Terminal selectors organize postmitotic chromatin accessibility for acquisition of serotonergic identity

Xinrui L. Zhang[#1], W. Clay Spencer[#1], Nobuko Tabuchi[1], Evan S. Deneris[1*]

[1]Department of Neurosciences, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, United States

[#]Contributed equally

*Correspondence to esd@case.edu

SUMMARY

Assembly of transcriptomes encoding unique neuronal identities requires selective accessibility of regulatory inputs to cis-regulatory sequences in nucleosome-embedded chromatin. Yet the mechanisms involved in shaping postmitotic neuronal chromatin are poorly understood. Here we used ATAC-seq, ChIPmentation, and single-cell analyses to show that unique distal enhancers and super-enhancers define the *Pet1* neuron lineage that generates serotonin (5-HT) neurons. Heterogeneous single cell chromatin landscapes are established early in postmitotic *Pet1* neurons and reveal the regulatory programs driving *Pet1* neuron subtype identities. Terminal selectors, Pet1 and Lmx1b, control chromatin accessibility in *Pet1* neurons to select enhancers for 5-HT neurotransmission and synaptogenesis. In addition, these factors are required to maintain chromatin accessibility during early maturation suggesting that postmitotic open chromatin is unstable and requires continuous terminal selector input. Together our findings reveal a previously unrecognized function of terminal selectors in organizing postmitotic accessible chromatin for the development of specialized neuronal identities.

INTRODUCTION

Diverse neuronal identities are acquired through the differential expression of genes that specify transmitter usage, build dendritic and axonal connectivity, and control synapse formation [1]. The acquisition of specialized neuronal identities is an extended process that plays out over several weeks of mid fetal to early postnatal life in mice and through the first few years of life in humans [2,3], as postmitotic neuronal transcriptomes mature to a stable state. Unique combinations of sequence-specific transcription factors selectively activate and repress genes to assemble the transcriptomes that encode individual neuron identities [4]. Genes with variants associated with neurodevelopmental disorders such as autism exhibit enriched expression in maturing neurons and are often involved in chromatin remodeling and transcriptional regulation; their developmental dysfunction is thought to contribute to disease pathogenesis [5,6].

Selective expression of neuronal identity genes depends on selective accessibility of promoter and enhancer sequences embedded in postmitotic nucleosomal-organized chromatin to regulatory factor inputs [7,8]. Chromatin accessibility is dynamic as brain regions [9,10] and neurons [11–14] mature. Yet, we have a poor understanding of how the dynamics of accessible chromatin in postmitotic neurons prefigures acquisition of neurotransmitter identities and myriad other molecular and morphological features of specialized neurons. Further, the regulatory factors involved in selecting postmitotic accessible chromatin regions to enable sequence-specific activation of unique combinations of identity features have not been identified.

Control of serotonin (5-HT) neuron identity is of particular interest as 5-HT has expansive modulatory effects on central neural circuitry [15]. Altered serotonergic gene expression, brought about by either genetic or environmental factors, has been implicated in neuropsychiatric disorders including depression, stress related anxiety disorders, autism, obsessive compulsive disorder, and schizophrenia [16]. Postmitotic 5-

HT neuron precursors terminally differentiate in the ventral hindbrain to acquire the capacity for 5-HT synthesis, reuptake, vesicular transport, autoreceptor regulation, and degradation through the coordinate expression of the 5-HT neurotransmission genes, *Tph2*, *Ddc*, *Gch1*, *Slc6a4* and *Slc22a3*, *Slc18a2*, and *Maoa/b* [2]. Subsequently, newborn 5-HT neurons enter a maturation stage during which serotonergic transcriptomes are highly dynamic as 5-HT neurons refine their terminal identities, migrate to populate the various raphe nuclei, acquire mature firing properties, and establish expansive connectivity throughout the brain and spinal cord [2].

The ETS domain factor Pet1 and the LIM HD factor Lmx1b are terminal selector transcription factors in 5-HT neurons as they are continuously expressed beginning at the postmitotic precursor stage, control their own expression through direct positive autoregulation, and are required for the acquisition of 5-HT neuron terminal identity through sequence-specific transcriptional activation of 5-HT pathway genes [17]. Moreover, these terminal selectors are required for the postmitotic maturation of other specialized features including serotonergic firing characteristics and formation of long-distance profusely arborized 5-HT axon architectures throughout the brain and spinal cord [18–20]. Pet1's terminal selector function in 5-HT neurons is clinically significant as recent studies identified biallelic loss of function of *FEV*, the human ortholog of *Pet1*, in autism cases suggesting that defects in transcriptional control of *Pet1* lineage neurons are a potential path to human neurodevelopmental disorders [21].

The low abundance of many postmitotic neuron-types has hindered the assay of their accessible chromatin landscapes and cis-regulatory architectures. Here, we adapted bulk and single cell ATAC-seq, scRNA-seq, and ChIPmentation protocols to map accessible chromatin and cis-regulatory elements of developing 5HT neurons. We then used these combined maps along with Pet1 and Lmx1b occupancy mapping to identify a critical role for Pet1 and Lmx1b in directly shaping chromatin accessibility of 5-

HT neurotransmission and synapse genes. Our findings show that terminal selectors control specialized neuronal identities not only through sequence-specific activation of select terminal effector genes but also by shaping postmitotic accessible chromatin at cis-regulatory elements.

RESULTS

The unique accessible chromatin landscape of maturing postmitotic *Pet1* neurons

5-HT neurons are generated from *Pet1*+ postmitotic precursors located in the embryonic ventral hindbrain from the isthmus to the caudal region of the myelencephalon [22]. Therefore we used the *Pet1::EYFP* transgenic mouse line to flow sort E14.5 Yfp+ neurons (*Pet1* neurons) from the entire longitudinal extent of the hindbrain for assay of accessible chromatin in maturing 5-HT neurons [23,24] (Supplementary Fig 1a-c). We generated three ATAC-seq biological replicates with a mean of 29M unique paired-end mapped reads per sample (Supplementary Fig 1d-g). We identified 68,871 peaks of transposase accessible chromatin (TAC) at E14.5, and 59,323 TACs in an equal number of control Yfp- (*Pet1*^{neg}) cells collected from the same hindbrain regions at the same developmental age, which comprise diverse non-serotonergic cell types. Principal component analysis (PCA) showed that cell type identity accounted for 93% of the variance when comparing *Pet1* neuron ATAC-seq replicates with *Pet1*^{neg} ATAC-seq samples and with previously published ATAC-seq datasets obtained from bulk tissue dissections (Fig 1a). The accessible chromatin landscape of *Pet1* neurons is more similar to that of bulk brain tissues than that of non-neuronal tissues such as the liver (Fig 1a). However, the accessible chromatin landscape of *Pet1* neurons is also distinct from previously published bulk hindbrain ATAC-seq[9], which indicates that bulk analyses did not capture the chromatin accessibility profiles of low abundance *Pet1* neurons (Fig 1a). Indeed, of the TACs we identified in *Pet1* neurons at E14.5, 31% are unique and not found in *Pet1*^{neg} cells of mouse hindbrain and 22% are not previously identified in any other cell type[9] (Fig 1b).

Serotonergic cis-regulatory elements

We next used ChIPmentation [25] to profile the genome-wide distribution of histone posttranslational modifications associated with transcriptional activation or repression in flow sorted E14.5 *Pet1* neurons. With the segmentation algorithm ChromHMM [26], we defined ten chromatin states by assessing the combinatorial pattern of histone marks at each genomic locus (Fig 1c; Supplementary Fig 1h). Both active and poised distal enhancer regions are over-represented among the unique *Pet1* neuron TACs while TSS-proximal promoter regions are more likely shared with other cell types (Fig 1c; Supplementary Fig 1i-j). Our ChIPmentation analysis also identified 1919 genomic regions that fulfill the criteria for super-enhancers based on the presence of unusually expansive H3K27ac signals (Fig 1d) [27,28]. Comparing these to previously annotated enhancers in other cell types [9], we found that a significant percentage (42%) of the super-enhancers we detected in *Pet1* neurons were not detected in bulk mouse forebrain, hindbrain, or liver tissues (Fig 1e), suggesting that they are highly specific to the *Pet1* lineage. As an example, a 82.9kb super-enhancer detected in *Pet1* neurons but not *Pet1*^{neg} neurons spans the *Lmx1b* locus (Fig 1f).

Notably, a distinct TAC was found immediately upstream of the *Pet1* promoter in *Pet1* neurons but not *Pet1*^{neg} cells (Fig 1g). This TAC shows enrichment of H3K4me1 and H3K27ac marks suggesting the presence of an active enhancer (Fig 1g). Indeed, previous transgenic assays, *in vivo*, demonstrated that this upstream region directs robust reporter expression to developing and adult 5-HT neurons present in each of the midbrain, pontine, and medullary raphe nuclei [23,24,29], hence providing functional validation of serotonergic enhancers we have defined with ATAC-seq and ChIPmentation. Further, we found unique TACs in *Pet1* neurons within 20kb of 5-HT neurotransmission genes (Fig 1h-j). The majority of these TACs are enriched in H3K4me1 and H3K27ac marks indicative of enhancers (Fig 1i-j). Together, our ATAC-

seq and ChIPmentation analyses have defined the open chromatin landscape of maturing 5-HT neurons and reveal distal cis-regulatory elements that define the *Pet1* postmitotic lineage.

Heterogeneous distal enhancer accessibility defines *Pet1* neuron subtypes

Adult 5-HT neurons possess diverse transcriptomes that encode functional subtypes [30–33], yet the developmental mechanisms underlying 5-HT neuron heterogeneity are not understood. To determine whether 5-HT neuron heterogeneity is coded early in accessible chromatin, we carried out single cell ATAC-seq (scATAC-seq) with E14.5 flow sorted *Pet1* neurons, which were processed using the 10x Genomics Chromium microfluidic droplet scATAC-seq method. The resulting library was sequenced to a depth sufficient to yield >50,000 paired-end reads per nucleus. After filtering based on standard quality control measures (Supplementary Fig 2a), we retained 1,692 nuclei for analysis. This led to the identification of 124,111 TACs, 60,117 of which were not identified in our bulk *Pet1* neuron ATAC-seq (Fig 2a). We observed high concordance (Spearman's correlation = 0.9) between our pseudo-bulk scATAC-seq dataset and our bulk ATAC-seq data as illustrated by the highly similar transposase accessible chromatin patterns in the *Pet1* upstream enhancer region (Fig 2b).

We grouped single *Pet1* neurons to identify 11 major subtypes with unique chromatin accessibility patterns (Fig 2d; Supplementary Fig 2b). TACs within *Pet1* and *Lmx1b* loci are robustly accessible in virtually all cells within the 11 identified subtypes (Fig 2e). Like *Pet1* and *Lmx1b*, two-thirds (63%) of TACs are broadly accessible among multiple subtypes, including TACs of 5-HT neurotransmission genes *Tph2*, *Slc22a3*, and *Slc6a4* (Fig 2e; Supplementary Fig 2c). In contrast, one-third of TACs (36.6%) are unique to a single cluster. Subtype-specific TACs are predominantly distal gene regulatory elements as previously defined by chromatin state (Fig 1c; Supplementary Fig

2d-e) and are often observed near genes that mark 5-HT neuron subtypes (Supplementary Fig 2b). For example, transcription factor En1 is enriched in the rostral rhombencephalic 5-HT neurons, whereas Hox gene cofactor Meis2 is selectively expressed in the caudal rhombencephalic 5-HT neurons [33]. Six of the scATAC-seq clusters share an unique TAC region immediately upstream of the TSS of the rostral marker gene *En1* while the remaining five clusters have more accessible TAC regions near the promoter of the caudal 5-HT neuron marker gene *Meis2* (Fig 2c-e). Thus, our data demonstrate that diverse single cell chromatin landscapes are established early in postmitotic *Pet1* neurons and likely prefigure transcriptomic and functional heterogeneity of adult 5-HT neurons.

To analyze the target genes of subtype-specific TACs, we performed parallel scRNA-seq using 10x Genomics and E14.5 flow sorted *Pet1* neurons (Supplementary Fig 2f-g). We obtained over 80,000 reads per cell for 1,612 cells, representing a median of 4,010 genes per cell and a total of 19,707 genes across all cells. 95% of the flow sorted *Pet1* neurons robustly expressed both *Pet1* and Yfp, validating that our cell sorting method yields relatively pure populations of *Pet1* neurons. The small numbers of Yfp- cells were removed from further analysis. We next co-embedded scATAC-seq with scRNA-seq data, then linked TACs to a gene if logistic regression detected a statistical association between gene expression and the binarized chromatin accessibility at a genomic region within 500kb [34]. Applying this method to our single cell datasets yielded a total of 41,191 promoter-enhancer pairings (Supplementary Table). Each gene is on average linked to only 1.8 TACs and most genes are associated with either none or a single TAC (Fig 2i). However, 10% of genes are associated with >5 TAC regions, and a select group of genes (0.045%) are associated with >20 TACs (Fig 2h-j; Supplementary Fig 3a). Gene ontology (GO) analysis of the top 1,500 genes with the highest number of linked cis-regulatory elements revealed that they are highly enriched for developmental

genes that define neuronal lineage identity (Supplementary Fig 3b), suggesting that a large number of cis-regulatory elements converge to fine-tune or safeguard the transcription of key neuronal master regulators [13,35]. Among these genes are *Pet1* and *Lmx1b* (Fig 2j).

We identified many subtype-specific TFs constituting a potential transcription factor code that underlies 5-HT neuron heterogeneity, such as the enrichment of Nkx6-1 activity, gene expression, and chromatin accessibility in the Nkx6-1/Ndnf cluster (Supplementary Fig 3c-d). TFs expressed in a *Pet1* neuron subtype have gene regions that are also more accessible in that subtype (Supplementary Fig 3d). Overall most subtype-specific TACs are correlated with the subtype-specific expression pattern of their linked genes; however, we noticed that the expression of 5-HT identity genes *Tph2*, *Slc6a4*, and *Slc22a3* are low or undetectable in most *Pet1* neurons at E14.5 despite their relatively uniform accessible chromatin in *Pet1* neurons at this stage (Figs 2e and 2g). The vast majority of *Pet1*-lineage neurons robustly express these 5-HT identity genes in the adult mouse brain (Supplementary Fig 3e) [18,30,32,36]. Immunofluorescence analysis of Tph2 protein levels showed that Tph2+ neurons emerge in the highest fractions in *Pet1* neurons of the dorsal raphe where 5-HT neurons are first born and have progressed further in their maturation than those in other raphe nuclei (Supplementary Fig 3f). This suggests that chromatin accessibility may be remodeled early during maturation to prime neuron identity before the transcription of terminal effector genes.

Identification of a highly dynamic stage of postmitotic accessible chromatin maturation

To investigate the dynamics of chromatin accessibility in maturing *Pet1* neurons, we compared TACs at E11.5 when postmitotic 5-HT neuron precursors are abundant,

E14.5 when *Pet1* neurons are actively extending dendrites and axons and continuing to acquire 5-HT identities, and E17.5 when *Pet1* neurons are coalescing to form the raphe nuclei and innervating distal target brain regions [18] (Supplementary Fig 4a-b). PCA showed high concordance of ATAC-seq replicates within each time point with developmental state accounting for the majority of variance (Fig 3a). Significant gains and losses in chromatin accessibility of TSS-distal TACs occurred throughout embryonic maturation (Fig 3b-e; Supplementary Fig 4c-d). Of the TACs detected in *Pet1* neurons, 16,493 “open” (gain accessibility; fold change>2 and FDR <0.01) between E11.5 and E14.5, while 4,148 open from E14.5 to E17.5 (Fig 3c). Concurrently, 9,402 E11.5 TACs “close” (lose accessibility, fold change>2 and FDR<0.01) by E14.5, and 1504 TACs close between E14.5 and E17.5 (Fig 3c). This suggests that chromatin remodeling is highly dynamic immediately following cell cycle exit and then gradually stabilizes as *Pet1* neurons acquire 5-HT identity. Assessment of chromatin states showed that active enhancers are over-represented among TACs opening between E11.5 and E17.5 (Fig 3b; Supplementary Fig 4e-g). *Pet1* neuron super-enhancer accessibility is particularly highly dynamic during postmitotic neuron maturation, with the vast majority (93%) of TACS either gaining or losing accessibility between E11.5 and E17.5 (Supplementary Fig 4h).

Next we compared our bulk ATAC-seq datasets at E11.5, E14.5, and E17.5 with our previously published bulk 5-HT neuron RNA-seq data obtained at E11.5, E15.5, and postnatal day 2-3 [20]. Using our TAC-to-gene assignments (Fig 2i-j), we found that changes in chromatin accessibility at cis-regulatory regions are significantly correlated with the expression trajectories of their target genes (Fig 3f-g). GO analysis revealed that genes linked to closing TACs are associated with embryonic neural precursor functions (Fig 3h). Conversely, genes linked to opening TACs are strongly enriched for synaptic organization and axon development (Fig 3h). Dynamic TACs are highly

enriched for diverse TF binding motifs, suggesting that combinatorial TF activities contribute to *Pet1* neuron maturation (Supplementary Fig 4i). Together these findings support a model in which the dynamic remodeling of serotonergic cis-regulatory element accessibility drives gene expression trajectories required for maturation of 5-HT identities while shutting down expression of genes required at earlier stages of development.

Pet1 controls chromatin accessibility of 5-HT neurotransmission genes

The highly dynamic remodeling of TACs in maturing 5-HT neurons predicts the existence of regulatory factors that shape postmitotic chromatin landscapes. However, such factors have not been described. Notably, TACs that are unique to *Pet1*-lineage neurons are enriched for ETS-factor binding sites (Supplementary Fig 5a). Thus, we hypothesized that Pet1 controls acquisition of 5-HT identity by controlling chromatin accessibility. *Pet1-/-* 5-HT neurons are an ideal model with which to examine terminal selector function in shaping the chromatin, as these neurons are retained in the brain in normal numbers and *Pet1::EYFP* transgene expression in these cells is robustly retained at fetal stages to permit their isolation and analysis [20]. PCA of ATAC-seq datasets obtained from *Pet1* neurons from *Pet1::EYFP+, Pet1-/-* hindbrain tissue at E11.5, E14.5, and E17.5 (Supplementary Fig 5b-c) showed that *Pet1-/- Pet1*-lineage neurons (*Pet1-/-* neurons) cluster closely with wildtype *Pet1* neurons in principal component analysis (Fig 4a), consistent with our observation that Pet1-deficient 5-HT neurons do not adopt another neuron-type identity. However, we found a striking overall reduction of Pet1 footprints and decreased accessibility of many TSS-distal TACs in *Pet1-/-* neurons compared to wildtype neurons at all three embryonic stages (Fig 4b-c; Supplementary Fig 5d). To determine whether Pet1 directly controls chromatin accessibility in these chromatin regions, we performed CUT&RUN with E14.5 flow sorted *Pet1* neurons to

generate a genomic occupancy map for Pet1. The CUT&RUN map significantly overlapped with our previous ChIP-seq of Pet1 but identified occupancy at several additional 5-HT neurotransmission genes whose expression is reduced in Pet1-deficient 5-HT neurons [20] (Supplementary Fig 5e). CUT&RUN data showed that Pet1 directly occupies at least 53% of Pet1-dependent TACs (Fig 4d-e). Moreover, 73% of all Pet1-regulated TACs and 80% of Pet1-regulated TACs with Pet1 occupancy were unique to *Pet1* neurons (Fig 4f).

GO analysis revealed that the Pet1 occupied TACs that close in *Pet1-/-* neurons are associated with genes required for serotonin synthesis (Fig 4g; Supplementary Fig 5f). Indeed, we found a significant loss of accessibility of enhancers either upstream of the TSS or within introns of *Tph2*, *Gch1*, and *Ddc*, suggesting that Pet1 remodels postmitotic neuronal chromatin to enable 5-HT synthesis in the brain (Fig 4h). TACs associated with the serotonin reuptake transporter genes *Slc22a3* and *Slc6a4* are also significantly reduced in accessibility (Fig 4i). Importantly, enhancers that require Pet1 for accessibility are unique to *Pet1*+ 5-HT precursors, while the non-*Pet1* neuron specific TACs of 5-HT pathway genes are not significantly affected by Pet1 deficiency (Fig 4h-i). These findings indicate that Pet1 controls the chromatin accessibility of cis-regulatory elements of 5-HT neurotransmission genes.

Interestingly, at the two separate enhancers associated with *Slc6a4* that are both directly occupied by Pet1, chromatin shows little accessibility to transposase at E11.5 but gradually gains accessibility from E11.5 to E17.5 in wildtype *Pet1* neurons (Fig 4i). In contrast, these TAC regions fail to fully open in *Pet1-/-* neurons (Fig 4i), strongly suggesting Pet1 is directly required for initial conversion of these regions to a euchromatic state. Among the Pet1-regulated TACs, we identified 47 regions that similarly fail to fully open between E11.5 and E17.5 in *Pet1*-/- neurons, 20 of which show direct Pet1 occupancy at E14.5 (Supplementary Fig 5g-h). TACs that require Pet1 for

chromatin opening are linked to genes enriched for synaptic products, suggesting that Pet1 opens condensed chromatin at cis-regulatory elements controlling synapse functions (Supplementary Fig 5g).

Continued Pet1 function is required to sustain 5-HT regulatory element accessibility

Pioneer factors in dividing stem cells induce "epigenetic memory" such that once "opened", pioneered sites preserve their accessible state through further rounds of cell division even in the absence of the pioneer factor [37,38]. To determine whether Pet1 is required to sustain stable chromatin accessibility after 5-HT identity gene TACs are already in an euchromatin state, we analyzed flow sorted E14.5 Pet1 conditional knockout (*Pet1*$^{fl/fl;\ Pet1\text{-}Cre;\ Ai9}$) *Pet1*-lineage neurons ("Pet1-cKO neurons") in which *Pet1* transcript level is maintained for about 2 days before it becomes lost at E12.5 [19] (Supplementary Fig 5i-k). We found that in Pet1-cKO neurons, many TACs lose accessibility as in *Pet1-/-* neurons (Fig 5a-c) at distal cis-regulatory elements (Fig 5d), including the enhancers of 5-HT neurotransmission genes *Tph2*, *Gch1*, *Ddc*, and *Slc22a3* (Fig 5e-f).

