Erythropoietin reshapes the adult hippocampal chromatin landscape to promote neurogenesis

Umut Cakir[1,2], Melanie Fritz[3], Umer Javed Butt[1,3], Riki Kawaguchi[4], Daniel Geschwind[4,5], Klaus-Armin Nave[6,7], Hannelore Ehrenreich[# 1,2,3] and Manvendra Singh[# 1,3]

[1]*Clinical Neuroscience, Max Planck Institute for Multidisciplinary Sciences, City Campus, Göttingen, Germany*

[2]*Georg-August-University, Göttingen, Germany*

[3]*Experimental Medicine, Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, J 5, Mannheim, Germany*

[4]*Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA, USA*

[5]*Institute of Precision Health, University of California Los Angeles, Los Angeles, CA, USA*

[6]*Department of Neurogenetics, Max Planck Institute for Multidisciplinary Sciences, City Campus, Göttingen, Germany*

[7]*Department of Neuropathology, Charite, Berlin, Germany*

Running Head: Erythropoietin and transposable elements in neurogenesis

Key Words: EPO, TE, pyramidal neurons, therapy, cognition, neuropsychiatric disorders

 # Correspondence:

Dr. Manvendra Singh, PhD
Experimental Medicine, Department of Psychiatry and Psychotherapy,
Central Institute of Mental Health,
Medical Faculty Mannheim, Heidelberg University, J 5,
68159 Mannheim, Germany
Phone: +49 (176) 670 23344
E-Mail: manvendra.singh@mpinat.mpg.de

Prof. Hannelore Ehrenreich, MD, DVM
Experimental Medicine, Department of Psychiatry and Psychotherapy
Central Institute of Mental Health
Medical Faculty Mannheim, Heidelberg University, J 5,
68159 Mannheim, Germany
Phone: +49-160-97000331
E-Mail: hannelore.ehrenreich@zi-mannheim.de

ABSTRACT

Understanding the molecular mechanisms by which erythropoietin (EPO) acts as neurotrophic factor that enhances hippocampal function and learning is essential to harness its therapeutic potential. Here, we employ single-nucleus ATAC-seq and RNA-seq to map the epigenomic and transcriptional landscapes of adult mouse hippocampus under recombinant human EPO (rhEPO) treatment. We discover significant lineage-specific chromatin remodeling predominantly in newly formed and immature excitatory neurons, highlighting a robust EPO-driven neurogenic response as the first direct evidence that an extrinsic factor can induce adult hippocampal neurogenesis. Notably, many EPO-induced accessible regions overlap ancient transposable elements, particularly ancient LINEs and SINEs, that are bound by key neurogenic transcription factors such as NEUROD1/2, NEUROG2, FOXG1, and ASCL1 and are linked to nearby genes governing neuronal differentiation and synaptic plasticity. Our findings uncover a previously unrecognized transposon-mediated mechanism underlying EPO-induced neurogenesis, highlighting an underappreciated role for TE-derived sequences in this process, and establish a publicly available single-nucleus multiomic atlas as a resource for understanding cell-type-specific gene regulation and neuroplasticity in the adult brain.

INTRODUCTION

Erythropoietin (EPO) is a hypoxia-inducible cytokine best known for its indispensable role in erythropoiesis[1]. In recent years, however, it has become evident that EPO is expressed by multiple cell types in the adult brain, with particularly high levels in hippocampus, where it has been implicated as a pleiotropic neurotrophic factor that can modulate synaptic plasticity, neuronal survival and circuit remodelling[2]. Indeed, there is growing therapeutic interest in EPO for neuropsychiatric and neurodegenerative conditions[3]. Recombinant human EPO (rhEPO) improves cognitive performance in several patient groups and confers neuroprotection in preclinical models of neurodegeneration and brain injury[4], supporting the existence of an endogenous "neuro-EPO" system in the adult central nervous system[2,5]. At the molecular level, several neuronal and glial lineages endogenously express EPO and its canonical receptor (EPOR)[6,7], suggesting that EPO signalling is embedded in local cellular networks rather than restricted to the classical hematopoietic axis.

The hippocampus is a heterogeneous structure composed of excitatory and inhibitory neurons, progenitor populations, and diverse glial and vascular cells. Their contributions to learning, memory and mood are controlled by cell-type-specific gene regulatory programs. Bulk genomics has revealed that EPO treatment modulates plasticity-related gene expression and improves hippocampus-dependent behaviour, but lacks the resolution to connect these changes to defined cell types and cis-regulatory elements. Recent single-cell transcriptomic work demonstrated that rhEPO "re-wires" hippocampal networks, promoting differentiation and integration of pyramidal neurons while dampening inhibitory interneuron activity[8,9]. These studies, together with electrophysiological evidence showing enhanced excitatory and reduced inhibitory input onto newly integrated CA1 neurons, argue that EPO can reshape hippocampal circuitry. At the same time, EPO induces plasticity-associated genes and chromatin regulators, including Bdnf, Gap43, Psd95, Egr1 and Ep300, and alters HDAC5 localisation, pointing towards a major epigenomic component[3,8,10]. However, how EPO signalling impinges on the chromatin landscape of specific hippocampal cell types, and how this relates to adult neurogenesis, has not been resolved.

Single-nucleus assay for transposase-accessible chromatin (snATAC-seq) enables genome-wide profiling of chromatin accessibility at single-cell resolution and thus the identification of

candidate cis-regulatory elements (cCREs): promoters and distal enhancers that underpin cell-type-specific transcriptional programs[11]. A key open question is how an extrinsic, clinically relevant modulator such as EPO reshapes chromatin accessibility in adult hippocampus *in vivo*, and which cell types and regulatory modules are most responsive. The snATAC-seq on hippocampi of mice treated with rhEPO allows us to capture dynamic regulatory events in rare hippocampal subpopulations and define the cCREs engaged by EPO, providing critical insights into how it reprograms neurogenic gene expression[8,12,13].

An additional layer of complexity comes from the pervasive presence of transposable elements (TEs) in mammalian genomes. Roughly half of the mouse and human genome consists of TEs, including LINEs, SINEs, endogenous retroviruses (ERVs) and DNA transposons[14]. Far from being inert, TEs are increasingly recognised as to have been co-opted into gene regulatory networks in development, immunity and the brain[14–17]. In fact, a substantial subset of brain-specific cCREs is derived from TEs. For instance, in the human cortex, around 80% of recently gained (human-specific) regulatory elements overlap TEs[18]. Notably, species-specific ERVs, SINEs, and even the DNA transposons have been shown to serve as cCREs in neural progenitors, providing novel binding sites for neurogenic TFs[19,20]. Whether adult hippocampal neurogenesis induced by an extrinsic stimulus such as EPO also engages TE-derived regulatory sequences has not been addressed yet.

Here, we combine snATAC-seq with matched single-nucleus RNA-seq (snRNA-seq) and external TF ChIP-seq datasets to map the epigenomic and transcriptional consequences of sustained rhEPO treatment in the adult mouse hippocampus. We first construct an integrated single-cell atlas of chromatin accessibility and gene expression across major hippocampal lineages under placebo and rhEPO conditions. We then define EPO-responsive differentially accessible regions (DARs) globally and in a lineage-resolved manner, and relate these to changes in gene expression, TF motif usage and GRNs, with particular focus on newly formed versus mature pyramidal neurons. Finally, we interrogate the contribution of TE-derived sequences to EPO-responsive cCREs and ask whether they are preferentially engaged by neurogenic TFs.

Our data show that rhEPO induces extensive, cell-type-specific chromatin remodelling in adult hippocampus, with the most pronounced accessibility gains in newly formed excitatory neurons. These chromatin changes are coupled to transcriptional activation of genes involved in neuronal

differentiation, synaptogenesis and synaptic plasticity, and are organised into regulons driven by canonical neurogenic TFs. A sizeable subset of EPO-responsive cCREs overlap TE insertions, particularly ancient LINE and SINE families, and these TE-derived sites are enriched for binding by neurogenic TFs in independent ChIP-seq datasets. Together, these findings place genome-wide chromatin accessibility changes rather than transcription alone at the centre of EPO's pro-neurogenic action in adult hippocampus and suggest that TE-derived sequences form part of the regulatory substrate that can be engaged by the EPO-induced pro-neurogenic chromatin-regulation mechanism. Beyond the mechanistic insights, the snATAC-seq atlas presented here provides a resource for exploring stimulus-responsive regulatory landscapes in the adult brain.

RESULTS

Single-cell chromatin accessibility profiling uncovers 35 cell clusters in adult mouse hippocampus

To discern how rhEPO impacts chromatin accessibility in adult hippocampus, we performed snATAC-seq on the right hippocampus of 8 adult C57BL/6N mice (4 rhEPO-treated, 4 placebo-treated) following an established dosing paradigm (11 injections every other day starting from P28 to P49)[21,22]. Libraries were generated using the 10x Genomics Chromium platform and processed with Cell Ranger ATAC. After stringent quality control (fragment number, TSS enrichment, nucleosome signal, blacklist fraction), ~35,000 high-quality nuclei were retained (Extended Data Fig. S1A-B, Table S1). Standard snATAC QC metrics were robust: we observed strong transcription-start-site (TSS) enrichment and characteristic nucleosomal banding, with <5% of fragments mapping to ENCODE blacklist regions (Extended Data Fig. S2A-B).

We constructed a unified peak set across samples using Signac framework[23], applied TF-IDF normalisation and latent semantic indexing (LSI), and integrated samples with Harmony to correct batch effects[24] (Fig. 1A). Dimensionality reduction and graph-based clustering resolved 35 distinct accessibility-defined clusters, visualised by UMAP (Fig. 1B). To assign cell identities, we integrated the snATAC-seq dataset with matched snRNA-seq profiles from the contralateral hippocampi of the same mice. Anchoring ATAC-derived gene activity scores to RNA expression via CCA yielded a joint embedding in which snATAC and snRNA nuclei overlapped well (Fig. 1C), and label transfer provided high-confidence cell-type annotations.

We distinguished 11 major hippocampal lineages: dentate gyrus granule neurons, excitatory (pyramidal) neurons, interneurons, microglia, astrocytes, oligodendrocytes, endothelial cells, pericytes, intermediate progenitors, ependymal cells and neuroimmune cells (Fig. 1B-C). Accessibility at canonical marker loci supported these assignments: for example, astrocyte clusters displayed strong accessibility at Slc1a2, oligodendrocyte-lineage clusters at Plp1/Mbp/Mog or Pdgfra, dentate gyrus neurons at Prox1, and excitatory neurons at Slc17a7 (Fig. 1G, Table S2–S4). These chromatin-defined identities closely matched the transcriptomic clusters in the companion snRNA-seq dataset (Fig. 1C-E and Extended Data Fig. S3, Fisher's Exact Test; p-value < 0.05), establishing a robust integrated atlas that we used as the basis for downstream analyses. Together, this integrated multi-omic atlas delineates the cell-type-specific regulatory landscape of the adult hippocampus, linking candidate cis-regulatory elements to

gene expression programs. In line with this, distribution of peaks across hippocampal lineages further highlights distinct cis-regulatory signatures for each cell type (Fig. 1F). These integrated multi-omic analyses provide a foundation for constructing gene regulatory networks underlying hippocampal cell identity and the effects of rhEPO treatment. For instance, we observed that the *Slc1a2* (*EAAT2*) promoter is broadly accessible across multiple hippocampal lineages, whereas its distal enhancers are selectively open in astrocytes, yielding astrocyte-specific expression (Fig. 1G). This fits the established principle that promoters are often ubiquitously accessible while enhancer deployment is cell-type specific, as shown by ENCODE-scale promoter cCRE analyses[25] and by astrocyte maturation studies documenting enhancer accessibility remodelling[26].

Characterization of hippocampal regulatory regions

To obtain an overview of regulatory diversity, we identified cell-type-specific DARs across the 11 lineages and performed motif enrichment analysis. We detected 135 TF motifs significantly enriched in accessibility peaks associated with specific clusters and their representative 11 lineages (log2 enrichment ratio > 2 and adjusted p-value < $1e^{-10}$), underscoring the breadth of regulatory programs operating in adult hippocampus (Fig. 2A and Extended Data Fig. S4, Table S2-3). As expected, well-known neurogenic and lineage-restricted TFs were prominent, including SOX TFs[27], members of the EGR family[28], the NEUROD1/2 and NEUROG1/2 factors, essential for neuronal differentiation[29,30] and NRF1, a metabolic regulator that activates neuronal and synaptic genes[31]. Notably, other motifs highlighted TFs less emphasised in hippocampal biology but linked to neurodevelopment, such as GLIS1/2, ZBTB14, RFX2/5, the NR4A2::RXRA heterodimer, ZNF423, a zinc-finger TF[32]. These motif enrichments showed lineage-specific patterns: OLIG-like and SOX-family motifs were predominant in glial clusters, consistent with oligodendrocyte/astrocyte programs[33], whereas EGR/ERG motifs dominated progenitor and early neuronal clusters except the interneurons, reflecting excitability-dependent genes and cell-cycle control in neural precursors[28,34]. In contrast, ETS/ELF motifs were highly enriched in endothelial and microglial clusters, aligning with their known association with immune regulatory programs[35] (Fig. 2A and Extended Data Fig. S4, Table S2-3). This integrated atlas therefore captures both the cellular composition and the underlying regulatory codes of the adult hippocampus under control and EPO conditions.

We next annotated snATAC-seq peaks by genomic context. Across all lineages, only ~10.4% of accessible regions overlapped annotated promoters (±1 kb from TSSs), whereas 44.5% fell

within introns, 36.9% in intergenic regions and 4.3% in exons or UTRs (Fig. 2B, Extended Data Fig. S5, and Table S4). This distribution is consistent with the notion that the majority of cCREs reside in intronic and intergenic space. Comparison with a recent single-cell atlas of adult mouse brain chromatin accessibility[36] revealed that ~10.8% of our hippocampal peaks were absent from prior annotations, indicating a set of unreported, potentially hippocampus-specific cCREs (Fig. 2B, Extended Data Fig. S5A, and Table S4). Given the prominence and the known contribution of TEs to regulatory landscapes, which constitute roughly half of the mammalian genome[17] and are increasingly recognized as a major source of regulatory innovation[37,38], we next intersected our snATAC-seq peaks with RepeatMasker annotations to quantify the extent of TE-derived regulatory activity. TEs frequently carry inducible enhancer or promoter activities, especially in the brain, where they have been shown to re-wire gene regulatory networks during development and evolution[19,37,39]. Approximately 20% of hippocampal cCREs overlapped annotated TEs, spanning LTR, LINE, SINE and other classes, consistent with previous reports that ~25% of mammalian cCREs are TE-derived[39] (Fig. 2C and Extended Data Fig. S5B). Notably, LTR elements exhibited disproportionately high accessibility, supporting the view that LTRs often function as self-contained regulatory modules for neighbor gene expression[37,39] (Fig. 2D and Extended Data Fig. S5C). Among these, subfamilies such as ERVK, ERVL, RMER1A, RMER1B, IAPEY4, LTR41 and ORR1E showed the most significant enrichment in cCRE repertoires (Fig. 2D-E, Extended Data Fig. S5C, and Table S5). These TE-derived cCREs may serve as lineage-specific regulatory elements influencing hippocampal gene programs[40,41]. Together, these findings reveal that TEs are extensively integrated into the adult hippocampal regulatory network, providing a pool of fresh cell type enriched cCREs that likely contribute to hippocampal function.

Epigenomic remodeling and enriched neuroplasticity regulators under rhEPO

To assess how rhEPO globally affects hippocampal chromatin accessibility, we compared aggregated peaks between EPO- and placebo-treated samples irrespective of cell identity. Logistic regression-based differential accessibility analysis identified 945 DARs (FDR < 0.01), of which 799 gained and 146 lost accessibility in rhEPO hippocampi (Fig. 3A). Upon mapping these DARs to annotated TSSs, we found that 639 of them were located within ±1 kb from TSS of known genes, out of which a substantial fraction of them overlapped with TE sequences (Fig. 3B). We then subjected these 639 genes to a gene ontology analysis, which indicated that these DAR-associated genes are significantly enriched in pathways for neuroplasticity, synapse organisation, and cognitive function, indicating a plausible mechanism of how EPO might re-

wire the neurogenesis-associated transcriptional networks (Fig. 3C and Table S6). For instance, we detected robust EPO-induced chromatin accessibility at the Ascl1 promoter, a keytranscriptional regulator that promotes neural progenitor differentiation [42,43]. Likewise, Egr2, an immediate-early gene of the EGR family is implicated in neuronal plasticity too[44] (Fig. 3A). This finding is further supported by TF binding motif enrichment analysis of these DARs, which revealed a convergence on TF families that govern neurogenesis and synaptic plasticity (Fig. 3D). In addition to enhancer marking TFs SP1/2, rhEPO-open chromatin regions were highly enriched for binding motifs of Kruppel-like factors (KLFs), which is consistent with KLFs' established role in neuronal maturation and associated with adult neurogenesis[45] as well as early growth response factors (EGR1–3), which are immediate-early TFs critical for learning and memory[44,46,47]. Notably, motifs for AP-2 (TFAP2) factors, CREB/ATF family members (e.g. Atf1) and GLIS factors were over-represented (Table S5). These families have established roles in neuronal maturation, learning and memory [48], and activity-dependent gene expression[49]. Thus, at a global level, EPO promotes chromatin opening at promoter-proximal and distal sites that harbour motifs for neurogenic and plasticity-associated TFs.

Importantly, upon overlaying the similar analysis of differentially expressed genes (DEGs) from our snRNA-seq data, we observed a strong correlation between DARs identified by snATAC-seq and DEGs from our parallel snRNA-seq analysis, particularly for DARs proximal to DEGs (Fig. 3E). This correlation indicates that chromatin remodeling mediated by rhEPO precedes and is functionally concordant with transcriptional changes in hippocampal lineages. Our parallel snRNA-seq dataset too shows higher expression of Egr2, Syt4, Nrg3, Nrp1 and many other DAR-associated genes in EPO-treated hippocampus than in controls (Fig. 3E and Table S7)[8,9]. Several synaptic genes exhibited enhanced accessibility only in EPO-treated nuclei, such as Synaptotagmin-4 (Syt4), an activity-inducible vesicle protein that modulates BDNF release and learning-related synaptic plasticity, and Neuregulin-3 (Nrg3), a trophic factor that promotes excitatory synapse formation on interneurons[50,51]. We also found differential peaks at the Neuropilin-1 (Nrp1) locus, a receptor for axonal guidance cues (semaphorins). Thus, rhEPO application not only activates these pro-neurogenic and plasticity-related genes at the chromatin level, but also overexpresses their transcription, suggesting a coordinated epigenome-transcriptome response underlying EPO-induced alterations in the neural network.

Intriguingly, nearly half of the regions gaining accessibility with EPO treatment overlapped TE sequences, suggesting that EPO's genome-wide epigenomic effects preferentially engage TE-

derived loci. These TEs were predominantly rodent-specific SINEs of the B1/Alu and B2 families (including subtypes such as *ID_B1*, *B3*, *B1_Mus1/2*) (Fig. 3B and 3F), in line with emerging evidence that SINE-derived DNA and transcripts can serve normal neuronal functions[52]. More generally, SINE retrotransposons across the mammalian clade can recruit transcriptional regulators or RNA-binding proteins to nearby genes, thereby modulating gene expression in cis, as shown in primate, bovine and murine species[52–55]. The enrichment of B1/B2 SINE elements within EPO-responsive chromatin regions, located in close vicinity to DEG-associated loci (Fig. 3F), implies that the retrotransposon sequences may be tapped to mediate the regulation of gene programs activated by EPO.

Lineage-Specific Chromatin Remodeling and TE-Driven GRN Activation in the Adult Mouse Hippocampus by EPO

To determine whether rhEPO exerts equal or lineage-specific effects on hippocampal chromatin states, we assessed differential chromatin accessibility across each of the 11 major hippocampal lineages provided in[8] from the same set of mice. Using the logistic regression (LR) test with adjusted $P < 0.01$ threshold, we identified a total of 39,404 genome-wide DARs in all 11 lineages taken together. Although a subset of these DARs (n = 7,608) was common to multiple lineages, the majority (n = 21,555) were lineage-specific (Fig. 4A and Table S8). While cataloguing the DARs in individual lineage, we found remarkably that pyramidal neurons alone accounted for approximately 50% (~22,000) of the total DARs, highlighting their pronounced chromatin responsiveness to rhEPO treatment (Fig. 4B and Table S8). To assess the differential promoter activity, we tested what fraction of DARs in each of 11 lineages falls within 1-kb range of the TSS of annotated genes. We found that promoter-proximal DARs followed the same lineage-specific distribution pattern observed for DARs overlapping TE sequences across the hippocampal lineages. Interestingly, the distribution pattern of DARs among TEs and TSS between the 11 lineages was consistently proportional. Pyramidal neurons exhibited by far the greatest number of DARs on TE sequences (n = 3,897) and on the promoters of genes (n=1849), with an inclined bias toward increased chromatin accessibility, as 1065 TSS were upregulated versus 784 downregulated regions (Fig. 4B and Extended Data Fig. S6A, Table S8).

Dentate gyrus showed the next highest DAR count (n = 6,131) but an opposite trend, as a majority of these regions had depleted accessibility for TEs (455 up vs. 726 down) and gene-TSSs (291 up vs. 305 down), whereas glial populations, oligodendrocytes (742 DARs),

astrocytes (581 DARs) and Microglia (308 DARs) displayed moderate numbers of EPO-responsive TSSs, skewed slightly toward openness. In contrast, interneurons (347 DARs) exhibited the reverse pattern, with predominantly downregulated regions (159 up versus 188 down). Endothelial, Pericytes, neuroimmune, intermediate and ependymal cells were least affected (Fig. 4B and Extended Data Fig. S6A, Table S8). Thus, the majority of DARs in pyramidal neurons, microglia, and oligodendrocytes were upregulated, whereas dentate gyrus cells and interneurons showed predominantly downregulated DARs. Overall, this distribution pattern suggests that EPO's chromatin remodelling effects are directed primarily toward excitatory lineage expansion and glial physiology, while concurrently limiting the accessibility changes associated with inhibitory neuronal programs.

We next tested a model, based on the above observations, in which EPO promotes open chromatin in excitatory neurons and glial cells by investigating if the key neurogenic gene promoter accessibilities are significantly upregulated in rhEPO samples. Expectedly, pyramidal neurons showed the activation of key neurogenic TFs and developmental regulators. For instance, Ascl1, Neurog2, and Pax6, the key factors driving neuronal commitment from precursors[56,57], and Nrg3, a neuregulin promoting excitatory synaptogenesis[51], all exhibited robust accessibility gains in pyramidal neurons, dentate gyrus, astrocytes, and oligodendrocytes (Fig. 4C and Extended Data Fig. S6B, Table S9). In addition to these established neurogenic genes, several less-characterised but pertinent regulators emerged among the top EPO-responsive promoters. For instance, the zinc-finger factors *Zbtb16* (Plzf) and *Zbtb18* (Rp58), both implicated in cortical neurogenesis and neuronal subtype development[58,59], showed increased accessibility, as did *Fabp5*, *Mdga2*, and *Dusp6* (Fig. 4C and Extended Data Fig. S6B, and Table S9). The enrichment of these diverse gene promoters suggests that EPO activates a broad pro-neurogenic chromatin landscape, encompassing classical neurogenesis drivers and novel regulatory elements linked to neuronal maturation and synaptic integration.

Following these leads, we tested whether rhEPO treatment might differentially re-wire GRNs in hippocampal pyramidal and glial lineages. To systematically test this hypothesis, we applied the Pando algorithm, integrating our matched snRNA-seq and snATAC-seq datasets, to identify key TFs, their target regulatory regions, and downstream effector genes[60]. This analysis revealed that while most regulons were conserved between rhEPO and placebo conditions, approximately a dozen regulons displayed specific enrichment under rhEPO (Fig. 4D and Extended Data Fig. S7, Table S10). Expectedly, these rhEPO-enriched regulons were

prominently driven by TFs involved in neuronal differentiation and maturation (Fig. 4D and Extended Data Fig. S7, Table S10). These observations support the model that rhEPO specifically engages neurogenic TFs to facilitate the elevation of gene expression essential for neuronal development and synaptic plasticity.

Additionally, we also found that a broad spectrum of TE families was significantly enriched within the DARs (Fig. 4E and Extended Data Fig. S8, Table S11). These enriched repeats included LTR retrotransposons of the ORR1E and MTD families, ancient LINEs L2/L2a and SINEs rodent B families B3, B4/B4A, the mammalian-wide interspersed repeat MIR/MIRb, which is an evolutionarily old SINE family[54,61–63]. This widespread TE activation is consistent with prior evidence that such elements can be co-opted as cCREs in neural cells[38,64,65]. Among many evidence of TE-derived cCREs, we observe increase accessibility of L2c, MER103C/MIR, and B1_Mur4 in rhEPO samples and seem plausible regulator of Phf24, Cpne5, and Haus6 genes, respectively (Fig. 4F). These patterns indicate that EPO treatment broadly unlocks repeat-rich genomic regions in a lineage-specific manner, suggesting that of TE families enriched in EPO-induced DARs act as cCREs that influence gene transcription in cis.

