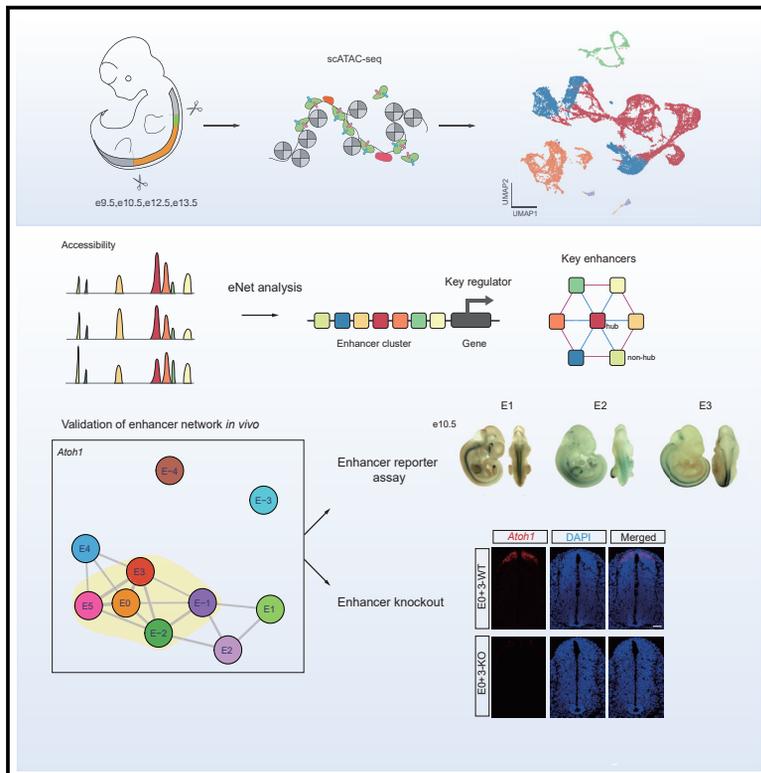# Developmental Cell

Single-cell chromatin accessibility identifies enhancer networks driving gene expression during spinal cord development in mouse

Graphical abstract

Authors

Muya Shu, Danni Hong, Hongli Lin, ...,
Falong Lu, Jialiang Huang, Jianwu Dai

Correspondence

fllu@genetics.ac.cn (F.L.),
jhuang@xmu.edu.cn (J.H.),
jwdai@genetics.ac.cn (J.D.)

In brief

Shu and Hong et al. provide single-cell chromatin accessibility analyses of developing mouse spinal cord and identify *cis*-regulatory elements and regulators for neural cells. Using these data, they define enhancer networks controlling master regulators during spinal cord development and experimentally validate *Atoh1* enhancers through *in vivo* perturbation as a proof of concept.

Highlights

- Single-cell chromatin accessibility profiling charts mouse spinal cord development

- Enhancer networks control master regulators during neuronal development

- Network hub enhancers are potentially key *cis*-elements to control gene expression

- *In vivo Atoh1* enhancer knockout demonstrates functional hierarchy of enhancer network

CellPress

Developmental Cell

CellPress

Resource

Single-cell chromatin accessibility identifies enhancer networks driving gene expression during spinal cord development in mouse

Muya Shu,[1,6] Danni Hong,[2,6] Hongli Lin,[2] Jixiang Zhang,[1] Zhengnan Luo,[3,4] Yi Du,[1,4] Zheng Sun,[1,4] Man Yin,[1,4] Yanyun Yin,[1] Lifang Liu,[2] Shilai Bao,[1,4] Zhiyong Liu,[3,4] Falong Lu,[1,4,*] Jialiang Huang,[2,5,*] and Jianwu Dai[1,4,7,*]

[1]State Key Laboratory of Molecular Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
[2]State Key Laboratory of Cellular Stress Biology, School of Life Sciences, Faculty of Medicine and Life Sciences, Xiamen University, Xiamen, Fujian 361102, China
[3]Institute of Neuroscience, State Key Laboratory of Neuroscience, Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China
[4]University of Chinese Academy of Sciences, Beijing 100101, China
[5]National Institute for Data Science in Health and Medicine, Xiamen University, Xiamen, Fujian 361102, China
[6]These authors contributed equally
[7]Lead contact
*Correspondence: fllu@genetics.ac.cn (F.L.), jhuang@xmu.edu.cn (J.H.), jwdai@genetics.ac.cn (J.D.)
https://doi.org/10.1016/j.devcel.2022.11.011

SUMMARY

Spinal cord development is precisely orchestrated by spatiotemporal gene regulatory programs. However, the underlying epigenetic mechanisms remain largely elusive. Here, we profiled single-cell chromatin accessibility landscapes in mouse neural tubes spanning embryonic days 9.5–13.5. We identified neuronal-cell-cluster-specific *cis*-regulatory elements in neural progenitors and neurons. Furthermore, we applied a novel computational method, eNet, to build enhancer networks by integrating single-cell chromatin accessibility and gene expression data and identify the hub enhancers within enhancer networks. It was experimentally validated *in vivo* for *Atoh1* that knockout of the hub enhancers, but not the non-hub enhancers, markedly decreased *Atoh1* expression and reduced dp1/dI1 cells. Together, our work provides insights into the epigenetic regulation of spinal cord development and a proof-of-concept demonstration of enhancer networks as a general mechanism in transcriptional regulation.

INTRODUCTION

The spinal cord links and allows communication between the brain and peripheral organs. Spinal cord formation begins around embryonic day 9 (e9) in mice.[1] Along the dorsal-ventral (DV) axis, the neural tube DV pattern is established mainly by opposing concentration gradients of bone morphogenic protein and sonic hedgehog (SHH) signals.[2–4] With the two signal morphogens, 11 types of neural progenitors are generated and occupy the defined partitions within the neural tube. They are categorized into dp1–dp6, p0–p2, pMN, and p3 along the DV axis. Then these cells mature and become post-mitotic neurons (dI1–dI6, V0–V2, MN, and V3).[5,6] Through intercellular signaling interactions, some neurons further diversify into distinct neuronal subtypes with different neurotransmitter properties.[5] The majority of the neurogenesis process lasts until e13.5, after which glial cells are produced from the remaining undifferentiated progenitors via gliogenesis.[7] Despite great efforts to understand the regulation of neural cell specification during spinal cord formation,[8,9] systematic characterization of the whole process from multiple molecular layers has not been reported, especially with respect to epigenetics.

The process of neurogenesis is tightly regulated by delicate gene regulatory networks. Cell-type- and cell-state-specific gene expression patterns are mediated by *cis*-regulatory elements (CREs) such as enhancers, promoters, silencers, and insulators. Among these elements, enhancers are the key drivers of spatiotemporal gene expression in a cell-type-specific manner.[10] It is a common phenomenon that developmentally important genes are controlled by enhancer clusters.[11] Enhancer clusters have been shown to work in an additive, redundant, or synergistic manner in diverse systems, such as limb[11,12] and retinal development,[13] hematopoiesis,[14,15] and embryonic stem cells differentiation.[16] Our previous work uncovered and dissected super-enhancer functional hierarchy during hematopoiesis.[17,18]

Single-cell multi-omics analyses such as single-cell RNA sequencing (scRNA-seq) and single-cell assay for transposase-accessible chromatin with high-throughput sequencing (scATAC-seq) can provide profound biological insights into
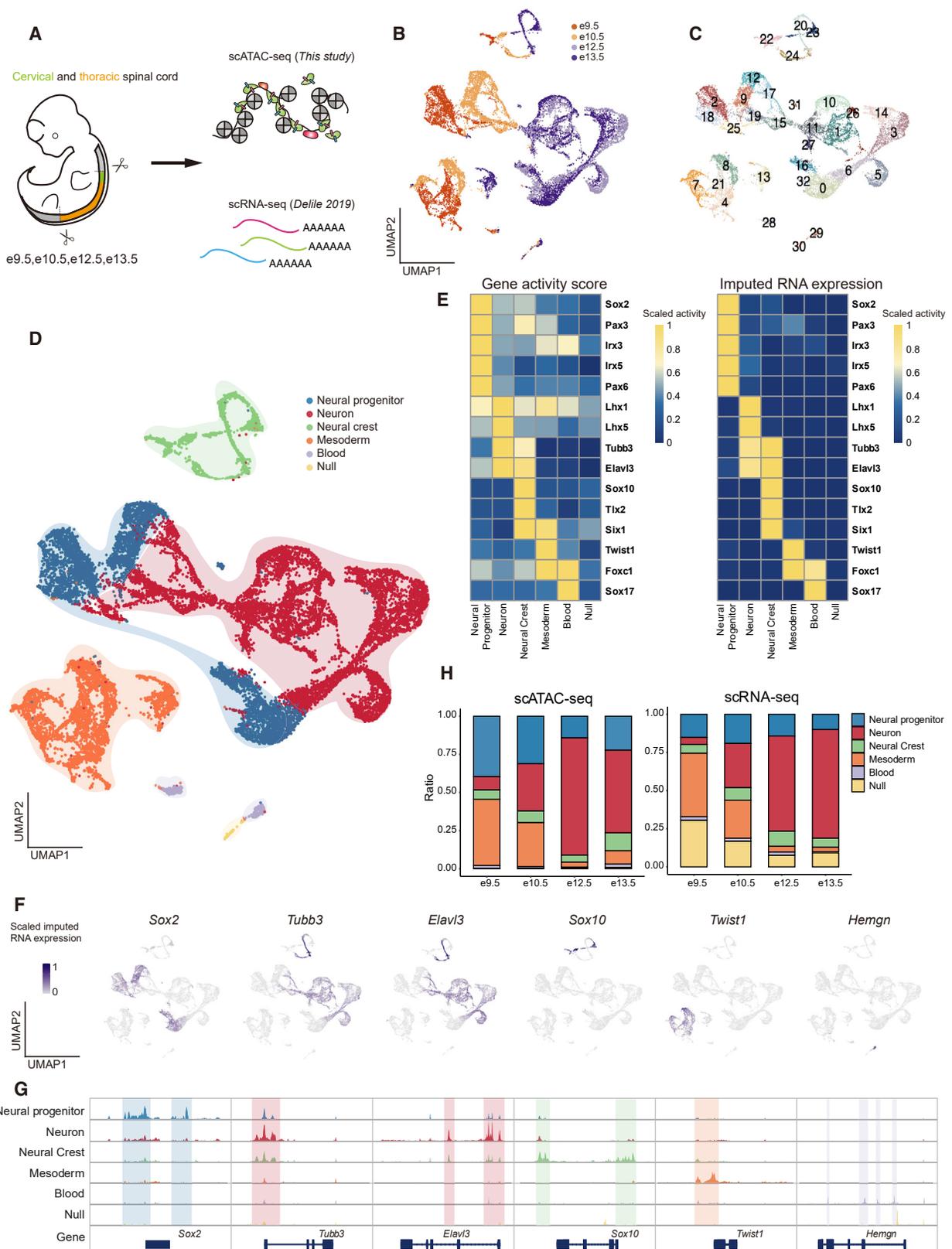
Figure 1. Chromatin accessibility profiling, clustering, and annotation of 19,715 cells from the developing mouse spinal cord

(A) Schematic of sample dissection strategy.

(B) Uniform manifold approximation and projection (UMAP) of scATAC-seq data colored by developmental stages.

(legend continued on next page)

cellular heterogeneity and underlying regulatory mechanisms at multiple molecular layers.[19] These approaches have been applied to a number of tissues, and they function as an extensible platform to study regulatory circuitry.[20–25] However, previous studies have largely focused on connecting enhancers with their target genes, but rarely on underlying mechanisms such as how multiple enhancers interact with each other to regulate precise gene expression.[26] By integrating single-cell chromatin accessibility and gene expression data, we have recently developed an algorithm, eNet,[27] to build an enhancer network that quantifies the underlying regulatory relationships among enhancers within an enhancer cluster. Furthermore, there remains a lack of single-cell chromatin accessibility profiling for spinal cord development.

In this study, we performed scATAC-seq to map single-cell chromatin accessibility during multiple stages of mouse neural tube development and identified key CREs and regulators. By integrating scATAC-seq and scRNA-seq data, beyond simply mapping enhancers to their target genes, we built enhancer networks, which allowed us to investigate how multiple enhancers work together to precisely orchestrate gene expression during spinal cord development. To provide a proof-of-concept demonstration of the important roles of enhancer networks, we performed functional validation of enhancers within the enhancer network in controlling *Atoh1* spatiotemporal expression during spinal cord development *in vivo*. Together, our analyses provide comprehensive resources for regulatory insight into spinal cord development and provide compelling evidence for the functional importance of enhancer networks in regulating key developmental genes.

