**The architecture and evolution of human and non-human primate brain transcriptomes**

**André M. M. Sousa1,2,3,\*, Ying Zhu1,2,\*, Robert R. Kitchen4,5,\*, Kyle A. Meyer1,2, Mingfeng Li1,2, Yuka Imamura Kawasawa1,2,6, Marta Mele7, Tiago Carvalho7, Andrew T. N. Tebbenkamp1,2, Mihovil Pletikos1,2, John J. Ely8, Patrick R. Hof9, Mark Reimers10, Richard P. Lifton11,12, Shrikant M. Mane13, James P. Noonan2,11, Matthew W. State14, Ed S. Lein15, James A. Knowles16, Tomas Marques-Bonet7,17,18, Chet C. Sherwood19, Mark B. Gerstein4, Nenad Šestan1,2,\***

1Department of Neurobiology, Yale School of Medicine, New Haven, Connecticut, USA

2Kavli Institute for Neuroscience, Yale School of Medicine, New Haven, Connecticut, USA

3Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical Sciences Institute, University of Porto, Porto, Portugal

4Program in Computational Biology and Bioinformatics, Departments of Molecular Biophysics and Biochemistry and Computer Science, Yale University, New Haven, Connecticut, USA 5Department of Psychiatry, Connecticut, New Haven, Connecticut, USA

6Department of Pharmacology, Penn State University College of Medicine, Hershey, Pennsylvania, USA

7Institut de Biologia Evolutiva, CSIC-Universitat Pompeu Fabra, PRBB, Doctor Aiguader 88, Barcelona, Catalonia, Spain

8Alamogordo Primate Facility, Holloman Air Force Base, New Mexico, USA

9Fishberg Department of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA

10Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia, USA

11Department of Genetics, Yale School of Medicine, New Haven, Connecticut, USA

12Howard Hughes Medical Institute, Yale University, Connecticut, USA

13Yale Center for Genomic Analysis, Yale School of Medicine, New Haven, Connecticut, USA

14Department of Psychiatry and Langley Porter Psychiatric Institute, University of California, San Francisco, San Francisco, California, USA

15Allen Institute for Brain Science, Seattle, Washington, USA

16Department of Psychiatry and Zilkha Neurogenetic Institute, Keck School of Medicine, University of Southern California, Los Angeles, California, USA

17Institució Catalana de Recerca i Estudis Avançats, Barcelona, Catalonia, Spain.

18Centro Nacional de Analisis Genomico, Barcelona, Catalonia, Spain

19Department of Anthropology, The George Washington University, Washington, District of Columbia, USA.

\*Co-first author

Correspondence should be addressed to N.S. (nenad.sestan@yale.edu)

**ABSTRACT**

**Phenotypic differences between human and non-human primate brains have been described at structural and functional levels, but the underlying transcriptional events remain elusive. Here, we deep sequenced mRNAs and small non-coding RNAs from 16 brain regions, including multiple neocortical areas of humans, chimpanzees, and rhesus macaques for comparative analyses. We found substantial species differences, including 19.2% of expressed mRNAs and 17.5% of expressed miRNAs displaying human-specific differential expression. Analyses of regional profiles revealed that the striatum had the highest number of human-specific changes. mRNA and miRNA were grouped into species-specific co-expression modules, including a module with human-specific upregulation of genes involved in the dopamine biosynthesis pathway. We also found that transcriptional patterns for glutamate and GABA neurotransmitter receptors are more conserved than for the neuromodulatory systems. This study provides a comprehensive resource on the human and non-human primate brain transcriptomes and insights into species-specific transcriptional programs with potential phenotypic differences.**

**INTRODUCTION**

The genetic differences between humans and closely related non-human primates (NHP), especially chimpanzees, have been characterized as relatively small1,2. Yet, there are immense phenotypic differences, especially in cognitive and behavioral capabilities3-9, indicating that unraveling the functionally meaningful genotype-to-phenotype correlations is particularly challenging. The transcriptome stands at the interface between the genome and downstream phenotypes, and its interrogation provides insights into the dynamics of biological processes and the emergence of novel phenotypic traits. Thus, critical questions in human evolutionary biology concern how the human brain and its transcriptional activity differ from those of NHP10,11.

The human brain is, on average, three times larger than that of our closest living relatives, the great apes3-9. However, despite this large difference and the concomitant expansion of certain brain regions, in particular the association areas of the cerebral cortex, there is limited evidence that any of them are unique to human8,9. Rather, multiple lines of evidence suggest that some of the brain regions have undergone microstructural and functional changes during human evolution12-14. Nevertheless, the transcriptional changes underlying these differences remain elusive.

Comparative transcriptome analyses of human and NHP tissues are rooted in the idea that changes in transcriptional profiles, more than coding sequences, underlie many phenotypic differences between humans and NHP10. Although considerable progress has been made in profiling gene expression in human and NHP brains15-18, a systematic survey of both coding and non-coding transcripts across many brain regions of high-quality post-mortem human and NHP brains is needed. Studies of the human brain transcriptome19-22 have revealed substantial regional differences, further emphasizing the importance of analyzing a wide range of brain regions in NHP. Here, we expand on the existing studies by using RNA-seq, the agnostic nature of which allows direct comparison of transcriptomes of different species, to profile coding and non-coding transcripts of 16 regions, including eleven areas of the neocortex (NCX), hippocampus (HIP), amygdala (AMY), striatum (STR), mediodorsal nucleus of the thalamus (MD), and cerebellar cortex (CBC), of adult post-mortem human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), and rhesus macaque (*Macaca mulatta*) brains. The analysis of these data allowed us to put transcriptional regional diversity into an evolutionary context.

