Supplemental Information

Supplemental Figures

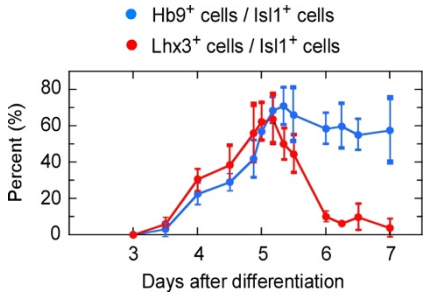

Figure S1. Temporal Changes in Lhx3 Expression during Hypaxial Motor Neuron Differentiation from ES cells, Related to Figure 1

Embryoid bodies were fixed at specified timepoints between day 3 and day 7 of differentiation, immunostained for spinal motor neuron markers Hb9, Lhx3, and Isl1, and immuno-positive cells were counted. Shown is the percentage of all Isl1-positive cells expressing Hb9 (blue) or Lhx3 (red). Data are presented as the mean ± SEM (average eight embryoid bodies per timepoint were counted; Supplemental Experimental Procedures).

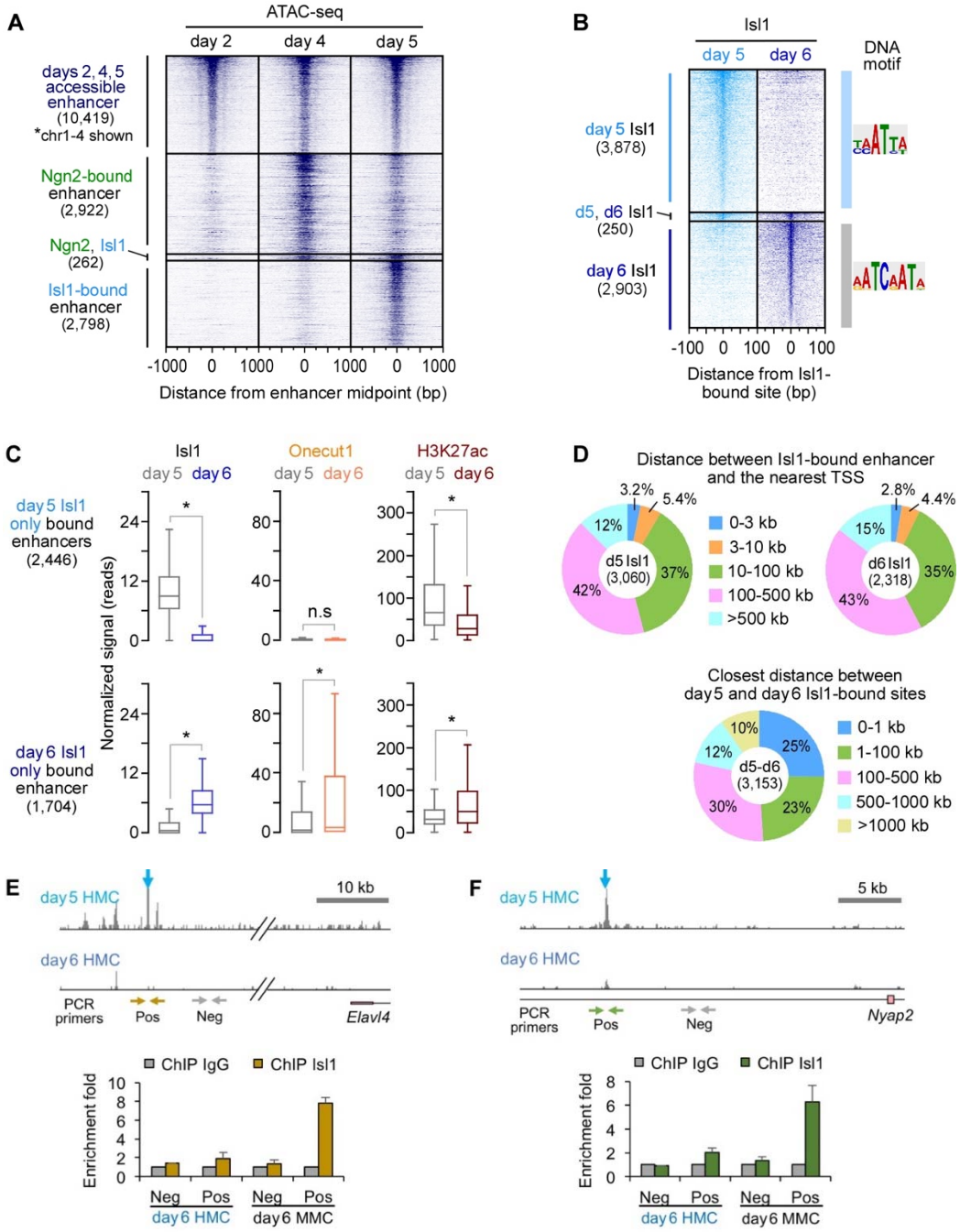

Figure S2. Dynamic Organization of TF Binding and Chromatin Accessibility in Progenitors and Postmitotic Motor Neurons, Related to Figure 2

(A) Chromatin accessibility on days 2, 4, and 5 of differentiated cells, detected by ATAC-seq relative to the midpoint of day 4 Ngn2-bound and day 5 Isl1-bound enhancers. Upper group shows a set of maintained distal enhancers (*n*=10,419) that are accessible on days 2, 4, and 5 (>2.5 kb from TSS; 2,907 enhancers on chromosomes 1 to 4 are shown). These maintained enhancers were used for the normalization of the ATAC-seq data sets across the three timepoints. Middle and lower groups show day 4 Ngn2-bound (*n*=262) and day 5 Isl1-bound enhancers (*n*=2,798), ordered as shown in **Figure 2A**. Day 2, primitive ectoderm; day 4, progenitor; day 5, nascent motor neuron (MN)

(B) Isl1 occupancy relative to the Isl1-bound sites (7,018 rows) in day 5 nascent and day 6 maturing postmitotic MNs, sorted and ordered by Isl1 occupancy, detected by ChIP-exo. Isl1-bound sites were identified by peak calls (Supplemental Experimental Procedures). Note that an Isl1-bound site is defined as 22-28 bp of protected region from exonuclease, whereas an Isl1-bound enhancer shown in **Figure 2C** represents 1 kb region around the Isl1-bound site. If multiple Isl1-bound sites reside within 1 kb, we merged them into a single enhancer. The right panel

