Young transposable elements rewired gene regulatory networks in human and chimpanzee hippocampal intermediate progenitors

Sruti Patoori[1], Samantha M. Barnada[1], Christopher Large[2], John I. Murray[2], and Marco Trizzino[1,*]

[1] Department of Biochemistry and Molecular Biology, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA
[2] Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

* Corresponding author: Marco Trizzino, Department of Biochemistry and Molecular Biology, Sidney Kimmel Medical College, Thomas Jefferson University, 233 S 10th Street, BLSB 826, Philadelphia, PA, 19104. E-mail: marco.trizzino@jefferson.edu

Abstract

The hippocampus is associated with essential brain functions such as learning and memory. Human hippocampal volume is significantly greater than expected when compared to non-human apes, suggesting a recent expansion. Intermediate progenitors, which are able to undergo multiple rounds of proliferative division before a final neurogenic division, may have played a role in the evolutionary hippocampal expansion. To investigate the evolution of gene regulatory networks underpinning hippocampal neurogenesis in apes, we leveraged the differentiation of human and chimpanzee induced Pluripotent Stem Cells into TBR2-positive hippocampal intermediate progenitors (hpIPCs). We find that the gene networks active in hpIPCs are significantly different between humans and chimpanzees, with ~2,500 genes differentially expressed. We demonstrate that species-specific transposon-derived enhancers contribute to these transcriptomic differences. Young transposons, predominantly Endogenous Retroviruses (ERVs) and SINE-Vntr-Alus (SVAs), were co-opted as enhancers in a species-specific manner. Human-specific SVAs provided substrates for thousands of novel TBR2 binding sites, and CRISPR-mediated repression of these SVAs attenuates the expression of ~25% of the genes that are upregulated in human intermediate progenitors relative to the same cell population in the chimpanzee.

Summary statement: Evolution of human and chimpanzee hippocampal development was mediated by co-option of young retrotransposons into species-specific enhancers.

Key words: hippocampal development, transposable elements, iPSCs, intermediate progenitors, TBR2, SVAs

Introduction

The hippocampus is associated with many traits relevant in the context of human evolution. These include traits such as tool use and language which require social cognition and learning, as well as spatial memory, navigation, and episodic memory (Burgess et al., 2002; Eichenbaum, 2017a; Eichenbaum, 2017b; Squire, 1992; Tomasello and Herrmann, 2010). This region of the brain is also greatly affected by Alzheimer's Disease (AD), a neurodegenerative disorder characterized by cell death, plaques and tangles of misfolded proteins, and cognitive decline (Duyckaerts et al., 2009). It has been hypothesized that the cognitive AD phenotype is uniquely human and that non-human primates, including chimpanzees, do not exhibit AD-related dementia (Edler et al., 2017; Finch and Austad, 2015; Walker and Jucker, 2017). If humans are uniquely susceptible to AD, it is crucial to understand how the human hippocampus differs from that of our closest biological relatives, the chimpanzees.

Human hippocampal volume is 50% greater than expected when compared to the hippocampal volumes of non-human apes, possibly indicating a recent hippocampal expansion specific to the human lineage (Barger et al., 2014). However, the evolution of the human hippocampus and the developmental mechanisms driving the human-specific volume increase have not yet been thoroughly studied.

Recent studies have suggested that evolutionary changes to neuronal progenitors may have had an impact on cortical volume in primates by increasing proliferative potential (Martínez-Cerdeño et al., 2006; Rétaux et al., 2013; Florio and Huttner, 2014). A specific class of neuronal progenitors known as intermediate progenitor cells (IPCs) or "transit amplifying cells" are able to undergo multiple rounds of proliferative division before a final neurogenic division (Englund, 2005; Arnold et al., 2008; Pontious et al., 2008; Hevner, 2019). These cells express the neurodevelopmental transcription factor TBR2 (EOMES) and are found in the sub-ventricular zone of the developing neocortex and hippocampus (Bulfone et al., 1999; Kimura et al., 1999; Englund, 2005; Cipriani et al., 2016). Genetic ablation of TBR2 in these progenitors results in reduced cortical thickness in mice (Sessa et al., 2008), abnormal cortical cell differentiation (Mihalas et al., 2016), and impaired neurogenesis in the hippocampal formation (Hodge et al., 2012). As these IPCs are hypothesized to play a role in neocortical expansion, they may also be involved in the lineage-specific hippocampal expansion seen in humans.

Many of the differences between humans and chimpanzees are due to diverging gene regulatory sequences (Agoglia et al., 2021; Enard et al., 2002; Gokhman et al., 2021; King and Wilson, 1975; Wray, 2007). A recent study comparing gene expression between adult human, chimpanzee and macaque brain regions identified several genes specifically upregulated in the human hippocampus (Sousa et al., 2017). However, transcriptomic differences between primate species during specific timepoints of hippocampal development have not been investigated.

As samples of developing human and chimpanzee brain tissue are extremely limited, Induced Pluripotent Stem Cells (iPSCs) are an ideal system to conduct comparative studies of human and chimpanzee hippocampal development. Previous studies have employed iPSC-derived cortical organoids and iPSC-derived neuronal progenitor cells (NPCs) from human and chimpanzee for comparative and developmental genomic purposes (Marchetto et al., 2019; Mora-Bermúdez et al., 2016). Here, we leverage human and chimpanzee iPSC-derived hippocampal progenitors as models for comparative developmental and genomic studies with the goal of identifying species-specific differences in gene regulation during hippocampal development.

Several recent papers have recently demonstrated that transposable elements (TEs) can alter existing regulatory elements or generate entirely novel ones, as well as expand in a species- or lineage-specific manner (reviewed in Sundaram and Wysocka, 2020). Species-specific TE expansion and co-option into gene regulatory networks have been demonstrated as a mechanism for evolutionary change (Chuong et al., 2016; Fuentes et al., 2018; Jacques et al., 2013; Lynch et al., 2011; Lynch et al., 2015; Miao et al., 2020; Mika et al., 2021; Pontis et al., 2019; Trizzino et al., 2017). Endogenous Retroviruses (ERVs)

and SINE-Vntr-Alus (SVAs) are among the TE families more frequently associated with gene regulatory activity in the human genome (Chuong et al., 2016; Fuentes et al., 2018; Pontis et al., 2019; Trizzino et al., 2017; Trizzino et al., 2018). SVAs encompass six subfamilies, denoted as SVA-A through -F. Of the ~3,000 SVA copies in the human genome, nearly half are human specific, including all the SVA-E and SVA-F (Quinn and Bubb, 2014; Wang et al., 2005). The remaining half are also found in other great apes. SVAs are still replication competent and thus able to transpose in the human genome. ERVs are retrotransposons belonging to the Long Terminal Repeat (LTR) group. They are remnants of past retroviral infection events, and make up ~8% of the human genome (Tokuyama et al., 2018). Both SVAs and ERVs were recently found to be enriched within the sequences of active cis-regulatory elements (enhancers, promoters) in hippocampal tissue as compared to other human brain regions where they are predominantly repressed (Trizzino et al., 2018). Therefore, we hypothesize that species-specific ERV and SVA transposon activity may influence the gene regulatory networks necessary for human and chimpanzee hippocampal development.

Given the key function that intermediate progenitors played in the evolution of the primate brain (Florio and Huttner, 2014; Martínez-Cerdeño et al., 2006), we sought to identify molecular differences between iPSC-derived human and chimpanzee hippocampal intermediate progenitors (hpIPCs) in terms of gene expression and the regulatory activity of non-coding regions. We specifically examined gene expression (RNA-seq and scRNA-seq), gene regulation (ATAC-seq) and functional TE activity via CRISPR-interference.

We leveraged singe-cell RNA-seq to examine the temporal trajectory of the hpIPCs during differentiation in both species. After confirming that the hpIPC differentiated cells express the appropriate neurodevelopmental markers, we conducted a transcriptomic comparison between human and chimpanzee hpIPCs. This analysis revealed over 2,500 differentially expressed genes. We then used ATAC-seq to conduct extensive analyses of differential chromatin accessibility between human and chimpanzee hpIPCs. In both species, differentially accessible chromatin regions were more likely than expected to overlap a TE insertion. Further, these regions were found as both enriched and depleted for specific TE families. Notably, species-specific enrichment for ERV and SVA sequences within differentially accessible genomic sites correlated with species-specific changes in nearby gene expression. This is likely driven by transcription factors binding to the TE-derived regulatory sequences, as we demonstrate for TBR2 and SVA-derived enhancers. Finally, we used CRISPR-interference to repress all the accessible SVAs in progenitor-like cells and demonstrated that such repression results in global changes in gene expression and affects hundreds of important neurodevelopmental genes.

This work demonstrates that two young TE families have contributed significantly to the gene regulatory differences between human and chimpanzee hippocampal development, providing insight into how the human hippocampus evolved both its unique cognitive capacity and its susceptibility to neurodegenerative disease.

Results

An iPSC-derived model for human and chimpanzee hippocampal intermediate progenitors

We modeled hpIPCs in humans and chimpanzees using three human and three chimpanzee iPSC lines. All six iPSC lines were validated as pluripotent in previous studies (Gallego Romero et al., 2015; Pagliaroli et al., 2021; Pashos et al., 2017; Ward et al., 2018; Yang et al., 2015; Zhang et al., 2015). For both human and chimpanzee iPSCs, we used two female cell lines and one male line.

We used a previously published method to generate hpIPCs from iPSCs (Yu et al., 2014). In this protocol, the stem cells are treated with a media containing anticaudalizing factors and a Sonic Hedgehog antagonist (DKK1, Nogging, SB431542) to generate forebrain progenitor cell types (Yu et al., 2014). It

145 should be noted that the hpIPCs are distinct from a more general neuronal progenitor cell (pan-NPC) as
146 the pan-NPC media is supplemented only with FGF2 and B27. Moreover, and the hpIPCs can be further
147 induced to generate mature hippocampal CA3 pyramidal or dentate gyrus granule neurons (Sarkar et al.,
148 2018; Yu et al., 2014).
149 As we were specifically interested in TBR2-positive intermediate hippocampal progenitors, we
150 differentiated one human male and one chimpanzee male iPSC line. To ensure that the differentiation
151 process proceeded similarly in each species, we compared the transcriptomes by conducting single-cell
152 RNA-seq (hereafter scRNA-seq) at 24 hour intervals from the iPSC stage (day 0) to the hpIPC (day 5). We
153 assayed a total of 18935 cells, with 4540 chimpanzee cells and 14395 human cells. These data
154 demonstrate that the differentiation follows a similar trajectory in both species (Figure 1A). Despite the
155 difference in final cell number between each species (Supplemental Figure 1A), all six time points overlap
156 closely between species (Supplemental Figure 1B, 1C). Chimpanzee cells from day 2 and day 3 cells overlap
157 more than expected with human cells from day 3 and day 4, respectively, suggesting that the chimpanzee
158 cells are farther along in the hpIPC differentiation at these time points. However, both species align closely
159 again by day 5, which is the time point chosen for all the genomic analyses conducted in the present study.
160 We further leveraged the scRNA-seq data to examine the expression of known
161 neurodevelopmental markers and saw a progressive increase in the expression of both PAX6 and OTX2
162 during the 5 day differentiation (Figure 1B, 1C), which remains consistent between species (Supplemental
163 Figure 1D, 1E). To investigate any noticeable changes to cell cycle or proliferation between species, we
164 quantified the number of at the G1, S, and G2M phases of the cell cycle within the scRNA-seq data set,
165 based on a set of known cell cycle genes and did not detect significant differences (Figure 1E).
166 To ensure that the differentiation resulted in hpIPCs specifically, we conducted a Western Blot in
167 both species and confirmed that TBR2 expression in human cells is greatest at days 4 and 5 (Figure 1D),
168 similarly to OTX2 and PAX6 expression (Figure 1B, 1C). In chimpanzee cells, TBR2 is expressed as early as
169 day 2, corroborating evidence from the scRNA-seq that chimpanzee differentiation initially proceeds
170 faster. However, we do not see evidence that TBR2 expression in chimpanzee cells is greater than TBR2
171 expression at day 5 in human cells. As OTX2 labels neuronal progenitors and TBR2 and PAX6 mark the
172 intermediate progenitor cell type (Florio and Huttner, 2014, Cipriani et al., 2016, Hevner, 2019), the data
173 demonstrate that the differentiation was successful and comparable between species. Lastly, we
174 conducted ChIP-seq against TBR2 and H3K27ac to determine whether TBR2 was bound at the same
175 genomic sites in both species (Figure 1G). We identified 3789 TBR2-bound regions in human hpIPCs and
176 observed that nearly all of them also exhibited the active histone mark H3K27ac. Upon translating the
177 coordinates of these regions to the chimpanzee genome, we observed that the 3781 orthologous regions
178 were bound by TBR2 in chimpanzee hpIPCs. Together, these data indicate that our iPSC-derived model is
179 suitable to study gene regulation within TBR2-positive hippocampal intermediate progenitors in both
180 human and chimpanzee.
181
182 **Important neurodevelopmental genes are differentially expressed between human and chimpanzee**
183 **hippocampal intermediate progenitors**
184 To investigate the differences between human and chimpanzee hippocampal intermediate
185 progenitors, we first aimed to characterize differential gene expression between the hpIPCs of the two
186 species. After five days of treatment with the hpIPC differentiation media, we collected cells for RNA
187 extraction (Fig. 2A). To ensure statistical power, we conducted bulk RNA-seq on two replicates each of all
188 six cell lines (i.e. three biological replicates and six technical replicates per species). The differentiation,
189 the harvesting, and the RNA processing were performed in mixed batches with samples from both species
190 to prevent batch effects. The libraries were sequenced using Illumina NextSeq500, generating 100 bp
191 Paired-End reads. Non-orthologous genes were omitted from the analysis and a total of 2,588 genes were
192 identified as differentially expressed (FDR <0.05 and log2foldChange > 1.5 or < -1.5). The genes with

193 log2foldChange > 1.5 were more highly expressed in the human hpIPCs ("Human UP") while the genes
194 with log2foldChange < -1.5 were more highly expressed in the chimpanzee hpIPCs ("Chimp UP"). In total,
195 1,686 (65.1%) of the differentially expressed genes (DEG) were "Human UP" and 901 (34.9%) of the DEG
196 were "Chimp UP" genes (Fig. 2B, Supplementary Table 1.1).
197 Hippocampal neurodevelopmental markers *PAX6*, *OTX2*, *NEUROD1* and *FOXG1* were found as
198 highly expressed in both species. This is consistent with the immunostaining in demonstrating that
199 the hpIPCs differentiation was successful and that the progenitors from both species are comparable.
200 The "Human UP" genes include *FOXP2*, *MTRNR2L8*, *DHX40*, *VPS13B*, *WDFY2*, and *PURB*, all of
201 which are associated with neurodevelopment or neurodegenerative disease (Abrajano et al., 2009;
202 Hickey et al., 2019; Kamboh et al., 2019; Kolehmainen et al., 2003; MacDermot et al., 2005; Mathys et al.,
203 2019; Sin et al., 2015; Taher et al., 2014). The "Chimp UP" genes include *HIST1H3A*, *BLC2L2*, and
204 *CLIC1* which have been reported as expressed in the hippocampus and associated with sleep, AD, and
205 neurite outgrowth (Averaimo et al., 2014; Datson et al., 2009; Wei, 2020).
206 To understand the transcriptional programs driving these differences in gene expression,
207 we conducted an Ingenuity Pathway Analysis. Three of the top five upstream regulators predicted by the
208 pathway analysis were the transcription factors CREB1, FOXA2 and TBR2 (EOMES) (Fig. 2C). CREB1 is
209 known to regulate genes involved in the nervous system and neurodevelopment (reviewed in Sakamoto
210 et al., 2011), FOXA2 controls dopaminergic neuronal development and disease (Kittappa et al., 2007), and
211 TBR2 plays a crucial role in cortical and hippocampal neurogenesis and is the signature marker of the
212 intermediate progenitor population (Cipriani et al., 2016; Englund, 2005). This pathway analysis was
213 consistent with the RNA-seq data as predicted targets of all three transcription factors were also found to
214 be among the 2,588 genes differentially expressed between human and chimpanzee hpIPCs (Fig. 2D,
215 Supplementary Table 1.2-1.4). The pathway analysis also determined that several of the differentially
216 expressed genes are involved in embryonic development (Fig. 2D, Supplementary Table 1.5). Overall,
217 these findings indicate that previously characterized neurodevelopmental gene regulatory networks are
218 utilized differently during human and chimpanzee hippocampal development.
219
220 **Human-specific chromatin accessibility patterns in hippocampal intermediate progenitors**
221 After identifying differentially expressed genes and the transcriptional networks that may be
222 involved, we sought to identify cis-regulatory differences between the human and chimpanzee hpIPCs. To
223 this end, we conducted ATAC-seq on the hpIPCs from both species. We used the same batches of
224 differentiated iPSCs for the ATAC-seq as we did for the RNA-seq (i.e. from the same batch of
225 differentiation), and generated 100 bp long Paired-End Illumina reads.
226 We first performed a "human-centric" analysis. We aligned the ATAC-seq reads from all six cell
227 lines to the respective reference genome assemblies (hg19 for the human cell lines, panTro5 for the
228 chimpanzee cell lines) and only retained uniquely mapped reads with high mapping quality (Samtools Q =
229 10 filtering). Next, we identified regions of accessible chromatin (peaks; FDR < 0.05) in all three human
230 cell lines. Only peaks replicated in all the three human lines were retained. To carry out a proper
231 comparison, we only retained replicated human ATAC-seq peaks with orthologs in the chimpanzee
232 genome (see methods; Fig. 3A). This filtering ultimately resulted in 82,235 human ATAC-seq peaks
233 replicated in all the human cell lines and with orthologs in the chimpanzee genome. These 82,235 regions
234 were used for downstream analysis. We found that the chromatin accessibility at these regions was highly
235 reproducible across all three human cell lines (Fig. 3B).
236 Next, we quantified the ATAC-seq read depth for each of the 82,235 regions and used DESeq2 to
237 identify sites exhibiting differential chromatin accessibility between the two species. In total, we identified
238 3,006 differentially accessible regions (FDR <0.05; log2foldChange > 1.5 or < -1.5; Supplementary Table
239 2.1-2.2). Of these regions, 92.3% were located at least 1 kb away from the closest transcription start site
240 (TSS), suggesting they could be putative enhancers, while the remaining were putative promoters (Fig.