RESULTS

Single-cell chromatin accessibility profiling of the developing mouse embryonic spinal cord

To explore the gene regulation during spinal cord development, we dissected the cervical and thoracic regions of mouse neural tubes at e9.5, e10.5, e12.5, and e13.5. For each stage of samples, we performed scATAC-seq using the Chromium platform (10x Genomics) (Figures 1A, S1A, and S1B; Table S1). After quality control, we obtained a total of 19,715 cells for further analysis (4,316 from e9.5, 4,741 from e10.5, 4,935 from e12.5, and 5,723 from e13.5) (Figure 1B), with a total of 228,689 peaks mapped to the nuclear genome and a median of 10,850 fragments per cell. To assess the similarities between individual cells, we performed unsupervised analyses including dimension reduction and clustering by using Signac,[28] which resulted in 33 cell clusters (Figure 1C). To infer cell cluster identities, we first calculated the gene activity scores by summing the fragments in gene promoter and gene body (Figure S1E). Then, we transferred cell-type labels by integrating the scATAC-seq dataset with a previously

published scRNA-seq dataset[29] by using Seurat package[30] (Figure S1F; see STAR Methods). We annotated cell clusters as neural progenitors, neurons, neural crest, mesoderm, and blood using the major cell-type labels from Delile et al.[29] (Figures 1D and S1I). Cells that did not fall into any of the above categories were named null, potentially because these cells were derived from other tissues of the embryo.[29] We observed a good correlation between gene activity scores and imputed RNA expression values (Figures 1E and S1G) and a concordance between cell-type labels derived from each dataset (Figure S1H). In addition, we examined the chromatin landscapes near several known cell-type-specific markers[29] (Figures 1F and 1G). As anticipated, the CREs around *Sox2* were uniquely accessible in neural progenitors. Cell-type-specific accessibility was also observed at the elements around *Tubb3* and *Elavl3* in neurons, *Sox10* in neural crest cells, *Twist1* in mesoderm, and *Hemgn* in blood cells (Figures 1E–1G). By comparing the proportions of cell types at different time points in the scATAC-seq dataset (Figure 1H), we found that the most drastic changes during development were the relative increase in neurons (marker genes: *Tubb3* and *Elavl3*) and the relative decrease in neural progenitors (marker gene: *Sox2*). Overall, the changes in cell composition observed in the scATAC-seq data were concordant with the corresponding scRNA-seq data (Figure 1H).

In summary, we generate the chromatin accessibility profiles for mouse embryonic spinal cord development, which provides a rich resource to explore the regulatory roles of CREs during spinal cord development.

Chromatin landscapes of neural progenitors and neurons encode the temporal and spatial patterns

To further study the gene regulatory program of lineage commitments during spinal cord development, we extracted the neural progenitors and neurons and re-clustered them into 23 clusters (Figure 2A). The cells from e9.5 were well separated from cells at other time points. This clear separation was also observed for cells from e10.5, while cells from e12.5 and e13.5 were closely located, nevertheless clearly separated (Figure 2B). Next, we sought to refine the subtypes of progenitors and neurons by transferring cell subtype annotation to scATAC-seq cells based on scRNA-seq data (Figures 2C, S2A, and S2B). We observed a mixture of subtype labels in most of the scATAC-seq clusters (Figures 2C, S2D, and S2E), which is consistent with the observation in the scRNA-seq data (Figures S2C and S2F).[29] This may be due to the overlap in neuron progenitor or neuron subtype markers (Table S3). However, we observed an obvious partition of neuron progenitors and neurons along the DV axis (Figure 2D), where each cluster mainly comprised neural progenitors or neurons from either the dorsal or ventral domain (Figure 2E). For example, we observed specific imputed RNA expression of *Pax7*, *Dbx1*, *Prrxl1*, and *En1* in dorsal neural progenitors, ventral

(C) UMAP embedding and clustering analysis of scATAC-seq data from 19,715 cells, identifying 33 clusters.

(D) UMAP of scATAC-seq data colored by major cell types.

(E) The expression level of genes used for the annotation of major cell types, quantified by gene activity scores from scATAC-seq and imputed RNA expression values from scRNA-seq.

(F) The scaled imputed RNA expression of marker genes colored in UMAP.

(G) Aggregated scATAC-seq tracks showing chromatin accessible peaks around the marker genes for each major cell types.

(H) Proportions of major cell types across spinal cord developmental stages as determined by scATAC-seq (left) and scRNA-seq (right).
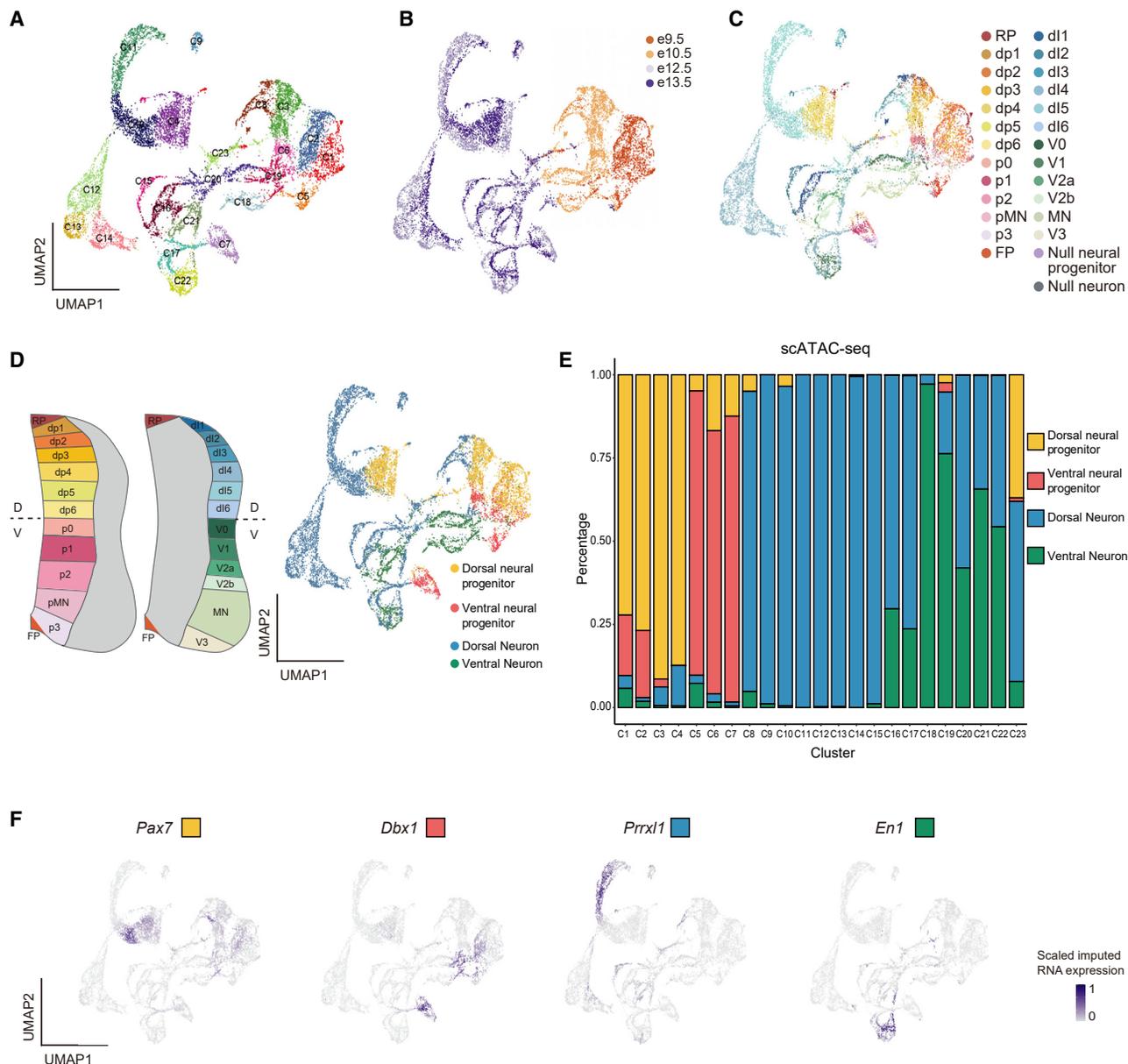
Figure 2. Chromatin landscapes of neural progenitors and neurons specify the temporal and spatial patterns
(A) UMAP embedding and clustering analysis of scATAC-seq data from 13,886 neural progenitors and neurons, identifying 23 clusters.
(B–D) UMAP of scATAC-seq neural progenitors and neurons, colored by (B) developmental stages, (C) cell subtypes, or (D) DV domains.
(E) Proportions of neural progenitors and neurons subdivided by DV domains in each cluster.
(F) The scaled imputed RNA expression of dorsal and ventral domains marker genes colored in UMAP.

neural progenitors, dorsal neurons, and ventral neurons, respectively (Figure 2F).[5] Together, these results show that the chromatin landscapes of neural progenitors and neurons can not only distinguish the developmental stages of the spinal cord but also encode spatial information along the DV domains.

Identifying cell-cluster-specific CREs in neural progenitors and neurons
Next, we sought to delineate the dynamics of neural cell CREs during spinal cord development. Majority of the CREs showed cell-type specificity (Figure S3A), which is consistent with pre-

vious observations.[31] Then, we performed differential analysis and refined 82,163 cell-cluster-specific CREs (cCREs), which were classified into 23 cCRE modules. Most of the cCRE modules showed highly cell-type-restricted chromatin accessibility during neural development (Figure 3A; Table S2). For example, SOX motifs were identified in neural-progenitor-specific cCRE modules (M2), where putative target genes of the CREs in M2 were enriched in biological processes (BPs) related to neural precursor cell proliferation (Figures 3A, S3D, and S3E). Interestingly, the cCREs in M7, which were specifically accessible in e12.5/e13.5 neural progenitors (cluster 7 [C7]), were highly
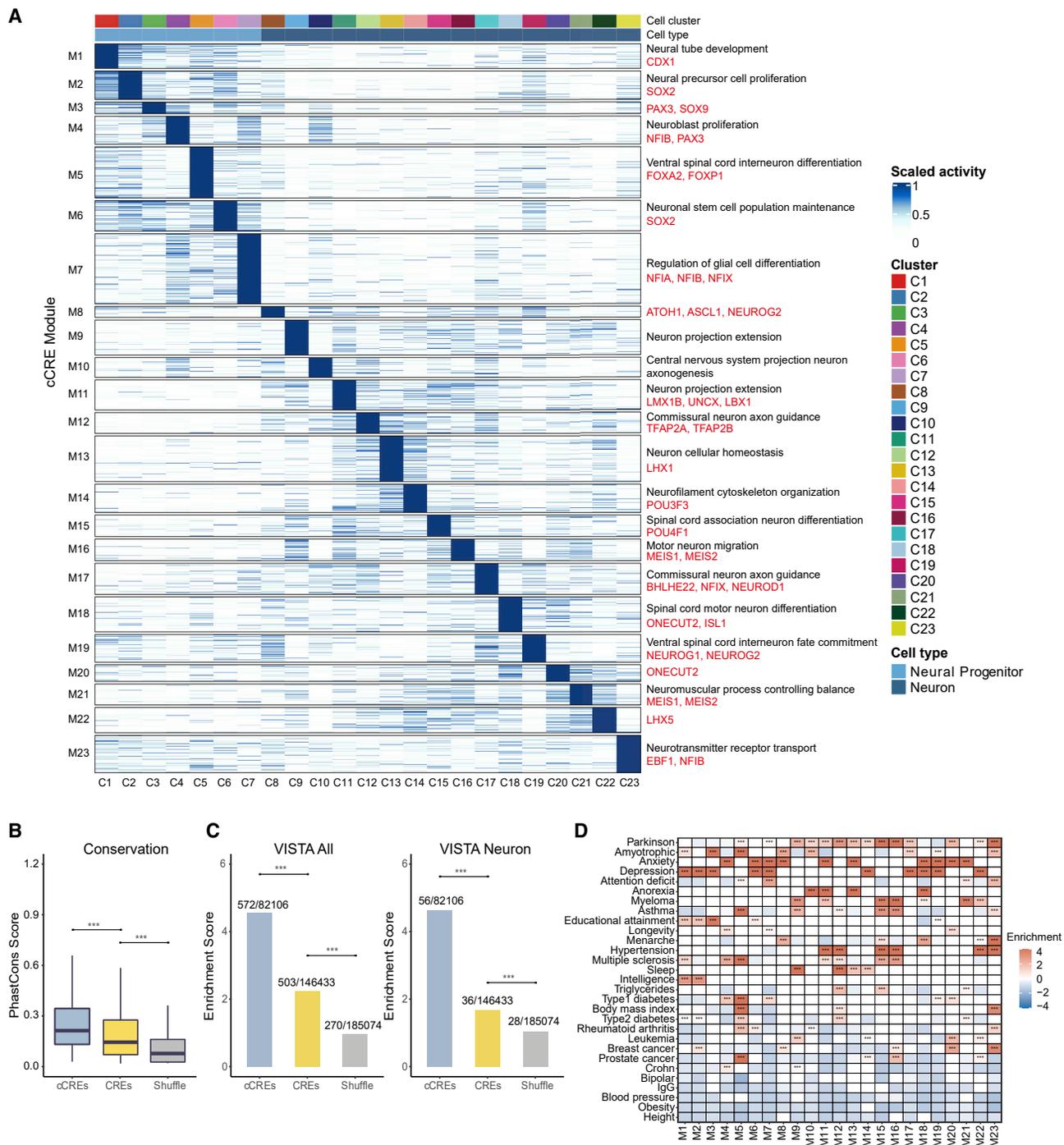
Figure 3. Identification and characterization of cell cluster-specific *cis*-regulatory elements (cCREs) in neural progenitors and neurons

(A) Heatmap showing the association of 23 cell clusters (columns) with 23 cCRE modules (rows). The representative gene ontology terms and transcription factor motifs significantly enriched in each cCRE module were shown at right.

(B) Box-Whisker plot showing the DNA sequence conservation phastCons score.

(C) Bar plot showing enrichment of the VISTA enhancers in cCREs and CREs, using shuffled CREs as background. The mouse enhancers with *in vivo* experimentally validated activity in all tissues (left) or neural tube (right) in the VISTA database were used.