shows DNA sequence motifs, which are the enriched DNA sequences located within Isl1-bound sites (22-28 bp width).
(C) Boxplots of TF occupancy measured by ChIP-exo, and H3K27ac intensity measured by ChIP-seq in day 5 nascent and day 6 maturing MNs. Enhancer groups were classified in **Figure 2C**. Occupancy or intensity between day 5 and day 6 MNs was normalized by making a median occupancies (or intensities) of the maintained enhancers as a normalizer. Boxplots show the median (line), second to third quartiles (box), 1.5x the interquartile range (whiskers). *$P<1\times10^{-5}$, n.s, not significant ($P>0.01$); two sample *t*-test.
(D) *Upper*, Distance (kb) between Isl1-bound enhancer and the nearest transcription start site (TSS). Isl1-bound enhancers in day 5 nascent MNs (n=3,060) and in day 6 maturing MNs (n=2,318), shown in **Figure 2C**. The majority of Isl1-bound enhancers reside at distal enhancer regions in both nascent and maturing MNs. *Lower*, Distance (kb) between the closest day 5 Isl1 and day 6 Isl1-bound sites (3,153 pairs of closest day 5 and day 6 Isl1-bound sites), shown in panel **B**. Only 25% (791/3,153) of day 6 Isl1-bound 'sites' (22-28 bp width) were located within 1 kb from day 5 Isl1-bound 'sites'. Note that 26% (614/2,318) of day 6 Isl1-bound 'enhancers' (1 kb width) were located within 1 kb from day 5 Isl1-bound 'enhancers', which were defined as the maintained (stable) day 5 and day 6 Isl1-bound enhancers in **Figure 2C**.
(E) Screenshot of ChIP-exo Isl1 data showing the location of day 5 nascent MN-specific enhancers proximal to a neuron-specific RNA binding protein gene, *Elavl4,* in hypaxial motor (HMC) neurons. ChIP assays for Isl1 binding at one of day 5-specific enhancers (cyan arrow, ~70 kb downstream of *Elavl4*) in maturing (day 6) HMC and maturing medial motor (MMC) neurons. Brown and grey arrows indicates positive and negative PCR primer sets, respectively. Error bars represent the s.d. n=3. Isl1 binding at the nascent MN enhancer is maintained in maturing MMC neurons, but not in maturing HMC neurons.
(F) Same as panel **E** except day 5 HMC specific enhancer at a neuronal tyrosine-phosphorylated phosphoinositide-3-kinase adaptor gene, *Nyap2* (n=3).

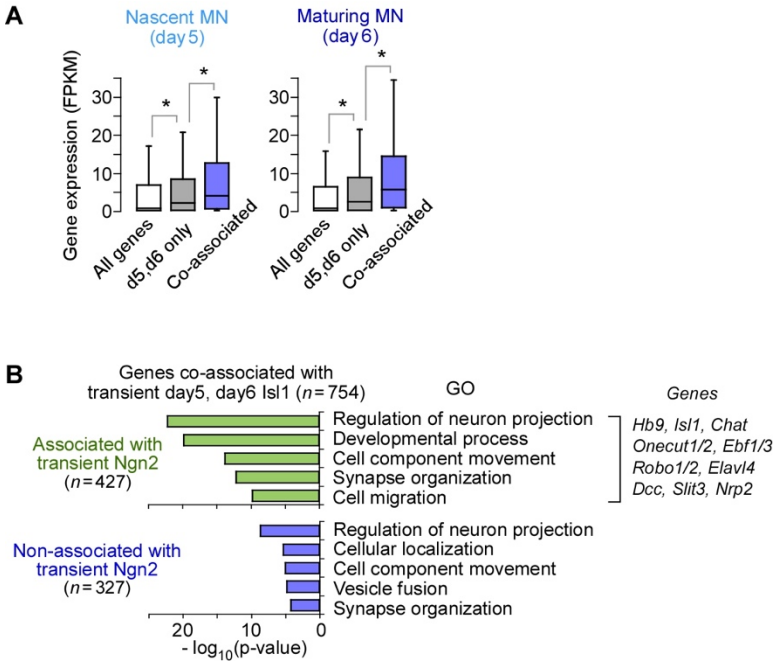

Figure S3. Stage-Specific Transient Enhancers Specify and Maintain Motor Neuron Effector Genes, Related to Figure 3

(A) Boxplots of RNA expression for genes associated only with day 5 or day 6 Isl1-bound enhancers (1,461 genes, class III), genes co-associated with both day 5 and day 6 Isl1-bound enhancers (885 genes, class I or II), shown in **Figure 3B**, and all genes (n=19,326) (Table S4). Boxplots show the median (line), second to third quartiles (box), 1.5x the interquartile range (whiskers), *$P<1\times10^{-5}$, n.s, not significant ($P>0.05$); Wilcoxon rank-sum test.

(B) Top five non-redundant GO terms for 427 genes that are co-associated with transient day 5 and day 6 Isl1- and transient day 4 Ngn2-bound enhancers, and for 327 genes that are associated with transient day 5 and day 6 Isl1-bound enhancers, but not associated with transient day 4 Ngn2-bound enhancers. Representative motor neuron effector genes were also shown.

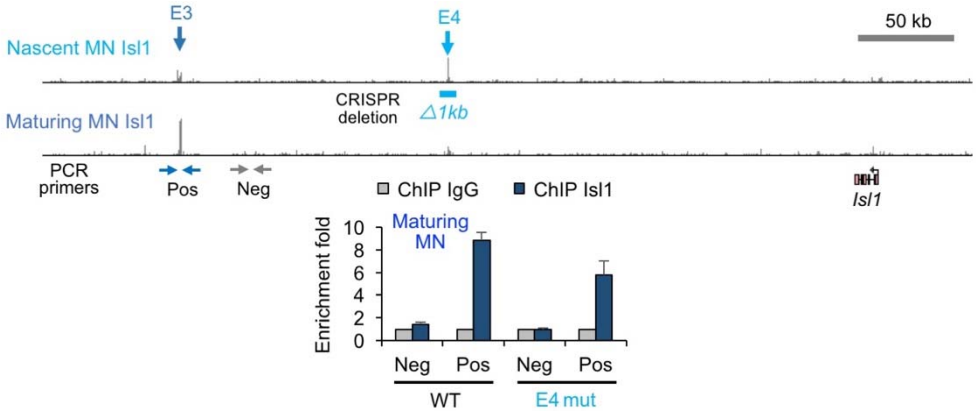

Figure S4. Functional Independence of Stage-Specific Motor Neuron Enhancers, Related to Figure 5
Independent Isl1 binding at nascent and maturing MN enhancers of the same gene. *Upper*, The locations of stage-specific transient enhancers proximal to *Isl1* gene were described in **Figure 5B**. Nascent MN-specific Isl1-E4 enhancer was deleted using the CRISPR genome editing. Screenshot shows intensity of ChIP-exo Isl1 in nascent and maturing WT MNs. *Lower*, ChIP assays for Isl1 binding at the maturing MN-specific Isl1-E3 enhancer in maturing MNs, differentiated from WT and the mutant ES cell line containing a targeted deletion of the Isl1-E4 enhancer. Horizontal arrows indicate positive (blue; Pos) and negative (gray; Neg) PCR primer sets, respectively. Error bars represent the s.d., *n*=3, two independent differentiations.

