < 19 PCW
6	Late mid-fetal	19 PCW ≤ Age < 24 PCW
7	Late fetal	24 PCW ≤ Age < 38 PCW
8	Neonatal and early infancy	0 M (birth) ≤ Age < 6 M
9	Late infancy	6 M ≤ Age < 12 M
10	Early childhood	1 Y ≤ Age < 6 Y
11	Middle and late childhood	6 Y ≤ Age < 12 Y
12	Adolescence	12 Y ≤ Age < 20 Y
13	Young adulthood	20 Y ≤ Age < 40 Y
14	Middle adulthood	40 Y ≤ Age < 60 Y
15	Late adulthood	60 Y ≤ Age

M, postnatal months; PCW, post-conceptual weeks; Y, postnatal years.

¹Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut 06510, USA. ²Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical Sciences Institute, University of Porto, 4099-003 Porto, Portugal. ³Graduate Program in Neuroscience, Croatian Institute for Brain Research, University of Zagreb School of Medicine, 10000 Zagreb, Croatia. ⁴Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia 23298, USA. ⁵MSc/PhD Molecular Biology Program, International Max Planck Research School for Molecular Biology, 37077 Göttingen, Germany. ⁶Yale Center for Genome Analysis, Yale University School of Medicine, New Haven, Connecticut 06510, USA. ⁷Institute of Genetic Medicine, Newcastle University, International Centre for Life, Newcastle upon Tyne NE1 3BZ, UK. ⁸Department of Pathology, Yale University School of Medicine, New Haven, Connecticut 06510, USA. ⁹Clinical Brain Disorders Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892, USA. ¹⁰The Lieber Institute for Brain Development, Johns Hopkins University Medical Campus, Baltimore, Maryland 21205, USA.

*These authors contributed equally to this work.

Table 2 | Ontology and nomenclature of analysed brain regions and NCX areas

Periods 1 and 2	Periods 3–15
FC, frontal cerebral wall	OFC, orbital prefrontal cortex DFC, dorsolateral prefrontal cortex VFC, ventrolateral prefrontal cortex MFC, medial prefrontal cortex M1C, primary motor (M1) cortex
PC, parietal cerebral wall	S1C, primary somatosensory (S1) cortex IPC, posterior inferior parietal cortex
TC, temporal cerebral wall	A1C, primary auditory (A1) cortex STC, superior temporal cortex ITC, inferior temporal cortex
OC, occipital cerebral wall	V1C, primary visual (V1) cortex
HIP, hippocampal anlage	HIP, hippocampus
—	AMY, amygdala
VF, ventral forebrain MGE, medial ganglionic eminence LGE, lateral ganglionic eminence CGE, caudal ganglionic eminence	STR, striatum
DIE, diencephalon DTH, dorsal thalamus	MD, mediodorsal nucleus of the thalamus —
URL, upper (rostral) rhombic lip	CBC, cerebellar cortex

(PCW) to 82 years; sex, 31 males and 26 females; post-mortem interval, 12.11 ± 8.63 (mean \pm s.d.) hours; pH, 6.45 ± 0.34 (mean \pm s.d.).

Transcriptome profiling was performed using total RNA extracted from a total of 1,340 dissected tissue samples (RNA integrity number, 8.83 ± 0.93 (mean \pm s.d.); Supplementary Tables 3 and 4). We used the Affymetrix GeneChip Human Exon 1.0 ST Array platform, which features comprehensive coverage of the human genome, with 1.4 million probe sets that assay expression across the entire transcript or individual exon, thereby providing redundancy and increased confidence in estimates of gene-level differential expression (DEX, differentially expressed) and differential exon usage (DEU). Descriptions of tissue sampling and quality control measures implemented throughout transcriptome data generation steps are provided in Supplementary Information, sections 2–5, and Supplementary Figs 5–8.

Global transcriptome dynamics

Spatio-temporal gene expression

After quality control assessments and quantile normalization, we summarized core and unique probe sets, representing 17,565 mainly

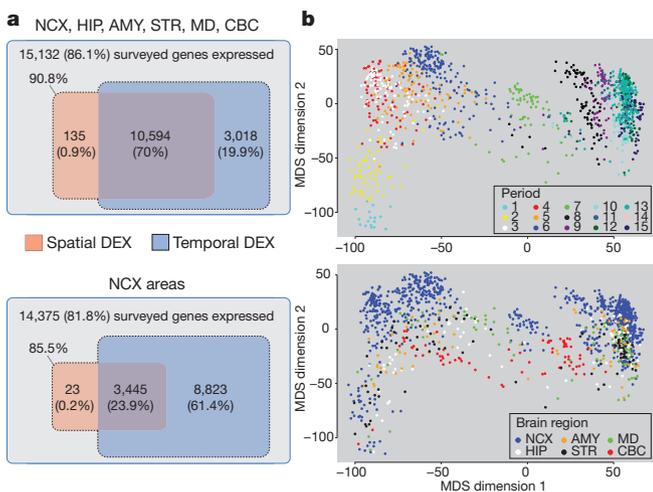


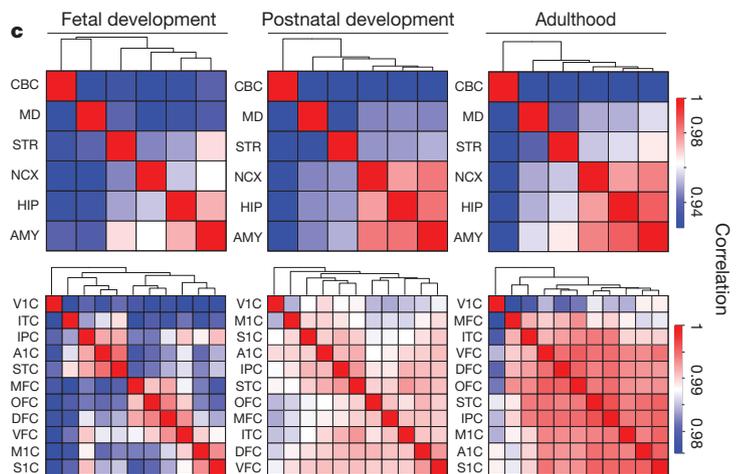
Figure 1 | Global spatio-temporal dynamics of gene expression. **a**, Venn diagrams representing the total number of genes considered to be expressed and the number of spatially and temporally DEX genes for brain regions (top) and NCX areas (bottom). **b**, Multidimensional scaling (MDS) plot showing transcriptional similarity, coloured by period (top) and region (bottom). Non-