In addition to selecting 5-HT identity genes for transcriptional activation, Pet1 maintains its expression through positive autoregulation [23]. CUT&RUN showed that Pet1 directly occupies the TAC that overlaps with its auto-regulatory enhancer [23,24], which is consistent with our previous Pet1 ChIP-seq findings [20] (Fig 5g; Supplementary Fig 5e). We found in *Pet1-/-* and Pet1-cKO neurons that Pet1 is needed to maintain the accessibility of this enhancer region (Fig 5g), suggesting that the regulation of chromatin accessibility is an important mechanism for terminal selector autoregulation. Together, these findings suggest the euchromatin landscape in postmitotic 5-HT neurons is a reversible event that requires continuous Pet1 activity for maintenance of accessibility.

Lmx1b co-regulates the enhancer chromatin accessibility of 5-HT identity genes with Pet1

To determine whether the capacity to regulate chromatin accessibility in postmitotic neurons is unique to Pet1, we assessed whether a second 5-HT terminal selector TF, Lmx1b, also controls chromatin architecture. Since Lmx1b occupancy in the brain has not been defined, we performed Lmx1b CUT&RUN in E14.5 *Pet1* neurons, which revealed that a significant portion of Pet1 and Lmx1b binding sites coincide within the same TAC regions (Supplementary Fig 6a-c). In Lmx1b-cKO (*Lmx1b*$^{fl/fl;\ Pet1\text{-}Cre;\ Ai9}$) *Pet1* neurons [18] (Supplementary Fig 6d-f), chromatin landscape is altered (Fig 6a), Lmx1b footprints are reduced (Fig 6b), and more TACs are differentially accessible than in Pet1-cKO neurons (Figs 5a and 6c), suggesting that Lmx1b has a broader impact than Pet1 on 5-HT chromatin. Lmx1b directly binds 19% of the differentially accessible TACs in Lmx1b-cKO (comparable to Pet1 which binds 20% of differentially accessible TACs in Pet1-cKO) indicating direct maintenance of accessibility by Lmx1b at those regions (Fig 6d; Supplementary Fig 5l), which are predominantly distal enhancers (Fig 6e). The Lmx1b-dependent TACs with direct Lmx1b occupancy, of which 65% are specific to *Pet1* neurons (Fig 6h), are most significantly associated with monoamine/serotonin synthesis (Fig 6c). Indeed, the conditional ablation of Lmx1b results in the even greater closing of 5-HT neurotransmission gene TACs than *Pet1* ablation (Fig 6i-j). Many TACs whose accessibility was not significantly reduced in Pet1-cKO neurons (fold change>2, FDR<0.01) significantly lose accessibility in the Lmx1b-cKO neurons (Supplementary Fig 6i), including the TACs of *Gchfr*, *Slc18a2*, and *Maoa* (Fig 6i-j). Lmx1b directly occupies the enhancers of many 5-HT neurotransmission genes (Fig 6i-j), whose expression is known to be critically dependent on Lmx1b [18,39].

We also found that the loss of chromatin accessibility in Lmx1b-cKO is highly similar to that of DKO (*Lmx1b*$^{fl/fl}$*; Pet1*$^{fl/fl;\ Pet1\text{-}Cre;\ Ai9}$) *Pet1* neurons (Figs 6b-c and 6f-g), suggesting that the ablation of both terminal selectors does not have greater effect than ablation of Lmx1b alone. PCA also indicates close clustering of Lmx1b-cKO neurons with DKO neurons (Supplementary Fig 6g). These findings are consistent with the model that Lmx1b is required for Pet1-dependent chromatin regulation. At a substantial subset of Lmx1b-dependent TACs such as those of *Tph2*, *Gch1*, *Ddc*, and *Slc6a4* (Fig 5e-f and 6i-j), Pet1 and Lmx1b non-redundantly co-regulate the accessibility of the cis-regulatory elements. In parallel, Lmx1b further sustains, independent of Pet1, an additional subset of TACs that are required for axon and synaptic formation, neurotrophin signaling, and neuronal growth (Supplementary Fig 6h-i).

Motif analysis on the differentially accessible regions in the cKO and DKO *Pet1* neurons showed that the binding sites of several TF families such as the Rfx-related factors are less accessible in Pet1- and Lmx1b-deficient neurons, demonstrating that together Pet1 and Lmx1b impact the global landscape of accessible TF binding sites (Supplementary Fig 6j).

Pet1 and Lmx1b regulate 5-HT synaptogenesis

In addition to controlling TACs associated with 5-HT transmission genes (Fig 6g), terminal selectors Pet1 and Lmx1b program the cis-regulatory element accessibility of many genes involved in axon growth and synapse formation (Fig 7a; Supplementary Fig 7a), including *Syt1*, which encodes a calcium sensor protein that mediates the release of synaptic vesicles in response to calcium [40], and *Slc17a8*, which encodes a vesicular glutamate transporter (vGlut3) that enables glutamatergic co-transmission in a subset of adult 5-HT neurons (Fig 7b) [41]. In support of the contribution of these Pet1- and Lmx1b-regulated enhancers to developmental gene expression, *Syt1* and *Slc17a8* expression is

decreased in Pet1-cKO, Lmx1b-cKO, and DKO neurons (Fig 7c) together with the expression of other synaptic and axonal genes linked to Pet1- and Lmx1b-regulated TAC regions (Fig 7d; Supplementary Fig 7b-f). The majority (59%) of the TACs maintained by Pet1 or Lmx1b are those that gain accessibility in *Pet1* neurons during maturation (Fig 7e).

To explore the biological significance of these Pet1 and Lmx1b-dependent TAC alterations, we analyzed the integrity of 5-HT pericellular baskets that encase postsynaptic neuronal dendrites and cell bodies in the medial septum [42]. By postnatal day 15 (P15), 5-HT pericellular baskets are readily identified in wildtype brains (Fig 7f; Supplementary Fig 7g). In contrast, in Pet1-cKO, Lmx1b-cKO, and DKO animals, characteristic 5-HT synaptic baskets failed to form (Fig 7f; Supplementary Fig 7g), which is likely attributable to the failure of most axonal projections to reach the medial septum (Fig 7f-g). Of the few axonal terminals that did wrap postsynaptic cell bodies, co-localization of Syt1 and vGlut3 immunostaining with TdTomato+ axon terminals is significantly reduced compared to wildtype animals (Fig 7h-j; Supplementary Fig 7h). This suggests that Pet1 and Lmx1b are required for both the innervation of postsynaptic neurons and the presynaptic machinery that enables 5-HT and glutamatergic co-transmission.

DISCUSSION

Specialized molecular, morphological and functional identities of neurons is largely acquired as postmitotic neurons mature [4]. How postmitotic epigenetic programs control gene expression trajectories to encode mature neuronal identities are poorly understood [7,43]. Here, we have defined the diversity and dynamics of accessible chromatin and the cis-regulatory architecture of developing *Pet1* neurons that generate mature 5-HT neurons [22]. We show that the cis-regulatory landscape of embryonic *Pet1* neurons is highly dynamic and strongly correlated with temporal expression trajectories as postmitotic *Pet1* neurons acquire serotonergic identities. Our scATAC-seq and scRNAseq analyses reveal that *Pet1* neuron chromatin landscapes are highly heterogeneous and established early to prefigure and likely encode 5-HT neuron subtypes including those expressing the glutamatergic marker, *Slc17a8*. We further show that terminal selectors, Pet1 and Lmx1b, play crucial roles in controlling the dynamics of chromatin accessibility to enable serotonergic neurotransmission and synapse formation. Thus, our findings identify key developmental regulators of chromatin accessibility in postmitotic neurons.

We found that distal transcriptional regulatory element accessibility is heterogeneous not only between *Pet1* neurons and other cell types but also within the *Pet1*-lineage neurons. Diversity in regulatory sequence accessibility is predictive of the transcriptomic outcome of single *Pet1* neurons. Furthermore, the temporal maturation of *Pet1* neuron cis-regulatory element accessibility is highly correlated with the expression trajectories of their target genes as serotonergic functional properties progressively emerge. TAC regions are gained near genes associated with synapse organization and axon development and lost near genes linked to cell cycle control and early embryonic morphogenesis. Importantly, the chromatin landscape is more dynamic between E11.5 and E14.5 than between E14.5 and E17.5, suggesting that the early postmitotic period

comprises a highly active stage of chromatin maturation. Since *Pet1* expression is initiated as 5-HT precursors exit the cell cycle, the temporal changes in open chromatin cannot be explained as an altered abundance of *Pet1* neuron subtypes. Rather, chromatin landscape within individual *Pet1* neurons is dynamic during postmitotic maturation.

The regulatory factors controlling the maturation of chromatin dynamics in postmitotic neurons are not well defined. TFs with known pioneer activity such as Ascl1 that are expressed in serotonergic progenitors are significantly downregulated or undetectable in postmitotic *Pet1* neurons [20,44], yet our data suggested that new postmitotic euchromatic regions dynamically appear while other euchromatic regions are lost as 5-HT neuron identity is being acquired. Several recent studies computationally mined genome-wide accessible chromatin sequences to identify TF binding sites that are enriched in regions of dynamic chromatin in neurons or brain regions [10,12,13,45]. Here we show that targeting of Pet1 and Lmx1b in the early stages of postmitotic maturation dramatically alters *Pet1* neuron accessible chromatin landscape. CUT&RUN assays revealed that Pet1 and Lmx1b directly control the accessibility of 5-HT specific cis-regulatory elements linked to genes encoding terminal effectors of 5-HT identity and neurotransmission. In fact, we provide evidence in support of a role for Pet1 in initiating the conversion to euchromatin of multiple nucleosome-protected cis-regulatory sequences in postmitotic neurons, including enhancers linked to the serotonin transporter gene, *Slc6a4*. Both Pet1 and Lmx1b are further required to maintain 5-HT specific open chromatin regions, suggesting that postmitotic accessible chromatin landscapes are unstable and require continuous terminal selector input at least during early postmitotic maturation. Finally, we show that Pet1 is required for accessibility of its own autoregulated enhancer region that directs 5-HT specific expression. Autoregulation is thought to boost and maintain terminal selector expression [46]. We found that in the

absence of *Pet1* expression, *Pet1* enhancer accessibility is dramatically reduced but is not fully closed. This suggests that Pet1 boosts its own expression through augmenting auto-enhancer accessibility that was previously initiated by upstream pioneers but was insufficient for proper expression levels of Pet1. Based on these multiple lines of evidence, we propose that Pet1 and Lmx1b define a previously unrecognized category of chromatin organizers, “postmitotic neuron chromatin organizers” (PnCOs), that act specifically during postmitotic neuron maturation to shape the accessible chromatin landscapes and thereby select enhancers required for acquisition of specialized neuronal identities. Other TFs in addition to Pet1 and Lmx1b may also perform similar function in other neuron types [11,12,47]. Whereas pioneer factors in progenitor cells can drive cell lineage commitment and induce, through transient input, stable “memories” of epigenetic states that are transmitted through further cell divisions [48], PnCOs function in differentiated cells that may require continuous input to sustain the euchromatin state of the regulatory sequences of terminal effector genes. The mechanism by which Pet1 and Lmx1b program neuronal chromatin is not understood, but could involve the formation of enhanceosome complexes, recruitment of chromatin remodelers, or direct cooperative displacement of histones [49].

While Lmx1b maintains accessibility of a distinct set of cis-regulatory elements independent of Pet1, chromatin regulation of many 5-HT TACs requires direct and non-redundant co-regulation by both terminal selectors, reminiscent of the cooperative activities of TFs and pioneer factors in other cell types [50,51]. Because a large number of cognate motif sites can exist in the genome for each individual TF, selective chromatin accessibility driven by collaborative PnCOs may serve to ensure that neuron-type-specific effector genes are restrictively activated only in the presence of multiple positive regulatory inputs. Altogether, our findings unearth a paradigm by which terminal

selectors not only activate sequence-specific transcription but also co-program the postmitotic neuronal epigenome to select active gene regulatory sites.

In addition to neurotransmitter synthesis, neurons must develop axonal and synaptic connectivity in order to transmit signals [52,53]. Pet1 and Lmx1b organize the access to cis-regulatory elements for both 5-HT neurotransmission and synaptic connectivity so that the acquisition of distinct functional properties is developmentally coupled. A growing body of evidence supports the idea that alterations in 5-HT neurotransmission cause behavioral changes related to those that characterize neuropsychiatric disorders such as anxiety and autism [21,54]. As many of these disorders are neurodevelopmental in origin, the study of the chromatin regulatory programs that control 5-HT neuron and other identities may shed light on the pathophysiology of brain disorders at critical stages of brain development.

METHODS

Animals

Mice were maintained in ventilated cages on a 12h light/dark cycle with access to food and water *ad libitum* with 2-5 mice per cage. The transgenic mouse lines used in this study have been described previously [18,20]. Conditional knockout mice (Lmx1b-cKO, Pet1-cKO, DKO) were generated by breeding animals bearing Ai9 Rosa Tomato ($Rosa^{Tom}$; Jackson Labs Stock #: 007909) and either the $Lmx1b^{fl/fl}$ [55], $Pet1^{fl/fl}$ [19], or both alleles with animals harboring *Pet1-Cre* (original name ePet-Cre) transgene. *Pet1+/+; Pet1::EYFP* and *Pet1-/-; Pet1::EYFP* mice were generated by crossing *Pet1+/-* animals bearing the *Pet1::EYFP* transgene. The *Pet1::EYFP* transgene expresses Yfp under the same *Pet1* enhancer as the *Pet1* enhancer region driving the *Pet1-Cre*. Genotyping was performed using ear tissues with the following primers: Pet-1: 5'-CGGTGGATGTGGAATGTGTGCG-3', 5'-CGCACTTGGGGGGTCATTATCAC-3', 5'-GCCTGATGTTCAAGGAAGACCTCGG-3'; floxed Pet-1: 5'-TAGGAGGGTCTGGTGTCTGG-3', 5'-GCGTCCTTGTGTGTAGCAGA-3'; Pet-EYFP: 5'-TATATCATGGCCGACAAGCAG-3', 5'-GAACTCCAGCAGGACCATGT-3'; floxed Lmx1b: 5'- AGGCTCCATCCATTCTTCTC-3', 5'- AGGCTCCATCCATTCTTCTC-3'; Rosa TdTomato: 5'-CCACAATAAGCAAGAGGCAC-3', 5'- CCGAAAATCTGTGGGAAG TC-3', 5'- CCGAAAATCTGTGGGAAG TC-3', 5'- CTGTTCCTGTACGGCATG G-3'; Pet1-Cre: 5'- CGGCATCAACGTTTTCTTTT -3', 5'- AGTCAGGGCAGAGCCATCTA -3'. Embryonic age was defined as the number of days following the appearance of a vaginal plug, designated as embryonic day (E) 0.5. For all procedures, animals were euthanized by anesthesia with isoflurane inhalation followed by cervical dislocation, and brain tissues were collected equally from male and female mice or embryos. All animal experiments were approved by the Institutional Animal Care and Use Committee of CWRU.

Neuron dissociation and flow cytometry

Isolation of *Pet1* neurons was performed as previously described with a few modifications [18]. Briefly, E11.5 or E14.5 embryos from *Pet1-Cre*$^{+}$*;Ai9* or *Pet1::EYFP+* animals were dissected in cold PBS under dissection microscope to collect all fluorescent cells in the fetal mouse hindbrain. Dissected tissues were transferred into cold 1X TrypLE Express (Gibco) and incubated at 37°C for 15 mins with gentle agitation. Following enzymatic trypsin digestion, samples were centrifuged for 1 min at 1500 rpm, TrypLE Express was removed, and samples were washed and resuspended in 500µL cold Leibovitz's L-15 media (Thermo Fisher) with 4µL DNAse I (Invitrogen) and 10% FBS (Gibco). Cells were then slowly triturated with fire-polished glass Pasteur pipettes of decreasing bore size until a single cell suspension was achieved.

E17.5 embryos were dissociated using a modified adult neuron dissociation method. *Pet1-Cre*$^{+}$*;Ai9* or *Pet1::EYFP+* embryos were dissected in cold aCSF solution (3.5mM KCl, 126mM NaCl, 20mM NaHCO3, 20µM Dextrose, 1.25 mM NaH_2PO_4, 2mM $CaCl_2$, 2mM $MgCl_2$, 50µM AP-V (Tocris), 29µM DNQX (Sigma), and 100nM TTX (Abcam)) to collect all fluorescent cells in the hindbrain. Dissected tissues were then incubated in bubbling (95% O_2, 5% CO_2) aCSF containing 1mg/mL protease from Streptomyces griseus (Sigma) for 15 mins at room temperature. Following enzymatic digestion, tissues were transferred to 500µL cold aCSF/10%FBS solution containing 4µL DNase I (Invitrogen) and gently triturated with fire-polished glass pipettes to generate single cell suspension.

Dissociated cells were filtered through a 40µm filter (BD Biosciences) and sorted on a Becton Dickinson FACS Aria or FACS Aria-SORP digital cell sorter with 85µm nozzle. Viability of flow sorted cells were routinely checked with Trypan blue staining followed by manual counting of intact cells on small portions of the FACS purified neurons.

Bulk *Pet1* neuron ATAC-sequencing library preparation

ATAC-seq was carried out as previously described [56] with a few modifications using rostral and caudal *Pet1* neurons flow sorted with either *Pet1::EYFP* (wildtype and *Pet1-/-* samples) or *Pet1-Cre;Ai9* (Pet1-cKO, Lmx1b-cKO, and DKO samples). Bulk ATAC-seq for each time-point/condition was performed with 3-5 biological replicates. Each biological replicate consisted of embryos flow sorted on the same day from either one or multiple dams that were then pooled to collect a total of 10,000-50,000 *Pet1* neurons. Neurons were sorted into L15/10%FBS solution, washed with ice cold PBS, and resuspended in 25µL ice cold ATAC-RSB buffer (10mM Tris-HCl, pH 7.4, 10mM NaCl, 3mM MgCl2, 0.1% NP40 (Roche cat#11332473001)). Cells were incubated in ATAC-RSB on ice for 2 mins, then centrifuged at 800 xg for 10 mins at 4°C. Pelleted nuclei were resuspended in 25µL Transposition mix (12.5µL 2x TD buffer (Illumina), 1.25µL transposase (100nM final) (Illumina), 8.25µL PBS, 3.25µL H_2O) for E11.5 and E14.5 samples and 15µL of Transposition mix for E17.5 samples. The transposition reactions were incubated at 37°C for 25 mins on a Thermomixer at 1000rpm. The resulting DNA fragments were purified using the DNA Clean and Concentrator-5 Kit (Zymo) and PCR amplified for 13 cycles with Illumina Nextera adapter primers using the NEBNext High Fidelity 2X Master Mix (NEB) with the following PCR program: (1) 5 mins at 72°C (2) 30s at 98°C (3) 10s at 98°C, (4) 30s at 63°C, (5) 1 min at 72°C, and (6) Repeat steps 3-5, 12X. Final PCR products were cleaned using PCRClean Dx beads (Aline Biosciences), assessed for quality on a Bioanalyzer, and sequenced for 75 cycles to an average depth of 29M paired-end reads per sample on an Illumina NextSeq 550.

ChIP-seq using Tagmentation (ChIPmentation)

Pet1 neurons (caudal and rostral) were flow sorted from *Pet1::EYFP* embryos at E14.5 into PBS/10% FBS. Neurons were frozen in PBS as cell pellets in -80°C until use.

50,000-300,000 frozen cells from multiple days were then thawed on ice, fixed with 1% formaldehyde in PBS for 10 mins at room temperature followed by quenching with 0.125M glycine, and pooled to generate one biological replicate. Two biological replicates were collected for each antibody. ChIP-seq using Tagmentation (ChIPmentation) was then performed as previously described [25]. Cells were washed once with ice cold PBS and lysed in 130µL Sonication Buffer (10mM Tris-HCl pH 8, 0.25% SDS, 2mM EDTA, 1x protease inhibitor cocktail (Active Motif # 37490), 1mM PMSF). Sonication was conducted in MicroTUBEs (130µL) on a Covaris S2 ultrasonicator until DNA fragments were in the range of 200-700bp. Lysates were diluted 1:1.5 with Equilibrium Buffer (10mM Tris-HCl pH 8, 233mM NaCl, 1.66% Triton X-100, 0.166% sodium deoxycholate, 1mM EDTA, 1x protease inhibitor cocktail (Active Motif)) and incubated with 3µg of antibody on a rotator overnight at 4°C. The following antibodies were used: H3K4me3 (Diagenode C15410003-50), H3K4me1 (Abcam ab8895), H3K27ac (Abcam ab4729), H3K27me3 (Active Motif 39055). The next day, 10µL of Protein A/G Dynabeads blocked overnight with 0.1% BSA in RIPA-LS Buffer (10mM Tris-HCl pH 8, 140mM NaCl, 1mM EDTA, 0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100) were transferred to each immunoprecipitated lysate and incubated for 2h at 4°C on a rotator. Following incubation, beads were washed twice with ice cold RIPA-LS Buffer, twice with cold RIPA-HS Buffer (10mM Tris-HCl pH 8, 500mM NaCl, 1mM EDTA, 0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100), twice with RIPA-LiCl Buffer (10mM Tris-HCl pH 8, 1mM EDTA, 250mM LiCl, 0.5% NP-40, 0.5% sodium deoxycholate), and twice with 10mM Tris-HCl pH8. Beads were then resuspended in 25µL of tagmentation reaction mix containing 12.5µL of 2x Tagmentation DNA Buffer (Illumina #15027866) and 1µL of Tagment DNA Enzyme (Illumina #15027865). Tagmentation reaction was incubated at 37°C for 10mins with gentle vortex every 5 mins. Tn5 reaction was inactivated by adding 500µL of ice-cold RIPA-LS Buffer.