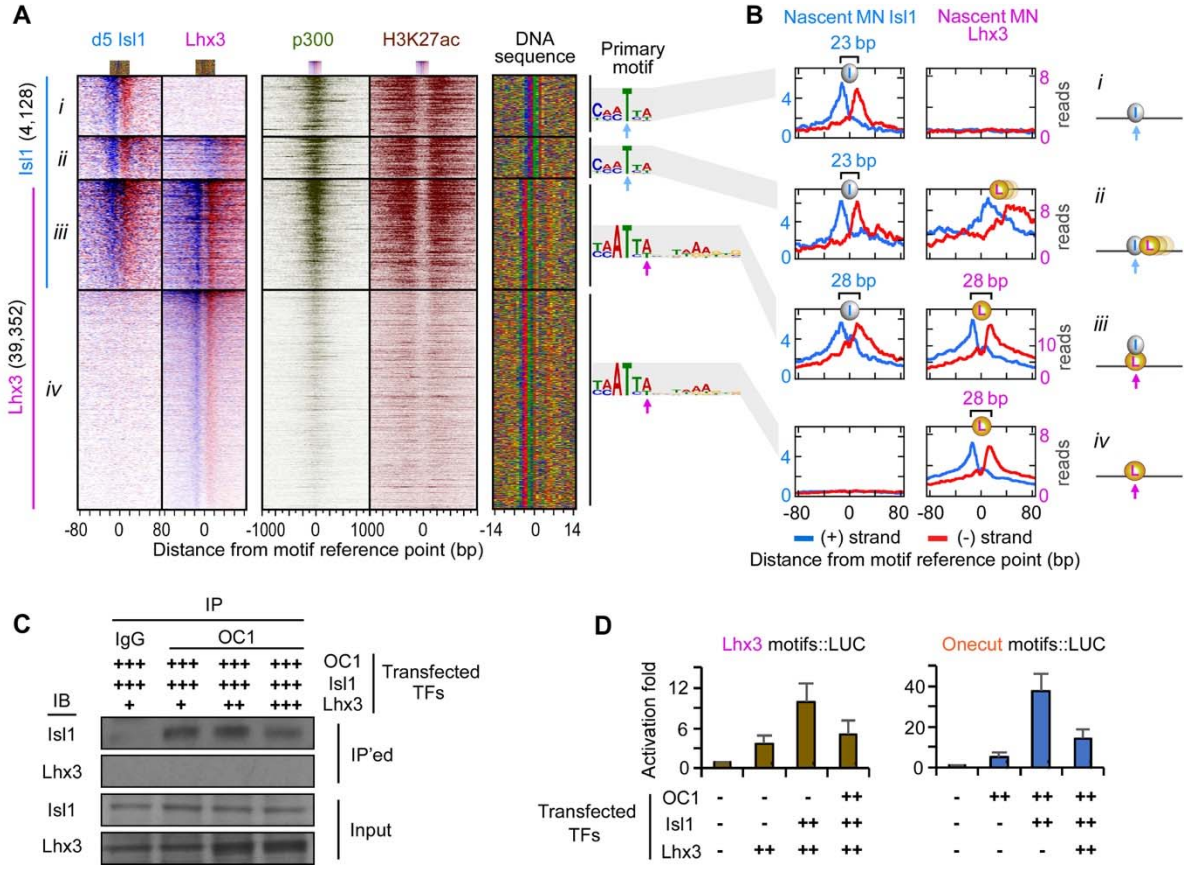

Figure S5. Isl1 is Recruited to DNA through Interactions with Lhx3 in Nascent Motor Neurons, Related to Figure 6

(A) Isl1, Lhx3, p300, and H3K27ac intensity relative to the TF motif reference point of Isl1-bound sites in day 5 nascent motor neurons (MNs) or Lhx3-bound sites in iNIL programmed spinal MNs (Mazzoni et al., 2013), sorted and ordered by TF occupancy (Table S6). TF occupancy on the (+) strand (blue, left border) and (-) strand (red, right border) is shown. We classified four subsets: (*i*) 1,127 Isl1 only bound sites, (*ii*) 805 Isl1 next to Lhx3-bound sites (Lhx3 was reoriented to the right side of Isl1 regardless the reference Isl motif orientation), (*iii*) 2,205 Isl1/Lhx3 co-bound sites (*iv*) 37,147 Lhx3 only bound sites (10,426 sites on chromosomes 1 to 4 are shown). Right panel shows a color chart representation of the DNA sequence located ±14 bp from the motif reference point, ordered as shown in the left panels. Cyan and magenta arrows indicate the motif reference point of Isl and Lhx3 DNA-binding motif, respectively.

(B) Composite (average read counts) plots of Isl1 and Lhx3 binding occupancy on the (+) and (-) strand shown in panel **A**. Right panel shows models of Isl1 and Lhx3 binding to DNA in nascent MN enhancers.

(C) Biochemical demonstration of Onecut1 (OC1) and Lhx3 competition for interaction with Isl1. ES cells were transfected with Onecut1- and Isl1-expressing vectors with increasing amount of Lhx3-expressing vector. 24 hours after transfection, immunoprecipitation (IP) was performed with Onecut1 and IgG antibodies, followed by an immunoblotting (IB) with Isl1 and Lhx3 antibodies, showing Isl1 bands (~39 kD), but not Lhx3 bands (~43 kD) in IP'ed samples. Input, 2.5% whole cell lysate; IgG, a negative control. Shown is a representative blot (*n*=2).

(D) *Left*, Luciferase reporter assays with the synthetic enhancer containing 3-Lhx3-binding motifs (indicated in **Figure 7E**) in ES cells, transfected with Lhx3- and Isl1-expressing vectors with a competitive expression of Onecut1. Error bars represent the s.d., *n*=3. *Right*, Luciferase reporter assays with the Isl1-E3 enhancer (286 bp; Supplemental Experimental Procedures) containing 3-Onecut-binding motifs in ES cells, transfected with OC1- and Isl1-expressing vectors with a competitive expression of Lhx3. Error bars represent the s.d. *n*=3. The empty vector (-) without containing gene encoding TF of interest was used to adjust total amount of vector same in each sample. Luciferase activity of cells transfected only with the empty vector was used as a normalizing control.