3C). As expected, given that this analysis was performed with a human-centric approach, 90.1% of the differentially accessible (DA) peaks were significantly more accessible in the human hpIPCs relative to chimpanzee hpIPCs ("Human UP"; Fig. 3C).

As TE insertions can be a source of cis-regulatory evolution, we examined whether the DA regions were more likely to overlap with a TE than those which were accessible to the same degree in both species (non-DA, Fig. 3D and Supplementary Table 2.4-2.5). We observed that 1,335/3,006 (44.4%) DA regions overlapped a TE (Supplementary Table 2.3). This is significantly higher than what was observed for the non-DA peaks (33.2% overlapped a TE; Two-sided Fisher's Exact Test, $p < 0.0001$; Fig. 3D). This indicates that chromatin regions with human-specific accessibility are significantly more likely to be TE-derived than the regions with accessibility levels conserved between human and chimpanzee.

Next, we associated the nearest gene to each DA region and found that TE-derived DA regions were significantly more likely to be near a differentially expressed gene relative to non-TE derived DA regions (Two-sided Fisher's Exact Test, $p = 0.016$; Fig. 3E). Finally, we investigated whether specific TE families were overrepresented among the TE-derived DA regions and found enrichment for LTRs. While LTRs account for ~16% of all human annotated TEs, they represented 33.9% of the TEs overlapping DA regions in our human-centric analysis (Two-Sided Fisher's Exact Test $p < 0.0001$; Fig. 3F and Supplementary Table 2.6). Of these enriched LTRs, 97.1% were ERVs. Notably, 31.7% of those LTR-derived DA regions were located near a differentially expressed gene (Two-Sided Fisher's Exact Test $p < 0.0001$).

Together, these data indicate that there are TE-derived cis-regulatory elements that have significantly greater accessibility in humans than chimpanzees during hippocampal neurogenesis. These TE insertions preceded the human-chimpanzee split, but the difference in accessibility is species-specific, suggesting that the co-option into gene regulatory networks took place after the human-chimpanzee divergence.

Chimpanzee-specific chromatin accessibility patterns in hippocampal intermediate progenitors

We repeated the ATAC-seq analysis as described above, but this time with a "chimpanzee-centric" approach. We started from a set of 72,211 peaks found as replicated in all the three chimpanzee lines and with orthologs in both species (Figs. 4A,B). With this approach, we identified 3,806 ATAC-seq peaks as differentially accessible (DA) between human and chimpanzee (FDR <0.05; log2foldChange > 1.5 or < -1.5; Supplementary Table 3.1-3.2), 82% of which displayed greater accessibility in chimpanzee as compared to humans (i.e. "Chimp UP"; Fig. 4C). Similar to what we observed with the human-centric analysis, 97.9% of the 3,806 DA regions were putative enhancers (distance > 1Kb from TSS; Fig. 4C).

As seen in the human-centric analysis, chimpanzee-specific DA peaks were more likely to be TE-derived than those similarly accessible across species (Two-sided Fisher's Exact Test, $p < 0.0001$; Fig. 4D and Supplementary Tables 3.3-3.5). Of the chimpanzee DA regions, 37.1% overlapped an annotated chimpanzee TE, compared to only 28.1% of the non-DA regions (Two-sided Fisher's Exact Test, $p < 0.0001$; Fig. 4D). However, the TE-derived enhancers in this chimpanzee-centric analysis were no more likely than the non-TE derived ones to be located near differentially expressed genes (Fig. 4E).

We found enrichment for LTRs, which account for approximately 16% of the chimpanzee TEs but represented 36.1% of the TE-derived DA regions (Two-sided Fisher's Exact Test, p-value < 0.0001; Fig. 4F and Supplementary Table 3.6; 98.2% were ERVs), and SVAs, which account for just 0.25% of annotated chimpanzee TEs but represented 1.1% of the TE-derived DA regions (Two-sided Fisher's Exact Test, $p < 0.0001$, Fig 4G and Supplementary Table 3.8). In particular, the SVA-B and C subfamilies were the most enriched (Fig. 4H). Together, these data indicate that chimpanzee ERV and SVA transposons were co-opted into regulatory elements important for the developing chimpanzee hippocampal intermediate progenitors.

289 **Genomic features underlying species-specific LTR enrichment at hpIPC enhancers**
290 We aimed to further investigate genomic features potentially underlying the LTR enrichment
291 among the differentially accessible hippocampal progenitor regions. To this end, we conducted a motif
292 analysis using the MEME suite (Bailey et al., 2015). Binding motifs for CTCF and EGR2 were detected as
293 enriched in the LTR-derived differentially accessible regions identified from both the human-centric and
294 chimpanzee-centric analyses (Figs. 5A, 5C). EGR2 is an early response gene involved in learning and
295 memory, in the brain's response to stimuli, and in hippocampal synaptic plasticity (Cheval et al., 2012;
296 Mukherjee et al., 2021; Poirier et al., 2007). CTCF, a well-known regulator of chromatin structure, has
297 been implicated in various neurodevelopmental disorders (reviewed in Davis et al., 2018)
298 We identified 63 differentially expressed genes located near the human-enriched LTRs (Fig. 5B
299 and Supplementary Table 2.7). These included *GLUL*, a glutamine synthetase hypothesized to provide
300 neuroprotection in AD patients, (Kohane and Wood, 2021) and *DLK1*, a Notch ligand involved in SVZ
301 neurogenesis (Ferrón et al., 2011), both of which are upregulated in human hpIPCs compared to
302 chimpanzee hpIPCs.
303 We identified 34 differentially expressed genes located near the chimp-enriched LTRs (Fig. 5D and
304 Supplementary Table 3.7). For the chimp-centric analysis the differentially expressed genes located near
305 TE-derived enhancers with species-specific accessibility included *HIST1H3A* and *MTRNR2L8* (Fig. 5D).
306 These genes were highly upregulated in the chimpanzee and human hpIPCs, respectively (Fig. 5D).
307 Importantly, *HIST1H3A* has been associated with autism spectrum disorders and sleep deprivation
308 (Crawley et al., 2016; Wei, 2020) while *MTRN2L8* has been reported as upregulated in Alzheimer's patients
309 (Mathys et al., 2019).
310
311 **Human-Specific SVAs play a major role in hippocampal neurogenesis**
312 As mentioned earlier, the chimpanzee-centric analysis also identified SVA transposons as enriched
313 within chimpanzee-specific enhancers (Fig. 6A and Supplementary Table 3.8). These SVAs were enriched
314 for the identified binding motifs of the neurodevelopmental factors ASCL1, ZIC1, and KLF8 (Andersen et
315 al., 2014; Aruga, 2004; Yi et al., 2014) as well as the JUN/FOS AP-1 dimer, which is a known enhancer
316 activator (Raivich, 2008; Raivich and Behrens, 2006) (Fig 6A).
317 It is important to remark that all the analyses shown so far were exclusively based on genomic
318 sites with characterized orthologs in both species, to ensure an "apple-to-apple" comparison. However,
319 nearly 2,000 SVA copies, including the entire SVA-E and F subfamilies, are exclusive to the human genome.
320 Given that previous studies found that SVAs are highly enriched in active enhancers and promoters of the
321 human hippocampus (Trizzino et al., 2018), we sought to investigate this further. We focused on the SVA
322 copies exclusively present in the human genome and not in any other primate genome (hereafter human-
323 specific SVAs). We performed sequence-based motif analysis for all the human-specific SVAs and
324 identified the binding motif for the hpIPCs signature factor TBR2 (EOMES) as the most enriched ($p = 10^{-}$
325 ^{2103}; Fig. 6B). Motifs for other transcription factors associated to hippocampal neurogenesis and function
326 were also recovered (SMAD4, VDR, PLAG1; Fig. 6B).
327 Next, we annotated all of the genes found within 50 kb from each human-specific SVA. A total of
328 2,216 genes were recovered using this approach. We performed Pathway Analysis on this set of genes
329 and found that Melatonin Degradation and Nicotine Degradation were the two most significantly enriched
330 pathways ($p = 7.6 \times 10^{-6}$ and $p = 2.5 \times 10^{-5}$ respectively; Fig. 6C) and PHF8 was recovered as the top
331 upstream regulator for the gene network (Fig. 6C). Notably, both Melatonin and Nicotine Degradation
332 pathways are strongly active in the human hippocampus. PHF8 is a histone demethylase that contributes
333 to the regulation of mTOR. The mTOR pathway is hyperactive in the human hippocampus where it
334 regulates the protein synthesis-dependent plastic changes underlying learning and memory (Bekinschtein
335 et al., 2007; Fortress et al., 2013; Graber et al., 2013). Mutations in the *PHF8* gene cause cognitive
336 impairment and intellectual disability (Chen et al., 2018).

337 Therefore, we sought to determine the contribution of human-specific SVAs to the TBR2-
338 mediated gene regulatory network in human hpIPCs. Using our ATAC-seq data, we identified 1,816 human
339 SVAs as accessible in human hpIPCs (Fig. 6D). Of these, nearly a quarter (434) displayed high accessibility,
340 while 1,382 were moderately accessible (Fig. 6D). Next, we used chromatin immunoprecipitation
341 followed by sequencing (ChIP-seq) to profile TBR2 binding in two human lines at day 5 of hpIPC
342 differentiation. As with the previous sequencing experiments, we generated 100 bp long Paired-End reads
343 and only retained uniquely mapping high quality reads (Samtools q=10 filtering) in order to maximize our
344 chance to properly map reads on repetitive regions. This experiment revealed that 739 of the accessible
345 SVAs showed TBR2 signal in the two human lines (Fig. 6D and Supplementary Table 2.8). Notably, 257 of
346 the (48.3%) TBR2-bound SVAs were human-specific. (Fig. 6E and Supplementary Table 2.9). TBR2-bound
347 human-specific SVAs were located near 37 genes that our RNA-seq analysis identified as differentially
348 expressed between human and chimpanzee (Fig. 6F and Supplementary Table 2.10). These genes include
349 *VPS13B* (up in humans), which is responsible for a rare developmental disease known as Cohen Syndrome
350 (Kolehmainen et al., 2003), *NR4A2* (down in humans), which has been implicated in neurodevelopmental
351 language impairment (Reuter et al., 2017), *DHX40* (down in humans), which is implicated in Alzheimer's
352 (Taher et al., 2014), and *E2F1* (up in humans), which is a cell cycle regulator associated with several
353 neurodegenerative diseases such as Alzheimer's (Zhang et al., 2010).
354 In summary, these data support a model in which human-specific SVAs provided a substrate for
355 binding sites of TBR2 and other important hippocampal regulators. Therefore, it is likely they were co-
356 opted in the gene regulatory networks active during hippocampal neurogenesis, which led to human-
357 specific regulation of key genes.
358
359 **CRISPR-mediated SVA repression has massive repercussions on global gene expression**
360 To further assess the contribution of SVA transposons to the gene regulation of hippocampal
361 intermediate progenitors, we leveraged CRISPR-interference to simultaneously repress most of the active
362 SVAs. We utilized NCCIT cells treated with retinoic acid (RA) as the experimental system for this purpose.
363 The NCCIT cell line is derived from embryonal carcinoma and thus exhibits a gene expression signature
364 highly similar to human embryonic stem cells (Fuentes et al., 2018; Barnada et al., 2022). Importantly,
365 NCCITs treated for 7 days with RA differentiate into intermediate neural progenitor-like cells expressing
366 both PAX6 and TBR2 (Mandal et al., 2015; Fig 7A). RNA-seq tpms suggest very strong correlation between
367 RA-treated NCCITs and human day-5 hpIPCs (Pearson correlation $p < 2.2 \times 10^{-16}$, R=0.99). We cloned a
368 stable NCCIT line with a permanently incorporated doxycycline-inducible, catalytically-dead Cas9 fused to
369 a repressive KRAB domain (dCas9-KRAB). The KRAB domain deposits repressive histone methylation
370 (H3K9me3) to the regions targeted by the dCas9 via single guide-RNAs (sgRNAs). Into this same line, we
371 also permanently knocked-in two single guide-RNAs (sgRNAs) that are able to target >80% of all SVAs
372 (Pontis et al. 2019). Hereafter, we refer to the RA-treated CRISPR line as RA-NCCITs. Remarkably, exposing
373 the RA-NCCITs to doxycycline for 72 hours was sufficient to induce dCas9 activation (Fig. 7B) and the
374 deposition of repressive histone methylation (H3K9me3) on over 2,500 previously unmethylated SVAs
375 (Fig. 7C). We next performed RNA-seq on the RA-NCCITs with or without doxycycline treatment (three
376 replicates per condition). First, we observed that the genome-wide expression levels (TPMs) of the RA-
377 NCCITs where highly correlated with those of the iPSC-derived human hpIPCs (Pearson correlation = 0.93;
378 $p < 2.2 \times 10^{-16}$). This indicates that RA-treated NCCITs are appropriate to model hippocampal intermediate
379 progenitors.
380 Then, we compared the expression levels of the RA-NCCITs with or without doxycycline treatment
381 and identified 5,795 differentially expressed genes (FDR < 0.05; Fig. 7D and Supplementary Table 4.1). Of
382 these genes, 677 were previously identified as differentially expressed when comparing human to
383 chimpanzee hpIPCs (Fig. 7E; Supplementary Table 4.2). In other words, the expression of over a quarter
384 (26.1%) of the genes that exhibited human-specific expression signature in hippocampal intermediate