(D) Scaled enrichment of sequence variants associated with the indicated traits/diseases in the human orthologs of cCREs. *p < 0.05, **p < 0.01, ***p < 0.001 (binomial test).
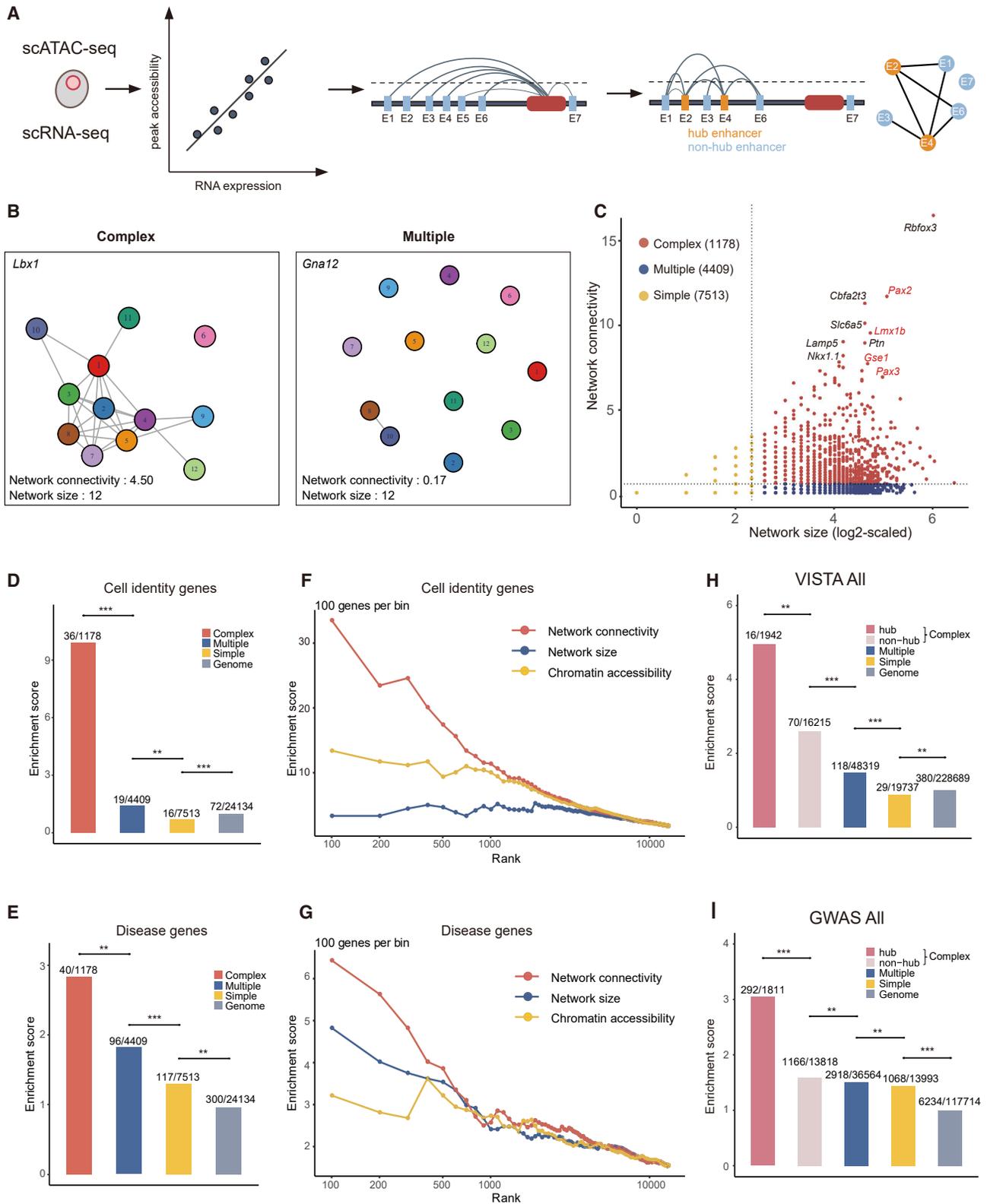
Figure 4. Enhancer networks during mouse spinal cord development
(A) Schematic depicting the framework of eNet for building enhancer networks.
(B) Representative enhancer networks in Complex (left) or Multiple (right) modes.

enriched in BP related to gliogenesis regulation and NFIA and NFIB motifs. NFIA and NFIB control the onset of gliogenesis in the developing spinal cord.[32] It suggests that cCREs in M7 may regulate the switch from neurogenesis to gliogenesis (Figure 3A). Therefore, the cCRE modules identified here lay a foundation for investigating the gene regulatory programs in different cell clusters of the developing spinal cord.

To further characterize these cCREs, we first observed that cCREs showed significantly higher phastCons conservation scores compared with CREs (p = 1.0E−4, Student's t test, Figure 3B). Next, we compared the cCREs with VISTA enhancers, which are experimentally validated mouse non-coding fragments with enhancer activity.[33] We found that the cCREs were significantly enriched more in VISTA enhancers (4.8-fold) than other CREs (2.4-fold), using the genome as background (p = 2.30E−55, binomial test, Figure 3C, left). Strikingly, cCREs displayed 4.7-fold enrichment in VISTA enhancers with activity in mouse neural tubes (Figure 3C, right). Last, we wondered whether cCREs in mouse neural progenitor and neuron clusters could assist in interpretation of non-coding risk variants related to neurological diseases. To this end, we identified human orthologs for ∼80% of the mouse cCREs (Figure S3C) and observed that the cCREs were significantly enriched more neuron-related GWAS SNPs (Figure S3F). Interestingly, the cCREs across various clusters are enriched with risk variants for neurological diseases but not other traits, such as blood pressure, obesity, or height, implying that these elements may be functionally important in the nervous system (Figure 3D). Together, our analyses delineate a comprehensive repertoire of cCREs during spinal cord development.

Identification of putative enhancer clusters

To further understand how these cCREs control gene expression programs, we examined their distribution in the genome. The majority of cCREs were located in regions more than 2 kb away from the nearest transcriptional start site (TSS) (Figure S3B). This is consistent with a previous finding that distal elements have a higher correlation with cell-type-specific gene expression than promoters do.[34] Therefore, we focused on the distal elements located at regions from 2 to 100 kb around the annotated TSS, hereafter referred to as putative enhancers (Figure S4A). Next, we mapped each putative enhancer to their target genes based on the correlation between gene expression and enhancer accessibility as previously described[21] (Figure S4A). In total, we identified 86,213 significant enhancer-gene associations, with a median of 5 enhancers per gene (Table S3). We termed a set of enhancers putatively regulating the same gene as a putative enhancer cluster. We observed that a subset of genes (2,573) associated with a large number

of enhancers (≥10) was highly enriched in cell identity genes (Figure S4B). In addition, GO analysis identified that the genes were enriched in neuron differentiation and neuron migration functions (Figure S4C). These results suggest that the putative enhancer clusters may play important roles in mediating the expression of neural genes, which is consistent with previous reports showing that multiple enhancers provide an effective regulatory buffer, conferring phenotypic robustness in development.[11,22]

Cell identity and disease genes tend to be regulated by complex enhancer networks

Many enhancers exist as clusters in the genome and control cell identity and disease genes.[11,22,35] However, previous studies largely focused on connecting enhancer clusters with their target genes rather than relationships among the enhancers. It remains challenging to understand how individual enhancers in an enhancer cluster interact with each other to precisely regulate gene expression.[26] To this end, we applied a novel computational method eNet to build enhancer networks, which is an unsupervised approach to explore the underlying enhancer regulation mechanism based on scATAC-seq and scRNA-seq data[27] (Figure 4A). Briefly, we took the single-cell gene expression matrix and enhancer accessibility matrix as input. For a given gene, we identified a putative enhancer cluster for each gene based on the correlation between gene expression and enhancer accessibility by adapting the method previously described.[21] Meanwhile, we identified the predicted enhancer interactions (PEIs) between enhancers within each putative enhancer cluster, where an enhancer pair is considered as a PEI if its co-accessibility is higher than 0.2, the default parameter recommended by Cicero.[36] Finally, we built an enhancer network for each gene, where nodes represent enhancers in the enhancer network and edges represent PEIs. We quantified the complexity of the enhancer networks by the network size and connectivity. Network size was the number of enhancers. Network connectivity was calculated as two times of the number of edges divided by the number of nodes (see STAR Methods).

Here, we applied eNet to build enhancer networks using all neural progenitors and neurons. For example, Lbx1 is controlled by an enhancer network consisting of 27 PEIs among 12 enhancers (Figure 4B, left). In contrast, Gna12 is controlled by an enhancer network containing a similar number of enhancers but only 1 PEI (Figure 4B, right). Interestingly, the number of cell types that a gene is expressed in showed no significant role in its network size or network connectivity (Figures S4D–S4F). We noticed that several neuron-related cell identity genes such as Pax2, Lmx1b, Gse1, and Pax3 displayed obviously large network size and high connectivity (Figure 4C). This led us to test

(C) Distribution of enhancer networks during mouse spinal cord development, where the x axis (log₂-scaled) represents the network size (the number of enhancers) of each enhancer network, the y axis represents the network connectivity, which was calculated as the average number of edges associated with each node in network. The top complex enhancer networks are labeled and the known cell identity gene enhancer networks are highlighted in red.

(D and E) Enrichment of cell identity (D) and disease genes (E) in genes in complex, multiple, and simple modes, using the whole genome as the background. *p < 0.05, **p < 0.01, ***p < 0.001; n.s., not significant (binomial test).

(F and G) Enrichment of cell identity (F) and disease genes (G) (y axis) was plotted for the top genes (x axis) ranked by different properties of enhancer networks, including network connectivity, network size, and chromatin accessibility.

(H and I) Enrichment of VISTA validated enhancers (H) and GWAS SNPs for enhancers in (I) complex (hub and non-hub), multiple, and simple modes, using randomly selected genomic regions as control. *p < 0.05, **p < 0.01, ***p < 0.001; n.s., not significant (binomial test).
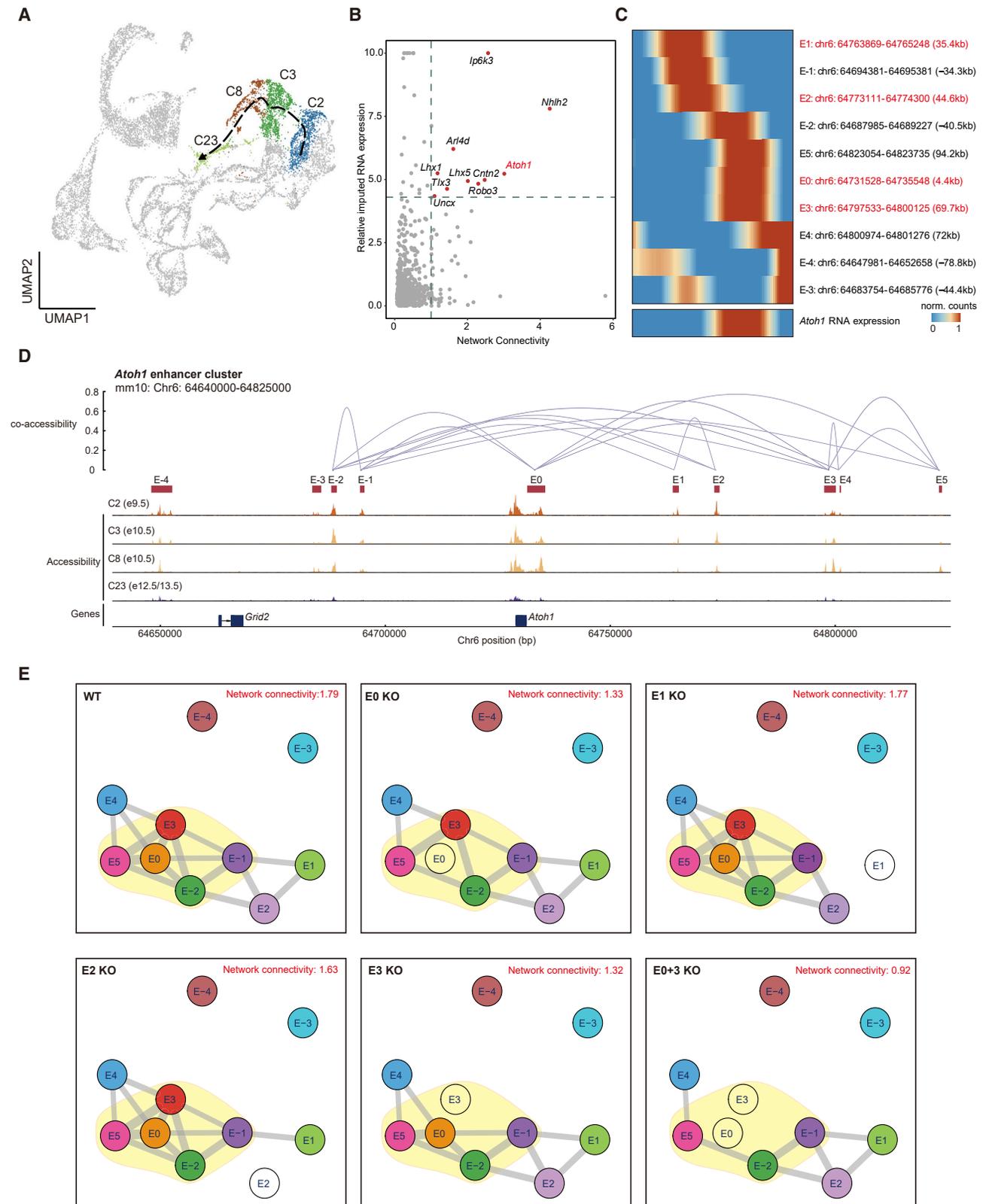
Figure 5. The *Atoh1* enhancer network

(A) UMAP of neural progenitors and neurons colored by the pseudotime of the differentiation of dI1 cell (clusters C2, C3, C8, and C23 in Figure 2A). The dashed line represents a double-spline fitted trajectory across pseudotime.

whether cell identity genes were more likely to be regulated by complex enhancer networks. To test this hypothesis, we classified the enhancer networks into three modes: simple (controlled by 1–5 enhancers, 7,513), multiple (multiple enhancers but limited PEIs, 4,409), and complex (multiple enhancers and frequent PEIs, 1,178) (Figure 4C; Table S3). We observed that genes in complex mode showed the highest enrichment in known cell identity genes (Table S4)[29] and in neuron-related disease genes (as curated from DisGeNET, Table S4)[37] (Figures 4D and 4E). These results indicate that the cell identity and disease genes are more likely to be regulated by complex enhancer networks.