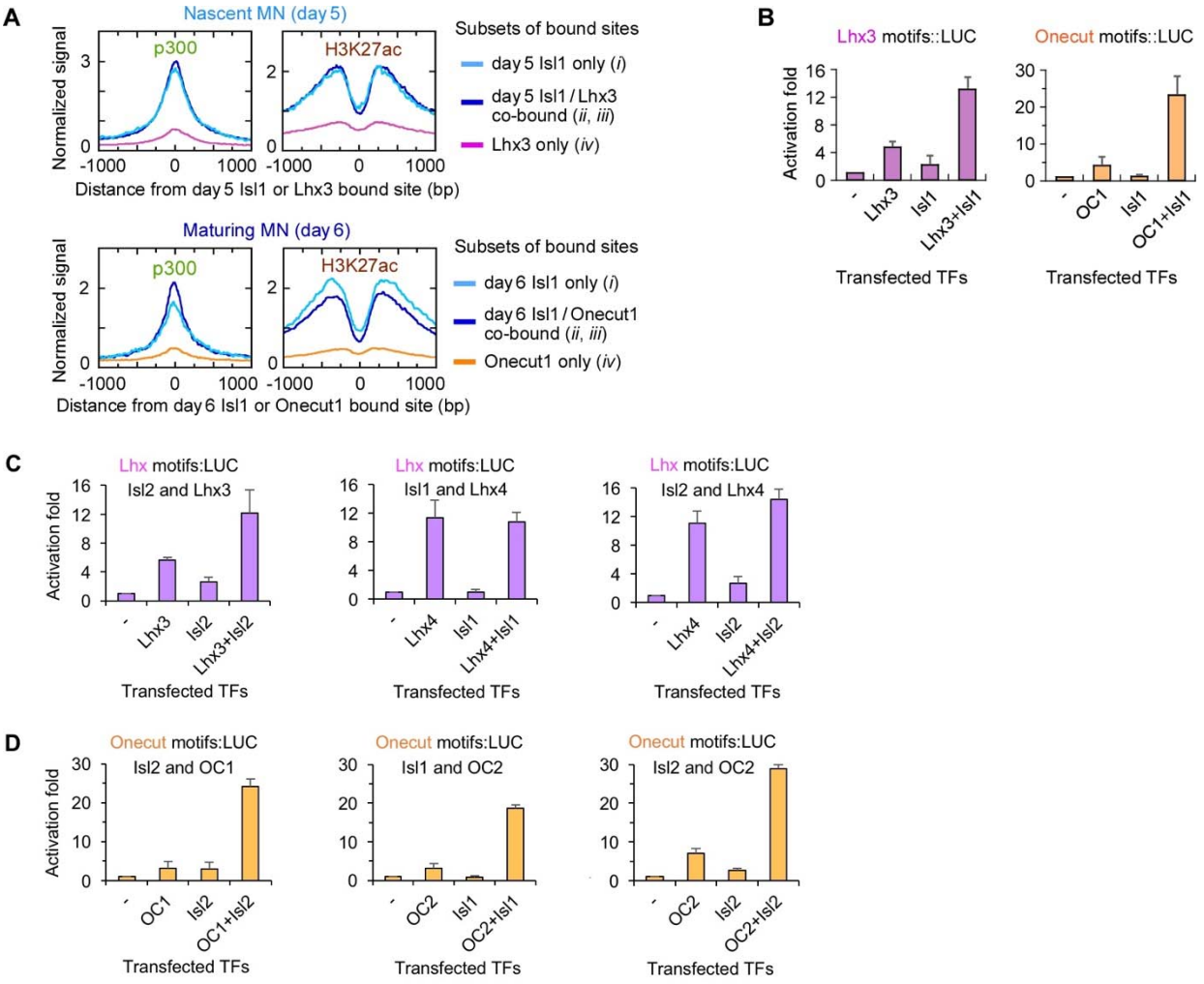

Figure S6. Isl1 and Acetyltransferase p300 Selectively Activate Enhancers Containing Clusters of Onecut1 or Lhx3 TFs, Related to Figure 7

(A) *Upper*, Composite plots of p300 and H3K27ac intensity for Isl1 only (subset *i*), Isl1/Lhx3 (subsets *ii* and *iii*), and Lhx3 only bound sites (subset *iv*) in nascent motor neurons (MNs) shown in Figure S5A. *Lower*, Composite plots for Isl1 only (subset *i*), Isl1/Onecut1 (subsets *ii* and *iii*), and Onecut1 only bound sites (subset *iv*) in maturing MNs shown in **Figure 6B**.

(B) *Left*, luciferase reporter assays with the synthetic enhancer containing 3-Lhx3-binding motifs (indicated in **Figure 7E**) in ES cells, transfected with Lhx3- and Isl1-expressing vectors, alone or together. Shown is fold activation relative to ES cells, transfected with empty vector (-) without containing a gene encoding TF of interest. Error bars represent the s.d. *n*=3. *Right*, Same as left panel, except luciferase reporter assays with the Isl1-E3 enhancer (286 bp; Supplemental Experimental Procedures) containing 3-Onecut-binding motifs in ES cells, transfected with Onecut1- and Isl1-expressing vectors. Error bars represent the s.d., *n*=3.

(C and D) Luciferase reporter assays with enhancers containing 3-Lhx3-binding motifs (panel **C**) and 3-Onecut-binding motifs (panel **D**) in ES cells, transfected with indicated TF-expressing vectors The empty vector without containing a gene encoding TF (-) was used as a normalizing control. Error bars represent the s.d. *n*=3.

Supplemental Tables

Excel files

Table S1 (1.5 MB)

Coordinates, H3K27ac Intensity, and Classification of 30,648 Active Enhancers for **Figure 1C**, Related to **Figure 1**

Table S2 (185 KB)

Coordinates, Isl1 Occupancy, and Classification of 4,764 Isl1-Bound Enhancers for **Figures 2C** and **3B**, Related to **Figures 2** and **3**

Table S3 (12 KB)

Correlation Coefficient (R) for **Figures 2F** and **7C**, Related to **Figures 2** and **7**

Table S4 (1.0 MB)

Gene List, RNA Expression FPKM Values, and Classification of 2,346 Isl1-Associated Genes for **Figure 3B**, and 5,133 Active Enhancer-Associated Genes for **Figure 3G**, Related to **Figure 3**

Table S5 (29 KB)

Gene Ontology Analysis (GO processes, *P*-Values) of Stable and Transient Isl1 Associated Genes for **Figure 3F** and Stable and Transiently Active Enhancer Associated Genes for **Figure 3J**, Related to **Figure 3**

Table S6 (1.1 MB)

Motif Reference Points, Motif Orientation, TF Occupancy, and Classification of 13,579 Isl1- or Onecut1-Bound Sites in Maturing Motor Neurons for **Figure 6B**, and 14,563 Isl1- or Lhx3-Bound Sites in Nascent Motor Neurons for Figure S5A, Related to **Figures 6** and S5

Table S7.

Number of Lhx3-Bound Sites in Clusters Recruiting Isl1 to Nascent Motor Neuron Enhancers (upper) and Number of Onecut1-Bound Sites in Clusters Recruiting Isl1 to Maturing Motor Neuron Enhancers (lower), Related to **Figure 7**.

Lhx3 or Onecut1 clusters were defined as genomic regions having multiple Lhx3- or Onecut1-bound sites within ±100 bp to each other. For example, if there are three Lhx3-bound sites within 100 bp to each other, they were considered as a single Lhx3 cluster. A Lhx3 or Onecut1 cluster recruiting Isl1 was defined as having Isl1-bound sites within ±100 bp from its constituent Lhx3- or Onecut1-bound sites. Frequency histograms of Isl1 recruitment to Lhx3 or Onecut1 clusters were shown in **Figure 7B**.