Lineage-specific chromatin remodeling predominantly affects pyramidal neurons

We had shown that in rhEPO-treated mice, excitatory neuron lineage, particularly CA1 pyramidal cells were markedly over-represented, while the proportion of GABAergic interneurons was reduced[8,9]. Notably, we observed a dramatic expansion of newly formed pyramidal neuron subpopulations, with up to a 5-fold increase (~500% of control levels) in rhEPO samples (Extended Data Fig. S9A-B). This dramatic expansion of immature excitatory neurons provides direct evidence that rhEPO treatment induces adult hippocampal neurogenesis, extending prior single-cell analyses showing ~2-fold enrichment (~200%) of immature pyramidal neurons and is consistent with earlier reports that prolonged EPO treatment increases mature CA1 pyramidal neuron numbers by ~20%[8,13]. Even a single EPO dose is known to acutely elevate immature glutamatergic precursors within hours[13]. To further delineate how EPO orchestrates transcriptional networks underlying neurogenesis, we performed high-resolution chromatin accessibility clustering of pyramidal neuron lineages from our snATAC-seq data. Initial clustering from snRNA-seq identified 20 distinct pyramidal clusters[8], whereas snATAC-seq revealed 29 distinct clusters (Extended Data Fig. S10A). Cross-modal integration allowed us to robustly annotate these clusters into 17 distinct cell types, showing substantial concordance between snRNA-seq and snATAC-seq modalities (Fig. 5A and Extended Data Fig. S10B). The two predominant pyramidal subgroups were newly formed (early-stage markers)

and mature neurons (late-stage markers) that exhibited complete congruence between modalities, clearly segregating into discrete UMAP clusters. Following these leads, we next sought to characterize chromatin accessibility differences between newly formed and mature pyramidal neurons. Differential accessibility analysis identified 7,420 DARs distinguishing these two populations, of which 1,705 DARs mapped within gene promoters and 935 overlapped TE sequences (Fig. 5B and Table S12). Remarkably, ~66 % of promoter-associated DARs displayed significantly higher chromatin accessibility in newly formed neurons, with gene ontology analyses revealing significant enrichment for neurogenesis and synapse formation (Fig. 5C). Of note, we observed that around 83% of DARs overlapping both promoter and TE-associated regions were upregulated in rhEPO samples.

We then investigated chromatin accessibility differences between rhEPO and placebo conditions within immature and mature pyramidal neurons. As expected, the mature pyramidal neurons exhibited limited chromatin remodeling (~7,000 DARs) compared to newly formed neurons (~11,000 DARs) with ~65% upregulated and ~35% downregulated in both cases (Fig. 5E and Extended Data Fig. S11A). These results suggest the diminished epigenetic plasticity potential of rhEPO upon neuronal maturation. Within newly formed neurons, alongside a plenty of promoter-proximal DARs (n=946, ±1 kb from TSS), we identified an additional ~4,000 distal DARs (within ±10 kb of gene TSS), marking putative enhancers and distal cCREs responsive to rhEPO (Fig. 5F, Extended Data Fig. S11B and Table S13). These distal regions prominently were located adjacent to immediate-early genes (Egr1, Egr2, Egr3), known mediators of neuronal activation and synaptic plasticity, as well as key neuronal specification genes (Ascl1, Bcl11a, Nrg3), and epigenetic regulators (Dnmt3a, Setd5) (Fig. 5G). Collectively, these data suggest that rhEPO preferentially promotes broad epigenomic remodeling and neurogenic transcriptional programs in newly formed neurons, while inducing limited, targeted adjustments in mature neurons.

TE-Linked Neurogenic Enhancer Activity on EPO-Induced Chromatin Remodeling

We subsequently examined whether the DARs induced by rhEPO overlapping TEs participate specifically in GRNs activated in newly formed neurons. In newly formed (immature) pyramidal neurons, we identified a substantially greater number of EPO-induced TE-derived differentially accessible regions (TE-DARs) compared to mature pyramidal neurons, at both promoter-proximal and distal genomic locations. Interestingly, mature neurons showed merely 1,253, whereas newly formed neurons featured 1,934 TE-DARs, with 844 located within a range of ±10

kb near gene promoters (Fig. 6A, Extended Data Fig. S11C-D and Table S13). To further investigate the regulatory context of these sites, we intersected the TE-DARs with ChIP-seq profiles for 12 neurogenic transcription factors (TFs) (FOXG1, PAX6, ASCL1, NEUROG1/2, NEUROD1/2, SOX2, TBR1, EGR1, FOXO3, FOS, and JUN). Using established permutation tests[66], we found hundreds of TE-DARs bound by one or more of these TFs - far more than expected by chance (Fig. 6B, Fig. 7A-C, Extended Data Fig. S12 and Table S14). Among TFs, NEUROD1 and FOXG1 exhibited the most extensive binding across TE families (295 and 256 TE families, respectively,pval < 0.001). A core set of 24 TE families, including ancient LINEs (L2, L3), LTR retrotransposons (LTR33, MER130, MLT1 series), and SINE elements (AmnSINE1, MamSINE1), was commonly enriched across multiple neurogenic TFs (Fig. 6C, Extended Data Fig. S13). Notably, many of these TE families have previously been exapted as gene regulatory enhancers in the brain. For instance, MER130 elements harbor consensus NEUROD/NEUROG-binding motifs and have been shown to function as enhancers during mouse neocortical development[67]. Likewise**,** AmnSINE1, an ancient SINE family, has well-documented enhancer activity crucial for mammalian brain development[68]. Together, our findings indicate that TE insertions from specific families have been specifically co-opted as part of the cis-regulatory repertoire of neurogenic TFs in mammalian brain.

To explore the functional relevance of these TE-derived DARs bound by neurogenic TFs, we examined the genes located nearby. We found 42 TE-containing DARs in rhEPO versus placebo comparison (filtered by >10% cells accessible at least in one condition) within 10 kb of range from genes TSS (Fig. 6D and Table S15). Consistent with our thesis, gene ontology analysis showed that the nearest genes to TF-bound TE-DARs were highly enriched for biological processes related to neurogenesis and synaptic integration (Extended Data Fig. S13 and Table S16). The genes nearest to these TE-DARs have well-documented roles in neuronal differentiation and synapse development, such as Hes5, Kcnma1, Pacsin1, Ankef1, Rab4a, Nptxr, and Nedd4l (Fig. 6D-E). These results demonstrate that ancient TE sequences serve as regulatory elements specifically during the developmental window of new neuron formation, thus enhancing the expression of genes critical for neuronal integration and synaptic plasticity. Furthermore, chromatin accessibility at these TF-bound TE-DARs was markedly increased in rhEPO samples (FDR < 0.01), which was accompanied by concordant transcriptional upregulation of nearby genes, specifically in newly formed neurons and dentate gyrus (Fig. 6E). For instance, our selected aforementioned candidates were exclusively accessible in rhEPO newly formed neurons, potentially driving the expression of several genes, as these genes are

upregulated in our snRNA-seq data (Fig. 6E-F). Together, leveraging publicly available ChIP-seq datasets, we confirmed that canonical neurogenic TFs extensively bind the EPO-induced TE-derived enhancers *in vivo*. Crucially, many of these TF-occupied transposon insertions lie adjacent to genes involved in neuronal differentiation and synapse formation, which show concordant increases in accessibility and expression under EPO treatment. In summary, our findings underscore an innovative regulatory mechanism whereby rhEPO engages TE-derived enhancers targeted by neurogenic TFs, facilitating epigenomic reprogramming essential for hippocampal neurogenesis and synaptic plasticity.

DISCUSSION

In this study, we constructed a high-resolution atlas of chromatin accessibility in the adult mouse hippocampus and demonstrated that EPO, a traditionally hematopoietic hormone, can profoundly reprogram the adult neural epigenome to drive neurogenesis. This expands upon prior single-cell brain atlases[11,69] by focusing on the hippocampus and incorporating an experimental perturbation (EPO), thereby providing new insights into how adult neurogenic niches can be modulated. Crucially, we found that EPO's effects are highly cell-type-specific, predominantly impacting excitatory neuron lineages (pyramidal neurons), which showed the most pronounced chromatin remodeling and population expansion in response to the treatment.

EPO as a potent extrinsic modulator of adult hippocampal neurogenesis

These findings enrich our understanding of adult hippocampal neurogenesis, a process first recognized in rodents in the 1960s[70,71] and later confirmed in humans using classical and current state-of-the-art methods[72,73]. The adult hippocampus retains progenitor cells that can differentiate into new neurons mainly in response to intrinsic or extrinsic stimuli. While we lack the temporal data, our results provide the first direct demonstration that EPO serves as an extrinsic trigger of adult hippocampal neurogenesis, which sets a mechanistic layer to earlier observations that EPO improves cognitive function and enhances neuroplasticity[2,13,74]. *In vivo* studies in healthy mice and clinical trials in patients have shown that EPO can improve learning and memory independent of its role in erythropoiesis[2,75–77]. We recently reported that EPO treatment increases mature pyramidal neuron numbers by ~20% and newly formed neurons by ~200% without altering cell proliferation, suggesting EPO promotes differentiation or maturation of pre-existing neural precursors. In this study, we observed a substantially greater relative expansion (~5-fold) of the young pyramidal neuron population under EPO. This striking difference may either reflect the sensitivity of our chromatin-based approach in capturing nascent neurons, as well as a possibly more robust EPO regimen, or a region bias in some samples. These hypotheses remain to be tested. Importantly, our data support the model that EPO drives the differentiation of progenitors rather than inducing widespread cell division, as we detected increased accessibility at neurogenic transcription factor loci (e.g., *Ascl1*, *Neurog2*, *Nrg3* etc.) but did not observe evidence of aberrant cell-cycle activation in non-neuronal lineages. This aligns with the notion that EPO biases cell fate choice toward neurons in the existing precursor pool. The net effect is an enrichment of pyramidal neurons at the cost of

interneurons. Pyramidal neurons integrate in the hippocampal circuitry, leading to potential implications for mood, memory, cognitive enhancement and recovery from brain injury.

TEs as a regulatory innovation in hippocampal gene regulation

Despite TE's known roles in other brain regions, the contribution of TEs to regulatory landscapes in the adult hippocampus has remained unexplored at single-cell resolution. One of the major highlights from our study is the extent to which TEs contribute to the gene regulatory changes underlying the EPO-induced phenomenon in hippocampi. We found that roughly one-fifth of all open chromatin regions in the adult hippocampus are derived from TEs, even under basal conditions. This observation is in line with large-scale epigenomic surveys showing that a substantial fraction of brain-specific cis-regulatory elements are TE-derived[36]. Notably, the hippocampus, particularly the dentate gyrus, a hub of adult neurogenesis, has been reported to exhibit high levels of TE activity and mosaicism[73]. TEs are thought to create genomic variability in neural progenitors and even generate long non-coding RNAs in the brain that fine-tune gene expression[64,78]. Our results extend these findings by pinpointing specific TE families that function as accessible chromatin sites in distinct hippocampal cell types. For example, we observed that certain LINE, SINE, and endogenous retrovirus (ERV) families, such as RMER1 retrotransposons and IAPEY4/ORR1 endogenous retroviral elements, are enriched in open chromatin unique to hippocampal neurons. These TE-derived sequences appear to serve as cell-type-specific enhancers or promoters, an interpretation supported by their overlap with known neuronal gene loci and motif enrichments.

TEs are increasingly recognized as drivers of gene regulatory network evolution. Many TEs carry binding site motifs for transcription factors and can thereby introduce novel regulatory modules. Classic examples include the co-option of a SINE element (*AmnSINE1*) as an enhancer essential for mammalian brain development[52,53] and the MER130 DNA transposon family, providing binding sites for neural transcription factors in the developing cortex[64,65]. Our data demonstrate that similar TE-mediated regulatory mechanisms are active in the adult hippocampus. We found motifs for neural fate determinants e.g., NeuroD**,** Sox, and Klf, embedded within several TE sequences that are accessible in neurogenic cell clusters. EPO's influence on the epigenome appears to leverage this TE repertoire. Remarkably, by aggregating all hippocampus cells, we observed that about 50% of the regions that gained DARs with EPO treatment were endogenous retroelement-derived sequences (Fig. 3), which is a proportion far exceeding their baseline representation. This enrichment suggests that EPO may activate a specific subclass of dormant regulatory elements to drive the nearby genes. One intriguing

possibility is that these TE-DARs carry latent binding motifs for inducible transcription factors and become functional enhancers only under certain stimuli, as in this case, EPO stimulation. Such a mechanism would represent a previously unappreciated mode of gene regulation: the *de novo* recruitment of ancient genomic elements to drive a contemporary cellular response.

Mechanisms of EPO-mediated neuroplasticity

Our integrative multi-omic analysis sheds light on the gene regulatory networks fine-tuned by EPO and how they intersect with TEs. By combining chromatin accessibility data with matched single-cell transcriptomes and leveraging a network inference algorithm, we identified key EPO-responsive regulons, which are sets of transcription factors and target genes that act in cascades. Prominently featured in these regulons were classical neurogenic transcription factors like NEUROD1**,** NEUROD2**,** FOXG1**,** PAX6**,** and JUN**,** known to orchestrate the progression of neural stem cells into mature neurons[79]. In EPO-treated hippocampi, these factors were not only upregulated at the mRNA level in pyramidal lineages, but their binding motifs were also significantly enriched in the EPO-gained accessible chromatin regions, and many of those were derived from TEs. This implies that EPO induces a feed-forward loop: it elevates neurogenic TF expression, and those TFs in turn bind to and activate TE-derived cCREs associated with neuron-specific genes, reinforcing the neurogenic program. Our incorporation of external ChIP-seq data from developing brain or neural culture systems provided orthogonal validation for this model, showing, for example, that NEUROD1/2 and NEUROG2 binding sites overlap many of the LINE/SINE/ERV-derived cCREs that EPO opened *in vivo*. While these findings resonate with reports that cell-type-specific TFs often exploit transposon-derived sequences for binding[80], here we demonstrate this phenomenon in an active, inducible context within the adult brain.

Taken together, we propose a conserved evolutionary mechanism at play: ancient TEs embedded in the genome have been repurposed as cCREs in neural cells, and during an EPO-induced neurogenic stimulus, the brain taps into this reservoir of regulatory elements to rapidly reshape gene expression. In evolutionary terms, this represents an efficient strategy – rather than evolving entirely new gene regulatory sequences from scratch, genome has preserved a large bank of TE-derived cCREs that lie dormant until the appropriate transcriptional signals such as those triggered by EPO/EPOR signaling, activate them. The outcome is a robust transcriptional activation of neurogenesis-promoting genes. This concept aligns with the hypothesis that TEs facilitate the acquisition of novel cis-regulatory elements over evolutionary time[37,81], and our data illustrate that such TE-facilitated regulatory innovation can be deployed

within a single organism's lifespan, enabling context-specific gene regulation to adapt to physiological stimuli.

Implications of EPO-mediated re-wiring of GRNs in regenerative therapy

Our study carries broad implications. First, it highlights a previously underappreciated molecular pathway by which extrinsic factors like rhEPO can induce neuroplastic changes: through large-scale epigenomic remodeling of chromatin accessibility. This complements classical signaling and gene expression analyses by demonstrating that changes in the "accessibility" of the genome are a critical intermediate step linking external signals to sustained cellular phenotypes. Of note, EPO is known to be upregulated in the brain under hypoxic or injury conditions, and it has been considered a natural neuroprotective agent[77,82]. Our findings raise the possibility that part of the brain's injury response may involve EPO-driven engagement of transposon-based cCREs to activate genes for repair, regeneration, or circuit reorganization. Second, our work highlights the functional importance of the vast repetitive genome in normal physiology. Whereas TEs have traditionally been studied in the context of genome instability or evolution, here we show they play a central role in acute, beneficial plasticity of the adult brain when challenged by external stimuli. This bridges the gap between transposon biology and neurobiology, suggesting that they are intimately involved in processes like learning, memory, and regeneration. Finally, from a translational perspective, identifying TE-derived cCREs that are activated by pro-neurogenic stimuli could open new therapeutic strategies. If certain transposon-derived cCREs are master switches for neurogenic gene programs, they might be targeted by gene editing or epigenetic drugs to stimulate neuron replacement in ageing or neurological disorders. Moreover, our single-cell multi-omic atlas of the EPO-treated and control hippocampus is publicly available as a resource for the community. It provides a catalogue of candidate regulatory elements for each major hippocampal cell type. In summary, our work not only provides the first mechanistic demonstration of EPO-induced adult hippocampal neurogenesis but also lays a foundation for future explorations into the "regulatory dark matter" of the genome in adult brain function and therapy.

CONCLUSIONS

While our present study on rhEPO-mediated epigenomic reorganization in the adult hippocampus is profound, it also confers obvious limitations. Firstly, our analysis captures a

snapshot of a single time point after rhEPO treatment, so we could not resolve the temporal dynamics of chromatin and transcriptional changes. While we identified a strong association between TFs-TEs interplay and chromatin accessibility as a proxy for regulatory activity maneuvered by EPO, the targeted perturbation e.g. CRISPR deletion of a candidate TE enhancer, and other epigenetic layers, such as histone modifications or 3D chromatin architecture, will be necessary to confirm the causality. Despite these limitations, our work reveals a compelling new paradigm in which an exogenous factor (rhEPO) leverages the genomic potential of transposable elements to execute a complex biological program such as neuronal development and plasticity. This insight expands our understanding of gene regulation in the adult brain and could inform novel strategies to enhance brain repair, cognitive function, and resilience against neurodegenerative conditions. By uncovering the latent regulatory power of transposons in neural plasticity, we pave the way for future studies to explore the therapeutic modulation of the “repeatome” in the service of regenerative medicine.

METHODS

Ethical approval

All mouse experiments were approved by the local Animal Care and Use Committee (Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit, LAVES – AZ 33.19-42502-04-17/2393) and conducted in strict accordance with the German Animal Protection Law. Every effort was made to minimize the number of mice used and their suffering.

Experimental model and treatments

Male C57BL/6N mice received 11 intraperitoneal injections of recombinant human erythropoietin (rhEPO, 5000 IU/kg body weight) or placebo (0.01 mL/g of solvent) every other day for 3 weeks, starting at postnatal day 28 (P28). In total, 23 male mice were used (rhEPO = 11, placebo = 12). Twenty-four hours after the final injection (P49), all animals were sacrificed by cervical dislocation. For each sample, 2 right hippocampi from mice of the same treatment group were pooled into a single tube (except for one rhEPO sample, which included only one hippocampus), resulting in 6 biological replicates per condition (N=6 rhEPO, N=6 placebo) for snRNA-seq (GSE220522). Likewise, the left hippocampi from 4 tubes of each group (N=4 rhEPO, N=4 placebo; tube labels are: A1, A2, A4, A6, B7, B8, B9, B11) were used for scATAC-seq. The snATAC-seq library construction is done for these 8 hippocampal samples using Chromium Single Cell ATAC Library & Gel Bead Kit v2 chemistry.

Single-cell ATAC-seq library processing and alignment

snATAC-seq libraries were generated and processed using the Cell Ranger ATAC pipeline (v2.1.0, 10x Genomics). Raw reads from 8 hippocampal samples were aligned to the mm10 reference genome (refdata-cellranger-arc-mm10-2020-A-2.0.0). Peak-barcode matrices and fragment files were generated using cellranger-atac count and used for downstream analysis.

Unified-peak set workflow

Downstream analysis was performed in R using the Signac (v1.13.0) and Seurat (v5.1.0) packages. Individual peak files were merged into a unified peak set using reduce() from the GenomicRanges package, excluding peaks <20 bp or >10,000 bp. Cells with fewer than 500 fragments in peaks were filtered out. Chromatin assays were constructed for each sample using FeatureMatrix().

Each Seurat object was annotated using EnsDb.Mmusculus.v79, and quality control (QC) metrics were computed: transcription start site (TSS) enrichment, nucleosome signal, percent reads in peaks, and blacklist ratio (based on blacklist_mm10). Cells were filtered using quantile-based thresholds, excluding outliers (below 2nd or above the 98th percentile for key metrics) to remove low-quality cells and potential multiplets.

Filtered objects were merged, then the merged object was processed with normalization with RunTFIDF(), dimensionality reduction via latent semantic indexing (RunSVD()), batch correction with harmony (RunHarmony() using LSI dims 2–30), and UMAP embedding. Clusters were identified with FindNeighbors() and FindClusters() with resolution 0.8 and algorithm 1. Gene activity scores were calculated via GeneActivity() and log-normalized with NormalizeData() (scale factor = median of nCount_Activity). A preprocessed single-nucleus RNA-seq dataset (Singh et al., 2023) served as a reference for cell type annotation. Unmatching samples (A3, A5, B10, B12) were excluded. Anchors were computed using canonical correlation analysis (CCA) to predict cells in scATAC-seq dataset.

Sample-specific peak set workflow

In addition to the unified peak set approach, we conducted a complementary analysis using sample-specific peak sets generated by Cell Ranger ATAC. For each sample, peak-barcode matrices and fragment files were independently loaded, and chromatin assays were constructed using a minimum cell threshold of min.cells =10. Seurat objects were generated separately for each sample and subsequently merged for integrated downstream analysis. Gene annotations were added using the EnsDb.Mmusculus.v79 database. QC metrics were calculated for each cell, including TSS enrichment, nucleosome signal, percent reads in peaks, and the blacklist ratio. To ensure high-quality samples, cells were filtered using the following strict criteria: nCount_peaks between 3,000 and 30,000, pct_reads_in_peaks >15%, blacklist_ratio < 0.05, nucleosome_signal <4, and TSS.enrichment >3.

Dimensionality reduction and clustering were performed: normalization with RunTFIDF(), feature selection via FindTopFeatures(min.cutoff ='q0'), latent semantic indexing (LSI) using RunSVD() on dimensions 2 through 8, and UMAP embedding via RunUMAP(). Clusters were identified using FindNeighbors() and FindClusters() with algorithm 1. Gene activity scores were computed using GeneActivity() and log-normalized with NormalizeData(). Cell type annotations previously inferred from the unified peak set analysis were transferred to the sample-specific peak set dataset by matching shared cell barcodes. Cells without annotation were excluded from further analysis.

For analysis related to the Pyramidal cluster, pyramidal neurons were subset from the unified peak set dataset and subjected to reclustering to refine subtype resolution. To facilitate annotation, these cells were integrated with a previously annotated single-nucleus RNA-seq reference dataset (Singh et al., 2023) using CCA, with 30 dimensions and 10 nearest neighbors employed to identify cross-modality anchors. Uncharacterized or ambiguous Seurat clusters were either removed or reannotated based on their transcriptomic profiles and proximity to well-defined clusters in the reference dataset. The finalized annotations were then transferred to the corresponding pyramidal neuron populations within the sample-specific peak set dataset by matching shared cell barcodes.

Differential Accessibility (DA) Analysis

All rhEPO- and placebo-derived cells were combined within each cluster (seurat cluster, higher-level cell-type annotation, or pyramidal subclusters). We then executed FindMarkers (test.use = "LR", min.pct = 0.01, latent.vars = "nCount_peaks") in a one-versus-rest fashion to identify peaks that are intrinsically enriched or depleted in a given cluster relative to all other clusters. We used 0.01 as min.pct threshold to capture accessibility shifts in smaller subfractions of a population. Peaks passing FDR < 0.01 were considered cell-type DA peaks. Furthermore, to pinpoint treatment-dependent changes, we stratified cells by condition inside each cluster and ran FindMarkers() with min.pct = 0.01 in the LR framework, and depth correction was applied. Peaks with FDR < 0.01 were classified as condition DA peaks. In differences between defined clusters (cell-type DA peaks) or treatment-dependent changes (condition DA peaks), summits (peak mid-points) were annotated via TxDb.Mmusculus.UCSC.mm10.knownGene to the closest gene's promoter, binned by genomic context (promoter, exon, intron, intergenic) and intersected with RepeatMasker (mm10) to identify and quantify overlap with TEs. Low-complexity, satellite, RNA, rRNA, tRNA, scRNA, snRNA, RC, srpRNA, and ambiguous repeats (entries flagged with "?" such as "LINE?") were excluded from all TE analyses unless explicitly stated.

Cluster-resolved motif enrichment

Chromatin dynamics were analyzed in two complementary tiers: chromatin accessibility differences between defined clusters and treatment-dependent changes. Vertebrate position-frequency matrices from the JASPAR-2020 CORE collection were added for motif analysis. For every Seurat cluster, higher-level cell-type annotation, or pyramidal subclusters, identified DA peaks using FindMarkers were filtered as min.pct threshold is 0.10, only positive, and FDR < 0.05. Clusters contributing ≥ 10 significant peaks were subjected to motif enrichment with

FindMotifs(). Likewise, within the same clustering framework, cells were split by condition, and DA peaks filtered with min.pct threshold is 0.10, only positive, and FDR < 0.05 parameters were used for motif analysis. For any cluster yielding ≥ 10 rhEPO-responsive peaks, FindMotifs() was run to identify transcription-factor motifs enriched in rhEPO-gained chromatin.

Transcription-factor ChIP-seq integration

Genome-wide binding profiles for 13 neurodevelopmental transcription factors (ASCL1, EGR1, FOS, FOXG1, FOXO3, JUN, NEUROD1, NEUROD2, NEUROG2, PAX6, SOX2, TBR1) were downloaded from GEO, (Table S17). Peak files were lifted over to the mm10 assembly where necessary (UCSC liftOver) and collapsed into non-redundant consensus regions with bedtools merge function. To ask whether individual TFs preferentially occupy repeat-derived regulatory DNA, we quantified intersections between each consensus ChIP peak set and RepeatMasker annotations using regioneR; 1000 permutations provided empirical P values and observed/expected overlap ratios. TE-associated peaks were assigned to the nearest transcription-start site with ChIPseeker (TxDb.Mmusculus.UCSC.mm10.knownGene) and subjected to GO-biological-process enrichment with clusterProfiler::enrichGO (BH-adjusted q ≤ 0.01). Next, TF peaks were overlapped with the 11 broad cell-class DA peak lists and the refined pyramidal-lineage DA peaks identified in our snATAC-seq analysis.

Data availability

All raw FASTQ files, fragment files, count matrices, and intermediate Seurat objects underlying the present study have been deposited in the GEO GSE and zenodo (available upon a reasonable request).

Code availability

Code for data processing, quality control, clustering, differential-accessibility testing, motif analysis, and TF-ChIP overlap is available on the following GitHub link: https://github.com/umutcakir/ATAC_EPO_vs_Placebo.