**RESULTS**

**Study design**

We used RNA-seq to survey the transcriptomes of 16 brain regions (Table 1) from six humans (3 males and 3 females; 31.2 ± 7.8 years old; postmortem interval (PMI) = 16.6 ± 7.7 hours; tissue pH = 6.62 ± 0.26; RNA integrity number (RIN) = 8.5 ± 0.6), five chimpanzees (3 males and 2 females; 27.0 ± 4.0 years old; postmortem interval (PMI) < 5 hours; tissue pH = 6.44 ± 0.06; RNA integrity number (RIN) = 8.84 ± 0.88), and five rhesus macaques (2 males and 3 females; 9.4 ± 2.2 years old; PMI < 2 hours; pH = 6.94 ± 0.09; RIN = 8.81 ± 1.02) (Supplementary Table 1). Poly-A enriched mRNA-seq was carried out on a total of 256 samples (Supplementary Table 1). Consistent and stringent quality control measures were applied at the level of tissue acquisition and dissection, sample preparation, control spike-in RNAs, sequencing, and data processing (see Supplementary Fig. 1 and Online Methods for full description). After these quality control procedures, 4.6 billion uniquely mapped reads (Supplementary Fig. 1a,b) from 247 samples were analyzed. Small-RNA sequencing was performed on exactly the same set of samples, producing 1.01 billion processed sequences that could be uniquely mapped to microRNAs (miRNAs; Supplementary Fig. 1d).

To minimize biases in comparative data analyses due to the quality of gene annotation for the three species, we developed a computational framework named XSAnno (see Online Methods) to create a common annotation set of 26,514 1:1:1 orthologous mRNAs in the three species, which included 16,531 protein-coding genes, 3,253 long noncoding RNAs (lncRNAs), and other transcript categories (see Supplementary Fig. 2 and Online Methods). In the small RNA dataset, we focused on miRNAs. To overcome the limitations of the chimpanzee and macaque miRNA annotations, which are less comprehensive than those available for human, we performed a re-annotation of all chimpanzee and macaque miRNAs based on the full suite of annotated human precursor sequences (see Supplementary Fig. 2c and Online Methods). Both mRNA and miRNA datasets were filtered based on RNA quality, sequencing error, and the number of uniquely mapped reads and analyzed by hierarchical clustering (Online Methods, and Supplementary Fig. 1e,f).

Principal component (PC) analysis was used to examine the contributions of various factors to transcriptional differences. For both the mRNA and miRNA data, the first two PC separated the samples by species. The third PC of miRNA separated the cerebellar cortex from other brain regions, while mRNA samples were not clustered by brain region in the space of the first three PC (Fig. 1). This analysis revealed that species differences contribute more to the global transcriptional differences than regions or inter-individual variation.

**Inter-species differential expression**

We first focused on genes that had expression differences among species for each of the 16 brain regions. For each region, differentially expressed (DEX) mRNAs and miRNAs were identified by comparing generalized linear models, with species as the main factor and batch as a cofactor. Of the expressed mRNA and miRNA orthologs considered in the present analyses, we found 6,389 (38.9%) mRNAs and 603 (40.6%) miRNAs that were DEX between at least two species in one or more regions. Based on *post hoc* comparisons (Supplementary Table 2), DEX mRNAs and miRNAs were classified by the relationship among species, for example human (H) and chimpanzee (C) expression levels similar to each other and greater than macaque (M) levels: H = C > M (Fig. 2a and 2f). As expected, based on evolutionary relationships, the two most frequent patterns of DEX mRNAs and miRNAs had similar expression levels in human and chimpanzee samples, which together differed from the expression levels in macaque samples (H = C > M and H = C < M).

This comparison between species also revealed substantial differences in the number of DEX genes in different brain regions, highlighting the importance of a multiregional analysis. For the 3,154 (19.2%) mRNAs with human-specific up- or downregulation (H > C = M or H < C = M), STR, followed by MD and certain neocortical areas, such as V1C and DFC, exhibited the highest number of human-specific DEX genes, whereas A1C, M1C, CBC, and HIP exhibited the lowest number (Fig. 2b). There was no obvious difference in within-species variation across all the analyzed brain regions, which indicates that the differences in DEX numbers are not likely to be caused by technical bias. Functional annotation of the genes with human-specific DEX (see Online Methods) revealed enrichment for categories such as “cell adhesion”, “plasma membrane part”, “extracellular matrix”, “glycosylation site:N-linked (GlcNAc...)”, and neurobiological processes like “neuron projection”, “dendrite”, and “synaptic transmission” (Bonferroni-adjusted P ≤ 2.79x10-5; Supplementary Table 4). This finding is consistent with previous studies implicating the evolution of these regions and NCX areas in the emergence of some of the most distinctively human aspects of cognition, and behavior3,5,9,23.

While there has been conflicting evidence of a human-specific increase in the number of up-regulated mRNAs24-27, in our study, 15 of the 16 regions show more genes with human-specific upregulation than downregulation, with human STR being the only region with more downregulated than upregulated genes. Alternatively, chimpanzee displayed an equal number of regions showing up- and downregulation (Supplementary Fig. 3a), while only 2 regions had more upregulated genes in macaque, with 14 showing more downregulated genes (Supplementary Fig. 3b). This result indicates that there is a global trend towards upregulation of gene expression in the evolution of the human brain.