protein-coding genes, into gene-level information. Using stringent criteria (\log_2 -transformed signal intensity of ≥ 6 in at least one sample and mean detection-above-background P value of < 0.01 in at least one region of at least one period) to define an ‘expressed’ gene, we found that 15,132 (86.1%) of 17,565 genes surveyed were expressed in at least one brain region during at least one period, and that 14,375 (81.8%) were expressed in at least one NCX area (Fig. 1a, Supplementary Information 6.1 and Supplementary Fig. 9). To investigate the contributions of different factors to the global transcriptome dynamics, we applied multidimensional scaling and principal-component analysis, which revealed that region and age (that is, spatio-temporal dynamics) contribute more to the global differences in gene expression than do other tested variables: sex, ethnicity and inter-individual variation (Fig. 1b; Supplementary Information, sections 6.2 and 6.3; and Supplementary Figs 10 and 11).

To identify genes that were spatially or temporally regulated, we used a conservative threshold (false-discovery-rate Q value of < 0.01 and ≥ 2 -fold \log_2 -transformed signal intensity difference), included post-mortem interval and RNA integrity number as technical covariates within all of our analysis-of-variance models of differential expression, considered the influence of dissection variation and applied a fivefold jackknife procedure (Supplementary Information, section 6.4, and Supplementary Figs 12 and 13). We found that 70.9% of expressed genes were spatially DEX between any two regions within at least one period, and that 24.1% were spatially DEX between any two NCX areas (Fig. 1a). By contrast, 89.9% of expressed genes were temporally DEX between any two periods across regions, and 85.3% were temporally DEX between any two periods across NCX areas. Moreover, 70.0% and 23.9% of expressed genes were both spatially and temporally DEX within brain regions and within NCX areas, respectively. The bulk of spatio-temporal regulation occurred during prenatal development. For instance, 57.7% of NCX-expressed genes were temporally DEX across fetal development (periods 3–7), whereas 9.1% were during postnatal development (periods 8–12) and 0.7% were during adulthood (periods 13–15). Together, these data indicate that the majority of brain-expressed protein-coding genes are temporally and, to a lesser extent, spatially regulated, and that this regulation occurs predominantly during prenatal development.

Transcriptional architecture of the human brain

To assess transcriptional relatedness between brain regions/NCX areas, we calculated correlation matrices of pairwise comparisons



metric; stress = 18.9%. Euclidean distance of \log_2 -transformed signal intensity was used to measure pairwise similarity. **c**, Heat map matrix of pairwise Spearman correlations between brain regions (top) and between NCX areas (bottom) during fetal development (periods 3–7), postnatal development (periods 8–12) and adulthood (periods 13–15).

(Fig. 1c) and performed unsupervised hierarchical clustering across periods 3–15, an interval during which all analysed regions/areas can be consistently followed across time. Among regions, this analysis showed distinct and developmentally regulated clustering of NCX (combination of 11 areas), HIP and AMY, with CBC having the most distinctive transcriptional profile. At the level of NCX areas, clustering formed the following groups during fetal periods: OFC, DFC and MFC; VFC and primary somatomotor cortex (S1C and M1C); and parietal-temporal perisylvian areas (IPC, A1C and STC). VIC had the most distinctive transcriptional profile of NCX areas throughout development and adulthood. The increased correlations between NCX, HIP, AMY and the majority of non-VIC NCX areas with age indicate that transcriptional differences are particularly pronounced during development.

Consistent with the clustering observed, CBC showed the greatest number of region-restricted or region-enriched DEX genes, with 516 (4.8%) of 10,729 genes spatially DEX (Supplementary Information, section 6.4, and Supplementary Table 5). By contrast, the numbers of genes highly enriched in the other regions were lower: NCX, 46 (0.43%); HIP, 48 (0.45%); AMY, 4 (0.04%); STR, 137 (1.28%); MD, 216 (2.01%). The majority of these spatially enriched genes were also temporally regulated, and some, such as those in Supplementary Figs 14 and 15 (NCX: *FLJ32063*, *KCNS1*; HIP: *CDC20B*, *METTL7B*; AMY: *TFAP2D*, *UTS2D*; STR: *C10orf11*, *PTPN7*; MD: *CEACAM21*, *SLC24A5*; CBC: *ESRRB*, *ZP2*), were transiently enriched during a narrow time window. These clustering and region-enrichment results reveal that regional transcriptomes are developmentally regulated and reflect anatomical differences.

Spatio-temporal differential exon usage

Alternative exon usage is an important mechanism for generating transcript diversity^{19,20}. Using a splicing analysis of variance and a splicing index algorithm with conservative criteria ($Q < 0.01$ with a minimum twofold splice index difference between at least two regions/areas or periods; Supplementary Information, section 6.5), we found that 13,647 (90.2%) of 15,132 expressed genes showed DEU across sampled regions (0.1%), periods (19.5%) or both (70.6%). Of 14,375 NCX genes, 88.7% showed DEU across sampled areas ($< 0.01\%$), periods (59.8%) or both (28.9%). The regulation of DEU also varied in time, with the majority of expressed genes (83.0%) showing temporal DEU across fetal development, whereas only 0.9% and 1.4% were temporally regulated across postnatal development and adulthood, respectively.

Focusing on *ANKRD32*, a gene we have previously shown to express an alternative variant in the late mid-fetal frontal cortex¹⁶, we

confirmed and extended our findings on DEU by showing that whereas the longer isoform (*ANKRD32a*) was equally expressed across fetal NCX areas, the shorter isoform (*ANKRD32b*), comprising the last three exons, exhibited dynamic areal patterns. *ANKRD32b* was transiently expressed in a gradient along the anterior–posterior axis of the mid-fetal frontal cortex, with the highest expression in OFC and the lowest in M1C. Before this, *ANKRD32b* was most highly enriched in the ITC and, to a lesser extent, the STC. These spatio-temporal patterns disappeared after birth, when only *ANKRD32a* was expressed, and were not observed in the mouse NCX of equivalent ages (Supplementary Fig. 16 and Supplementary Table 6). These findings illustrate the complexity of DEU in the human brain and demonstrate how specific alternative transcripts can be spatially restricted during a narrow developmental window and with interspecies differences.

