Supplementary Information for

Accelerated evolution of oligodendrocytes in human brain

Stefano Berto, Isabel Mendizabal, Noriyoshi Usui, Kazuya Toriumi , Paramita Chatterjee, Connor Douglas, Carol Tamminga, Todd M. Preuss, Soojin V. Yi, and Genevieve Konopka

Correspondence: Genevieve.Konopka@utsouthwestern.edu (G.K.), soojinyi@gatech.edu (S.V.Y.)

This PDF file includes:

Supplementary Figures: S1 to S4

Supplementary materials for this manuscript include the following:

- **Table S1. Demographic data.** Demographic data with life traits, RNA-seq QC metrices, technical and biological covariates.
- **Table S2. Differential gene expression summary statistics for NeuN and OLIG2 data for Human, Chimpanzee, and Rhesus macaque.** All statistics for genes differentially expressed in each species, functional enrichment, and database for association with previous studies.
- **Table S3. WGCNA statistics for NeuN and OLIG2 data.** All statistics for module detection, overlap with differentially expressed genes, and functional enrichment.
- **Table S4. MAGMA summary statistics for NeuN and OLIG2 modules.** All statistics for GWAS enrichment in NeuN and OLIG2 modules.
- **Table S5. Differential gene expression summary statistics for NeuN and OLIG2 data for Schizophrenia vs Control.** Genes differentially expressed in schizophrenia compared with controls in NeuN and OLIG2 with relative statistics.

Supplementary figures:

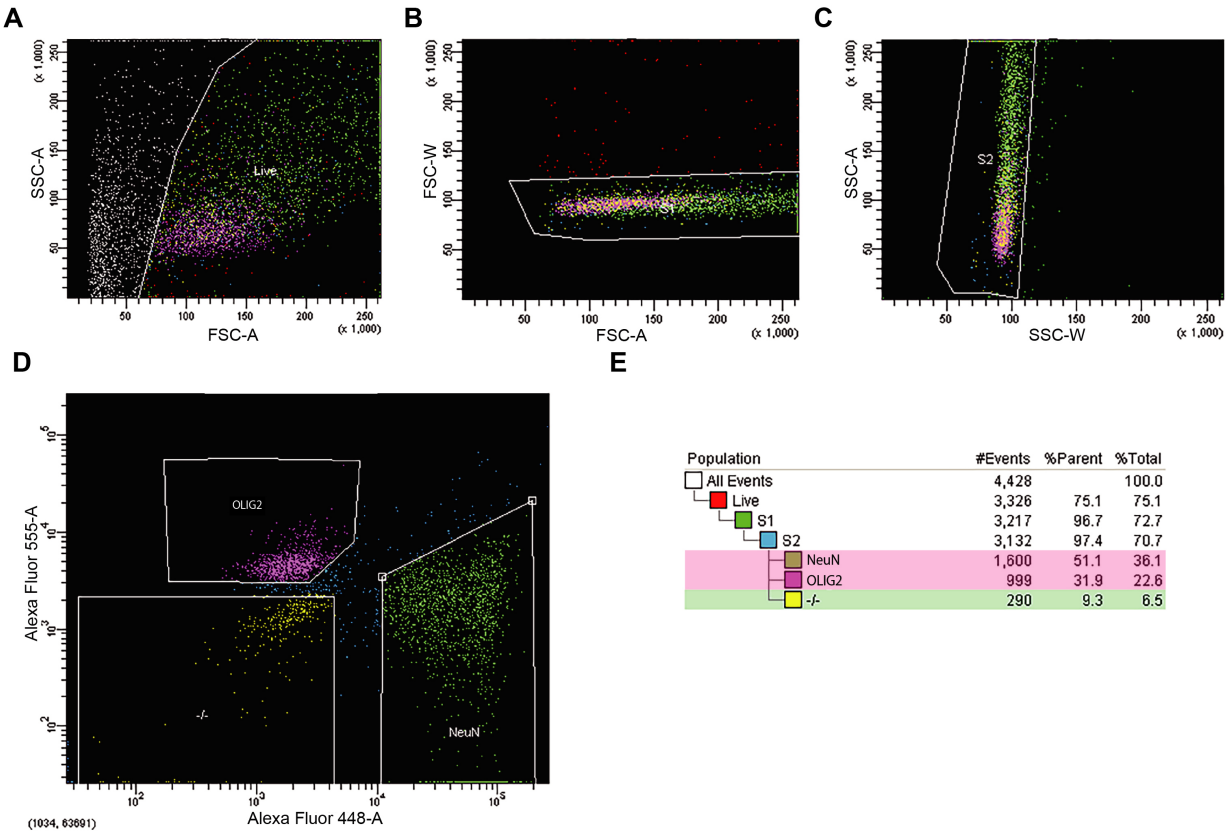

Fig. S1. **Generation cell-type gene expression profiles.** A-E) An example isolation of nuclei expressing either NeuN conjugated to Alexa 488 or OLIG2 conjugated to Alexa 555. Nuclei were first sorted for size and complexity for removing dead cells (A), followed by gating to exclude doublets that indicate aggregates of nuclei B-C), and then further sorted to isolate nuclei based on fluorescence (D). "-/-" nuclei are those that are neither NeuN+ nor OLIG2+. (E) An example of percentage nuclei at each selection step during FANS of a chimpanzee sample. Note that while in this example more nuclei were NeuN+, in other samples, the proportions might be reversed. NeuN-positive (NeuN+) nuclei represent neurons within the cerebral cortex as few NeuN-negative (NeuN-) cells in the mammalian cortex are neurons (e.g. Cajal-Retzius neurons). OLIG2-positive (OLIG2+) nuclei represent oligodendrocytes and their precursors.