385 progenitors seem to be under control of SVA transposons. Importantly, one of these genes was *FOXP2*,
386 which has been associated with the evolution of language and is implicated in several speech-disorders
387 (Enard, 2011; Liégeois et al., 2016; MacDermot et al., 2005).
388 Since the gRNAs used for this experiment were originally designed to target a DNA sequence
389 shared by the SVAs with the LTR5H family (Pontis et al. 2019), we restricted the analysis to the genes that
390 are associated to human SVAs (i.e. only considering the genes that represent the closest gene to an
391 annotated SVA: hereafter SVA-genes). By doing so, we found 611 SVA-genes as differentially expressed in
392 RA-NCCITs upon SVA repression (Supplementary Table 4.3). Of these, 90 were previously identified as
393 differentially expressed when comparing human to chimpanzee hpIPCs (Supplementary Table 4.4). Thus,
394 the expression of these 90 genes can be bona-fide considered as directly regulated by SVA-derived
395 enhancers in both primary hippocampal progenitors and in the NCCIT cell line. Remarkably, the large
396 majority of these genes (72.5%) had decreased expression upon SVA repression. These include *SOX2*,
397 *FGF2*, *PRDM1*, *NTRK2*, and *TFAP2B* (Supplementary Table 4.4).
398 Finally, we examined the genes previously identified as being near a TBR2-bound human-specific
399 SVA in iPSC-derived hpIPCs, and found that nearly a third of them lose expression in RA-NCCITs upon SVA
400 repression (Fig. 7F). In summary, our functional experiments indicate a widespread role for human-
401 specific SVA transposons as cis-regulatory elements during hippocampal neurogenesis.
402
403 **Discussion**
404 The hippocampus is susceptible to specific neurodegenerative disorders such as Alzheimer's
405 Disease but may have also played an important role in the evolution of human cognition. Spatial memory,
406 which is attributed to the hippocampus, may have contributed to the geographic expansion of ancient
407 humans. Characterizing the human-specific gene regulatory networks of hippocampal development
408 provides insight into its role in human evolution. Though there is no consensus on whether the cognitive
409 phenotypes seen in AD are uniquely human, understanding the unique properties of the human
410 hippocampus may lead the way to potential treatments. Thus, the work described here is relevant both
411 in terms of evolutionary developmental biology and evolutionary medicine.
412 To study the evolution of the human hippocampus from a developmental standpoint, we
413 investigated the extent to which TEs contributed to gene expression profiles of human and chimpanzee
414 hippocampal intermediate progenitors. TEs account for nearly ~50% of the human genome composition
415 and many elegant studies have established that at least a fraction of the TEs can regulate the host genes
416 in humans and other primates (Chuong et al., 2013; Chuong et al., 2016; Cosby et al., 2021; del Rosario et
417 al., 2014; Du et al., 2016; Fuentes et al., 2018; Jacques et al., 2013; Judd et al., 2021; Lynch et al., 2011;
418 Lynch et al., 2015; Mika et al., 2021; Modzelewski et al., 2021; Okhovat et al., 2020; Rayan et al., 2016;
419 Schmidt et al., 2012; Sundaram et al., 2014; Trizzino et al., 2017; Trizzino et al., 2018; Ward et al., 2018;
420 Xie et al., 2013).
421 The hpIPCs are a transient progenitor population during a critical developmental stage in the Sub-
422 ventricular zone (SVZ) of both the hippocampus and neocortex (Bulfone et al., 1999; Cipriani et al., 2016;
423 Englund, 2005; Kimura et al., 1999), and there is consensus that this progenitor population may have
424 played a role in the evolution of brain volume in mammals (Florio and Huttner, 2014; Martínez-Cerdeño
425 et al., 2006). To study the developmental evolution of human hpIPCs, we leveraged a comparative
426 approach centered on differentiating human and chimpanzee iPSCs into a neuronal population which
427 closely recapitulates hpIPCs, as demonstrated by the high expression of signature markers such as TBR2
428 and PAX6.
429 The transcriptomes of human and chimpanzee hpIPCs have not been previously compared, largely
430 due to the limited availability of primary tissue. Here, we carried out this comparison using our iPSC-
431 derived system and identified profound differences between the two species, with over 2,500 genes
432 differentially expressed at this stage. These genes include several which were previously associated with

433 cognitive function, language, neurodevelopment and neurodegeneration, many of which are upregulated
434 in humans relative to the chimpanzee (e.g. *MTRNR2L8*, *DHX40*, *VPS13B*, *WDFY2*, *PURB*).
435 We demonstrate that species-specific enhancers significantly contributed to the gene expression
436 differences we identified. This is consistent with recent studies that used primary brain tissues from
437 several species to profile species-specific cis-regulatory activity in mammals (Emera et al., 2016; Reilly et
438 al., 2015). Importantly, we find that these species-specific enhancers are enriched for young transposable
439 elements. Several studies have identified evolutionarily young L1 LINEs as active in the brain during
440 different developmental stages, suggesting that they could serve as alternative promoters for many genes
441 involved in neuronal functions (Coufal et al., 2009; Jönsson et al., 2019; Sur et al., 2017; Thomas et al.,
442 2012; Zhao et al., 2019). Here, we identified other young transposable elements, ERVs and SVAs, as
443 regulators of intermediate progenitor gene expression. The identification of ERVs as candidate enhancers
444 in human and chimpanzee intermediate progenitors is consistent with previous studies that
445 demonstrated that ERVs heavily impact the gene regulatory programs during immune response (Chuong
446 et al., 2013; Chuong et al., 2016), in pluripotency maintenance and development (Coluccio et al., 2018;
447 Fuentes et al., 2018; Miao et al., 2020), in the mammalian placenta (Lynch et al., 2011; Lynch et al., 2015;
448 Mika et al., 2021), in the primate liver (Trizzino et al., 2017), and in many cancer types (Ito et al.; Ivancevic
449 and Chuong, 2020; Shah et al., 2021). Similar to the L1s, the ERVs also possess a well-defined cis-regulatory
450 architecture (e.g. they have their own promoter), and this may have had a role in the co-option of these
451 TEs as functioning cis-regulatory elements.
452 The SVAs are particularly interesting from a human evolution standpoint, given that half of the
453 known copies are exclusively present in our species. Moreover, SVAs are among the few transposable
454 elements that still exhibit active transposition in the human genome. We and others have previously
455 revealed that SVAs can be source of enhancers in primates (Playfoot et al., 2021; Pontis et al., 2019;
456 Trizzino et al., 2017; Trizzino et al., 2018). The repression of some SVAs by specific zinc-finger proteins at
457 specific stages of neuronal development is also a crucial mechanism for successful neurogenesis (Playfoot
458 et al., 2021; Pontis et al., 2019; Turelli et al., 2020). Here, we demonstrate that SVAs are pervasive
459 regulators of hippocampal neurogenesis and they act as enhancers in the hippocampal intermediate
460 progenitor population. By using CRISPR-interference (CRISPR-i), we show that repressing hundreds of
461 normally "de-repressed" SVAs alters the expression of thousands of genes. Intriguingly, our CRISPR-i
462 experiments revealed that global SVA repression leads to the attenuation of the expression levels of over
463 a quarter of the ~2,500 genes previously identified as with human-specific gene expression signal in
464 hippocampal intermediate progenitors. These include critical neurodevelopmental regulators such as
465 *FOXP2*, *HAND2*, *MEF2C*, *SOX2*, and *SOX4*. In normal conditions, these genes are more highly expressed in
466 humans relative to the chimpanzee, but upon SVA repression these differences were diminished.
467 In conclusion, our findings indicate that the development of hippocampal neurons has been
468 profoundly affected by the domestication of young transposable elements. These young TEs have been
469 co-opted as functional enhancers and promoters and ultimately rewired the expression of hundreds of
470 critically important neuro-regulators in the developing human hippocampus. The human-specific gene
471 expression and the associated TE-derived enhancers we have identified here may play important roles
472 both in human evolution and neurodegenerative disease.
473
474
475 **Materials and Methods**
476
477 **Antibodies**

Antibody	Catalog number	Use	Dilution
TBR2	Abcam ab216870	ChIP	15 ug per ChIP
TBR2	Santa Cruz sc - 293481	Western Blot	1:500

H3K9me3	Abcam ab8898	ChIP	3 ug per ChIP
GAPDH	CST #5174	Western Blot	1:1000
H3K27ac	Abcam ab4729	ChIP	3 ug per ChIP
HRP-conjugated anti-rabbit	Cell Signaling 7074S	Western Blot	1:10000
HRP-conjugated anti-mouse	Cell Signaling 7076S	Western Blot	1:10000

478
479
480 **Human and Chimpanzee iPSC cultures**
481 The human male iPSC line denoted as SV20 was obtained from the University of Pennsylvania, where it
482 was generated, and validated by the expression of pluripotency markers and differentiation into various
483 cell types in multiple studies (Pagliaroli et al., 2021; Pashos et al., 2017; Yang et al., 2015; Zhang et al.,
484 2015). The human female iPSC line GM 23716 was obtained from the Coriell Institute for Medical Research
485 (Camden, NJ) and validated by the expression of pluripotency markers and differentiation into cranial
486 neural crest cells in a previous study (Pagliaroli et al., 2021). The human female iPSC line 21792 and all
487 three of the chimpanzee iPSC lines were obtained from the laboratory of Yoav Gilad at the University of
488 Chicago and validated in previous studies (Gallego Romero et al., 2015; Ward et al., 2018)
489 The iPSC lines were expanded in feeder-free, serum-free mTeSR™1 medium (STEMCELL Technologies).
490 Cells were passaged ~1:10 at 80% confluency using ReLeSR (STEMCELL Technologies) and small cell
491 clusters (50–200 cells) were subsequently plated on tissue culture dishes coated overnight with Geltrex™
492 LDEV-Free hESC-qualified Reduced Growth Factor Basement Membrane Matrix (Fisher-Scientific).
493
494 **hpIPC Differentiation**
495 The iPSC lines were differentiated into hpIPCs as previously described (Yu et al. 2014). 3 batches comprised
496 of one human and one chimpanzee iPSC line each were cultured until approximately 50-70% confluence
497 was reached and then treated with the hpIPC media for five days prior to collection for RNA-seq, ATAC-
498 seq, ChIP-seq or immunofluorescence. The hpIPC media consists of DMEM/F12, 0.5X N2, 0.5X B27, DKK-
499 1, cyclopamine, Noggin, and SB431542 as well as the antibiotic Penn-Strep.
500
501 **Western Blot**
502 For total lysate, cells were harvested and washed three times in 1X PBS and lysed in RIPA buffer (50mM
503 Tris-HCl pH7.5, 150mM NaCl, 1% Igepal, 0.5% sodium deoxycholate, 0.1% SDS, 500uM DTT) with proteases
504 inhibitors. Twenty µg of whole cell lysate were loaded in Novex WedgeWell 4-20% Tris-Glycine Gel
505 (Invitrogen) and separated through gel electrophoresis (SDS-PAGE) Tris-Glycine-SDS buffer (Invitrogen).
506 The proteins were then transferred to ImmunBlot PVDF membranes (ThermoFisher) for antibody probing.
507 Membranes were incubated with 10% BSA in TBST for 30 minutes at room temperature (RT), then
508 incubated for variable times with the suitable antibodies diluted in 5% BSA in 1X TBST, washed with TBST
509 and incubated with a dilution of 1:10000 of secondary antibody for one hour at RT. The antibody was then
510 visualized using Super Signal West Dura Extended Duration Substrat (ThermoFisher) and imaged with
511 Amersham Imager 680. Used antibodies are listed in the "Antibodies" section of the methods.
512
513 **Immunofluorescence**
514 Upon fixation (4% PFA for 10 minutes), cells were first treated with 10 mM sodium citrate (pH = 6) for 10
515 minutes at 95 C. The cells were then permeabilized in blocking solution (0.1% Triton X-100, re PBS, 5%
516 normal donkey serum) for one hour at room temperature and then incubated overnight at 4 C with the
517 primary antibody of interest. Subsequently, the cells were treated with blocking solution for 30 minutes
518 at room temperature and then fluorophore-conjugated secondary antibodies for 2 hours at room

519 temperature. The cell nuclei were stained with 4,6-diamidine-2-phenylindole dihydrochloride (DAPI;
520 Sigma-Aldrich; 50 mg/ml in PBS for 15 min at RT). Immunostained cells analyzed with confocal microscopy,
521 using a Nikon A1R+. Images were captured with a 40X objectives and a pinhole of 1.2 Airy unit. Analyses
522 were performed in sequential scanning mode to rule out cross-bleeding between channels. Used
523 antibodies are listed in the "Antibodies" section of the methods.
524
525 **Real-time quantitative polymerase chain reaction (RT-qPCR)**
526 Cells were lysed in Tri-reagent and RNA was extracted using the Direct-zol RNA MiniPrep kit (Zymo
527 research). 600ng of template RNA was retrotranscribed into cDNA using RevertAid first strand cDNA
528 synthesis kit (Thermo Scientific) according to manufacturer directions. 15ng of cDNA were used for each
529 real-time quantitative PCR reaction with 0.1 µM of each primer, 10 µL of PowerUp™ SYBR™ Green Master
530 Mix (Applied Biosystems) in a final volume of 20 µl, using QuantStudio 3 Real-Time PCR System (Applied
531 Biosystem). Thermal cycling parameters were set as following: 3 minutes at 95°C, followed by 40 cycles of
532 10 s at 95°C, 20 s at 63°C followed by 30 s at 72°C. Each sample was run in triplicate. 18S rRNA was used
533 as normalizer. Primer sequences are reported in Supplementary Data 1.
534
535 **ChIP-Seq and ChiP-qPCR**
536 Samples from different conditions were processed together to prevent batch effects. 15 million cells were
537 cross-linked with 1% formaldehyde for 5 min at room temperature, quenched with 125mM glycine,
538 harvested and washed twice with 1× PBS. The pellet was resuspended in ChIP lysis buffer (150 mM NaCl,
539 1% Triton X-100, 0,7% SDS, 500 µM DTT, 10 mM Tris-HCl, 5 mM EDTA) and chromatin was sheared to an
540 average length of 200–500 bp, using a Covaris S220 Ultrasonicator. The chromatin lysate was diluted with
541 SDS-free ChIP lysis buffer. For ChIP-seq, 10 µg of antibody (3 µg for H3K27ac) was added to 5 µg of
542 sonicated chromatin along with Dynabeads Protein A magnetic beads (Invitrogen) and incubated at 4 °C
543 overnight. On day 2, beads were washed twice with each of the following buffers: Mixed Micelle Buffer
544 (150 mM NaCl, 1% Triton X-100, 0.2% SDS, 20 mM Tris-HCl, 5 mM EDTA, 65% sucrose), Buffer 500 (500
545 mM NaCl, 1% Triton X-100, 0.1% Na deoxycholate, 25 mM HEPES, 10 mM Tris-HCl, 1 mM EDTA),
546 LiCl/detergent wash (250 mM LiCl, 0.5% Na deoxycholate, 0.5% NP-40, 10 mM Tris-HCl, 1 mM EDTA) and
547 a final wash was performed with 1× TE. Finally, beads were resuspended in 1× TE containing 1% SDS and
548 incubated at 65 °C for 10 min to elute immunocomplexes. Elution was repeated twice, and the samples
549 were further incubated overnight at 65 °C to reverse cross-linking, along with the untreated input (5% of
550 the starting material). On day 3, after treatment with 0.5 mg/ml Proteinase K for 1h at 65 °C, DNA was
551 purified with Zymo ChIP DNA Clear Concentrator kit and quantified with QUBIT.
552 For all ChIP-seq experiments, barcoded libraries were made with NEB ULTRA II DNA Library Prep Kit for
553 Illumina, and sequenced on Illumina NextSeq 500, producing 100 bp PE reads.
554 For ChIP-qPCR, on day 1 the sonicated lysate was aliquot into single immunoprecipitations of 2.5×10^6
555 cells each. A specific antibody or a total rabbit IgG control was added to the lysate along with Protein A
556 magnetic beads (Invitrogen) and incubated at 4 °C overnight. On day3, ChIP eluates and input were
557 assayed by real-time quantitative PCR in a 20 µl reaction with the following: 0.4 µM of each primer, 10 µl
558 of PowerUp SYBR Green (Applied Biosystems), and 5 µl of template DNA (corresponding to 1/40 of the
559 elution material) using the fast program on QuantStudio qPCR machine (Applied Biosystems). Thermal
560 cycling parameters were: 20sec at 95 °C, followed by 40 cycles of 1sec at 95°C, 20sec at 60°C. Used
561 antibodies are listed in the "Antibodies" section of the methods.
562
563 **ChIP-seq Analyses**
564 After removing the adapters with TrimGalore!, the sequences were aligned to the reference hg19, using
565 Burrows Wheeler Alignment tool (BWA), with the MEM algorithm[64] . Uniquely mapping aligned reads
566 were filtered based on mapping quality (MAPQ > 10) to restrict our analysis to higher quality and likely