Sex differences in the transcriptome

Sex-biased gene expression

Previous studies have identified sexually dimorphic gene expression in the developing and adult human brain^{11–13}. Analysis of our data set using a sliding-window algorithm and t -test model ($Q < 0.01$ with > 2 -fold difference in \log_2 -transformed signal intensity; Supplementary Information, section 6.6) identified 159 genes, including a number of previously reported and newly uncovered genes with male or female bias in expression located on the Y (13 genes), X (9 genes) and autosomal (137 genes) chromosomes. A large fraction (76.7%) had male-biased expression (Fig. 2a and Supplementary Table 7). Notable spatial differences were observed, and more genes had sex-biased expression during prenatal development than during postnatal life, with the adult brain characterized by having the fewest.

Consistent with previous findings^{12,13}, we found that the largest differences were attributable to Y-chromosome genes, especially *PCDH11Y*, *RPS4Y1*, *USP9Y*, *DDX3Y*, *NLGN4Y*, *UTY*, *EIF1AY* and *ZFY*, which showed constant expression across regions and periods, with the exception of *PCDH11Y* downregulation in the postnatal CBC (Fig. 2b). Notably, the functional homologues of these genes on the X chromosome, barring *ZFX* during fetal development (*PCDH11X*, *RPS4X*, *USP9X*, *DDX3X*, *NLGN4X*, *UTX* and *EIF1AX*), were not upregulated in a compensatory manner in female brains (Supplementary Fig. 17).

We also found other X-linked and autosomal genes with sex-biased expression and distinct spatio-temporal patterns, including functionally uncharacterized transcripts (*LOC554203*, *C3orf62*, *FLJ35409* (also known as *MIR137HG*) and *DKFZP58611420*), *S100A10* (which has been linked to depression²¹) and *IGF2* (an imprinted autosomal gene previously implicated in embryonic growth and cognitive function^{22,23}), that showed population-level male-biased expression (Fig. 2c).

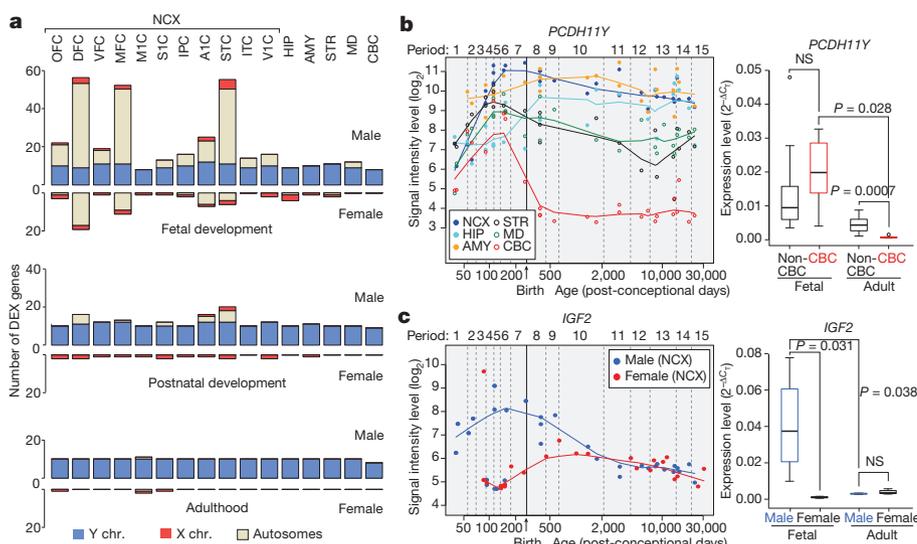


Figure 2 | Sex-biased gene expression.

a, Number of sex-biased DEX genes in brain regions/NCX areas during fetal development (periods 3–7), postnatal development (periods 8–12) and adulthood (periods 13–15). **b**, *PCDH11Y* exon array signal intensity (left) and validation by quantitative PCR with reverse transcription (qRT-PCR) with reverse transcription (qRT-PCR; right) ($N = 5$ male brains per period). **c**, *IGF2* exon array signal intensity (left) and qRT-PCR (right) validation in NCX ($N = 4$ per sex and period). P values were calculated by unpaired t -test. Whiskers indicate fifth and ninety-fifth percentiles, respectively. NS, not significant.

Sex-biased exon usage

We next explored sex-biased DEU using a sliding-window algorithm with a splicing *t*-test model ($Q < 0.01$ and splicing index > 2 ; Supplementary Information, section 6.6). We identified 155 genes (145 autosomal) that showed sex-biased expression of probe sets encoding one or a subset of exons (Supplementary Table 8) in one or multiple regions/NCX areas. These included several members of the collagen family of genes (*COL1A1*, *COL1A2*, *COL3A1*, *COL5A2* and *COL6A3*), *C3*, *KCNH2* (a gene associated with schizophrenia²⁴), *NOTCH3* (a gene mutated in a common form of hereditary stroke disorder²⁵), *ELN* (a gene located within the Williams syndrome critical region²⁶) and *NLGN4X* (an X-chromosome gene implicated in synapse function and associated with ASD and moderate X-linked intellectual disability^{10,27}). Although comparably expressed in males and females at the population and gene levels (Supplementary Fig. 17), *NLGN4X* had a significant male bias in expression of exon 7 and, to a lesser extent, exons 1, 5 and 6 in a developmentally regulated manner (Fig. 3). Together, these findings show that developmentally and spatially regulated differences in gene- and exon-level expression exist between male and female brains.