Beads with washed twice with RIPA-LS Buffer and twice with TE Buffer (10mM Tris-HCl pH 8, 1mM EDTA) and resuspended in 48µL of ChIP Elution Buffer (10mM Tris-HCl pH 8, 5mM EDTA, 300mM NaCl, 0.4% SDS) and 2µL of Proteinase K solution (ThermoFisher). Digestion occurred at 55°C for 1h followed by sample de-crosslinking at 65°C overnight. Supernatant was transferred the next day to a new tube, and Proteinase K digestion was repeated on remaining bead-bound chromatin by resuspending the beads in 1µL Proteinase K in 19µL ChIP elution buffer and eluting for 1h at 55°C. Supernatants from both elutions were combined and purified using DNA Clean and Concentrator-5 Kit (Zymo). Library preparation was performed as described for ATAC-seq. ChIPmentation libraries were analyzed on a Bioanalyzer for quality control and then sequenced with 2x75 paired-end reads per sample on an Illumina NextSeq 550.

Single cell ATAC-seq library construction

Pet1 neurons from *Pet1::EYFP* embryos were flow sorted at E14.5 into PBS/10% FBS, washed twice with PBS/0.04%BSA, and resuspended in 100µL freshly prepared Lysis Buffer (10mM Tris-HCl pH 7.4, 10mM NaCl, 3mM $MgCl_2$, 0.1% Tween-20, 0.1% IGEPAL). After 3 mins of cell lysis on ice, 1mL of chilled Wash Buffer (10mM Tris-HCl pH 7.4, 10mM NaCl, 3mM $MgCl_2$, 1% BSA, 0.1% Tween-20) was added to the reaction and nuclei were pelleted and resuspended in 1X Nuclei Buffer (10x Genomics). Nuclei concentration and quality were determined using Trypan blue staining followed by cell counting and inspection on a hemocytometer to ensure that nuclei were free of nuclear membrane blebbing and that there was complete lysis. Nuclei were then added to a transposition mix to perform the Tn5 transposase reaction. The reaction was incubated at 37ºC for 60 mins. Transposed nuclei were then loaded onto the 10x Genomics Chromium controller to perform nuclei encapsulation with barcoded gel beads, followed by library amplification for 13 cycles. After purification, the library was analyzed on an

Agilent Fragment Analyzer to determine fragment size. Library was sequenced on Illumina Nextseq 550 Sequencing System to a depth sufficient to yield >50,000 paired-end reads per nuclei.

Single cell RNA-seq library construction

Pet1 neurons from *Pet1::EYFP* embryos were flow sorted at E14.5 into PBS/10% FBS. FACS purified neuron viability was confirmed to be >80% by staining a portion of the cells with Trypan blue and counting the number of live cells on a hemocytometer. Single cells were loaded in the 10x Single Cell 3' Chip at a concentration of 1000 cells/µL. Library was prepared according to 10x Genomics Single Cell Protocols Cell Preparation Guide on the 10x Genomics Chromium controller and sequenced to a depth of over 80,000 reads per cell for ~2,700 cells on Illumina Nextseq 550 Sequencing System.

CUT&RUN library preparation

CUT&RUN was performed based on the Epicypher protocol v1.6, which is derived from the method developed by Skene, Henikoff, and Henikoff (2018) [57]. Briefly, Concanavalin A-coated (ConA) beads were washed twice with CUT&RUN Bead Activation Buffer (20mM HEPES pH 7.9, 10mM KCl, 1mM $CaCl_2$, 1mM $MnCl_2$). FACS purified *Pet1* neurons were washed twice with Wash Buffer (20mM HEPES pH 7.5, 150mM NaCl, 0.5mM spermidine, protease inhibitors (Roche)) and then incubated with activated ConA beads for 15 mins at room temperature to allow the cells to absorb to the beads. ConA-bead bound cells were incubated overnight in 50µL Antibody Buffer (Wash Buffer, 0.01% digitonin, 2mM EDTA) with 1:100 rabbit anti-FEV (Sigma-Aldrich #HPA067679), guinea pig anti-Lmx1b (Carmen Birchmeier, Max Delbruck Center Berlin), or rabbit monoclonal IgG antibody (Jackson ImmunoResearch Lab 115-007-003). One replicate was collected for each antibody. After overnight incubation, the ConA-bound cells were washed twice

with Digitonin Wash Buffer (Wash Buffer, 0.01% digitonin), resuspended in 50µL Digitonin Buffer containing 700ng/mL CUTANA pAG-MNase (Epicypher), and incubated for 10 mins at room temperature. After binding of pAG-MNase, ConA bead bound cells were washed twice with Digitonin Buffer, then resuspended in 50µL Digitonin Buffer containing 2mM $CaCl_2$ and incubated at 4°C for 2h on a rotator to induce digestion of the chromatin. Reaction was stopped by the addition of 33µL STOP Buffer to each sample (340mM NaCl, 20mM EDTA, 4mM EGTA, 50µg/mL RNase A (ThermoFisher), 50µg/mL GlycoBlue (ThermoFisher)) and incubated at 37°C for 15 mins to release cleaved chromatin into the supernatant and degrade RNA. DNA was purified from the supernatant using the DNA Clean and Concentrator-5 kit (Zymo). Sequencing library preparation was performed using the NEBNext Ultra II Library Prep Kit for Illumina (NEB) according to the manufacturer's instructions, with the following modification: following adapter ligation DNA was purified using 1.75x PCRClean Dx beads (Aline Biosciences) and libraries were amplified using the following CUT&RUN specific PCR cycling conditions to enrich for 100-700bp fragments: (1) 45s at 98°C, (2) 15s at 98°C, (3) 10s at 60°C, (4) repeat steps 2-3 for a total of 14X, and (5) 1 min at 72°C. PCRClean Dx beads were used to purify the final libraries, which were then sequenced at 75bp x2 on an Illumina NextSeq 550.

RNA-seq library preparation

Rostral *Pet1* neurons were dissected from E17.5 DKO (*Pet1-Cre; Pet1*$^{fl/fl}$*; Lmx1b*$^{fl/fl}$*; Ai9)* or control *Pet1-Cre; Pet1*$^{+/+}$*; Lmx1b*$^{+/+}$*; Ai9* embryos and flow sorted into Trizol LS (Invitrogen) on a Becton Dickonson FACS Aria-SORP sorter. Experiment was performed with three control and two DKO biological replicates. Each biological replicate represented 1000-3000 pool flow sorted neurons from multiple embryos of one or more litters. Total RNA was isolated using chloroform extraction and purified using RNA Clean

& Concentrator-5 Kit (Zymo). RNA concentration and quality were measured using Quantifluor RNA system (Promega) and Agilent Fragment Analyzer (Advanced Analytics). DNase I treatment and library construction was performed using the Ovation SoLo RNA-Seq Systems V2 according to manufacturer's instructions. Final libraries were sequenced on a NextSeq 550 (Illumina) with single-end sequencing for 75 cycles.

Immunohistochemistry, image acquisition, and processing

Mice were deeply anesthetized by isoflurane inhalation (chamber atmosphere containing 4% isoflurane) and perfused for 20 mins with cold 4% paraformaldehyde (PFA) in PBS. Brains were post-fixed in 4% PFA overnight and transferred to 30% sucrose/PBS overnight. Tissues were then frozen in Optimal Cutting Temperature (OCT) solution and sectioned on a cryostat at 25μm thickness. Tissue sections were mounted on SuperFrost Plus slides (Thermo Fisher Scientific) and dried. For antigen retrieval, sections were heated up in Sodium Citrate Buffer for 5 min in the microwave at low power. Sections were blocked with 10% NGS in PBS-T for 1 hr, followed by blocking of endogenous biotin with Avidin/Biotin Blocking kit (Vector) only for Syt1 and vGlut3 stained samples. Sections were incubated with primary antibody (1:1000) at room temperature overnight, then treated the next day with secondary antibodies for 1hr, followed by visualization (Tph2, GFP, and RFP) or further processing with the VECTASTAIN ABC kit (VECTOR) according to manufacturer's instructions before visualization. Primary antibodies used were: mouse anti-vGlut3 antibody (Abcam ab134310), mouse anti-Synaptotagmin1 (Sigma, SAB1404433), rabbit anti-Tph2 (Millipore Sigma, ABN60), rabbit anti-RFP (Rockland 600-401-379), and chicken anti-GFP (Abcam, ab13970) Secondary antibodies used were goat anti-mouse, Alexa Fluor 488 or 647 (Invitrogen, A10029 or A28181; 1:500), goat anti-chicken, Alexa Fluor 488 (Invitrogen, A11039 #11039, 1:500), goat anti-rabbit, Alexa Fluor 594 (Invitrogen,

A11037, 1:500), and biotinylated goat anti-guinea pig IgG antibody (Vector, BA7000; 1:200).

Stained slides were imaged on an LSM800 confocal microscope (Carl Zeiss). In each experiment two animals are analyzed per genotype. Images were acquired per animal using 20X and 63X objective lens and processed in ImageJ. Signal brightness and contrast were edited across whole images equally among genotypes for all pericellular basket synapse and axon images. Signal intensity calculations were normalized for each section. Two-way ANOVA with Welch's correction analysis was performed.

ATAC-seq data processing and analysis

ATAC-seq data were processed using the standardized software pipeline from the ENCODE consortium to perform quality-control, read alignment, and peak calling [58]. FASTQ files from ATAC-seq reads were mapped to UCSC mm10 with Bowtie v2.3.4.3. Samtools v1.9 was used to remove PCR duplicates and chrM reads. Peak calling was performed using MACS2 with the following parameters: --nomodel --shift 37 --ext 73 --pval 1e-2 -B --SPMR --callsummits. Then by intersecting called peaks in replicates using BEDTools v2.29.0, we defined replicated peaks as the set of peaks present independently in each replicated peak call set and also called in the pooled set.

Differential accessibility analysis

A comprehensive set of ATAC-seq peaks (TACs) from control datasets (E11.5, E14.5, and E17.5) was generated to be used as the reference for assessing differential accessibility between control time points and for control to TF knockout comparisons. All control peak sets were merged using 1bp overlap to create the union of all control peaks (E11.5, E14.5, and E17.5). Any potential batch effects were corrected using surrogate variable analysis (SVA) [59]. The number of significant surrogate variables was determined

using an iteratively re-weighted least squares approach. One significant surrogate variable was found and added to the model design in DESeq2 [60]. Differentially accessible peaks were determined using DESeq2 v1.32.0 with a fold change cutoff of 2-fold and FDR adjusted p-value ≤0.01 for all wildtype and conditional knockout comparisons and ≤0.05 for *Pet1-/-* comparisons. BEDTools v2.30.0 was used to find shared intersections between identified peaks (1bp minimum overlap). Peak annotations were performed using ChIPseeker v1.28.3. Principal component analysis was performed using the prcomp function within the 'plotPCA' wrapper in DESeq2 on variance stabilizing transformed count data to correct for the underlying mean-variance relationship and plotted with ggplot2 [60].

Volcano plots were generated using the EnhancedVolcano package in R (v1.10.0, https://github.com/kevinblighe/EnhancedVolcano).

Motif enrichment

TF motifs were identified using monaLisa v0.1.50 (https://fmicompbio.github.io/monaLisa/index.html). Differentially accessible peak regions were binned into "loss", "no change", and "gain" categories and motif enrichment for each bin was compared against an equal number of sequences randomly sampled from the genome matched by GC composition. Statistically significant enriched motifs were determined using a one-tailed Fisher's exact test with Benjamini-Hochberg adjusted p-values ≤ 1e-8.

Gene ontology enrichment

Gene ontology was performed using Webgestalt requiring >5 and <2000 genes per category and FDR ≤0.05 [61].

ChIPmentation data processing and analysis

ChIPmentation data were processed using the standard ENCODE ChIP-seq software pipeline from the ENCODE Consortium. Sequencing reads were mapped to the UCSC mm10 by Bowtie v2.3.4.3. PCR duplicates were removed using samtools v1.9. The unique mapped reads were used to call peaks comparing immunoprecipitated chromatin with input chromatin with MACS2 v2.2.4 with FDR adjusted p-value cutoff of 0.01. Peak lists were filtered to remove peaks overlapping ENCODE blacklisted regions. BEDTools was used to find shared intersections between identified peaks (1bp minimum overlap). Peak annotations were performed using ChIPseeker v1.28.3.

Chromatin state analysis

Chromatin state analysis was performed using ChromHMM v1.22 [26]. Four histone mark ChIPmentation datasets were binarized at a resolution of 200 bp with a signal threshold of 4. We trained a 10-state model as it characterized the key interactions between the histone marks. The posterior probability state of each genomic bin was calculated using the learned model.

Super-Enhancer analysis

The rank-ordering of super-enhancers (ROSE) package was used to identify super-enhancers based on the ranking of H3K27ac signal intensities from E14.5 *Pet1* neurons [27,28]. Clusters of E14.5 ATAC-seq peaks within 12.5 kb were grouped and ranked for their total input background subtracted H3K27ac signal with super-enhancers called above the inflection point of 6391.9.

scATAC-seq analysis

Raw scATAC-seq data was processed using the CellRanger ATAC v1.2.0 pipeline (10X Genomics). Reads were aligned to the GRCm38 (mm10) mouse genome. Peaks were called using MACS2 v2.2.4. The peak/cell matrix and all unique single-cell fragments were imported into R using the Signac package [62]. Several QC metrics were calculated to assess the quality of the data. Cells with >3000 to <100000 peak region fragments, >40% reads in peaks, <0.025 blacklist ratio, <4 nucleosome signal, >2 TSS enrichment were retained for further analysis. The data was normalized and dimensionality reduced using latent semantic indexing (LSI) [63]. First, term frequency-inverse document frequency (TF-IDF) normalizes across cells for sequencing depth and across peaks to give higher values to rare peaks. Next, we ran singular value decomposition (SVD) on the TF-IDF matrix using all variable features. A gene activity matrix was constructed by counting the number of fragments that map within the 2 kb upstream region of each gene. Motif enrichment and TF activity for scATAC-seq clusters were quantified using the "RunChromVAR" wrapper function within the Signac package (v1.14.0) [62,64]. The filtered cell (n=1692) by peak (n=124111) matrix was used as input with the curated database of mm10 cisBP motifs (n=884). Default settings were used to compute motif accessibility deviation Z-scores.

scRNA-seq analysis

Raw scRNA-seq data was processed using the CellRanger v3.1.0 pipeline (10X Genomics). Reads were aligned to the GRCm38 (mm10) mouse genome. The unique molecular index (UMI) count matrix was imported into R using Seurat [65–68]. Cells with < 200 or > 7500 expressed genes or > 10% mitochondrial genes were excluded from further analysis. Gene counts were log normalized, then the dataset was scaled and centered and mitochondrial gene levels were regressed out. Dimensionality of the data

was determined using an elbow plot and the first 20 principal components were used to reduce dimensionality using principal component analysis.

Integration of scATAC-seq and scRNA-seq data

scATAC-seq and scRNA-seq datasets were integrated using SCTransform from Seurat using 3000 variable features. To perform clustering, we first computed a principal components analysis (PCA) on the normalized count matrix and used the first 20 PCs to generate a shared nearest neighbor (SNN) graph with the k parameter set to 20. Then we clustered the data with a resolution of 0.8 using the Louvain algorithm. Next, we performed uniform manifold approximation and projection (UMAP) to project the clustering onto low dimensional space and plot the first two dimensions for visualizations (30 nearest neighbors, minimum distance = 0.3) [69]. Cluster markers were determined using the Wilcoxon Rank Sum test with a minimum fraction of cells expressing the gene at 0.25 and a log fold-change threshold of 0.25 for each cluster versus all other clusters. The top two marker genes for each cluster are used to identify each cluster.

Enhancer-gene association

scATAC-seq and scRNA-seq seurat objects were imported into SnapATAC [34]. A gene expression profile for each cell based on the weighted average of the nearest neighboring cells (k = 15) was imputed based on the scRNA-seq dataset. From the integrated scATAC-seq and scRNA-seq datasets, a “pseudo” multiomics cell is made that contains both gene expression and chromatin accessibility information. Then, logistic regression was performed to quantify the association between gene expression and the binarized accessibility state at putative enhancer elements in a 500 kb window centered at the TSS of each gene. The gene to putative enhancer associations were filtered with p-value ≤ 0.05.

RNA-seq analysis

RNA-seq reads were aligned to UCSC mm10 using Hisat2 (v2.1.0) with a penalty of 12 for a non-canonical splice site, maximum and minimum penalty for mismatch of 1,0, and maximum and minimum penalty for soft-clipping of 3,0 parameters [70]. Gene expression quantification and differential expression were calculated using Cufflinks v2.2.2 with a fold change cutoff of 1.5 fold and FDR cutoff of 0.05 [71–73].

CUT&RUN data processing and analysis

CUT&RUN data were processed using the standard ENCODE ChIP-seq software pipeline from the ENCODE Consortium. Sequencing reads were mapped to the UCSC mm10 by Bowtie v2.3.4.3. PCR duplicates were removed using samtools v1.9. The unique mapped reads were used to call peaks comparing TF-bound chromatin with an IgG control with SEACR using both stringent and relaxed threshold settings [74]. The stringent threshold uses the peak of the total signal curve and the relaxed threshold uses the total signal of the “knee” to the peak of the total signal curve. Both threshold settings have been shown to have high sensitivity and specificity [74]. BEDTools was used to find shared intersections between identified peaks (1bp minimum overlap). Peak annotations were performed using ChIPseeker v1.28.3. CUT&RUN data for Pet1 and Lmx1b were normalized together for visualization using the ChIPseqSpikeInFree method to determine the scaling factor for adjusting signal coverage across the genome [75].

Footprinting analysis

Transcription factor footprints were identified using RGT-HINT [76,77]. First, we searched for motifs from HOCOMOCO v11 within the set of differentially accessible ATAC-seq peaks for each TF knockout condition [78]. Then, differential footprints were called on

uniquely-mapped paired-end ATAC-seq reads with bias correction for the Tn5 transposase within the set of differentially accessible ATAC-seq peaks comparing control to each TF knockout condition.

Genomic visualization

ATAC-seq, ChIPmentation, and CUT&RUN data were visualized in Integrative Genomic Viewer (IGV, v2.9.3). Bam files for all ATAC-seq and ChIPmentation datasets were normalized to 1x genome coverage using deeptools (v3.5.0) bamCoverage.

Statistical analysis

No statistical methods were used to predetermine sample size. Measurements for all experiments were taken from distinct samples. No methods were used to determine whether the data conformed to the assumptions of the statistical method. Experimenters were blinded to genotype during immunofluorescence analysis of 5-HT pericellular baskets. No other blinding of experiments was performed. Clustering of scATAC-seq and scRNA-seq was performed in an unbiased manner. Cell-subtypes were assigned after unbiased clustering. Nuclei that did not meet quality thresholds were removed from further analyses as described in the results and methods.

Data visualization

Data plots were generated using ggplot2 (v3.3.5) in R (v4.0.4) [79] and Microsoft Excel. 5-HT neurotransmission diagram was created using Biorender.com.

Code availability

No custom code was used in this study. Open-source algorithms were used as described in the analysis methods. Details on how these algorithms were used to analyze data are available from the corresponding author upon request.

ACKNOWLEDGMENTS

This research was supported by NIH grants RO1 MH117643 and RO1 MH062723 to E.S.D., F30 MH122173 and T32 GM007250 to X.L.Z., and the Uehara Memorial Foundation Fellowship (201940009) to N.T. We thank Katherine Lobur for assistance with mouse breeding and Laura Marsland for assisting with immunohistochemistry analysis of Tph2 in embryonic mouse hindbrain. We thank Dr. Polyxeni Philippidou for helpful comments on the manuscript. We thank Dr. Carmen Birchmeier for the anti-Lmx1b antibody. This work was supported by the Cytometry and Microscopy Shared Resource at Case Comprehensive Cancer Center, the CWRU Genomics Core, and the CWRU Applied Functional Genomics Core.

AUTHOR CONTRIBUTIONS

W.C.S and E.S.D conceived the original project. X.L.Z., W.C.S., and E.S.D, designed the experiments. X.L.Z. performed all experiments except for immunofluorescence analysis and quantification of pericellular baskets, which was performed by N.T. W.C.S. performed all data processing and bioinformatics analyses. X.L.Z, W.C.S, N.T, and E.S.D analyzed and interpreted the findings. X.L.Z, W.C.S, and E.S.D wrote the manuscript.

Competing Interests Statement

The authors declare no competing interests.