Among the mRNAs with human-specific up- or downregulation, we found 31 that were DEX across all analyzed regions (Supplementary Table 3). In contrast, only 3 genes were significantly DEX (H > C = M or H < C = M) exclusively in all eleven NCX areas: *TWIST1, RP11-364P22.1* and *CTB-78F1.1*. *TWIST1* (Fig. 2c) encodes a transcription factor mutated in Saethre–Chotzen syndrome, a disorder associated with multi-organ congenital abnormalities and intellectual disability28. TWIST1 also inhibits the function of the MEF2 transcription factor 29. According to a recent study30, MEF2Ais likely a key regulator of a module of co-expressed genes involved in synaptogenesis. Prompted by these reports, we validated human-specific *TWIST1* expression by droplet digital PCR (ddPCR) and analyzed its cellular expression pattern in the human NCX by immunohistochemistry. This analysis revealed that TWIST1 protein is localized to the nuclei of NCX neurons (Supplementary Fig. 4a,b), which is consistent with the possibility that the protein plays an upstream, regulatory role in neuronal and synaptic function in the human brain. In contrast to *TWIST1*, *RP11-364P22.1*,and *CTB-78F1.1* encode putative non-coding RNAs whose functions have not yet been characterized.

We also found several genes that displayed human-specific DEX in a subset of NCX areas (Supplementary Table 3). For example, the expression of *MET*, a gene implicated in autism spectrum disorders31, is enriched in the human prefrontal cortex (Fig. 2d). There were also several genes that show human-specific DEX in only one brain region (Supplementary Table 3), such as *ZP2* (Fig. 2e), which encodes a protein that mediates sperm-egg recognition32 and is specifically enriched in the human CBC. The human- and CBC-specific expression pattern was confirmed by ddPCR and immunohistochemistry revealed that *ZP2* is predominantly expressed in cerebellar granular cells (Supplementary Fig. 4c,d). However, the function of ZP2 in human granular cells is unknown.

Of the 202 miRNAs exhibiting statistically significant and more than 2-fold human-specific differential expression, 155 were upregulated and 47 were downregulated in the human brain (Fig. 2f). These include miRNAs previously reported as being more abundant in human, for example miR-941, which has been implicated in neurotransmitter signaling33. However, many have not been previously reported to have species-specific expression profiles. We used ddPCR to validate three miRNAs (miR-4284-5p, miR-31, and miR-107) based on their expression profiles (Fig. 2g and Supplementary Fig. 4e,f). miR-4284-5p is consistently expressed across all human brain regions but is consistently non-detected in all chimpanzee and macaque samples. The results of the ddPCR are in agreement with the RNA-seq, showing a human-specific upregulation of miR-4284-5p in all the analyzed brain regions (Fig. 2g).

Whereas miR-4284-5p was highly expressed in humans, miR-31 was downregulated specifically in human telencephalic regions (i.e., NCX, HIP, AMY and STR), and miR-107 was more abundant in macaque and slightly less abundant in chimpanzee, compared to human (Supplementary Fig. 4e,f). miR-31 has repeatedly been implicated in the repression of several mRNAs that are involved in tumour progression, including metastasis34, but no neural-specific function has yet been reported. In contrast, miR-107 has been reported to regulate several transcripts important for neural function and diseases such as *NOTCH2* in gliomas35, granulin, which is mutated in frontotemporal dementia36, BACE1, which plays an important role in the pathogenesis of Alzheimer’s disease37, as well as the global miRNA biosynthesis pathway itself by targeting *DICER*38.

**Species-specific regional transcriptional architecture**

In addition to species differences at the level of individual brain regions, we investigated how the regional transcriptional architecture differed among species. For each species, we determined the number of DEX genes between regions and found a smaller number of DEX genes in human brains regions, with macaque showing the highest number. This pattern was especially pronounced in NCX areas. Analysis of within-region variation indicated this higher number of DEX genes might partly result from smaller within-region variation in gene expression among macaque samples (Supplementary Fig. 5a,c). To test if this smaller within-region variation in macaque samples was not due to smaller genetic distances between rhesus macaques, we calculated genetic distances inferred from single-nucleotide polymorphisms in mRNA expressed in the dorsolateral prefrontal cortex (DFC) samples (see Online Methods and Supplementary Fig. 5d). As expected, based on whole genome data2,39, this analysis revealed that the human specimens exhibited the least intra-species variation while macaques had the greatest, indicating that the smaller within-region variation in gene expression among macaque samples was not due to a smaller genetic distance caused by inbreeding of macaques.

We also calculated pairwise correlation matrices of brain regions, using the union of the regional DEX genes in each species, and performed hierarchical clustering (Fig. 3). Comparing the clusters among species revealed several differences in the transcriptional correlation of brain regions. As expected20,26, the CBC was the most divergent area of the brain in all species, followed by the STR and the MD; HIP and AMY also clustered in the three species, thus clearly separating NCX areas from non-NCX regions. NCX areas were more transcriptionally similar in humans than NCX areas were in either chimpanzees or macaques (Fig. 3). Among these areas, several species differences in clustering were present. Interestingly, the human clustering is more similar to the macaque clustering, with a slight difference in the most distinct NCX area (V1C in human and MFC in macaque), than it is to the chimpanzee clustering. We confirmed the robustness of the clustering by multiscale bootstrap resampling (Supplementary Fig. 4). The miRNA clustering pattern of all the analyzed brain regions was very similar in the three species and, as seen in the mRNA clustering, CBC was the most distinct region, while NCX areas were highly correlated with each other.

**Species differences in co-expression of intronic miRNAs and their host mRNA**

Given the large number of samples in our dataset, we were able to confidently assess the extent to which certain miRNAs that lie within introns multi-exon genes are positively correlated, and may therefore be co-expressed, with their ‘host’ gene (see Online Methods). Following the liftover of human precursor miRNAs to the chimpanzee and macaque genomes, of the 26,514 orthologous mRNAs and lncRNAs analyzed, 814, 777, and 756 contain at least one miRNA within one of their introns in human, chimpanzee, and macaque, respectively (614, 586, and 574, respectively, are expressed in our datasets; Supplementary Table 5). Just under half (274) of the expressed intronic miRNAs in our data were also DEX in at least one species and 81 of these exhibited strong positive correlation with their host gene (Pearson correlation coefficient > 0.5). Six of these 274 miRNAs are intronic to genes that serve exclusively as hosts for their contained miRNAs, and most of these show consistently strong positive correlation in all three species.