Based on this observation, we next asked whether the complexity of enhancer networks can be used to predict cell identity and disease genes. To systematically test the prediction performance, we ranked all genes by the properties of their enhancer networks, including network connectivity, network size, or overall chromatin accessibility. We then calculated the enrichment in cell identity genes during spinal cord development and neural-associated disease genes within the list of top-ranked genes, using the whole genome as background (Figures 4F and 4G). The genes ranking by network connectivity showed significantly higher correlation with enrichment of cell identity genes and disease genes than those by network size or chromatin accessibility. For example, the top 100 genes ranked by network connectivity showed the highest enrichment in cell identity and disease genes, with 33.5- and 6.4-fold enrichment, respectively. This indicated that network connectivity has better performance in predicting cell identity and disease genes than network size and chromatin accessibility. In summary, this analysis suggests that cell identity and disease genes associated with spinal cord development tend to be regulated by complex enhancer networks.

Network hub enhancers enrich in validated enhancers and disease-associated variants

Next, we asked whether enhancer networks might provide an opportunity to study how enhancers control gene expression. Our previous work uncovers the hierarchical organization of super-enhancers, where the enhancers with frequent chromatin interactions detected by Hi-C (hub enhancers) play distinct roles in chromatin organization and gene activation.[18] Inspired by our previous work, we defined the enhancers with high frequent PEIs as network hub enhancers in enhancer networks in this study (see STAR Methods). To gain insight into the function of network hub enhancers, we first compared the enrichment of VISTA enhancers in complex (hub and non-hub), multiple, and simple modes. We found that the network hub enhancers showed significantly higher enrichment in VISTA enhancers than the non-hub enhancers (5.0- versus 2.6-fold, p = 6.4E−3

binomial test) (Figure 4H), in particular for VISTA enhancers with experimentally validated activity in mouse neural tubes (12.2- versus 1.5-fold, p = 5.4E−3 binomial test, Figure S4J). In addition, network hub enhancers displayed higher enrichment in disease- and phenotype-associated GWAS SNPs than non-hub enhancers (3.0- versus 1.6-fold, p = 1.9E−26, binomial test, Figure 4I), especially GWAS SNPs associated with neural traits (5.9- versus 2.2-fold, p = 4.5E−3, binomial test, Figure S4K). Taken together, we find that network hub enhancers enrich in experimentally validated enhancers and disease-associated variants, suggesting that they may play important roles in gene regulation.

The *Atoh1* enhancer network is hierarchically organized

To further explore the regulatory roles of enhancer networks during spinal cord development, we focused on the dp1/dI1 cell development trajectory,[5,9] which starts from neural progenitors dp1 (C2) at e9.5, commits to dp1 (C3), and dI1 interneurons (C8) at e10.5, then finally matures to neuron cells (C23) at e12.5/e13.5 (Figures 5A and S5A). To identify the key regulators functionally important for dI1 cell development, we built dp1/dI1 cell-specific enhancer networks using cells along the dp1/dI1 development trajectory (Figures 5A, S5B, and S5C; Table S5). Then, we identified candidate dp1/dI1 key regulators based on two criteria. The first one is the network connectivity calculated using the cells in dp1/dI1 trajectory (C2, C3, C8, and C23), given our finding that network connectivity is associated with cell identity genes (Figure 4). The other is the relative gene expression, which considered both the gene expression level and cell-type specificity in dp1/dI1 trajectory. Based on these two criteria, we identified the top 10 candidate regulators for dp1/dI1 (Figures 5B, S5D, and S5E), among which seven candidates (*Nhlh2*, *Atoh1*, *Lhx1*, *Lhx5*, *Cntn2*, *Robo3*, and *Tlx3*) were identified as markers of neuronal subtypes.[29]

Next, we asked how enhancer networks control master regulators during lineage commitment. As a proof-of-concept study, we focused on the *Atoh1* enhancer network, which showed hierarchical organization as well as a reasonable number of constituent enhancers for further experimental validation. *Atoh1* was expressed at the highest level in dI1 interneurons (C8) at e10.5 specifically (Figure S5B) and putatively regulated by an enhancer cluster of 10 enhancers (Figures 5C and 5D). Along dI1 differentiation pseudotime, *Atoh1* enhancers can be separated into three temporal groups based on their accessibility: early (E1, E-1, and E2), middle (E-2, E5, E0, and E3), and late (E4, E-4, and E-3) (Figure 5C). Interestingly, the middle enhancers became accessible simultaneously with the *Atoh1* imputed RNA expression (Figure 5C). Meanwhile, they are associated with more PEIs than early or late enhancers (Figure 5D). Network hub enhancers showed significantly higher correlation with gene expression and

(B) Scatter plot showing the predicted factors that might be functionally important for dp1/dI1 specification, considering both enhancer network connectivity (x axis) and relative imputed RNA expression (y axis). Top 10 shortlist was highlighted.

(C) Dynamics of *Atoh1* gene expression and chromatin accessibility of individual enhancers along dI1 differentiation pseudotime.

(D) Genome browser tracks showing aggregated chromatin accessibility profiles for clusters along dI1 differentiation around the *Atoh1* gene. Loops denote the predicted enhancer interaction (PEI) of each enhancer pair. Loop height represents the co-accessibility score of each PEI.

(E) *In silico* perturbation of the *Atoh1* weighted enhancer network. Nodes represent enhancers in the *Atoh1* weighted enhancer network, whereas the blank node represents the removed enhancer. The edges represent the PEIs and edge width represents the co-accessibility between the PEIs. The weighted network connectivity score is shown in the upright in each graph.

lower time residuals (defined as the difference of chromatin accessibility and expression of the gene) than non-hub enhancers (Figure S5F), indicating that hub enhancers are more synchronous with gene expression than non-hub enhancers. To quantitatively estimate the function of enhancers in *Atoh1* enhancer networks, we built a weighted enhancer network using the cells in dp1/dI1 trajectory following eNet analysis, except assigning chromatin co-accessibility as the weight of edges (Figures 5D and 5E). We observed a high overlap between middle enhancers and network hub enhancers (Figure 5E), suggesting that they may be important in regulating the *Atoh1* expression program. To test this hypothesis, we performed *in silico* perturbation[38] on the *Atoh1* enhancer network by assessing the effect of removing individual enhancers on network connectivity. Interestingly, removal of individual network hub enhancers (e.g., E0 or E3) attenuated the network connectivity more severely than the removal of non-hub enhancers (e.g., E1 or E2) (Figures 5E and S5G). Moreover, after removing E0 and E3 at the same time, the network was severely impaired, with the weighted network connectivity decreasing almost 50% (Figure 5E). These results suggest that *Atoh1* enhancers are hierarchically organized; network hub enhancers may play important roles in regulating the expression of *Atoh1* during dp1/dI1 lineage specification.

Transgenic reporter assays identify distinct spatiotemporal activity of *Atoh1* enhancers during mouse spinal cord development

We next experimentally assessed the function of network hub enhancers (E0 and E3) and non-hub enhancers (E1 and E2) in the *Atoh1* enhancer network (Figure 5E). Of note, E0 and E3 were chosen as the representative network hub enhancers for experimental validation because their removal caused the most severe attenuation of the network connectivity in *in silico* perturbation, and E0 is also a known enhancer for *Atoh1* expression.[39] E1 and E2 were chosen as the representative network non-hub enhancers because they are located between E0 and E3 (Figure 5D). We found that these putative enhancers, E0–E3, are evolutionarily conserved (Figure S6A). The activity of E0 during spinal cord development has previously been reported,[40] so we tested the activity of the three identified putative enhancers (E1–E3) of *Atoh1* in the spinal cord by creating transgenic reporter mice for these enhancers with the PiggyBac system (Figures 6A and S6B). If an enhancer was functional in the neural tube, when placed close to a LacZ gene downstream of a heat shock protein 68 kDa mini promoter (Hsp68), β-galactosidase (β-gal) activity could be detected after X-gal staining in neural tube.[41–43] These *Atoh1* enhancers showed distinct spatiotemporal activity patterns in the neural tube. E1 conferred β-gal activity from e9.5 to e10.5, whereas E2 and E3 conferred β-gal activity from e10.5 to e13.5 (Figure 6B). Similar results were also observed in another independent transgenic mouse line for each of the enhancers (Figure S6C). These results suggest that the activity of *Atoh1* enhancers drive spatiotemporal expression of *Atoh1* during mouse spinal cord development.

Functional hierarchy of the *Atoh1* enhancer network demonstrated by genetic knockout of enhancers

Next, we used CRISPR-Cas9 genome editing to individually delete E0–E3 to study their function in regulating proper *Atoh1*

expression during spinal cord development *in vivo* (Figures 7A and S7A–S7E). Because *Atoh1* expression peaks at e10.5 in the neural tube,[9] we measured *Atoh1* expression in the spinal cord of the KO mice at e10.5 (Figures 7B–7E). The mice with E1 or E2 non-hub enhancers deletion showed minimal changes in *Atoh1* expression compared with their wild-type (WT) siblings (Figures 7C and 7D). This supported the results of the *in silico* perturbation (Figure 5E), where removal of E1 or E2 only mildly impaired the enhancer network. In contrast, mice with E0 or E3 network hub enhancers deletion showed significant downregulation of *Atoh1* compared with their WT siblings (Figures 7B and 7E). These suggest that network hub enhancers (E0 and E3) play more important roles than non-hub enhancers (E1 and E2) in regulating expression of *Atoh1* in the developing spinal cord. Moreover, we tested the effects of deletion of two hub enhancers by generating E0+E3 double-KO mice thorough targeting E3 in E0 heterozygotes with CRISPR-Cas9 genome editing (Figures 7A and S7F). Consistent with the observation in the *in silico* perturbation, where dual removal of E0+E3 led to severe collapse of the enhancer network (Figure 5E), the E0+E3 dual KO mice showed significantly decreased level of *Atoh1* mRNA (Figure 7F). We additionally quantified expression of *Barhl2*, which is a well-established marker for *Atoh1* lineage cells in the spinal cord.[44,45] *Barhl2* showed similar gene expression changes to *Atoh1* as measured in the e13.5 developing spinal cord after genetic manipulation of these enhancers (Figures 7G–7K). In summary, these experiments supported our *in silico* model of the *Atoh1* enhancer network (Figure 5E). Removal of network hub enhancers attenuated network connectivity much more than removal of non-hub enhancers. Together, our results provide compelling evidence demonstrating the functional hierarchy of the *Atoh1* enhancer network during mouse spinal cord development, and network hub enhancers appear to be the major functional constituents in modulating the precise expression of target genes.

DISCUSSION

Changes in chromatin accessibility prime a cell for a particular lineage prior to changes in gene expression, foreshadowing the lineage choice.[22] Here, we applied scATAC-seq analysis to dissect the CREs and regulators spanning mouse spinal cord development in each cell cluster and cataloged the cells based on the similarity of chromatin accessibility profiles. Interestingly, these cell clusters are not composed of uniform cell types, but the cell types in each cell cluster are spatially neighboring cells, indicating that even the same types of cells can possess different chromatin landscapes that encode necessary spatial information. These cell-cluster-restricted modules were enriched in distinct sets of TF motifs with GO term annotations related to neuron specification, which allow us to expand the molecular description of cell types.