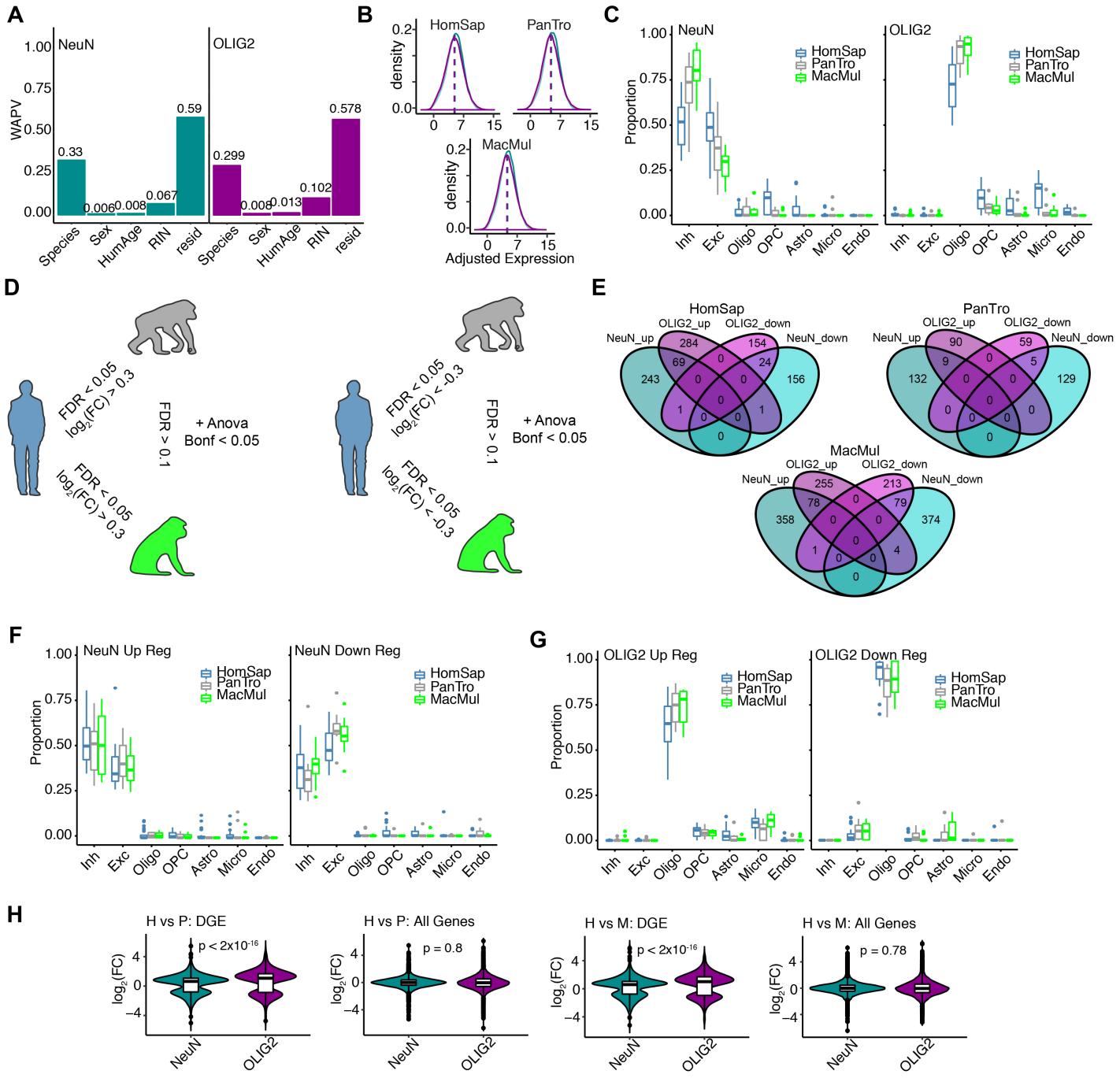

Fig. S2. **Analysis of species-specific cell-type gene expression profiles.** A) Variance explained by covariates weighted across the first 5 principal components (WAPV = Weighted average proportion variance) for NeuN or OLIG2. Gene expression patterns showed small variance explained by biological (Sex and Humanized Age) and technical (RIN) covariates. B) Density plot of gene expression for NeuN and OLIG2 in each species. Y-axis represents the density. X-axis represents the adjusted gene expression. C) Deconvolution based on Allen Brain Institute single nuclei data from the human middle temporal gyrus (MTG) (1). NeuN is explained by a high proportion of inhibitory and excitatory neurons.

OLIG2 is explained by oligodendrocytes. Y-axis represents weighted proportion. X-axis represents the cell-type identified in the Allen Brain Institute single nuclei data from MTG (HomSap = *Homo sapiens*, PanTro = *Pan troglodytes*, MacMul = *Macaca mulatta*, Inh = inhibitory neurons, Exc = excitatory neurons, Oligo = oligodendrocytes, OPC = oligodendrocyte precursor cells, Astro = astrocytes, Micro = microglia, Endo = endothelial cells). D) A parsimony approach based on linear model statistics from pairwise comparisons. Left panel: Upregulated genes were considered if FDR < 0.05 in species 1 vs species 2 and species 1 vs species 3 with $\log_2$(FC) > 0.3 in both comparisons while not significantly differentially expressed (defined as FDR > 0.1) in species 2 vs species 3. Similarly, right panel: Downregulated genes were considered if FDR < 0.05 in species 1 vs species 2 and species 1 vs species 3 with $\log_2$(FC) < -0.3 in both comparisons with FDR > 0.1 in species 2 vs species 3. Additionally, we added a Bonferroni corrected ANOVA < 0.05 cutoff in the model across the three species analyzed. Blue = human, grey = chimpanzee, green = rhesus macaque. E) Venn diagrams showing the intersections between species-specific differentially expressed genes in human (HomSap), chimpanzee (PanTro) and rhesus macaque (MacMul). F-G) Deconvolution of human-specific up-/down-regulated genes in NeuN+ (F) and OLIG2 (G) based on Allen Brain Institute single nuclei data from the human middle temporal gyrus (MTG). G) Violin plot showing differences between $\log_2$(Fold Change) of Human specific genes (DGE) and All expressed genes in NeuN and OLIG2 for human (*H*) compared with chimpanzee (*P*) and human (*H*) compared with rhesus macaque (*M*). Shown are the p-values from Kolmogorov–Smirnov test (K-S test).