We noticed that a striking number of important neuronal and signalling genes, such as *PDE2A, HTR2C, CLCN5, CALCR, CAMK1D****,*** *GABRA3, MAP2K4,* and *KALRN*, exhibit very strong positive correlation with their host miRNA(s), suggesting that related pathways may also be modified between species through differences in mRNA-host miRNA co-expression relationship (Supplementary Table 5). One miRNA, *mir-767*, lies within an intron of *GABRA3* and shows reduced expression and reduced correlation in human compared to both chimpanzee and macaque. Both miRNAs intronic to *CALCR* (*mir-489* and *mir-653*) are more abundant in macaque and appear more strongly expressed in the AMY and MD than any other tissue. Both mature products of *mir-139*, which are intronic and positively correlated to *PDE2A*, are very abundant in all three species, more than 2-fold depleted in MD, and more than 8-fold depleted in CBC compared to other tissues. Two of the three mature miRNAs intronic to *HTR2C* (miR-1264and miR-1912) are not highly expressed, but show consistently greater abundance in HIP, AMY, STR, and MD compared to the other tissues, except miR-1912, which is not expressed in macaque (Supplementary Table 5).

**Species differences in gene co-expression modules**

As the DEX analysis focused on single genes and on a limited number of expression patterns, we next applied weighted gene co-expression correlation network analysis (WGCNA; see Online Methods) to identify clusters (modules) of highly co-expressed genes by searching for genes with similar patterns of variation across samples and species. We identified 229 modules of co-expressed genes, many with distinct regional and/or species expression patterns (Supplementary Table 6; see http://medicine.yale.edu/lab/sestan/resources/index.aspx). We then clustered all miRNAs based on their individual Pearson correlations to the average mRNA module expression profile across all regions and species (Supplementary Fig. 6 and Supplementary Table 7; see Online Methods). Using this approach, we identified 24 stable clusters of miRNAs (Fig. 4a), which we further refined using HiTS-CLIP brain data obtained from Boudreau et al., 201440. The inclusion of the HiTS-CLIP targets greatly improved the specificity of the miRNA/mRNA module assignments based on those pairs with significant target enrichment (Fig. 4b and Supplementary Table 8). We also found that several miRNA/mRNA modules exhibited species-specific HiTS-CLIP target enrichment (Fig. 4c). Furthermore, each miRNA module averaged expression across the 16 regions exhibited distinct profiles highlighting the widespread species- and/or tissue-specific behavior (Fig. 4d).

Modules with gene expression patterns that are shared across all three species include a module (M) containing mRNAs that are enriched in NCX areas (M33) and some miRNA clusters (miRNA M8, miRNA M11, and miRNA M31) with reduced expression in NCX regions (Fig. 4 and Supplementary Table 6). The evolutionarily conserved modules, such as M33, are enriched for categories such as “Homeobox, conserved site” (Bonferroni-adjusted P = 0.02) and “transcription factor” (P = 0.023). Hub genes in this module include *TBR1* and *SATB2*, two genes that have previously been shown to be critical for NCX development41. We also found examples of region-specific networks across all species, such as M16 enriched in CBC and M129 in MD (Supplementary Table 6; see http://medicine.yale.edu/lab/sestan/resources/index.aspx for visualization of expression patterns of all modules). Genes in M129 are enriched in “cell-cell signaling” (P = 0.04). Genes in M16 are associated with “transcription factor activity” (Bonferroni-adjusted P = 2.9x10-4), “DNA-binding” (P = 3.9x10-4), and “developmental protein” (P = 8.5x10-4). Genes in M16 also anti-correlate very strongly with miRNAs in cluster miRNA M10, which includes miR-137, recently associated with schizophrenia in a GWAS study42, and *miR-222-3p* andmiR-221-3p, which both include the Fragile X syndrome gene *FMR1* amongst predicted targets43.

In contrast to these modules, other mRNA modules were distinguished by species differences. Among human-specific modules, M81, M162, M192, and M229 were upregulated in all human brain regions (Supplementary Table 6). Genes in M192 were associated with “oxidative phosphorylation” (P = 0.03), and in M229 with “ribosome” (P = 1.2x10-3). We also found modules that show human-specific downregulation in all regions, such as M69 and M173 (Supplementary Table 6). No significant pathways were identified for M81, M69, and M173 after Bonferroni adjustment.

Interestingly, we found several miRNA/mRNA module pairs that highlight the extensive species- and region-specific expression differences in these data. For example mRNA M34, which contains genes that are generally down-regulated in the macaque, we find that the miRNAs in miRNA M29 exhibit the opposite species-level expression profile and are significantly enriched for HiTS-CLIP targets. Another example of a region-specific miRNA/mRNA pair are mRNA M53 and miRNA M24 which exhibit opposite regional expression profiles, notably in the V1C, MD, and CBC and are also enriched for HiTS-CLIP targets (Supplementary Table 8).

Some modules that showed species differences in only a subset of regions, such as M92 and M215, consisted of genes selectively upregulated in human NCX regions (Fig. 5a and 5d; Supplementary Table 6). In contrast, M32 is composed of genes, including *TWIST1*, that are upregulated in chimpanzee and macaque, but not human, cortical areas (Fig. 5b and 5e). M130 is composed of genes that were upregulated in human AMY, HIP, and STR, which is the region with the most interspecies DEX genes (Fig. 5c and 5f). Interestingly, two genes in this module are involved in the synthesis of dopamine: *TH* and *DDC*. Due to this finding, we carried out a more detailed analysis of genes involved in dopamine metabolism.