1. Hobert, O. Regulation of Terminal Differentiation Programs in the Nervous System. *Annu. Rev. Cell Dev. Biol.* **27**, 681–696 (2011).
2. Deneris, E. & Gaspar, P. Serotonin neuron development: shaping molecular and structural identities. *Wiley Interdiscip. Rev. Dev. Biol.* **7**, (2018).
3. Stiles, J. & Jernigan, T. L. The Basics of Brain Development. *Neuropsychol. Rev.* **20**, 327–348 (2010).
4. Polleux, F., Ince-Dunn, G. & Ghosh, A. Transcriptional regulation of vertebrate axon guidance and synapse formation. *Nat. Rev. Neurosci.* **8**, 331–340 (2007).
5. De Rubeis, S. *et al.* Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**, 209–215 (2014).
6. Heavner, W. E. & Smith, S. E. P. Resolving the Synaptic versus Developmental Dichotomy of Autism Risk Genes. *Trends Neurosci.* **43**, 227–241 (2020).
7. Gallegos, D. A., Chan, U., Chen, L.-F. & West, A. E. Chromatin Regulation of Neuronal Maturation and Plasticity. *Trends Neurosci.* **41**, 311–324 (2018).
8. Perino, M. & Veenstra, G. J. C. Chromatin Control of Developmental Dynamics and Plasticity. *Dev. Cell* **38**, 610–620 (2016).
9. Gorkin, D. U. *et al.* An atlas of dynamic chromatin landscapes in mouse fetal development. *Nature* **583**, 744–751 (2020).
10. de la Torre-Ubieta, L. *et al.* The Dynamic Landscape of Open Chromatin during Human Cortical Neurogenesis. *Cell* **172**, 289-304.e18 (2018).
11. Frank, C. L. *et al.* Regulation of chromatin accessibility and Zic binding at enhancers in the developing cerebellum. *Nat. Neurosci.* **18**, 647–656 (2015).
12. Preissl, S. *et al.* Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat. Neurosci.* **21**, 432 (2018).

13. Trevino, A. E. *et al.* Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution. *Cell* **184**:5053-5069.e23 (2021).
14. Stroud, H. *et al.* An activity-mediated transition in transcription in early postnatal neurons. *Neuron* **107**, 874-890.e8 (2020).
15. Celada, P., Puig, M. V. & Artigas, F. Serotonin modulation of cortical neurons and networks. *Front. Integr. Neurosci.* **7**, 25 (2013).
16. Deneris, E. S. & Wyler, S. C. Serotonergic transcriptional networks and potential importance to mental health. *Nat. Neurosci.* **15**, 519–527 (2012).
17. Hobert, O. Regulatory logic of neuronal diversity: Terminal selector genes and selector motifs. *Proc. Natl. Acad. Sci.* **105**, 20067–20071 (2008).
18. Donovan, L. J. *et al.* Lmx1b is required at multiple stages to build expansive serotonergic axon architectures. *eLife* **8**, e48788 (2019).
19. Liu, C. *et al.* Pet-1 is required across different stages of life to regulate serotonergic function. *Nat. Neurosci.* **13**, 1190–1198 (2010).
20. Wyler, S. C. *et al.* Pet-1 Switches Transcriptional Targets Postnatally to Regulate Maturation of Serotonin Neuron Excitability. *J. Neurosci.* **36**, 1758–1774 (2016).
21. Doan, R. N. *et al.* Recessive gene disruptions in autism spectrum disorder. *Nat. Genet.* **51**, 1092 (2019).
22. Hendricks, T., Francis, N., Fyodorov, D. & Deneris, E. S. The ETS Domain Factor Pet-1 Is an Early and Precise Marker of Central Serotonin Neurons and Interacts with a Conserved Element in Serotonergic Genes. *J. Neurosci.* **19**, 10348–10356 (1999).
23. Scott, M. M., Krueger, K. C. & Deneris, E. S. A Differentially Autoregulated Pet-1 Enhancer Region Is a Critical Target of the Transcriptional Cascade That Governs Serotonin Neuron Development. *J. Neurosci.* **25**, 2628–2636 (2005).

24. Scott, M. M. *et al.* A genetic approach to access serotonin neurons for in vivo and in vitro studies. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16472–16477 (2005).
25. Schmidl, C., Rendeiro, A. F., Sheffield, N. C. & Bock, C. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat. Methods* **12**, 963–965 (2015).
26. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* **12**, 2478–2492 (2017).
27. Lovén, J. *et al.* Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. *Cell* **153**, 320–334 (2013).
28. Whyte, W. A. *et al.* Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* **153**, 307–319 (2013).
29. Krueger, K. C. & Deneris, E. S. Serotonergic transcription of human FEV reveals direct GATA factor interactions and fate of Pet-1 deficient serotonin neuron precursors. *J. Neurosci. Off. J. Soc. Neurosci.* **28**, 12748–12758 (2008).
30. Okaty, B. W. *et al.* Multi-Scale Molecular Deconstruction of the Serotonin Neuron System. *Neuron* **88**, 774–791 (2015).
31. Okaty, B. W. *et al.* A single-cell transcriptomic and anatomic atlas of mouse dorsal raphe Pet1 neurons. *eLife* **9**, e55523 (2020).
32. Ren, J. *et al.* Single-cell transcriptomes and whole-brain projections of serotonin neurons in the mouse dorsal and median raphe nuclei. *eLife* **8**, e49424 (2019).
33. Wylie, C. J. *et al.* Distinct transcriptomes define rostral and caudal serotonin neurons. *J. Neurosci.* **30**, 670–684 (2010).
34. Fang, R. *et al.* Comprehensive analysis of single cell ATAC-seq data with SnapATAC. *Nat. Commun.* **12**, 1337 (2021).
35. Ma, S. *et al.* Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin. *Cell* **183**, 1103-1116.e20 (2020).

36. Okaty, B. W., Commons, K. G. & Dymecki, S. M. Embracing diversity in the 5-HT neuronal system. *Nat. Rev. Neurosci.* **20**, 397 (2019).

37. Mayran, A. & Drouin, J. Pioneer transcription factors shape the epigenetic landscape. *J. Biol. Chem.* **293**, 13795–13804 (2018).

38. Pataskar, A. *et al.* NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. *EMBO J*. **35**:24-45(2016).

39. Ding, Y.-Q. *et al.* Lmx1b is essential for the development of serotonergic neurons. *Nat. Neurosci.* **6**, 933–938 (2003).

40. Yoshihara, M. & Littleton, J. T. Synaptotagmin I Functions as a Calcium Sensor to Synchronize Neurotransmitter Release. *Neuron* **36**, 897–908 (2002).

41. Vigneault, É. *et al.* Distribution of vesicular glutamate transporters in the human brain. *Front. Neuroanat.* **9**, 23 (2015).

42. Senft, R. A. & Dymecki, S. M. Neuronal pericellular baskets: neurotransmitter convergence and regulation of network excitability. *Trends Neurosci.* (2021) doi:10.1016/j.tins.2021.08.006.

43. Bradke, F., Giovanni, S. D. & Fawcett, J. Neuronal Maturation: Challenges and Opportunities in a Nascent Field. *Trends Neurosci.* **43**, 360–362 (2020).

44. Aydin, B. *et al.* Proneural factors Ascl1 and Neurog2 contribute to neuronal subtype identities by establishing distinct chromatin landscapes. *Nat. Neurosci.* **22**, 897–908 (2019).

45. Gray, L. T. *et al.* Layer-specific chromatin accessibility landscapes reveal regulatory networks in adult mouse visual cortex. *eLife* **6**, e21883 (2017).

46. Leyva-Díaz, E. & Hobert, O. Transcription factor autoregulation is required for acquisition and maintenance of neuronal identity. *Dev. Camb. Engl.* **146**, dev177378 (2019).

47. Allaway, K. C. *et al.* Genetic and epigenetic coordination of cortical interneuron development. *Nature* **597**, 693–697 (2021).

48. D'Urso, A. & Brickner, J. H. Mechanisms of epigenetic memory. *Trends Genet. TIG* **30**, 230–236 (2014).

49. Spitz, F. & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012).

50. Holmberg, J. & Perlmann, T. Maintaining differentiated cellular identity. *Nat. Rev. Genet.* **13**, 429–439 (2012).

51. Zaret, K. S. & Carroll, J. S. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* **25**, 2227–2241 (2011).

52. Spencer, W. C. & Deneris, E. S. Regulatory Mechanisms Controlling Maturation of Serotonin Neuron Identity and Function. *Front. Cell. Neurosci.* **11**, (2017).

53. Kratsios, P. *et al.* Transcriptional Coordination of Synaptogenesis and Neurotransmitter Signaling. *Curr. Biol.* **25**, 1282–1295 (2015).

54. Suri, D., Teixeira, C. M., Cagliostro, M. K. C., Mahadevia, D. & Ansorge, M. S. Monoamine-Sensitive Developmental Periods Impacting Adult Emotional and Cognitive Behaviors. *Neuropsychopharmacology* **40**, 88–112 (2015).

55. Zhao, Z.-Q. *et al.* Lmx1b Is Required for Maintenance of Central Serotonergic Neurons and Mice Lacking Central Serotonergic System Exhibit Normal Locomotor Activity. *J. Neurosci.* **26**, 12781–12788 (2006).

56. Buenrostro, J., Wu, B., Chang, H. & Greenleaf, W. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr. Protoc. Mol. Biol. Ed. Frederick M Ausubel Al* **109**, 21.29.1-21.29.9 (2015).

57. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019 (2018).

58. Landt, S. G. *et al.* ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* **22**, 1813–1831 (2012).

59. Leek, J. T. svaseq: removing batch effects and other unwanted noise from sequencing data. *Nucleic Acids Res.* **42**, e161 (2014).

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

61. Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z. & Zhang, B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* **47**, W199–W205 (2019).

62. Stuart, T., Srivastava, A., Lareau, C. & Satija, R. *Multimodal single-cell chromatin analysis with Signac*. 2020.11.09.373613 https://www.biorxiv.org/content/10.1101/2020.11.09.373613v1 (2020) doi:10.1101/2020.11.09.373613.

63. Cusanovich, D. A. *et al.* Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910–914 (2015).

64. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).

65. Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902.e21 (2019).

66. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

67. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).

68. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).

69. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv.org* arXiv:1802.03426 (2018).

70. Kim, D., Kim, D., Lee, G., Cho, S.-I. & Kim, J.-S. Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat. Biotechnol.* **37**, 430–435 (2019).

71. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).

72. Roberts, H. C. *et al.* A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. *Age Ageing* **40**, 423–429 (2011).

73. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).

74. Meers, M. P., Tenenbaum, D. & Henikoff, S. Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. *Epigenetics Chromatin* **12**, 42 (2019).

75. Jin, H. *et al.* ChIPseqSpikeInFree: a ChIP-seq normalization approach to reveal global changes in histone modifications without spike-in. *Bioinformatics* **36**, 1270–1272 (2020).

76. Gusmao, E. G., Allhoff, M., Zenke, M. & Costa, I. G. Analysis of computational footprinting methods for DNase sequencing experiments. *Nat. Methods* **13**, 303–309 (2016).

77. Li, Z. *et al.* The bread wheat epigenomic map reveals distinct chromatin architectural and evolutionary features of functional genetic elements. *Genome Biol.* **20**, 139 (2019).

78. Kulakovskiy, I. V. *et al.* HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **46**, D252–D259 (2018).

79. Wickham, H., Chang, W. & Wickham, M. H. Package 'ggplot2'. *Create Elegant Data Vis. Using Gramm. Graph. Version* **2**, 1–189 (2016).

80. Wei, G.-H. *et al.* Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J.* **29**, 2147–2160 (2010).

81. Jolma, A. *et al.* DNA-Binding Specificities of Human Transcription Factors. *Cell* **152**, 327–339 (2013).

Figure 1: Unique distal enhancers and super-enhancers define the *Pet1* lineage.

(a) Principal component analysis reveals the unique accessible chromatin landscape of *Pet1* neurons. Dots of the same color represent biological replicates.

(b) Bar plot showing percentage of unique ATAC-seq peaks (transposase accessible chromatin "TAC" regions) from E14.5 *Pet1* neurons when compared to cis-regulatory elements annotated by ENCODE datasets (Gorkin et al., 2020) (left bar) or with TACs in E14.5 *Pet1*^{neg} cells from the same hindbrain region (right bar).

(c) Emission probabilities for histone modifications in ten ChromHMM states and the percent genomic coverage of each chromatin state for E14.5 *Pet1* neuron specific TACs ("*Pet1* specific") and *Pet1* neuron TACs shared with other tissues in ENCODE ATAC-seq datasets.

(d) Distribution of H3K27ac signal across all enhancers identified in E14.5 *Pet1* neurons. Enhancers with exceptionally high read counts (bracket) were designated super-enhancers.

(e) Bar plot of *Pet1* neuron super-enhancers that are unique to *Pet1* neurons or overlap ("shared") with annotated enhancers from ENCODE datasets.

(f) Genome browser view of the 5-HT neuron specific super-enhancer across the *Lmx1b* locus defined as an extended stretch of H3K27ac and H3K4me1 enrichment.

(g) Genome browser views of ATAC-seq and ChIPmentation signals at the *Pet1/Fev* locus. Red bar denotes the conserved *Pet1* enhancer region that is sufficient to drive 5-HT neuron-specific transgene expression [23,24].

(h) Diagram of 5-HT neurotransmission.

(i-j) Genome browser views of ATAC-seq and ChIPmentation peaks at 5-HT synthesis (*Tph2, Gch1, Gchfr, Ddc*), vesicular transport (*Slc18a2*), synaptic reuptake (*Slc22a3, Slc6a4*), and metabolism (*Maoa/b*) genes. Light blue shading highlights *Pet1* neuron specific TACs.

Supplementary Fig 1, related to Figure 1: ATAC-seq quality control and analysis.

(a) FACS plots of collecting Yfp+ neurons (*Pet1* neurons) from dissociated E14.5 mouse hindbrain cells using forward vs. side-scattered light gating followed by removal of cell doublets. A unique population of cells appear in the hindbrain of *Pet1::EYFP*+ mice compared to *Pet1::EYFP*- mice.

(b) Tapestation bioanalysis showing the nucleosome laddering of a representative E14.5 ATAC-seq sample using 10,000 FACS-purified *Pet1* neurons. Mono, di, and tri nucleosome sizes (sequencing adapters included) are labeled.

(c) Fragment size distribution of ATAC-seq library.

(d) The number of total read pairs per cell type.

(e) ATAC-seq quality metrics showing the number of usable read pairs per cell type after filtering for mapping quality and PCR duplicates.

(f) Library complexity plotted using three ENCODE data standard metrics (NRF, PBC1, PBC2) [58]. Dotted lines indicate threshold cutoff for very high, high, and medium library complexity.

(g) Correlation of ATAC-seq signals at replicated peaks between biological replicated based on Pearson's correlation coefficient (left) or Spearman's correlation coefficient (right).

(h) Annotation of ATAC-seq peak genomic features for E14.5 *Pet1* neurons and *Pet1*^{neg} hindbrain cells.

(i) Average chromatin accessibility at different chromatin states in E14.5 *Pet1* neurons, validating that promoters and enhancers have strong open chromatin signals.

(j) Emission probabilities for histone modifications in ten ChromHMM states and the genomic coverage of each chromatin state for E14.5 *Pet1* neuron specific TACs and *Pet1* neuron TACs shared with E14.5 *Pet1*^{neg} hindbrain cells.

Figure 2: Heterogeneity in chromatin accessibility and transcriptomes reveals transcriptional programs driving *Pet1* neuron subtype identities.

(a) Venn diagram showing overlap of single cell *Pet1* neuron TACs versus bulk ATAC-seq TACs.

(b) Genome browser view of *Pet1* locus showing high concordance between the pseudobulk scATAC-seq and bulk FACS-purified *Pet1* neuron ATAC-seq at E14.5.

(c) scATAC-seq tracks showing the aggregated chromatin accessibility peaks of *En1* (left) and *Meis2* (right) for each cluster.

(d) Uniform manifold approximation and projection (UMAP) visualization of chromatin accessibility of single E14.5 *Pet1*-lineage neurons. scATAC-seq is projected onto scRNA-seq UMAP space. Cells are colored by scATAC-seq cluster assignment. Genes associated with the top two marker TACs for each cluster are labeled.

(e) The chromatin accessibility profiles of select genes across UMAP visualized scATAC-seq clusters.

(f) UMAP visualization of single E14.5 *Pet1* neurons based on gene expression. Cells are colored by scRNA-seq cluster assignment.

(g) The expression profile of example 5-HT neuron identity genes across all scRNA-seq clusters in UMAP.

(h) Representative gene with high number of associated TACs. Loops represent statistically significant TAC-gene association. Loop height represents the p-value of TAC-gene correlation.

(i) Number of TAC-gene associations for all genes.

(j) Number of significantly correlated TACs for each gene. Genes are ranked by the numbers of associated peaks. Select transcription factors with known or potential functions in the *Pet1* lineage are labeled.

Supplementary Figure 2, related to Figure 2: Identification of *Pet1* neuron subtype-specific TACs.

(a) scATAC-seq quality metrics showing the distribution of percent of reads in peaks, number of fragments, transcriptional start site enrichment, blacklist ratio, and nucleosome signal per cell.

(b) Heatmap showing the marker TACs/chromatin accessibility peaks across scATAC-seq clusters.

(c) scATAC-seq tracks showing the aggregated chromatin accessibility peaks at the Tph2 locus for each cluster.

(d) scATAC-seq peaks (TACs) sorted by distance to TSS for all TACs, TACs shared by multiple clusters, and TACs that are unique to an individual cluster. A greater fraction of subtype specific TACs are 50-500kb from the TSS than shared TACs.

(e) Genomic coverage of ChromHMM chromatin states for scATAC-seq TACs that are specific to a single E14.5 *Pet1* subtype/cluster (red) versus shared among multiple subtypes (blue).

(f) Heatmap showing the marker gene expression across scRNA-seq clusters.

(g) scRNA-seq quality control metrics showing the distribution of the number of reads, number of genes, and mitochondrial gene fraction per cell.

Supplementary Figure 3, related to Figure 2: *Pet1* neuron subtype specific TF activity, expression, and accessibility

(a) TACs associated with *Lmx1b*. Loops represent statistically significant TAC-gene association. Loop height represents the p-value of TAC-gene correlation.

(b) Top biological functions gene ontology terms for the 1500 genes with the highest number of associated TACs.

(c) TF motif enrichment of *Pet1* neuron subtype-specific TACs showing subtype-specific TF activities.

(d) Dot plots of gene expression and chromatin accessibility profiles of TFs across *Pet1* neuron subtypes showing a correlation between TF expression and accessibility. Dot size indicates the percentage of cells expressing the gene (left) or exhibiting accessible promoter-proximal chromatin accessibility (right). Color saturation represents the average normalized expression or accessibility level. TFs were selected by their high cluster-dependent variance in expression.

(e) Tph2 immunohistochemistry in adult *Pet1:EYFP* mouse brain showing Tph2 and Yfp overlap in dorsal raphe.

(f) Tph2 immunohistochemistry in E14.5 *Pet1:EYFP* mouse brain showing partial overlap of Tph2 and Yfp in the dorsal raphe and minimal Tph2 protein levels in the median raphe and within the caudal *Pet1* neurons.

Figure 3: Gene regulatory dynamics driving *Pet1* neuron maturation

(a) Principal component analysis showing the distribution of *Pet1* neuron ATAC-seq data over two principal components.

(b) Distribution of chromatin states that exhibit gain (green), loss (red), or no change (blue) in chromatin accessibility from E11.5 to E17.5.

(c) Number of TACs detected at each developmental time point (circles) and the number of TACs that are gained, lost, or propagated to the next time point (arrows).

(d) Volcano plots showing the differential chromatin accessibility of TACs between consecutive time points, with fold change cutoff of 2-fold and significance cutoff of FDR <0.01. Each dot represents one ATAC-seq peak (TAC).

(e) Genome browser tracks of genes showing loss (left) or gain (right) of chromatin accessibility from E11.5 to E17.5.

(f) Box plots showing the distribution of the log2 expression values of the linked genes for TACs that exhibit gain, no change, or loss in accessibility between E11.5 and E17.5.

Significance of accessibility to gene expression correlation is calculated by Wilcoxon ranks sum test.

(g) Genome browser view of ATAC-seq and E14.5 ChIPmentation signals at representative genes that decrease (e.g. *Ascl1*) or increase (e.g. *Ntrk2*) in accessibility between E11.5 and E17.5. Parallel trends are seen in gene expression trajectories (panels on the right of tracks).

(h) Top five gene ontology biological process enrichment terms for genes linked to TACs that lose (left) or gain (right) accessibility from E11.5 to E17.5.

Supplementary Figure 4, related to Figure 3: Dynamic developmental remodeling of distal cis-regulatory elements

(a) ATAC-seq quality metrics showing the distribution of total reads (left) and usable reads after filtering for mapping quality and PCR duplicates (right) for each developmental time point.

(b) ATAC-seq quality metrics showing library complexity plotted using three ENCODE data standards (NRF, PBC1, PBC2) for each developmental time point [58]. Dotted lines indicate threshold cutoff for very high, high, and medium library complexity.

(c) Distribution of gene feature annotations for all TAC regions at each developmental time point.

(d) Distribution of gene feature annotations for TACs that show gain or loss in accessibility between developmental time points.

(e) Histone ChIPmentation signals (E14.5) at TACs regions that show gain in chromatin accessibility from E11.5 to E17.5.

(f) Histone ChIPmentation signals (E14.5) at TACs regions that show loss in chromatin accessibility from E11.5 to E17.5.

(g) Histone ChIPmentation signals (E14.5) at TACs regions that show no change in chromatin accessibility from E11.5 to E17.5.

(h) Overlap of all *Pet1* neuron super-enhancers (green) with gained (blue) or lost (red) TACs from E11.5 to E17.5.

(i) Motif enrichment within differentially accessible TACs. Highly similar motif hits are combined and assigned to their closest match of a TF class or family. TF families are sorted into three groups by whether their motifs are most significantly enriched in LOSS, NO CHANGE, or GAIN TACs. The TF families are then ranked by fold enrichment within each group. Fold enrichment and p-value are calculated and visualized by the Bioconductor package monaLisa.

Figure 4: Pet1 organizes postmitotic chromatin accessibility

(a) Principal component analysis showing the distribution of ATAC-seq data for wildtype and *Pet1-/-* *Pet1*-lineage neurons (*Pet1-/-* neurons) over two principal components.

(b) *In vivo* Pet1 footprints derived from Tn5 insertion frequency (ATAC-seq reads) over representative ETS-factor motifs within *Pet1* neuron accessible chromatin regions. *Pet1-/-* neurons show reduced Pet1 footprints at E11.5 (left), E14.5 (middle), and E17.5 (right) compared to wildtype neurons.