Enhancers existing in a cluster to regulate a single gene is a prevalent phenomenon, which are named as shadow enhancers,[46] super-enhancers,[35] stretch enhancers,[47] multiple enhancers,[11] activity-by-contact (ABC) model,[48] and domains of regulatory chromatin (DORC)[22] etc. These models mainly explored the relationship between enhancers and their target genes; however, it is less clear how enhancers interact with
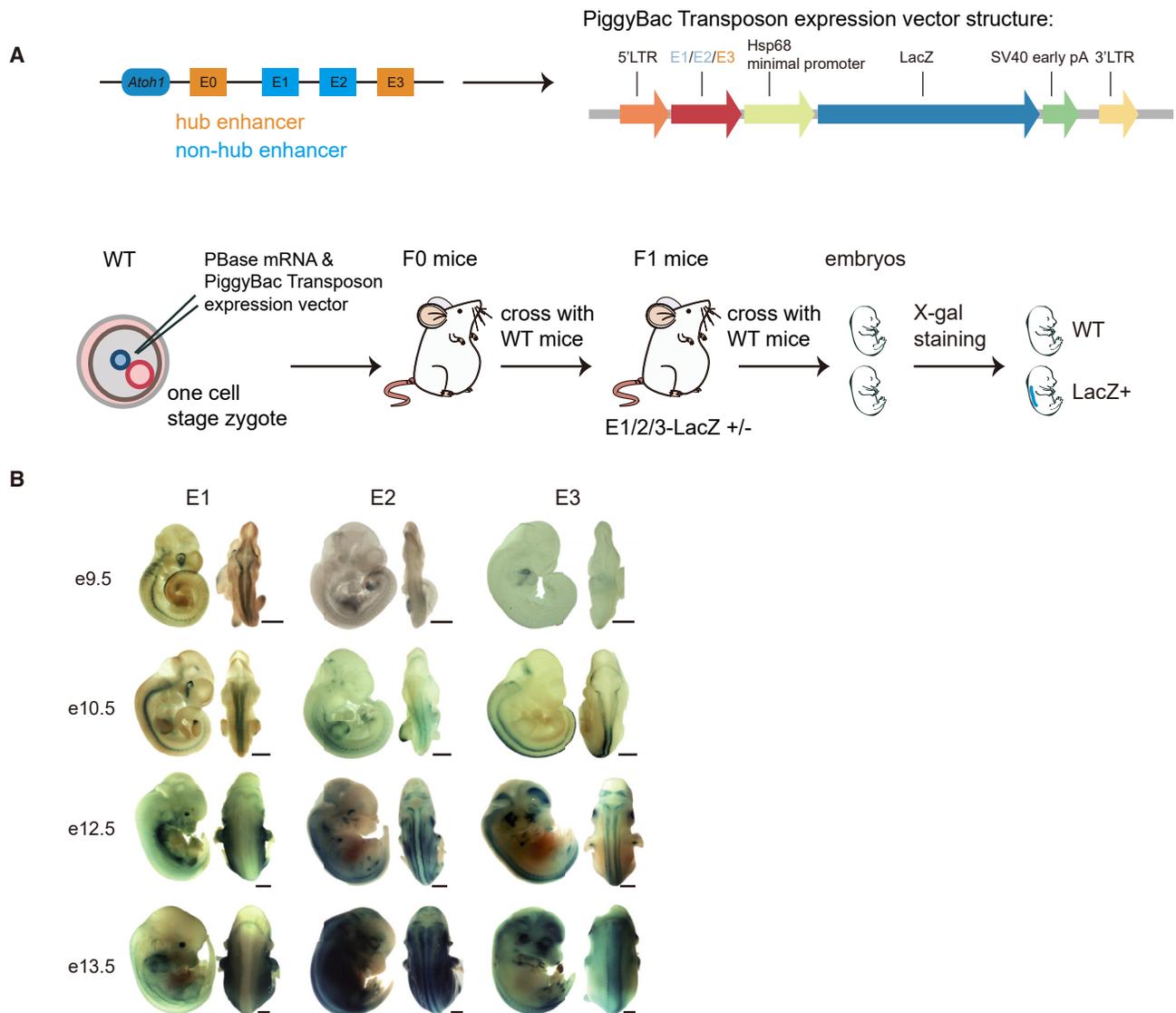
Figure 6. The spatiotemporal activity of *Atoh1* enhancers during spinal cord development identified by transgenic reporter assays

(A) Strategy of E1/E2/E3-lacZ transgenic mouse generation and X-Gal staining. E1, E2, and E3 were cloned upstream of an hsp68 minimal promoter and the reporter gene *lacZ*, then used to generate transgenic mice (top). After obtaining transgenic mice, whole-mount β-galactosidase histochemistry was conducted on mouse embryos (bottom).

(B) Whole-mount X-gal staining of E1, E2, and E3 reporter embryos at e9.5–e13.5. The blue signal displays β-gal activity of the LacZ reporter driven by the indicated enhancers, representing the activities of the three *Atoh1* enhancers. Scale bars, 500 μm.

each other in the same cluster to orchestrate target gene expression. We and others have uncovered the functional hierarchy among constituent enhancers of each super-enhancer.[17,18,49–52] However, these studies relied on high-resolution genome-wide chromatin interactions, which remain difficult to capture for each cell type.

Here, we built enhancer networks to represent the regulatory relationship among a set of enhancers regulating the same gene. The enhancer networks can be computed based on all the analyzed cells or individual cluster of cells, using all cells can capture the major enhancer networks in the tissue, while using individual cluster of cells might be able to provide additional information specific for that cluster of cells (Figure S5H). In addi-

tion, enhancers within a network can be differentially accessible in different cell types/clusters, suggesting that different cell types expressing the same gene might rely on different regulatory elements (Figures S7I–S7K). We inferred PEIs based on chromatin co-accessibility among single cells, allowing us to see how individual elements interact with each other to regulate gene expression. Within an enhancer network, we found that enhancers were hierarchically organized as demonstrated that hub enhancers play more important roles in gene expression than non-hub enhancers by using *in vivo* KO analyses. Understanding how hub enhancers within the enhancer network sustain the hierarchical structure warrants further investigate in the future. Together, the enhancer network model provides a
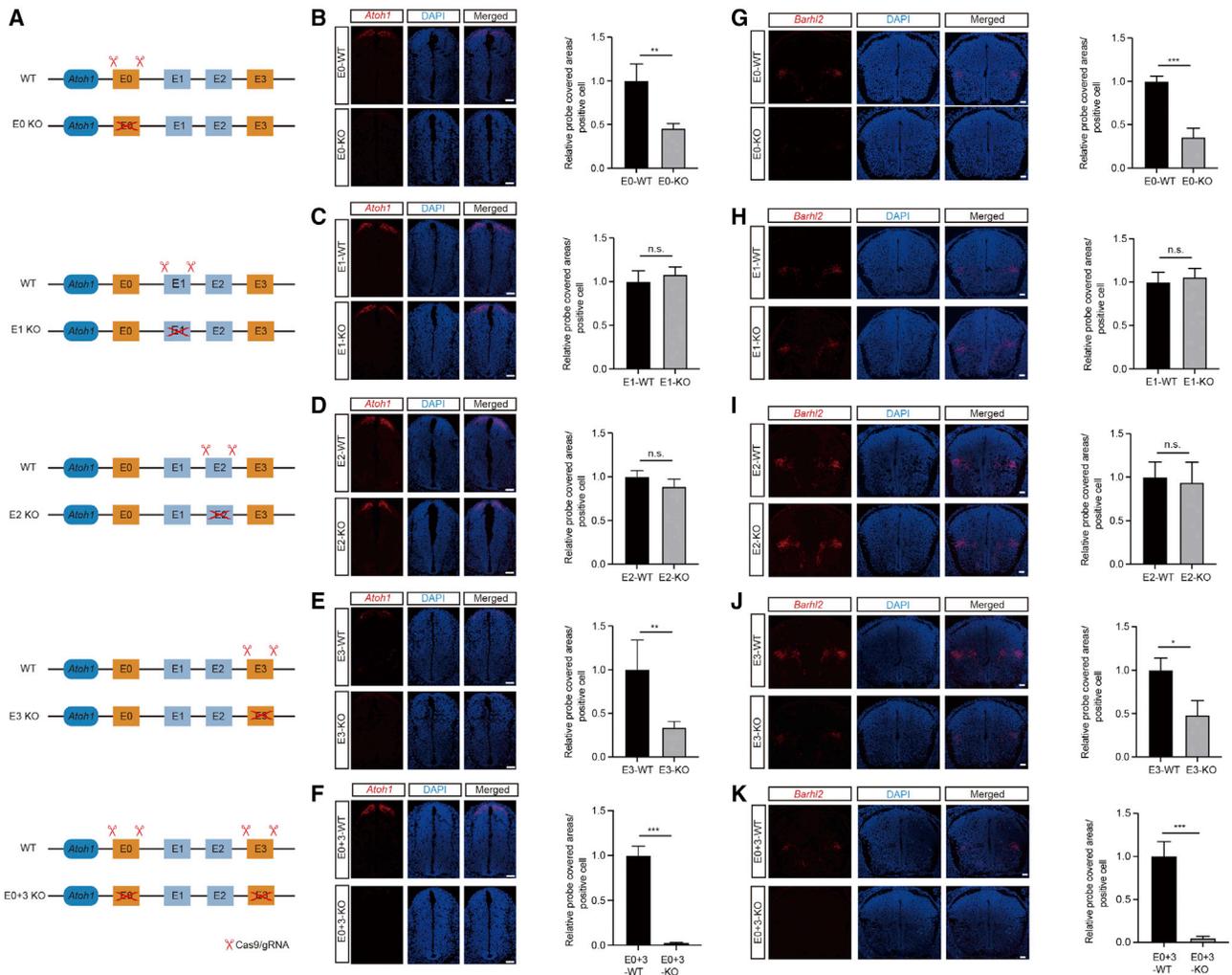
Figure 7. The functional hierarchy of *Atoh1* enhancer network validated by KO mice

(A) Strategy for creating enhancer KO mice with the CRISPR-Cas9 genome editing system.

(B–F) RNAscope detection of *Atoh1* mRNA in E0/E1/E2/E3/E0+E3 KO or WT mice at e10.5. Left: representative images show staining for *Atoh1* RNAscope probes (red) and DAPI (blue). Right: quantification of *Atoh1* RNAscope probe signal. Values shown are the mean ± standard error of the mean; $n \geq 3$ independent experiments. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; n.s., not significant (Student's t test). Scale bars, 50 μm. WT, wild type; KO, knockout.

(G–K) RNAscope detection of *Barhl2* mRNA in E0/E1/E2/E3/E0+E3 KO or WT mice at e13.5. Left: representative images show staining for *Barhl2* RNAscope probes (red) and DAPI (blue). Right: quantification of *Barhl2* RNAscope probe signal. Values shown are the mean ± standard error of the mean; $n \geq 3$ independent experiments. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; n.s., not significant (Student's t test). Scale bars, 50 μm. WT, wild type; KO, knockout.

systematic approach to simultaneously identify candidate genes important for cellular development and integral functional regulatory elements such as network hub enhancers.

We use *Atoh1*, a master regulator of dp1/dI1 neural development in the spinal cord,[44,53] as a model to dissect the regulatory role of its enhancer network. We proved that E1, E2, and E3 are *Atoh1* enhancers besides the known E0 in the spinal cord. Moreover, E0 and E3 are hub enhancers that are functionally more important than non-hub enhancers. We note a difference between the pattern of the *Atoh1* enhancer activity (Figure 6B). The combinatorial contributions of transcriptional regulation by enhancers and promoters, post-transcriptional regulation, and post-translational regulation collectively establish the expression pattern of Atoh1 protein. When endogenous *Atoh1* coding region is replaced by *β-gal*, the β-gal activity pattern,[44,53] repre-

senting the collective activity of endogenous *Atoh1* transcriptional activity, might show difference from Atoh1 protein pattern although largely overlap. Interestingly, E1, E2, and E3 activities from our transgenic LacZ reporters showed spinal cord activities, which supported that E1, E2, and E3 are *Atoh1* enhancers.

Atoh1 not only regulates the dp1/dI1 development in the spinal cord but also serves as the master regulator for hair cell development in cochlea and granule cell development in cerebellum.[54–56] E0 showed enhancer activity in the spinal cord, cochlea, and cerebellum.[40] However, the functional importance of E0 has never been experimentally studied. We here found that deletion of E0 caused mild downregulation of *Atoh1* in the spinal cord and severe cerebellum development defect (Figures S7G and S7H). However, deletion of E0 had no obvious effect on *Atoh1* expression or hair cell development in the cochlea.[57] These findings not only identify

that the functional importance of enhancers can be tissue specific for the same gene but also suggest that other enhancers besides E0 may cooperatively regulate the expression of *Atoh1* in a tissue-specific manner. Indeed, we found that the novel E3 could regulate *Atoh1* expression spatiotemporally in the spinal cord, both alone and in cooperation with E0. In contrast, E3 knockouts showed no obvious defect in cerebellum development (Figures S7G and S7H). These observations suggest that the enhancer network hierarchy can be tissue specific, pointing comparative analysis of enhancer networks in different tissues a promising direction for future research.

Limitations of the study

There are some limitations for this study. First, this study is based on the scATAC-seq and scRNA-seq analyses; however, these analyses lack the precise spatial information to fully characterize spatiotemporal gene regulation, which warrant further investigations in the future. Second, *Atoh1* plays important roles in spinal cord, cochlea, and cerebellum very likely using different enhancer networks. How enhancer networks are remodeled in different tissues or organs remains an interesting direction to be explored. Meanwhile, the molecular principles that build the enhancer networks are not clear either. Third, studies of more key regulatory elements identified in this study besides enhancer network for *Atoh1* are needed to better understand the regulation of mammalian spinal cord development.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - scATAC-seq
  - LacZ transgenic mice generation
  - X-Gal staining
  - Haematoxylin and eosin (H&E) and Immunohistochemistry (IHC) staining
  - Enhancer KO mice generation
  - Immunofluorescent staining, RNAscope and prepare of spinal cord sections
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - RNAscope data processing
  - scATAC-seq data pre-processing
  - Integration of scATAC-seq and scRNA-seq data as well as cell type annotation
  - Identification of cCREs
  - Characterization of cCREs
  - Build enhancer networks
  - Define network hub enhancers
  - Curation of cell identity and disease genes
  - Enrichment analysis of cell identity and disease genes

- Performance evaluation in predicting cell identity and disease genes
- Enrichment analysis of GWAS SNPs and validated enhancers
- Trajectory analysis of dI1 differentiation
- dI1 differentiation specific enhancer network
- *In silico* perturbation of the *Atoh1* enhancer network

AUTHOR CONTRIBUTIONS

J.D., F.L., J.H., M.S., and D.H. conceived and designed the experiments. J.Z. and M.S. performed scATAC-seq experiment. Z. Luo, Y.D., Z. Liu, and M.S. generated transgenic and knockout mice. Z.S., M.Y., and Y.Y. maintained the mouse strains. M.S. and S.B. performed staining experiments. D.H., H.L., and L.L. analyzed scATAC-seq and scRNA-seq data. M.S., D.H., H.L., J.H., and F.L. wrote the manuscript with the input from other authors.

REFERENCES

1. Ybot-Gonzalez, P., Gaston-Massuet, C., Girdler, G., Klingensmith, J., Arkell, R., Greene, N.D.E., and Copp, A.J. (2007). Neural plate morphogenesis during mouse neurulation is regulated by antagonism of Bmp signalling. Development *134*, 3203–3211. https://doi.org/10.1242/dev.008177.