Number of Lhx3-bound sites in cluster	Total number of Lhx3 clusters	Total number of Lhx3 sites in cluster	Number of Lhx3 clusters recruiting Isl1	% of Lhx3 clusters recruiting Isl1
1	22,231	22,231	1,114	5.0
2	3,911	7,822	600	15.3
3	1,373	4,119	389	28.3
≥4	1,050	5,180	502	47.8

Number of Onecut1-bound sites in cluster	Total number of Onecut1 clusters	Total number of Onecut1 sites in cluster	Number of Onecut1 clusters recruiting Isl1	% of Onecut1 clusters recruiting Isl1
1	62,564	62,564	766	1.2
2	3,344	6,688	467	14.0
3	740	2,220	263	35.5
≥4	370	1,734	205	55.4

Supplemental Experimental Procedures

Experimental Methods

Cell culture. The embryonic stem cell (ESC) line was derived from the inbred mouse strain 129/Ola ES-E14 male embryo. ESCs were cultured on 0.1% gelatin coated plate in EmbryoMax D-MEM (Millipore, SLM-220) supplemented with 10% ESC-grade fetal bovine serum (Thermo, SH3007003E), L-glutamine (Thermo, SC:25030081), 0.1 mM *beta*-mercaptoethanol and leukemia inhibitory factor (LIF, 100 U/mL, Millipore, ESG1107). Motor neuron differentiation of ESCs was performed as described previously (Tan et al., 2016; Wichterle and Peljto, 2008) with some modifications. Briefly, ESCs were trypsinized (Thermo, SC25300054) and seeded at 5 x 10^4 cells per mL in ADFNK medium (Advanced DMEM/F12:Neurobasal (1:1) Medium, 10% Knockout Serum Replacement, 2 mM L-glutamine, and 0.1 mM 2-mercaptoethanol) to initiate formation of embryoid bodies (day 0). Medium was exchanged on days 2 and 5 of differentiation. Patterning of embryoid bodies was induced by supplementing media on day 2 with 1 μM all-trans retinoic acid (Sigma, R2625) and 0.25 μM Smoothened agonist (SMO) of hedgehog signaling (SAG, Millipore, 566660). DAPT (Selleckchem, S2215) was added to the culture medium at 5 μM on day 4 and day 5. The same conditions were used for inducible iNgn2 and iNIL lines (Mazzoni et al., 2013), except 3 μg/mL doxycycline (Clontech, NC0424034) was supplemented to the culture medium on day 3 and day 2, respectively.

ChIP-exo and ChIP-seq. Approximately 30 million cells were crosslinked with 1% formaldehyde, then processed through the ChIP-exo assay as described previously (Rhee and Pugh, 2011, 2012), using Illumina adaptors. Briefly, cells were lysed, and chromatin pellets were isolated and then solubilized and fragmented by sonication. Fragmented chromatin was then subject to immunoprecipitation using magnetic beads coupled with Protein G (Life Technologies, 10004D) and antibodies (~5 μg) against protein of interest (Isl1, RRID:AB_1157901, gifts from T. Jessell; Onecut1, RRID:AB_2251852, Lot# J04505; H3K27ac, RRID:AB_2118291, Lot# GR00563-2; p300, RRID:AB_2293429, Lot# I2712) or V5 epitope-tagged Ngn2 (using iNgn2 cell line) and Lhx3 (using iNIL cell line) (V5, RRID:AB_2556564, Lot# 1652168). After washing the beads to remove un-bound proteins and DNA, ChIP-seq samples were eluted from the magnetic beads. While ChIP-exo samples were still on the beads, the immunoprecipitated DNA fragments were polished, and ligated to an appropriate sequencing library adaptor. Samples were then subjected to *lambda* exonuclease digestion (NEB, Cat# M0262L), which processively removes nucleotides from the 5' end of double-stranded DNA until it encounters a protein-DNA crosslink induced by formaldehyde treatment. The resulting single-stranded DNA was eluted from the magnetic beads and converted to double-stranded DNA by primer annealing and extension. A second sequencing adaptor was ligated to exonuclease treated ends, PCR amplified, gel purified, and sequenced by Illumina HiSeq 2000.

ATAC-seq. ATAC-seq libraries were prepared as described previously (Buenrostro et al., 2013). Briefly, approximately 50,000 cells were lysed and spun down. The nuclei pellet was resuspended in 50 μL of the transposase reaction mix (Illumina DNA Library Preparation Kit, FC-121-1030) containing 2.5 μL Tn5 transposase, 25 μL 2xTD buffer, 22.5 uL nuclease-free water, and then incubated for 30 min at 37°C, followed by a column purification using MinElute PCR Purification Kit (Qiagen, Cat# 28004) The transposase-treated DNA was PCR amplified with sequencing adaptors, purified, and sequenced by Illumina HiSeq 2000.

RNA-seq. RNA was extracted from days 0, 4, 5, and 6 of differentiated cells using TRIzol reagent (Thermo, 15596026). Illumina Ribo-Zero rRNA Removal Kit (MRZH116) was used with DNase-treated total RNA sample to remove ribosomal RNA. RNA-seq libraries were prepared for sequencing using standard Illumina protocols, and sequenced by Illumina HiSeq 2000.

Immunocytochemistry. Embryoid bodies were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS), embedded in OCT (Tissue-Tek) sectioned on cryostat and stained overnight at 4°C with primary antibodies, followed by washes and staining with secondary antibodies for 1 hour at room temperature. After staining, samples were mounted with Aqua Poly Mount. Images were acquired with a LSM 510 Carl Zeiss confocal microscope. We used antibodies to Isl1 (RRID:AB_2126323, gift from T. Jessell, 1:1000); Hb9 (RRID:AB_2145209, gifts from T. Jessell, 1:1000). For quantification, average eight embryoid bodies per timepoint (average 380 cells per embryoid body) were randomly selected and scored for expression of motor neuron markers using MetaMorph Software (Molecular Devices, LLC).

Electroporation. The GFP enhancer reporter plasmid containing a minimal promoter was cloned by replacing *Luciferase* gene in pGL4.23[*luc2*/minP] vector (Promega, Part# 9PIE841A) with destabilized GFP (ZsGreen1) reporter sequence from pZsGreen1-DR Vector (Clontech, Cat# 632428). Enhancer DNA fragments were subcloned

into the reporter plasmid and co-electroporated with a ubiquitously expressed CMV-mCherry reporter as a positive control into the developing neural tube of the *Hamburger Hamilton* (HH) stage 13 chick embryos(Hamburger and Hamilton, 1951; Muhr et al., 1999), using ECM 830 Square Wave Electroporation System (BTX, 45-0002). Chick embryos were incubated in 39°C incubator, and analyzed at stage 18 and 23, 24 and 48 hours after electroporation, respectively. Plasmid-injected chicks were dissected, fixed in 4% paraformaldehyde, embedded in OCT, sectioned and then processed for immunocytochemistry as described above.