(c) Box plots showing the distribution of the log2 fold change in accessibility of all TACs in *Pet1-/-* vs. wildtype *Pet1* neurons at three developmental time points.

(d) The top significantly enriched motif within Pet1 CUT&RUN peaks as determined by *de novo* MEME motif analysis significantly matches (p=5.33E-06) the Pet1/FEV high-affinity binding site previously defined *in vitro* [80] based on analysis by the TOMTOM Motif Comparison Tool (top panel).

(e) The fraction of Pet1-regulated TACs containing Pet1 CUT&RUN peaks.

(f) The fraction of all *Pet1* neuron TACs, *Pet1* neuron TACs that are dependent on Pet1, and Pet1-dependent TACs occupied by Pet1 that are specific to *Pet1*/5-HT lineage or shared with *Pet1*^{neg} cells. A significant fraction of Pet1-regulated TACs is unique to *Pet1* neurons and absent from *Pet1*^{neg} cells.

(g) Volcano plot showing the differential chromatin accessibility of TACs between *Pet1-/-* and wildtype *Pet1* neurons at E14.5, with fold change cutoff of 2-fold and significance cutoff of FDR <0.01 (top). Each dot represents one TAC. Highest FDR GO terms for

Pet1-dependent TACs (in red) that are occupied by Pet1 show strong enrichment for serotonin biosynthesis. Select genes required for 5-HT synthesis are labeled on the volcano plot.

(h-i) Genome browser tracks showing the gene-enhancer pairings and the ATAC-seq, Pet1 CUT&RUN, and histone modification ChIPmentation signals at 5-HT identity genes required for 5-HT synthesis (*Tph2, Gch1, Gchfr, Ddc*), transport (*Slc18a2*), synaptic reuptake (*Slc22a3, Slc6a4*), and metabolism (*Maoa/b*). Zoomed-in inserts display the changes in chromatin accessibility between wildtype (red) and *Pet1-/-* (green) *Pet1* neurons within the highlighted (blue) regions at three developmental time points. Two TACs of the gene *Slc6a4* (red dashed line box) gain accessibility in wildtype 5-HT neurons from E11.5 to E17.5 but fail to open to the same extent in *Pet1-/-* neurons. Asterisks denote FDR<0.01 and fold change >2 relative to wildtype at the same developmental time point.

Supplementary Fig 5, related to Figures 4 and 5: Pet1 is required for chromatin accessibility at *Pet1* neuron specific TACs

(a) Motif enrichment within *Pet1* neuron specific TACs using the Bioconductor package monaLisa. Similar motif hits are combined and assigned to their closest match of a TF class or family.

(b) ATAC-seq quality metrics showing the distribution of total reads (left) and usable reads after filtering for mapping quality and PCR duplicates (right) for *Pet1-/-* neurons at E11.5, E14.5, and E17.5.

(c) ATAC-seq quality metrics showing library complexity plotted using three ENCODE data standards (NRF, PBC1, PBC2) [58] for *Pet1-/-* neurons assayed at E11.5, E14.5, and E17.5. Dotted lines indicate threshold cutoff for very high, high, and medium library complexity.

(d) The ChromHMM chromatin states genomic coverage for TACs that lose accessibility in *Pet1-/-* neurons at E11.5, E14.5, or E17.5.

(e) Genome browser tracks of Pet1 CUT&RUN and previously published ChIP-seq using myc-Pet1 [20] showing the similarity of Pet1 occupancy peaks at Pet1-regulated genes.

(f) GO terms for Pet1-occupied differentially accessible TACs between *Pet1-/-* and wildtype *Pet1* neurons at E11.5 (top) and E17.5 (bottom) show enrichment for serotonin biosynthesis.

(g) The fraction of TACs that require Pet1 for chromatin opening from E11.5 to E17.5 that contains Pet1 CUT&RUN peaks (top). Gene ontology analysis on TACs that require Pet1 for initial chromatin accessibility shows enrichment of synapse-related terms (bottom).

(h) Representative genome browser tracks of TACs that require Pet1 for euchromatin formation.

(i) Schematic of conditional Pet1 targeting and labeling using *Pet1-Cre* transgene followed by FACS collection of TdTomato+ neurons for ATAC-seq.

(j) ATAC-seq quality metrics showing the distribution of total reads (left) and usable reads after filtering for mapping quality and PCR duplicates (right) for Pet1-cKO 5-HT precursors at E14.5.

(k) ATAC-seq quality metrics showing library complexity plotted using three ENCODE data standards (NRF, PBC1, PBC2) (Landt et al., 2012) for Pet1-cKO 5-HT precursors at E14.5.

(l) Fraction of Pet1-maintained TACs (differentially accessible TACs in Pet1-cKO) containing Pet1 CUT&RUN peaks.

Figure 5: Sustained Pet1 input is required for chromatin accessibility

(a) Volcano plot showing the differential chromatin accessibility of TACs between Pet1-cKO and wildtype *Pet1* neurons, with fold change cutoff of 2-fold and significance cutoff of FDR <0.01 (top). Highest FDR GO terms for the Pet1-maintained TACs (in red) that are occupied by Pet1 show enrichment for serotonin biosynthesis (bottom).

(b) Overlap between differential TACs in *Pet1-/-* and Pet1-cKO 5-HT neurons.

(c) Comparison of *in vivo* Pet1 footprints derived from Tn5 insertion frequency (ATAC-seq reads) over representative ETS-factor motif between Pet1-cKO and wildtype 5-HT precursors within all *Pet1*/5-HT-lineage accessible chromatin regions.

(d) The ChromHMM chromatin states genomic coverage for TACs that are maintained by Pet1 at E14.5.

(e-f) Genome browser tracks showing the gene-enhancer pairings and the ATAC-seq, Pet1 CUT&RUN, and histone modification ChIPmentation signals at 5-HT identity genes. Zoomed-in inserts display the changes in chromatin accessibility between wildtype (red) and Pet1-cKO (yellow) 5-HT precursors within the highlighted (blue) regions at E14.5. Asterisks denote FDR<0.01 and fold change >2 compared to wildtype.

(g) Genome browser tracks showing the ATAC-seq, Pet1 CUT&RUN, and histone modification ChIPmentation signals at the *Pet1* gene locus.

Figure 6: Lmx1b co-regulates chromatin accessibility with Pet1 at 5-HT neurotransmission genes but has broader impact on accessibility landscape

(a) Principal component analysis showing the distribution of wildtype, *Pet1-/-*, Pet1-cKO, Lmx1b-cKO, and DKO neuron ATAC-seq data over two principal components.

(b) Comparison of *in vivo* Lmx1b footprints, derived from Tn5 insertion frequency (ATAC-seq reads) over representative LIM-homedomain motif, between Lmx1b-cKO and wildtype *Pet1*/5-HT neurons within accessible chromatin regions.

(c) Volcano plot showing the differential chromatin accessibility of TACs between Lmx1b-cKO and wildtype *Pet1* neurons, with fold change cutoff of 2-fold and significance cutoff of FDR <0.01 (top). Highest FDR GO terms for Lmx1b-maintained TACs (in red) that are occupied by Lmx1b show enrichment of serotonin biosynthesis (bottom).

(d) The top most significantly enriched motif within Pet1 CUT&RUN peaks as determined by *de novo* MEME motif analysis matches (*p=4.03E-02*) the Lmx1b LIM-homeodomain binding site previously defined *in vitro* [81] based on analysis by the TOMTOM Motif Comparison Tool (left panel). Fraction of Lmx1b-maintained TACs containing Lmx1b CUT&RUN peaks (right panel).

(e) The ChromHMM chromatin states genomic coverage for TACs that lose accessibility in Lmx1b-cKO (blue) and DKO (red) in E14.5 *Pet1* neurons.

(f) Comparison of the *in vivo* Pet1 (top) and Lmx1b (bottom) footprints between DKO and wildtype *Pet1*/5-HT lineage neurons.

(g) Volcano plot showing the differential chromatin accessibility of TACs between DKO and wildtype 5-HT precursors, with fold change cutoff of 2-fold and significance cutoff of FDR <0.01 (top). Highest FDR GO terms for Pet1- and Lmx1b-maintained TACs (in red) that are occupied by the two TFs show enrichment for serotonin biosynthesis (bottom).

(h) The fraction of Pet1-cKO, Lmx1b-cKO, and DKO differentially accessible TACs that is specific to *Pet1*/5-HT lineage. A significant fraction of the Pet1- and Lmx1b-regulated TACs is unique to *Pet1* neurons and absent from *Pet1*^{neg} cells.

(i-j) Genome browser tracks showing the gene-enhancer pairings and the ATAC-seq, Pet1 CUT&RUN, and histone modification ChIPmentation signals at 5-HT identity genes. Display data ranges for the zoomed-out panels are scaled to the peak in the blue highlighted regions. Zoomed-in inserts display the changes in chromatin accessibility between wildtype (red), Lmx1b-cKO (yellow), and DKO (yellow) 5-HT precursors within

the highlighted (blue) regions at E14.5. Asterisks denote FDR<0.01 and fold change >2 compared to wildtype.

Supplementary Fig 6, related to Figure 6: Pet1 and Lmx1b control the global landscape of TF binding sites

(a) Distribution of gene features for E14.5 Pet1 and Lmx1b CUT&RUN peaks using either “stringent” or “relaxed” peak calling parameters [74].

(b) Overlap of Pet1 and Lmx1b CUT&RUN binding sites defined by either “stringent” or “relaxed” peak selection mode [74].

(c) Overlap of “stringent” and “relaxed” E14.5 Pet1 and Lmx1b CUT&RUN peaks with E14.5 ATAC-seq TACs showing that Pet1 and Lmx1b mostly occupy open chromatin.

(d) Schematic of conditional Lmx1b targeting and labeling using *Pet1-Cre* transgene followed by FACS collection of TdTomato+ neurons for ATAC-seq.

(e) ATAC-seq quality metrics showing the distribution of total reads (left) and usable reads after filtering for mapping quality and PCR duplicates (right) for Lmx1b-cKO and DKO at E14.5.

(f) ATAC-seq quality metrics showing library complexity plotted using three ENCODE data standards (NRF, PBC1, PBC2) [58] for Lmx1b-cKO and DKO 5-HT precursors assayed at E14.5.

(g) Principal component analysis showing the distribution of E14.5 wildtype, Pet1-cKO, Lmx1b-cKO, and DKO neuron ATAC-seq data over two principal components.

(h) Significant gene ontology terms for the LOSS TACs in Lmx1b-cKO that are not significantly decreased in accessibility in Pet1-cKO.

(i) The number of TACs that are differentially accessible between Pet1-cKO, Lmx1b-cKO, and DKO 5-HT neurons by fold change >2, FDR <0.01 (left). Overlap of TACs showing loss (middle) or gain (right) of chromatin accessibility in Pet1-cKO, Lmx1b-cKO, and DKO *Pet1* neurons relative to wildtype.

(j) Top enriched motifs within TACs showing loss of accessibility in Pet1-cKO (top), Lmx1b-cKO (middle), or DKO (bottom) neurons relative to wildtype. Highly similar motif hits are combined and assigned to their closest match of a TF class or family. The TF families are then ranked by fold enrichment. Fold enrichment and p-value are calculated and visualized by the Bioconductor package monaLisa.

Figure 7: Loss of Pet1 and Lmx1b alters chromatin accessibility at synaptic gene enhancers and disrupts 5-HT neuron connectivity

(a) Gene ontology analysis of all Pet1- and Lmx1b-maintained TAC regions, depicted as volcano plot of enrichment (x-axis) vs. FDR statistical significance (y-axis), showing many neuronal biological processes among the top enriched annotations.

(b) Genome browser tracks showing the gene-enhancer pairings and the ATAC-seq, Pet1 and Lmx1b CUT&RUN, and histone modification ChIPmentation signals at synaptic

genes *Syt1* and *Slc17a8* show changes in chromatin accessibility between wildtype (red), Pet1-cKO (yellow), Lmx1b-cKO (yellow), and DKO (yellow) *Pet1* neurons within the highlighted (blue) regions at E14.5. Display data ranges for the zoomed-out panels are optimized based on the blue highlighted regions.

(c) Relative expression (FPKMs) of *Syt1* and *Slc17a8* in rostral Pet1-cKO, Lmx1b-cKO, and DKO neurons at E17.5. Data presented as mean ± SEM.

(d) Differential expression (from E11.5 to E17.5) of the genes associated with Pet1-occupied TACs in Pet1-cKO vs. wildtype 5-HT neurons (left), genes associated with Pet1-occupied TACs in Lmx1b-cKO vs. wildtype 5-HT neurons (middle), and genes associated with Pet1 and Lmx1b co-occupied TACs in DKO vs. wildtype 5-HT neurons (right).

(e) Fraction of the Pet1- and/or Lmx1b-maintained TACs that change (blue) or gain (red) in accessibility in wildtype *Pet1* neurons from E11.5 to E17.5.

(f) TdTomato (red) and DAPI (blue) co-staining of pericellular baskets in the medial septum of P15 *Pet1-Cre; Ai9* control, Pet1-cKO, Lmx1b-cKO, and DKO mice.

(g) Diagram showing the analyzed brain area and quantification of TdTomato axons (pixels/μm^2) (Two-way ANOVA; n = three animals per genotype and two quantified brain sections per animal) (bottom panels). Images were generated through automatic stitching of individual 20x images.

(h) Representative immunohistochemistry images of synaptotagmin1 (*Syt1*), TdTomato, and DAPI in the medial septum of P15 *Pet1-Cre; Ai9* mice showing the decreased overlap of Syt1 (green) and TdTomato (red) signals in KO animals compared to wildtype, Images were generated through automatic stitching of individual 63x images.

(i) Representative immunohistochemistry images of vGlut3 (*Slc17a8*), TdTomato, and DAPI in the medial septum of adult *Pet1-Cre; Ai9* mice showing the decreased overlap of vGlut3 (green) and TdTomato (red) signals in KO animals compared to wildtype. Images were generated through automatic stitching of individual 63x images.

(j) Quantification of Syt1 (left) and vGlut3 (right) co-localization with TdTomato in the medial septum as a percentage of the total TdTomato signal. (Two-way ANOVA; n = two animals per genotype and 3-4 quantified images per animal).

Supplementary Fig 7, related to Figure 7: Pet1 and Lmx1b control chromatin access to synapse and axon gene TACs

(a) Top gene ontology terms for the differentially accessible TACs in Pet1-cKO (top), Lmx1b-cKO (middle), and DKO *Pet1* neurons (bottom) ranked by fold enrichment.

(b) Genome browser tracks showing the gene-enhancer pairings and the ATAC-seq, Pet1 and Lmx1b CUT&RUN, and histone modification ChIPmentation signals at synaptic gene *Sorcs3.* Display data range is scaled to the peak in the blue highlighted region.

(c) Relative expression (FPKMs) of *Sorcs3* in rostral Pet1-cKO, Lmx1b-cKO, and DKO neurons at E17.5. Data presented as mean ± SEM.

(d) Genome browser tracks showing the gene-enhancer pairings and the ATAC-seq, Pet1 and Lmx1b CUT&RUN, and histone modification ChIPmentation signals at select axon related genes *Dclk1, Cdh13,* and *Raph1.* Display data ranges are scaled to the peaks in the blue highlighted regions.

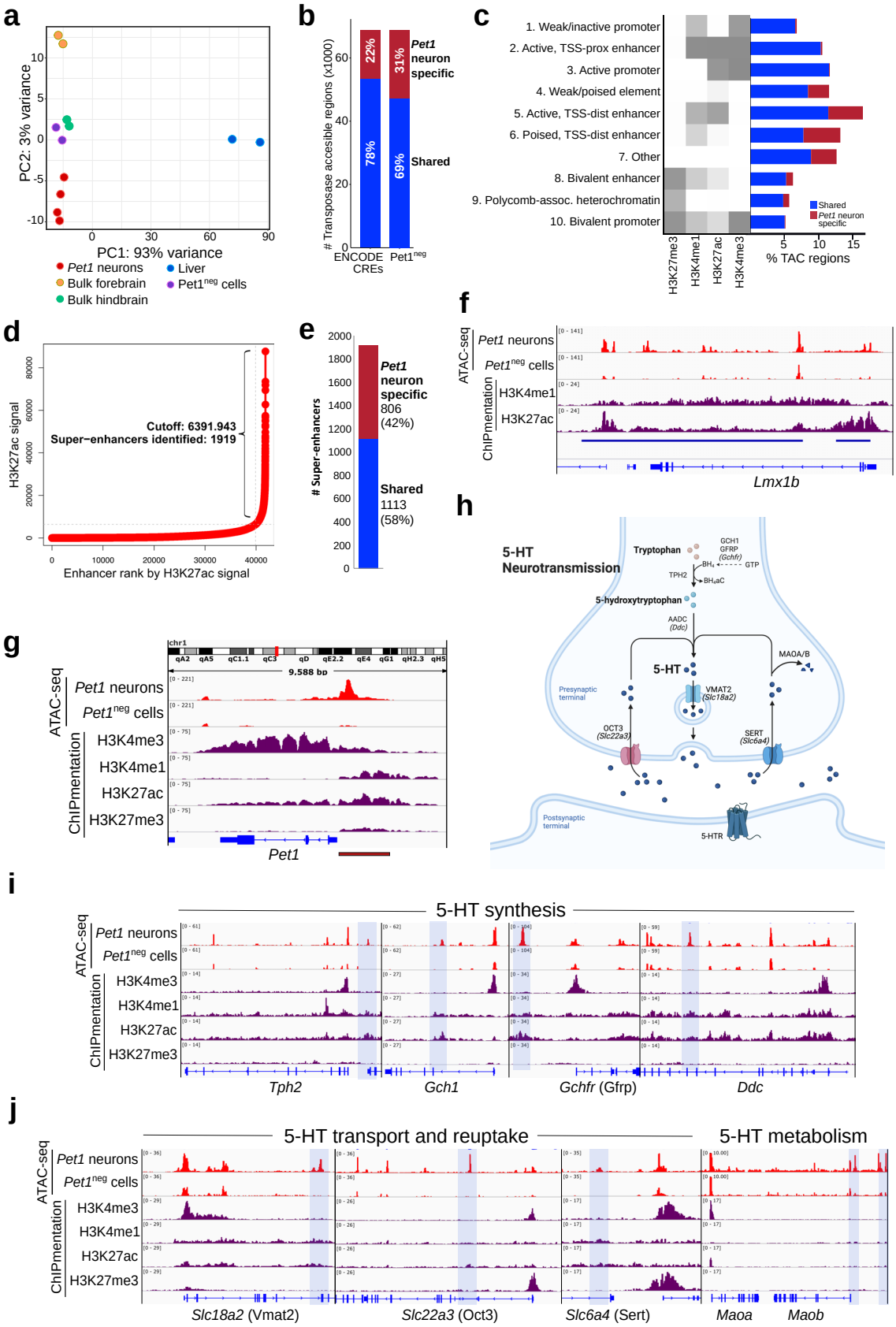
(e) Relative expression (FPKMs) of *Dclk1, Cdh13,* and *Raph1* in rostral Pet1-cKO, Lmx1b-cKO, and DKO neurons at E17.5. Data presented as mean ± SEM.

(f) Genome browser tracks showing the gene-enhancer pairings and the ATAC-seq, Pet1 CUT&RUN, and histone modification ChIPmentation signals at synaptic genes that require Pet1 for chromatin opening. Zoomed-in displays show changes in chromatin accessibility between wildtype (red) and *Pet1-/-* (green) 5-HT precursors within the highlighted (blue) regions across developmental time points. Display data ranges for the top zoomed-out panels are scaled to the peak in the blue highlighted regions.

(g) Additional immunohistochemistry images of synaptotagmin1 (*Syt1*), TdTomato, and DAPI in the medial septum of P15 *Pet1-Cre; Ai9* mice showing the decreased overlap of Syt1 (green) and TdTomato (red) signals in KO animals compared to wildtype, Images were generated through automatic stitching of individual 63x images.

(h) Additional immunohistochemistry images of vGlut3 (*Slc17a8*), TdTomato, and DAPI in the medial septum of P15 *Pet1-Cre; Ai9* mice showing the decreased overlap of vGlut3 (green) and TdTomato (red) signals in KO animals compared to wildtype, Images were generated through automatic stitching of individual 63x images.