References

1. Byts N, Sirén AL. Erythropoietin: a multimodal neuroprotective agent. *Exp Transl Stroke Med*. 2009;1:4. doi:10.1186/2040-7378-1-4

2. Adamcio B, Sargin D, Stradomska A, et al. Erythropoietin enhances hippocampal long-term potentiation and memory. *BMC Biol*. 2008;6(1):37. doi:10.1186/1741-7007-6-37

3. Miskowiak KW, Vinberg M, Christensen EM, et al. Recombinant human erythropoietin for treating treatment-resistant depression: a double-blind, randomized, placebo-controlled phase 2 trial. *Neuropsychopharmacology*. 2014;39(6):1399-1408. doi:10.1038/npp.2013.335

4. Grimm G, Stockenhuber F, Schneeweiss B, Madl C, Zeitlhofer J, Schneider B. Improvement of brain function in hemodialysis patients treated with erythropoietin. *Kidney Int*. 1990;38(3):480-486. doi:10.1038/ki.1990.229

5. Sargin D, El-Kordi A, Agarwal A, et al. Expression of constitutively active erythropoietin receptor in pyramidal neurons of cortex and hippocampus boosts higher cognitive functions in mice. *BMC Biol*. 2011;9:27. doi:10.1186/1741-7007-9-27

6. Ott C, Martens H, Hassouna I, et al. Widespread Expression of Erythropoietin Receptor in Brain and Its Induction by Injury. *Mol Med*. 2015;21(1):803-815. doi:10.2119/molmed.2015.00192

7. Butt UJ, Çakır U, Wildenburg AF, et al. Forebrain-specific loss of erythropoietin provokes compensatory upregulation of different EPO receptors. *Mol Psychiatry*. Published online September 30, 2025. doi:10.1038/s41380-025-03230-7

8. Singh M, Zhao Y, Gastaldi VD, et al. Erythropoietin re-wires cognition-associated transcriptional networks. *Nat Commun*. 2023;14(1):4777. doi:10.1038/s41467-023-40332-8

9. Curto Y, Carceller H, Klimczak P, et al. Erythropoietin restrains the inhibitory potential of interneurons in the mouse hippocampus. *Mol Psychiatry*. Published online April 15, 2024. doi:10.1038/s41380-024-02528-2

10. Liu B, Wen JK, Li BH, et al. Celecoxib and acetylbritannilactone interact synergistically to suppress breast cancer cell growth via COX-2-dependent and -independent mechanisms. *Cell Death Dis*. 2011;2(7):e185. doi:10.1038/cddis.2011.64

11. Li YE, Preissl S, Hou X, et al. An atlas of gene regulatory elements in adult mouse cerebrum. *Nature*. 2021;598(7879):129-136. doi:10.1038/s41586-021-03604-1

12. Rey F, Balsari A, Giallongo T, et al. Erythropoietin as a Neuroprotective Molecule: An Overview of Its Therapeutic Potential in Neurodegenerative Diseases. *ASN Neuro*. 2019;11:1759091419871420. doi:10.1177/1759091419871420

13. Wakhloo D, Scharkowski F, Curto Y, et al. Functional hypoxia drives neuroplasticity and neurogenesis via brain erythropoietin. *Nat Commun*. 2020;11(1):1313. doi:10.1038/s41467-020-15041-1

14. Sundaram V, Wysocka J. Transposable elements as a potent source of diverse cis-regulatory sequences in mammalian genomes. *Philos Trans R Soc Lond B Biol Sci*. 2020;375(1795):20190347. doi:10.1098/rstb.2019.0347

15. Jönsson ME, Garza R, Johansson PA, Jakobsson J. Transposable Elements: A Common Feature of Neurodevelopmental and Neurodegenerative Disorders. *Trends Genet*. 2020;36(8):610-623. doi:10.1016/j.tig.2020.05.004

16. Monsen Ø, Grønvold L, Datsomor A, et al. The role of transposon activity in shaping cis-regulatory element evolution after whole-genome duplication. *Genome Res*. 2025;35(3):475-488. doi:10.1101/gr.278931.124

17. Wells JN, Feschotte C. A Field Guide to Eukaryotic Transposable Elements. *Annu Rev Genet*. 2020;54(1):539-561. doi:10.1146/annurev-genet-040620-022145

18. Zemke NR, Armand EJ, Wang W, et al. Conserved and divergent gene regulatory programs of the mammalian neocortex. *Nature*. 2023;624(7991):390-402. doi:10.1038/s41586-023-06819-6

19. Patoori S, Barnada SM, Large C, Murray JI, Trizzino M. Young transposable elements rewired gene regulatory networks in human and chimpanzee hippocampal intermediate progenitors. *Development*. 2022;149(19). doi:10.1242/dev.200413

20. Pulver C, Grun D, Duc J, et al. Statistical learning quantifies transposable element-mediated cis-regulation. *Genome Biol*. 2023;24(1):258. doi:10.1186/s13059-023-03085-7

21. Sinnamon JR, Torkenczy KA, Linhoff MW, et al. The accessible chromatin landscape of the murine hippocampus at single-cell resolution. *Genome Res*. 2019;29(5):857-869. doi:10.1101/gr.243725.118

22. Preissl S, Fang R, Huang H, et al. Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat Neurosci*. 2018;21(3):432-439. doi:10.1038/s41593-018-0079-3

23. Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. Single-cell chromatin state analysis with Signac. *Nat Methods*. 2021;18(11):1333-1341. doi:10.1038/s41592-021-01282-5

24. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*. 2019;16(12):1289-1296. doi:10.1038/s41592-019-0619-0

25. Fan K, Moore JE, Zhang XO, Weng Z. Genetic and epigenetic features of promoters with ubiquitous chromatin accessibility support ubiquitous transcription of cell-essential genes. *Nucleic Acids Res*. 2021;49(10):5705-5725. doi:10.1093/nar/gkab345

26. Lattke M, Goldstone R, Ellis JK, et al. Extensive transcriptional and chromatin changes underlie astrocyte maturation in vivo and in culture. *Nat Commun*. 2021;12(1):4335. doi:10.1038/s41467-021-24624-5

27. Favaro R, Valotta M, Ferri ALM, et al. Hippocampal development and neural stem cell maintenance require Sox2-dependent regulation of Shh. *Nat Neurosci*. 2009;12(10):1248-1256. doi:10.1038/nn.2397

28. O'Donovan KJ, Tourtellotte WG, Millbrandt J, Baraban JM. The EGR family of transcription-regulatory factors: progress at the interface of molecular and systems neuroscience. *Trends Neurosci*. 1999;22(4):167-173. doi:10.1016/s0166-2236(98)01343-5

29. Toda T, Parylak SL, Linker SB, Gage FH. The role of adult hippocampal neurogenesis in brain health and disease. *Mol Psychiatry*. 2019;24(1):67-87. doi:10.1038/s41380-018-0036-2

30. Noack F, Vangelisti S, Raffl G, et al. Multimodal profiling of the transcriptional regulatory landscape of the developing mouse cortex identifies Neurog2 as a key epigenome remodeler. *Nat Neurosci*. 2022;25(2):154-167. doi:10.1038/s41593-021-01002-4

31. Cheng A, Wan R, Yang JL, et al. Involvement of PGC-1α in the formation and maintenance of neuronal dendritic spines. *Nat Commun*. 2012;3:1250. doi:10.1038/ncomms2238

32. Deshpande O, Lara RZ, Zhang OR, Concepcion D, Hamilton BA. ZNF423 patient variants, truncations, and in-frame deletions in mice define an allele-dependent range of midline brain abnormalities. *PLoS Genet*. 2020;16(9):e1009017. doi:10.1371/journal.pgen.1009017

33. Zhou Q, Anderson DJ. The bHLH transcription factors OLIG2 and OLIG1 couple neuronal and glial subtype specification. *Cell*. 2002;109(1):61-73. doi:10.1016/s0092-8674(02)00677-3

34. Fong BC, Chakroun I, Iqbal MA, et al. The Rb/E2F axis is a key regulator of the molecular signatures instructing the quiescent and activated adult neural stem cell state. *Cell Rep*. 2022;41(5):111578. doi:10.1016/j.celrep.2022.111578

35. Jantaratnotai N, Ling A, Cheng J, Schwab C, McGeer PL, McLarnon JG. Upregulation and expression patterns of the angiogenic transcription factor ets-1 in Alzheimer's disease brain. *J Alzheimers Dis*. 2013;37(2):367-377. doi:10.3233/JAD-122191

36. Zu S, Li YE, Wang K, et al. Single-cell analysis of chromatin accessibility in the adult mouse brain. *Nature*. 2023;624(7991):378-389. doi:10.1038/s41586-023-06824-9

37. Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet*. 2017;18(2). doi:10.1038/nrg.2016.139

38. Fueyo R, Judd J, Feschotte C, Wysocka J. Roles of transposable elements in the regulation of mammalian transcription. *Nat Rev Mol Cell Biol*. 2022;23(7):481-497. doi:10.1038/s41580-022-00457-y

39. Du AY, Chobirko JD, Zhuo X, Feschotte C, Wang T. Regulatory transposable elements in the encyclopedia of DNA elements. *Nat Commun*. 2024;15(1):7594. doi:10.1038/s41467-024-51921-6

40. Jönsson ME, Garza R, Sharma Y, et al. Activation of endogenous retroviruses during brain development causes an inflammatory response. *EMBO J*. 2021;40(9):e106423. doi:10.15252/embj.2020106423

41. Sankowski R, Strohl JJ, Huerta TS, et al. Endogenous retroviruses are associated with hippocampus-based memory impairment. *Proc Natl Acad Sci U S A*. 2019;116(51):25982-25990. doi:10.1073/pnas.1822164116

42. Du H, Wang Z, Guo R, et al. Transcription factors Bcl11a and Bcl11b are required for the production and differentiation of cortical projection neurons. *Cereb Cortex*. 2022;32(17):3611-3632. doi:10.1093/cercor/bhab437

43. Kim EJ, Leung CT, Reed RR, Johnson JE. In vivo analysis of Ascl1 defined progenitors reveals distinct developmental dynamics during adult neurogenesis and gliogenesis. *J Neurosci*. 2007;27(47):12764-12774. doi:10.1523/JNEUROSCI.3178-07.2007

44. Beckmann AM, Wilce PA. Egr transcription factors in the nervous system. *Neurochem Int*. 1997;31(4):477-510. doi:10.1016/S0197-0186(96)00136-2

45. Moore DL, Apara A, Goldberg JL. Krüppel-like transcription factors in the nervous system: novel players in neurite outgrowth and axon regeneration. *Mol Cell Neurosci*. 2011;47(4):233-243. doi:10.1016/j.mcn.2011.05.005

46. Gallo FT, Katche C, Morici JF, Medina JH, Weisstaub N V. Immediate Early Genes, Memory and Psychiatric Disorders: Focus on c-Fos, Egr1 and Arc. *Front Behav Neurosci*. 2018;12. doi:10.3389/fnbeh.2018.00079

47. O'Donovan KJ, Tourtellotte WG, Millbrandt J, Baraban JM. The EGR family of transcription-regulatory factors: progress at the interface of molecular and systems neuroscience. *Trends Neurosci*. 1999;22(4):167-173. doi:10.1016/S0166-2236(98)01343-5

48. Loureiro-Campos E, Mateus-Pinheiro A, Patrício P, et al. Constitutive deficiency of the neurogenic hippocampal modulator AP2γ promotes anxiety-like behavior and cumulative memory deficits in mice from juvenile to adult periods. *Elife*. 2021;10. doi:10.7554/eLife.70685

49. Pittenger C, Huang YY, Paletzki RF, et al. Reversible inhibition of CREB/ATF transcription factors in region CA1 of the dorsal hippocampus disrupts hippocampus-dependent spatial memory. *Neuron*. 2002;34(3):447-462. doi:10.1016/s0896-6273(02)00684-0

50. Dean C, Liu H, Dunning FM, Chang PY, Jackson MB, Chapman ER. Synaptotagmin-IV modulates synaptic function and long-term potentiation by regulating BDNF release. *Nat Neurosci*. 2009;12(6):767-776. doi:10.1038/nn.2315

51. Müller T, Braud S, Jüttner R, et al. Neuregulin 3 promotes excitatory synapse formation on hippocampal interneurons. *EMBO J*. 2018;37(17). doi:10.15252/embj.201798858

52. Podbevšek P, Fasolo F, Bon C, et al. Structural determinants of the SINE B2 element embedded in the long non-coding RNA activator of translation AS Uchl1. *Sci Rep*. 2018;8(1):3189. doi:10.1038/s41598-017-14908-6

53. Linker SB, Randolph-Moore L, Kottilil K, et al. Identification of bona fide B2 SINE retrotransposon transcription through single-nucleus RNA-seq of the mouse hippocampus. *Genome Res*. 2020;30(11):1643-1654. doi:10.1101/gr.262196.120

54. Horton I, Kelly CJ, Dziulko A, Simpson DM, Chuong EB. Mouse B2 SINE elements function as IFN-inducible enhancers. *Elife*. 2023;12. doi:10.7554/eLife.82617

55. Kelly CJ, Chitko-McKown CG, Chuong EB. Ruminant-specific retrotransposons shape regulatory evolution of bovine immunity. *Genome Res*. 2022;32(8):1474-1486. doi:10.1101/gr.276241.121

56. Nacher J, Varea E, Blasco-Ibañez JM, et al. Expression of the transcription factor Pax 6 in the adult rat dentate gyrus. *J Neurosci Res*. 2005;81(6):753-761. doi:10.1002/jnr.20596

57. Andersen J, Urbán N, Achimastou A, et al. A transcriptional mechanism integrating inputs from extracellular signals to activate hippocampal stem cells. *Neuron*. 2014;83(5):1085-1097. doi:10.1016/j.neuron.2014.08.004

58. Lin HC, Ching YH, Huang CC, et al. Promyelocytic leukemia zinc finger is involved in the formation of deep layer cortical neurons. *J Biomed Sci*. 2019;26(1):30. doi:10.1186/s12929-019-0519-8

59. Lu C, Garipler G, Dai C, et al. Essential transcription factors for induced neuron differentiation. *Nat Commun*. 2023;14(1):8362. doi:10.1038/s41467-023-43602-7

60. Fleck JS, Jansen SMJ, Wollny D, et al. Inferring and perturbing cell fate regulomes in human brain organoids. *Nature*. 2023;621(7978):365-372. doi:10.1038/s41586-022-05279-8

61. Nishihara H, Smit AFA, Okada N. Functional noncoding sequences derived from SINEs in the mammalian genome. *Genome Res*. 2006;16(7):864-874. doi:10.1101/gr.5255506

62. McCarthy EM, McDonald JF. Long terminal repeat retrotransposons of Mus musculus. *Genome Biol*. 2004;5(3):R14. doi:10.1186/gb-2004-5-3-r14

63. Jurka J, Zietkiewicz E, Labuda D. Ubiquitous mammalian-wide interspersed repeats (MIRs) are molecular fossils from the mesozoic era. *Nucleic Acids Res*. 1995;23(1):170-175. doi:10.1093/nar/23.1.170

64. Notwell JH, Chung T, Heavner W, Bejerano G. A family of transposable elements co-opted into developmental enhancers in the mouse neocortex. *Nat Commun*. 2015;6:6644. doi:10.1038/ncomms7644

65. Wenger AM, Clarke SL, Notwell JH, et al. The enhancer landscape during early neocortical development reveals patterns of dense regulation and co-option. *PLoS Genet*. 2013;9(8):e1003728. doi:10.1371/journal.pgen.1003728

66. Singh M, Leddy SM, Iñiguez LP, Bendall ML, Nixon DF, Feschotte C. Transposable elements may enhance antiviral resistance in HIV-1 elite controllers. *Genome Biol*. 2025;26(1):28. doi:10.1186/s13059-025-03484-y

67. Sekine K, Onoguchi M, Hamada M. Transposons contribute to the acquisition of cell type-specific cis-elements in the brain. *Commun Biol*. 2023;6(1):631. doi:10.1038/s42003-023-04989-7

68. Hirakawa M, Nishihara H, Kanehisa M, Okada N. Characterization and evolutionary landscape of AmnSINE1 in Amniota genomes. *Gene*. 2009;441(1-2):100-110. doi:10.1016/j.gene.2008.12.009

69. Lein ES, Hawrylycz MJ, Ao N, et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature*. 2007;445(7124):168-176. doi:10.1038/nature05453

70. Altman J, Das GD. Postnatal neurogenesis in the guinea-pig. *Nature*. 1967;214(5093):1098-1101. doi:10.1038/2141098a0

71. Altman J, Das GD. Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *J Comp Neurol*. 1965;124(3):319-335. doi:10.1002/cne.901240303

72. Eriksson PS, Perfilieva E, Björk-Eriksson T, et al. Neurogenesis in the adult human hippocampus. *Nat Med*. 1998;4(11):1313-1317. doi:10.1038/3305

73. Dumitru I, Paterlini M, Zamboni M, et al. Identification of proliferating neural progenitors in the adult human hippocampus. *Science*. 2025;389(6755):58-63. doi:10.1126/science.adu9575

74. Mengozzi M, Cervellini I, Villa P, et al. Erythropoietin-induced changes in brain gene expression reveal induction of synaptic plasticity genes in experimental stroke. *Proc Natl Acad Sci U S A*. 2012;109(24):9617-9622. doi:10.1073/pnas.1200554109

75. Mennen S, Franta M, Begemann M, et al. Tolerability and first hints for potential efficacy of motor‐cognitive training under inspiratory hypoxia in health and neuropsychiatric disorders: A translational viewpoint. *Neuroprotection*. Published online June 4, 2024. doi:10.1002/nep3.47

76. Miskowiak K, O'Sullivan U, Harmer CJ. Erythropoietin Enhances Hippocampal Response during Memory Retrieval in Humans. *Journal of Neuroscience*. 2007;27(11):2788-2792. doi:10.1523/JNEUROSCI.5013-06.2007

77. Sakanaka M, Wen TC, Matsuda S, et al. *In vivo* evidence that erythropoietin protects neurons from ischemic damage. *Proceedings of the National Academy of Sciences*. 1998;95(8):4635-4640. doi:10.1073/pnas.95.8.4635

78. Coufal NG, Garcia-Perez JL, Peng GE, et al. L1 retrotransposition in human neural progenitor cells. *Nature*. 2009;460(7259):1127-1131. doi:10.1038/nature08248

79. Gonçalves JT, Schafer ST, Gage FH. Adult Neurogenesis in the Hippocampus: From Stem Cells to Behavior. *Cell*. 2016;167(4):897-914. doi:10.1016/j.cell.2016.10.021

80. Lin WT, Chao CM, Lin HL, Hung MC, Lai CC. Emergence of antibiotic-resistant bacteria in patients with Fournier gangrene. *Surg Infect (Larchmt)*. 2015;16(2):165-168. doi:10.1089/sur.2013.118

81. Feschotte C. Transposable elements and the evolution of regulatory networks. *Nat Rev Genet*. 2008;9(5). doi:10.1038/nrg2337

82. Wang L, Zhang Z, Wang Y, Zhang R, Chopp M. Treatment of Stroke With Erythropoietin Enhances Neurogenesis and Angiogenesis and Improves Neurological Function in Rats. *Stroke*. 2004;35(7):1732-1737. doi:10.1161/01.STR.0000132196.49028.a4

Acknowledgements

This work has been supported by the European Research Council (ERC) Advanced Grant to HE under the European Union's Horizon Europe research and innovation programme (acronym *BREPOCI;* grant agreement No 101054369), as well as by the Max Planck Society, the Max Planck Förderstiftung, the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) via DFG-Center for Nanoscale Microscopy & Molecular Physiology of the Brain (CNMPB). Research in the labs of HE and KAN is funded by SFB TRR 274/2 – 408885537 project C01 (to H.E. and K.A.N). UC received support from the IMPRS-Genome Science PhD program. Research in the labs of DG, RK and KAN is supported by the Adelson Medical Research Foundation.

Author Contributions

Concept, design, supervision: MS, HE

Funding acquisition: HE, RK, DG, KAN

Drafting manuscript and display items: UC, MS, MF, UJB, HE

Data acquisition/generation: MS, RK, DG, KAN

Data analyses & interpretation: UC, RK, MS, HE

All authors read and approved the final version of the manuscript.

Competing Interests:

The authors declare no competing financial or other interests in connection with this article.

FIGURE LEGENDS

Figure 1.

Integrated single-nucleus transcriptomic and chromatin-accessibility atlas of mouse hippocampus after rhEPO treatment

(A) C57BL/6 mice received recombinant human erythropoietin (rhEPO) or placebo every other day from postnatal day (P) 28 to P49. Hippocampal regions were dissected, and nuclei were isolated for single-nucleus RNA-seq (snRNA-seq) and single-nucleus ATAC-seq (snATAC-seq). The resulting datasets were jointly analysed to integrate gene expression and chromatin accessibility. Data processing included Cell Ranger peak calling, Signac-based quality control and preprocessing, Harmony batch correction, co-embedding snRNA-seq and snATAC-seq, cell-type annotation, differential accessibility testing, motif enrichment, gene regulatory network reconstruction, and identification of gene- and transposable element (TE)-associated candidate cis-regulatory elements (cCREs).

(B) UMAP projection of the snATAC-seq dataset colored by 35 unbiased Seurat clusters.

(C) Joint embedding of snATAC-seq and snRNA-seq using gene activity and expression. Left: UMAP projection of integrated nuclei, with snATAC-seq (red) and snRNA-seq (blue) showing strong overlap following CCA-based integration of ATAC-derived gene activity scores and RNA expression. Right: same embedding colored by major cell types assigned via label transfer and canonical markers, including pyramidal neurons (Pyr), dentate gyrus neurons (DG), interneurons (In), oligodendrocytes (Oli), astrocytes (Ast), microglia (Mic), endothelial cells (End), intermediate progenitors (Int), pericytes (Per), ependymal cells (Epd), and neuroimmune cells (NI).

(D) Violin plots of prediction scores for cell-type assignments obtained by label transfer from snRNA-seq to snATAC-seq. Each violin represents a major hippocampal cell type, while the final column in the lower panel shows the overall distribution across all types. The dashed line marks the median overall prediction score, illustrating high cross-modality annotation accuracy.

(E) UMAP projection of the snATAC-seq dataset, colored by cell-type identities transferred from the matched snRNA-seq reference shown in panel C.

(F) Heatmap of differentially accessible chromatin regions (DARs, rows) across hippocampal cell types (columns). Scaled accessibility values (red/yellow = high, blue = low) highlight distinct cis-regulatory signatures for each lineage. For visualization, the top 20 peaks per cell type are shown.

(G) Genome browser tracks showing chromatin accessibility at the Slc17a7, Mog, Plp1, and Slc1a2 loci across major hippocampal cell types. Highlighted regions mark cell-type-restricted accessible regions, illustrating lineage-specific regulatory landscapes.

Figure 2.

Cell-type TF motif usage and TE architecture of accessible chromatin

(A) Heatmap of significantly enriched transcription factor (TF) motifs (fold-enrichment >1.5; FDR $<1\times10^{-1}$) in cell-type–specific DARs. Rows correspond to 7 major hippocampal cell types and columns to representative motifs (color scale = $-\log_{10}$ adjusted p-value). The remaining 4 cell types are not shown due to limited motif detection resulting from their comparatively lower cell numbers.

(B) Genomic distribution of (i) all annotated TEs in the mouse genome, (ii) all snATAC-seq peaks identified in this study (225,618 summits), and (iii) the subset of peaks overlapping TE sequences (39,009 summits). Bars indicate the percentage of features located in distal intergenic, intronic, exonic, promoter (<1 kb, 1–2 kb, or 2–3 kb), downstream, or UTR regions. Right: Overlap of snATAC-seq peaks with CATlas, showing 201,210 overlapping peaks (black) and 24,408 non-overlapping peaks (grey).

(C) Venn diagram summarising the intersection between peaks that map to gene bodies (112,456), TEs (17,728), or both (21,281).

(D) Bar chart of the TE families at peak summits, colored by TE class (LTR, SINE, LINE, simple repeat, DNA, other, and unknown).

(E) The scatter plot shows $\log_2$(observed/expected) overlaps for each TE family versus $-\log_{10}$ p-value. The right plot compares observed to expected counts for the top significantly enriched families highlighted in the left plot.

Figure 3.

rhEPO remodels chromatin accessibility and engages neurogenic regulators.

(A) Volcano plot of DARs between rhEPO- and placebo-treated hippocampi, identifying 945 DARs (799 gained, 146 lost). Neurogenic regulators (Ascl1, Dnmt3a, Nrg3, Egr2) are highlighted.

(B) A total of 1,058 DARs (FDR < 0.01) were identified and grouped into 3 categories: promoter-proximal (within ±1 kb of a TSS), TE-associated (DARs overlapping TEs within ±1 kb of a TSS), and distal cis-regulatory elements (>500 bp from a TSS). The bar plot and accompanying UpSet

diagram summarize the counts per category, with schematic illustrations on the left indicating classification criteria.

(C) Gene ontology enrichment of DAR-associated genes, with significant terms related to synaptic plasticity, neurogenesis, and neurite morphology.

(D) Motif enrichment analysis of DARs, revealing over-representation of TFs including KLFs, EGR1–3, ASCL1, NRF1, SP-family factors, and TBR1.

(E) Scatter plot of differentially expressed genes from snATAC-seq inferred from gene activity and differentially expressed genes (DEGs, from snRNA-seq). Log_2 fold changes (rhEPO vs placebo) show strong concordance ($R = 0.68$, $p = 1.5 \times 10^{-3}$), with representative neurogenic genes (Nrg3, Syt4, Egr2, Nrp1) highlighted.

(F) Bar plot showing the number of unique genes with TSS located within ±1 kb of TE-derived DARs. Rodent-specific B1 SINE subfamilies (B1_Mus1, B1_Mur1, B1_Mur4, ID_B1) are the most frequent contributors, linking TE-derived accessibility to gene regulation.

Figure 4.

rhEPO selectively remodels chromatin accessibility, transposon usage, and TF programs across hippocampal cell types.

(A) Global view of DARs (FDR < 0.01) between rhEPO- and placebo-treated nuclei. Horizontal bars indicate the number of DARs identified in each lineage (color legend, left); the UpSet matrix to the right shows how many DARs are unique or shared across lineages.

(B) Stacked bars show, for each hippocampal lineage, the total number of DARs (light blue), the subset overlapping TEs (green), and the subset located within ±1 kb of gene TSSs (dark blue). Bars extending upward represent DARs with increased accessibility in rhEPO-treated samples, while downward bars indicate reduced accessibility compared with placebo.

(C) Heatmap of rhEPO-induced promoter accessibility changes at selected neurogenic genes across hippocampal lineages. Colors indicate log_2 fold change (EPO vs placebo), with red showing increased and blue decreased accessibility. Asterisks denote significant changes (FDR < 0.01).