2. Liem, K.F., Jessell, T.M., and Briscoe, J. (2000). Regulation of the neural patterning activity of sonic hedgehog by secreted BMP inhibitors expressed by notochord and somites. Development *127*, 4855–4866.

3. Mizutani, C.M., Meyer, N., Roelink, H., and Bier, E. (2006). Threshold-dependent BMP-mediated repression: a model for a conserved mechanism that patterns the neuroectoderm. PLoS Biol. *4*, e313. https://doi.org/10.1371/journal.pbio.0040313.

4. Zagorski, M., Tabata, Y., Brandenberg, N., Lutolf, M.P., Tkačik, G., Bollenbach, T., Briscoe, J., and Kicheva, A. (2017). Decoding of position in the developing neural tube from antiparallel morphogen gradients. Science *356*, 1379–1383. https://doi.org/10.1126/science.aam5887.

5. Alaynick, W.A., Jessell, T.M., and Pfaff, S.L. (2011). SnapShot: spinal cord development. Cell *146*, 178–178.e1. https://doi.org/10.1016/j.cell.2011.06.038.

6. Lee, S.-K., and Pfaff, S.L. (2001). Transcriptional networks regulating neuronal identity in the developing spinal cord. Nat. Neurosci. *4*, 1183–1191. https://doi.org/10.1038/nn750.

7. Delaunay, D., Heydon, K., Cumano, A., Schwab, M.H., Thomas, J.-L., Suter, U., Nave, K.-A., Zalc, B., and Spassky, N. (2008). Early neuronal and glial fate restriction of embryonic neural stem cells. J. Neurosci. *28*, 2551–2562. https://doi.org/10.1523/JNEUROSCI.5497-07.2008.

8. Briscoe, J., and Small, S. (2015). Morphogen rules: design principles of gradient-mediated embryo patterning. Development *142*, 3996–4009. https://doi.org/10.1242/dev.129452.

9. Lai, H.C., Seal, R.P., and Johnson, J.E. (2016). Making sense out of spinal cord somatosensory development. Development *143*, 3434–3448. https://doi.org/10.1242/dev.139592.

10. Snetkova, V., and Skok, J.A. (2018). Enhancer talk. Epigenomics *10*, 483–498. https://doi.org/10.2217/epi-2017-0157.

11. Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y., Lee, E.A., Zhu, Y., Plajzer-Frick, I., Pickle, C.S., et al. (2018). Enhancer redundancy provides phenotypic robustness in mammalian development. Nature *554*, 239–243. https://doi.org/10.1038/nature25461.

12. Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., and Duboule, D. (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. Science *340*. 1234167. https://doi.org/10.1126/science.1234167.

13. Honnell, V., Norrie, J.L., Patel, A.G., Ramirez, C., Zhang, J., Lai, Y.-H., Wan, S., and Dyer, M.A. (2022). Identification of a modular super-enhancer in murine retinal development. Nat. Commun. *13*, 253. https://doi.org/10.1038/s41467-021-27924-y.

14. Bahr, C., von Paleske, L., Uslu, V.V., Remeseiro, S., Takayama, N., Ng, S.W., Murison, A., Langenfeld, K., Petretich, M., Scognamiglio, R., et al. (2018). A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. Nature *553*, 515–520. https://doi.org/10.1038/nature25193.

15. Hay, D., Hughes, J.R., Babbs, C., Davies, J.O.J., Graham, B.J., Hanssen, L.L.P., Kassouf, M.T., Marieke Oudelaar, A.M., Sharpe, J.A., Suciu, M.C., et al. (2016). Genetic dissection of the α-globin super-enhancer in vivo. Nat. Genet. *48*, 895–903. https://doi.org/10.1038/ng.3605.

16. Thomas, H.F., Kotova, E., Jayaram, S., Pilz, A., Romeike, M., Lackner, A., Penz, T., Bock, C., Leeb, M., Halbritter, F., et al. (2021). Temporal dissection of an enhancer cluster reveals distinct temporal and functional contributions of individual elements. Mol. Cell *81*, 969–982.e13. https://doi.org/10.1016/j.molcel.2020.12.047.

17. Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T.V., Zon, L.I., Yuan, G.-C., et al. (2016). Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. Dev. Cell *36*, 9–23. https://doi.org/10.1016/j.devcel.2015.12.014.

18. Huang, J., Li, K., Cai, W., Liu, X., Zhang, Y., Orkin, S.H., Xu, J., and Yuan, G.C. (2018). Dissecting super-enhancer hierarchy based on chromatin interactions. Nat. Commun. *9*, 943. https://doi.org/10.1038/s41467-018-03279-9.

19. Chappell, L., Russell, A.J.C., and Voet, T. (2018). Single-cell (multi)omics technologies. Annu. Rev. Genomics Hum. Genet. *19*, 15–41. https://doi.org/10.1146/annurev-genom-091416-035324.

20. Kelsey, G., Stegle, O., and Reik, W. (2017). Single-cell epigenomics: recording the past and predicting the future. Science *358*, 69–75. https://doi.org/10.1126/science.aan6826.

21. Li, Y.E., Preissl, S., Hou, X., Zhang, Z., Zhang, K., Qiu, Y., Poirion, O.B., Li, B., Chiou, J., Liu, H., et al. (2021). An atlas of gene regulatory elements in adult mouse cerebrum. Nature *598*, 129–136. https://doi.org/10.1038/s41586-021-03604-1.

22. Ma, S., Zhang, B., LaFave, L.M., Earl, A.S., Chiang, Z., Hu, Y., Ding, J., Brack, A., Kartha, V.K., Tay, T., et al. (2020). Chromatin potential identified by shared single-cell profiling of RNA and chromatin. Cell *183*, 1103–1116.e20. https://doi.org/10.1016/j.cell.2020.09.056.

23. Sarropoulos, I., Sepp, M., Frömel, R., Leiss, K., Trost, N., Leushkin, E., Okonechnikov, K., Joshi, P., Giere, P., Kutscher, L.M., et al. (2021). Developmental and evolutionary dynamics of cis-regulatory elements in mouse cerebellar cells. Science *373*, abg4696.

24. Shema, E., Bernstein, B.E., and Buenrostro, J.D. (2019). Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution. Nat. Genet. *51*, 19–25. https://doi.org/10.1038/s41588-018-0290-x.

25. Trevino, A.E., Müller, F., Andersen, J., Sundaram, L., Kathiria, A., Shcherbina, A., Farh, K., Chang, H.Y., Paşca, A.M., Kundaje, A., et al. (2021). Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution. Cell *184*, 5053–5069.e23. https://doi.org/10.1016/j.cell.2021.07.039.

26. Blobel, G.A., Higgs, D.R., Mitchell, J.A., Notani, D., and Young, R.A. (2021). Testing the super-enhancer concept. Nat. Rev. Genet. *22*, 749–755. https://doi.org/10.1038/s41576-021-00398-w.

27. Hong, D., Lin, H., Liu, L., Shu, M., Dai, J., Lu, F., Tong, M., and Huang, J. (2022). Complexity of enhancer networks predicts cell identity and disease genes revealed by single-cell multi-omics analysis. Preprint at bioRxiv. https://doi.org/10.1093/bib/bbac508.

28. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat. Methods *18*, 1333–1341. https://doi.org/10.1038/s41592-021-01282-5.

29. Delile, J., Rayon, T., Melchionda, M., Edwards, A., Briscoe, J., and Sagner, A. (2019). Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. Development *146*. dev173807. https://doi.org/10.1242/dev.173807.

30. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

31. Mo, A., Mukamel, E.A., Davis, F.P., Luo, C., Henry, G.L., Picard, S., Urich, M.A., Nery, J.R., Sejnowski, T.J., Lister, R., et al. (2015). Epigenomic signatures of neuronal diversity in the mammalian brain. Neuron *86*, 1369–1384. https://doi.org/10.1016/j.neuron.2015.05.018.

32. Deneen, B., Ho, R., Lukaszewicz, A., Hochstim, C.J., Gronostajski, R.M., and Anderson, D.J. (2006). The transcription factor NFIA controls the onset of gliogenesis in the developing spinal cord. Neuron *52*, 953–968. https://doi.org/10.1016/j.neuron.2006.11.019.

33. Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser–a database of tissue-specific human enhancers. Nucleic Acids Res. *35*, D88–D92. https://doi.org/10.1093/nar/gkl822.

34. Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., et al. (2016). Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat. Genet. *48*, 1193–1203. https://doi.org/10.1038/ng.3646.

35. Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934–947. https://doi.org/10.1016/j.cell.2013.09.053.

36. Pliner, H.A., Packer, J.S., McFaline-Figueroa, J.L., Cusanovich, D.A., Daza, R.M., Aghamirzaie, D., Srivatsan, S., Qiu, X., Jackson, D., Minkina, A., et al. (2018). Cicero predicts cis-regulatory DNA interactions from single-cell chromatin accessibility data. Mol. Cell *71*, 858–871.e8. https://doi.org/10.1016/j.molcel.2018.06.044.

37. Piñero, J., Bravo, À., Queralt-Rosinach, N., Gutiérrez-Sacristán, A., Deu-Pons, J., Centeno, E., García-García, J., Sanz, F., and Furlong, L.I. (2017). DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. Nucleic Acids Res. *45*, D833–D839. https://doi.org/10.1093/nar/gkw943.

38. Barabási, A.-L., and Pósfai, M. (2016). Network Science (Cambridge University Press).

39. Mulvaney, J., and Dabdoub, A. (2012). Atoh1, an essential transcription factor in neurogenesis and intestinal and inner ear development: function, regulation, and context dependency. J. Assoc. Res. Otolaryngol. 13, 281–293. https://doi.org/10.1007/s10162-012-0317-4.

40. Helms, A.W., Abney, A.L., Ben-Arie, N., Zoghbi, H.Y., and Johnson, J.E. (2000). Autoregulation and multiple enhancers control Math1 expression in the developing nervous system. Development 127, 1185–1196.

41. Kothary, R., Clapoff, S., Darling, S., Perry, M.D., Moran, L.A., and Rossant, J. (1989). Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. Development 105, 707–714.

42. Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D., et al. (2006). In vivo enhancer analysis of human conserved non-coding sequences. Nature 444, 499–502. https://doi.org/10.1038/nature05295.

43. Rossant, J., Zirngibl, R., Cado, D., Shago, M., and Giguère, V. (1991). Expression of a retinoic acid response element-hsplacZ transgene defines specific domains of transcriptional activity during mouse embryogenesis. Genes Dev. 5, 1333–1344. https://doi.org/10.1101/gad.5.8.1333.

44. Bermingham, N.A., Hassan, B.A., Wang, V.Y., Fernandez, M., Banfi, S., Bellen, H.J., Fritzsch, B., and Zoghbi, H.Y. (2001). Proprioceptor pathway development is dependent on MATH1. Neuron 30, 411–422. https://doi.org/10.1016/S0896-6273(01)00305-1.

45. Yuengert, R., Hori, K., Kibodeaux, E.E., McClellan, J.X., Morales, J.E., Huang, T.-W.P., Neul, J.L., and Lai, H.C. (2015). Origin of a non-Clarke's Column Division of the dorsal spinocerebellar tract and the role of caudal proprioceptive neurons in motor function. Cell Rep. 13, 1258–1271. https://doi.org/10.1016/j.celrep.2015.09.064.

46. Hong, J.-W., Hendrix, D.A., and Levine, M.S. (2008). Shadow enhancers as a source of evolutionary novelty. Science 321, 1314. https://doi.org/10.1126/science.1160631.

47. Parker, S.C.J., Stitzel, M.L., Taylor, D.L., Orozco, J.M., Erdos, M.R., Akiyama, J.A., van Bueren, K.L., Chines, P.S., Narisu, N., et al.; NISC Comparative Sequencing Program (2013). Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc. Natl. Acad. Sci. USA 110, 17921–17926. https://doi.org/10.1073/pnas.1317023110.

48. Fulco, C.P., Nasser, J., Jones, T.R., Munson, G., Bergman, D.T., Subramanian, V., Grossman, S.R., Anyoha, R., Doughty, B.R., Patwardhan, T.A., et al. (2019). Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. Nat. Genet. 51, 1664–1669. https://doi.org/10.1038/s41588-019-0538-0.

49. Cai, W., Huang, J., Zhu, Q., Li, B.E., Seruggia, D., Zhou, P., Nguyen, M., Fujiwara, Y., Xie, H., Yang, Z., et al. (2020). Enhancer dependence of cell-type-specific gene expression increases with developmental age. Proc. Natl. Acad. Sci. USA 117, 21450–21458. https://doi.org/10.1073/pnas.2008672117.

50. Kai, Y., Li, B.E., Zhu, M., Li, G.Y., Chen, F., Han, Y., Cha, H.J., Orkin, S.H., Cai, W., Huang, J., et al. (2021). Mapping the evolving landscape of super-enhancers during cell differentiation. Genome Biol. 22, 269. https://doi.org/10.1186/s13059-021-02485-x.

51. Liu, X., Chen, Y., Zhang, Y., Liu, Y., Liu, N., Botten, G.A., Cao, H., Orkin, S.H., Zhang, M.Q., and Xu, J. (2020). Multiplexed capture of spatial configuration and temporal dynamics of locus-specific 3D chromatin by biotinylated dCas9. Genome Biol. 21, 59. https://doi.org/10.1186/s13059-020-01973-w.

52. Shin, H.Y., Willi, M., HyunYoo, K., Zeng, X., Wang, C., Metser, G., and Hennighausen, L. (2016). Hierarchy within the mammary STAT5-driven Wap super-enhancer. Nat. Genet. 48, 904–911. https://doi.org/10.1038/ng.3606.