Gene co-expression networks

To extract additional biological information embedded in the multi-dimensional transcriptome data set, we performed weighted gene co-expression network analysis²⁸, which allowed us to identify modules of co-expressed genes. We identified 29 modules associated with distinct spatio-temporal expression patterns and biological processes (Fig. 4a; Supplementary Information, section 6.7; Supplementary Tables 9–11; and Supplementary Figs 18–20). Among modules corresponding to specific spatio-temporal patterns, M8 consisted of 24 genes with a common developmental trend that showed the highest expression levels in early fetal NCX and HIP (period 3), followed by a progressive decline in expression levels with age until infancy (period 9) (Fig. 4b). By contrast, M15 contained 310 genes showing changes in the opposite direction (relative to those in M8) in the NCX, HIP, AMY and STR (Fig. 4c). Gene ontology enrichment analysis showed that genes in M8 were enriched for gene ontology categories related to neuronal differentiation (Bonferroni-adjusted $P = 7.7 \times 10^{-3}$) and

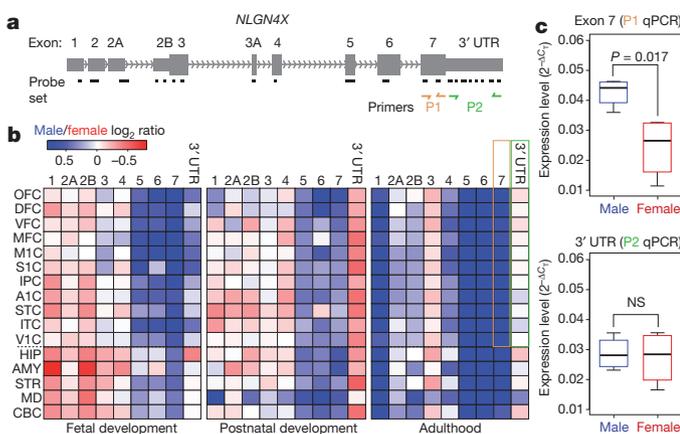


Figure 3 | Sex-biased differential exon usage. **a**, Gene structure and probe set composition of *NLGN4X*. Yellow and green arrows depict primers used for qRT-PCR validation. **b**, Heat map of the \log_2 male/female signal intensity ratio of each exon for fetal development (periods 3–7), postnatal development (periods 8–12) and adulthood (periods 13–15). Differences in expression of exon 7 (yellow frame) and the 3' untranslated region (UTR; green frame) in adult NCX are highlighted. Note that exons 2 and 3A did not meet our expression criteria and are not represented. **c**, qRT-PCR validation of expression of exon 7 and the 3' UTR in adult NCX ($N = 4$ per sex). *P* values were calculated by unpaired *t*-test. Whiskers indicate fifth and ninety-fifth percentiles, respectively.

transcription factors ($P = 5.2 \times 10^{-3}$) (Supplementary Information, section 6.8, and Supplementary Table 9). Conversely, M15 gene ontology categories included ionic channels ($P = 8.0 \times 10^{-8}$) and neuroactive ligand–receptor interaction ($P = 4.0 \times 10^{-14}$).

Genes with the highest degree of connectivity within a module are termed hub genes and are expected to be functionally important within the module. M8 hub genes included transcription factors *TBR1*, *FEZF2*, *FOXP1*, *SATB2*, *NEUROD6* and *EMX1* (Fig. 4b), which have been functionally implicated in the development of NCX and HIP projection neurons^{29–38}. Furthermore, *FOXP1* variants have been linked to Rett syndrome and intellectual disability³⁴. Sequence variants in M15 hub genes (Fig. 4c) have been linked to major depression³⁹ (*GDA*) and to schizophrenia and affective disorders^{6,40} (*NRGN* and *RGS4*).

We also identified two large-scale, temporally regulated modules (M20 and M2) with opposite developmental trajectories of genes co-expressed across regions: expression in M20 gradually decreased with age and expression in M2 gradually increased (Supplementary Figs 21 and 22). M20 was enriched for gene ontology categories related to zinc-finger proteins ($P = 7.3 \times 10^{-48}$) and transcription factors ($P = 4.8 \times 10^{-50}$), including many ZNF and SOX family members. M2 was enriched for gene ontology categories related to membrane proteins ($P = 1.8 \times 10^{-21}$), calcium signalling ($P = 8.1 \times 10^{-10}$), synaptic transmission ($P = 1.6 \times 10^{-6}$) and neuroactive ligand–receptor interaction ($P = 4.1 \times 10^{-4}$), reflecting processes important in postnatal brain maturation. Their hub genes encoded transcriptional factors, modulators of chromatin state and signal transduction proteins, all of which are likely to be involved in driving the co-expression networks. Drastic expression shifts in M20 and M2 in the opposite direction just before birth indicate that this period is associated with global transcriptional changes that probably reflect environmental influences on brain development and intrinsic changes in cellular composition and functional processes.

Expression trajectories of neurodevelopment

One important use for the generated data set is to gain insight into normal and abnormal human neurodevelopment by analysing trajectories of individual genes or groups of genes associated with a particular neurobiological category or disease. To test this strategy, we compared our expression data for *DCX* (a gene expressed in neuronal progenitor cells and immature migrating neurons), as well as for genes associated with dendrite (*MAP1A*, *MAPT*, *CAMK2A*) and synapse (*SYN*, *SYPL1*, *SYPL2*, *SYN1*) development, with independently generated, non-transcriptome human data sets. The *DCX* expression trajectory was remarkably reminiscent of the reported changes in the density of *DCX*-immunopositive cells in the postnatal human HIP^{36,40} ($r = 0.946$, Pearson correlation; Fig. 5a). In our transcriptome data set, *DCX* expression increased until early mid-fetal development (period 5) and then gradually declined with age until early childhood (period 10). Likewise, expression trajectories of dendrite and synapse development gene groups closely paralleled the growth of basal dendrites of DFC pyramidal neurons⁴¹ ($r = 0.810$ for layer 3 and $r = 0.700$ for layer 5; Fig. 5b) and DFC synaptogenesis⁴² ($r = 0.940$; Fig. 5c), respectively. Steep increases in both processes occurred between the late mid-fetal period and late infancy, indicating that a considerable portion of these two processes occurs before birth and reaches a plateau around late infancy.