Figure 1


a
PC2: 3% variance
PC1: 93% variance
Pet1 neurons
Liver
Bulk forebrain
Pet1neg cells
Bulk hindbrain
b
Transposase accesible regions (x1000)
22%
78%
31%
69%
Pet1 neuron specific
Shared
ENCODE CREs
Pet1neg
c
1. Weak/inactive promoter
2. Active, TSS-prox enhancer
3. Active promoter
4. Weak/poised element
5. Active, TSS-dist enhancer
6. Poised, TSS-dist enhancer
7. Other
8. Bivalent enhancer
9. Polycomb-assoc. heterochromatin
10. Bivalent promoter
H3K27me3
H3K4me1
H3K27ac
H3K4me3
Shared
Pet1 neuron specific
% TAC regions
d
H3K27ac signal
Cutoff: 6391.943
Super−enhancers identified: 1919
Enhancer rank by H3K27ac signal
e
Super-enhancers
Pet1 neuron specific
806
(42%)
Shared
1113
(58%)
f
ATAC-seq
Pet1 neurons
Pet1neg cells
ChIPmentation
H3K4me1
H3K27ac
Lmx1b
g
chr1
9.588 bp
ATAC-seq
Pet1 neurons
Pet1neg cells
ChIPmentation
H3K4me3
H3K4me1
H3K27ac
H3K27me3
Pet1
h
5-HT Neurotransmission
Tryptophan
GCH1
GFRP
(Gchfr)
GTP
TPH2
5-hydroxytryptophan
AADC
(Ddc)
5-HT
MAOA/B
Presynaptic terminal
VMAT2
(Slc18a2)
OCT3
(Slc22a3)
SERT
(Slc6a4)
Postsynaptic terminal
5-HTR
i
5-HT synthesis
ATAC-seq
Pet1 neurons
Pet1neg cells
ChIPmentation
H3K4me3
H3K4me1
H3K27ac
H3K27me3
Tph2
Gch1
Gchfr (Gfrp)
Ddc
j
5-HT transport and reuptake
5-HT metabolism
ATAC-seq
Pet1 neurons
Pet1neg cells
ChIPmentation
H3K4me3
H3K4me1
H3K27ac
H3K27me3
Slc18a2 (Vmat2)
Slc22a3 (Oct3)
Slc6a4 (Sert)
Maoa
Maob

Supplementary Fig 1

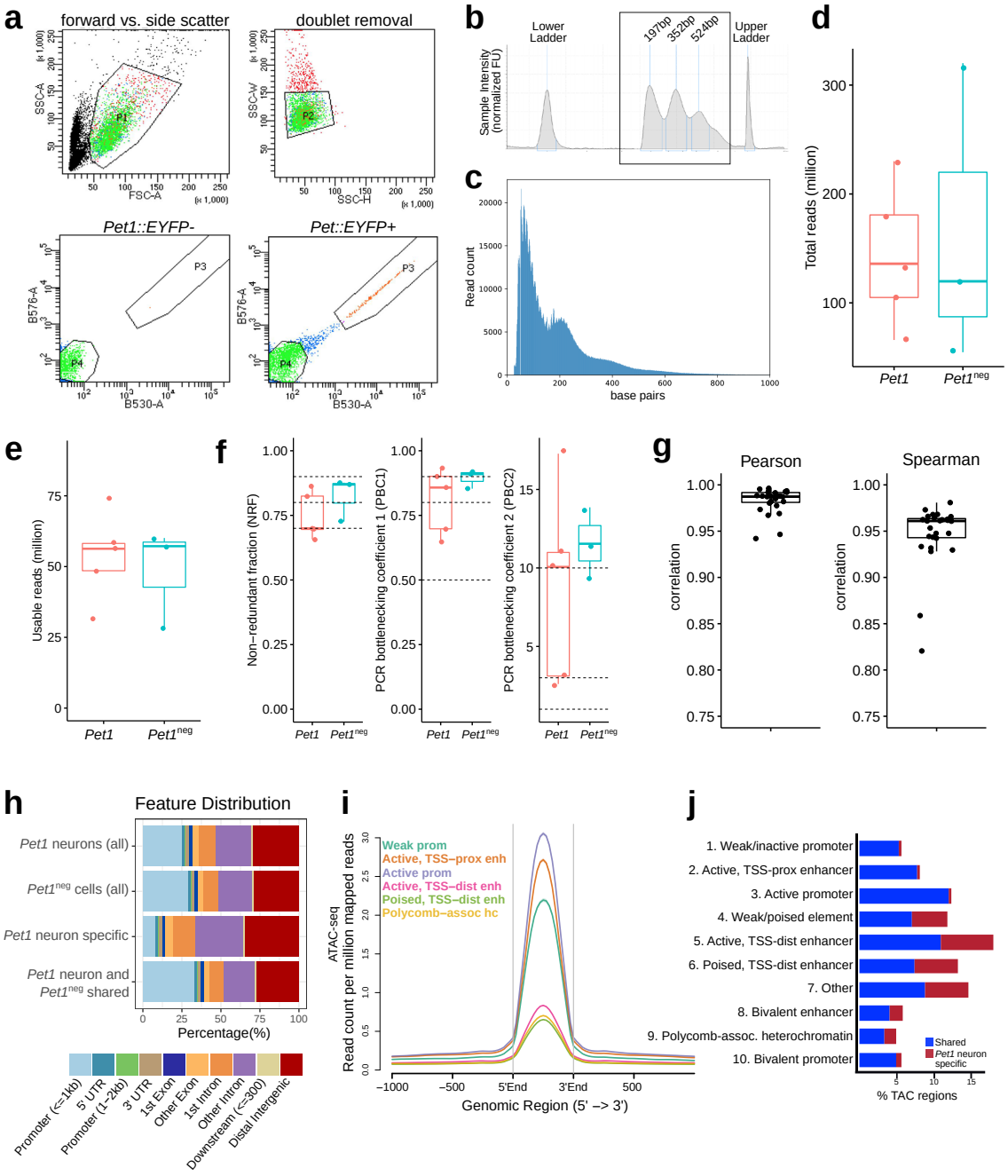
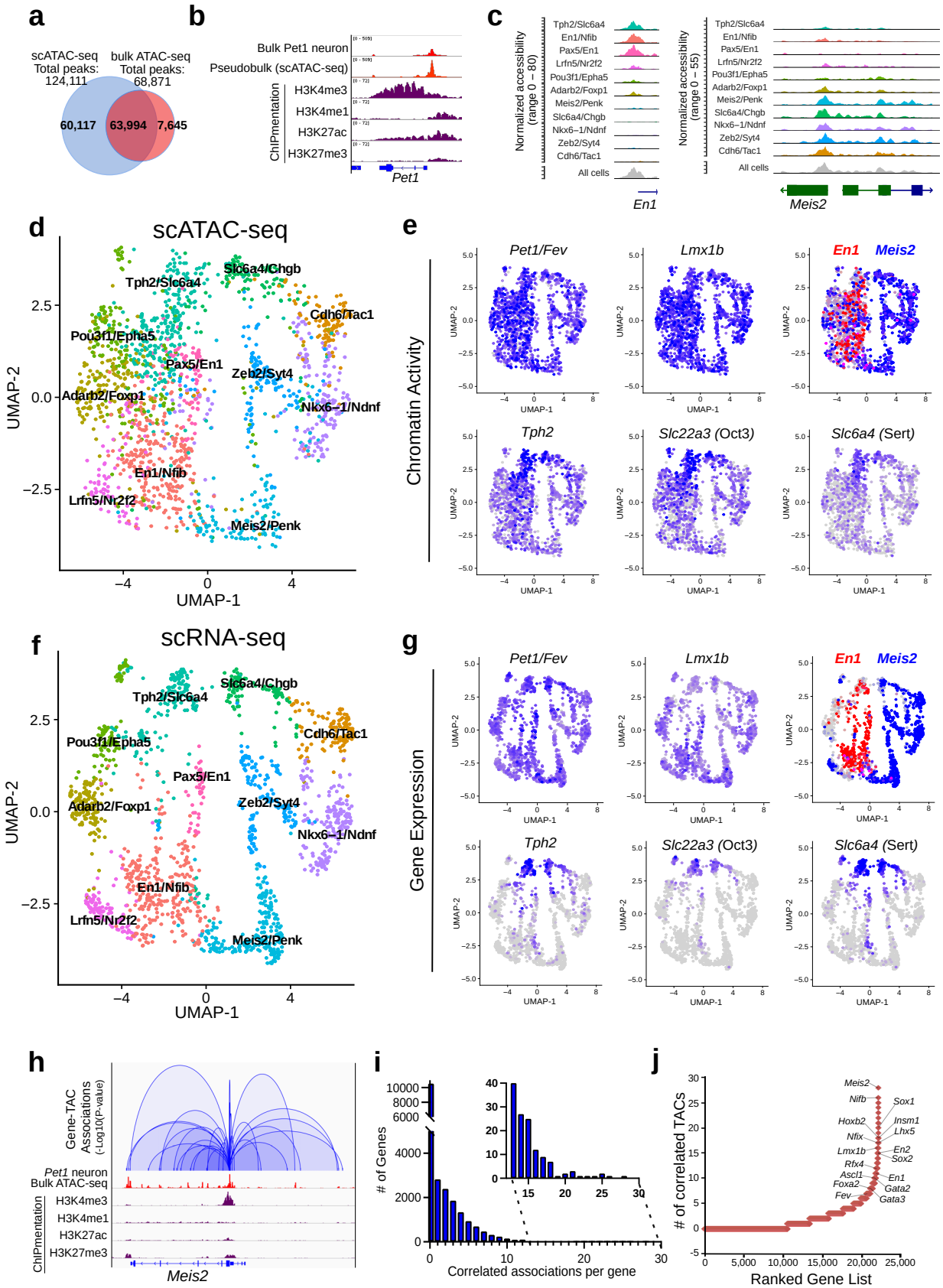
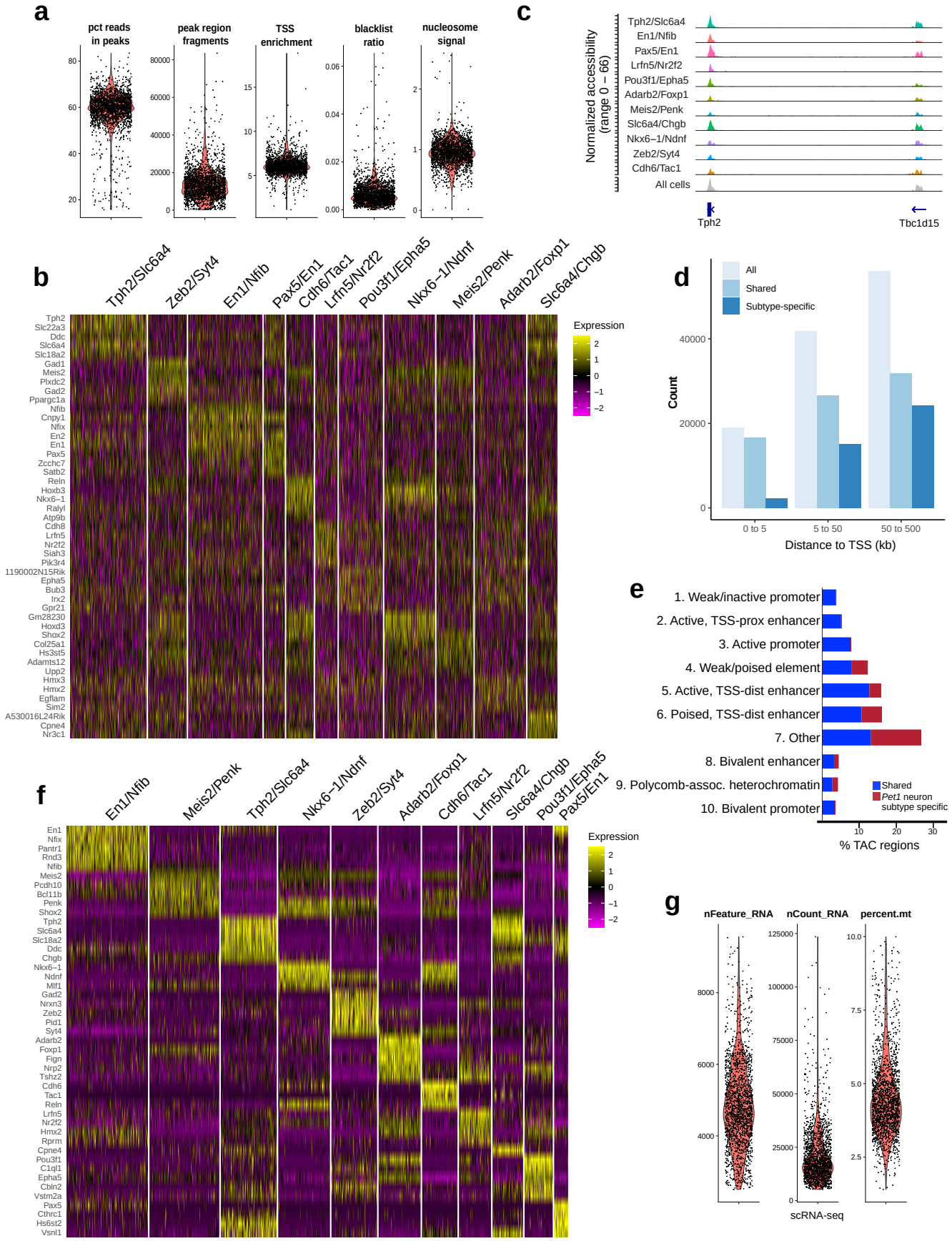


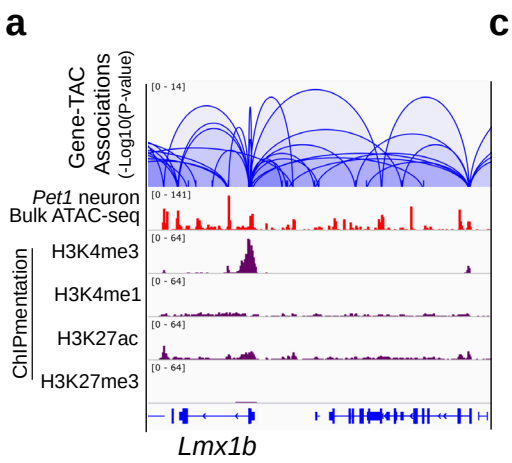
Figure 2


a
scATAC-seq Total peaks: 124,111
bulk ATAC-seq Total peaks: 68,871
60,117
63,994
7,645
b
Bulk Pet1 neuron
Pseudobulk (scATAC-seq)
ChIPmentation
H3K4me3
H3K4me1
H3K27ac
H3K27me3
Pet1
c
Normalized accessibility (range 0 – 80)
Normalized accessibility (range 0 – 55)
Tph2/Slc6a4
En1/Nfib
Pax5/En1
Lrfn5/Nr2f2
Pou3f1/Epha5
Adarb2/Foxp1
Meis2/Penk
Slc6a4/Chgb
Nkx6–1/Ndnf
Zeb2/Syt4
Cdh6/Tac1
All cells
En1
Meis2
d
scATAC-seq
UMAP-2
UMAP-1
e
Chromatin Activity
Pet1/Fev
Lmx1b
En1
Meis2
Tph2
Slc22a3 (Oct3)
Slc6a4 (Sert)
f
scRNA-seq
g
Gene Expression
h
Gene-TAC Associations (-Log10(P-value))
Pet1 neuron Bulk ATAC-seq
ChIPmentation
Meis2
i
# of Genes
Correlated associations per gene
j
# of correlated TACs
Ranked Gene List
Meis2
Nifb
Sox1
Hoxb2
Insm1
Nfix
Lhx5
Lmx1b
En2
Sox2
Rfx4
Ascl1
En1
Foxa2
Gata2
Fev
Gata3

Supplementary Fig 2

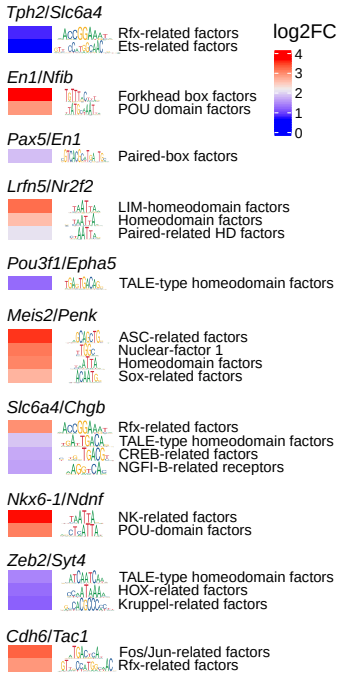


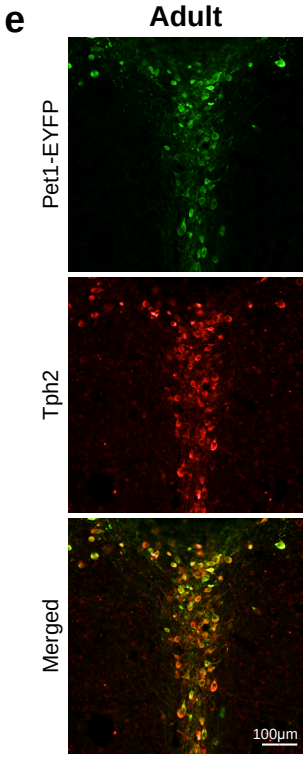
Supplementary Fig 3

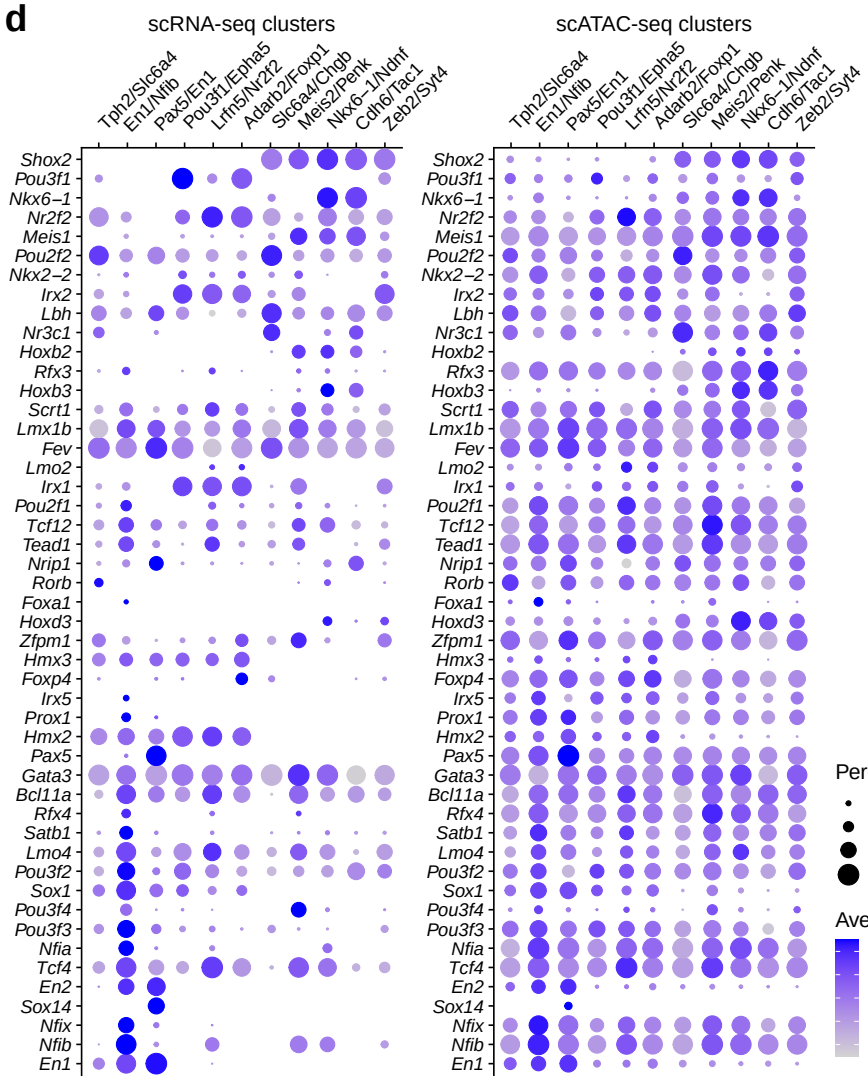


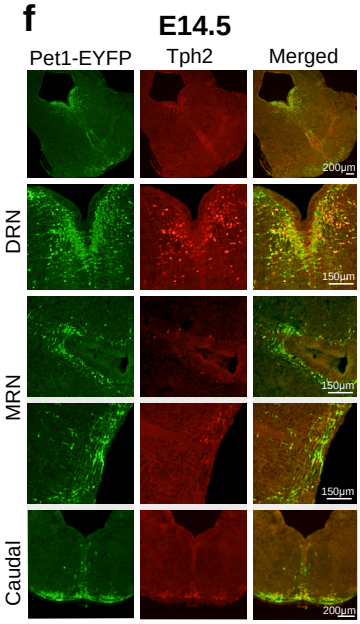
b

Gene Set	Description	FDR
GO:0051961	neg regulation of nervous system development	1.96E-09
GO:0061564	axon development	1.96E-09
GO:0030902	hindbrain development	1.96E-09
GO:0010721	negative regulation of cell development	7.78E-09
GO:0097485	neuron projection guidance	1.34E-08
GO:0099177	regulation of trans-synaptic signaling	1.40E-07