567 uniquely mapped reads, and PCR duplicates were removed. We called peaks for each individual using
568 MACS2[65] (H3K27ac) or Homer[66], at 5% FDR, with default parameters.
569
570 **RNA-Seq**
571 Cells were lysed in Tri-reagent (Zymo research) and total RNA was extracted using Quick-RNA Miniprep
572 kit (Zymo research) according to the manufacturer's instructions. RNA was further quantified using
573 DeNovix DS-11 Spectrophotometer while the RNA integrity was checked on Bioanalyzer 2100 (Agilent).
574 Only samples with RIN value above 8.0 were used for transcriptome analysis. RNA libraries were prepared
575 using 1 µg of total RNA input using NEBNext® Poly(A) mRNA Magnetic Isolation Module, NEBNext®
576 UltraTM II Directional RNA Library Prep Kit for Illumina® and NEBNext® UltraTM II DNA Library Prep Kit for
577 Illumina® according to the manufacturer's instructions (New England Biolabs).
578
579 **Single-cell RNA-Seq**
580 Cells from both species at each time were first incubated with Accutase at 37 C for 7 minutes. The cells
581 were collected with DMEM/F12 and centrifuged for 5 minutes at 150 xg. The cells were resuspended in
582 10% FBS in DMEM and strained with a 40 um cell strainer to create a single-cell suspension. After
583 confirming >90% viability with a Countess III, the cells were processed with the 10X Genomics Cell
584 Multiplexing Oligo protocol.
585
586 **Single-cell RNA-Seq analyses**
587 10X Cell Ranger (Zheng et al., 2017) was used to demultiplex and map the scRNA-seq data, with the tools
588 cellranger multi and cellranger mkfastq. Seurat 4 (Hao et al., 2021) was used for individual analysis of
589 scRNA-seq data, as well as integration of human and chimpanzee datasets. The human genes used to
590 classify the cells into the G1/S/G2M phases of the cell cycle were obtained from
591 https://github.com/hbctraining/scRNA-seq/blob/master/lessons/06_SC_SCT_and_integration.md
592
593 **RNA-Seq Analyses**
594 After removing the adapters with TrimGalore!, Kallisto (Bray et al., 2016) was used to count reads mapping
595 to each gene. We analyzed differential gene expression levels with DESeq2 (Love et al., 2014) , with the
596 following model: design = ~condition, where condition indicates either Human or Chimpanzee.
597
598 **ATAC-Seq**
599 For ATAC-Seq experiments, 50,000 cells per condition were processed as described in the original ATAC-
600 seq protocol paper (Buenrostro et al. 2013). ATAC-seq data were processed with the same pipeline
601 described for ChIP-seq, with one modification: all mapped reads were offset by +4 bp for the forward-
602 strand and -5 bp for the reverse-strand. Peaks were called using MACS2 (Zhang et al., 2008, 2).
603
604 **Generation of the NCCIT-dCas9KRAB-SVAsgRNA Stable Cell Line**
605 This cell line was generated in our recent study (Barnada et al., 2022). Briefly, a dCas9-KRAB was cloned
606 into a piggyBac transposon containing ampicillin and puromycin resistance (Addgene) which was obtained
607 from the Wysocka Lab at Stanford University. The piggyBac dCas9-KRAB doxycycline-inducible plasmid,
608 along with a piggyBac transposase (Cell Signaling), were transfected into NCCIT cells (ATCC) at ~70%
609 confluency using a 6:1 ratio of Fugene HD (Promega) for 48 hours in ATCC-formulated RPMI media. Two
610 days post-transfection, the media was changed and the transfected cells were selected for using 1µg
611 puromycin per 1mL media. A piggyBac transposon plasmid containing two sgRNAs (SVAsgRNA1:
612 5'CTCCCTAATCTCAAGTACCC 3' and SVAsgRNA2: 5' TGTTTCAGAGAGCACGGGGT 3'; Integrated DNA
613 Technologies) targeting ~80% of all annotated SVAs in humans (Pontis et al. 2019), along with a piggyBac
614 transposase, were transfected into the NCCIT-dCas9KRAB cells using a 6:1 ratio of Fugene HD for 48 hours

615 in ATCC-formulated RPMI media. Two days post-transfection, the media was changed and the transfected
616 cells were selected for using 400µg geneticin per 1mL of media in addition to 1µg puromycin per 1mL
617 media. The NCCIT-dCas9KRAB-SVAsgRNA cell line was maintained in ATCC-formulated RPMI media
618 supplemented with 10% tet-free FBS, 1% L-glutamine, 1µg/mL puromycin, and 400µg/mL geneticin and
619 incubated at 5% CO_2, 20% O_2 at 37°C.
620
621 **Retinoic Acid-induced neuronal differentiation and CRISPR-interference of NCCIT-dCas9KRAB-**
622 **SVAsgRNA cells**
623 The NCCIT-dCas9KRAB-SVAsgRNA cells, at ~20% confluency, were treated with 10µM retinoic acid (RA)
624 per 10mL media for 1 week to induce neuronal differentiation. At day 4, the media was refreshed and the
625 cells were additionally treated with 2µg doxycycline per 1mL of media for 3 days. The cells were collected
626 on day 7 of RA treatment (day 3 of doxycycline treatment) for qPCR and genomic experiments. Expression
627 of the doxycycline-inducible dCas9 was verified via western blot.
628
629 **Statistical and genomic analyses**
630 All statistical analyses were performed using R v3.3.1 or Graphpad Prism version 9.2.0 for Mac OS X .
631 BEDtools v2.27.1 (Quinlan and Hall, 2010) was used for genomic analyses. Pathway analysis was
632 performed with Ingenuity Pathway Analysis Suite (QIAGEN Inc.,
633 https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis). Motif analyses were
634 performed using the MEME-Suite (Bailey et al., 2015), and specifically with the Meme-ChIP application.
635 Fasta files of the regions of interest were produced using BEDTools v2.27.1. Shuffled input sequences
636 were used as background. E-values < 0.001 were used as threshold for significance. Orthologous ATAC-
637 seq regions were identified using the University of California Santa Cruz Genome Browser tool LiftOver.
638
639 **Acknowledgements**
640 The authors are grateful to Dr. Yoav Gilad (University of Chicago) and with the Yerkes National Primate
641 Research Center (Emory University, Atlanta) for providing all the chimpanzee lines and one of the human
642 lines, and to Dr. Joanna Wysocka's lab, and in particular Dr. Raquel Fueyo (Stanford University), for
643 providing CRISPR reagents and critical advice. The authors thank the Thomas Jefferson Stem Cell and
644 Regenerative Neuroscience Center for the support in the process of optimization of the iPSC
645 differentiation. The authors are grateful to Trizzino lab members Luca Pagliaroli, PhD, Chiara Scopa, PhD
646 and Connor Ott for help with specific analyses and experiments. The authors thank the Genomic Facility
647 at The Wistar Institute (Philadelphia, PA) for the Next Generation Illumina Sequencing. The PAX6 antibody
648 developed by Kawakami, A. was obtained from the Developmental Hybridoma Bank, created by the NICHD
649 of the NIH and maintained at The University of Iowa, Department of Biology, Iowa City, IA 52242.
650
651 **Competing interests**
652 The authors declare no competing interests.
653
654 **Funding**
655 For this work, M.T. was funded by the National Institute of Health (NIH-NIGMS R35GM138344) and by the
656 G. Harold and Leila Y. Mathers Foundation.
657
658 **Data availability**
659 The original genome-wide data generated in this study have been deposited in the GEO database under
660 accession code GSE189347.
661
662

663 **Author contributions**
664 MT and SP designed the project. SP performed most of the experiments. SMB carried out the CRISPR
665 experiments in NCCIT cells. SP and MT analyzed the data and wrote the manuscript. CL and JIM supervised
666 SP in the scRNA-seq analyses. All the authors read and approved the manuscript.
667
668 **References**

669 Abrajano, J. J., Qureshi, I. A., Gokhan, S., Zheng, D., Bergman, A. and Mehler, M. F. (2009). REST and
670 CoREST Modulate Neuronal Subtype Specification, Maturation and Maintenance. *PLOS ONE* 4,
671 e7936.

672 Agoglia, R. M., Sun, D., Birey, F., Yoon, S.-J., Miura, Y., Sabatini, K., Pașca, S. P. and Fraser, H. B. (2021).
673 Primate cell fusion disentangles gene regulatory divergence in neurodevelopment. *Nature* 592,
674 421–427.

675 Andersen, J., Urbán, N., Achimastou, A., Ito, A., Simic, M., Ullom, K., Martynoga, B., Lebel, M., Göritz, C.,
676 Frisén, J., et al. (2014). A Transcriptional Mechanism Integrating Inputs from Extracellular Signals
677 to Activate Hippocampal Stem Cells. *Neuron* 83, 1085–1097.

678 Arnold, S. J., Huang, G.-J., Cheung, A. F. P., Era, T., Nishikawa, S.-I., Bikoff, E. K., Molnar, Z., Robertson, E.
679 J. and Groszer, M. (2008). The T-box transcription factor Eomes/Tbr2 regulates neurogenesis in
680 the cortical subventricular zone. *Genes & Development* 22, 2479–2484.

681 Aruga, J. (2004). The role of Zic genes in neural development. *Mol Cell Neurosci* 26, 205–221.

682 Averaimo, S., Gritti, M., Barini, E., Gasparini, L. and Mazzanti, M. (2014). CLIC1 functional expression is
683 required for cAMP-induced neurite elongation in post-natal mouse retinal ganglion cells. *Journal*
684 *of Neurochemistry* 131, 444–456.

685 Bailey, T. L., Johnson, J., Grant, C. E. and Noble, W. S. (2015). The MEME Suite. *Nucleic Acids Research*
686 43, W39–W49.

687 Barger, N., Hanson, K. L., Teffer, K., Schenker-Ahmed, N. M. and Semendeferi, K. (2014). Evidence for
688 evolutionary specialization in human limbic structures. *Front. Hum. Neurosci.* 8,.

689 Barnada, S. M., Isopi, A., Tejada-Martinez, D., Goubert, C., Patoori, S., Pagliaroli, L., Tracewell, M. and
690 Trizzino, M. (2022). Genomic features underlie the co-option of SVA transposons as cis-
691 regulatory elements in human pluripotent stem cells. *PLOS Genetics* 18, e1010225.

692 Bekinschtein, P., Katche, C., Slipczuk, L. N., Igaz, L. M., Cammarota, M., Izquierdo, I. and Medina, J. H.
693 (2007). mTOR signaling in the hippocampus is necessary for memory formation. *Neurobiology of*
694 *Learning and Memory* 87, 303–307.

695 Bray, N. L., Pimentel, H., Melsted, P. and Pachter, L. (2016). Near-optimal probabilistic RNA-seq
696 quantification. *Nat Biotechnol* 34, 525–527.

697 Bulfone, A., Martinez, S., Marigo, V., Campanella, M., Basile, A., Quaderi, N., Gattuso, C., Rubenstein, J.
698 L. R. and Ballabio, A. (1999). Expression pattern of the Tbr2 (Eomesodermin) gene during mouse
699 and chick brain development. *Mechanisms of Development* 84, 133–138.

700 Burgess, N., Maguire, E. A. and O'Keefe, J. (2002). The Human Hippocampus and Spatial and Episodic
701 Memory. *Neuron* 35, 625–641.

702 Chen, X., Wang, S., Zhou, Y., Han, Y., Li, S., Xu, Q., Xu, L., Zhu, Z., Deng, Y., Yu, L., et al. (2018). Phf8
703 histone demethylase deficiency causes cognitive impairments through the mTOR pathway. *Nat*
704 *Commun* 9, 114.

705 Cheval, H., Chagneau, C., Levasseur, G., Veyrac, A., Faucon-Biguet, N., Laroche, S. and Davis, S. (2012).
706 Distinctive features of Egr transcription factor regulation and DNA binding activity in CA1 of the
707 hippocampus in synaptic plasticity and consolidation and reconsolidation of fear memory.
708 *Hippocampus* 22, 631–642.

709 Chuong, E. B., Rumi, M. A. K., Soares, M. J. and Baker, J. C. (2013). Endogenous retroviruses function as
710 species-specific enhancer elements in the placenta. *Nat Genet* 45, 325–329.

711 Chuong, E. B., Elde, N. C. and Feschotte, C. (2016). Regulatory evolution of innate immunity through co-
712 option of endogenous retroviruses. *Science* 351, 1083–1087.

713 Cipriani, S., Nardelli, J., Verney, C., Delezoide, A.-L., Guimiot, F., Gressens, P. and Adle-Biassette, H.
714 (2016). Dynamic Expression Patterns of Progenitor and Pyramidal Neuron Layer Markers in the
715 Developing Human Hippocampus. *Cereb. Cortex* 26, 1255–1271.

716 Coluccio, A., Ecco, G., Duc, J., Offner, S., Turelli, P. and Trono, D. (2018). Individual retrotransposon
717 integrants are differentially controlled by KZFP/KAP1-dependent histone methylation, DNA
718 methylation and TET-mediated hydroxymethylation in naïve embryonic stem cells. *Epigenetics &*
719 *Chromatin* 11, 7.

720 Cosby, R. L., Judd, J., Zhang, R., Zhong, A., Garry, N., Pritham, E. J. and Feschotte, C. (2021). Recurrent
721 evolution of vertebrate transcription factors by transposase capture. *Science* 371, eabc6405.

722 Coufal, N. G., Garcia-Perez, J. L., Peng, G. E., Yeo, G. W., Mu, Y., Lovci, M. T., Morell, M., O'Shea, K. S.,
723 Moran, J. V. and Gage, F. H. (2009). L1 retrotransposition in human neural progenitor cells.
724 *Nature* 460, 1127–1131.

725 Crawley, J. N., Heyer, W.-D. and LaSalle, J. M. (2016). Autism and Cancer Share Risk Genes, Pathways,
726 and Drug Targets. *Trends in Genetics* 32, 139–146.

727 Datson, N. a., Morsink, M. c., Steenbergen, P. j., Aubert, Y., Schlumbohm, C., Fuchs, E. and de Kloet, E. r.
728 (2009). A molecular blueprint of gene expression in hippocampal subregions CA1, CA3, and DG is
729 conserved in the brain of the common marmoset. *Hippocampus* 19, 739–752.

730 Davis, L., Onn, I. and Elliott, E. (2018). The emerging roles for the chromatin structure regulators CTCF
731 and cohesin in neurodevelopment and behavior. *Cell. Mol. Life Sci.* 75, 1205–1214.

732 del Rosario, R. C. H., Rayan, N. A. and Prabhakar, S. (2014). Noncoding origins of anthropoid traits and a
733 new null model of transposon functionalization. *Genome Res* 24, 1469–1484.

734 Du, J., Leung, A., Trac, C., Lee, M., Parks, B. W., Lusis, A. J., Natarajan, R. and Schones, D. E. (2016).
735 Chromatin variation associated with liver metabolism is mediated by transposable elements.
736 *Epigenetics & Chromatin* 9, 28.

737 Duyckaerts, C., Delatour, B. and Potier, M.-C. (2009). Classification and basic pathology of Alzheimer
738 disease. *Acta Neuropathol* 118, 5–36.

739 Edler, M. K., Sherwood, C. C., Meindl, R. S., Hopkins, W. D., Ely, J. J., Erwin, J. M., Mufson, E. J., Hof, P. R.
740 and Raghanti, M. A. (2017). Aged chimpanzees exhibit pathologic hallmarks of Alzheimer's
741 disease. *Neurobiology of Aging* 59, 107–120.

742 Eichenbaum, H. (2017a). The role of the hippocampus in navigation is memory. *Journal of*
743 *Neurophysiology* 117, 1785–1796.

744 Eichenbaum, H. (2017b). Prefrontal–hippocampal interactions in episodic memory. *Nat Rev Neurosci* 18,
745 547–558.

746 Emera, D., Yin, J., Reilly, S. K., Gockley, J. and Noonan, J. P. (2016). Origin and evolution of
747 developmental enhancers in the mammalian neocortex. *PNAS* 113, E2617–E2626.

748 Enard, W. (2011). FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution.
749 *Current Opinion in Neurobiology* 21, 415–424.

750 Enard, W., Khaitovich, P., Klose, J., Zöllner, S., Heissig, F., Giavalisco, P., Nieselt-Struwe, K., Muchmore,
751 E., Varki, A., Ravid, R., et al. (2002). Intra- and Interspecific Variation in Primate Gene Expression
752 Patterns. *Science* 296, 340–343.

753 Englund, C. (2005). Pax6, Tbr2, and Tbr1 Are Expressed Sequentially by Radial Glia, Intermediate
754 Progenitor Cells, and Postmitotic Neurons in Developing Neocortex. *Journal of Neuroscience* 25,
755 247–251.

756 Ferrón, S. R., Charalambous, M., Radford, E., McEwen, K., Wildner, H., Hind, E., Morante-Redolat, J. M.,
757 Laborda, J., Guillemot, F., Bauer, S. R., et al. (2011). Postnatal loss of Dlk1 imprinting in stem
758 cells and niche astrocytes regulates neurogenesis. *Nature* 475, 381–385.