**Human-specific expression patterns of dopamine synthesis enzymes**

Dopamine is synthetized from L-tyrosine by two enzymes in consecutive steps: tyrosine hydroxylase (TH) converts L-tyrosine into L-dihydroxyphenylalanine (L-DOPA), and then, DOPA decarboxylase (DDC) converts L-DOPA into dopamine (Fig. 6a). We found that *TH* and *DDC* showed higher expression levels in both mRNA (Benjamini-Hochberg-adjusted P = 3.5×10-4 and 1.4×10-3, respectively) and protein in the STR of humans compared to chimpanzees and macaques (Fig. 6b-e). Furthermore, the genes encoding two of the enzymes involved in dopamine degradation, *COMT* and *MAOA*, did not show any significant expression differences between species (Supplementary Fig. 7a-d). Together, these findings indicate that humans may have higher levels of dopamine in STR, because humans produce more dopamine synthesis enzymes, which is then exposed to the same level of degrading enzymes as in the other species. Dopamine is involved in several cognitive processes, especially working memory, reasoning, and overall intelligence44, all of which show significant differences in humans. It is, therefore, possible that these differences might be partly explained by this putative increase in dopamine levels. Surprisingly, we found that 3 genes that encode dopamine receptors – *DRD1*, *DRD2*, and *DRD3* – were downregulated in the human STR (Supplementary Fig. 7e-h), suggesting that a homeostatic plasticity may exist in the expression of genes associated with dopamine signaling.

**Evolution of transcriptional patterns for neurotransmitter receptors**

Although differences in neurotransmitter levels may alter physiological properties of neural circuits, the composition of receptors is also essential. We generated co-expression networks between neuroreceptor subunits and measured the conservation of these networks among the three species. A Z-statistic and empirical p-values were calculated to evaluate the preservation of the network based on permutation tests (see Online Methods). We limited further analysis to four systems that involved enough genes to provide reliable comparisons: glutamatergic, GABAergic, serotoninergic, and cholinergic (Fig. 7a-d). The major excitatory and inhibitory neurotransmitter systems, namely the glutamatergic and the GABAergic systems, have similar transcriptional correlation networks among species. Hierarchical clustering of receptor subunits revealed conserved co-expression of *GABRA1*, *GABRB2*,and *GABRG2*, representing the most common assembly of GABAA receptors45 (α1β2γ2; Fig. 7g). Similarly, *GRIA2* and *GRIA3*, subunits of AMPA receptors, were co-expressed in all three species (Fig. 7h). This conservation is unsurprising given that these two systems are ubiquitous and have a vital role in brain function. On the other hand, among the neuromodulatory systems analyzed, the cholinergic and the serotoninergic systems have different networks among the three species: human and chimpanzee serotoninergic and cholinergic networks are different from macaque (Fig. 7e and 7f). The role of serotonin in cognition and its involvement in some neurological diseases, such as schizophrenia, are well reported46. Acetylcholine has also been implicated in the processes of learning and memory, as well as neurodegenerative diseases47. Importantly, interactions between these two neuromodulatory systems seem to play a critical role in cognitive processes48, reinforcing the idea that the differences in the receptor networks of these two neuromodulatory systems revealed in our study might be relevant to species differences in cognition.

**DISCUSSION**

Here, we profiled the adult human, chimpanzee, and rhesus macaque brain transcriptomes by deep sequencing both mRNAs and small RNAs in sixteen regions. All generated data are publicly available in the Non-Human Primate Brain Transcriptome database (http://medicine.yale.edu/lab/sestan/resources/index.aspx), providing the basis for a variety of future investigations and comparisons with other genomic-level data sets.

Our analysis revealed several novel differences among species, including many that are human-specific. However, in order to provide valuable insights from this data, we have had to overcome some limitations associated with our samples. Stringent criteria in all the analyses were applied in order to minimize false positives. This may have precluded us from detecting more subtle, but potentially important, differences. Also, all of these analyses were performed in postmortem specimens with different ante- and postmortem conditions, which could alter RNA integrity. Although RNA-seq allowed for more in-depth analysis of the transcriptome, including some low-expressing transcripts and their associated regional patterns, it also presented some challenges. We had to address differences in the quality of genome annotation to accurately calculate differential expression between species by using a new pipeline – XSAnno. Finally, while we made attempts to link transcriptional patterns with specific biological processes, it is important to note that differences in mRNA levels do not necessarily correspond to differences at the protein level, and often may not have functional consequences49. As these concerns are addressed in the future, it will be possible to uncover additional insights into the transcriptional foundations of human and NHP brain evolution. .

Despite these caveats, we found substantial differences in the number of number of genes exhibiting species-specific expression across the analyzed brain regions. In particular, the STR, followed by the MD and NCX areas such as V1C and DFC, was the region with the highest number of human-specific DEX genes. This is in contrast to previous reports suggesting that the prefrontal NCX17 was the region with the highest number of human-specific DEX genes. While the greatest number of species-specific differences in gene expression were found in these structures, the effects of these transcriptional changes are likely to have much broader functional and phenotypic consequences, due to the extensive connectivity between regions. For example, the MD projects mainly to the prefrontal cortex and transcriptional changes in this structure can, therefore, affect specific NCX areas. Similar to previous reports24-26, we also found that a slightly higher number of DEX genes with human-specific expression exhibited upregulation, a pattern not seen in chimpanzee or macaques.