(D) Stacked bars show the proportion of regulons inferred by Pando that are shared, rhEPO-specific, or placebo-specific within each lineage. The dot plot on the right summarizes the absolute number of differential regulons, with orange indicating rhEPO-specific and green placebo-specific regulons.

(E) TE-enrichment analysis for rhEPO-responsive DARs. The left scatter plot depicts log_2(observed/expected) overlaps versus $-log_{10}P$ for every annotated TE family (points colored

by TE class; dashed boxes mark the top significantly enriched families). The right scatter plot compares observed to expected counts for the top significant families.

(F) Genome browser snapshots of rhEPO-induced chromatin accessibility at TE-derived regulatory elements. In pyramidal neurons, increased accessibility was observed at an L2c element near Phf24 and at MIR and MER103C elements near Cpne5. In astrocytes, a B1_Mur4 element near Haus6 showed enhanced accessibility. Tracks display snATAC-seq signal for rhEPO (red) and placebo (blue), with genes (grey) and TEs (red) shown below.

Figure 5.

Pyramidal cell type-specific chromatin remodelling provoked by rhEPO

(A) Left: Joint UMAP embedding of pyramidal-lineage nuclei profiled by snATAC-seq (red) and snRNA-seq (blue). Right: UMAP of the same nuclei colored by clusters annotated through label transfer and marker gene activity, resolving newly formed neurons, mature pyramidal neurons, and other subtypes (e.g., mossy, NP SUB, Cajal-like cells).

(B) Volcano plot of DARs between newly formed and mature pyramidal neurons (FDR < 0.01). A total of 7,420 DARs were identified, including 4,315 with increased accessibility and 3,105 with reduced accessibility in newly formed neurons.

(C) Gene ontology enrichment analysis of promoter DARs upregulated in newly formed neurons. Significant terms (-log10p > 10) highlight processes linked to synapse organization, neurogenesis, learning, memory, and membrane potential and neuronal compartments (dendrite, synapse, distal axon, and synaptic membrane).

(D) Numbers of DARs in newly formed and mature neurons (rhEPO vs placebo), shown for all DARs (left), promoter-proximal gene-TSS DARs (middle), and TE-derived DARs (right). Newly formed neurons exhibit higher counts across all categories.

(E) Proportional distribution of DARs from panel D. In both newly formed and mature neurons, the majority of DARs are upregulated (pink), while a smaller fraction is downregulated (blue), indicating that rhEPO predominantly increases chromatin accessibility.

(F) Genomic distribution of EPO-responsive DARs in newly formed neurons relative to the nearest TSS. DARs are classified as promoter-proximal (≤1 kb, blue), distal promoter (1–10 kb, green), or intergenic/intronic (>10 kb, grey).

(G) Distance of significant DAR summits to the nearest TSS for selected neurogenic and synaptic genes in newly formed neurons. Each dot represents the nearest DAR to a gene TSS, colored by $\log_2$ fold change (red = increased, blue = decreased accessibility).

Figure 6.

TE-derived DARs and their association with neurogenic TFs and gene regulation

(A) Numbers of TE-derived DARs (rhEPO vs placebo) in newly formed and mature neurons, separated into proximal cis-regulatory elements (cCREs within 10 kb of a gene TSS, blue) and distal cCREs (>10 kb from a gene TSS, grey). Newly formed neurons harbor more TE-derived DARs than mature neurons.

(B) Bar chart ranking neurogenic TFs by the number of TE families significantly enriched within their ChIP-seq binding sites. Each bar represents the total count of enriched TE families associated with the indicated TF.

(C) Matrix of significantly enriched TE families across TF ChIP-seq datasets. Black squares denote significant enrichment ($p < 0.001$), white squares denote no significant enrichment ($p > 0.001$). TE families such as L2b, L2, MIRb, and MIR3 are strongly associated with multiple neurogenic TFs.

(D) Neighboring genes within ±10 kb of selected TE-derived DARs (rhEPO vs placebo) that were accessible in >10% of pyramidal nuclei. Each circle represents the closest DAR relative to the TSS, with position on the y-axis showing distance (kb) from the TSS, and color indicating $\log_2$ fold change (red = increased, blue = decreased accessibility under rhEPO).

(E) Concordance between TE-derived DARs bound by neurogenic TFs and their nearest genes. Left, genomic coordinates of DARs identified by snATAC-seq, with color scale indicating $\log_2$ fold change (rhEPO vs placebo). Right, heatmaps of matched gene expression from snRNA-seq across hippocampal cell types. Asterisks denote significant transcriptional changes (FDR < 0.01), showing coordinated chromatin accessibility and expression responses to rhEPO.

(F) Genome browser snapshots showing TE-derived DARs near neurogenic genes. Representative loci include Nptxr, Ckb, and Nfx1. Tracks display chromatin accessibility across major hippocampal lineages, with dashed boxes marking TE-overlapping elements that show lineage-restricted accessibility. Genes are shown in black, with TEs highlighted in red.

Figure 7.

Neurogenic transcription factors FOXG1, NEUROD1, and NEUROD2 exhibit selective binding to transposable-element–derived sequences

(A) FOXG1 ChIP-seq shows aggregate ChIP-seq signal profiles (top) and corresponding heatmaps (bottom) centered on representative repeat element families (MER58A, L3, MIR, MER130). “Positive” (Pos, blue) represents transposable elements overlapping FOXG1 ChIP-seq peaks, whereas “Negative” (Neg, green) indicates transposable elements without peaks. Each heatmap displays signal density within ±2 kb of the element center, revealing distinct FOXG1 enrichment over specific TE subfamilies.

(B) NEUROD1 ChIP-seq shows aggregate signal profiles and heatmaps demonstrating pronounced NEUROD1 binding across the same TE families (MER58A, L3, MIR, MER130).

(C) NEUROD2 ChIP-seq shows meta-profiles and heatmaps depicting NEUROD2 binding over the same TE families. Altogether, these analyses demonstrate that ancient transposable-element subfamilies function as conserved cis-regulatory modules co-opted into the neurogenic transcriptional program.

EXTENDED DATA

Extended Data Fig. S1. Quality control metrics derived from Cell Ranger ATAC output

(A) Summary of snATAC-seq quality metrics obtained from Cell Ranger ATAC output. For each sample, the table shows the number of high-quality nuclei passing filtering, mean raw read pairs per nucleus, fraction of high-quality fragments overlapping called peaks, fraction of high-quality fragments confidently assigned to nuclei, median high-quality fragments per nucleus, duplicate fragment fraction, and total number of peaks detected.

(B) Violin plots showing (left) the distribution of total peaks detected per nucleus and (right) the percentage of sequencing reads falling within peaks across nuclei for each sample.

Extended Data Fig. S2. Additional quality control profiles for snATAC-seq libraries

(A) Density distributions of three core snATAC-seq quality metrics across samples: TSS enrichment scores, nucleosome signal estimates, and ENCODE blacklist fractions. TSS enrichment reflects the characteristic accumulation of fragments at transcription start sites, while nucleosome signal is the ratio of mononucleosome fragments to nucleosome-free fragments per

cell. ENCODE blacklist fractions indicate the proportion of fragments mapping to problematic genomic regions.

(B) Aggregate ATAC-seq signal profile centered on transcription start sites (top), showing the expected sharp increase in accessibility at TSSs. The heatmap (bottom) displays ATAC-seq fragment density within ±3 kb of TSSs for all peaks, illustrating strong signal localization around promoter regions and consistent accessibility patterns across the dataset.

Extended Data Fig. S3. Cross-modality cell-type annotation and correspondence between snATAC-seq and snRNA-seq datasets.

(A) Stacked bar plot showing the percentage composition of annotated cell types within each Seurat snATAC-seq cluster. Each bar represents one chromatin-accessibility-defined Seurat cluster, and colors correspond to major hippocampal lineages: pyramidal neurons, dentate gyrus neurons, astrocytes, oligodendrocytes, interneurons, microglia, endothelial cells, pericytes, intermediate cells, neuroimmune cells, and ependymal cells. The distribution illustrates the dominant cell identity assigned to each Seurat snATAC-seq cluster.

(B) Sankey diagram comparing cell-type annotations derived from snATAC-seq markers (left) with those obtained from snRNA-seq markers (right). Flow widths represent the degree of correspondence between modalities, with thicker connections indicating stronger agreement between ATAC-based and RNA-based cell-type assignments.

Extended Data Fig. S4. Transcription factor motif enrichment across Seurat snATAC-seq clusters.

Heatmap showing enriched transcription factor motifs across Seurat-defined snATAC-seq clusters. Motifs were filtered for fold enrichment > 1.5 and adjusted p-value < 1×10^{-2}. Each column represents an individual Seurat cluster and is annotated by its corresponding cell-type identity (top color bar). Each row shows a transcription factor motif, with color intensity indicating the $-\log_{10}$ adjusted p-value of enrichment. The pattern of enrichment reveals distinct regulatory signatures across clusters, with related cell types sharing similar motif profiles and exhibiting coordinated accessibility at lineage-specific regulatory elements.

Extended Data Fig. S5. Genomic distribution of snATAC-seq peaks and enrichment of transposable element families.

(A) Pie chart summarizing the genomic annotation of all snATAC-seq peak summits, showing the proportion of peaks located in promoter regions (stratified by distance to the TSS), untranslated regions, exons, introns, downstream regions, and distal intergenic regions. Distal intergenic and intronic regions constitute the largest fractions of accessible chromatin.

(B) Bar plot showing the number of peak summits overlapping annotated genes, annotated transposable elements (TEs), both genes and TEs, or neither genomic feature. A substantial subset of peaks overlaps TEs, either exclusively or in combination with gene bodies.

(C) Z-scores for the enrichment of the top 50 TE families overlapping snATAC-seq peak summits. TE families are ranked by their enrichment, revealing pronounced overrepresentation of several LTR, SINE, and LINE subfamilies within accessible chromatin regions.

Extended Data Fig. S6. EPO-responsive promoter accessibility across hippocampal lineages.

(A) Heatmap showing $\log_2$ fold-change in promoter accessibility (FDR < 0.01) between rhEPO- and placebo-treated samples across major hippocampal cell types. Each row corresponds to a lineage and each column to a promoter-associated differentially accessible region (DAR). Red indicates increased accessibility under rhEPO, and blue indicates decreased accessibility. The total number of upregulated and downregulated promoter DARs for each lineage is shown on the right. Pyramidal neurons exhibit the largest number of EPO-responsive promoters.

(B) Differential accessibility heatmaps summarizing genome-wide DARs (left) and promoter-specific DARs for selected neurogenesis-related genes (right) across hippocampal cell types. The genome-wide heatmap depicts hierarchical clustering of all DARs between rhEPO and placebo, with colors indicating $\log_2$ fold-changes. The promoter-focused heatmap highlights accessibility changes at key neurogenic regulators; these genes are Ascl1, Neurog2, Pax6, Tbr1, Nrg3, Zbtb16, Zbtb18, Fabp5, Mdga2, and Dusp6. Asterisks denote significant change upon the rhEPO-treatment (FDR < 0.01).

Extended Data Fig. S7. Regulon activity across hippocampal lineages under rhEPO and placebo conditions.

Binary heatmap showing the presence or absence of Pando-inferred regulons across major hippocampal cell types for rhEPO-treated and placebo-treated conditions. Each row represents a regulon, and each column corresponds to a cell type, annotated with color-coded lineage identities. Black filled boxes indicate regulons that are active (present) in a given lineage, while white filled boxes indicate the absence of regulon activity. The comparison highlights regulons that are active or absent under rhEPO or placebo treatment and reveals lineage-specific differences in regulatory program activation across the hippocampus.

Extended Data Fig. S8. Transposable element families enriched within EPO-responsive chromatin regions.

(A) Bar plot showing the number of transposable element (TE) loci overlapping differentially accessible regions (DARs) identified between rhEPO- and placebo-treated samples. TE families are ranked by the number of overlapping loci, revealing strong contributions from L1, ERVL-MaLR, B4, ERVK, Alu, B2, and other repeat classes to EPO-responsive changes in chromatin accessibility.

(B) Heatmap displaying the $\log_2$ fold-change in accessibility for TE-associated DARs across major hippocampal cell types. Rows represent TE families, and columns represent cell lineages. Color intensity reflects mean chromatin accessibility differences (rhEPO vs. placebo), with red indicating increased accessibility and blue indicating decreased accessibility. The heatmap highlights lineage-specific patterns of TE-derived accessibility remodeling in response to rhEPO treatment.

Extended Data Fig. S9. Expansion of newly formed pyramidal neuron populations under rhEPO treatment.

(A) Bar plot showing the relative fraction of newly formed pyramidal neuron subtypes in placebo- and rhEPO-treated hippocampi, normalized to the total number of nuclei. Across multiple early-stage neuronal populations, including superficial migrating, deep-layer Foxp2$^+$ and migrating deep Sox5$^+$ neurons, ventral CA1 migrating, CA1 retrohippocampal, and ventral CA3 serotonin neuron subtypes, rhEPO treatment leads to a substantial increase in cell abundance compared with placebo conditions.

(B) Table summarizing the absolute number of nuclei assigned to each newly formed pyramidal neuron cluster under placebo and rhEPO conditions. The numerical increases further illustrate the strong expansion of newly formed pyramidal lineages following rhEPO treatment.

Extended Data Fig. S10. Integration of snATAC-seq and snRNA-seq datasets and annotation of pyramidal neuron lineages.

(A) UMAP representation of Seurat-defined chromatin accessibility clusters from snATAC-seq, with each color indicating a distinct cluster ID.

(B) Joint UMAP embedding of snATAC-seq (red) and snRNA-seq (blue) datasets, showing strong overlap between modalities following integration and confirming successful alignment of chromatin accessibility and transcriptomic profiles.

(C) UMAP visualization of chromatin accessibility profiles for annotated pyramidal neuron sublineages.

(D) Violin plots of label transfer prediction scores for pyramidal neuron subtypes, reflecting the confidence of cross-modality cell identity assignments.

Extended Data Fig. S11. Genomic distribution and Transposable element (TE) association of differentially accessible regions (DARs) in newly formed and mature pyramidal neurons.

(A) Bar plot showing the number of DARs located in promoter-proximal regions (≤1 kb), distal regions (1–10 kb), and intergenic regions (>10 kb) for newly formed and mature pyramidal neurons (rhEPO vs. placebo). Newly formed neurons exhibit a larger number and higher proportion of distal and intergenic DARs.

(B) Density plot of distances from DAR summits to the nearest transcription start site (TSS) on a log-scaled x-axis. Newly formed and mature neurons show broadly similar distance profiles, with both distributions peaking around 10 kilobases from the nearest TSS

(C) Heatmap of $\log_2$ fold-change values for TE-overlapping DARs in newly formed and mature neurons (rhEPO vs. placebo). Rows represent TE-associated peaks, illustrating distinct groups of TE-derived regions that become more or less accessible under EPO treatment depending on neuronal maturation state.

(D) Bar plot summarizing the number of TE-overlapping DARs that increase (pink) or decrease (blue) in accessibility in response to rhEPO within newly formed and mature neurons. Newly formed neurons exhibit a greater number of rhEPO-induced accessibility gains or losses compared to mature neurons.

Extended Data Fig. S12. Transposable element families enriched within neurogenic transcription factor ChIP-seq peaks.

Matrix showing TE families significantly enriched ($p < 0.001$) within ChIP-seq binding sites of neurogenic transcription factors. Rows correspond to transcription factors (TBR1, SOX2, PAX6, NEUROG2, NEUROD2, NEUROD1, JUN, FOS, FOXO3, FOXG1, EGR1, ASCL1), and columns represent individual TE families. Black squares indicate significant enrichment ($p < 0.001$), while white squares indicate non-significant enrichment ($p > 0.001$). Only TE families enriched in ≥2 independent ChIP-seq datasets are shown. The matrix highlights recurrent engagement of diverse TE families by key neurogenic TFs, supporting widespread co-option of TE-derived sequences in neuronal regulatory networks.

Extended Data Fig. S13. Gene ontology enrichment for genes located within ±10 kb of TE-overlapping DARs.

Dot plots showing the top 30 enriched Gene Ontology (GO) terms for genes whose transcription start sites lie within ±10 kb of transposable element–overlapping differentially accessible regions (TE-DARs; rhEPO vs. placebo). Enrichment results are shown for Biological Process and Cellular Component. Dot size indicates the number of genes associated with each term, and dot color represents statistical significance ($-\log_{10}$ q-value). The enriched categories highlight pathways related to neurogenesis, synapse organization, neuronal signaling, ion transport, and membrane specialization, etc., reflecting functional processes linked to TE-derived regulatory remodeling under rhEPO treatment.

SUPPLEMENTARY INFORMATION

Supplementary Table S1: Summary of sequencing and alignment metrics for snATAC–seq samples.

Supplementary Table S2: Transcription factor motif enrichment across Seurat-defined hippocampal clusters without treatment separation.

Supplementary Table S3: Transcription factor motif enrichment across the 11 major hippocampal lineages without treatment separation.

Supplementary Table S4: Summary of significant accessible chromatin peaks across the 11 hippocampal lineages and detailed annotation of peaks not present in the Mouse Brain Atlas.

Supplementary Table S5: Transposable element family enrichment within all accessible chromatin regions.

Supplementary Table S6: Globally defined differentially accessible regions, motif enrichments, TE-associated regulatory elements, and matched snRNA-seq comparisons for rhEPO and placebo conditions.

Supplementary Table S7: Correlation between differentially accessible regions and differentially expressed genes after filtering for significant values.

Supplementary Table S8: Significantly differentially accessible regions between rhEPO- and placebo-treated samples across the 11 hippocampal lineages.

Supplementary Table S9: Significantly differentially accessible regions between EPO- and placebo-treated samples across the 11 hippocampal lineages, annotated for proximal genes and overlapping transposable elements.

Supplementary Table S10: Gene regulatory network modules inferred for each of the 11 hippocampal lineages under EPO and placebo conditions.

Supplementary Table S11: Enrichment of transposable element families within differentially accessible regions identified between EPO- and placebo-treated samples.

Supplementary Table S12: Significant accessible chromatin peaks distinguishing newly formed from mature pyramidal neuron lineages.

Supplementary Table S13: Differentially accessible regions between EPO- and placebo-treated samples within newly formed and mature pyramidal neurons, annotated for nearby genes and transposable elements.

Supplementary Table S14: Enrichment of transposable element families within ChIP–seq peak sets for neurodevelopmental transcription factors.

Supplementary Table S15: Transposable element–associated differentially accessible regions located within 10 kb of annotated transcription start sites.

Supplementary Table S16: Gene Ontology analysis of genes associated with TE-overlapping differentially accessible regions bound by transcription factors.

Supplementary Table S17: List of neurogenic transcription factors included in TF-binding and enrichment analyses.

Figure 1

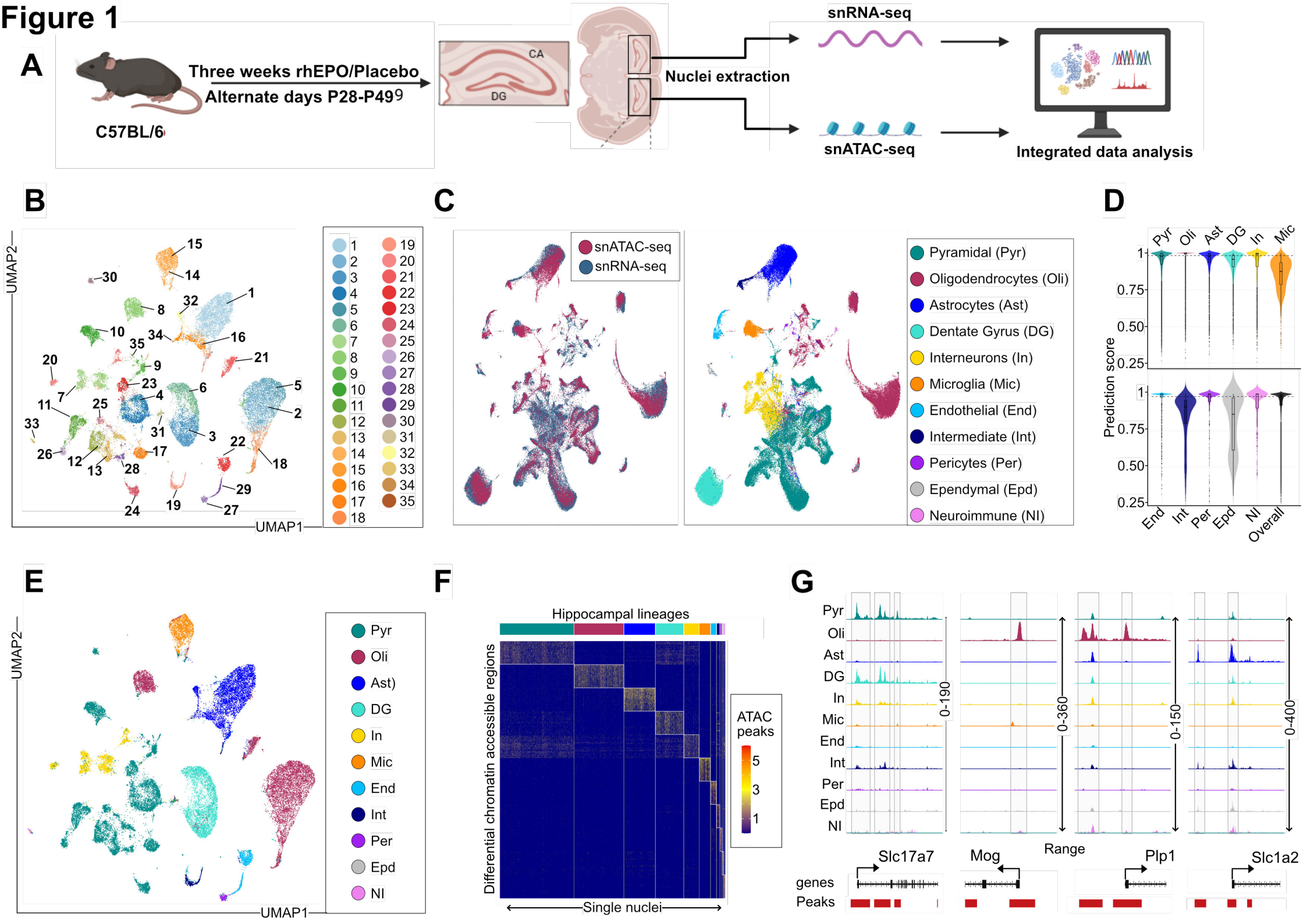

A
Three weeks rhEPO/Placebo
Alternate days P28-P49
C57BL/6
CA
DG
Nuclei extraction
snRNA-seq
snATAC-seq
Integrated data analysis
B
UMAP2
UMAP1
C
snATAC-seq
snRNA-seq
Pyramidal (Pyr)
Oligodendrocytes (Oli)
Astrocytes (Ast)
Dentate Gyrus (DG)
Interneurons (In)
Microglia (Mic)
Endothelial (End)
Intermediate (Int)
Pericytes (Per)
Ependymal (Epd)
Neuroimmune (NI)
D
Pyr
Oli
Ast
DG
In
Mic
Prediction score
End
Int
Per
Epd
NI
Overall
E
UMAP2
UMAP1
Pyr
Oli
Ast)
DG
In
Mic
End
Int
Per
Epd
NI
F
Hippocampal lineages
Differential chromatin accessible regions
ATAC peaks
Single nuclei
G
Pyr
Oli
Ast
DG
In
Mic
End
Int
Per
Epd
NI
0-190
0-360
0-150
0-400
Range
Slc17a7
Mog
Plp1
Slc1a2
genes
Peaks