759 Finch, C. E. and Austad, S. N. (2015). Commentary: is Alzheimer's disease uniquely human? *Neurobiology*
760 *of Aging* 36, 553–555.

761 Florio, M. and Huttner, W. B. (2014). Neural progenitors, neurogenesis and the evolution of the
762 neocortex. *Development* 141, 2182–2194.

763 Fortress, A. M., Fan, L., Orr, P. T., Zhao, Z. and Frick, K. M. (2013). Estradiol-induced object recognition
764 memory consolidation is dependent on activation of mTOR signaling in the dorsal hippocampus.
765 *Learn. Mem.* 20, 147–155.

766 Fuentes, D. R., Swigut, T. and Wysocka, J. (2018). Systematic perturbation of retroviral LTRs reveals
767 widespread long-range effects on human gene regulation. *eLife* 7, e35989.

768 Gallego Romero, I., Pavlovic, B. J., Hernando-Herraez, I., Zhou, X., Ward, M. C., Banovich, N. E., Kagan, C.
769 L., Burnett, J. E., Huang, C. H., Mitrano, A., et al. (2015). A panel of induced pluripotent stem
770 cells from chimpanzees: a resource for comparative functional genomics. *eLife* 4, e07103.

771 Gokhman, D., Agoglia, R. M., Kinnebrew, M., Gordon, W., Sun, D., Bajpai, V. K., Naqvi, S., Chen, C., Chan,
772 A., Chen, C., et al. (2021). Human–chimpanzee fused cells reveal cis -regulatory divergence
773 underlying skeletal evolution. *Nature Genetics* 53, 467–476.

774 Graber, T. E., McCamphill, P. K. and Sossin, W. S. (2013). A recollection of mTOR signaling in learning and
775 memory. *Learn. Mem.* 20, 518–530.

776 Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C.,
777 Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573-
778 3587.e29.

779 Hevner, R. F. (2019). Intermediate progenitors and Tbr2 in cortical development. *J. Anat.* 235, 616–625.

780 Hickey, S. L., Berto, S. and Konopka, G. (2019). Chromatin Decondensation by FOXP2 Promotes Human
781 Neuron Maturation and Expression of Neurodevelopmental Disease Genes. *Cell Reports* 27,
782 1699-1711.e9.

783 Hodge, R. D., Nelson, B. R., Kahoud, R. J., Yang, R., Mussar, K. E., Reiner, S. L. and Hevner, R. F. (2012).
784 Tbr2 Is Essential for Hippocampal Lineage Progression from Neural Stem Cells to Intermediate
785 Progenitors and Neurons. *Journal of Neuroscience* 32, 6275–6287.

786 Ito, J., Kimura, I., Soper, A., Coudray, A., Koyanagi, Y., Nakaoka, H., Inoue, I., Turelli, P., Trono, D. and
787 Sato, K. Endogenous retroviruses drive KRAB zinc-finger protein family expression for tumor
788 suppression. *Science Advances* 6, eabc3020.

789 Ivancevic, A. and Chuong, E. B. (2020). Transposable elements teach T cells new tricks. *PNAS* 117, 9145–
790 9147.

791 Jacques, P.-É., Jeyakani, J. and Bourque, G. (2013). The Majority of Primate-Specific Regulatory
792 Sequences Are Derived from Transposable Elements. *PLOS Genetics* 9, e1003504.

793 Jönsson, M. E., Ludvik Brattås, P., Gustafsson, C., Petri, R., Yudovich, D., Pircs, K., Verschuere, S.,
794 Madsen, S., Hansson, J., Larsson, J., et al. (2019). Activation of neuronal genes via LINE-1
795 elements upon global DNA demethylation in human neural progenitors. *Nat Commun* 10, 3182.

796 Judd, J., Sanderson, H. and Feschotte, C. (2021). Evolution of mouse circadian enhancers from
797 transposable elements. *Genome Biol* 22, 193.

798 Kamboh, M. I., Fan, K.-H., Yan, Q., Beer, J. C., Snitz, B. E., Wang, X., Chang, C.-C. H., Demirci, F. Y.,
799 Feingold, E. and Ganguli, M. (2019). Population-based genome-wide association study of
800 cognitive decline in older adults free of dementia: identification of a novel locus for the
801 attention domain. *Neurobiol Aging* 84, 239.e15-239.e24.

802 Kimura, N., Nakashima, K., Ueno, M., Kiyama, H. and Taga, T. (1999). A novel mammalian T-box-
803 containing gene, Tbr2, expressed in mouse developing brain. *Developmental Brain Research*
804 115, 183–193.

805 King, M. and Wilson, A. (1975). Evolution at two levels in humans and chimpanzees. *Science* 188, 107–
806 116.

807 Kittappa, R., Chang, W. W., Awatramani, R. B. and McKay, R. D. G. (2007). The foxa2 Gene Controls the
808 Birth and Spontaneous Degeneration of Dopamine Neurons in Old Age. *PLOS Biology* 5, e325.

809 Kohane, A. A. and Wood, T. R. (2021). *Neurodevelopmental clustering of gene expression identifies lipid*
810 *metabolism genes associated with neuroprotection and neurodegeneration*.

811 Kolehmainen, J., Black, G. C. M., Saarinen, A., Chandler, K., Clayton-Smith, J., Träskelin, A.-L., Perveen, R.,
812 Kivitie-Kallio, S., Norio, R., Warburg, M., et al. (2003). Cohen syndrome is caused by mutations in
813 a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-
814 mediated sorting and intracellular protein transport. *Am J Hum Genet* 72, 1359–1369.

815 Liégeois, F. J., Hildebrand, M. S., Bonthrone, A., Turner, S. J., Scheffer, I. E., Bahlo, M., Connelly, A. and
816 Morgan, A. T. (2016). Early neuroimaging markers of FOXP2 intragenic deletion. *Sci Rep* 6,
817 35192.

818 Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for
819 RNA-seq data with DESeq2. *Genome Biology* 15, 550.

820 Lynch, V. J., Leclerc, R. D., May, G. and Wagner, G. P. (2011). Transposon-mediated rewiring of gene
821 regulatory networks contributed to the evolution of pregnancy in mammals. *Nat Genet* 43,
822 1154–1159.

823 Lynch, V. J., Nnamani, M. C., Kapusta, A., Brayer, K., Plaza, S. L., Mazur, E. C., Emera, D., Sheikh, S. Z.,
824 Grützner, F., Bauersachs, S., et al. (2015). Ancient Transposable Elements Transformed the
825 Uterine Regulatory Landscape and Transcriptome during the Evolution of Mammalian
826 Pregnancy. *Cell Rep* 10, 551–561.

827 MacDermot, K. D., Bonora, E., Sykes, N., Coupe, A.-M., Lai, C. S. L., Vernes, S. C., Vargha-Khadem, F.,
828 McKenzie, F., Smith, R. L., Monaco, A. P., et al. (2005). Identification of FOXP2 Truncation as a
829 Novel Cause of Developmental Speech and Language Deficits. *The American Journal of Human*
830 *Genetics* 76, 1074–1080.

831 Mandal, C., Jung, K. H., Kang, S. C., Choi, M. R., Park, K. S., Chung, I. Y. and Chai, Y. G. (2015). Knocking
832 down of UTX in NCCIT cells enhance cell attachment and promote early neuronal cell
833 differentiation. *BioChip J* 9, 182–193.

834 Marchetto, M. C., Hrvoj-Mihic, B., Kerman, B. E., Yu, D. X., Vadodaria, K. C., Linker, S. B., Narvaiza, I.,
835 Santos, R., Denli, A. M., Mendes, A. P., et al. (2019). Species-specific maturation profiles of
836 human, chimpanzee and bonobo neural cells. *eLife* 8, e37527.

837 Martínez-Cerdeño, V., Noctor, S. C. and Kriegstein, A. R. (2006). The role of intermediate progenitor cells
838 in the evolutionary expansion of the cerebral cortex. *Cereb Cortex* 16 Suppl 1, i152-161.

839 Mathys, H., Davila-Velderrain, J., Peng, Z., Gao, F., Mohammadi, S., Young, J. Z., Menon, M., He, L.,
840 Abdurrob, F., Jiang, X., et al. (2019). Single-cell transcriptomic analysis of Alzheimer's disease.
841 *Nature* 570, 332–337.

842 Miao, B., Fu, S., Lyu, C., Gontarz, P., Wang, T. and Zhang, B. (2020). Tissue-specific usage of transposable
843 element-derived promoters in mouse development. *Genome Biol* 21, 255.

844 Mihalas, A. B., Elsen, G. E., Bedogni, F., Daza, R. A. M., Ramos-Laguna, K. A., Arnold, S. J. and Hevner, R.
845 F. (2016). Intermediate Progenitor Cohorts Differentially Generate Cortical Layers and Require
846 Tbr2 for Timely Acquisition of Neuronal Subtype Identity. *Cell Reports* 16, 92–105.

847 Mika, K., Marinić, M., Singh, M., Muter, J., Brosens, J. J. and Lynch, V. J. (2021). Evolutionary
848 transcriptomics implicates new genes and pathways in human pregnancy and adverse
849 pregnancy outcomes. *eLife* 10, e69584.

850 Modzelewski, A. J., Shao, W., Chen, J., Lee, A., Qi, X., Noon, M., Tjokro, K., Sales, G., Biton, A., Anand, A.,
851 et al. (2021). A mouse-specific retrotransposon drives a conserved Cdk2ap1 isoform essential for
852 development. *Cell* 184, 5541-5558.e22.

853 Mora-Bermúdez, F., Badsha, F., Kanton, S., Camp, J. G., Vernot, B., Köhler, K., Voigt, B., Okita, K., Maricic,
854 T., He, Z., et al. (2016). Differences and similarities between human and chimpanzee neural
855 progenitors during cerebral cortex development. *eLife* 5, e18683.

856 Mukherjee, D., Gonzales, B. J., Ashwal-Fluss, R., Turm, H., Groysman, M. and Citri, A. (2021). Egr2
857 induction in spiny projection neurons of the ventrolateral striatum contributes to cocaine place
858 preference in mice. *eLife* 10, e65228.

859 Okhovat, M., Nevonen, K. A., Davis, B. A., Michener, P., Ward, S., Milhaven, M., Harshman, L., Sohota, A.,
860 Fernandes, J. D., Salama, S. R., et al. (2020). Co-option of the lineage-specific LAVA
861 retrotransposon in the gibbon genome. *PNAS* 117, 19328–19338.

862 Pagliaroli, L., Porazzi, P., Curtis, A. T., Scopa, C., Mikkers, H. M. M., Freund, C., Daxinger, L., Deliard, S.,
863 Welsh, S. A., Offley, S., et al. (2021). Inability to switch from ARID1A-BAF to ARID1B-BAF impairs
864 exit from pluripotency and commitment towards neural crest formation in ARID1B-related
865 neurodevelopmental disorders. *Nat Commun* 12, 6469.

866 Pashos, E. E., Park, Y., Wang, X., Raghavan, A., Yang, W., Abbey, D., Peters, D. T., Arbelaez, J., Hernandez,
867 M., Kuperwasser, N., et al. (2017). Large, Diverse Population Cohorts of hiPSCs and Derived
868 Hepatocyte-like Cells Reveal Functional Genetic Variation at Blood Lipid-Associated Loci. *Cell*
869 *Stem Cell* 20, 558-570.e10.

870 Playfoot, C. J., Duc, J., Sheppard, S., Dind, S., Coudray, A., Planet, E. and Trono, D. (2021). Transposable
871 elements and their KZFP controllers are drivers of transcriptional innovation in the developing
872 human brain. *Genome Res.* 31, 1531–1545.

873 Poirier, R., Cheval, H., Mailhes, C., Charnay, P., Davis, S. and Laroche, S. (2007). Paradoxical Role of an
874 Egr Transcription Factor Family Member, Egr2/Krox20, in Learning and Memory. *Front Behav*
875 *Neurosci* 1, 6.

876 Pontious, A., Kowalczyk, T., Englund, C. and Hevner, R. F. (2008). Role of intermediate progenitor cells in
877 cerebral cortex development. *Dev Neurosci* 30, 24–32.

878 Pontis, J., Planet, E., Offner, S., Turelli, P., Duc, J., Coudray, A., Theunissen, T. W., Jaenisch, R. and Trono,
879 D. (2019). Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic
880 Genome Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* 24, 724-
881 735.e5.

882 Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic
883 features. *Bioinformatics* 26, 841–842.

884 Quinn, J. P. and Bubb, V. J. (2014). SVA retrotransposons as modulators of gene expression. *Mobile*
885 *Genetic Elements* 4, e32102.

886 Raivich, G. (2008). c-Jun expression, activation and function in neural cell death, inflammation and
887 repair. *J Neurochem* 107, 898–906.

888 Raivich, G. and Behrens, A. (2006). Role of the AP-1 transcription factor c-Jun in developing, adult and
889 injured brain. *Prog Neurobiol* 78, 347–363.

890 Rayan, N. A., Del Rosario, R. C. H. and Prabhakar, S. (2016). Massive contribution of transposable
891 elements to mammalian regulatory sequences. *Semin Cell Dev Biol* 57, 51–56.

892 Reilly, S. K., Yin, J., Ayoub, A. E., Emera, D., Leng, J., Cotney, J., Sarro, R., Rakic, P. and Noonan, J. P.
893 (2015). Evolutionary genomics. Evolutionary changes in promoter and enhancer activity during
894 human corticogenesis. *Science* 347, 1155–1159.

895 Rétaux, S., Bourrat, F., Joly, J.-S. and Hinaux, H. (2013). Perspectives in Evo-Devo of the Vertebrate Brain.
896 In *Advances in Evolutionary Developmental Biology*, pp. 151–172. John Wiley & Sons, Ltd.

897 Reuter, M. S., Krumbiegel, M., Schlüter, G., Ekici, A. B., Reis, A. and Zweier, C. (2017). Haploinsufficiency
898 of NR4A2 is associated with a neurodevelopmental phenotype with prominent language
899 impairment. *Am J Med Genet A* 173, 2231–2234.

900 Sakamoto, K., Karelina, K. and Obrietan, K. (2011). CREB: a multifaceted regulator of neuronal plasticity
901 and protection. *Journal of Neurochemistry* 116, 1–9.

902 Sarkar, A., Mei, A., Paquola, A. C. M., Stern, S., Bardy, C., Klug, J. R., Kim, S., Neshat, N., Kim, H. J., Ku, M.,
903 et al. (2018). Efficient Generation of CA3 Neurons from Human Pluripotent Stem Cells Enables
904 Modeling of Hippocampal Connectivity In Vitro. *Cell Stem Cell* 22, 684-697.e9.

905 Schmidt, D., Schwalie, P. C., Wilson, M. D., Ballester, B., Gonçalves, A., Kutter, C., Brown, G. D., Marshall,
906 A., Flicek, P. and Odom, D. T. (2012). Waves of retrotransposon expansion remodel genome
907 organization and CTCF binding in multiple mammalian lineages. *Cell* 148, 335–348.

908 Sessa, A., Mao, C.-A., Hadjantonakis, A.-K., Klein, W. H. and Broccoli, V. (2008). Tbr2 directs conversion
909 of radial glia into basal precursors and guides neuronal amplification by indirect neurogenesis in
910 the developing neocortex. *Neuron* 60, 56–69.

911 Shah, A. H., Gilbert, M., Ivan, M. E., Komotar, R. J., Heiss, J. and Nath, A. (2021). The role of human
912 endogenous retroviruses in gliomas: from etiological perspectives and therapeutic implications.
913 *Neuro-Oncology* 23, 1647–1655.

914 Sin, C., Li, H. and Crawford, D. A. (2015). Transcriptional Regulation by FOXP1, FOXP2, and FOXP4
915 Dimerization. *Journal of Molecular Neuroscience* 55, 437–448.

916 Sousa, A. M. M., Zhu, Y., Raghanti, M. A., Kitchen, R. R., Onorati, M., Tebbenkamp, A. T. N., Stutz, B.,
917 Meyer, K. A., Li, M., Kawasawa, Y. I., et al. (2017). Molecular and cellular reorganization of neural
918 circuits in the human lineage. *Science* 358, 1027–1032.

919 Squire, L. R. (1992). Memory and the Hippocampus: A Synthesis From Findings With Rats, Monkeys, and
920 Humans.

921 Sundaram, V. and Wysocka, J. (2020). Transposable elements as a potent source of diverse cis-
922 regulatory sequences in mammalian genomes. *Philosophical Transactions of the Royal Society B:*
923 *Biological Sciences* 375, 20190347.