Second, our analyses revealed that each species has a distinct regional transcriptional brain architecture. The multi-regional study design allowed us to specifically analyze species differences in intra-species clustering of regions based on DEX mRNA and miRNA genes. Interestingly, the human NCX areas appear to be more transcriptionally similar than chimpanzee and macaque NCX. However, we observed smaller within-region variation in gene expression between macaque samples than between both human and chimpanzee samples, which may have increased the statistical power of differential expression analysis in macaque. The inter-species differences between brain regions and variation within brain regions may result from differences in environment and lifestyle, as well as ante- and post-mortem factors.

Third, we also found several human-specific modules of co-expressed genes, including two modules containing hub genes coding for essential enzymes of the dopamine biosynthesis pathway. Since the genes involved in the anabolic pathway were upregulated and the genes involved in the catabolic pathway did not show statistically different expression, we hypothesize that humans may have higher levels of dopamine than the other species analyzed. However, this effect was accompanied by a downregulation of the genes coding the dopamine receptors D1, D2, and D3. It will be important to analyze the protein levels of the receptors to confirm that humans have lower amounts of these receptors in the STR. If confirmed, this downregulation of dopamine receptors might balance the upregulation of dopamine biosynthesis. Nevertheless, further studies on this balance between neurotransmitter and their receptors will be needed.

Fourth, we extended our study to other neurotransmitter systems through the analysis of co-expression networks of genes encoding neurotransmitter receptors. We found that both the glutamatergic and the GABAergic systems were well conserved among species, whereas the serotoninergic and cholinergic systems were not. We reveal that the organization of the correlation networks for these neuromodulators show some differences among humans, chimpanzees, and macaques. These differences might be relevant to the differences among species in cognition and behavior. Moreover, some of the human-specific DEX mRNAs and miRNAs revealed by our data have previously been functionally linked to neurological and psychiatric disorders. Together, these data highlight differences that are likely crucial for both diseases specific to the human lineage50, as well as some aspects of species-specific phenotypes.

In summary, this study provides a unique resource on human, chimpanzee, and macaque regional brain transcriptomes and new insights into species- and human-specific transcriptional programs that may underlie important phenotypic differences.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Accession codes.** Human sequencing data are freely accessible and available at dbGAP (accession number: phs000731.v1.p1). Chimpanzee and macaque sequencing data are freely accessible and available at NCBI BioProjects (accession number: PRJNA236446).

**ACKNOWLEDGEMENTS**

We thank A. Bauernfeind, M. Horn, B. Wicinski, G. Terwilliger, and S. Wilson for assistance with tissue acquisition and processing, D. Singh for technical assistance with preparing sequencing libraries, and the Alamogordo Primate Facility for providing chimpanzee tissue. We also thank members of the Yale University Biomedical High Performance Computing Center and Yale Center for Genome Analysis. Support for predoctoral fellowships was provided by the Portuguese Foundation for Science and Technology (A.M.M.S.), the China Scholarship Council (Y.Z.), and National Science Foundation Graduate Research Fellowship under Grant DGE-1122492 (K.A.M). This project was supported by grants MH089929 and MH081896 from the National Institute of Mental Health (N.S.), Kavli Foundation, James S. McDonnell Foundation Scholar Award (P.R.H., C.S., N.S.), NARSAD Distinguished Investigator Award (N.S.), and the Foster-Davis Foundation (N.S.), and ERC Starting Grant (260372) and MICINN BFU2011-28549 (T.M.-B.).

**AUTHOR CONTRIBUTIONS**

A.M.M.S., M.P., J.J.E., C.C.S., P.R.H., and N.S. managed tissue procurement and examination, A.M.M.S., M.P., and N.S. contributed to tissue and sample processing, Y.Z., R.R.K., M.L., M.M., T.C., M.R., and T.M-B, contributed to sequencing data analysis and interpretation. A.M.M.S., K.A.M., Y.I.K., and A.T.N.T. contributed to validation data generation, analysis, and interpretation, J.A.K., E.S.L., S.M.M., J.N.P., M.B.G., M.W.S., N.S., and R.P.L. contributed to overall project design and management. A.M.M.S., Y.Z., and N.S. designed the study. A.M.M.S., Y.Z., R.R.K. and N.S. wrote the manuscript with input from all of the other authors.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.

**REFERENCES**

1. Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69-87 (2005).

2. Rhesus Macaque Genome Sequencing and Analysis Consortium *et al.* Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**, 222-234 (2007).

3. Hill, R. S. & Walsh, C. A. Molecular insights into human brain evolution. *Nature* **437**, 64-67 (2005).

4. Kennedy, H. & Dehay, C. Self-organization and interareal networks in the primate cortex. *Prog. Brain Res.* **195**, 341-360 (2012).

5. Lui, J. H., Hansen, D. V. & Kriegstein, A. R. Development and evolution of the human neocortex. *Cell* **146**, 18-36 (2011).

6. Gazzaniga, M. S. *Human: The science behind what makes us unique*. (Harper Perennial, New York, 2009).

7. Sherwood, C. C., Subiaul, F. & Zawidzki, T. W. A natural history of the human mind: tracing evolutionary changes in brain and cognition. *J Anat* **212**, 426-454 (2008).

8. Preuss, T. M. The human brain: rewired and running hot. *Ann N Y Acad Sci* **1225 Suppl 1**, E182-E191 (2011).

9. Geschwind, D. H. & Rakic, P. Cortical evolution: judge the brain by its cover. *Neuron* **80**, 633-647 (2013).

10. King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107-116 (1975).

11. Carroll, S. B. Genetics and the making of Homo sapiens. *Nature* **422**, 849-857 (2003).

12. Preuss, T. M., Qi, H. & Kaas, J. H. Distinctive compartmental organization of human primary visual cortex. *Proc Natl Acad Sci U S A* **96**, 11601-11606 (1999).