After demonstrating the accuracy and viability of using the data set to profile human neurodevelopment, we manually curated lists of genes associated with over 80 categories, including various neurodevelopmental processes, neural cell types and neurotransmitter systems (Supplementary Information 6.9 and Supplementary Table 12). Notable trajectories and differences in their onset times, rates of increase and decrease, and shapes were observed within and between brain regions for categories including major neurodevelopmental processes (neural cell proliferation and migration, dendrite and synapse development, and myelination; Fig. 5d), cortical GABAergic inhibitory interneurons (*CALB1*, *CALB2*, *NOS1*, *PVALB* and *VIP*) and

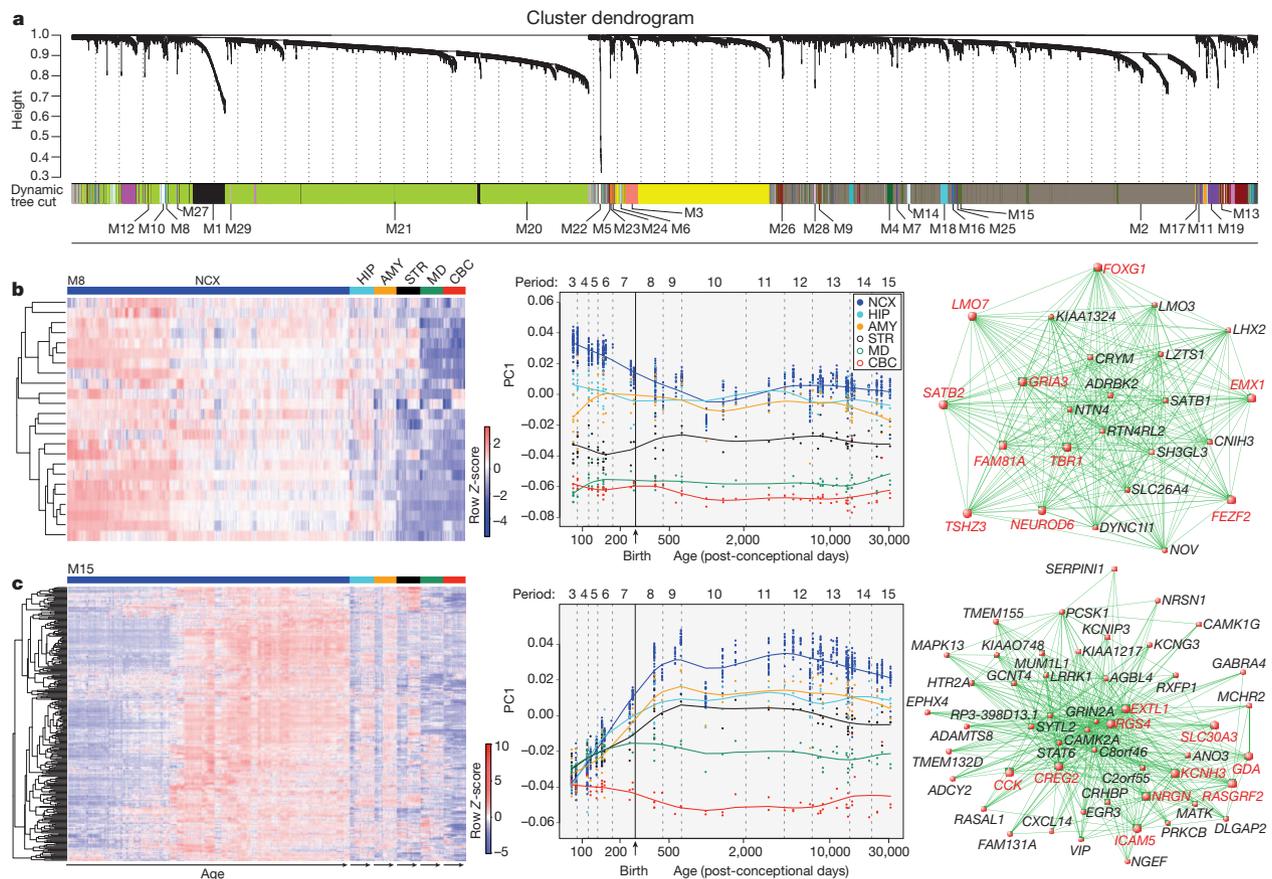


Figure 4 | Global co-expression networks and gene modules. **a**, Dendrogram from gene co-expression network analysis of samples from periods 3–15. Modules of co-expressed genes were assigned a colour and number (M1 to M29). **b**, Left: heat map of genes in M8 showing the spatio-temporal expression pattern after hierarchical clustering. The expression values for each gene are arranged in the heat map, ordered first by brain region, then by age and last by

NCX area. Middle: spatio-temporal pattern of M8 summarized by the first principal component (PC1) for expression of genes in the module across age. Right: 24 M8 genes; the top ten hub genes are shown in red. **c**, Same analyses as in **b**, but for M15; the top 50 genes defined by the highest intramodular connectivity are shown in right panel. Results for other modules are available in Supplementary Information.

glutamate receptors (Supplementary Figs 23 and 24). Two expected patterns were observed in neurodevelopmental trajectories: changes in expression of cell proliferation genes preceded the increase in expression of *DCX*, and expression of each decreased during perinatal development whereas synapse development, dendrite development and myelination trajectories increased. Notably, the NCX trajectory for synapse development did not drastically decline during late childhood or adolescence (Fig. 5c, d) as previously reported for synapse density⁴². We also identified co-expression network modules and additional genes that are highly correlated with the categories (Supplementary Tables 10, 13 and 14). For example, M20 and M2 were strongly correlated with neuron migration ($r = 0.894$) and myelination ($r = 0.972$), respectively.

In addition, our data set enabled us to generate expression trajectories of genes commonly associated with ASD and schizophrenia. We investigated a number of genes previously linked to these disorders (Supplementary Information, section 6.10) and observed distinct and dynamic expression patterns, especially among NCX areas (Supplementary Fig. 25 shows examples for *CNTNAP2*, *MET*, *NLGN4X* and *NRGN*). To gain insight into potential biological functions of ASD- and schizophrenia-associated genes in human neurodevelopment, we identified other genes with significantly correlated spatio-temporal expression profiles and performed gene ontology enrichment analysis (Supplementary Tables 15 and 16). These findings reveal associated spatio-temporal differences in these expression trajectories and provide additional co-expressed genes that can be interrogated for their role in the respective processes or disorders.

Expression quantitative trait loci

Previous studies have identified expression quantitative trait loci (eQTLs) in the adult human brain, primarily in the cerebral cortex^{43–47}. Our multiregional developmental data set enabled us to search for association between SNP genotypes and spatio-temporal gene expression. We tested only for *cis*-eQTLs, restricting the search to SNPs within 10 kilobases of either a transcription start site or a transcription end site, as opposed to *trans*-eQTLs, which would require much larger sample sizes.