924 Sundaram, V., Cheng, Y., Ma, Z., Li, D., Xing, X., Edge, P., Snyder, M. P. and Wang, T. (2014). Widespread
925 contribution of transposable elements to the innovation of gene regulatory networks. *Genome*
926 *Res* 24, 1963–1976.

927 Sur, D., Kustwar, R. K., Budania, S., Mahadevan, A., Hancks, D. C., Yadav, V., Shankar, S. K. and Mandal, P.
928 K. (2017). Detection of the LINE-1 retrotransposon RNA-binding protein ORF1p in different
929 anatomical regions of the human brain. *Mobile DNA* 8, 17.

930 Taher, N., McKenzie, C., Garrett, R., Baker, M., Fox, N. and Isaacs, G. D. (2014). Amyloid-β Alters the DNA
931 Methylation Status of Cell-fate Genes in an Alzheimer's Disease Model. *Journal of Alzheimer's*
932 *Disease* 38, 831–844.

933 Thomas, C. A., Paquola, A. C. M. and Muotri, A. R. (2012). LINE-1 Retrotransposition in the Nervous
934 System. *Annual Review of Cell and Developmental Biology* 28, 555–573.

935 Tokuyama, M., Kong, Y., Song, E., Jayewickreme, T., Kang, I. and Iwasaki, A. (2018). ERVmap analysis
936 reveals genome-wide transcription of human endogenous retroviruses. *PNAS* 115, 12565–
937 12572.

938 Tomasello, M. and Herrmann, E. (2010). Ape and Human Cognition: What's the Difference? *Current*
939 *Directions in Psychological Science* 19, 3–8.

940 Trizzino, M., Park, Y., Holsbach-Beltrame, M., Aracena, K., Mika, K., Caliskan, M., Perry, G. H., Lynch, V. J.
941 and Brown, C. D. (2017). Transposable elements are the primary source of novelty in primate
942 gene regulation. *Genome Res.*

943 Trizzino, M., Kapusta, A. and Brown, C. D. (2018). Transposable elements generate regulatory novelty in
944 a tissue-specific fashion. *BMC Genomics* 19, 468.

945 Turelli, P., Playfoot, C., Grun, D., Raclot, C., Pontis, J., Coudray, A., Thorball, C., Duc, J., Pankevich, E. V.,
946 Deplancke, B., et al. (2020). Primate-restricted KRAB zinc finger proteins and target
947 retrotransposons control gene expression in human neurons. *Science Advances* 6, eaba3200.

948 Walker, L. C. and Jucker, M. (2017). The Exceptional Vulnerability of Humans to Alzheimer's Disease.
949 *Trends in Molecular Medicine* 23, 534–545.

950 Wang, H., Xing, J., Grover, D., Hedges, D. J., Han, K., Walker, J. A. and Batzer, M. A. (2005). SVA Elements:
951 A Hominid-specific Retroposon Family. *Journal of Molecular Biology* 354, 994–1007.

952 Ward, M. C., Zhao, S., Luo, K., Pavlovic, B. J., Karimi, M. M., Stephens, M. and Gilad, Y. (2018). Silencing
953 of transposable elements may not be a major driver of regulatory evolution in primate iPSCs.
954 *eLife* 7, e33084.

955 Wei, Y. (2020). Comparative transcriptome analysis of the hippocampus from sleep-deprived and
956 Alzheimer's disease mice. *Genet Mol Biol* 43, e20190052.

957 Wray, G. A. (2007). The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* 8, 206–216.

958 Xie, M., Hong, C., Zhang, B., Lowdon, R. F., Xing, X., Li, D., Zhou, X., Lee, H. J., Maire, C. L., Ligon, K. L., et
959 al. (2013). DNA hypomethylation within specific transposable element families associates with
960 tissue-specific enhancer landscape. *Nat Genet* 45, 836–841.

961 Yang, W., Liu, Y., Slovik, K. J., Wu, J. C., Duncan, S. A., Rader, D. J. and Morrisey, E. E. (2015). Generation
962 of iPSCs as a Pooled Culture Using Magnetic Activated Cell Sorting of Newly Reprogrammed
963 Cells. *PLoS One* 10, e0134995.

964 Yi, R., Chen, B., Zhao, J., Zhan, X., Zhang, L., Liu, X. and Dong, Q. (2014). Krüppel-like Factor 8 Ameliorates
965 Alzheimer's Disease by Activating β-Catenin. *J Mol Neurosci* 52, 231–241.

966 Yu, D. X., Di Giorgio, F. P., Yao, J., Marchetto, M. C., Brennand, K., Wright, R., Mei, A., Mchenry, L., Lisuk,
967 D., Grasmick, J. M., et al. (2014). Modeling Hippocampal Neurogenesis Using Human Pluripotent
968 Stem Cells. *Stem Cell Reports* 2, 295–310.

969 Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M.,
970 Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* 9,
971 R137.

972 Zhang, J., Li, H., Yabut, O., Fitzpatrick, H., D'Arcangelo, G. and Herrup, K. (2010). Cdk5 Suppresses the
973 Neuronal Cell Cycle by Disrupting the E2F1–DP1 Complex. *J. Neurosci.* 30, 5219–5228.

974 Zhang, H., Xue, C., Shah, R., Bermingham, K., Hinkle, C. C., Li, W., Rodrigues, A., Tabita-Martinez, J.,
975 Millar, J. S., Cuchel, M., et al. (2015). Functional Analysis and Transcriptomic Profiling of iPSC-
976 Derived Macrophages and Their Application in Modeling Mendelian Disease. *Circulation*
977 *Research* 117, 17–28.

978 Zhao, B., Wu, Q., Ye, A. Y., Guo, J., Zheng, X., Yang, X., Yan, L., Liu, Q.-R., Hyde, T. M., Wei, L., et al.
979 (2019). Somatic LINE-1 retrotransposition in cortical neurons and non-brain tissues of Rett
980 patients and healthy individuals. *PLOS Genetics* 15, e1008043.

981 Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D.,
982 McDermott, G. P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of
983 single cells. *Nat Commun* 8, 14049.

984
985
986 **Figure legends**
987
988 **Figure 1** – Differentiation of human and chimpanzee iPSCs into hippocampal intermediate progenitors (A)
989 single-cell RNA-seq data of human and chimpanzee cells from iPSC to day 5 of hpIPC differentiation,
990 grouped by time point and split by species. (B, C) PAX6 and OTX2 expression across the hpIPC
991 differentiation time points. Darker purple indicates greater expression. (D) Western Blot of human and
992 chimpanzee differentiated iPSCs against TBR2 from day 0 to 5 of differentiation. GAPDH is shown as the
993 loading control. (E) Quantification of cells in S, G1 and G2M phases of the cell cycle from single-cell RNA-
994 seq data. (F) Heatmaps of ChIP-seq against TBR2 and H3K27ac in day 5 hpIPCs from human and
995 chimpanzee. Red indicates signal.
996
997 **Figure 2** – Differential Gene Expression In Human and Chimpanzee hpIPCs. (A) Schematic of iPSC
998 differentiation followed by RNA-seq library generation and analysis. (B) Volcano plot depicting "Human
999 UP" genes (right) and "Chimp UP" genes (left) with a log2FoldChange threshold of 1.5 and -1.5,
1000 respectively, and a p-value threshold of 0.05. (C) Top upstream regulators of the differentially expressed
1001 genes predicted by Ingenuity Pathway Analysis, ranked by -log10(p-value). (D) Heatmaps depicting the
1002 expression of genes predicted to be under the control of the transcription factors CREB1, FOXA2 and TBR2
1003 or predicted to play a role in embryonic development. The rows of the heatmaps indicate the transcript
1004 names, columns indicate the species, sample number and replicate. "HS1 Rep 1" indicates Homo sapiens
1005 sample 1 replicate 1 and "PT2 Rep 2" indicates Pan troglodytes sample 2 replicate 2
1006
1007 **Figure 3** – Human Centric Chromatin Accessibility Analysis. (A) Schematic of iPSC differentiation followed
1008 by ATAC-seq library generation and analysis. (B) UCSC Genome Browser visualization of human ATAC-seq
1009 libraries (HS Rep 1, Rep 2, Rep 3) aligned to hg19 genome assembly. (C) Distribution of the differentially
1010 accessible (DA) peaks into enhancers or promoters, and into Human UP (greater accessibility in humans)
1011 or Chimp UP (greater accessibility in chimpanzees). (D) Differentially accessible chromatin regions ($p <$
1012 0.05, $n = 3006$) are more likely to overlap with transposable elements (TE) than non-differentially
1013 accessible regions ($p > 0.9$, $n = 3{,}006$). (E) The 1335 TE-derived DA regions are more likely to be near a
1014 differentially expressed gene, as compared to DA regions that do not overlap a TE. (F) Distribution of the
1015 five major TE classes in the human genome, compared with their distribution among the TE-derived DA
1016 chromatin regions, and among the TE-derived DA chromatin regions proximal to a differentially expressed
1017 gene
1018
1019 **Figure 4** – Chimpanzee Centric Chromatin Accessibility Analysis. (A) Schematic of iPSC differentiation
1020 followed by ATAC-seq library generation and analysis. (B) UCSC Genome Browser visualization of human
1021 ATAC-seq libraries (PT Rep 1, Rep 2, Rep 3) aligned to hg19 genome assembly. (C) Distribution of the
1022 differentially accessible (DA) peaks into enhancers or promoters, and into Human UP (greater accessibility
1023 in humans) or Chimp UP (greater accessibility in chimpanzees). (D) Differentially accessible chromatin
1024 regions ($p < 0.05$, $n = 3{,}806$) are more likely to overlap with transposable elements than other accessible
1025 regions ($p > 0.9$, $n = 3806$. (E) Distribution of DA regions that overlap a TE and proximity to genes
1026 differentially expressed between human and chimpanzee hpIPCs. (F) Distribution of the five major TE
1027 classes in the chimpanzee genome, compared with their distribution among the 1410 TE-derived DA

1028 chromatin regions (G) Breakdown of the 20 Miscellanous TEs represented in the 1,410 TE-derived DA
1029 chromatin regions. (H) Distribution of SVA family of TEs in the human genome as compared to their
1030 distribution in the TE-derived DA chromatin regions (n = 16).
1031
1032 **Figure 5** – LTRs Are Enriched Among Human and Chimpanzee Differentially Accessible Transposons. (A)
1033 Distribution of LTRs compared to non-LTRs in the human genome and in the 1,335 human TE-derived DA
1034 regions, with predicted binding motifs. (B) Differentially expressed genes close to the human-enriched DA
1035 LTRs. (C) Distribution of LTRs compared to non-LTRs in the chimpanzee genome and in the 1,410
1036 chimpanzee TE-derived DA regions, with predicted binding motifs. (D) Differentially expressed genes close
1037 to the chimpanzee-enriched DA LTRs.
1038
1039 **Figure 6** – Human-specific SVAs Bind Tbr2 and Influence Neurodevelopment (A) The SVAs enriched within
1040 the differentially accessible chimpanzee peaks are enriched for binding motifs of neurodevelopmental
1041 transcription factors KLF8, ZIC1, and ASCL1 as well as the AP-1 dimer. (B) Human SVAs are enriched for
1042 the binding motif of TBR2. (C) Genes proximal to human-specific SVAs are predicted to be regulated by
1043 neurodevelopmental TFs and function in important neuronal pathways. (D) 1,816 human SVAs exhibit
1044 ATAC-seq signal in hpIPCs (red = signal, white = noise) and of these, 739 exhibit TBR2 ChIP signal (black =
1045 signal, yellow = background). (E) Breakdown of 739 accessible, TBR2 bound SVAs into "human-specific"
1046 and "conserved in chimpanzees". (F) Differentially expressed genes close to the human-specific TBR2-
1047 bound SVAs.
1048
1049 **Figure 7** – CRISPR-interference on human SVAs. (A) Treating NCCITs with Retinoic Acid for 7 days leads to
1050 induction of TBR2 and PAX6, as shown in qRT-PCR. (B) Treating the RA-NCCIT stable CRISPR line for 72
1051 hours with doxycycline leads to dCas9 activation, as shown in the Cas9 immunoblot. (C) dCas9 activation
1052 results in deposition of H3K9me3 at most human SVAs as shown in the H3K9me3 ChIP-seq heatmap. (D)
1053 Volcano plot depicting genes differentially expressed upon doxycycline treatment. (E) Venn diagram
1054 showing overlap between genes differentially expressed between human and chimpanzee's hippocampal
1055 intermediate progenitors and between RA-NCCITs with and without doxycycline treatment (i.e. with and
1056 without SVA repression). (F) Pie chart illustrating the fraction of TBR2-controlled genes that are
1057 differentially expressed upon SVA repression in RA-NCCITs.
1058
1059 **Supplemental Figure 1 –** Single Cell RNA-seq Analysis of Human and Chimpanzee hpIPC Differentiation.
1060 (A) Number of cells assayed at each time point in each species (B) Overlap of all human and chimpanzee
1061 cells in a single plot. Chimpanze cells are depicted in red and human cells are depicted in blue. (C) Overlap
1062 of all human and chimpanzee cells across all six time points in a single plot. (D) Expression of PAX6 across
1063 all six time points, human cells are depicted on the right and chimpanzee cells are depicted on the left. (E)
1064 Expression of OTX2 across all six time points, human cells are depicted on the left and chimpanzee cells
1065 are depicted on the right (F) Overlap of human and chimpanzee cells at G1, G2M and S phases of the cell
1066 cycle.
1067
1068
1069
1070
1071

Figure 1

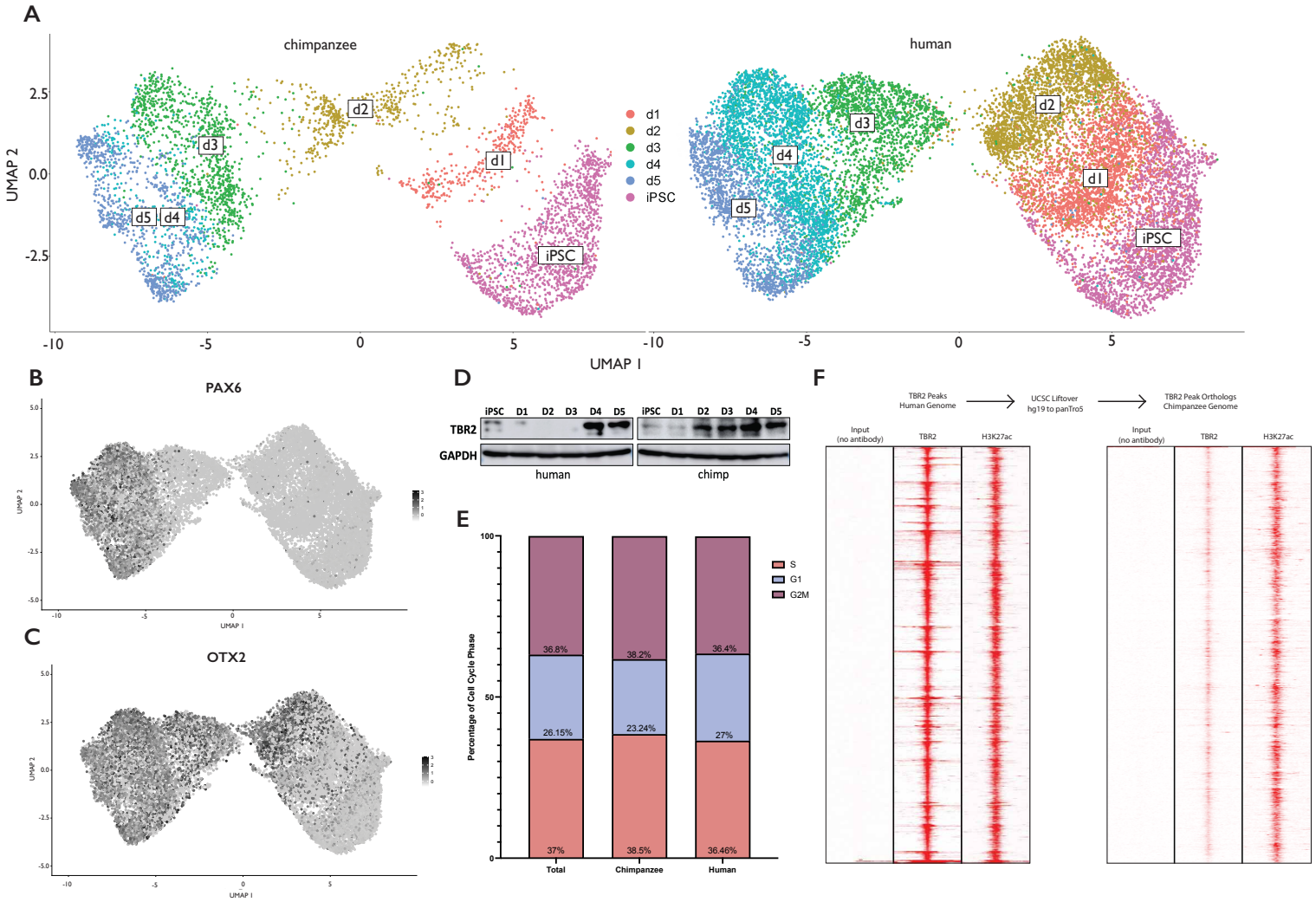

A
chimpanzee
human
d1
d2
d3
d4
d5
iPSC
UMAP 2
UMAP I
B
PAX6
C
OTX2
D
iPSC D1 D2 D3 D4 D5 iPSC D1 D2 D3 D4 D5
TBR2
GAPDH
human
chimp
E
S
G1
G2M
Percentage of Cell Cycle Phase
36.8%
26.15%
37%
38.2%
23.24%
38.5%
36.4%
27%
36.46%
Total
Chimpanzee
Human
F
TBR2 Peaks Human Genome
UCSC Liftover hg19 to panTro5
TBR2 Peak Orthologs Chimpanzee Genome
Input (no antibody)
TBR2
H3K27ac