Figure 2

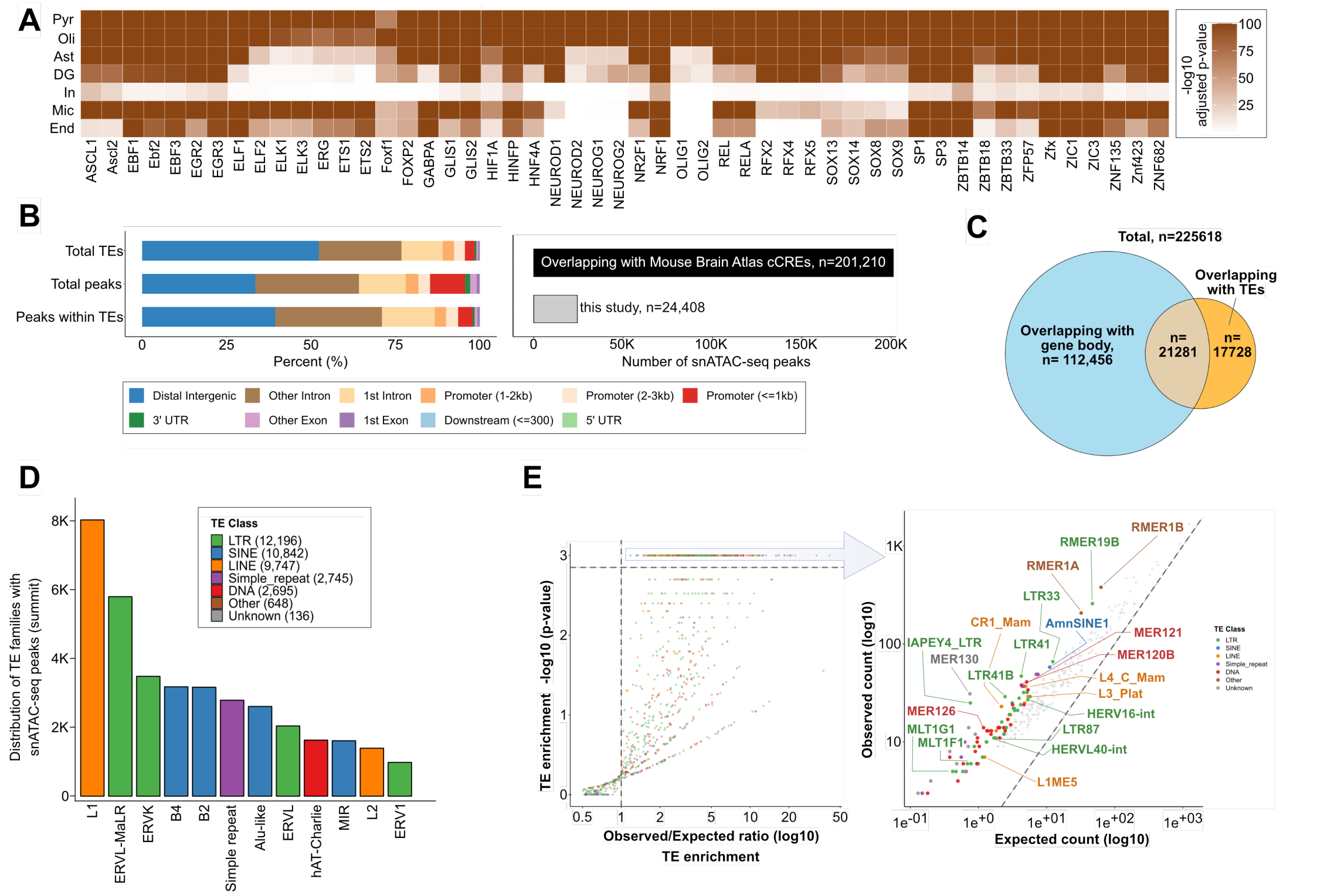

A
Pyr
Oli
Ast
DG
In
Mic
End
ASCL1 Ascl2 EBF1 Ebf2 EBF3 EGR2 EGR3 ELF1 ELF2 ELK1 ELK3 ERG ETS1 ETS2 Foxf1 FOXP2 GABPA GLIS1 GLIS2 HIF1A HINFP HNF4A NEUROD1 NEUROD2 NEUROG1 NEUROG2 NR2F1 NRF1 OLIG1 OLIG2 REL RELA RFX2 RFX4 RFX5 SOX13 SOX14 SOX8 SOX9 SP1 SP3 ZBTB14 ZBTB18 ZBTB33 ZFP57 Zfx ZIC1 ZIC3 ZNF135 Znf423 ZNF682
-log10 adjusted p-value
100
75
50
25
B
Total TEs
Total peaks
Peaks within TEs
0
25
50
75
100
Percent (%)
Overlapping with Mouse Brain Atlas cCREs, n=201,210
this study, n=24,408
0
50K
100K
150K
200K
Number of snATAC-seq peaks
Distal Intergenic
Other Intron
1st Intron
Promoter (1-2kb)
Promoter (2-3kb)
Promoter (<=1kb)
3' UTR
Other Exon
1st Exon
Downstream (<=300)
5' UTR
C
Total, n=225618
Overlapping with TEs
Overlapping with gene body, n= 112,456
n= 21281
n= 17728
D
Distribution of TE families with snATAC-seq peaks (summit)
8K
6K
4K
2K
0
TE Class
LTR (12,196)
SINE (10,842)
LINE (9,747)
Simple_repeat (2,745)
DNA (2,695)
Other (648)
Unknown (136)
L1
ERVL-MaLR
ERVK
B4
B2
Simple repeat
Alu-like
ERVL
hAT-Charlie
MIR
L2
ERV1
E
TE enrichment -log10 (p-value)
3
2
1
0
0.5
1
2
5
10
20
50
Observed/Expected ratio (log10)
TE enrichment
Observed count (log10)
1K
100
10
1e-01
1e+0
1e+01
1e+02
1e+03
Expected count (log10)
RMER1B
RMER19B
RMER1A
LTR33
CR1_Mam
AmnSINE1
MER121
IAPEY4_LTR
LTR41
MER120B
MER130
LTR41B
L4_C_Mam
L3_Plat
MER126
HERV16-int
MLT1G1
LTR87
MLT1F1
HERVL40-int
L1ME5
TE Class
LTR
SINE
LINE
Simple_repeat
DNA
Other
Unknown

Figure 3

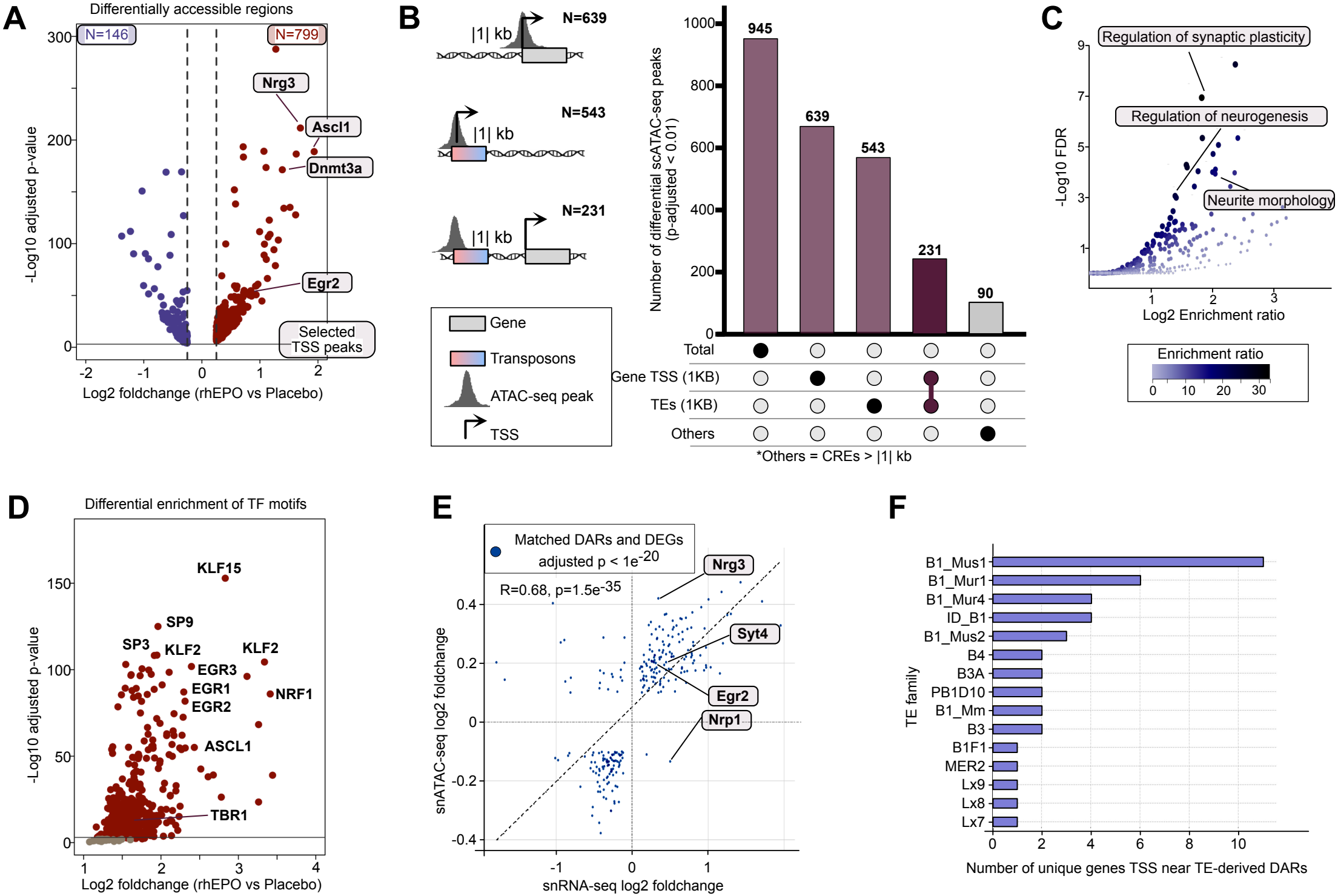

A
Differentially accessible regions
N=146
N=799
Nrg3
Ascl1
Dnmt3a
Egr2
Selected TSS peaks
-Log10 adjusted p-value
Log2 foldchange (rhEPO vs Placebo)
B
|1| kb
N=639
N=543
N=231
Gene
Transposons
ATAC-seq peak
TSS
Number of differential scATAC-seq peaks (p-adjusted < 0.01)
945
639
543
231
90
Total
Gene TSS (1KB)
TEs (1KB)
Others
*Others = CREs > |1| kb
C
Regulation of synaptic plasticity
Regulation of neurogenesis
Neurite morphology
-Log10 FDR
Log2 Enrichment ratio
Enrichment ratio
D
Differential enrichment of TF motifs
KLF15
SP9
SP3
KLF2
KLF2
EGR3
EGR1
EGR2
NRF1
ASCL1
TBR1
-Log10 adjusted p-value
Log2 foldchange (rhEPO vs Placebo)
E
Matched DARs and DEGs
adjusted p < 1e-20
R=0.68, p=1.5e-35
Nrg3
Syt4
Egr2
Nrp1
snATAC-seq log2 foldchange
snRNA-seq log2 foldchange
F
B1_Mus1
B1_Mur1
B1_Mur4
ID_B1
B1_Mus2
B4
B3A
PB1D10
B1_Mm
B3
B1F1
MER2
Lx9
Lx8
Lx7
TE family
Number of unique genes TSS near TE-derived DARs

Figure 4

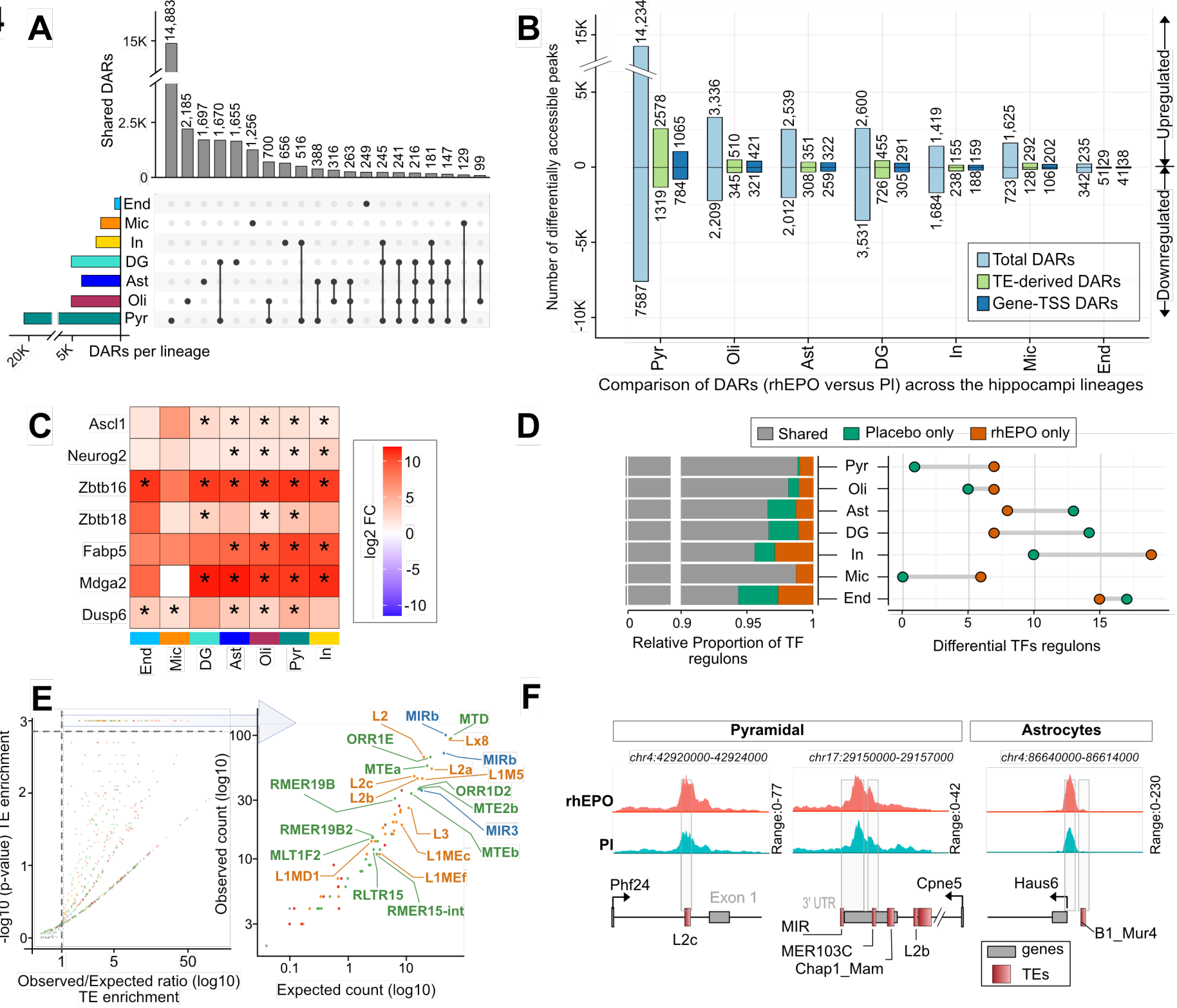

A
Shared DARs
15K
2.5K
0
14,883
2,185
1,697
1,670
1,655
1,256
700
656
516
388
316
263
249
245
241
216
181
147
129
99
End
Mic
In
DG
Ast
Oli
Pyr
20K
5K
DARs per lineage
B
Number of differentially accessible peaks
15K
5K
0
-5K
-10K
Upregulated
Downregulated
Total DARs
TE-derived DARs
Gene-TSS DARs
Pyr
Oli
Ast
DG
In
Mic
End
Comparison of DARs (rhEPO versus PI) across the hippocampi lineages
C
Ascl1
Neurog2
Zbtb16
Zbtb18
Fabp5
Mdga2
Dusp6
log2 FC
10
5
0
-5
-10
End
Mic
DG
Ast
Oli
Pyr
In
D
Shared
Placebo only
rhEPO only
Pyr
Oli
Ast
DG
In
Mic
End
0
0.9
0.95
1
Relative Proportion of TF regulons
0
5
10
15
Differential TFs regulons
E
-log10 (p-value) TE enrichment
Observed/Expected ratio (log10) TE enrichment
Observed count (log10)
Expected count (log10)
L2
MIRb
MTD
Lx8
ORR1E
MIRb
MTEa
L2a
L1M5
RMER19B
L2c
L2b
ORR1D2
MTE2b
MIR3
RMER19B2
L3
MTEb
MLT1F2
L1MEc
L1MD1
L1MEf
RLTR15
RMER15-int
F
Pyramidal
Astrocytes
chr4:42920000-42924000
chr17:29150000-29157000
chr4:86640000-86614000
rhEPO
PI
Range:0-77
Range:0-42
Range:0-230
Phf24
Exon 1
L2c
3' UTR
Cpne5
MIR
MER103C
Chap1_Mam
L2b
Haus6
B1_Mur4
genes
TEs

Figure 5

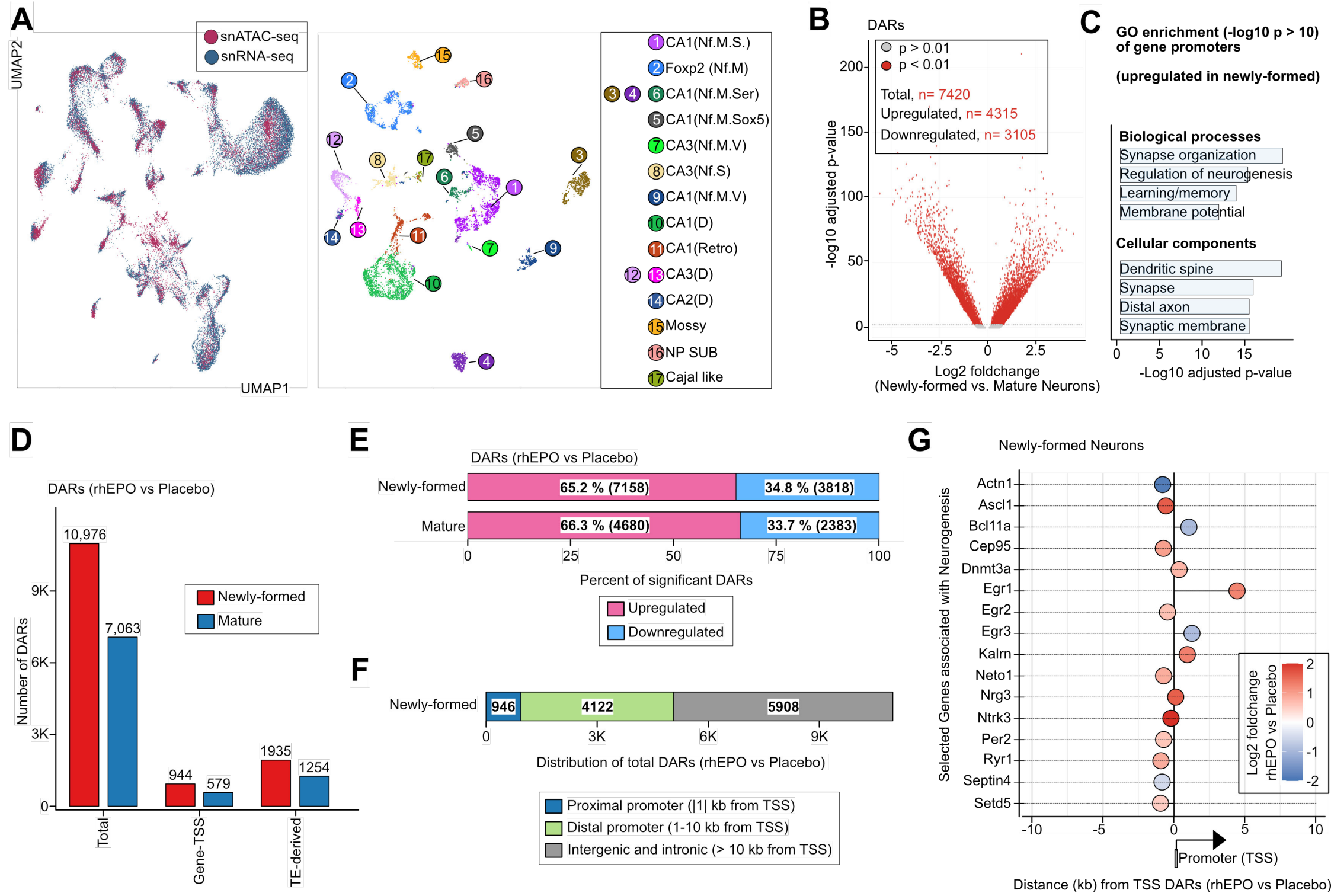
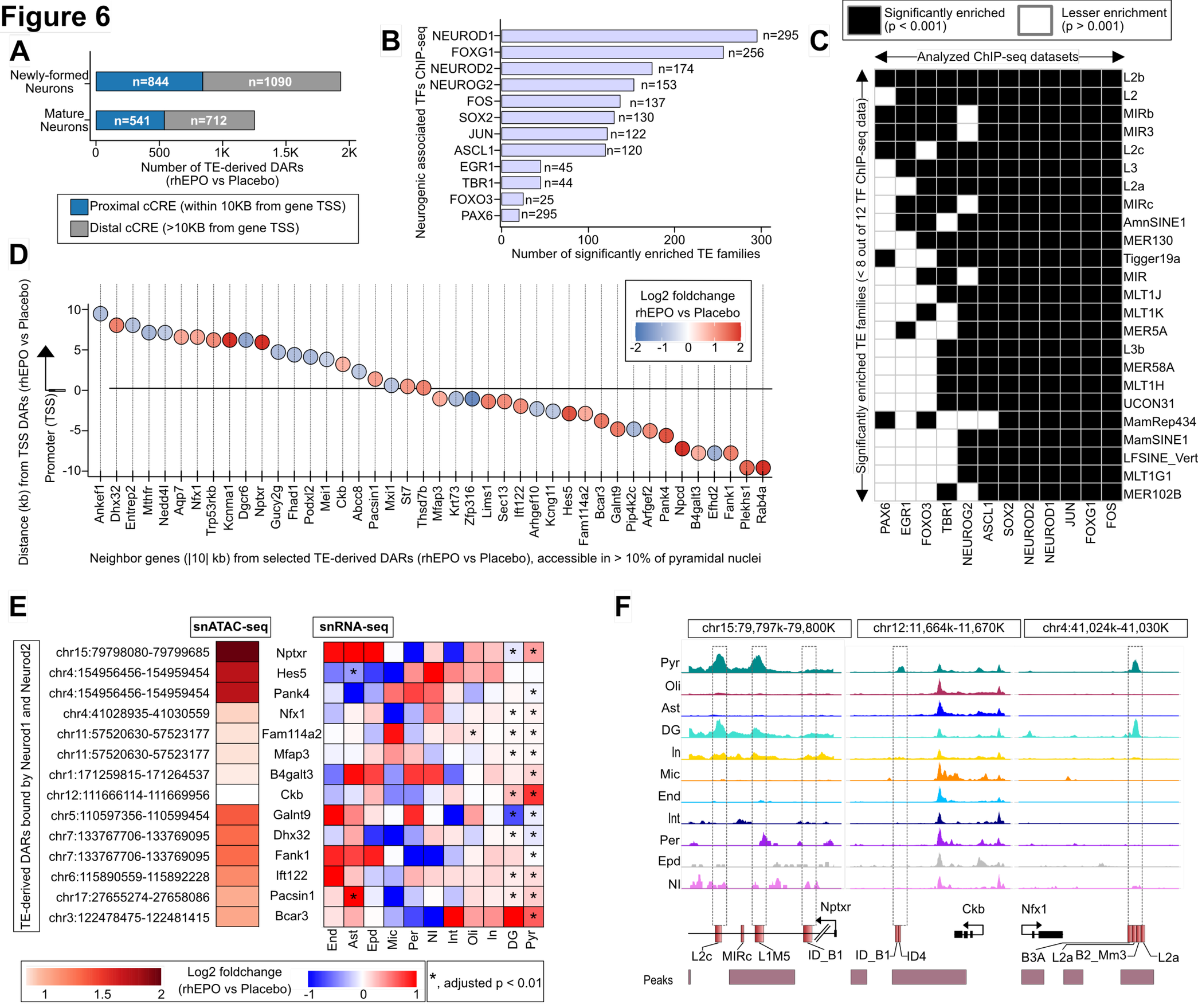

A
snATAC-seq
snRNA-seq
UMAP2
UMAP1
1 CA1(Nf.M.S.)
2 Foxp2 (Nf.M)
3 4 6 CA1(Nf.M.Ser)
5 CA1(Nf.M.Sox5)
7 CA3(Nf.M.V)
8 CA3(Nf.S)
9 CA1(Nf.M.V)
10 CA1(D)
11 CA1(Retro)
12 13 CA3(D)
14 CA2(D)
15 Mossy
16 NP SUB
17 Cajal like
B
DARs
p > 0.01
p < 0.01
Total, n= 7420
Upregulated, n= 4315
Downregulated, n= 3105
-log10 adjusted p-value
Log2 foldchange
(Newly-formed vs. Mature Neurons)
C
GO enrichment (-log10 p > 10)
of gene promoters
(upregulated in newly-formed)
Biological processes
Synapse organization
Regulation of neurogenesis
Learning/memory
Membrane potential
Cellular components
Dendritic spine
Synapse
Distal axon
Synaptic membrane
-Log10 adjusted p-value
D
DARs (rhEPO vs Placebo)
Number of DARs
Newly-formed
Mature
10,976
7,063
944
579
1935
1254
Total
Gene-TSS
TE-derived
E
DARs (rhEPO vs Placebo)
Newly-formed
65.2 % (7158)
34.8 % (3818)
Mature
66.3 % (4680)
33.7 % (2383)
Percent of significant DARs
Upregulated
Downregulated
F
Newly-formed
946
4122
5908
Distribution of total DARs (rhEPO vs Placebo)
Proximal promoter (|1| kb from TSS)
Distal promoter (1-10 kb from TSS)
Intergenic and intronic (> 10 kb from TSS)
G
Newly-formed Neurons
Selected Genes associated with Neurogenesis
Actn1
Ascl1
Bcl11a
Cep95
Dnmt3a
Egr1
Egr2
Egr3
Kalrn
Neto1
Nrg3
Ntrk3
Per2
Ryr1
Septin4
Setd5
Log2 foldchange
rhEPO vs Placebo
Promoter (TSS)
Distance (kb) from TSS DARs (rhEPO vs Placebo)

Figure 6


A
Newly-formed Neurons
n=844
n=1090
Mature Neurons
n=541
n=712
0
500
1K
1.5K
2K
Number of TE-derived DARs (rhEPO vs Placebo)
Proximal cCRE (within 10KB from gene TSS)
Distal cCRE (>10KB from gene TSS)
B
Neurogenic associated TFs ChIP-seq
NEUROD1 n=295
FOXG1 n=256
NEUROD2 n=174
NEUROG2 n=153
FOS n=137
SOX2 n=130
JUN n=122
ASCL1 n=120
EGR1 n=45
TBR1 n=44
FOXO3 n=25
PAX6 n=295
0
100
200
300
Number of significantly enriched TE families
C
Significantly enriched (p < 0.001)
Lesser enrichment (p > 0.001)
Analyzed ChIP-seq datasets
Significantly enriched TE families (< 8 out of 12 TF ChIP-seq data)
L2b
L2
MIRb
MIR3
L2c
L3
L2a
MIRc
AmnSINE1
MER130
Tigger19a
MIR
MLT1J
MLT1K
MER5A
L3b
MER58A
MLT1H
UCON31
MamRep434
MamSINE1
LFSINE_Vert
MLT1G1
MER102B
PAX6
EGR1
FOXO3
TBR1
NEUROG2
ASCL1
SOX2
NEUROD2
NEUROD1
JUN
FOXG1
FOS
D
Distance (kb) from TSS DARs (rhEPO vs Placebo)
Promoter (TSS)
10
5
0
-5
-10
Log2 foldchange rhEPO vs Placebo
-2
-1
0
1
2
Ankef1
Dhx32
Entrep2
Mthfr
Nedd4l
Aqp7
Nfx1
Trp53rkb
Kcnma1
Dgcr6
Nptxr
Gucy2g
Fhad1
Podxl2
Mei1
Ckb
Abcc8
Pacsin1
Mxi1
St7
Thsd7b
Mfap3
Krt73
Zfp316
Lims1
Sec13
Ift122
Arhgef10
Kcng11
Hes5
Fam114a2
Bcar3
Galnt9
Pip4k2c
Arfgef2
Pank4
Npcd
B4galt3
Efhd2
Fank1
Plekhs1
Rab4a
Neighbor genes (|10| kb) from selected TE-derived DARs (rhEPO vs Placebo), accessible in > 10% of pyramidal nuclei
E
snATAC-seq
snRNA-seq
TE-derived DARs bound by Neurod1 and Neurod2
chr15:79798080-79799685
chr4:154956456-154959454
chr4:154956456-154959454
chr4:41028935-41030559
chr11:57520630-57523177
chr11:57520630-57523177
chr1:171259815-171264537
chr12:111666114-111669956
chr5:110597356-110599454
chr7:133767706-133769095
chr7:133767706-133769095
chr6:115890559-115892228
chr17:27655274-27658086
chr3:122478475-122481415
Nptxr
Hes5
Pank4
Nfx1
Fam114a2
Mfap3
B4galt3
Ckb
Galnt9
Dhx32
Fank1
Ift122
Pacsin1
Bcar3
End
Ast
Epd
Mic
Per
NI
Int
Oli
In
DG
Pyr
Log2 foldchange (rhEPO vs Placebo)
1
1.5
2
-1
0
1
*, adjusted p < 0.01
F
chr15:79,797k-79,800K
chr12:11,664k-11,670K
chr4:41,024k-41,030K
Pyr
Oli
Ast
DG
In
Mic
End
Int
Per
Epd
NI
Nptxr
Ckb
Nfx1
L2c
MIRc
L1M5
ID_B1
ID_B1
ID4
B3A
L2a
B2_Mm3
L2a
Peaks