13. Rilling, J. K. *et al.* The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat Neurosci* **11**, 426-428 (2008).

14. Semendeferi, K. *et al.* Spatial organization of neurons in the frontal pole sets humans apart from great apes. *Cereb Cortex* **21**, 1485-1497 (2011).

15. Khaitovich, P. *et al.* Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309**, 1850-1854 (2005).

16. Babbitt, C. C. *et al.* Both noncoding and protein-coding RNAs contribute to gene expression evolution in the primate brain. *Genome Biol Evol* **2**, 67-79 (2010).

17. Konopka, G. *et al.* Human-specific transcriptional networks in the brain. *Neuron* **75**, 601-617 (2012).

18. Somel, M. *et al.* Transcriptional neoteny in the human brain. *Proc Natl Acad Sci U S A* **106**, 5743-5748 (2009).

19. Johnson, M. B. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494-509 (2009).

20. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483-489 (2011).

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

22. Pletikos, M. *et al.* Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron* **81**, 321-332 (2014).

23. Lieberman, P. On the nature and evolution of the neural bases of human language. *Am J Phys Anthropol* **Suppl 35**, 36-62 (2002).

24. Caceres, M. *et al.* Elevated gene expression levels distinguish human from non-human primate brains. *Proc Natl Acad Sci U S A* **100**, 13030-13035 (2003).

25. Gu, J. & Gu, X. Induced gene expression in human brain after the split from chimpanzee. *Trends Genet* **19**, 63-65 (2003).

26. Khaitovich, P. *et al.* Regional patterns of gene expression in human and chimpanzee brains. *Genome Res* **14**, 1462-1473 (2004).

27. Uddin, M. *et al.* Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles. *Proc Natl Acad Sci U S A* **101**, 2957-2962 (2004).

28. Howard, T. D. *et al.* Mutations in TWIST, a basic helix-loop-helix transcription factor, in Saethre-Chotzen syndrome. *Nat Genet* **15**, 36-41 (1997).

29. Hamamori, Y., Wu, H. Y., Sartorelli, V. & Kedes, L. The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist. *Mol Cell Biol* **17**, 6563-6573 (1997).

30. Liu, X. *et al.* Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques. *Genome Res* **22**: 611-622 (2012).

31. Campbell, D. B. *et al.* A genetic variant that disrupts MET transcription is associated with autism. *Proc Natl Acad Sci U S A* **103**, 16834-16839 (2006).

32. Hoodbhoy, T. & Dean, J. Insights into the molecular basis of sperm-egg recognition in mammals. *Reproduction* **127**, 417-422 (2004).

33. Hu, H. Y. *et al.* Evolution of the human-specific microRNA miR-941. *Nat Commun* **3**, 1145, doi:10.1038/ncomms2146 (2012).

34. Valastyan, S. *et al.* A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. *Cell* **137**, 1032-1046 (2009).

35. Chen, L. *et al.* MicroRNA-107 inhibits glioma cell migration and invasion by modulating Notch2 expression. *J Neuro-Oncol* **112**, 59-66 (2013).

36. Wang, W. X. *et al.* miR-107 regulates granulin/progranulin with implications for traumatic brain injury and neurodegenerative disease. *Am J Pathol* **177**, 334-345 (2010).

37. Wang, W. X. *et al.* The expression of microRNA miR-107 decreases early in Alzheimer's disease and may accelerate disease progression through regulation of beta-site amyloid precursor protein-cleaving enzyme 1. *J Neurosci* **28**, 1213-1223 (2008).

38. Martello, G. *et al.* A MicroRNA Targeting Dicer for Metastasis Control. *Cell* **141**, 1195-1207 (2010).

39. Prado-Martinez, J. *et al.* Great ape genetic diversity and population history. *Nature* **499**, 471-475 (2013).

40. Boudreau, R. L. et al. Transcriptome-wide discovery of microRNA binding sites in human brain. *Neuron* **81**, 294-305 (2014).

41. Kwan, K. Y., Sestan, N. & Anton, E. S. Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development* **139**, 1535-1546, doi:10.1242/dev.069963 (2012).

42. Kim, A. H. *et al.* Experimental validation of candidate schizophrenia gene ZNF804A as target for hsa-miR-137. *Schizophr Res* **141**, 60-64 (2012).

43. Zongaro, S. *et al.* The 3' UTR of FMR1 mRNA is a target of miR-101, miR-129-5p and miR-221: implications for the molecular pathology of FXTAS at the synapse. *Hum Mol Genet* **22**, 1971-1982 (2013).

44. Nieoullon, A. Dopamine and the regulation of cognition and attention. *Prog Neurobiol* **67**, 53-83 (2002).

45. Sigel, E. & Steinmann, M. E. Structure, function, and modulation of GABA(A) receptors. *J Biol Chem* **287**, 40224-40231 (2012).

46. Roth, B. L., Hanizavareh, S. M. & Blum, A. E. Serotonin receptors represent highly favorable molecular targets for cognitive enhancement in schizophrenia and other disorders. *Psychopharmacology* **174**, 17-24 (2004).

47. Hasselmo, M. E. & Sarter, M. Modes and models of forebrain cholinergic neuromodulation of cognition. *Neuropsychopharmacology* **36**, 52-73 (2011).

48. Steckler, T. & Sahgal, A. The role of serotonergic-cholinergic interactions in the mediation of cognitive behaviour. *Behav Brain Res* **67**, 165-199 (1995).

49. Khan, Z. *et al.* Primate transcript and protein expression levels evolve under compensatory selection pressures. *Science* **342**, 1100-1104 (2013).