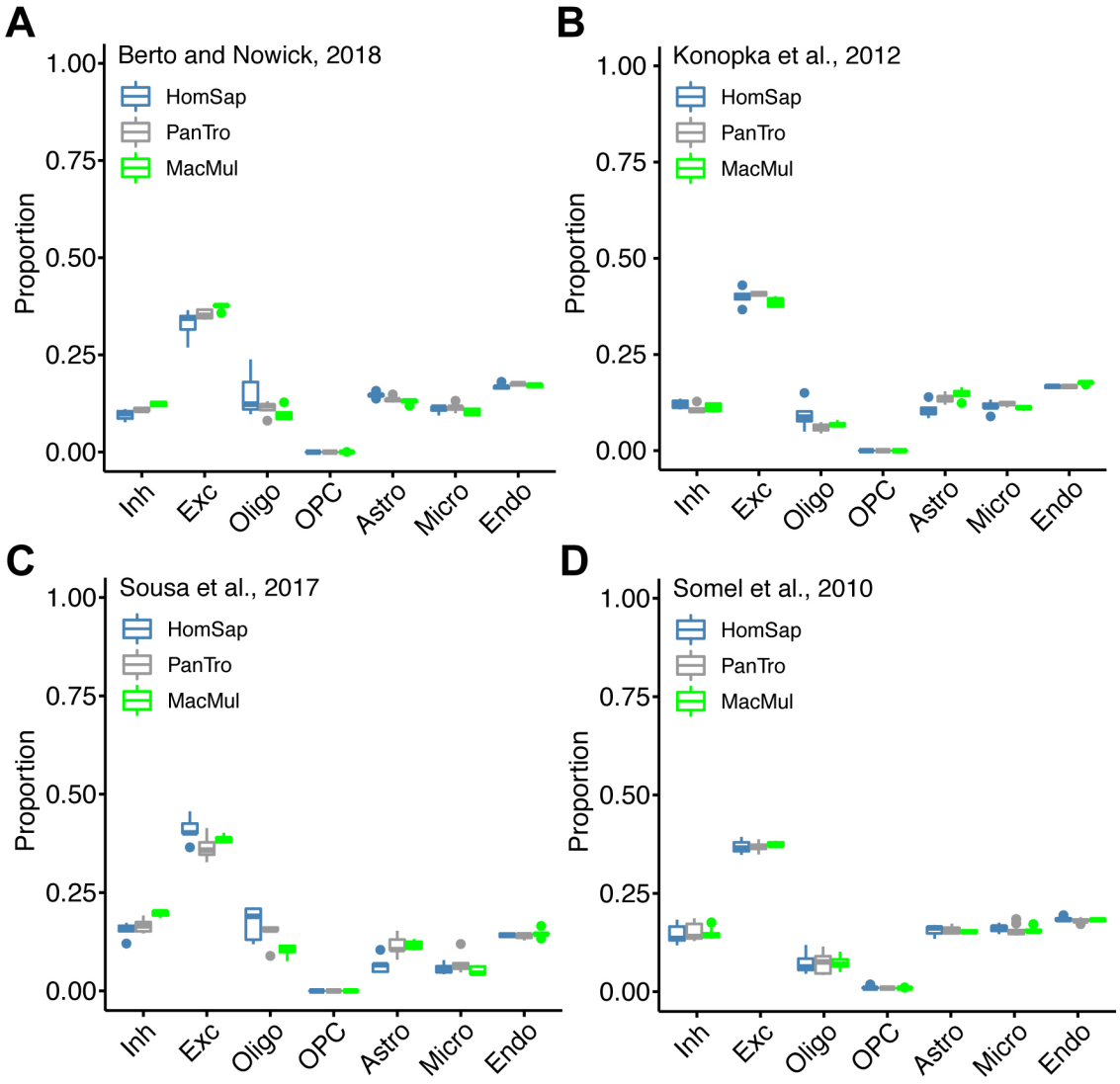

Fig. S3. Deconvolution based on the Allen Brain Institute single nuclei data from human middle temporal gyrus (MTG) (1). A) Primate transcriptomic meta-analysis data from Berto et al. (2). B) Primate transcriptomic data from Konopka et al. (3). C) Primate transcriptomic data from Sousa et al. (4). D) Primate transcriptomic data from Somel et al. (5). All data show a high proportion of excitatory neurons and sparse proportion of other cell-types (< 0.25). Y-axis represents weighted proportion. X-axis represents the cell-type identified in the single nuclei data from MTG (1). (HomSap = *Homo sapiens*, PanTro = *Pan troglodytes*, MacMul = *Macaca mulatta*, Inh = inhibitory neurons, Exc = excitatory neurons, Oligo = oligodendrocytes, OPC = oligodendrocyte precursor cells, Astro = astrocytes, Micro = microglia, Endo = endothelial cells).