Figure 2

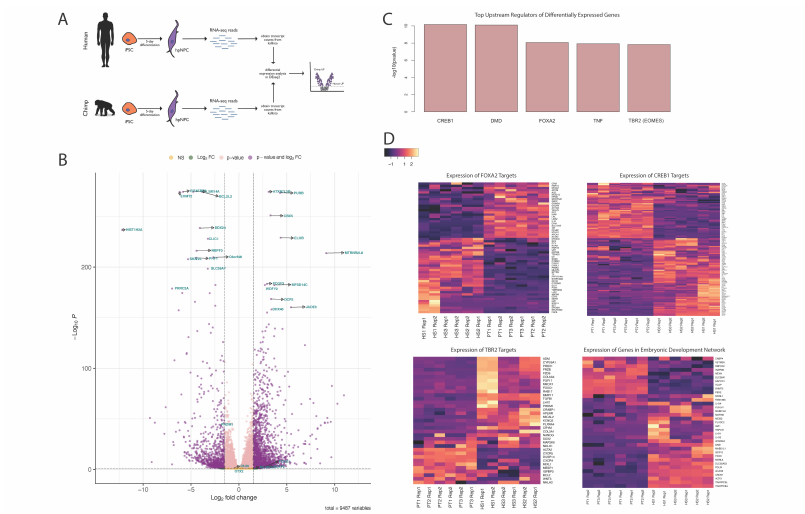

A
Human
Chimp
PSC
hpNPC
RNA-seq reads
B
Log2 fold change
C
Top Upstream Regulators of Differentially Expressed Genes
CREB1
DMD
FOXA2
TNF
TBR2 (EOMES)
D
Expression of FOXA2 Targets
Expression of CREB1 Targets
Expression of TBR2 Targets
Expression of Genes in Embryonic Development Network

Figure 3

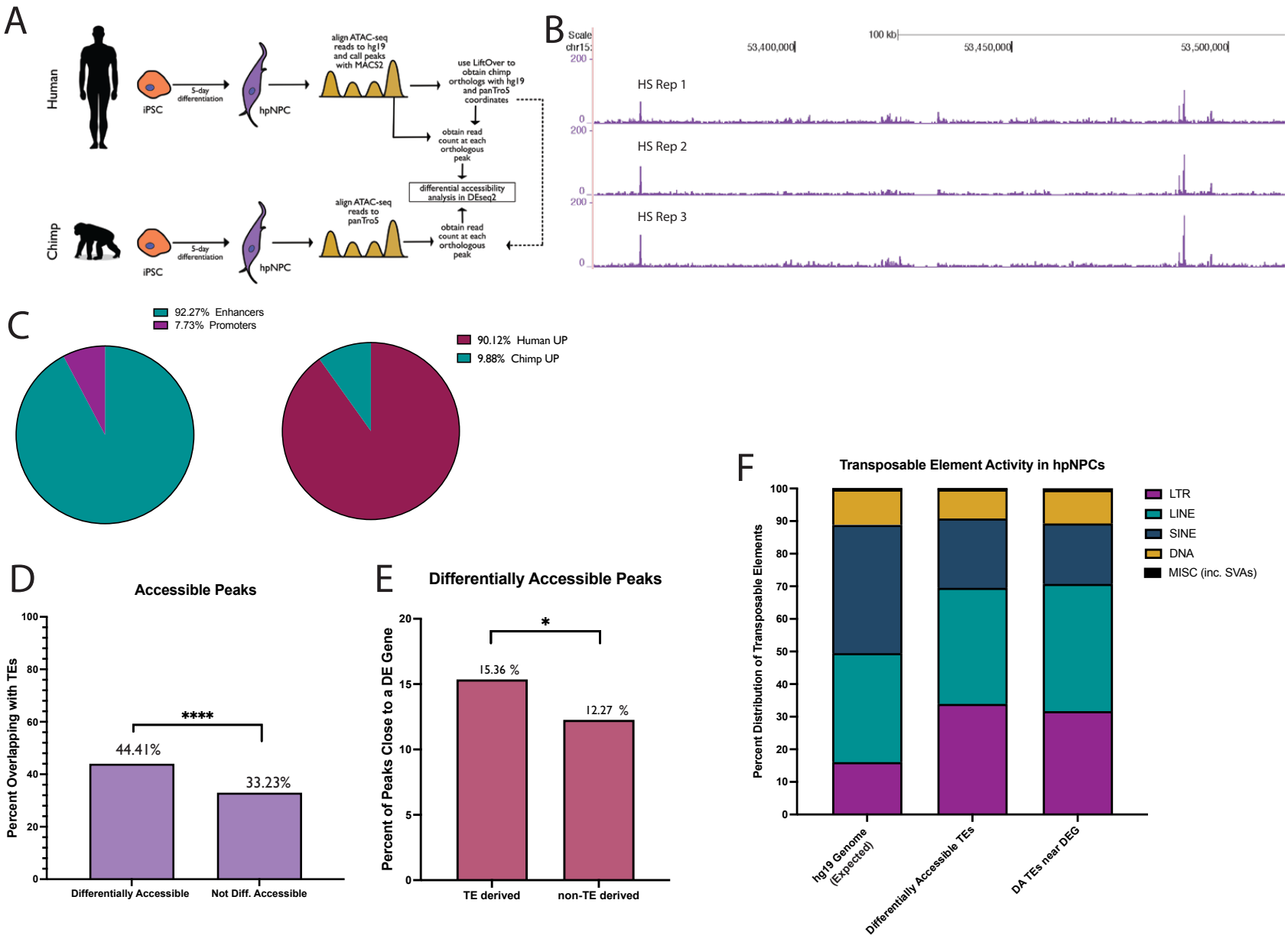

A
Human
iPSC
5-day differentiation
hpNPC
align ATAC-seq reads to hg19 and call peaks with MACS2
use LiftOver to obtain chimp orthologs with hg19 and panTro5 coordinates
obtain read count at each orthologous peak
differential accessibility analysis in DEseq2
Chimp
iPSC
5-day differentiation
hpNPC
align ATAC-seq reads to panTro5
obtain read count at each orthologous peak
B
Scale
chr15:
100 kb
53,400,000
53,450,000
53,500,000
HS Rep 1
HS Rep 2
HS Rep 3
C
92.27% Enhancers
7.73% Promoters
90.12% Human UP
9.88% Chimp UP
D
Accessible Peaks
Percent Overlapping with TEs
****
44.41%
33.23%
Differentially Accessible
Not Diff. Accessible
E
Differentially Accessible Peaks
Percent of Peaks Close to a DE Gene
*
15.36 %
12.27 %
TE derived
non-TE derived
F
Transposable Element Activity in hpNPCs
Percent Distribution of Transposable Elements
LTR
LINE
SINE
DNA
MISC (inc. SVAs)
hg19 Genome (Expected)
Differentially Accessible TEs
DA TEs near DEG

Figure 4

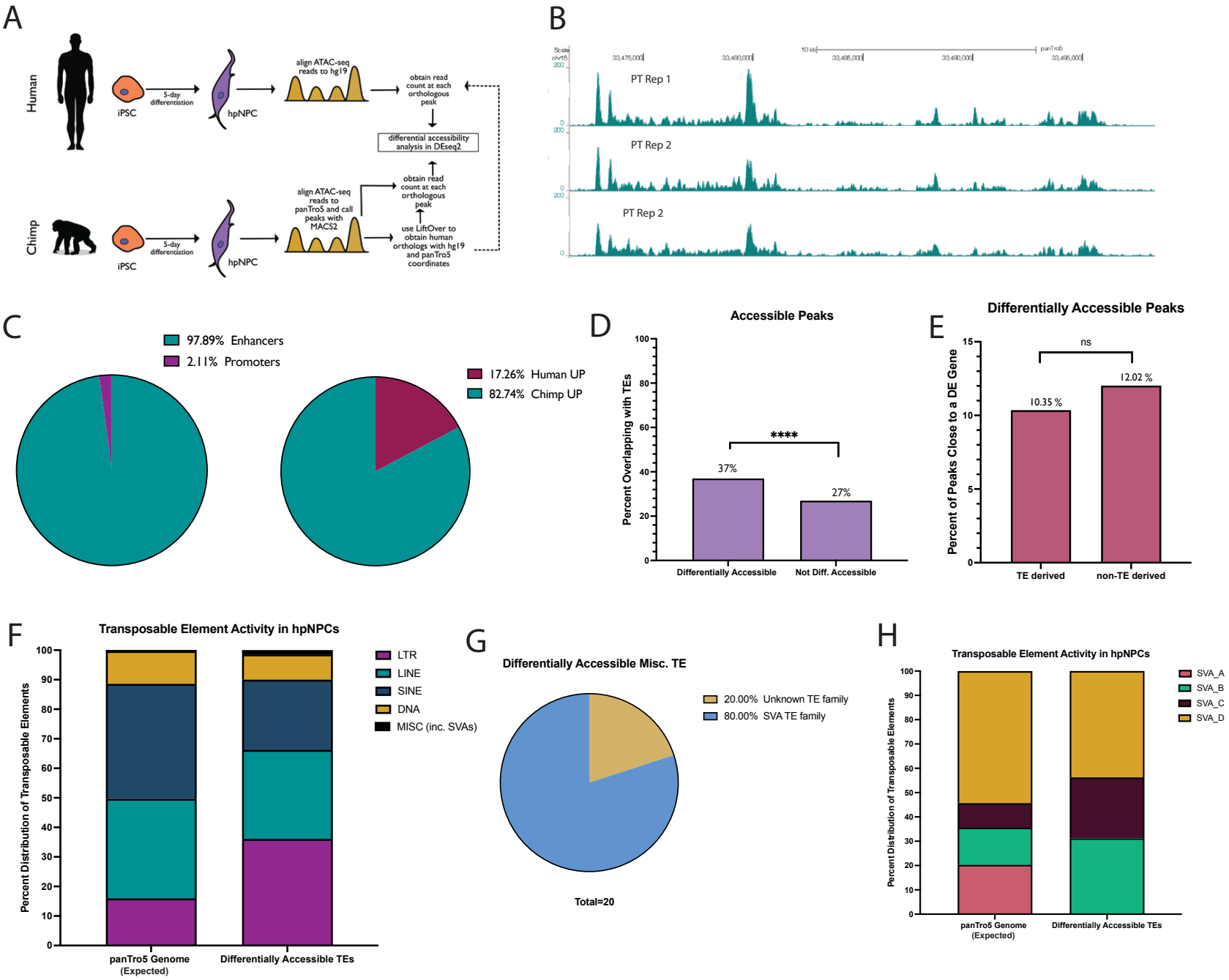

A
Human
iPSC
5-day differentiation
hpNPC
align ATAC-seq reads to hg19
obtain read count at each orthologous peak
differential accessibility analysis in DEseq2
Chimp
iPSC
5-day differentiation
hpNPC
align ATAC-seq reads to panTro5 and call peaks with MACS2
obtain read count at each orthologous peak
use LiftOver to obtain human orthologs with hg19 and panTro5 coordinates
B
PT Rep 1
PT Rep 2
PT Rep 2
C
97.89% Enhancers
2.11% Promoters
17.26% Human UP
82.74% Chimp UP
D
Accessible Peaks
Percent Overlapping with TEs
****
37%
27%
Differentially Accessible
Not Diff. Accessible
E
Differentially Accessible Peaks
Percent of Peaks Close to a DE Gene
ns
10.35 %
12.02 %
TE derived
non-TE derived
F
Transposable Element Activity in hpNPCs
Percent Distribution of Transposable Elements
LTR
LINE
SINE
DNA
MISC (inc. SVAs)
panTro5 Genome (Expected)
Differentially Accessible TEs
G
Differentially Accessible Misc. TE
20.00% Unknown TE family
80.00% SVA TE family
Total=20
H
Transposable Element Activity in hpNPCs
Percent Distribution of Transposable Elements
SVA_A
SVA_B
SVA_C
SVA_D
panTro5 Genome (Expected)
Differentially Accessible TEs