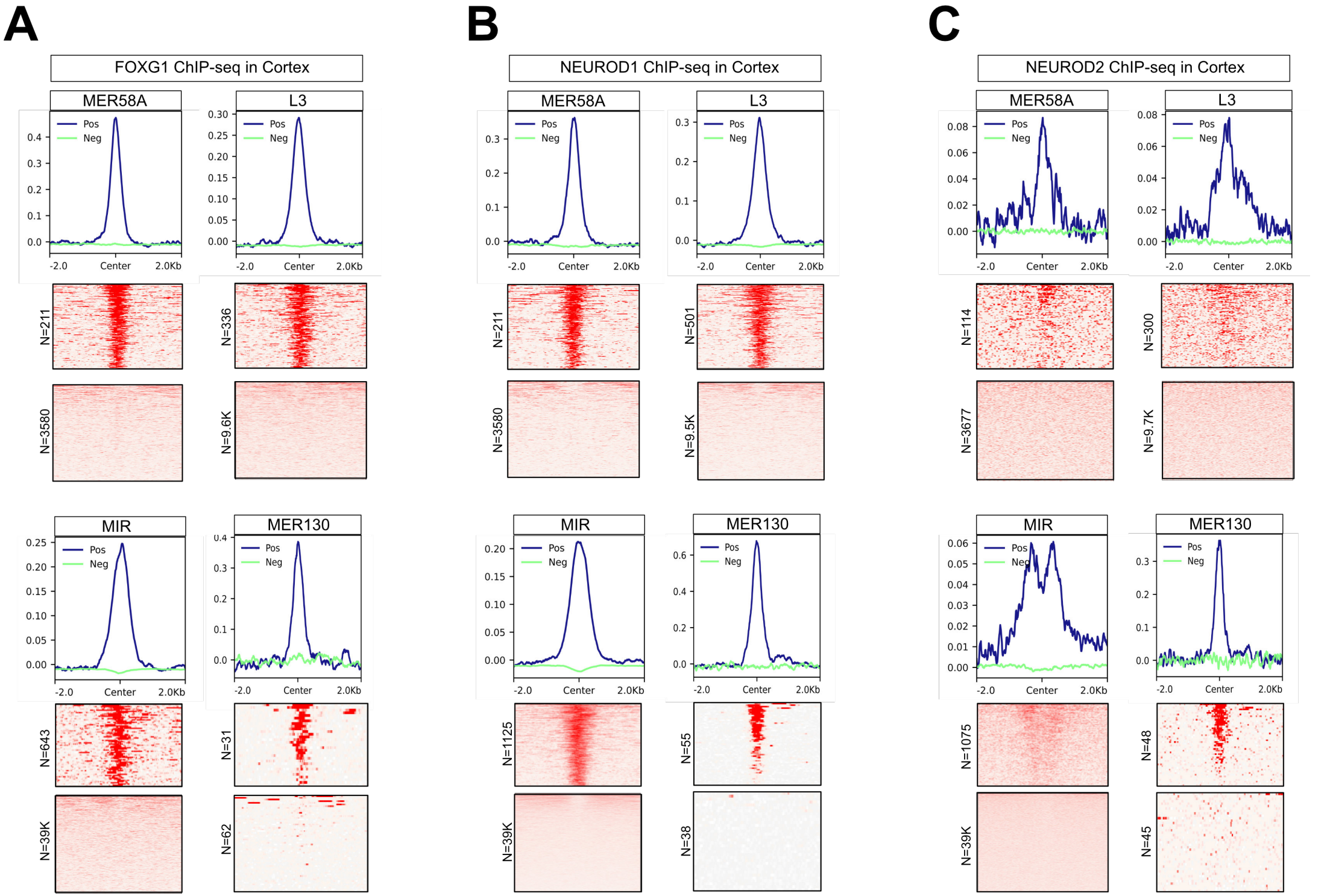

A
FOXG1 ChIP-seq in Cortex
MER58A
L3
MIR
MER130
Pos
Neg
-2.0
Center
2.0Kb
N=211
N=336
N=3580
N=9.6K
N=643
N=31
N=39K
N=62
B
NEUROD1 ChIP-seq in Cortex
N=211
N=501
N=3580
N=9.5K
N=1125
N=55
N=39K
N=38
C
NEUROD2 ChIP-seq in Cortex
N=114
N=300
N=3677
N=9.7K
N=1075
N=48
N=39K
N=45

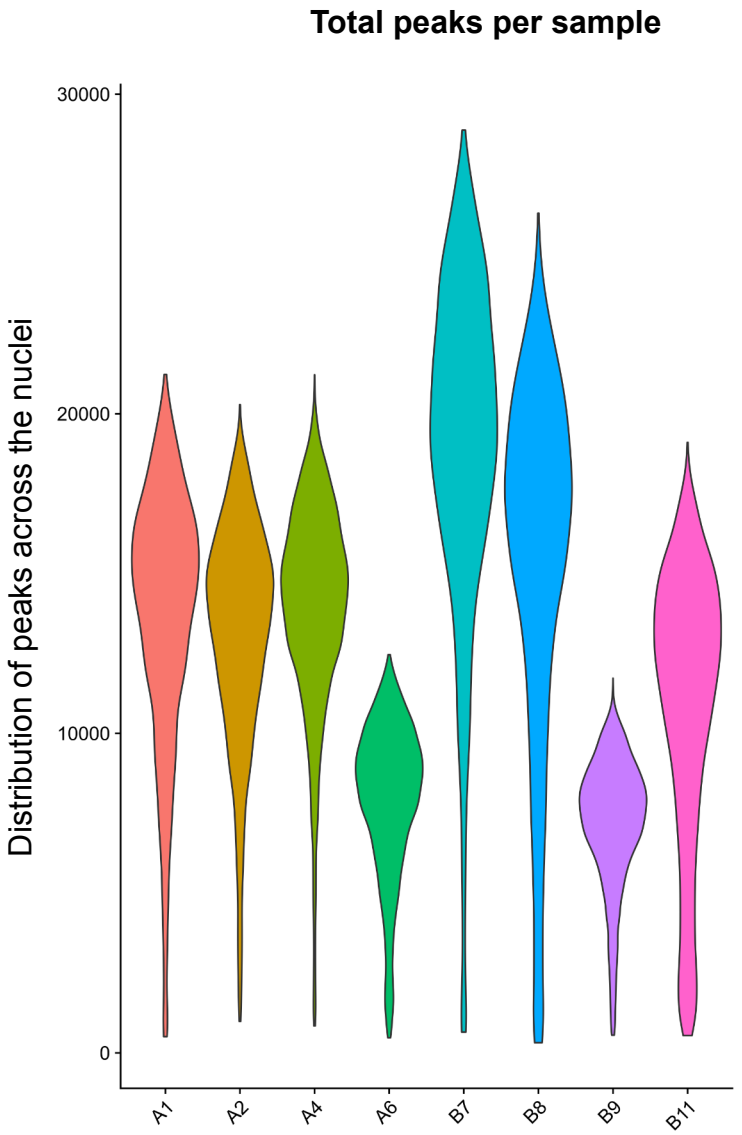
Figure S1

A

Alignment summary matrix (Cell Ranger ATAC)

Sample	# of cells	Mean raw read pairs per cell	Fraction of high-quality fragments overlapping peaks	Fraction of high-quality fragments in cells	Median high-quality fragments per cell	Fraction duplicates	# of peaks
Hipp A1	3383	33966.8265	0.4638	0.9586	24570	0.1794	159631
Hipp A2	2981	29247.2918	0.5279	0.9629	21243	0.18	162143
Hipp A4	3639	33372.0179	0.4922	0.9701	23830	0.1907	148187
Hipp A6	7750	19238.2001	0.5349	0.9103	14030.5	0.1349	172184
Hipp B7	1968	47319.7205	0.4962	0.9573	31900.5	0.2281	153405
Hipp B8	1910	35333.3021	0.5278	0.9636	25325.5	0.2084	144765
Hipp B9	6011	17002.5327	0.4835	0.9314	12673	0.1316	145402
Hipp B11	6586	27354.2473	0.5347	0.927	19867	0.1783	171436

B



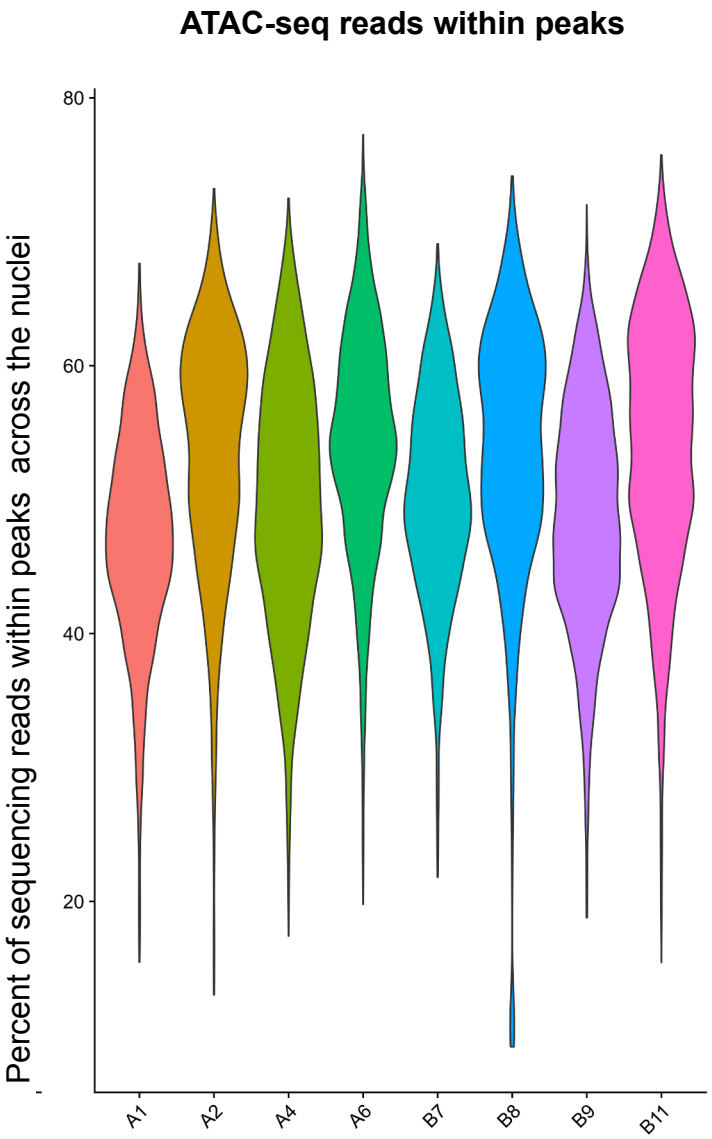


Figure S2

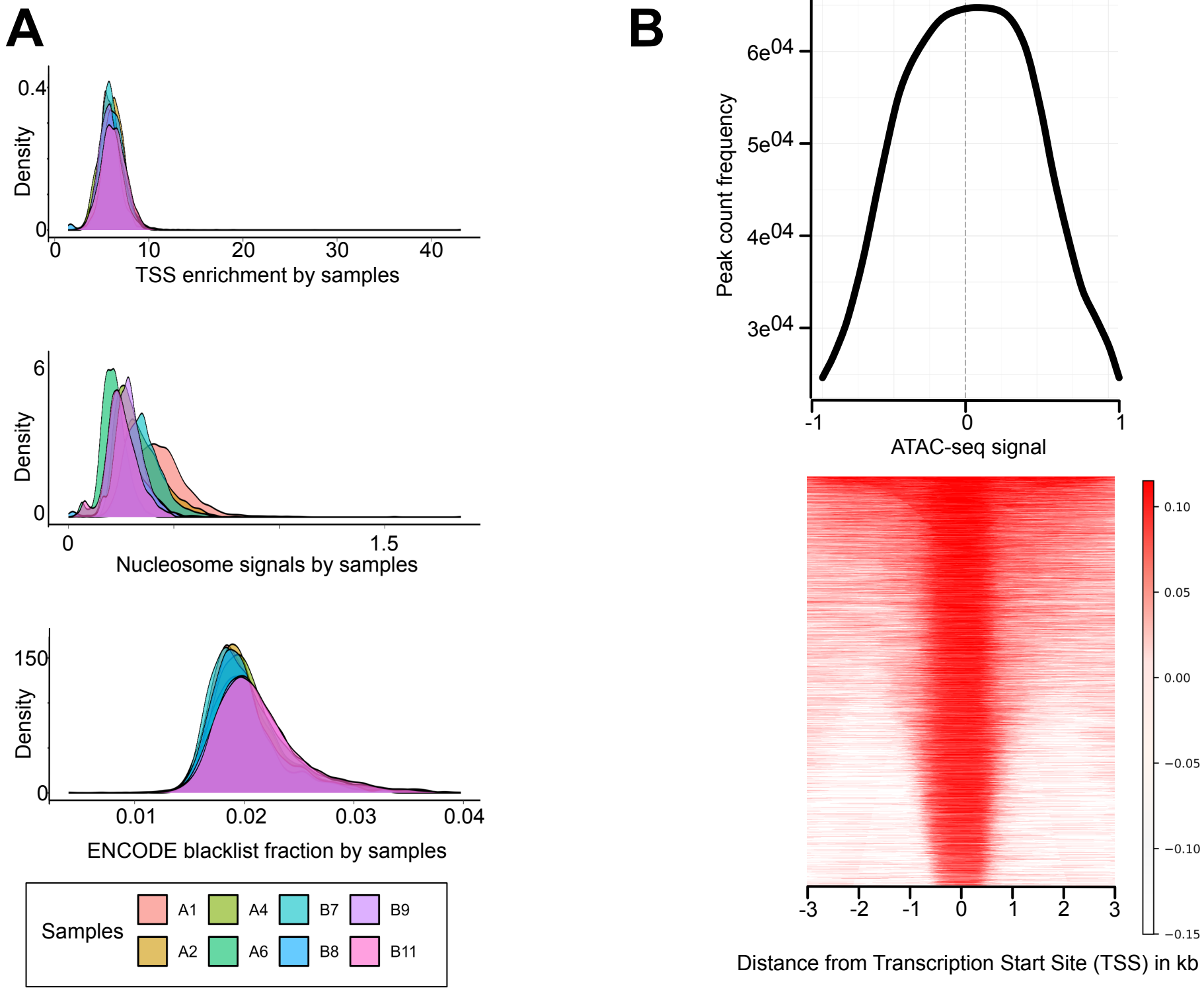

A
Density
0.4
0
0
10
20
30
40
TSS enrichment by samples
Density
6
0
0
1.5
Nucleosome signals by samples
Density
150
0
0.01
0.02
0.03
0.04
ENCODE blacklist fraction by samples
Samples
A1
A2
A4
A6
B7
B8
B9
B11
B
Peak count frequency
6e04
5e04
4e04
3e04
-1
0
1
ATAC-seq signal
0.10
0.05
0.00
−0.05
−0.10
−0.15
-3
-2
-1
0
1
2
3
Distance from Transcription Start Site (TSS) in kb

Figure S3

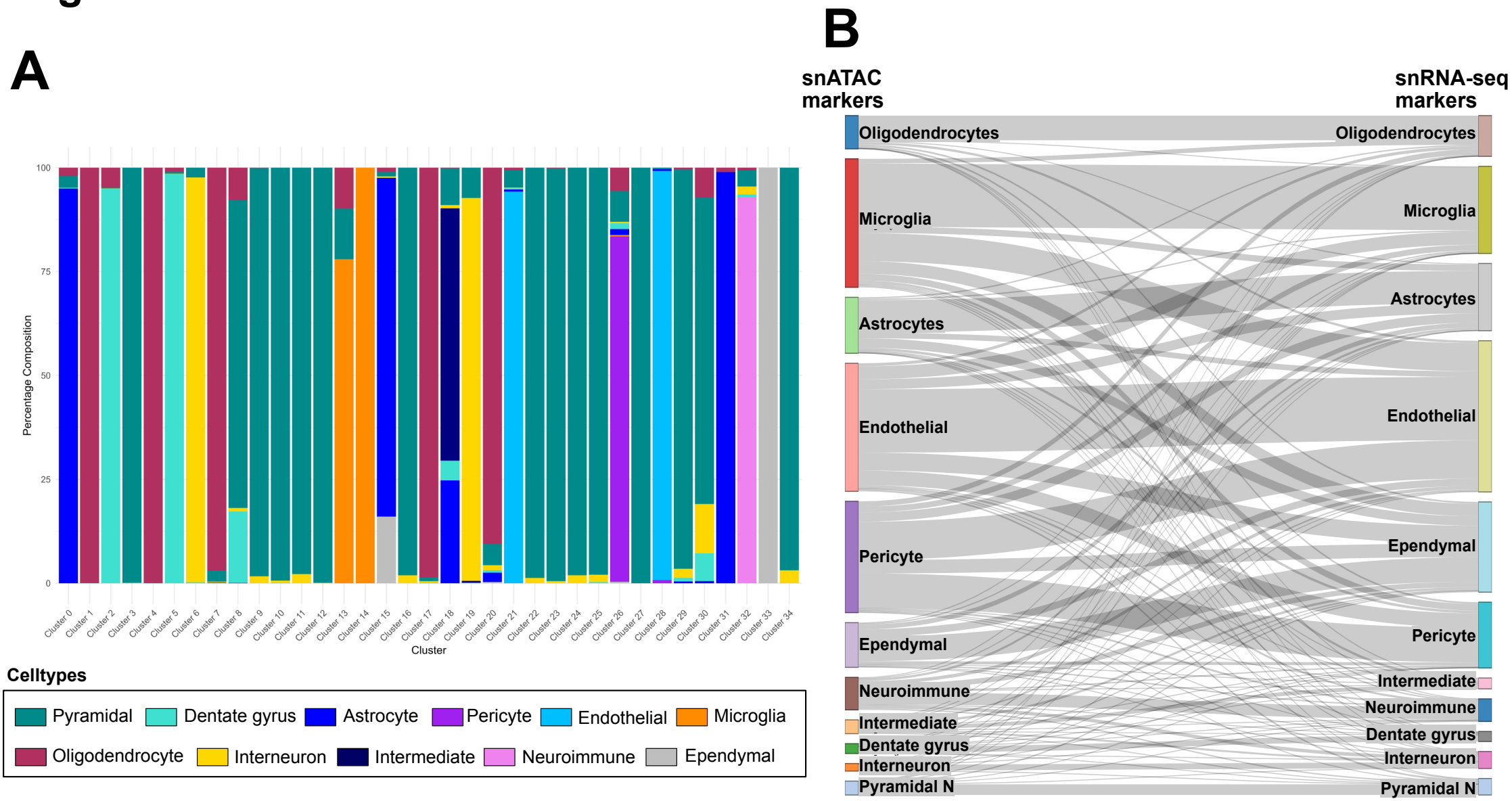

A
Percentage Composition
100
75
50
25
0
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Cluster 11
Cluster 12
Cluster 13
Cluster 14
Cluster 15
Cluster 16
Cluster 17
Cluster 18
Cluster 19
Cluster 20
Cluster 21
Cluster 22
Cluster 23
Cluster 24
Cluster 25
Cluster 26
Cluster 27
Cluster 28
Cluster 29
Cluster 30
Cluster 31
Cluster 32
Cluster 33
Cluster 34
Cluster
Celltypes
Pyramidal
Dentate gyrus
Astrocyte
Pericyte
Endothelial
Microglia
Oligodendrocyte
Interneuron
Intermediate
Neuroimmune
Ependymal
B
snATAC markers
snRNA-seq markers
Oligodendrocytes
Microglia
Astrocytes
Endothelial
Pericyte
Ependymal
Neuroimmune
Intermediate
Dentate gyrus
Interneuron
Pyramidal N
Oligodendrocytes
Microglia
Astrocytes
Endothelial
Ependymal
Pericyte
Intermediate
Neuroimmune
Dentate gyrus
Interneuron
Pyramidal N

Figure S4

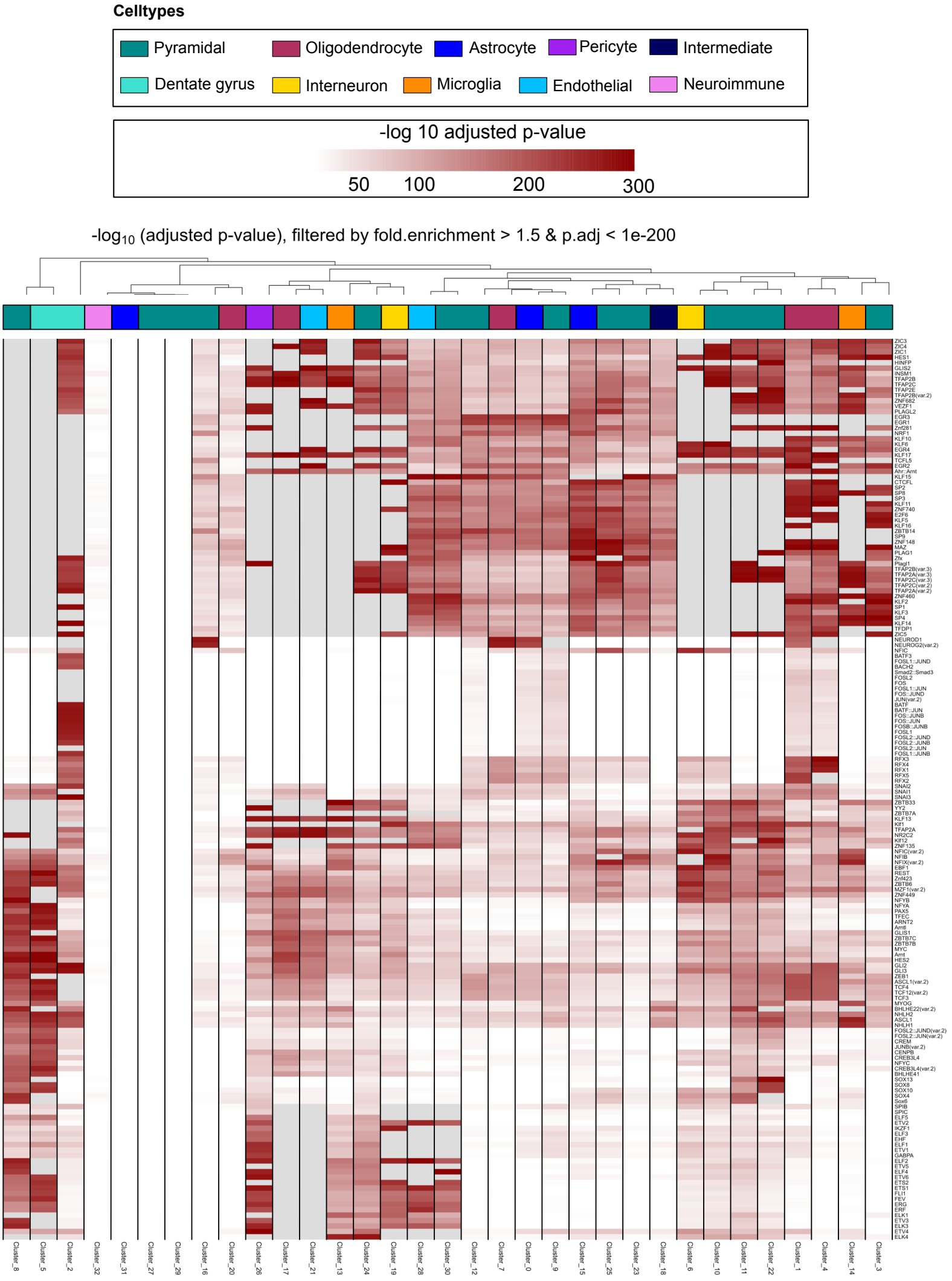

Celltypes
Pyramidal
Oligodendrocyte
Astrocyte
Pericyte
Intermediate
Dentate gyrus
Interneuron
Microglia
Endothelial
Neuroimmune
-log 10 adjusted p-value
50
100
200
300
-log10 (adjusted p-value), filtered by fold.enrichment > 1.5 & p.adj < 1e-200
Cluster_8
Cluster_5
Cluster_2
Cluster_32
Cluster_31
Cluster_27
Cluster_29
Cluster_16
Cluster_20
Cluster_26
Cluster_17
Cluster_21
Cluster_13
Cluster_24
Cluster_19
Cluster_28
Cluster_30
Cluster_12
Cluster_7
Cluster_0
Cluster_9
Cluster_15
Cluster_25
Cluster_23
Cluster_18
Cluster_6
Cluster_10
Cluster_11
Cluster_22
Cluster_1
Cluster_4
Cluster_14
Cluster_3