Implementing a conservative strategy (gene-wide Bonferroni correction followed by genome-wide $Q < 0.1$; Supplementary Information, section 9), we identified 39 NCX, eight HIP, four AMY, two STR, six MD and five CBC genes (Supplementary Table 17) with evidence of *cis*-eQTL, including two previously reported genes^{45,47} (*ITGB3BP* and *ANKRD27*). Consistent with previous studies⁴⁸, associated SNPs were enriched near transcription start and termination sites (Fig. 6a, b).

An example of a significant association in NCX, MD and CBC is that between SNP rs10785190 and *GLIPR1L2*, a member of the glioma pathogenesis-related 1 family of genes⁴⁹. The expression differences were observed at the level of the whole transcript and exons 1 and 2, the only exons we observed to be expressed at appreciable levels in the NCX (Fig. 6c, d). The NCX probably had more *cis*-eQTLs than other regions owing to its smaller variation in gene expression resulting from the averaged expression of 11 areas. Many eQTLs identified as significant in NCX also have similar associations in other regions, although they were not statistically significant after the conservative genome-wide correction (Supplementary Table 17). Thus, we have

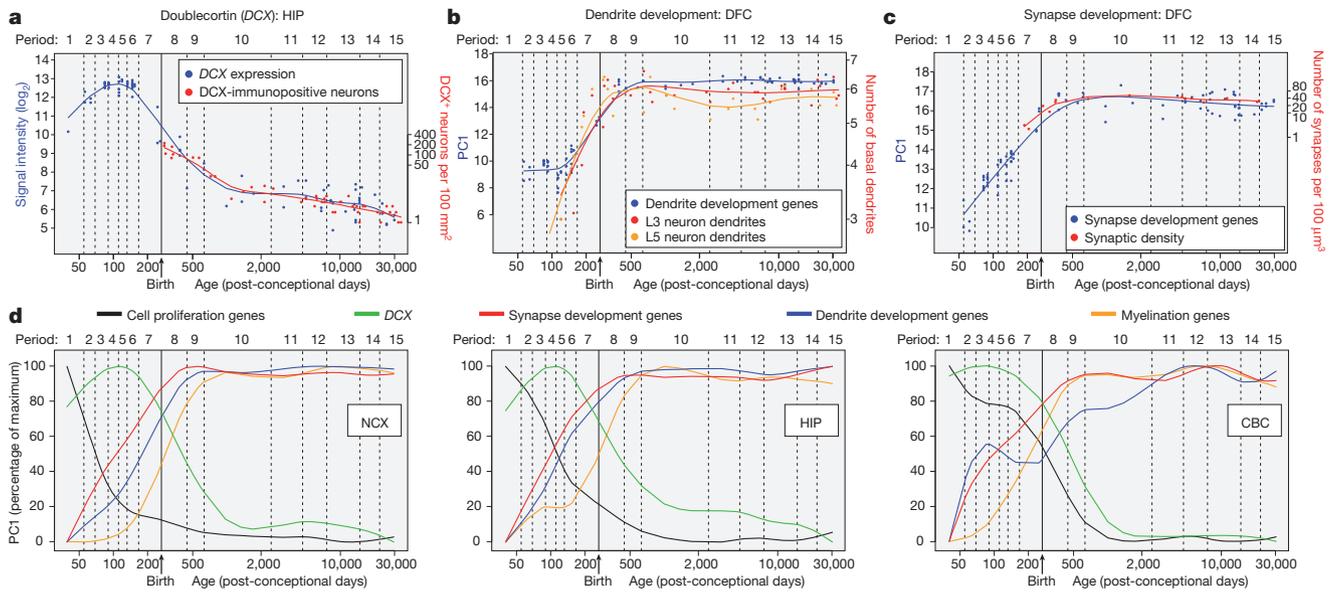


Figure 5 | Trajectories of genes associated with neurodevelopmental processes. **a**, Comparison between *DCX* expression in HIP and the density of *DCX*-immunopositive cells in the human dentate gyrus³⁶. **b**, Comparison between transcriptome-based dendrite development trajectory in DFC and Golgi-method-based growth of basal dendrites of layer 3 (L3) and 5 (L5) pyramidal neurons in the human DFC⁴¹. **c**, Comparison between transcriptome-based synapse development trajectory in DFC and density of

DFC synapses calculated using electron microscopy⁴². For **b** and **c**, PC1 for gene expression was plotted against age to represent the developmental trajectory of genes associated with dendrite (**b**) or synapse (**c**) development. Independent data sets were centred, scaled and plotted on a logarithmic scale. **d**, PC1 value for the indicated sets of genes (expressed as percentage of maximum) plotted against age to represent general trends and regional differences in several neurodevelopmental processes in NCX, HIP and CBC.

identified polymorphic regulators of transcription in different regions across development, potentially providing insights into inter-individual differences and genetic control of the brain transcriptome.

Discussion

Our analysis reveals several features of the human brain transcriptome, and increases our knowledge of the transcriptional events in human neurodevelopment. We show that gene expression and exon usage have complex and dynamically regulated patterns, some of which may not be evident in the transcriptomes of commonly studied model organisms. Moreover, these patterns differ more prominently across time and space than they do between sexes, ethnicities or individuals, despite their underlying genetic differences. Transcriptome differences between males and females also included several disease-related genes, offering possible mechanisms underlying the sex differences in the incidence, prevalence and severity of some brain disorders. We also found that some of the inter-individual variations in the regional and developmental transcriptomes were associated with specific SNP genotypes, which may have altered expression-regulating elements. Thus, the present data set (available at <http://www.humanbraintranscriptome.org>), along with an accompanying study⁵⁰, provides a basis for a variety of further investigations and comparisons with other transcriptome-related data sets of both healthy and diseased states.