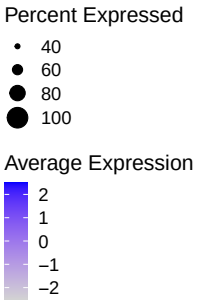
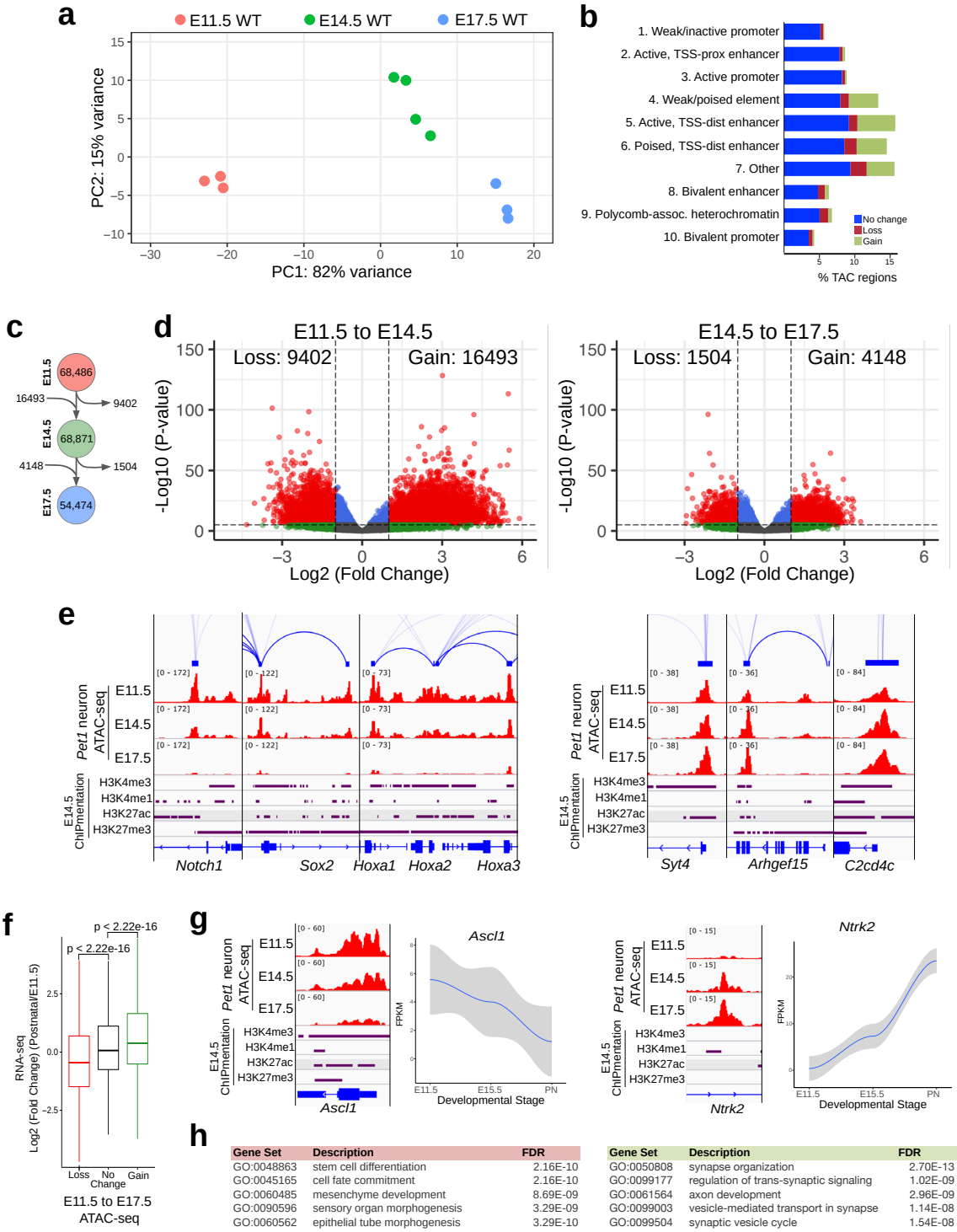


Figure 3



Gene Set	Description	FDR
GO:0048863	stem cell differentiation	2.16E-10
GO:0045165	cell fate commitment	2.16E-10
GO:0060485	mesenchyme development	8.69E-09
GO:0090596	sensory organ morphogenesis	3.29E-09
GO:0060562	epithelial tube morphogenesis	3.29E-10

Gene Set	Description	FDR
GO:0050808	synapse organization	2.70E-13
GO:0099177	regulation of trans-synaptic signaling	1.02E-09
GO:0061564	axon development	2.96E-09
GO:0099003	vesicle-mediated transport in synapse	1.14E-08
GO:0099504	synaptic vesicle cycle	1.54E-08

Supplementary Fig 4

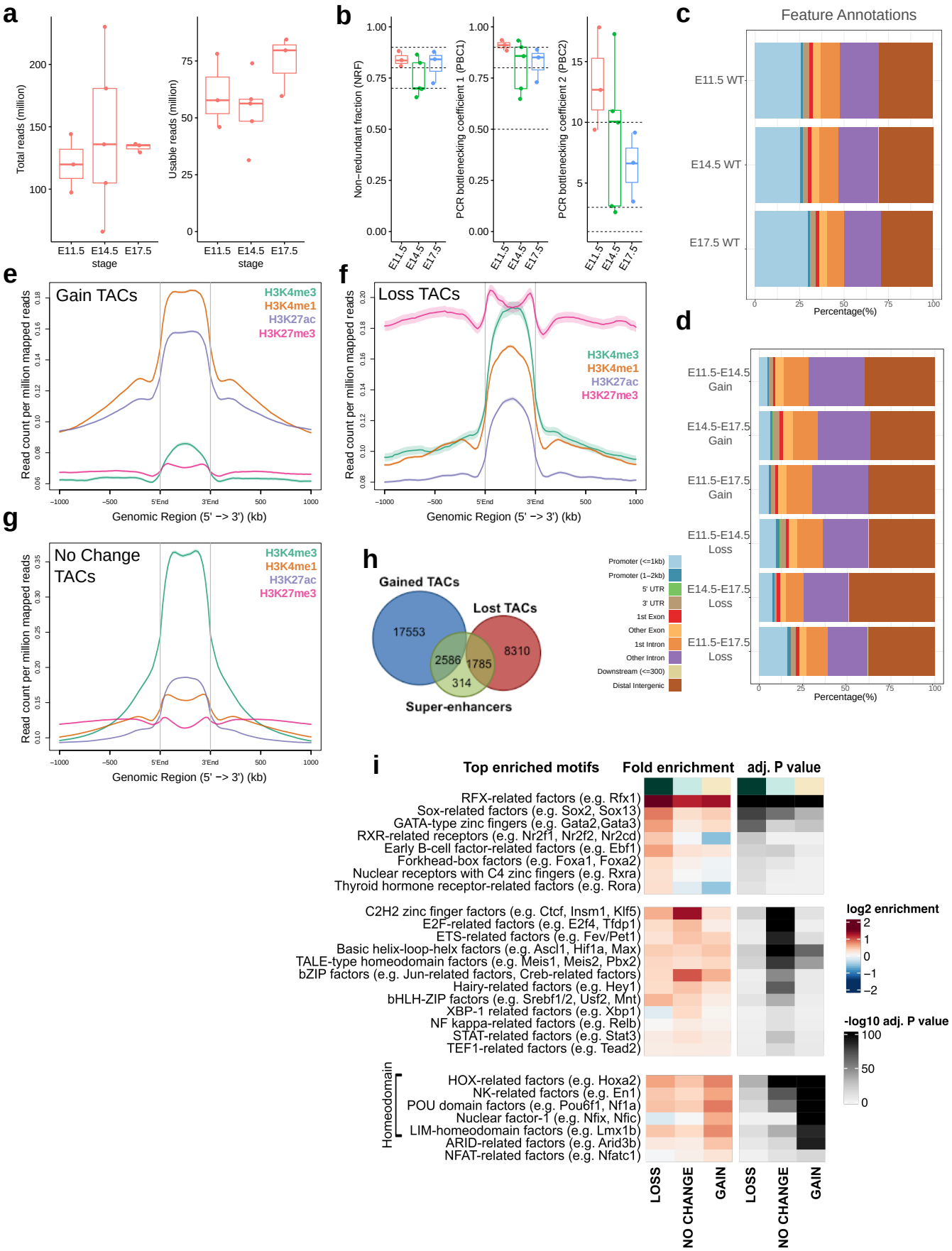
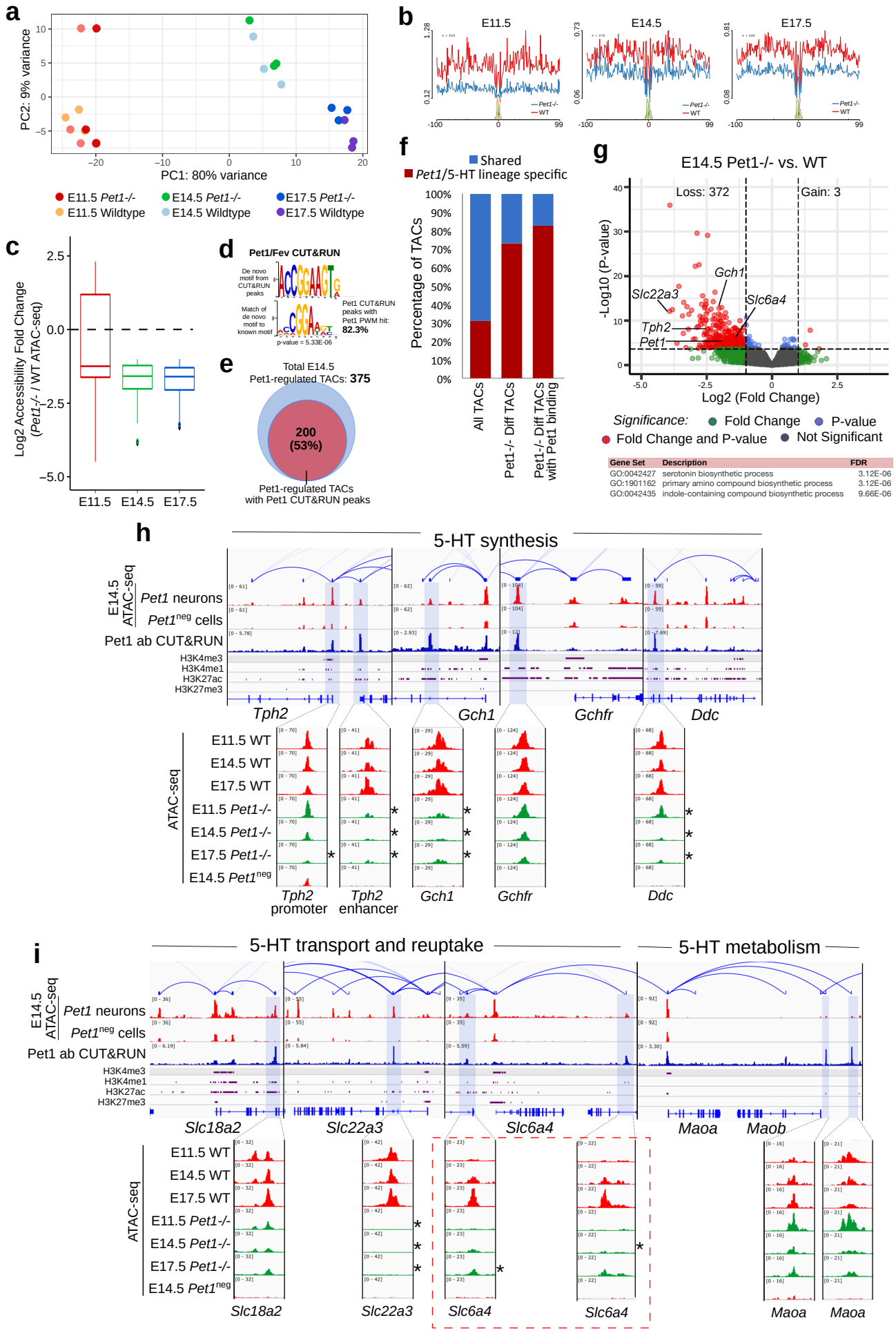


Figure 4


a
PC2: 9% variance
PC1: 80% variance
E11.5 Pet1-/-
E11.5 Wildtype
E14.5 Pet1-/-
E14.5 Wildtype
E17.5 Pet1-/-
E17.5 Wildtype
b
E11.5
E14.5
E17.5
Pet1-/-
WT
c
Log2 Accessibility Fold Change (Pet1-/- / WT ATAC-seq)
E11.5
E14.5
E17.5
d
Pet1/Fev CUT&RUN
De novo motif from CUT&RUN peaks
Match of de novo motif to known motif
Pet1 CUT&RUN peaks with Pet1 PWM hit: 82.3%
p-value = 5.33E-06
e
Total E14.5 Pet1-regulated TACs: 375
200 (53%)
Pet1-regulated TACs with Pet1 CUT&RUN peaks
f
Shared
Pet1/5-HT lineage specific
Percentage of TACs
All TACs
Pet1-/- Diff TACs
Pet1-/- Diff TACs with Pet1 binding
g
E14.5 Pet1-/- vs. WT
Loss: 372
Gain: 3
-Log10 (P-value)
Log2 (Fold Change)
Gch1
Slc22a3
Slc6a4
Tph2
Pet1
Significance:
Fold Change
P-value
Fold Change and P-value
Not Significant
Gene Set | Description | FDR
GO:0042427 | serotonin biosynthetic process | 3.12E-06
GO:1901162 | primary amino compound biosynthetic process | 3.12E-06
GO:0042435 | indole-containing compound biosynthetic process | 9.66E-06
h
5-HT synthesis
E14.5 ATAC-seq
Pet1 neurons
Pet1neg cells
Pet1 ab CUT&RUN
H3K4me3
H3K4me1
H3K27ac
H3K27me3
Tph2
Gch1
Gchfr
Ddc
ATAC-seq
E11.5 WT
E14.5 WT
E17.5 WT
E11.5 Pet1-/-
E14.5 Pet1-/-
E17.5 Pet1-/-
E14.5 Pet1neg
Tph2 promoter
Tph2 enhancer
i
5-HT transport and reuptake
5-HT metabolism
Slc18a2
Slc22a3
Slc6a4
Maoa
Maob

Supplementary Fig 5

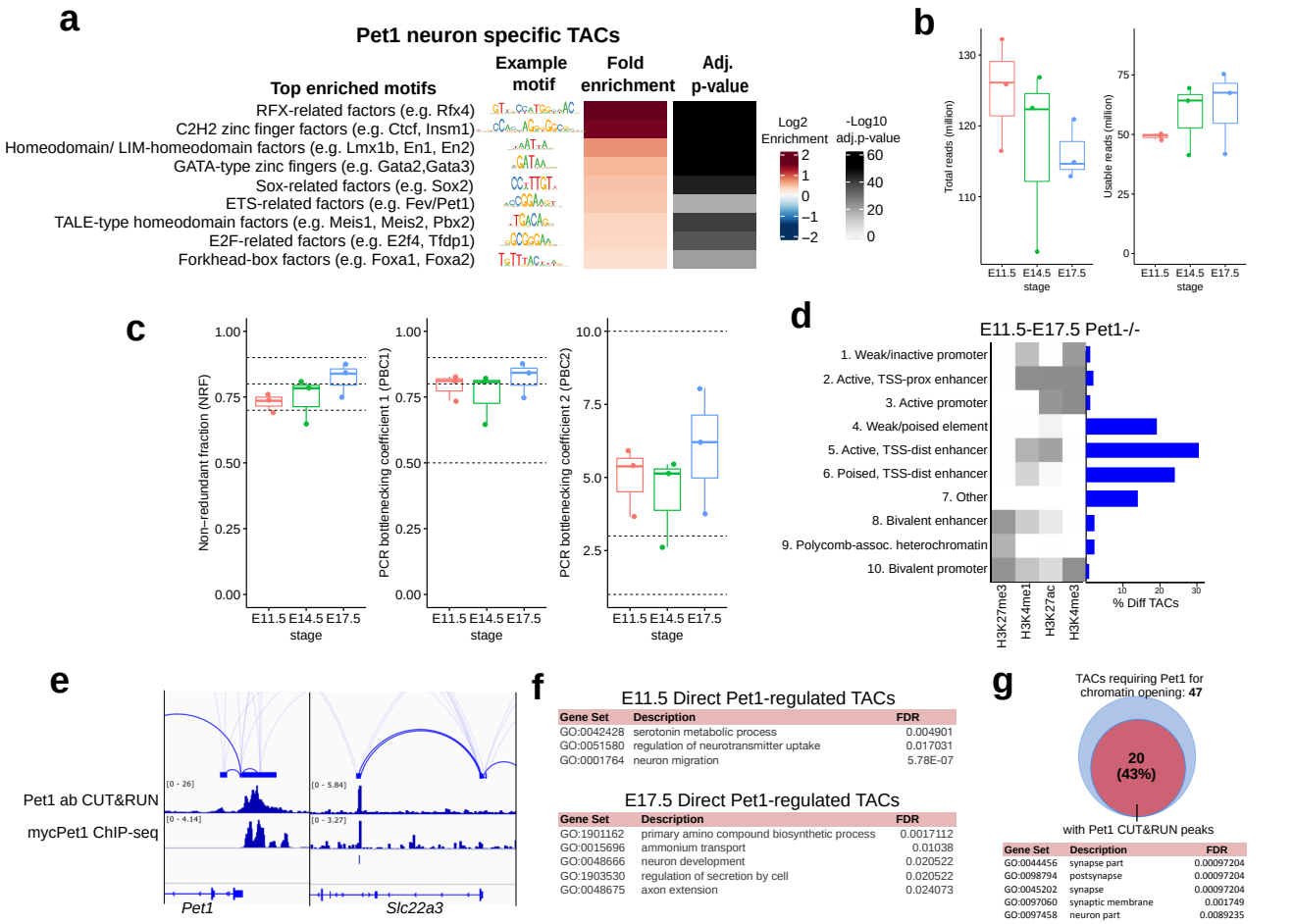
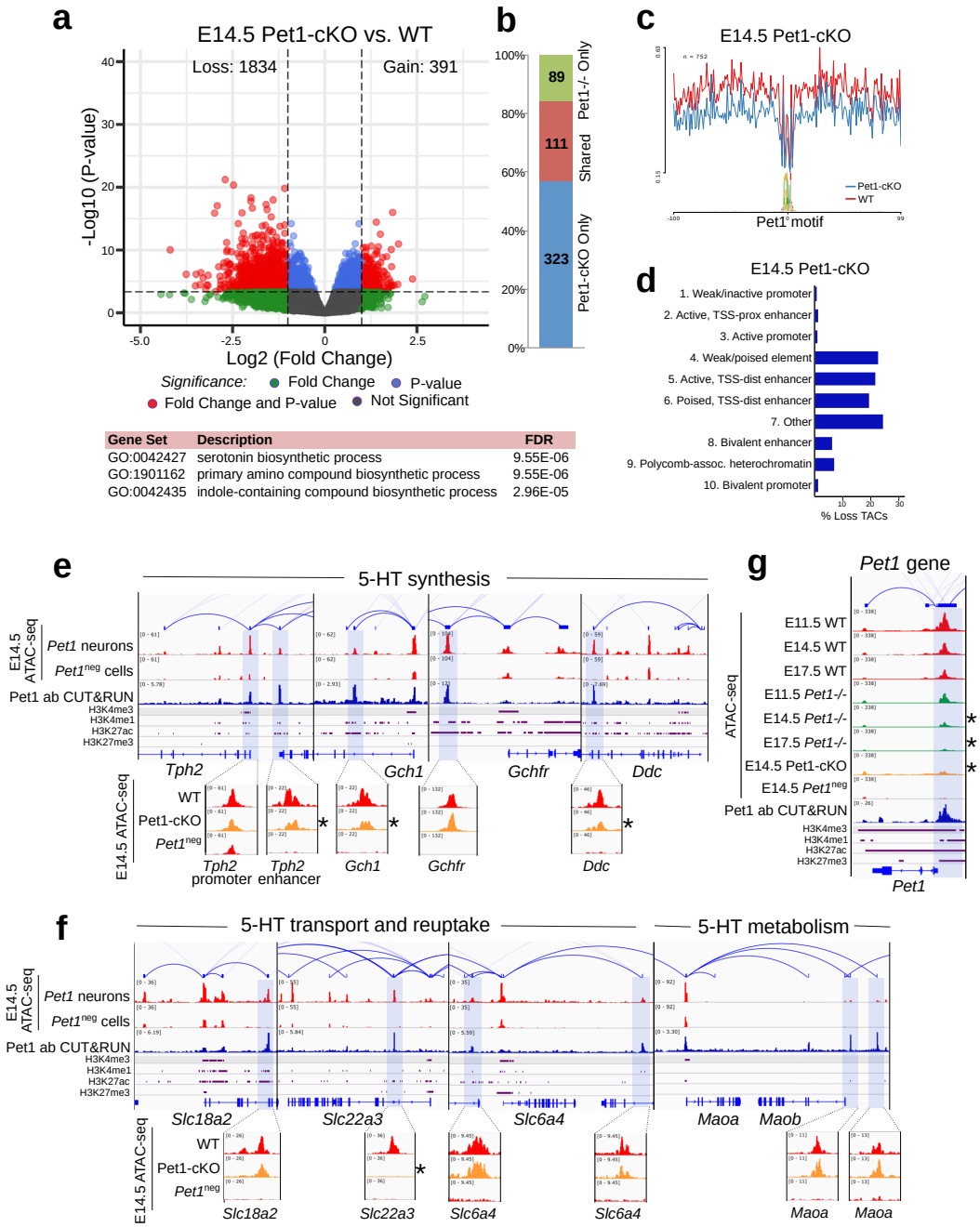
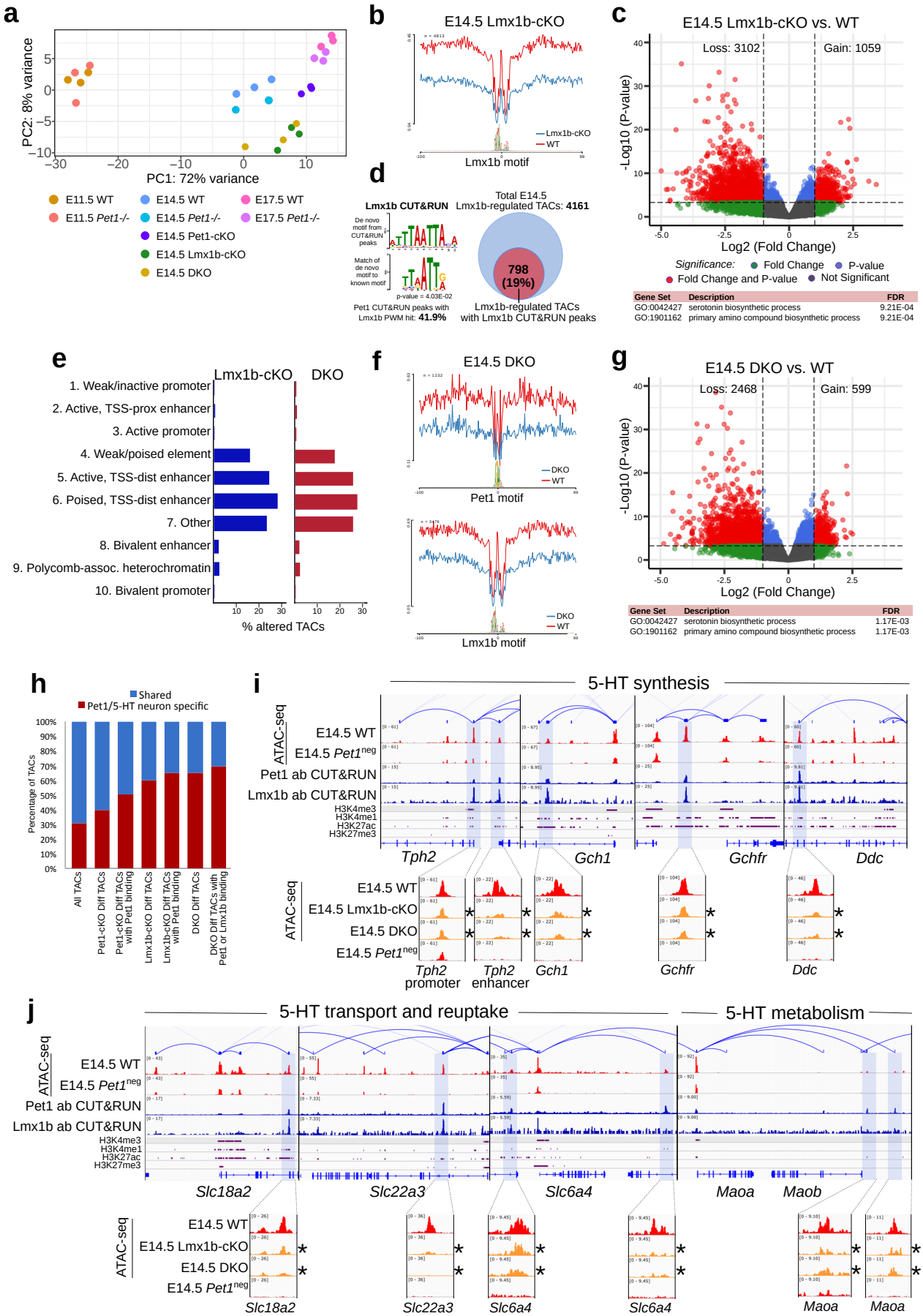