The 274 bp *Nrp2* enhancer DNA fragment corresponds to the genomic location (chromosome 1):
TTTCATGATGCCGCGCACACACTGAGTAGCCAAGAAGTGCTTACTTAATTGCCTTTCATCAGTTTGCAC
ACAGTAGCTCACATAGAGGCTGCCATTAAGAAAGACCATAGCCAGCAGAGGATCAACATCCAGTAAT
TTCCAAATGGGTTTGGCATTGTCAGAGGGGAGGTACCTCAACTGTTGGCTGTCCTCCTTGTAGGGTGAA
AGGTCAGGACTTGGTGCATAAGCCTAATTACTTGGGCAGACTCTGCATGTCCTTTGACCCCTTAGTGAC

The 143 bp *Gripap1* enhancer DNA fragment corresponds to the genomic location (chromosome X):
GAGGGGACACAGCACTGCATCATTCCCCGTGGCACTCCGCTCTGGGGCCATTAACAGACGCCTTCCCA
TCGATTGGTTCTGCTAAGTGCTGCTGATTCAGACGGGTGGGGGTGAGGGCCCCTAGGGGCAGTGGGGG
TAGGAAC

Co-immunoprecipitation. Cells were cultured, crosslinked, lysed, and sonicated as described for ChIP experiments (n=2) as described above. Protein complexes were immunoprecipitated using magnetic beads coupled with Protein G (Life Technologies, 10004D) and antibodies (~5 µg) against mouse Isl1 (RRID:AB_2126323, gift from T. Jessell) or goat IgG (RRID:AB_354267, Lot#ES3411101) overnight at 4°C. Immunoprecipitates were washed, and boiled in SDS sample loading buffer (4x NuPAGE LDS Buffer, Thermo, NP0008). After gel running on 4-12% Bis-Tris Protein Gel (NuPAGE Novex, Invitrogen, NP0323BOX), proteins were dry-transferred to PVDF membrane using iBlot Gel Transfer System (Invitrogen, IB21001) according to manufacturer's instruction. Western blots were then incubated using antibodies against rabbit Onecut1 (Santa Cruz, sc-13050), and developed using ECL Western Blotting Substrate (Pierce, Prod#32209) and Amersham Hyperfilm ECL (GE Healthcare, 28-9068-38).

Generation of ESC lines containing enhancer deletion. The RNA-guided Cas9 nuclease from the microbial clustered regularly interspaced short palindromic repeats (CRISPR/Cas9) system was used to delete enhancers in ESCs. The CRISPR/Cas9 genome editing was processed as described in Ran et al.(Ran et al., 2013) with some modifications. Briefly, guide RNAs (gRNAs) were designed with the available software at http://crispr.mit.edu. A pair of oligos encoding the 22-23 bp guide sequences containing 5'-NGG PAM sequence (underlined) to target the left and right sides of intended deletion of enhancers were synthesized from IDT (Integrated DNA Technologies, Inc).

Guide sequences to target the 629 bp region of the Hb9-E1 enhancer (chromosome 5):
Forward 5' ACTTATGCGCCAGACCAGCGAGG 3'
Reverse 5' GTTGGTGAGCGTCCGGCATTGGG 3'

Guide sequences to target the 847 bp region of the Hb9-E5 enhancer (chromosome 5):
Forward 5' TAGATGAGAGGTAGTAACCCAGG 3'
Reverse 5' GGACTTGGCATACAACCTAGTGG 3

Guide sequences to target the 1,413 bp region of the Isl1-E3 enhancer (chromosome 13):
Forward 5' AGTATGACATACGATGCTTGCGG 3'
Reverse 5' GTGGTTATTCATTCCTGATCAGG 3'

Guide sequences to target the 1,519 bp region of the Isl1-E4 enhancer (chromosome 13),
Forward 5' TGTGGTGCCATGGGTACACAGG 3'
Reverse 5' ACTAAATAGGACCGTGCTTTGG 3'

To synthesize gRNAs, we cloned 22-23 bp genomic sites of the form 5'-$N_{19\text{-}20}$NGG-3' (N=A, T, G, or C) near the intended target enhancer into U6 target gRNA expression vector containing U6 promoter (pgRNA, addgene, Cat# 41824) by following the available cloning protocol at https://www.addgene.org/static/cms/files/hCRISPR_gRNA_Synthesis.pdf. To generate ESC lines containing enhancer deletion, ESCs were seeded and incubated for 48 hours on mouse feeder cells (CF-1 MEFs, MTI-GlobalStem, Cat# GSC-6301G) to maintain stem cell pluripotency. Cells were transiently transfected with a pair of pgRNA vectors and Cas-mCherry plasmid containing a GAGGS promoter with Cas9 DNA endonuclease and mCherry reporter using Mouse Neural Stem Cell Nucleofector Kit (Lonza, Cat# VPG-1004) and Primary Cell Nucleofector II/2B (Lonza), according to manufacturer's instruction. ESCs were dissociated 24-48 hours after transfection, mCherry-expressing ESCs were FACS sorted. Approximately 2,000 of sorted cells were plated on a dish with mouse feeder cells. ESCs formed clones in 5-7 days after FACS sorting and plating. Individual colony was picked, followed by genotyping using PCR primers around genomic regions of intended enhancer deletion. The ESC clones containing enhancer deletion were sub-cultured to mouse feeder cells and then differentiated to postmitotic hypaxial motor neurons. To generate an ESC line containing deletion of two enhancers, we deleted the second enhancer after generating an ESC line containing the first enhancer deletion as described above.

Luciferase reporter assay. Enhancer DNA fragments were cloned into pGL4.23[*luc2*/minP] Vector, containing a minimal promoter with *luc2* (*Photinus pyralis*, known as firefly) luciferase reporter gene (Promega, Part# 9PIE841). This cloned firefly luciferase plasmid was co-transfected with pRL-TK Vector, containing the the herpes simplex virus thymidine kinase (HSV-TK) promoter with renilla (*Renilla reniformis*) luciferase reporter gene (Promega, Cat# E2241) for normalization of transfection efficiency. For transcription factor (TF)-expressing vectors, the phosphoglycerate kinase (PGK) promoter and the cDNA of interest (e.g. mouse Isl1, mouse Lhx3, mouse Onecut1) were cloned into pminiTol2 (pDB739, addgene, Cat# 31829). Empty vectors (pDB739-PGK) were used to equalize the total amount of transfected DNA for TF-expressing vectors. Mouse ESCs were cultured as described above. For luciferase assays, ESCs were seeded and incubated for 24 hours, and transiently transfected using Lipofectamine 3000 (Thermo Fisher Scientific, Cat# L3000008), according to manufacturer's instruction. Cells were harvested 24 hrs after transfection. Cell extracts were assayed for luciferase activity, according to the instruction of Dual-Glo Luciferase Assay System (Promega, Cat# E2940), using 20/20 Luminometer (Tuner Biosystems #2030-102). The firefly luciferase intensity was normalized with the renilla luciferase intensity. The Isl1-E3 enhancer DNA sequences (286 bp) containing 3-Oneuct-binding motifs were like the following (chromosome 13, Onecut binding motifs are underlined):