53. Wang, V.Y., Hassan, B.A., Bellen, H.J., and Zoghbi, H.Y. (2002). Drosophila atonal Fully Rescues the Phenotype of Math1 Null mice. Curr. Biol. 12, 1611–1616. https://doi.org/10.1016/S0960-9822(02)01144-2.

54. Ben-Arie, N., Bellen, H.J., Armstrong, D.L., McCall, A.E., Gordadze, P.R., Guo, Q., Matzuk, M.M., and Zoghbi, H.Y. (1997). Math1 is essential for genesis of cerebellar granule neurons. Nature 390, 169–172. https://doi.org/10.1038/36579.

55. Ben-Arie, N., Hassan, B.A., Bermingham, N.A., Malicki, D.M., Armstrong, D., Matzuk, M., Bellen, H.J., and Zoghbi, H.Y. (2000). Functional conservation of atonal and Math1 in the CNS and PNS. Development 127, 1039–1048.

56. Bermingham, N.A., Hassan, B.A., Price, S.D., Vollrath, M.A., Ben-Arie, N., Eatock, R.A., Bellen, H.J., Lysakowski, A., and Zoghbi, H.Y. (1999). Math1: an essential gene for the generation of inner ear hair cells. Science 284, 1837–1841. https://doi.org/10.1126/science.284.5421.1837.

57. Luo, Z., Du, Y., Li, S., Zhang, H., Shu, M., Zhang, D., He, S., Wang, G., Lu, F., and Liu, Z. (2022). Three distinct Atoh1 enhancers cooperate for sound receptor hair cell development. Proc. Natl. Acad. Sci. USA 119. e2119850119. https://doi.org/10.1073/pnas.2119850119.

58. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag).

59. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. 32, 381–386. https://doi.org/10.1038/nbt.2859.

60. Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat. Methods 14, 975–978. https://doi.org/10.1038/nmeth.4401.

61. Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. InterJournal. Complex Syst. 1695, 1–9.

62. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinform. Oxf. Engl. 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

63. Xue, Y., Bao, Y., Zhang, Z., Zhao, W., Xiao, J., He, S., Zhang, G., Li, Y., Zhao, G., Chen, R., et al. (2022). Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2022. Nucleic Acids Res. 50, D27–D38. https://doi.org/10.1093/nar/gkab951.

64. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods 16, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0.

65. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. 28, 495–501. https://doi.org/10.1038/nbt.1630.

66. Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15, 1034–1050. https://doi.org/10.1101/gr.3715005.

67. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., et al. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 42, D1001–D1006. https://doi.org/10.1093/nar/gkt1229.

68. Satpathy, A.T., Granja, J.M., Yost, K.E., Qi, Y., Meschi, F., McDermott, G.P., Olsen, B.N., Mumbach, M.R., Pierce, S.E., Corces, M.R., et al. (2019). Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. Nat. Biotechnol. 37, 925–936. https://doi.org/10.1038/s41587-019-0206-z.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-PRRT2	Sigma-Aldrich	Cat# HPA014447; RRID: AB_1855786
IHC secondary antibody	Servicebio	Cat# GB23303; RRID: AB_2811189
DBA solution	Servicebio	Cat# G1211
Goat Anti-Rabbit IgG H&L	Abcam	Cat# ab150077; RRID: AB_2630356
Goat Anti-Chicken IgY H&L	Abcam	Cat# ab150173; RRID: AB_2827653
Chemicals, peptides, and recombinant proteins		
Hanks Balanced Solution	Life Technologies	Cat# 14175095
Hibernate E low fluorescence	Brain Bits	HELF5
Accutase	Sigma Aldrich	Cat# T200100
X-Gal	Vetec	Cat# V900468
$K_3Fe(CN)_6$	Sigma Aldrich	Cat# 244023
$K_4Fe(CN)_6 3H_2O$	Sigma Aldrich	Cat# P3289
Sodium deoxycholate	Sigma Aldrich	Cat# 30970
Paraformaldehyde	Sigma Aldrich	Cat# 158127
OCT Compound	SAKURA	Cat# 4583
Sucrose	Vetec	Cat# V900116
Formalin	Solarbio	Cat# G2161
mCas9	Sudgen	Cat# 91010
Critical commercial assays		
Chromium Next GEM Single Cell ATAC Library & Gel Bead Kit v1.1	10x Genomics	Cat# PN1000175
RNAscope Multiplex Fluorescent V2 Assay	Advanced Cell Diagnostics	Cat# 323100
2 × Phanta Max Master Mix (Dye Plus)	Vazyme	Cat# P525
Deposited data		
scATAC-seq	This paper	CRA005358
Code	This paper	https://github.com/xmuhuanglab/eNet
Experimental models: Organisms/strains		
E1-lacZ transgenic mice	This paper	N/A
E2-lacZ transgenic mice	This paper	N/A
E3-lacZ transgenic mice	This paper	N/A
E0 KO mice	This paper	N/A
E1 KO mice	This paper	N/A
E2 KO mice	This paper	N/A
E3 KO mice	This paper	N/A
E0+3 KO mice	This paper	N/A
Oligonucleotides		
Primers for transgenic mice, see Table S6	This paper	N/A
sgRNA, see Table S6	This paper	N/A
RNAscope probe Mm-Barhl2	Advanced Cell Diagnostics	492331
RNAscope probe Mm-Atoh1	Advanced Cell Diagnostics	408791
Recombinant DNA		
pHsp68-E1-lacZ plasmid	This paper	N/A
pHsp68-E2-lacZ plasmid	This paper	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pHsp68-E3-lacZ plasmid	This paper	N/A
Software and algorithms		
R(version 3.6.2)	N/A	https://www.r-project.org/
ggplot2(version 3.3.3)	Wickham[58]	https://cran.r-project.org/web/packages/ggplot2
Signac(version 1.1.1)	Stuart et al.[30]	https://satijalab.org/signac/
Seurat(version 3.2.3)	Stuart et al.[28]	https://satijalab.org/seurat/v4_changes
cicero(version 1.3.4.10)	Pliner et al.[36]	https://cole-trapnell-lab.github.io/cicero-release/
monocle3(version 0.2.1)	Trapnell et al.[59]	https://cole-trapnell-lab.github.io/monocle3/
ComplexHeatmap(version 2.7.1.1008)	N/A	https://github.com/jokergoo/ComplexHeatmap
chromVAR(version 1.8.0)	Schep et al.[60]	https://greenleaflab.github.io/chromVAR/
pheatmap(version 1.0.12)	N/A	https://CRAN.R-project.org/package=pheatmap
igraph(version 1.2.6)	Csardi and Nepusz[61]	https://igraph.org
rGREAT(version 3.0.0)	N/A	https://github.com/jokergoo/rGREAT
bedtools(version 2.30.0)	Quinlan et al.[62]	https://github.com/arq5x/bedtools2
GraphPad Prism	GraphPad Software	http://www.graphpad.com/; RRID: SCR_002798
Fiji	Fiji Software	https://fiji.sc; RRID: SCR_002285

RESOURCE AVAILABILITY

Lead contact
Further information and reagents' requests should be addressed to the lead contact, Jianwu Dai (jwdai@genetics.ac.cn).

Materials availability
All unique materials generated in this paper are available upon request to the lead contact without restriction.

Data and code availability
- All data reported in this paper will be shared by the lead contact upon request.
- All data generated in this study are freely accessible in GSA with the accession codes CRA005358 and OMIX (https://ngdc.cncb.ac.cn/omix/releaseList)[63] with the accession codes OMIX001822.
- All code used to analyze the data is available online at Github (https://github.com/xmuhuanglab/eNet), and archived at Zenodo (https://doi.org/10.5281/zenodo.7320541).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All mouse experiments were performed under the Chinese Ministry of Public Health Guidelines and the United States National Institutes of Health Guidelines for the care and use of animals. All the *in vivo* experiments were performed according to the ethical guidelines established by the Animal Care and Use Committee of the Institute of Genetics and Development Biology, Chinese Academy of Sciences. 3-4 mice were housed in a standard ventilated cage with a 12h light/12h dark cycle. The ambient temperature is 25°C with relative humidity of 50 %. All mice ad libitum access to water and chow.

METHOD DETAILS

scATAC-seq
The mouse embryos at the specified stages were generated from outbred CD1 mice. Observation of the vaginal plug was defined as e0.5. For neural tube dissection, cervical and thoracic sections of single mouse embryos were washed in Hanks Balanced Solution without calcium and magnesium (HBSS, Life Technologies, 14175095) three times, then dissected in Hibernate E low fluorescence (Brain Bits, HELF5). The samples were then incubated with Accutase (Sigma, T200100) for 15 min at 37 °C to dissociate the cells. To generate a single-cell suspension, an equivalent volume of phosphate-buffered saline (PBS) was added to the sample to stop the reaction. Single nuclei suspensions were prepared by following the nuclei isolation protocol for single cell ATAC sequencing (10x Genomics, CG000169), and scATAC-seq libraries were prepared according to the Chromium Next GEM Single Cell ATAC Library & Gel Bead Kit v1.1 (10x Genomics, PN1000175). To minimize potential multiplets, we typically aimed to capture ~8,000 nuclei per channel. Indexed libraries were pooled and sequenced on the Illumina Novaseq 6000 SP by Berry Genomics Corporation.

LacZ transgenic mice generation

LacZ transgenic mice were generated from CD1 background mice. Primers were designed to flank the candidate enhancer elements. Enhancer element constructs were PCR amplified and cloned into the Hsp68-LacZ vector (Vectorbuilder) to create pHsp68-E1-lacZ plasmid, pHsp68-E2-lacZ plasmid and pHsp68-E3-lacZ plasmid. DNA was isolated from mouse tails, boiled for 5 min in lysis solution (50 mM Tris HCl pH 8.0, 20 mM NaCl, 1 mM EDTA at pH 8.0, and 1% SDS), then screened via PCR with *lacZ* primers to identify transgenic animals. Plasmids, primers and transgenic mouse strain are listed in Table S6.

X-Gal staining

Embryos were dissected at e9.5, e10.5, e12.5, and e13.5. A small piece of tissue was removed from each embryo for genotyping by PCR examination of the *lacZ* gene. The embryos were fixed in cold 0.125% glutaraldehyde on ice (e9.5 for 20 min, e10.5 for 30 min, e12.5 for 60 min, and e13.5 for 70 min). Embryos were washed three times for 5 min each in PBS. Embryos were then moved to a freshly made X-Gal staining solution containing 100 mM phosphate buffer (pH 7.2), 2 mM $MgCl_2$, 5 mM EGTA, 0.02% NP40, 0.01% sodium-deoxycholate, 50 mM $K_3Fe(CN)_6$, 50 mM $K_4Fe(CN)_6$, and 1 mg/ml X-Gal. Samples were incubated in X-Gal staining solution at 37 °C. Incubation times varied from several hours to overnight, depending on the strength of β-Gal expression. After staining, embryos were washed three times with PBS and stored in freshly made 4% formaldehyde. The embryos were again washed with PBS three times prior to photographing. Primers and transgenic mouse strain are listed in key resources table.

Haematoxylin and eosin (H&E) and Immunohistochemistry (IHC) staining

Formaldehyde-fixed embryos were paraffin-embedded and sectioned transversely for 5-7 μm thick transverse sections. The sections were then deparaffinized and rehydrated. H&E histological staining was performed on these sections following standard protocols. For IHC, after antigen retrieval and blocking, sections were incubated overnight at 4 °C with anti-PRRT2 primary antibody (Sigma-Aldrich, HPA014447). Samples were then incubated with secondary antibody (Servicebio, GB23303) at room temperature for 30 minutes, followed by treatment with DBA solution (Servicebio, G1211).

Enhancer KO mice generation

The strains of *Atoh1* enhancer knockout mice were generated *via* CRISPR/Cas9 approach (Figure S7A). In brief, the Cas9 mRNA, sgRNA-1, and sgRNA-2 were microinjected into one-cell stage zygotes to generate E0 +/- mice (Figure S7B); sgRNA-3 and sgRNA-4 were used for generation of E1 +/- mice (Figure S7C); sgRNA-5 and sgRNA-6 were used for construction of E2 +/- mice (Figure S7D); sgRNA-7 and sgRNA-8 were used for construction of E3 +/- mice (Figure S7E). F0 mice with expected gene editing were screened by junction PCR (Vazyme, P525). After germline transmission by crossing with WT mice, the F1 mice were further identified by junction PCR and confirmed by Sanger sequencing. E0+3 KO mice were generated from E0 KO mice (Figure S7B), and sgRNA-9 and sgRNA-10 were used to generate the dual KO mice (Figure S7F). Sequences of sgRNAs and primers are listed in the Table S6.

Immunofluorescent staining, RNAscope and prepare of spinal cord sections

Mouse spinal cord tissues were fixed in fresh 4% paraformaldehyde in PBS for 16–24 h, cryoprotected in 30% sucrose, embedded in Optimal Cutting Temperature (O.C.T) compound, and 10-μm sections taken. Sections were stored at -80 °C. Primary GFP antibodies (Abcam, ab13970 and ab290) and secondary antibodies (Abcam, ab150077 and ab150173) were diluted 1:500. Images were acquired on a Leica SP8 microscope. Individual optical slices were obtained. RNAscope was performed as instructed by the manufacturer (ACD, CA, USA).