50. Varki, A. Nothing in medicine makes sense, except in the light of evolution. *J Mol Med* **90**, 481-494 (2012).

**Figure Legends**

**Figure 1** Transcriptome-based mRNA and miRNA clustering clearly discriminates samples by species. Three-dimensional plot of the top three Principal Components based on (**a**, **b**) mRNA expression or (**c**, **d**) miRNA expression reveal distinct separation of samples based on (**a**, **c**) species and (**b**, **d**) brain region, especially between subcortical structures. Each point represents one sample. The percentage of variation that is explained by each Principal Component is shown.

**Figure 2** Interspecies differential expression. (**a**) Heat map showing the number of mRNA genes that are conserved (grey) and differentially expressed (blue), according to post-hoc comparisons explained in Supplementary Table 2. H: human (*Homo sapiens*), C: chimpanzee (*Pan troglodytes*), M: rhesus monkey (*Macaca mulatta*) (**b**) Number of mRNA genes that are significantly up- (H>C=M; tan) or downregulated (H<C=M; black) across the 16 regions of the human brain emphasizes that the striatum is the region with most human-specific expression. (**c, d**, **e**) Expression [log2 (RPKM+1)] levels of *TWIST1*, *MET*, and *ZP2*, respectively. Significant differences are labeled with an asterisk. The boxes represent quartiles of the data and the whiskers represent 1.5 times interquartile range. (**f**) Inter-species patterns of miRNA expression are illustrated based on log2 fold-change between human and chimpanzee (x-axis) vs human and macaque (y axis). Guiding lines indicate ± 2-fold difference in abundance. (**g**) Digital RT-PCR validation of miR-4284-5p expression. Error bars represent standard deviation.

**Figure 3** Transcriptional architecture of human, chimpanzee, and macaque brains. Heat maps represent the unsupervised hierarchical clustering of brain regions of human, chimpanzee, and macaque based on expression correlations, using the union of regional DEX genes in each species. Colors represent Pearson correlation between all pairs of brain regions in each species. The lower left half of each heat map shows mRNA correlations, while corresponding miRNA values are plotted in the upper right. Clustering was performed relative to only the mRNA expression data; miRNA row and column ordering was adjusted accordingly for each species.

**Figure 4** miRNAs co-expression modules exhibit both conserved and species-specific expression patterns**.** (**a**) Pearson correlation of all pairs of miRNA modules (rows) and mRNA modules (columns), yellow represents strong positive correlation and blue corresponds strong negative correlation. (**b**) Pearson correlation of only those miRNA-mRNA cluster pairs with a significant enrichment of HiTS-CLIP targets, correlation colour scheme is the same as a). (**c**) miRNA and mRNA module pairs that exhibit species-specific enrichment of miRNA-mRNA target pairs as defined by HiTS-CLIP; red triangles/squares highlight significant enrichment in human but not chimpanzee and/or macaque; blue represents enrichment in chimpanzee but not human; and green represents enrichment in macaque but not human. (**d**) Scaled, average miRNA module expression profiles highlight inter-species and inter-region differences (red – human; blue – chimpanzee; green – macaque). Modules M1 and M37 exemplify human-specific up- and downregulation, respectively, across all regions and modules M4 and M21 contain miRNAs that are up- and downregulated in CBC equally in all species.

**Figure 5** mRNA co-expression modules highlight patterns of human-specific expression. (**a, b, c**)Bar plots showing the spatial expression pattern of eigengenes of modules 92, 32, and 130, respectively. These three modules were selected for their distinct human-specific inter-region mRNA expression (**d, e, f**) Co-expression network illustration of the intramodular hub genes of these modules. (**a**, **d**) M92 contains genes that are upregulated in human cortical areas, including *TWIST2*. (**b**, **e**) M32 is composed of genes that are upregulated in both chimpanzee and macaque NCX areas, including *TWIST1*. (**c**, **f**) M130 contains genes that are upregulated mainly in the human STR, AMY, HIP and, to a lesser extent, NCX areas. *TH* and *DDC*, two genes involved in the biosynthesis of dopamine are hub genes of this module. The size of nodes represents the intramodular connectivity of the gene. Dark green represents higher connectivity between genes.

**Figure 6** Human-specific upregulation of genes associated with dopamine synthesis**.** (**a**) Schematic figure of the dopamine biosynthesis pathway. (**b**, **d**) Expression [log2 (RPKM+1)] pattern of *TH* and *DDC*, respectively, showing higher expression in the human striatum. The boxes represent quartiles of the data and the whiskers represent 1.5 times interquartile range. Asterisk represents human-specific differential expression (Benjamini-Hochberg adjusted P < 0.01). (**c**, **e**) Representative ­­­­­­immunoblot of TH and DDC, respectively, in striatal lysate showing higher levels of protein in human than in chimpanzee and macaque. GAPDH serves as a control for equal protein loading.

**Figure 7** Evolution of neurotransmitter receptor gene co-expression networks**.** Illustration of co-expression networks of mRNAs encoding proteins in (**a**) glutamatergic (**b**) GABAergic,(**c**) cholinergic, and (**d**) serotoninergic receptor subunits. Positive correlation (red) and negative correlation (blue) between genes are shown. (**e**)Z statistics of the Maximum Adjacency Ratio (MAR; see Methods) and (**f**)the corresponding p-values of neurotransmitter receptor networks show that the GABAergic and glutamatergic systems are the most conserved (P < 0.05). Cholinergic and serotoninergic systems show lower conservation, especially in macaque (P > 0.05). All other neurotransmitter receptor networks analyzed show low conservation among the three species. (**g**) Hierarchical clustering of genes encoding GABA receptors. The co-expression of *GABRA1*, *GABRG2* and *GABRB2* (red box), which represent the most common assembly of GABAA receptors (α1β2γ2), is conserved across species. (**h**) Hierarchical clustering of genes encoding glutamate receptors. The co-expression of *GRIA2* and *GRIA3* (red box), subunits of AMPA receptors, is conserved across species.