Although our study has uncovered many intricacies in gene expression and exon usage in the human brain, there are potential limitations of our study that warrant discussion. Foremost, we used stringent criteria to minimize false positives and faithfully characterize general transcriptional patterns, rather than to capture all the changes that may occur. Also, we analysed dissected tissue that contained multiple cell types, thus diluting the transcriptional contribution and dynamic range of expression of any one specific cell type. Current limitations prevent us from using cell-type-specific approaches in systematically analysing the spatio-temporal transcriptome. Furthermore, the number of brains and regions analysed so far is not sufficient to investigate the full magnitude of transcriptional changes or the full range of eQTLs. Application of sequencing technology will allow more in-depth analyses of the transcriptome, and aid in discovery of novel or

low-expressing transcripts. Finally, although specific patterns of expression are often linked to specialized biological processes, it is important to remember that the relationship between messenger RNA and protein levels is not always linear nor translated into apparent phenotypic differences. As these concerns are addressed in future

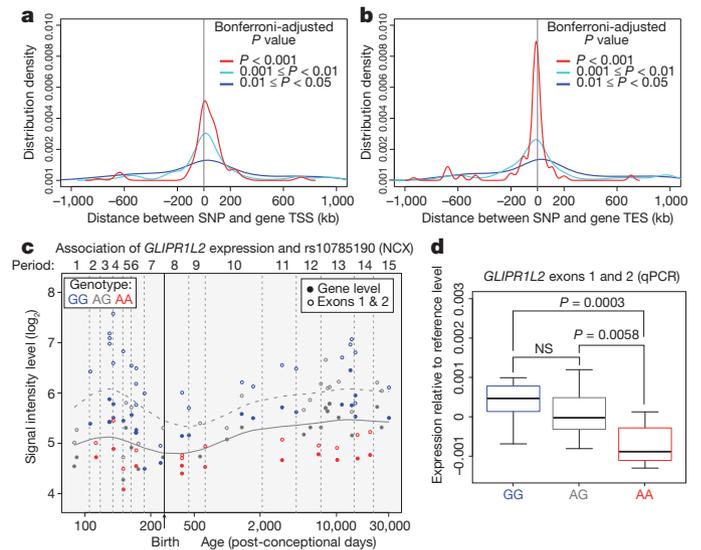


Figure 6 | Association between SNPs and gene expression. **a**, SNP distribution around transcription start sites (TSS); **a** and transcription end sites (TES); **b**) of the associated genes, based on several Wald test *P*-value cut-offs after gene-wide Bonferroni correction. **c**, *GLIPR1L2* expression association with rs10785190 genotype, a SNP located in exon 1. The solid and dashed curves, calculated from locally weighted scatter-plot smoothing (LOWESS), show the developmental trends of gene expression and exon-1 and exon-2 expression, respectively. **d**, qRT-PCR validation of exon-1 and exon-2 expression in NCX for each genotype (*N* = 14 GG, 14 AG, 8 AA), plotted relative to the LOWESS curve in **c** to facilitate comparison across developmental periods. *P* values were calculated by unpaired *t*-test. Whiskers indicate fifth and ninety-fifth percentiles, respectively.

with more samples and new data sets from human and non-human primate brains, it will be possible to uncover deeper insights into the transcriptional foundations of human brain development and evolution.

METHODS SUMMARY

Supplementary Information, sections 3–9, provides a full description of tissue acquisition and processing, data generation, validation and analyses.

Received 11 December 2010; accepted 30 August 2011.