Figure S5

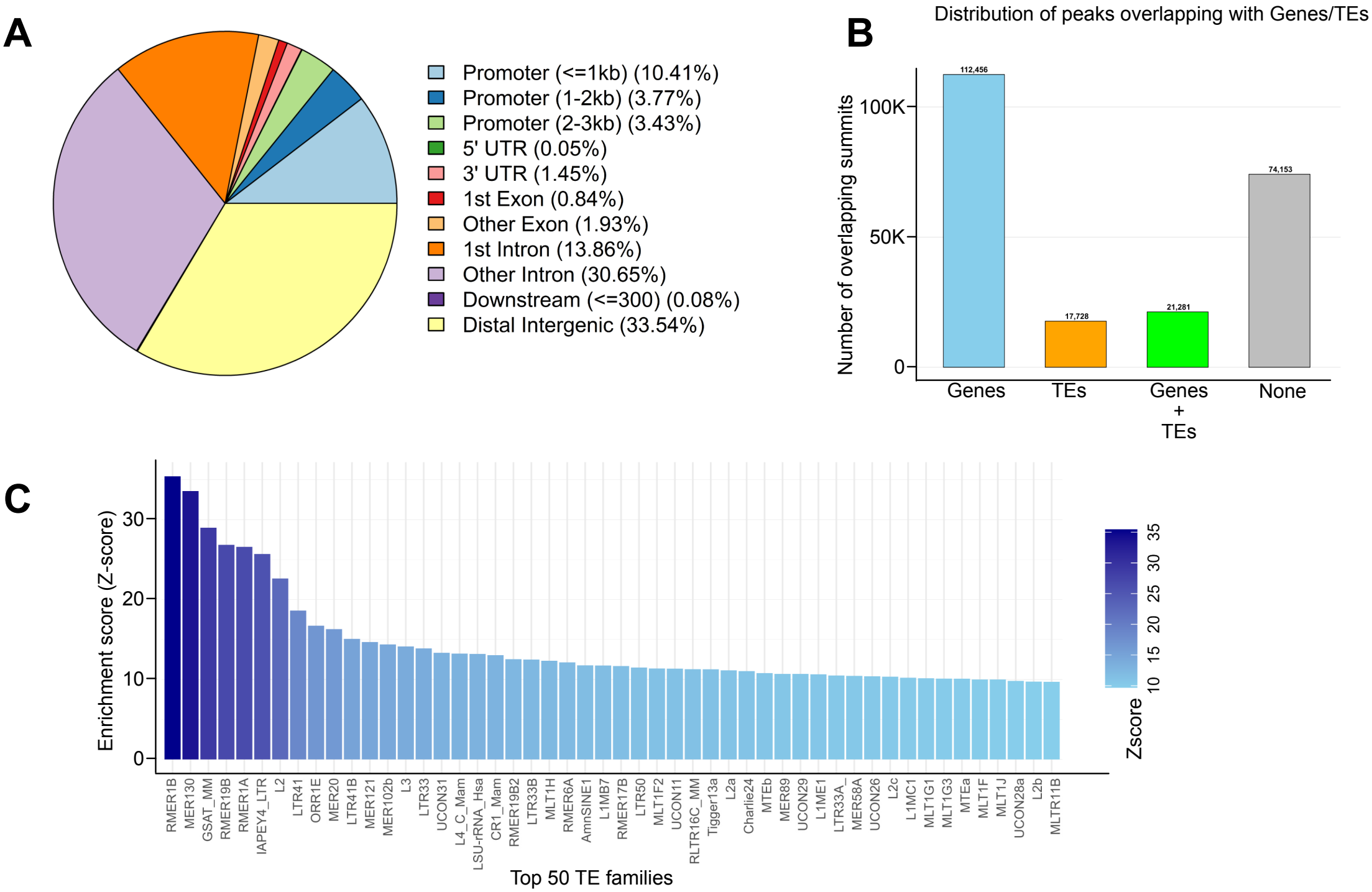
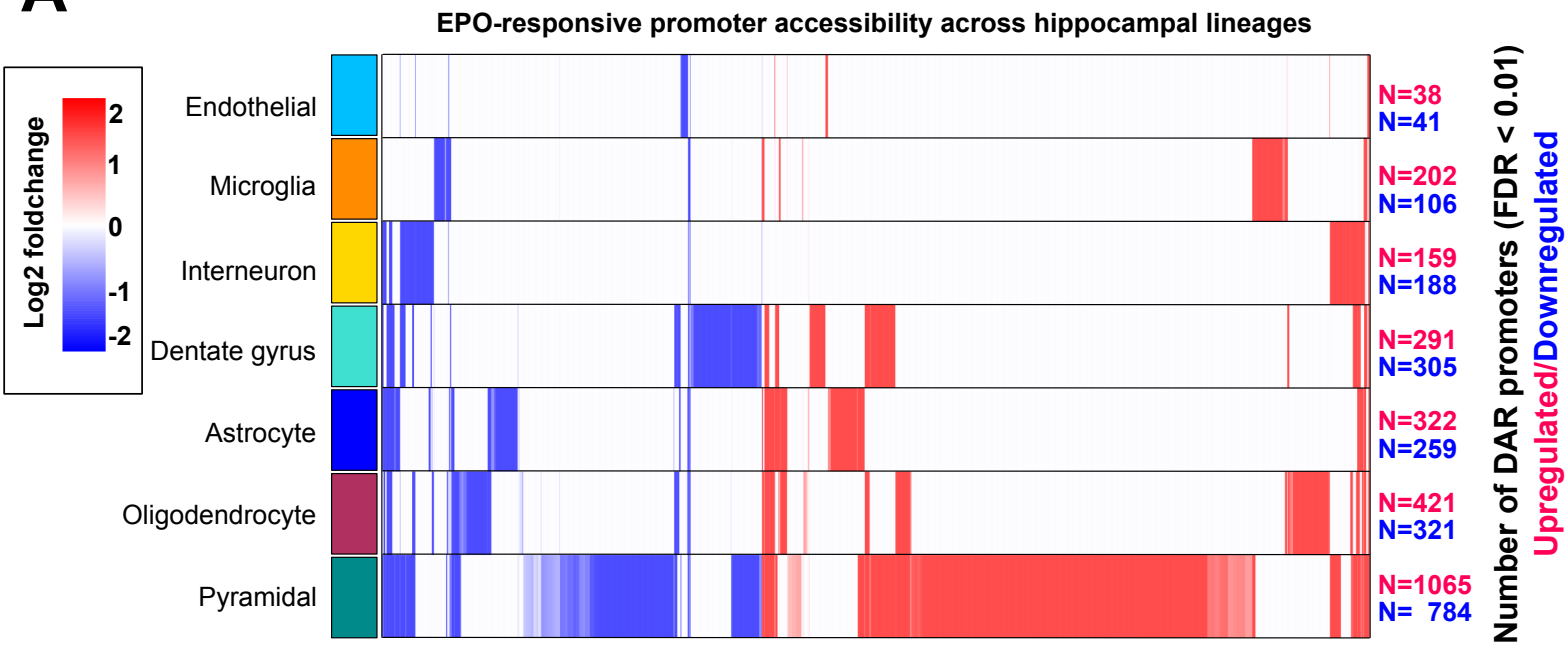

A
Promoter (<=1kb) (10.41%)
Promoter (1-2kb) (3.77%)
Promoter (2-3kb) (3.43%)
5' UTR (0.05%)
3' UTR (1.45%)
1st Exon (0.84%)
Other Exon (1.93%)
1st Intron (13.86%)
Other Intron (30.65%)
Downstream (<=300) (0.08%)
Distal Intergenic (33.54%)
B
Distribution of peaks overlapping with Genes/TEs
Number of overlapping summits
100K
50K
0
112,456
17,728
21,281
74,153
Genes
TEs
Genes + TEs
None
C
Enrichment score (Z-score)
30
20
10
0
RMER1B
MER130
GSAT_MM
RMER19B
RMER1A
IAPEY4_LTR
L2
LTR41
ORR1E
MER20
LTR41B
MER121
MER102b
L3
LTR33
UCON31
L4_C_Mam
LSU-rRNA_Hsa
CR1_Mam
RMER19B2
LTR33B
MLT1H
RMER6A
AmnSINE1
L1MB7
RMER17B
LTR50
MLT1F2
UCON11
RLTR16C_MM
Tigger13a
L2a
Charlie24
MTEb
MER89
UCON29
L1ME1
LTR33A_
MER58A
UCON26
L2c
L1MC1
MLT1G1
MLT1G3
MTEa
MLT1F
MLT1J
UCON28a
L2b
MLTR11B
Top 50 TE families
Zscore
35
30
25
20
15
10

Figure S6

A


EPO-responsive promoter accessibility across hippocampal lineages
Log2 foldchange
2
1
0
-1
-2
Endothelial
Microglia
Interneuron
Dentate gyrus
Astrocyte
Oligodendrocyte
Pyramidal
N=38
N=41
N=202
N=106
N=159
N=188
N=291
N=305
N=322
N=259
N=421
N=321
N=1065
N= 784
Number of DAR promoters (FDR < 0.01)
Upregulated/Downregulated

B

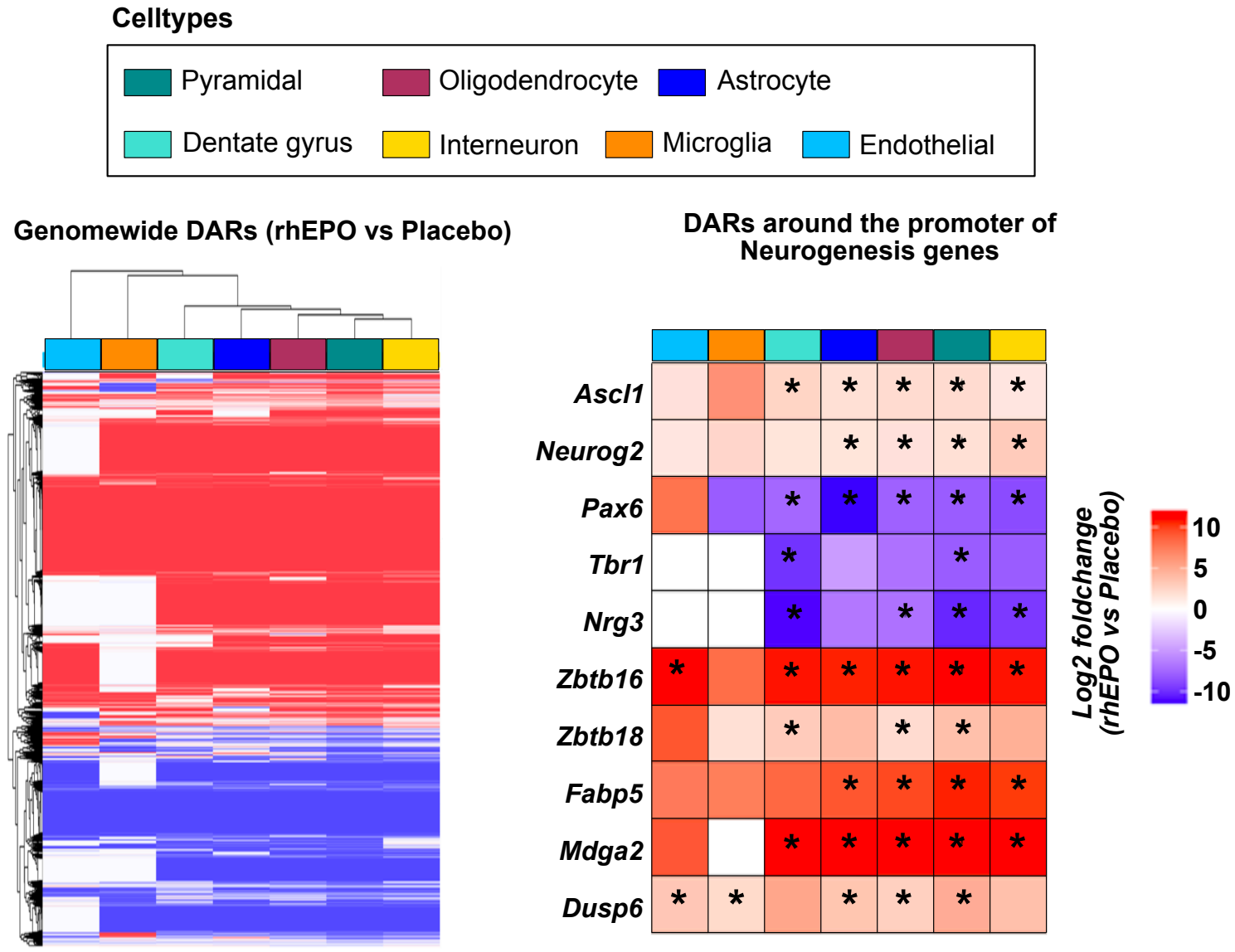

Celltypes
Pyramidal
Oligodendrocyte
Astrocyte
Dentate gyrus
Interneuron
Microglia
Endothelial
Genomewide DARs (rhEPO vs Placebo)
DARs around the promoter of Neurogenesis genes
Ascl1
Neurog2
Pax6
Tbr1
Nrg3
Zbtb16
Zbtb18
Fabp5
Mdga2
Dusp6
Log2 foldchange (rhEPO vs Placebo)
10
5
0
-5
-10

Figure S7

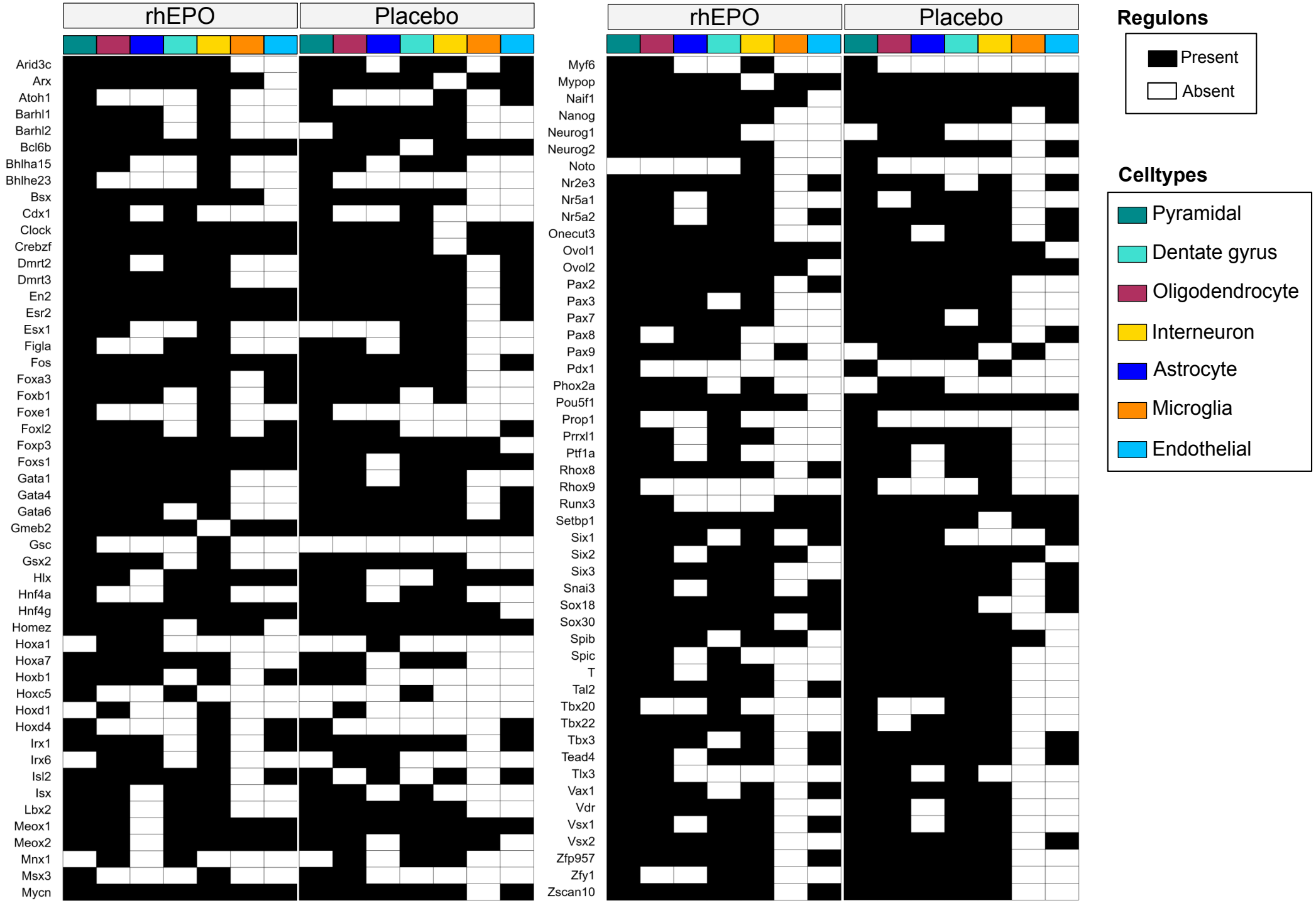

rhEPO
Placebo
Arid3c
Arx
Atoh1
Barhl1
Barhl2
Bcl6b
Bhlha15
Bhlhe23
Bsx
Cdx1
Clock
Crebzf
Dmrt2
Dmrt3
En2
Esr2
Esx1
Figla
Fos
Foxa3
Foxb1
Foxe1
Foxl2
Foxp3
Foxs1
Gata1
Gata4
Gata6
Gmeb2
Gsc
Gsx2
Hlx
Hnf4a
Hnf4g
Homez
Hoxa1
Hoxa7
Hoxb1
Hoxc5
Hoxd1
Hoxd4
Irx1
Irx6
Isl2
Isx
Lbx2
Meox1
Meox2
Mnx1
Msx3
Mycn
rhEPO
Placebo
Myf6
Mypop
Naif1
Nanog
Neurog1
Neurog2
Noto
Nr2e3
Nr5a1
Nr5a2
Onecut3
Ovol1
Ovol2
Pax2
Pax3
Pax7
Pax8
Pax9
Pdx1
Phox2a
Pou5f1
Prop1
Prrxl1
Ptf1a
Rhox8
Rhox9
Runx3
Setbp1
Six1
Six2
Six3
Snai3
Sox18
Sox30
Spib
Spic
T
Tal2
Tbx20
Tbx22
Tbx3
Tead4
Tlx3
Vax1
Vdr
Vsx1
Vsx2
Zfp957
Zfy1
Zscan10
Regulons
Present
Absent
Celltypes
Pyramidal
Dentate gyrus
Oligodendrocyte
Interneuron
Astrocyte
Microglia
Endothelial

Figure S8

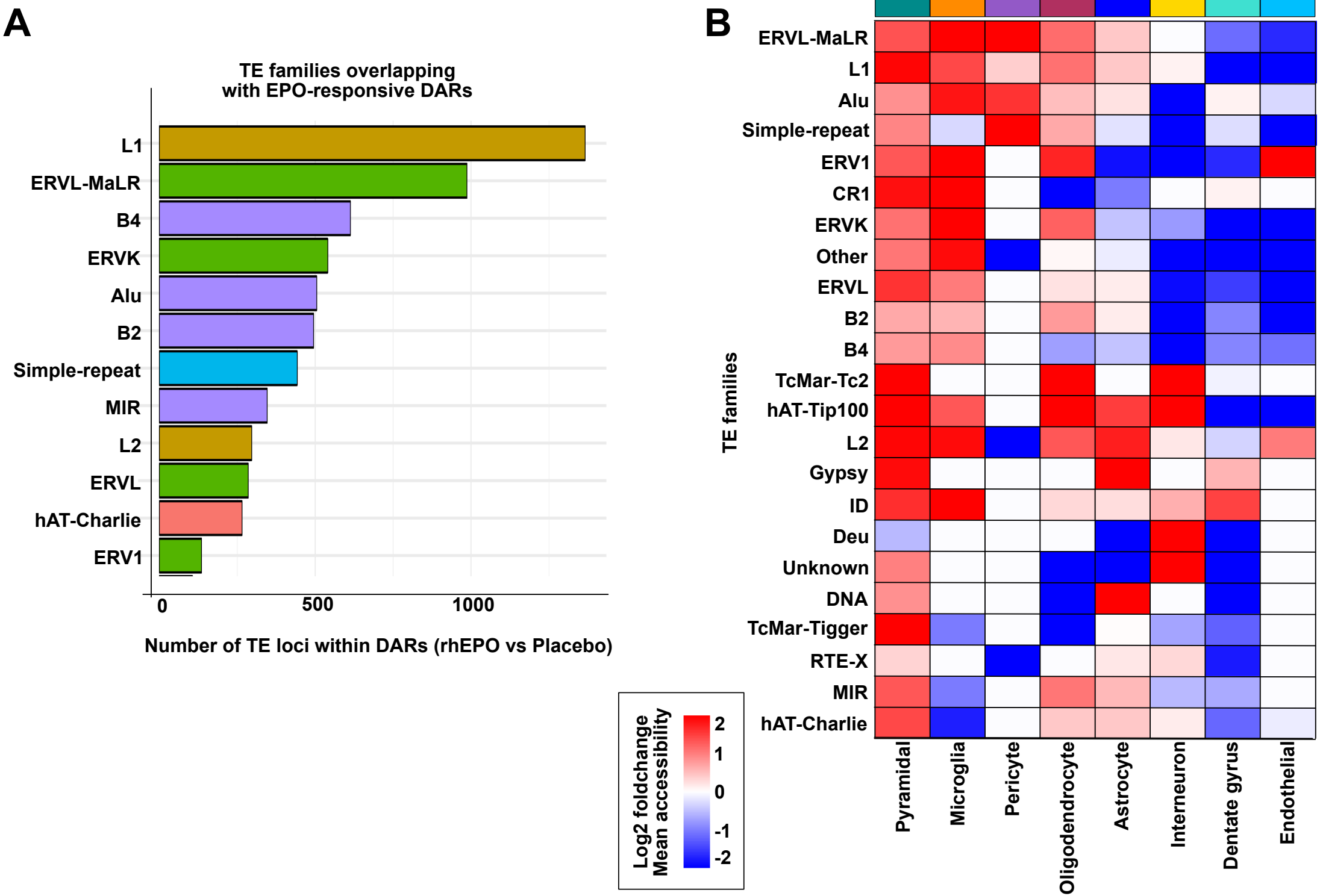
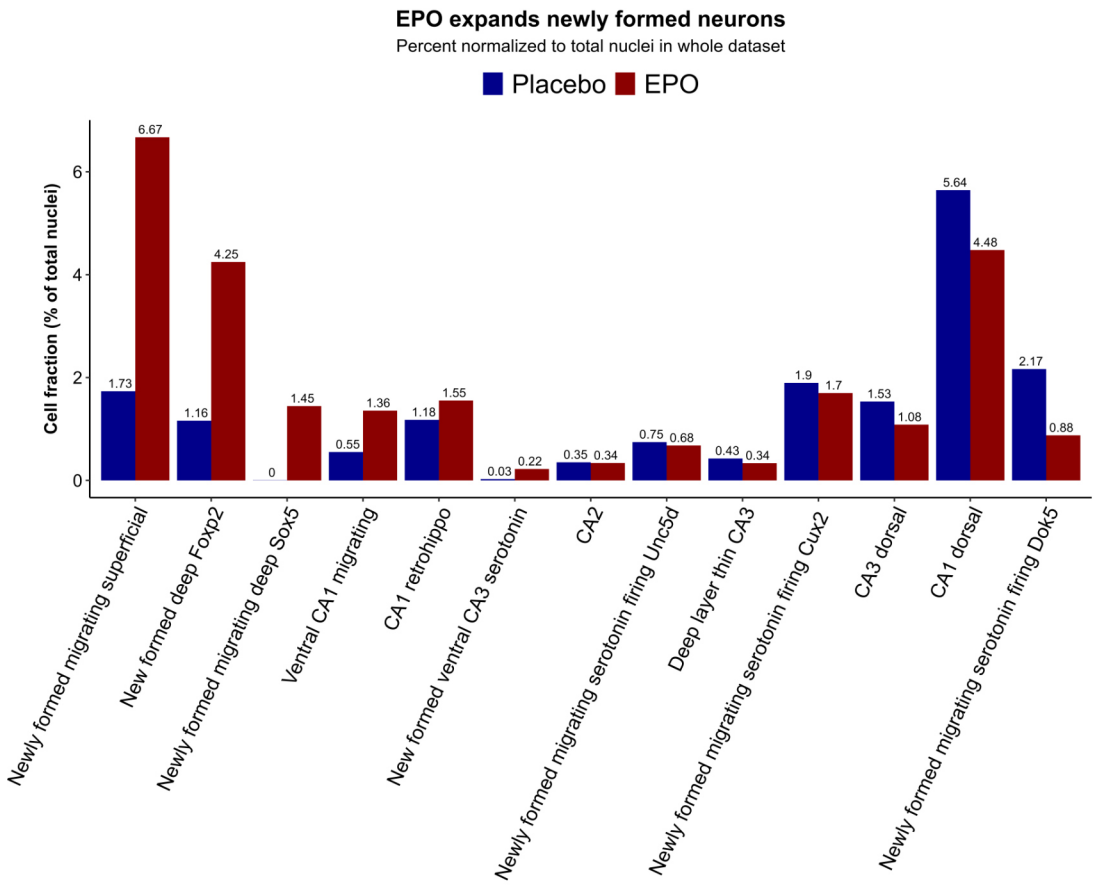

A
TE families overlapping
with EPO-responsive DARs
L1
ERVL-MaLR
B4
ERVK
Alu
B2
Simple-repeat
MIR
L2
ERVL
hAT-Charlie
ERV1
0
500
1000
Number of TE loci within DARs (rhEPO vs Placebo)
B
TE families
ERVL-MaLR
L1
Alu
Simple-repeat
ERV1
CR1
ERVK
Other
ERVL
B2
B4
TcMar-Tc2
hAT-Tip100
L2
Gypsy
ID
Deu
Unknown
DNA
TcMar-Tigger
RTE-X
MIR
hAT-Charlie
Pyramidal
Microglia
Pericyte
Oligodendrocyte
Astrocyte
Interneuron
Dentate gyrus
Endothelial
Log2 foldchange
Mean accessibility
2
1
0
-1
-2