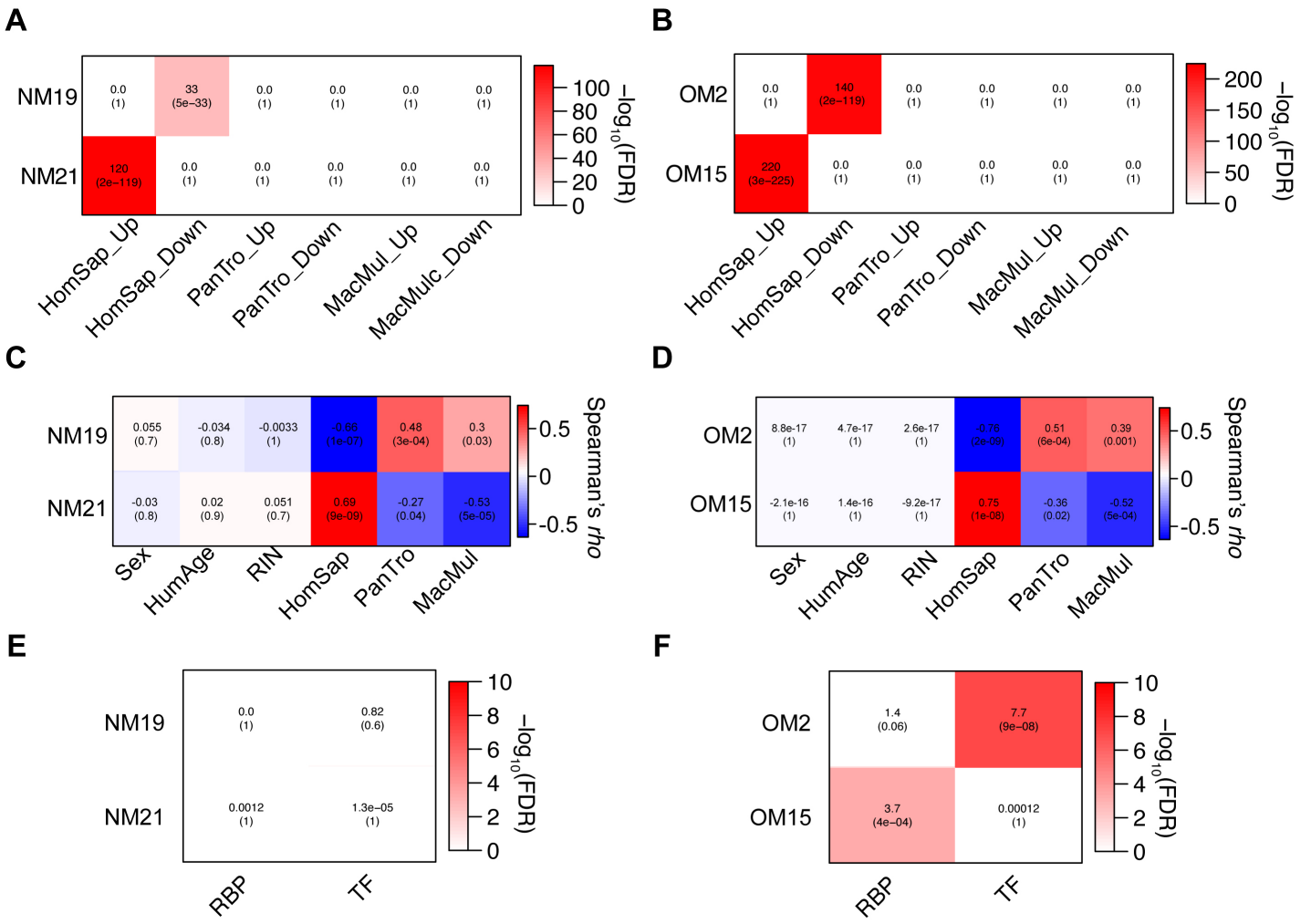

Fig. S4. **Correlations among module eigengene and associated factors.** A-B) Fisher's exact test Odds Ratios with associated FDR adjusted p-values (within parentheses) representing enrichment of species specific DEGs in (A) NeuN modules and (B) OLIG2 modules. C-D) Spearman's rank correlations with associated p-values (within parentheses) between covariates and module eigengene of the module detected in (C) NeuN modules and (D) OLIG2 modules. (HumAge = Humanized Age, HomSap = *Homo sapiens*, PanTro = *Pan troglodytes*, MacMul = *Macaca mulatta*). None of the modules selected showed significant association with technical or biological covariates. E-F) Fisher's exact test Odds Ratios with associated FDR adjusted p-values (within parentheses) representing enrichment of RNA-binding proteins (RBP) and transcription factors (TF) in (E) NeuN modules and (F) OLIG2 modules.

Supplementary References

1. Boldog E*, et al.* (2018) Transcriptomic and morphophysiological evidence for a specialized human cortical GABAergic cell type. *Nat Neurosci* 21(9):1185-1195.
2. Berto S & Nowick K (2018) Species-Specific Changes in a Primate Transcription Factor Network Provide Insights into the Molecular Evolution of the Primate Prefrontal Cortex. *Genome Biol Evol* 10(8):2023-2036.
3. Konopka G*, et al.* (2012) Human-specific transcriptional networks in the brain. *Neuron* 75(4):601-617.
4. Sousa AMM*, et al.* (2017) Molecular and cellular reorganization of neural circuits in the human lineage. *Science* 358(6366):1027-1032.
5. Somel M*, et al.* (2010) MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res* 20(9):1207-1218.