QUANTIFICATION AND STATISTICAL ANALYSIS

RNAscope data processing

Quantification was performed using Fiji software. Statistical analysis was performed using GraphPad Prism software. Data are presented as the mean ± SEM. *p* value was calculated using GraphPad Prism software by two-tailed Student's t-test. *p* values are presented in indicated figures as appropriate. Statistical methods are indicated in figure legends.

scATAC-seq data pre-processing

Sequencing reads were aligned to the mm10 reference genome, and a cell-by-peak matrix was generated from fastq files using the software CellRanger (version 1.2.0) with default parameters. Cell barcode matrices from different time points (e9.5, e10.5, e12.5 and e13.5) were then pooled using CellRanger ATAC aggr to aggregate these into a single aggregated peak-barcode matrix. All analyses described below were performed on the aggregated dataset. Quality control at the library level was performed by checking the expected periodicity in the frequency of fragment sizes (Figure S1A) and enrichment of insertions around annotated TSSs (Figure S1B). After converting the cell-by-peak matrix into a Seurat object with Signac, we performed quality control for cell barcodes. Barcodes corresponding to cell nuclei were distinguished from empty droplets by requiring at least 3,500 fragments and a minimum TSS enrichment of 2. A total of 19,715 barcodes passed this step. Dimensionality reduction, clustering and gene activity scores were determined using standard processes in Seurat and Signac.[28] To control for batch effects in the different developmental stages, we applied the Signac implementation of Harmony[64] using the RunHarmony function with default parameters (Figure S1C) and compared the cell composition between the clusters with and without removing batch effect (Figure S1D).

Integration of scATAC-seq and scRNA-seq data as well as cell type annotation

We integrated our scATAC-seq and scRNA-seq from a previous study and annotated cell types and subtypes using Seurat package. The process steps were shown as following: 1) To improve the computational efficiency and accuracy, we performed the integration in a stage-wise manner, matching developmental stages between studies (Figures S2A and S2B). We first obtained gene-by-barcode matrices of UMI counts (ArrayExpress accession number: E-MTAB-7320) and reprocessed the data separately for each stage using a standard Seurat pipeline (v3.0.1). Briefly, we used data only from high quality cells as previously filtered by the authors, applied SCTransform to scale the data and identify highly variable genes, and reduced the dimensions of the data to 50 principal components. We processed our scATAC-seq data separately for each developmental stage as described above (see "scATAC-seq pre-processing"). 2) We performed canonical correlation analysis (CCA) to generate a shared dimensionality reduction of the 'query' scATAC-seq gene activity calculated by Signac and the 'reference' scRNA-seq gene expression. 3) We identified pairs of corresponding cells using the highly variable genes in two datasets that anchor the two datasets together. 4) To transfer cell type and subtype annotations from scRNA-seq to scATAC-seq cell populations, we used Seurat's label-transfer algorithm to leverage these anchors to predict cell types in scATAC-seq data. 5) To generate gene expression and chromatin accessibility in the same cell, we transferred the UMI counts from scRNA-seq to the scATAC-seq by using Seurat's label-transfer algorithm. 6) To assess the utility of gene scores for predicting gene expression in our data, we focused on the genes used for integration (highly variable genes in the scRNA-seq data with gene score estimates in the scATAC-seq data) and averaged expression estimates across clusters in the scATAC-seq data. We then calculated Pearson correlation coefficient per gene between the gene score and imputed RNA expression across clusters. We observed robust correlations between the two estimates (median Pearson's r across all genes and stages: 0.51; Figure S1G).

Identification of cCREs

After determining cell clusters from the scATAC-seq data, we identified cell cluster-specific *cis*-regulatory elements (cCREs) that showed specific chromatin accessibility in each cluster. To this end, we used the function FindMarkers() of Signac(v1.1.1), which performed a differential accessibility (DA) test between each cluster of cells and the rest of cells. The output of each test is log fold-change of the average accessibility (avg_logFC) between the two groups and adjusted *p-value* based on bonferroni correction using all peaks in the dataset. We selected the CREs specific accessible in each cell cluster by using the cutoff avg_logFC > 0.25 (recommended by Signac). In total, we generated 82,163 cCREs and grouped them into 23 modules. To determine the mapping between cCREs and modules, we calculated the average chromatin accessibility of each cCRE in each cluster, then assigned the cluster where the cCRE showed the highest accessibility to group the cCRE to the corresponding module. Thus, each cCRE module represented a cluster, e.g., cCREs in M1 were specific accessible in C1, etc.

Characterization of cCREs

To associate CREs in a cCRE module with the biological processes (BP) of their putative target genes, we used the R implementation of GREAT (rGREAT, 1.18.0)[65] to identify enriched gene ontologies. To determine peak overlap with TF motifs, we then performed a hypergeometric test to determine the probability of observing the motif at the given frequency by chance using Signac (v1.1.1), comparing with a background set of peaks matched for GC content or the rest of the genome. To calculate the cCRE motif activity for each cell, we applied the Signac implementation of chromVAR[60] using the RunChromVAR function with default parameters. To assess sequence constraint in CREs, we used estimates of evolutionary conservation (phastCons scores) based on multiple alignments. We downloaded phastCons 100-way vertebrate conservation scores for mm10 as bigWig files from the UCSC table browser.[66] The mean phastCons score within the cCRE was considered as a metric of constraint for each cCRE.

Build enhancer networks

To explore the underlying mechanism of enhancer clusters in regulating gene expression, we developed eNet to build enhancer networks based on scATAC-seq and scRNA-seq datasets.[27] Briefly, it contains the following steps: 1) We prepared the enhancer accessibility and gene expression matrix in single cells as the input data of eNet. 2) We identified a putative enhancer cluster for each gene based on the correlation between gene expression and enhancer accessibility. To this end, we first calculated correlation of gene expression and enhancer accessibility which are within gene TSS ±100 kb. Then we identified significantly corelated enhancers as the putative enhancers for each gene by adapting the method previously described,[21] with some modifications that we changed a 1-Mb window centred on the TSS for each gene to a 200 kb window. 3) We identified the predicted enhancer interactions (PEIs, if the co-accessibility of an enhancer pair is more than 0.2, the default parameter recommended by Cicero, the enhancer pair is considered as an PEI) between enhancers within each putative enhancer cluster using Cicero.[36] Then, we built an enhancer network, where nodes represent enhancers in the enhancer network and edges represent PEIs. 4) We quantified the complexity of the enhancer networks by the network size and connectivity. Network size was calculated as the number of enhancers. Network connectivity was calculated as two-fold of the number of edges divided by the number of nodes. 5) We classified the enhancer networks into several modes based on their complexity: Simple, Multiple, Complex and others; the last was discarded due to the very small number of networks it contained (Figure 4C). Intuitively, a gene in Simple mode was regulated simply by 1-5 number of enhancers; a gene in Multiple mode was regulated by more than 5 enhancers but limited PEIs below the threshold (network connectivity < 1); and a gene in Complex mode was regulated by more than 5 enhancers and frequent PEIs pass the threshold (network connectivity ≥ 1).

Define network hub enhancers

Briefly, for enhancers in Complex mode, we counted the number of PEIs associated with each enhancer (node degree, refers to the number of edges associated with the node) and divided them by the total number of PEIs in the enhancer network, termed as normalized node degree. By applying a threshold value of the normalized node degree 1/3, the 0.95 quantile of all the normalized node degree, we divided the enhancers into two groups, termed as network hub enhancers (normalized node degree $\geq$ 1/3) and non-hub (normalized node degree < 1/3) enhancers. Of note, hub enhancers exhibit greater accessibility and larger size, but comparable distance to target genes (Figure S4G). Comparing hub and non-hub enhancers with similar chromatin accessibility or genomic spanning regions, the hub enhancers show significantly higher PEIs than non-hub enhancers, indicating that the enhancer interactions provide additional information beyond the degree of chromatin accessibility and length of regions (Figures S4H and S4I).

Curation of cell identity and disease genes

Cell identity genes related to neurons were obtained from the "Table S1. Knowledge matrix used to identify cell types" file from Delile et al.[29] The disease genes related to neurons were downloaded from DisGeNET.[37] These genes are all shown in Table S4.

Enrichment analysis of cell identity and disease genes

We performed enrichment analysis for cell identity and disease genes in each of the three modes (Complex, Multiple and Simple). In short, given a group of genes in each mode, an enrichment score was calculated as the fold-enrichment relative to the whole genome as a background. The calculation method was as follows:

$$(m\,/\,n)/(M\,/\,N)$$

where m and M represent the number of cell identity genes in the group and the whole genome, respectively, and n and N represent the number of genes in the group and the whole genome, respectively.

Performance evaluation in predicting cell identity and disease genes

To assess the enhancer network performance in estimating cell identity and disease-related genes, we used various scoring methods to rank all genes, including network connectivity, network size, and overall chromatin accessibility. Then, using the whole genome as background, we applied a moving window of 100 to calculate the fold-enrichment of cell identity or disease genes among the top-ranked genes.

Enrichment analysis of GWAS SNPs and validated enhancers

GWAS SNPs

The GWAS Catalog[67] was downloaded from the UCSC Table Browser (http://genome.ucsc.edu/). Furthermore, based on a semi-automatic text mining method as described in our previous work,[17] a list of cell type-specific GWAS SNPs was curated.

Neuron-related GWAS SNPs

The subset of neuron-related GWAS SNPs was selected as those associated with at least one of the following keywords in the 'trait' field: 'Amyotrophic lateral sclerosis', 'Parkinson's disease', 'Attention deficit', 'Anorexia', 'Type 1 diabetes', 'Ulcerative colitis', 'Menarche', 'Depressed affect', 'Intelligence', 'sclerosis', 'Insomnia', 'Menopause', 'Artery disease', 'Educational attainment', 'Cerebral', 'Ischemic', 'Spastic Diplegia', 'Malaria', 'Aneurysm', 'Cortex', 'Spastic Quadriplegia', 'Band Heterotopia', 'Cerebrovascular Disease', 'Arteriovenous Malformations of the Brain', 'Spastic Hemiplegia', 'Intracranial Embolism', 'Brain Edema', 'Brain Injury', 'Adrenoleukodystrophy', 'Intracranial Thrombosis', 'Seizure Disorder', 'Depression', 'Encephalopathy', 'Arteriovenous Malformation', 'Cardiac Arrest', 'Cerebritis', 'Mitochondrial Dna Depletion Syndrome 4a', 'Hypoxia', 'Thrombosis', 'Developmental and Epileptic Encephalopathy 39', 'Hemorrhage, Intracerebral', 'Schizophrenia', and 'Spasticity'.

VISTA enhancers

The VISTA enhancer database curates experimentally validated human and mouse noncoding fragments with gene enhancer activity as assessed in transgenic mice.[33] We downloaded the mouse VISTA regions (mm9 coordinates) from the VISTA database (https://enhancer.lbl.gov/), then curated a list of neuron-associated VISTA enhancers that show activity in the neural tube.

Enrichment analysis

We converted the enhancers from mm10 to hg38 or mm9 genomic coordinates using liftOver software from UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgLiftOver). The overlap between loci and GWAS SNPs or VISTA enhancers was performed using bedtools intersect.[62] In short, for enhancers in each group, the enrichment score was calculated as the fold-enrichment relative to the genome background. The computing method was as follows:

$$(m\,/\,n)/(M\,/\,N)$$

where m and M represent the number of within-group and genome-wide SNPs, respectively, and n and N represent the number of within-group and genome-wide loci, respectively. The genome-wide background was generated from a list of loci obtained by randomly shuffling the list of regular enhancers.

Trajectory analysis of dI1 differentiation

We performed trajectory analysis for dI1 cell differentiation using a previously described method[68] to order cells in pseudotime. In brief, we first described candidate trajectories by ordering clusters (C2, C3, C8 and C23). Next, for each cluster, we computed the mean coordinates in both UMAP1 and UMAP2 dimensions, cells in the top 5% Euclidean distance to the mean coordinates were filtered. We then calculated the UMAP distance for each cell from the cluster i to the mean coordinates of cluster i+1 along the trajectory. In the next step, we computed a pseudotime vector by calculating the quantiles for each cell by their distance to the next cluster, then fitted a successive trajectory to both UMAP coordinates using the function 'smooth.spline'. Then, we aligned all cells to the trajectory by their Euclidean distance to the nearest point along the manifold. Finally, we scaled this alignment to 100, and that was used as pseudotime for further analyses.

We assessed the significance of the trajectory by its ordering of clusters for further supporting longer trajectories in inferred pseudotime. In brief, we extracted the lattermost cluster, ranked the top 10,000 accessible elements, and calculated the Euclidean distance between this cluster and the remaining clusters using logCPM. Next, we applied this computing method to all other clusters in the reverse trajectory, excluding the previous clusters for directionality. To calculate the significance of the ordering, we counterchanged the order 5,000 times; we then calculated the average rank of the ordering for the counterchanged and input trajectories, which allowed for computation of the empirical *p-value* that we could assign to every reduced dimension trajectory from the initial accessibility matrix (Figure S5A).

dI1 differentiation specific enhancer network

We built the enhancer networks using the cells in dI1 differentiation (C2, C3, C8, C23) following eNet analysis the method (see 'build enhancer network').

In silico perturbation of the *Atoh1* enhancer network

To quantitatively estimate the function of enhancers in *Atoh1* enhancer networks, we first built the weighted enhancer networks using the cells in dI1 differentiation (C2, C3, C8, C23) following eNet analysis the STAR Methods (see 'build enhancer network'), except assigning chromatin co-accessibility as weight of edges considering limited cells involved in here. In a weighted enhancer network, as illustrated in Figure 5E, the nodes represented the putative enhancers, and the edges represented PEIs, the width of the edges indicated the co-accessibility of the PEI. The network connectivity score is calculated by the sum of PEIs' co-accessibility divided by number of enhancers. Then we performed *in silico* analysis to mimic the effect of enhancer knockout by assessing changes of network connectivity score after removing a specific enhancer. For example, we compared the network connectivity score of the *Atoh1* enhancer networks before (WT) and after (E0 KO) removing E0 and associated PEIs (Figures 5E and S5G).