Figure S9

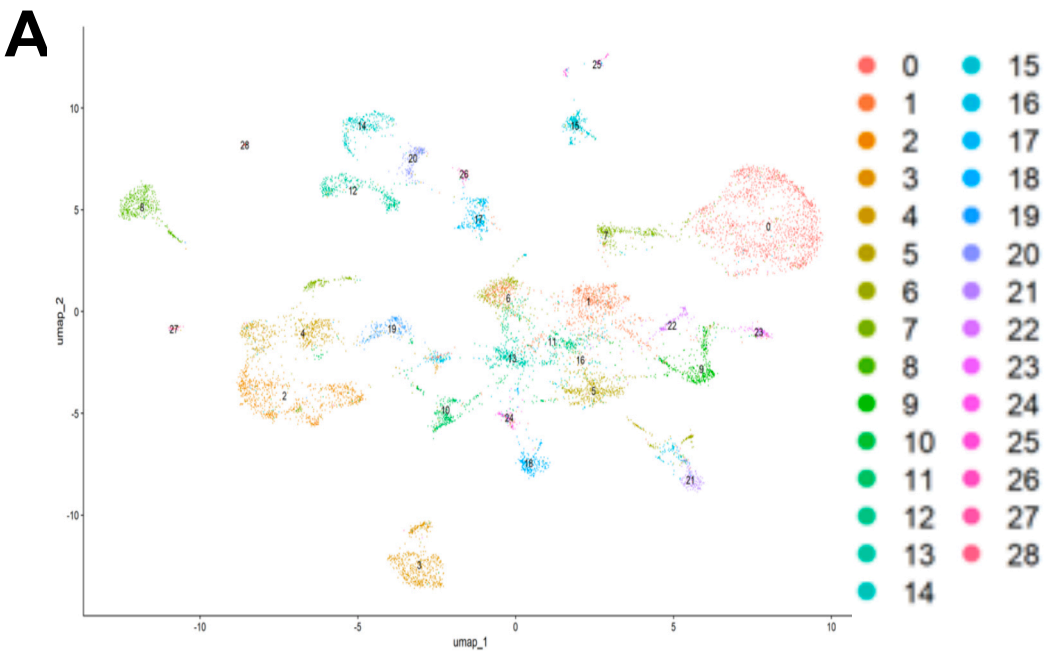
A

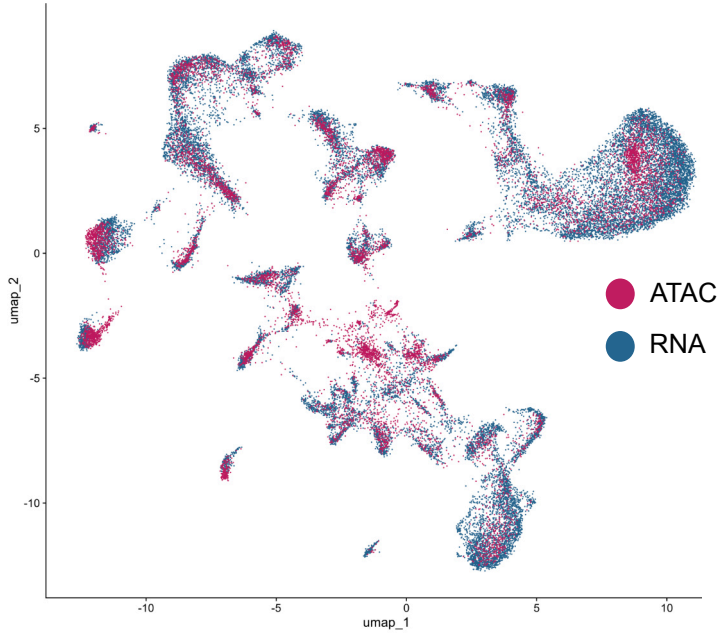


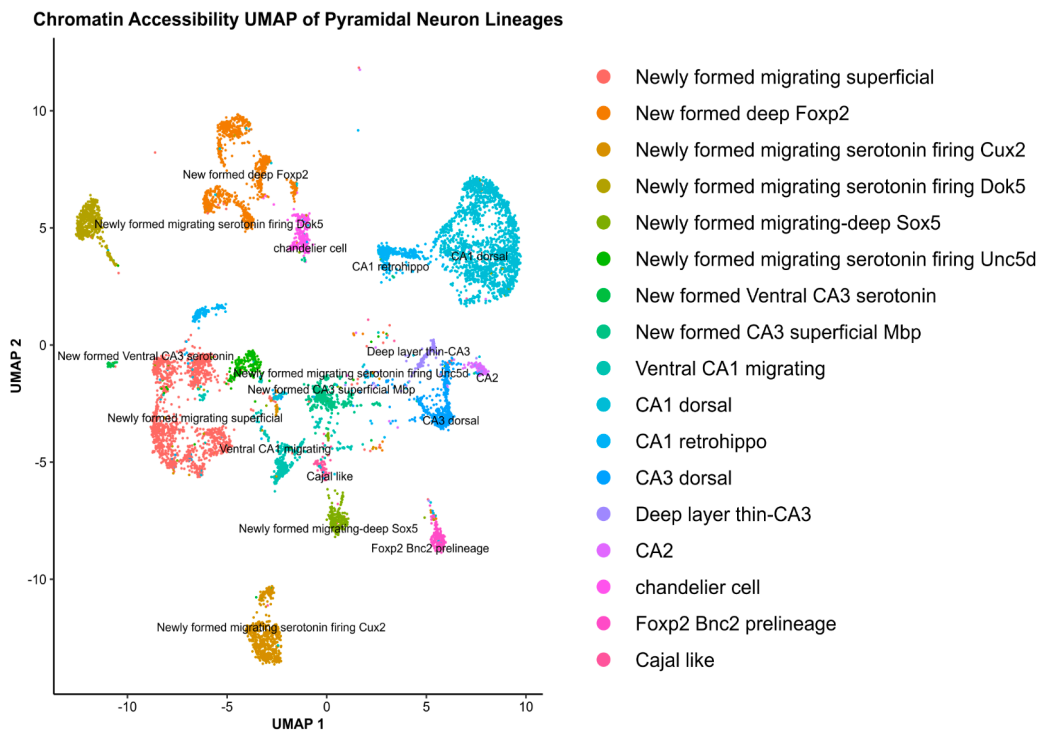
B

Cluster	Placebo	EPO
Newly formed migrating superficial	195	1079
New formed deep Foxp2	153	691
Newly formed migrating deep Sox5	1	253
Ventral CA1 migrating	88	246
CA1 retrohippo	198	273
New formed ventral CA3 serotonin	2	33
CA2	45	52
Newly formed migrating serotonin firing Unc5d	105	119
Deep layer thin CA3	69	54
Newly formed migrating serotonin firing Cux2	273	302
CA3 dorsal	255	191
CA1 dorsal	810	790
Newly formed migrating serotonin firing Dok5	359	110

Figure S10


A
umap_1
umap_2
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28


B
umap_1
umap_2
ATAC
RNA


C
Chromatin Accessibility UMAP of Pyramidal Neuron Lineages
UMAP 1
UMAP 2
New formed deep Foxp2
Newly formed migrating serotonin firing Dok5
chandelier cell
CA1 retrohippo
CA1 dorsal
New formed Ventral CA3 serotonin
Deep layer thin-CA3
Newly formed migrating serotonin firing Unc5d
CA2
New formed CA3 superficial Mbp
Newly formed migrating superficial
CA3 dorsal
Ventral CA1 migrating
Cajal like
Newly formed migrating-deep Sox5
Foxp2 Bnc2 prelineage
Newly formed migrating serotonin firing Cux2
Newly formed migrating superficial
New formed deep Foxp2
Newly formed migrating serotonin firing Cux2
Newly formed migrating serotonin firing Dok5
Newly formed migrating-deep Sox5
Newly formed migrating serotonin firing Unc5d
New formed Ventral CA3 serotonin
New formed CA3 superficial Mbp
Ventral CA1 migrating
CA1 dorsal
CA1 retrohippo
CA3 dorsal
Deep layer thin-CA3
CA2
chandelier cell
Foxp2 Bnc2 prelineage
Cajal like

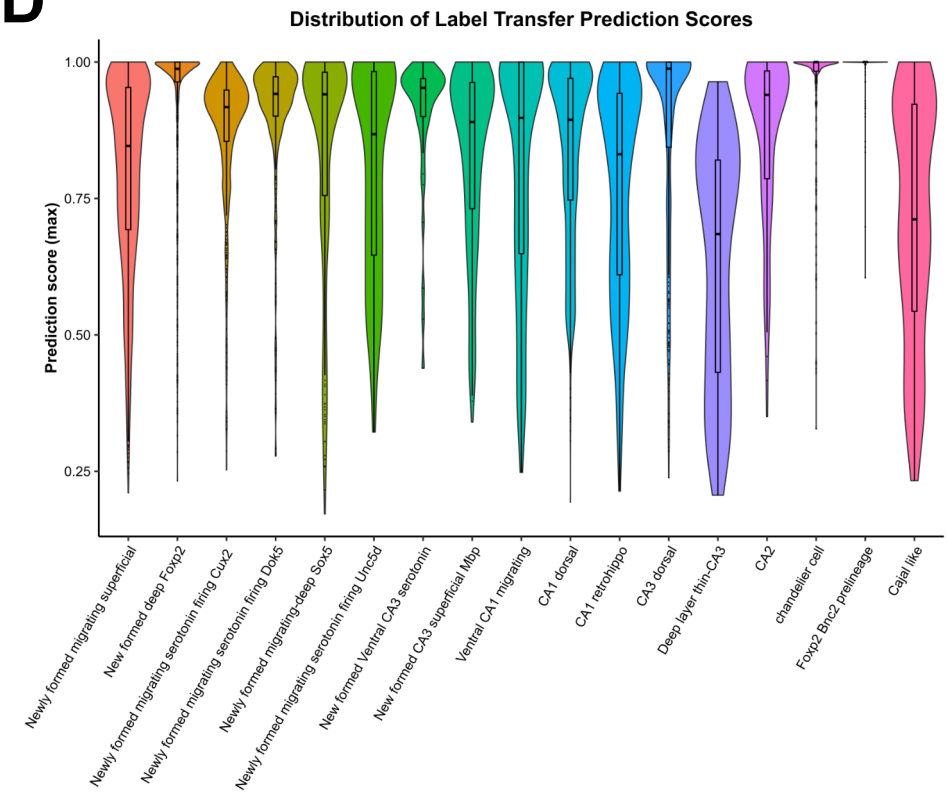

D
Distribution of Label Transfer Prediction Scores
Prediction score (max)
1.00
0.75
0.50
0.25
Newly formed migrating superficial
New formed deep Foxp2
Newly formed migrating serotonin firing Cux2
Newly formed migrating serotonin firing Dok5
Newly formed migrating-deep Sox5
Newly formed migrating serotonin firing Unc5d
New formed Ventral CA3 serotonin
New formed CA3 superficial Mbp
Ventral CA1 migrating
CA1 dorsal
CA1 retrohippo
CA3 dorsal
Deep layer thin-CA3
CA2
chandelier cell
Foxp2 Bnc2 prelineage
Cajal like

Figure S11

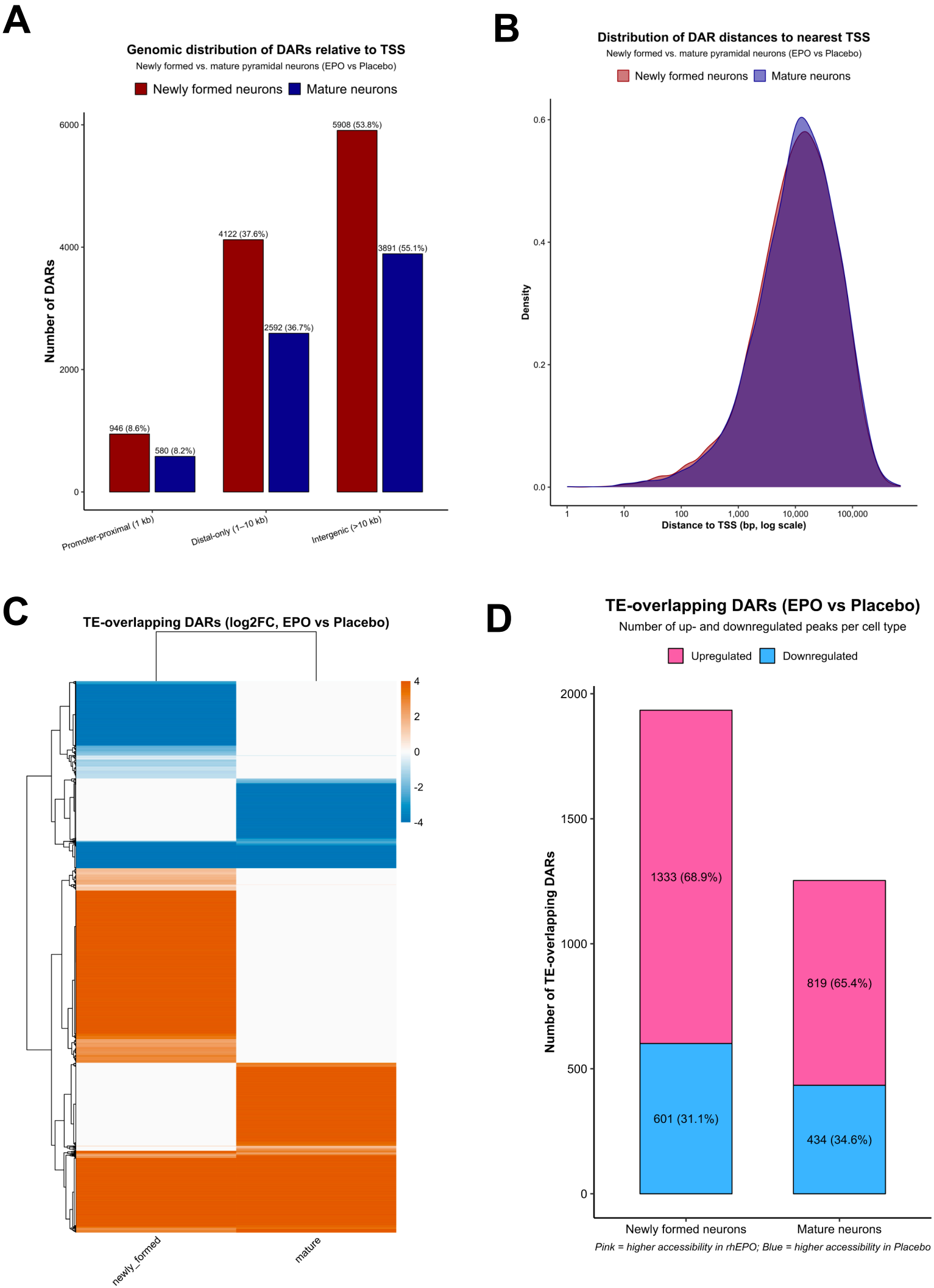

A
Genomic distribution of DARs relative to TSS
Newly formed vs. mature pyramidal neurons (EPO vs Placebo)
Newly formed neurons
Mature neurons
Number of DARs
6000
4000
2000
0
946 (8.6%)
580 (8.2%)
4122 (37.6%)
2592 (36.7%)
5908 (53.8%)
3891 (55.1%)
Promoter-proximal (1 kb)
Distal-only (1–10 kb)
Intergenic (>10 kb)
B
Distribution of DAR distances to nearest TSS
Newly formed vs. mature pyramidal neurons (EPO vs Placebo)
Newly formed neurons
Mature neurons
Density
0.6
0.4
0.2
0.0
1
10
100
1,000
10,000
100,000
Distance to TSS (bp, log scale)
C
TE-overlapping DARs (log2FC, EPO vs Placebo)
4
2
0
-2
-4
newly_formed
mature
D
TE-overlapping DARs (EPO vs Placebo)
Number of up- and downregulated peaks per cell type
Upregulated
Downregulated
Number of TE-overlapping DARs
2000
1500
1000
500
0
1333 (68.9%)
601 (31.1%)
819 (65.4%)
434 (34.6%)
Newly formed neurons
Mature neurons
Pink = higher accessibility in rhEPO; Blue = higher accessibility in Placebo

Figure S12

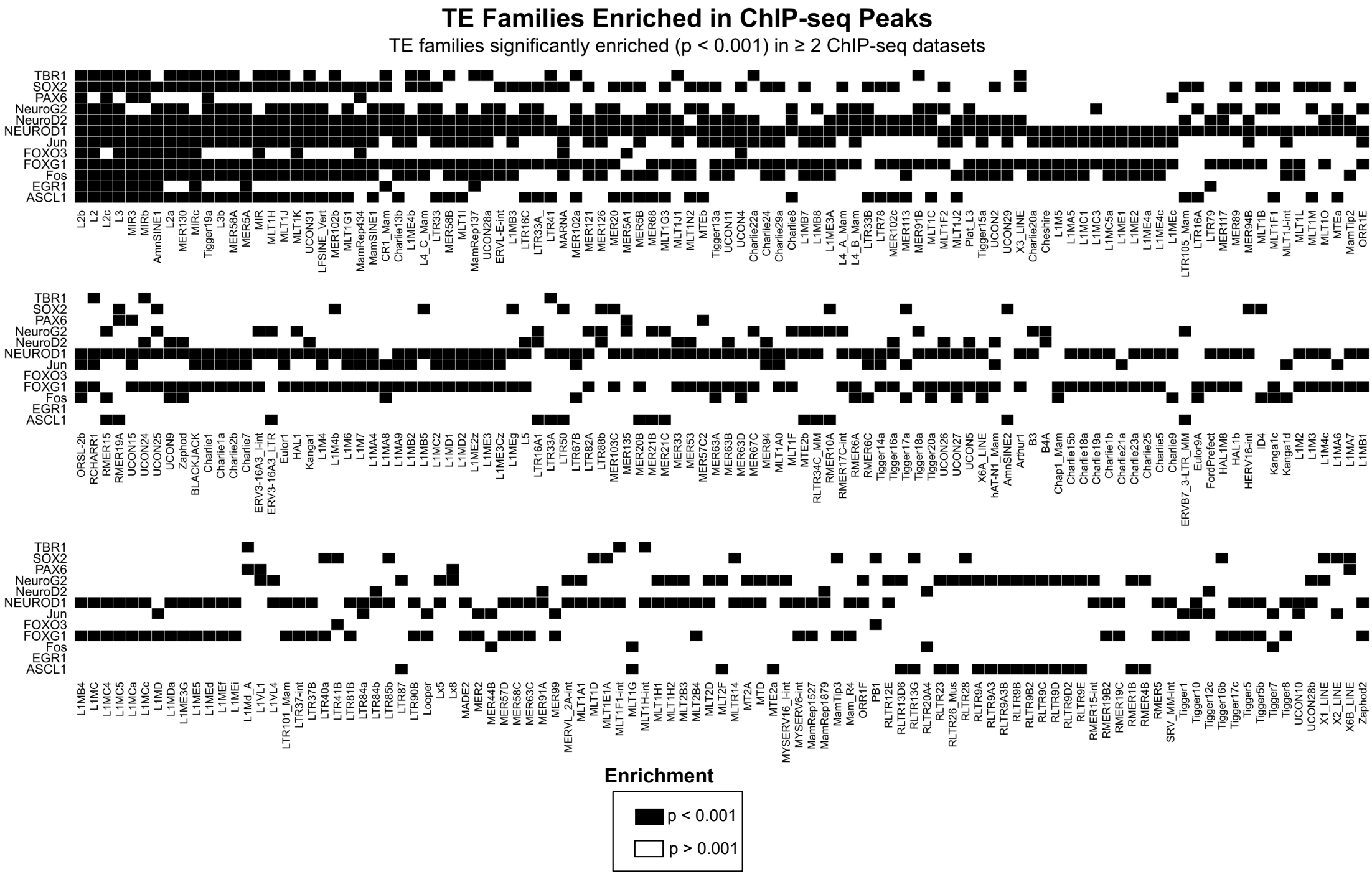

TE Families Enriched in ChIP-seq Peaks
TE families significantly enriched (p < 0.001) in ≥ 2 ChIP-seq datasets
TBR1
SOX2
PAX6
NeuroG2
NeuroD2
NEUROD1
Jun
FOXO3
FOXG1
Fos
EGR1
ASCL1
L2b L2 L2c L3 MIR3 MIRb AmnSINE1 L2a MER130 MIRc Tigger19a L3b MER58A MER5A MIR MLT1H MLT1J MLT1K UCON31 LFSINE_Vert MER102b MLT1G1 MamRep434 MamSINE1 CR1_Mam Charlie13b L1ME4b L4_C_Mam LTR33 MER58B MLT1I MamRep137 UCON28a ERVL-E-int L1MB3 LTR16C LTR33A_ LTR41 MARNA MER102a MER121 MER126 MER20 MER5A1 MER5B MER68 MLT1G3 MLT1J1 MLT1N2 MTEb Tigger13a UCON11 UCON4 Charlie22a Charlie24 Charlie29a Charlie8 L1MB7 L1MB8 L1ME3A L4_A_Mam L4_B_Mam LTR33B LTR78 MER102c MER113 MER91B MLT1C MLT1F2 MLT1J2 Plat_L3 Tigger15a UCON2 UCON29 X3_LINE Charlie20a Cheshire L1M5 L1MA5 L1MC1 L1MC3 L1MC5a L1ME1 L1ME2 L1ME4a L1ME4c L1MEc LTR105_Mam LTR16A LTR79 MER117 MER89 MER94B MLT1B MLT1F1 MLT1J-int MLT1L MLT1M MLT1O MTEa MamTip2 ORR1E
ORSL-2b RCHARR1 RMER15 RMER19A UCON15 UCON24 UCON25 UCON9 Zaphod BLACKJACK Charlie1 Charlie1a Charlie2b Charlie7 ERV3-16A3_I-int ERV3-16A3_LTR Eulor1 HAL1 Kanga1 L1M4 L1M4b L1M6 L1M7 L1MA4 L1MA8 L1MA9 L1MB2 L1MB5 L1MC2 L1MD1 L1MD2 L1ME2z L1ME3 L1ME3Cz L1MEg L5 LTR16A1 LTR33A LTR50 LTR67B LTR82A LTR88b MER103C MER135 MER20B MER21B MER21C MER33 MER53 MER57C2 MER63A MER63B MER63D MER67C MER94 MLT1A0 MLT1F MTE2b RLTR34C_MM RMER10A RMER17C-int RMER6A RMER6C Tigger14a Tigger16a Tigger17a Tigger18a Tigger20a UCON26 UCON27 UCON5 X6A_LINE hAT-N1_Mam AmnSINE2 Arthur1 B3 B4A Chap1_Mam Charlie15b Charlie18a Charlie19a Charlie1b Charlie21a Charlie23a Charlie25 Charlie5 Charlie9 ERVB7_3-LTR_MM Eulor9A FordPrefect HAL1M8 HAL1b HERV16-int ID4 Kanga1c Kanga1d L1M2 L1M3 L1M4c L1MA6 L1MA7 L1MB1
L1MB4 L1MC L1MC4 L1MC5 L1MCa L1MCc L1MD L1MDa L1ME3G L1ME5 L1MEd L1MEf L1MEi L1Md_A L1VL1 L1VL4 LTR101_Mam LTR37-int LTR37B LTR40a LTR41B LTR81B LTR84a LTR84b LTR85b LTR87 LTR90B Looper Lx5 Lx8 MADE2 MER2 MER44B MER57D MER58C MER63C MER91A MER99 MERVL_2A-int MLT1A1 MLT1D MLT1E1A MLT1F1-int MLT1G MLT1H-int MLT1H1 MLT1H2 MLT2B3 MLT2B4 MLT2D MLT2F MLTR14 MT2A MTD MTE2a MYSERV16_I-int MYSERV6-int MamRep1527 MamRep1879 MamTip3 Mam_R4 ORR1F PB1 RLTR12E RLTR13D6 RLTR13G RLTR20A4 RLTR23 RLTR26_Mus RLTR28 RLTR9A RLTR9A3 RLTR9A3B RLTR9B RLTR9B2 RLTR9C RLTR9D RLTR9D2 RLTR9E RMER15-int RMER19B2 RMER19C RMER1B RMER4B RMER5 SRV_MM-int Tigger1 Tigger10 Tigger12c Tigger16b Tigger17c Tigger5 Tigger5b Tigger7 Tigger8 UCON10 UCON28b X1_LINE X2_LINE X6B_LINE Zaphod2
Enrichment
p < 0.001
p > 0.001

Figure S13

GO enrichment for genes near TE-DARs (Top 30 terms)
Biological Process
regulation of metal ion transport
regulation of cation transmembrane transport
synapse organization
small GTPase mediated signal transduction
calcium ion transport
positive regulation of ion transport
regulation of neurogenesis
regulation of membrane potential
regulation of GTPase activity
muscle system process
cell junction assembly
regulation of transporter activity
regulation of transmembrane transporter activity
regulation of ion transmembrane transporter activity
regulation of cell morphogenesis
protein localization to cell periphery
positive regulation of cell projection organization
calcium ion transmembrane transport
potassium ion transport
positive regulation of GTPase activity
peptidyl-serine phosphorylation
peptidyl-serine modification
muscle contraction
hormone transport
hormone secretion
regulation of calcium ion transport
dendrite development
cellular carbohydrate metabolic process
amide transport
negative regulation of ion transport
2% 3% 4% 5% 6%
Cellular Component
synaptic membrane
postsynaptic specialization
neuron to neuron synapse
postsynaptic density
asymmetric synapse
cell leading edge
distal axon
postsynaptic membrane
cell cortex
membrane raft
membrane microdomain
transport vesicle
secretory granule
exocytic vesicle
basal part of cell
intrinsic component of organelle membrane
extrinsic component of membrane
transporter complex
transmembrane transporter complex
synaptic vesicle
presynaptic membrane
integral component of organelle membrane
spindle
ion channel complex
basolateral plasma membrane
basal plasma membrane
intrinsic component of synaptic membrane
focal adhesion
cell-substrate junction
cation channel complex
2% 3% 4% 5% 6%
Gene ratio
Gene count
10
20
30
− log10 q
6
4
2