Figure 5

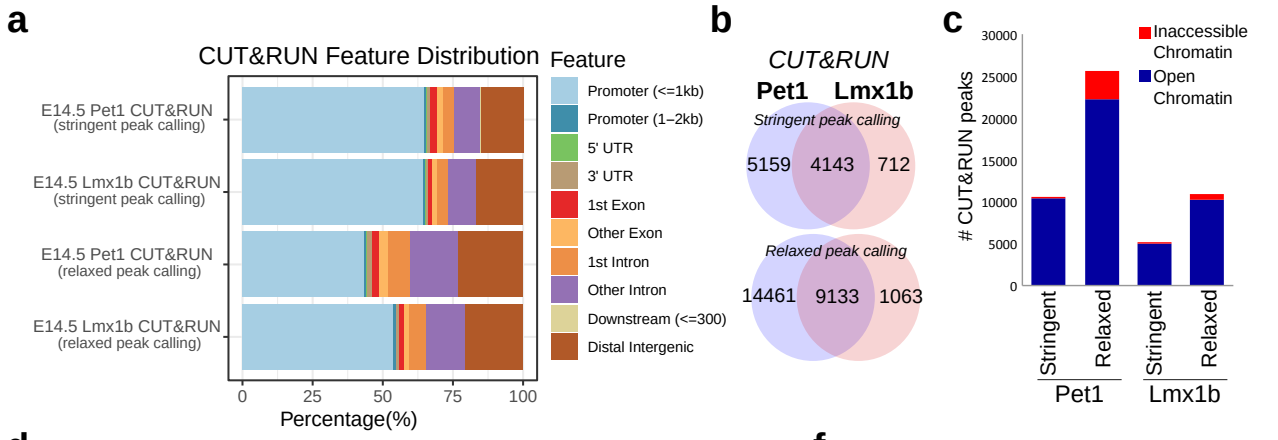


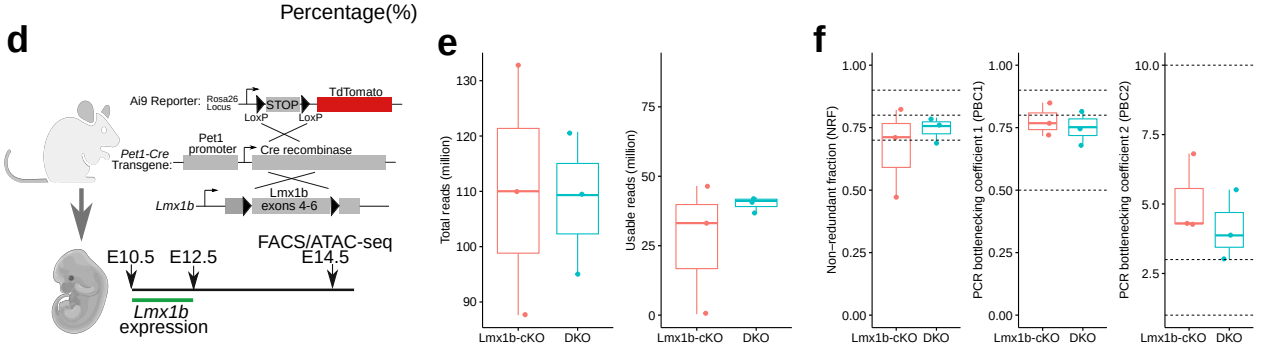
Gene Set	Description	FDR
GO:0042427	serotonin biosynthetic process	9.55E-06
GO:1901162	primary amino compound biosynthetic process	9.55E-06
GO:0042435	indole-containing compound biosynthetic process	2.96E-05

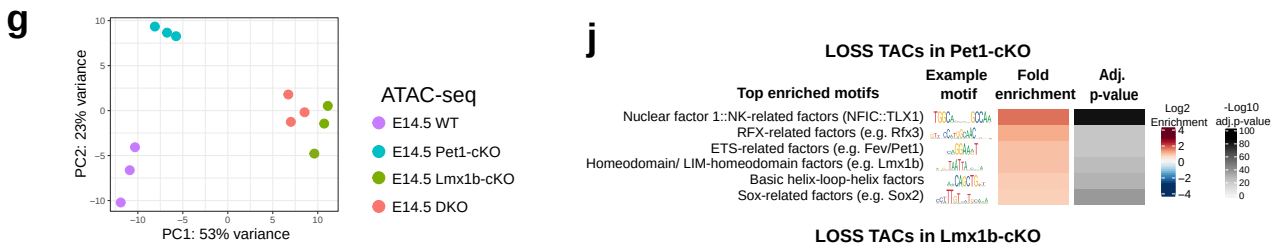
Figure 6


a
PC2: 8% variance
PC1: 72% variance
E11.5 WT
E11.5 Pet1-/-
E14.5 WT
E14.5 Pet1-/-
E14.5 Pet1-cKO
E14.5 Lmx1b-cKO
E14.5 DKO
E17.5 WT
E17.5 Pet1-/-
b
E14.5 Lmx1b-cKO
Lmx1b-cKO
WT
Lmx1b motif
c
E14.5 Lmx1b-cKO vs. WT
Loss: 3102
Gain: 1059
-Log10 (P-value)
Log2 (Fold Change)
Significance:
Fold Change
P-value
Fold Change and P-value
Not Significant
Gene Set Description FDR
GO:0042427 serotonin biosynthetic process 9.21E-04
GO:1901162 primary amino compound biosynthetic process 9.21E-04
d
Lmx1b CUT&RUN
De novo motif from CUT&RUN peaks
Match of de novo motif to known motif
p-value = 4.03E-02
Pet1 CUT&RUN peaks with Lmx1b PWM hit: 41.9%
Total E14.5 Lmx1b-regulated TACs: 4161
798 (19%)
Lmx1b-regulated TACs with Lmx1b CUT&RUN peaks
e
Lmx1b-cKO
DKO
1. Weak/inactive promoter
2. Active, TSS-prox enhancer
3. Active promoter
4. Weak/poised element
5. Active, TSS-dist enhancer
6. Poised, TSS-dist enhancer
7. Other
8. Bivalent enhancer
9. Polycomb-assoc. heterochromatin
10. Bivalent promoter
% altered TACs
f
E14.5 DKO
DKO
WT
Pet1 motif
Lmx1b motif
g
E14.5 DKO vs. WT
Loss: 2468
Gain: 599
-Log10 (P-value)
Log2 (Fold Change)
Gene Set Description FDR
GO:0042427 serotonin biosynthetic process 1.17E-03
GO:1901162 primary amino compound biosynthetic process 1.17E-03
h
Shared
Pet1/5-HT neuron specific
Percentage of TACs
All TACs
Pet1-cKO Diff TACs
Pet1-cKO Diff TACs with Pet1 binding
Lmx1b-cKO Diff TACs
Lmx1b-cKO Diff TACs with Pet1 binding
DKO Diff TACs
DKO Diff TACs with Pet1 or Lmx1b binding
i
5-HT synthesis
ATAC-seq
E14.5 WT
E14.5 Pet1neg
Pet1 ab CUT&RUN
Lmx1b ab CUT&RUN
H3K4me3
H3K4me1
H3K27ac
H3K27me3
Tph2
Gch1
Gchfr
Ddc
E14.5 Lmx1b-cKO
E14.5 DKO
Tph2 promoter
Tph2 enhancer
j
5-HT transport and reuptake
5-HT metabolism
Slc18a2
Slc22a3
Slc6a4
Maoa
Maob

Supplementary Fig 6







h TACs dependent on only Lmx1b (not Pet1)

Gene Set	Description	FDR
GO:0061564	axon development	0.0021383
GO:0038179	neurotrophin signaling pathway	0.013515
GO:0050808	synapse organization	0.01633
GO:0060560	developmental growth involved in morphogenesis	0.027331

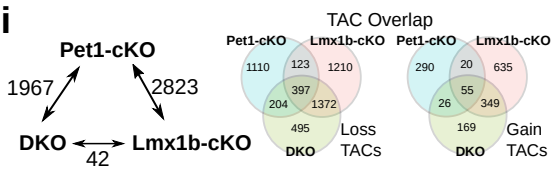
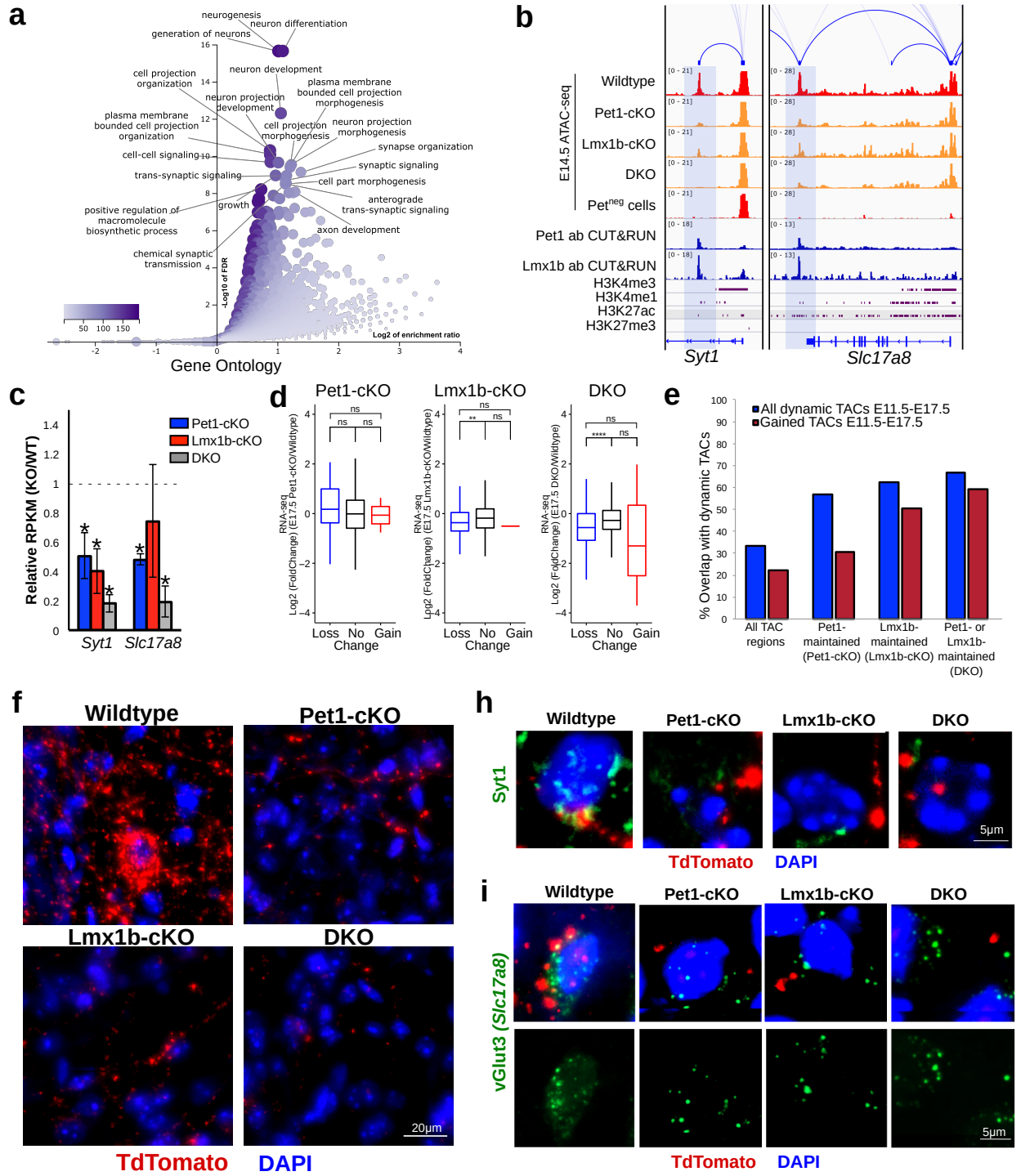


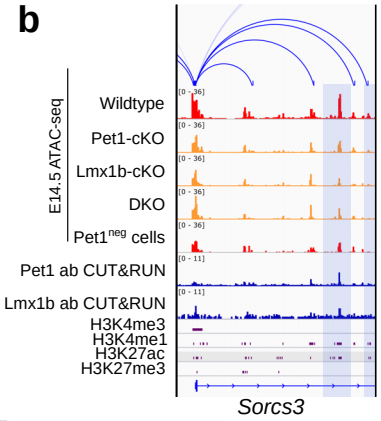
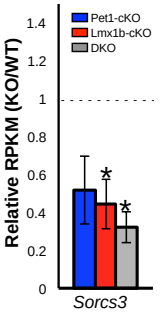
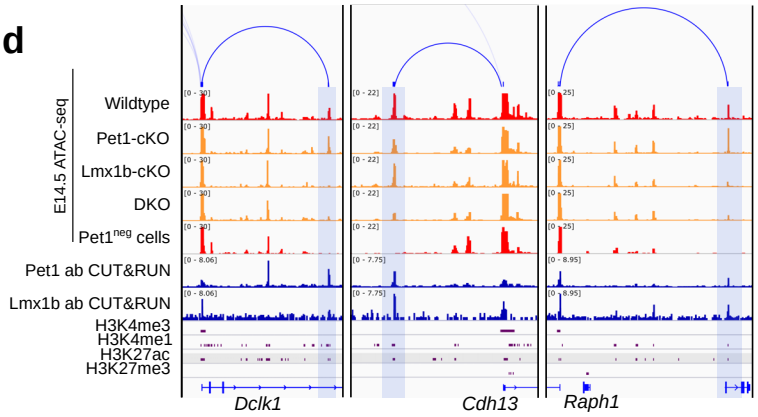
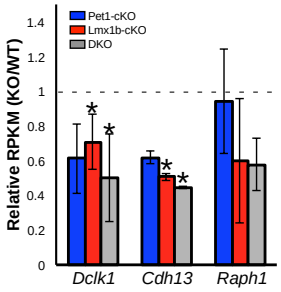
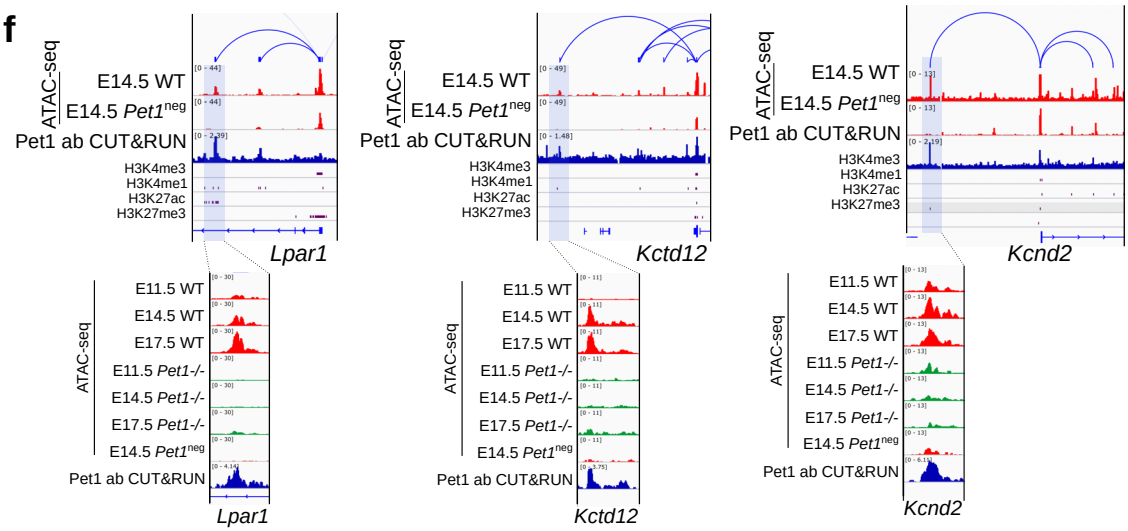
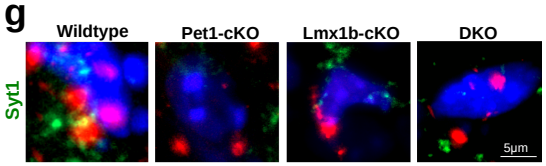
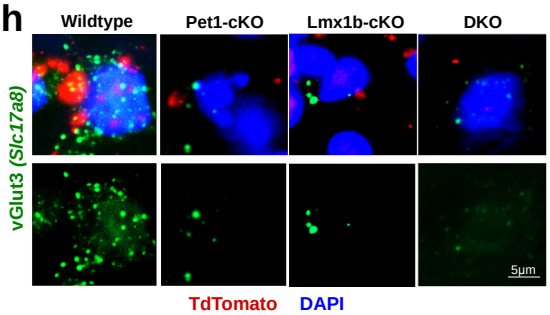
Figure 7


a
neurogenesis
neuron differentiation
generation of neurons
neuron development
cell projection organization
plasma membrane bounded cell projection morphogenesis
neuron projection development
plasma membrane bounded cell projection organization
cell projection morphogenesis
neuron projection morphogenesis
cell-cell signaling
synapse organization
trans-synaptic signaling
synaptic signaling
cell part morphogenesis
growth
anterograde trans-synaptic signaling
positive regulation of macromolecule biosynthetic process
axon development
chemical synaptic transmission
-Log10 of FDR
Log2 of enrichment ratio
Gene Ontology
b
E14.5 ATAC-seq
Wildtype
Pet1-cKO
Lmx1b-cKO
DKO
Pet^neg cells
Pet1 ab CUT&RUN
Lmx1b ab CUT&RUN
H3K4me3
H3K4me1
H3K27ac
H3K27me3
Syt1
Slc17a8
c
Pet1-cKO
Lmx1b-cKO
DKO
Relative RPKM (KO/WT)
Syt1
Slc17a8
d
Pet1-cKO
Lmx1b-cKO
DKO
RNA-seq Log2 (FoldChange) (E17.5 Pet1-cKO/Wildtype)
RNA-seq Log2 (FoldChange) (E17.5 Lmx1b-cKO/Wildtype)
RNA-seq Log2 (FoldChange) (E17.5 DKO/Wildtype)
Loss
No Change
Gain
e
All dynamic TACs E11.5-E17.5
Gained TACs E11.5-E17.5
% Overlap with dynamic TACs
All TAC regions
Pet1-maintained (Pet1-cKO)
Lmx1b-maintained (Lmx1b-cKO)
Pet1- or Lmx1b-maintained (DKO)
f
Wildtype
Pet1-cKO
Lmx1b-cKO
DKO
20μm
TdTomato
DAPI
h
Wildtype
Pet1-cKO
Lmx1b-cKO
DKO
Syt1
5μm
TdTomato
DAPI
i
Wildtype
Pet1-cKO
Lmx1b-cKO
DKO
vGlut3 (Slc17a8)
5μm
TdTomato
DAPI
g
Medial Septum
Axon density
Axon density (%Area)
WT
Pet1-cKO
Lmx1b-cKO
DKO
j
Syt1
Colocalization of TdTomato and Syt1 (KO /CON(%))
vGlut3 (Slc17a8)
Colocalization of TdTomato and vGlut3 (KO /CON (%))

Supplementary Fig 7

a

Gene Set	Description	FDR
GO:0048812	neuron projection morphogenesis	2.90E-10
GO:0050808	synapse organization	7.40E-10
GO:0099537	trans-synaptic signaling	1.12E-09
GO:0099536	synaptic signaling	2.00E-09
GO:0007268	chemical synaptic transmission	3.18E-09

Gene Set	Description	FDR
GO:0097485	neuron projection guidance	6.83E-06
GO:0007268	chemical synaptic transmission	0.002641
GO:0060291	long-term synaptic potentiation	0.002693
GO:0016050	vesicle organization	0.002693
GO:0055075	potassium ion homeostasis	0.004719

Gene Set	Description	FDR
GO:0031175	neuron projection development	0
GO:0050808	synapse organization	2.67E-12
GO:0061564	axon development	8.54E-12
GO:0007409	axonogenesis	1.30E-11
GO:0050807	regulation of synapse organization	2.09E-08