ATATCATCAGCCTTAAAAACTCTCACTAAGTTATTTTAATCAATATACTTGTTTTAAGTAAAAGCCATT
TACTCTCCTGACTTGTACATTAACCTTCCATTAAGTATTGGTGAAGAAGTTAAATGTACCTTAGTCAAT
AATGCTAATGCAGGAGCTGTTGTATTACCTATGGAGAAATGCGTATTCATATTATGTTAAAGTGGCTGT
TCTTAACGATTAGCCATTATGTGTCAGGAAACTGATAATAAATTTGAAATATTCCACGAGACAATGAA
TTATGATGTTG

Data Analyses

Deep sequencing and alignment to the genome. ChIP-seq and ChIP-exo DNA samples were analyzed by single-end sequencing, and ATAC-seq DNA samples were analyzed using a paired-end sequencing (HiSeq 2000 Sequencing Systems, Illumina). All sequencing data sets were aligned using Bowtie (version 0.12.7) (Langmead et al., 2009) to build version NCBI37/mm9 assembly of the mouse genome. All data were deposited to GEO database.

Identifying and displaying ChIP-Seq, ChIP-exo, ATAC-Seq enriched regions. We analyzed ChIP-seq and ChIP-exo data sets using the GEM peak caller (Guo et al., 2012) and the strand separate peak finding algorithms in the GeneTrack (Albert et al., 2008; Rhee and Pugh, 2011) to identify regions enriched by transcription factors (TFs) over background, independent of DNA motif information. For peak calling of ChIP-seq H3K27ac and ATAC-seq data sets, we used the MACS software ("---broad"; version2.0.9) (Zhang et al., 2008), the HOMER software ("-size 100 -fdr 0.01 -style factor"; v4.7.2) (Heinz et al., 2010), or the GeneTrack software ("-s 50 -e 1000"). For ES cells, we used the published ChIP-seq H3K27ac data set (Gene Expression Omnibus, GSE27844) (Whyte et al., 2012).

For ChIP-seq sequencing reads, unfiltered tags on each strand were shifted by half of the estimated fragment length in the 3' direction (60-90 bp depending on samples), to adjust tags to the center of sonicated DNA fragments. All figures displaying ChIP-exo sequencing reads except **Figures 6** and S5 represent unfiltered tags after shifting them on each strand in the 3' direction by a fixed distance (half length of the protected region from exonuclease, ~12 bp depending on TFs), so as to reflect better the points of protein-DNA crosslinking (Rhee and Pugh, 2011). ChIP-exo sequencing reads in **Figures 6** and S5 were displayed without shifting them on each strand in the 3' direction. ATAC-seq sequencing reads were not shifted. Tags from biological replicates were merged after demonstrating their reproducibility.

Definition of TF-bound sites, TF-bound enhancers, and active enhancers. TF-bound sites were defined as 22-28 bp genomic regions flanked by ChIP-exo enriched peaks demarcating a site protected from *lambda* exonuclease digestion. A TF-bound enhancer was defined as a genomic region centered on the midpoint of the identified TF-bound site and extending ±1 kb. If multiple bound sites for the same TF at a given cell stage resided within 1 kb of each other, they were considered a single enhancer and were centered on the site exhibiting the highest TF occupancy. In all of our analysis we used a ±1 kb region as an enhancer boundary based on our observation that H3K27Ac histone modifications and accessible chromatin regions typically spread over regions extending 0.5-1 kb from the TF-bound sites.

An H3K27ac-enriched enhancer was defined as a genomic region >2.5 kb from the nearest transcription start site (TSS) of known mRNA genes (19,326 of annotated mm9 NCBI Build 37 RefSeq mRNA genes having completed coding start and end sequences) and extending ±1 kb centered on the midpoint of the identified H3K27-enriched region. If multiple H3K27ac-enriched regions at a given cell stage resided within 1 kb to each other, they were considered as a single enhancer and centered on an enhancer having higher H3K27ac occupancy. Accessible enhancers enriched by ATAC-seq were defined as the same criteria as H3K27ac-enriched regions (Figure S2A). For

active H3K27ac-enriched/accessible enhancers, we identified all genomic regions where midpoints of H3K27ac-enriched and chromatin accessible (ATAC-seq) enhancers at a given cell stage (days 4, 5, or 6) fell within 1 kb of each other. We used an ATAC-seq peak midpoint as a reference point of these active enhancers. For ES cells, DNase-seq data set was used instead of ATAC-seq.

Maintained and dynamic stage-specific enhancers were defined based on the presence or absence of active enhancers within 1 kb of each other at the compared developmental stages. For example, if an Isl1-bound enhancer in nascent motor neurons resides within 1 kb from an Isl1-bound enhancer in maturing motor neurons, they were classified as a single maintained (stable) enhancer in nascent and maturing motor neurons. If they were further than 1 kb to each other, Isl1-bound enhancers were annotated as stage-specific Isl1-bound enhancers.

Assignment of enhancers to associated genes. To link enhancers with their associated genes (**Figures 3** and S3), each enhancer midpoint was assigned to the nearest TSS of 19,326 annotated mm9 NCBI Build 37 RefSeq genes. To avoid assigning to uncharacterized or predicted genes, we used genes having completed coding start and end sequences for gene assignment. Genes associated with intragenic enhancers had precedence over genes associated with intergenic enhancers.

RNA-Seq analysis. RNA-seq was performed using a paired-end sequencing (HiSeq 2000 Sequencing Systems, Illumina), and aligned to the mouse genome using RNA-seq alignment algorithm in the STAR (Spliced Transcripts Alignment to a Reference) software (version2.3.1) (Dobin et al., 2013). To assemble transcripts and estimate the relative abundances of these transcripts, we used the Cufflinks software (release 2.2.1) (Trapnell et al., 2013). FPKM (Fragments Per Kilobase of exon per Million fragments mapped) was calculated per individual gene. For $\log_2$ RNA fold-changes (**Figures 1** and **3**) if an FPKM value less than $1.5x10^{-3}$, it was recoded to $1.5x10^{-3}$ to avoid producing denominators with a zero value and FPKM values were then $\log_2$ transformed. To determine genes that significantly change their expression between day 4 and day 5 of differentiated cells, we classified all genes detected from RNA-seq profiling (n=18,614) into three groups: genes with more than 2-fold expression change between day 4 and day 5 (2,349 downregulated genes and 1,907 upregulated genes), maintained genes (<2-fold expression change between day 4 and day 5, *n*=6,793), and non-expressed genes (*n*=7,565).