Figure 5

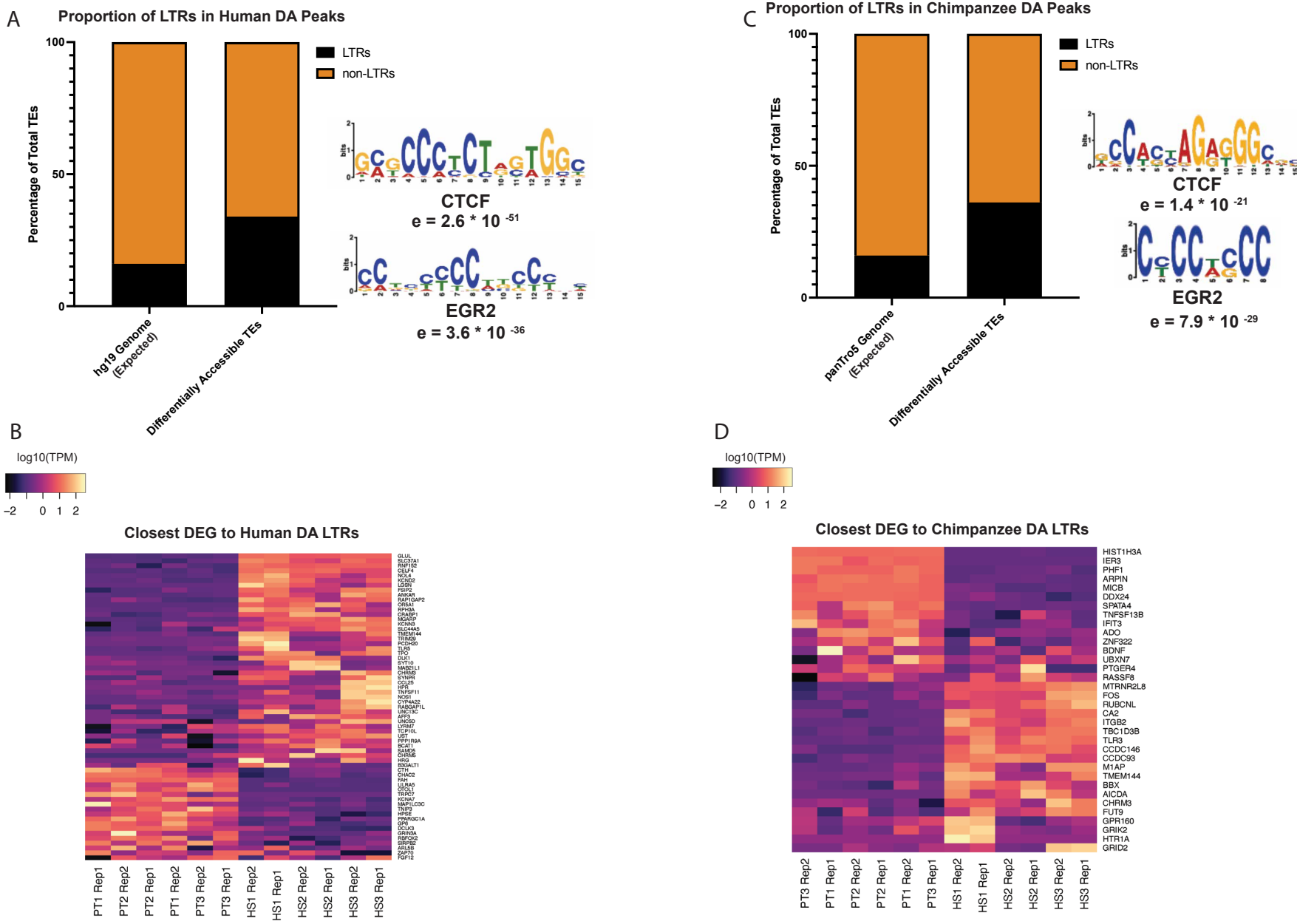

A
Proportion of LTRs in Human DA Peaks
Percentage of Total TEs
100
50
0
LTRs
non-LTRs
hg19 Genome (Expected)
Differentially Accessible TEs
CTCF
e = 2.6 * 10 -51
EGR2
e = 3.6 * 10 -36
C
Proportion of LTRs in Chimpanzee DA Peaks
Percentage of Total TEs
100
50
0
LTRs
non-LTRs
panTroS Genome (Expected)
Differentially Accessible TEs
CTCF
e = 1.4 * 10 -21
EGR2
e = 7.9 * 10 -29
B
log10(TPM)
−2 0 1 2
Closest DEG to Human DA LTRs
PT1 Rep1
PT2 Rep2
PT2 Rep1
PT1 Rep2
PT3 Rep2
PT3 Rep1
HS1 Rep2
HS1 Rep1
HS2 Rep2
HS2 Rep1
HS3 Rep2
HS3 Rep1
D
log10(TPM)
−2 0 1 2
Closest DEG to Chimpanzee DA LTRs
HIST1H3A
IER3
PHF1
ARPIN
MICB
DDX24
SPATA4
TNFSF13B
IFIT3
ADO
ZNF322
BDNF
UBXN7
PTGER4
RASSF8
MTRNR2L8
FOS
RUBCNL
CA2
ITGB2
TBC1D3B
TLR3
CCDC146
CCDC93
M1AP
TMEM144
BBX
AICDA
CHRM3
FUT9
GPR160
GRIK2
HTR1A
GRID2
PT3 Rep2
PT1 Rep1
PT2 Rep2
PT2 Rep1
PT1 Rep2
PT3 Rep1
HS1 Rep2
HS1 Rep1
HS2 Rep2
HS2 Rep1
HS3 Rep2
HS3 Rep1

Figure 6

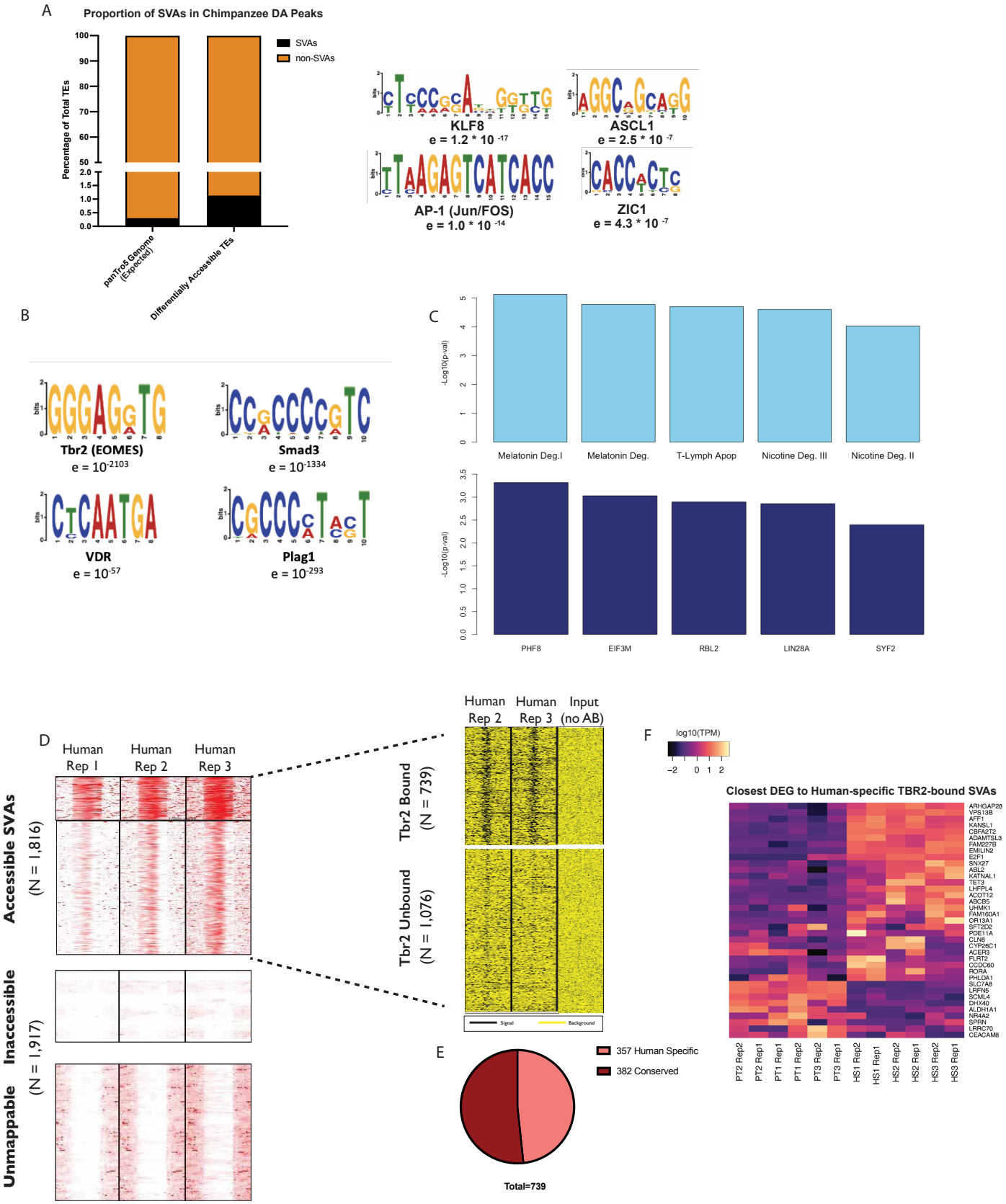
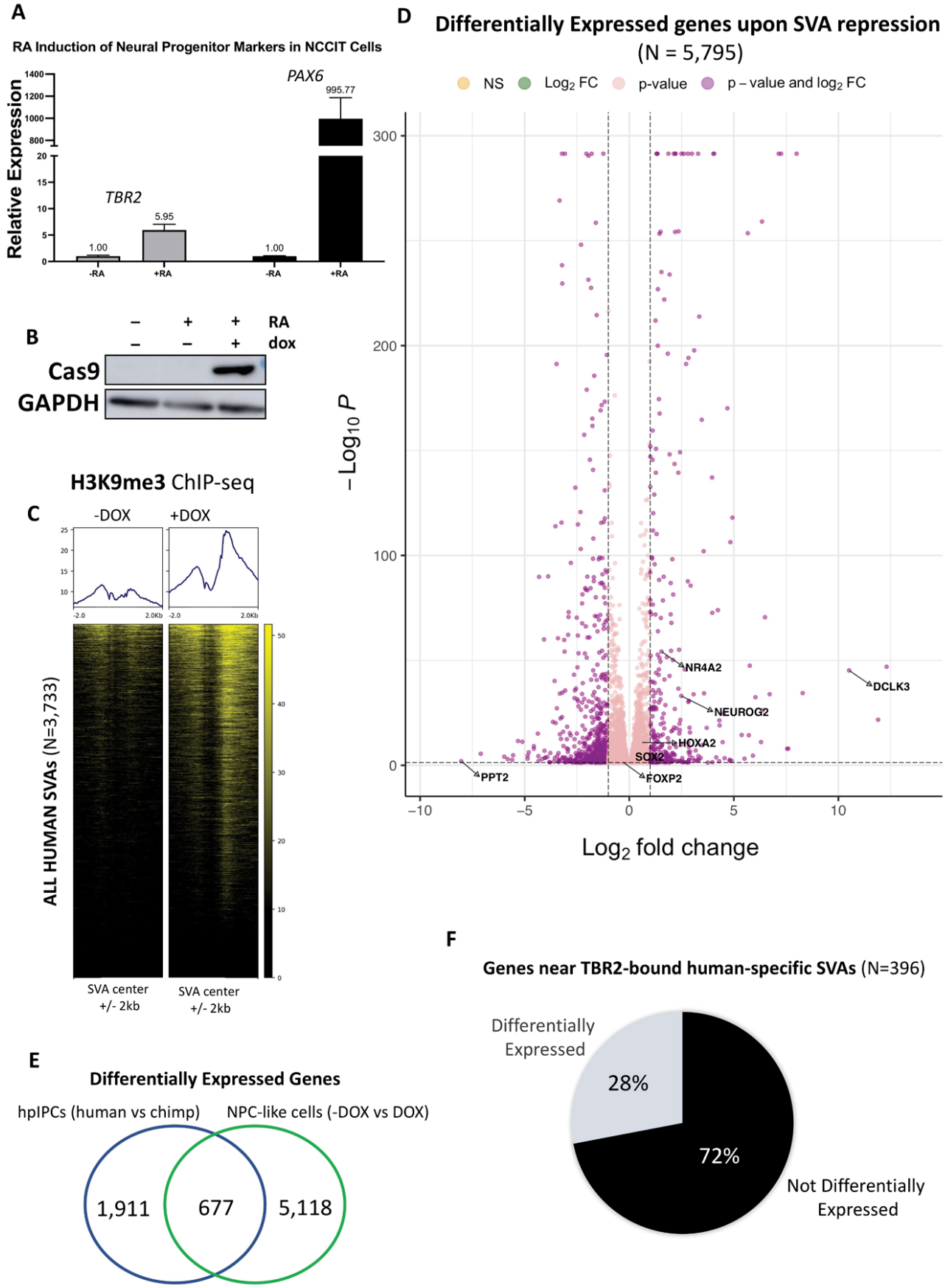
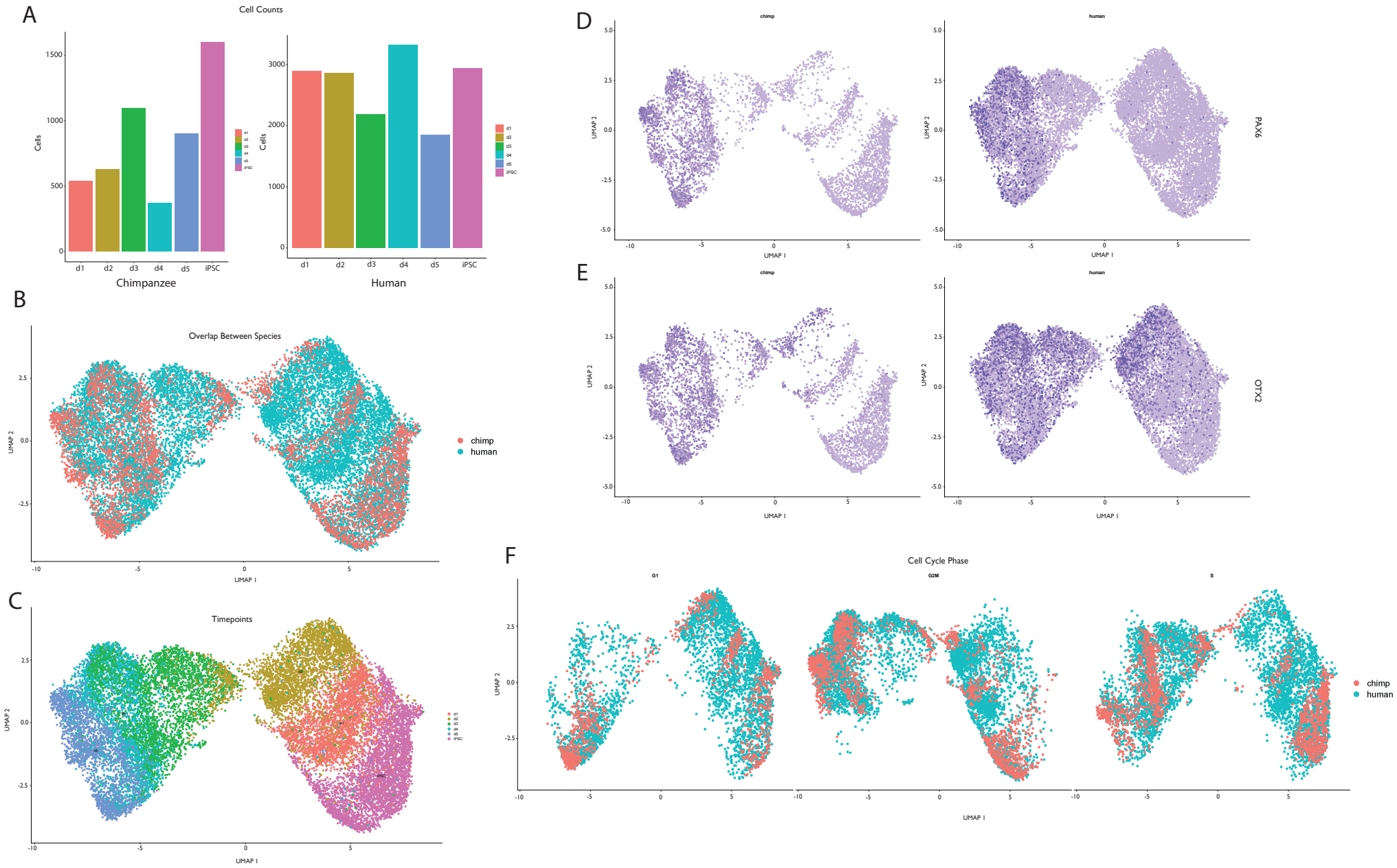

A
Proportion of SVAs in Chimpanzee DA Peaks
SVAs
non-SVAs
Percentage of Total TEs
panTro5 Genome (Expected)
Differentially Accessible TEs
KLF8
e = 1.2 * 10 -17
ASCL1
e = 2.5 * 10 -7
AP-1 (Jun/FOS)
e = 1.0 * 10 -14
ZIC1
e = 4.3 * 10 -7
B
Tbr2 (EOMES)
e = 10-2103
Smad3
e = 10-1334
VDR
e = 10-57
Plag1
e = 10-293
C
-Log10(p-val)
Melatonin Deg.I
Melatonin Deg.
T-Lymph Apop
Nicotine Deg. III
Nicotine Deg. II
PHF8
EIF3M
RBL2
LIN28A
SYF2
D
Human Rep 1
Human Rep 2
Human Rep 3
Accessible SVAs (N = 1,816)
Inaccessible (N = 1,917)
Unmappable
Tbr2 Bound (N = 739)
Tbr2 Unbound (N = 1,076)
Human Rep 2
Human Rep 3
Input (no AB)
Signal
Background
E
357 Human Specific
382 Conserved
Total=739
F
log10(TPM)
Closest DEG to Human-specific TBR2-bound SVAs
PT2 Rep2
PT2 Rep1
PT1 Rep1
PT1 Rep2
PT3 Rep2
PT3 Rep1
HS1 Rep2
HS1 Rep1
HS2 Rep2
HS2 Rep1
HS3 Rep2
HS3 Rep1

Figure 7


A
RA Induction of Neural Progenitor Markers in NCCIT Cells
Relative Expression
TBR2
PAX6
1.00
5.95
1.00
995.77
-RA
+RA
-RA
+RA
B
RA
dox
Cas9
GAPDH
H3K9me3 ChIP-seq
C
-DOX
+DOX
ALL HUMAN SVAs (N=3,733)
SVA center +/- 2kb
SVA center +/- 2kb
D
Differentially Expressed genes upon SVA repression
(N = 5,795)
NS
Log2 FC
p-value
p – value and log2 FC
– Log10 P
Log2 fold change
NR4A2
DCLK3
NEUROG2
HOXA2
SOX2
FOXP2
PPT2
E
Differentially Expressed Genes
hpIPCs (human vs chimp)
NPC-like cells (-DOX vs DOX)
1,911
677
5,118
F
Genes near TBR2-bound human-specific SVAs (N=396)
Differentially Expressed
28%
72%
Not Differentially Expressed

Supplemental Figure 1