- Kostovic, I. & Judas, M. Prolonged coexistence of transient and permanent circuitry elements in the developing cerebral cortex of fetuses and preterm infants. *Dev. Med. Child Neurol.* **48**, 388–393 (2006).
- Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nature Rev. Neurosci.* **10**, 724–735 (2009).
- Rubenstein, J. L. Annual Research Review: Development of the cerebral cortex: implications for neurodevelopmental disorders. *J. Child Psychol. Psychiatry* **52**, 339–355 (2011).
- Preuss, T., Cáceres, M., Oldham, M. & Geschwind, D. Human brain evolution: insights from microarrays. *Nature Rev. Genet.* **5**, 850–860 (2004).
- Hill, R. S. & Walsh, C. A. Molecular insights into human brain evolution. *Nature* **437**, 64–67 (2005).
- Lewis, D. A. & Levitt, P. Schizophrenia as a disorder of neurodevelopment. *Annu. Rev. Neurosci.* **25**, 409–432 (2002).
- Meyer-Lindenberg, A. & Weinberger, D. R. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Rev. Neurosci.* **7**, 818–827 (2006).
- Insel, T. Rethinking schizophrenia. *Nature* **468**, 187–193 (2010).
- State, M. The genetics of child psychiatric disorders: focus on autism and Tourette syndrome. *Neuron* **68**, 254–269 (2010).
- Jamain, S. *et al.* Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nature Genet.* **34**, 27–29 (2003).
- Vawter, M. *et al.* Gender-specific gene expression in post-mortem human brain: localization to sex chromosomes. *Neuropsychopharmacology* **29**, 373–384 (2004).
- Weickert, C. *et al.* Transcriptome analysis of male-female differences in prefrontal cortical development. *Mol. Psychiatry* **14**, 558–561 (2009).
- Reinius, B. & Jazin, E. Prenatal sex differences in the human brain. *Mol. Psychiatry* **14**, 988–989 (2009).
- Abrahams, B. *et al.* Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc. Natl Acad. Sci. USA* **104**, 17849–17854 (2007).
- Sun, T. *et al.* Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science* **308**, 1794–1798 (2005).
- Johnson, M. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494–509 (2009).
- Somel, M. *et al.* MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res.* **20**, 1207–1218 (2010).
- Ip, B. *et al.* Investigating gradients of gene expression involved in early human cortical development. *J. Anat.* **217**, 300–311 (2010).
- Licatalosi, D. D. & Darnell, R. B. Splicing regulation in neurologic disease. *Neuron* **52**, 93–101 (2006).
- Blencowe, B. J. Alternative splicing: new insights from global analyses. *Cell* **126**, 37–47 (2006).
- Svenningsson, P. *et al.* Alterations in 5-HT_{1B} receptor function by p11 in depression-like states. *Science* **311**, 77–80 (2006).
- Chen, D. Y. *et al.* A critical role for IGF-II in memory consolidation and enhancement. *Nature* **469**, 491–497 (2011).
- Lehtinen, M. K. *et al.* The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. *Neuron* **69**, 893–905 (2011).
- Huffaker, S. J. *et al.* A primate-specific, brain isoform of KCNH2 affects cortical physiology, cognition, neuronal repolarization and risk of schizophrenia. *Nature Med.* **15**, 509–518 (2009).
- Joutel, A. *et al.* Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* **383**, 707–710 (1996).
- Ewart, A. K. *et al.* Hemizyosity at the elastin locus in a developmental disorder, Williams syndrome. *Nature Genet.* **5**, 11–16 (1993).
- Südhof, T. C. Neuroligins and neuexins link synaptic function to cognitive disease. *Nature* **455**, 903–911 (2008).
- Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, article17 (2005).
- Hansen, D. V., Rubenstein, J. L. & Kriegstein, A. R. Deriving excitatory neurons of the neocortex from pluripotent stem cells. *Neuron* **70**, 645–660 (2011).
- Heyner, R. *et al.* Tbr1 regulates differentiation of the preplate and layer 6. *Neuron* **29**, 353–366 (2001).
- Molyneaux, B. J., Arlotta, P., Hirata, T., Hibi, M. & Macklis, J. D. Fez1 is required for the birth and specification of corticospinal motor neurons. *Neuron* **47**, 817–831 (2005).
- Chen, B., Schaevitz, L. & McConnell, S. Fez1 regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. *Proc. Natl Acad. Sci. USA* **102**, 17184–17189 (2005).
- Chen, J., Rasin, M., Kwan, K. & Sestan, N. Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. *Proc. Natl Acad. Sci. USA* **102**, 17792–17797 (2005).
- Ariani, F. *et al.* FOXP1 is responsible for the congenital variant of Rett syndrome. *Am. J. Hum. Genet.* **83**, 89–93 (2008).
- Kwan, K. *et al.* SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. *Proc. Natl Acad. Sci. USA* **105**, 16021–16026 (2008).
- Knott, R. *et al.* Murine features of neurogenesis in the human hippocampus across the lifespan from 0 to 100 years. *PLoS ONE* **5**, e8809 (2010).
- Han, W. *et al.* TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract. *Proc. Natl Acad. Sci. USA* **108**, 3041–3046 (2011).
- McKenna, W. L. *et al.* Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development. *J. Neurosci.* **31**, 549–564 (2011).
- Perroud, N. *et al.* Genome-wide association study of increasing suicidal ideation during antidepressant treatment in the GENDEP project. *Pharmacogenomics J.* advance online publication, (<http://dx.doi.org/10.1038/tpj.2010.70>) (2010).
- Stefansson, H. *et al.* Common variants conferring risk of schizophrenia. *Nature* **460**, 744–747 (2009).
- Petanjek, Z., Judas, M., Kostović, I. & Uylings, H. Lifespan alterations of basal dendritic trees of pyramidal neurons in the human prefrontal cortex: a layer-specific pattern. *Cereb. Cortex* **18**, 915–929 (2008).
- Huttenlocher, P. R. & Dabholkar, A. S. Regional differences in synaptogenesis in human cerebral cortex. *J. Comp. Neurol.* **387**, 167–178 (1997).
- Stranger, B. E. *et al.* Population genomics of human gene expression. *Nature Genet.* **39**, 1217–1224 (2007).
- Heinzen, E. L. *et al.* Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol.* **6**, e1 (2008).
- Liu, C. *et al.* Whole-genome association mapping of gene expression in the human prefrontal cortex. *Mol. Psychiatry* **15**, 779–784 (2010).
- Gibbs, J. R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6**, e1000952 (2010).
- Myers, A. J. *et al.* A survey of genetic human cortical gene expression. *Nature Genet.* **39**, 1494–1499 (2007).
- Webster, J. A. *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet.* **84**, 445–458 (2009).
- Ren, C., Ren, C. H., Li, L., Goltsov, A. A. & Thompson, T. C. Identification and characterization of RTVP1/GLIPR1-like genes, a novel p53 target gene cluster. *Genomics* **88**, 163–172 (2006).
- Colantuoni, C. *et al.* Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* doi:10.1038/nature10524 (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Belanger, V. Imamovic, R. Johnson, P. Larton, S. Lindsay, B. Poulos, J. Rajan, D. Rimm and R. Zielke for assistance with tissue acquisition, D. Singh for technical assistance, I. Kostovic and Z. Petanjek for dendrite measurements, P. Levitt for suggesting the inclusion of ITC in the study, and D. Karolchik and A. Zweig for help in creating tracks for the UCSC Genome Browser. We also thank A. Beckel-Mitchener, M. Freund, N. Gerstein, D. Geschwind, T. Insel, M. Judas, J. Knowles, E. Lein, P. Levitt, N. Parikhshak and members of the Sestan laboratory for discussions and criticism. Tissue was obtained from several sources including the Human Fetal Tissue Repository at the Albert Einstein College of Medicine, the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, the Laboratory of Developmental Biology at the University of Washington (supported by grant HD000836 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development) and the Joint MRC/Wellcome Trust Human Developmental Biology Resource (<http://hdb.org>) at the IGH, Newcastle Upon Tyne (UK funding awards G0700089 and GR082557). Support for predoctoral fellowships was provided by the China Scholarship Council (Y.Z.), the Portuguese Foundation for Science and Technology (A.M.M.S.), the Samsung Scholarship Foundation (Y.S.), a Fellowship of the German Academic Exchange Service – DAAD (S. Mayer) and NIDA grant DA026119 (T.G.). This work was supported by grants from the US National Institutes of Health (MH081896, MH089929, NS054273), the Kavli Foundation and NARSAD, and by a James S. McDonnell Foundation Scholar Award (N.S.).

Author Contributions H.J.K., Y.I.K., A.M.M.S., M.P., K.A.M., G.S., Y.S., M.B.J., Z.K., S. Mayer, S.F., S.U., S. Mane and N.S. performed and analysed the experiments. F.C., Y.Z., X.X., M.L., T.G. and M.R. analysed the data. Z.K., A.M.M.S., M.P., G.S., S.N.L., A.V., D.R.W., T.M.H., A.H., J.E.K. and N.S. participated in tissue procurement and examination. N.S. designed the study and wrote the manuscript, which all authors commented and edited.

Author Information Exon array have been deposited in the NCBI Gene Expression Omnibus under accession number GSE25219 and genotyping data have been deposited in the NCBI database of Genotypes and Phenotypes under accession number phs000406.v1.p1. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to N.S. (nenad.sestan@yale.edu).