Classification of associated genes. Developmentally regulated and housekeeping genes included in **Figure 1D** were defined based on the following criteria. Developmentally downregulated genes (groups 1, 2 and 3): FPKM values are more than 6.0 on day 0 (third quartile FPKM value on day 0; groups 1, 2 and 3), and the expression decreases from day 0 to day 4 by >1.25-fold (groups 1 and 2), decreases from day 4 to day 5 by >1.25-fold (groups 2 and 3), or decreases from day 5 to day 6 by >1.25-fold (group 3). Developmentally regulated genes in groups 4-8 were defined by the following criteria: FPKM values increase from day 0 to day 4 by more than 1.25-fold (groups 4 and 5), and decrease from day 4 to day 5 by >1.25-fold (group 4). We defined groups 6, 7 and 8 of the regulated genes meeting the following criteria: FPKM values are less than 6.0 on day 0 (groups 6, 7 and 8), increase from day 0 to day 4 more than 1.25-fold (groups 6 and 7), increase from day 4 to day 5 more than 1.25-fold (groups 6, 7 and 8), and increase from day 5 to day 6 more than 1.25-fold (groups 7 and 8).

Non-expressed were defined as all genes having FPKM values less than a median in all four timepoints (0.7 on day 0; 0.8 on day 4; 0.9 on day 5; 0.9 on day 6). All other genes that do not belong to the regulated or non-expressed category were defined as constitutively expressed housekeeping genes. Most of housekeeping genes (3,127 out of 5,215) had FPKM value within the top third quartile (6.0 on day 0; 7.5 on day 4; 7.1 on day 5; 6.8 on day 6) in at least three timepoints.

The calculation of percentages of differentially expressed genes were based on the population of developmentally regulated and housekeeping genes (excluding non-expressed genes) to make this comparison more relevant to the analysis of differentially active enhancers.

Motif enrichment analysis. We used the Multiple Em for Motif Elicitation (MEME, v4.11) algorithm (Bailey et al., 2015) for de novo motif discovery. We searched the enriched motifs occurring proximal to the TF-bound sites (20-40 bp windows centered on peak midpoints) identified from ChIP-exo for randomly selected 300 bound sites of each data set. We used then the Find Individual Motif Occurrences (FIMO) (Grant et al., 2011) to scan the enriched motif from MEME for all bound sites of each data set. The newly discovered motifs were compared against a database of known motifs (i.e. UniPROBE and JASPAR) using the Motif Comparison Tool (Tomtom) (Gupta et al., 2007).

Correlation analysis. For correlation plots depicted in **Figure 2F**, Isl1-bound enhancers were filtered to include 2,446 enhancers that contain only day 5-specific Isl1 binding and 1,704 enhancers that contain only day 6-specific Isl1 binding shown in **Figures 2C** and **2D** (stably occupied enhancers were excluded from the analysis). Occupancy levels were normalized to the median occupancy for each group of enhancers and values were $\log_2$ transformed. The Pearson product-moment correlation values (R) were calculated using the ‘‘correl’’ function in Excel (Table S3). Heat map plots were generated using Java Treeview. Correlation values (R) in **Figure 7C** were calculated as same

as above except using Isl1-bound sites shown in **Figures 6B** and S5A, which include Isl1-containing sites (subsets *i*, *ii*, and *iii*) versus Onecut1 or Lhx3 only bound sites (subset *iv*).

Gene ontology analysis. To find the Gene Ontology (GO) terms enriched in enhancer-associated genes, we used the Gene Ontology Enrichment Analysis and Visualization (GOrilla) system(Eden et al., 2007; Eden et al., 2009). To remove redundant GO terms (Figure S3), we used the Reduce and Visualize Gene Ontology (REViGO) software with 0.5 of the *SimRel* similarity score based on Resnik's and Lin's similarity measures (Supek et al., 2011). For Table S5, the Protein Annotation through Evolutionary Relationship (PANTHER) classification system was used with Bonferroni correction (Mi et al., 2013).

Significance test. Where comparison of occupancy/intensity levels (read counts) or expression levels (FPKM) between bound regions, genes, timepoints, or data sets is reported in box plots, the *P*-value for Wilcoxon rank-sum test (Mann-Whitney *U* test) was computed using the Excel functions, "rank" and "normdist", and the *P*-value for two sample *t*-test was computed using the Excel function "ttest".

Supplementary References

Albert, I., Wachi, S., Jiang, C., and Pugh, B.F. (2008). GeneTrack--a genomic data processing and visualization framework. Bioinformatics *24*, 1305-1306.

Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME Suite. Nucleic Acids Res *43*, W39-49.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods *10*, 1213-1218.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Eden, E., Lipson, D., Yogev, S., and Yakhini, Z. (2007). Discovering motifs in ranked lists of DNA sequences. PLoS Comput Biol *3*, e39.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017-1018.

Guo, Y., Mahony, S., and Gifford, D.K. (2012). High resolution genome wide binding event finding and motif discovery reveals transcription factor spatial binding constraints. PLoS Comput Biol *8*, e1002638.

Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W.S. (2007). Quantifying similarity between motifs. Genome Biol *8*, R24.

Hamburger, V., and Hamilton, H.L. (1951). A series of normal stages in the development of the chick embryo. J Morphol *88*, 49-92.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Mazzoni, E.O., Mahony, S., Closser, M., Morrison, C.A., Nedelec, S., Williams, D.J., An, D., Gifford, D.K., and Wichterle, H. (2013). Synergistic binding of transcription factors to cell-specific enhancers programs motor neuron identity. Nat Neurosci *16*, 1219-1227.

Mi, H., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2013). Large-scale gene function analysis with the PANTHER classification system. Nat Protoc *8*, 1551-1566.

Muhr, J., Graziano, E., Wilson, S., Jessell, T.M., and Edlund, T. (1999). Convergent inductive signals specify midbrain, hindbrain, and spinal cord identity in gastrula stage chick embryos. Neuron *23*, 689-702.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281-2308.

Rhee, H.S., and Pugh, B.F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell *147*, 1408-1419.

Rhee, H.S., and Pugh, B.F. (2012). ChIP-exo method for identifying genomic location of DNA-binding proteins with near-single-nucleotide accuracy. Curr Protoc Mol Biol *Chapter 21*, Unit 21 24.

Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One *6*, e21800.

Tan, G.C., Mazzoni, E.O., and Wichterle, H. (2016). Iterative Role of Notch Signaling in Spinal Motor Neuron Diversification. Cell Rep *16*, 907-916.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol *31*, 46-53.

Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature *482*, 221-225.

Wichterle, H., and Peljto, M. (2008). Differentiation of mouse embryonic stem cells to spinal motor neurons. Curr Protoc Stem Cell Biol *Chapter 1*, Unit 1H 1 1-1H 1 9.

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