# Supporting Information

## Lee et al. 10.1073/pnas.1302676110

### SI Materials and Methods

**DNA Constructs.** Rat Islet-1 (*Isl1*), human *STAT1*, mouse LIM homeobox 3 (*Lhx3*), nuclear LIM interactor (*NLI*), Neurogenin2 (*Ngn2*), *STAT3*, and *LacZ* genes were cloned in pCS2 and/or pcDNA3 (Invitrogen) containing a HA, FLAG, or myc-epitope tag for expression in mammalian cells and chicken embryos. These vectors were previously described (1–3). Constitutively active STAT3 (STAT3-CA) and dominant negative forms of STAT3 (STAT3-DNs) were cloned in pCAGGS vector. STAT3-CA is a fusion of STAT3 mutant, in which two cysteine residues within the C-terminal loop of the SH2 domain are substituted for A661 and N663, and a VP16 transcriptional activation domain. Two cysteine residues in STAT3-CA allow for sulfhydryl bonds to form between STAT3 monomers, enabling dimerization of STAT3 without a phosphorylation on Y705 residue (4). STAT3-F and STAT3-D were used as STAT3-DNs (5). STAT3-F, in which Tyr705 of STAT3 is mutated to Phe, does not respond to extracellular signaling that requires STAT3 phosphorylation on Tyr705. STAT3-D, in which Glu434 and Glu435 within the DNA-binding domain are mutated to Ala, is unable to bind to STAT3-binding motif. Isl1-Lhx3 is a fusion protein of Isl1 full length and Lhx3 full length linked by a flexible linker GGSG-GSGGSGG (6). Hb9-MNe (an MN-specific enhancer in an MN-specific gene *Hb9*):LUC was described previously (3).

To generate shRNA knockdown constructs for chick STAT1 and STAT3 and mouse STAT3, double-stranded oligonucleotides were formed by annealing top and bottom strands for shRNA targets and cloned to EFU6-300 vector. The following oligonucleotides were used for the cloning. Short hairpin (sh)-chick STAT1, top: 5′-GAT CCG GCT GTG GAT AGA GGG AAT TTC AAG AGA ATT CCC TCT ATC CAC AGC CTT TTT A, bottom: 5′-AGC TTA AAA AGG CTG TGG ATA GAG GGA ATT CTC TTG AAA TTC CCT CTA TCC ACA GCC G; sh-chick STAT3, top: 5′-GAT CCG GAA GAG ACG GCA ACA AAT TTC AAG AGA ATT TGT TGC CGT CTC TTC CTT TTT A, bottom: 5′-AGC TTA AAA AGG AAG AGA CGG CAA CAA ATT CTC TTG AAA TTT GTT GCC GTC TCT TCC G; sh-mouse STAT3 #1, top: 5′-GAT CCG GCA TCA ATC CTG TGG TAT TTC AAG AGA ATA CCA CAG GAT TGA TGC CTT TTT A, bottom: 5′-AGC TTA AAA AGG CAT CAA TCC TGT GGT ATT CTC TTG AAA TAC CAC AGG ATT GAT GCC G; and sh-mouse STAT3 #2, top: 5′-GAT CCG GCA TAT GCA GCC AGC AAA TTC AAG AGA TTT GCT GGC TGC ATA TGC CTT TTT A, bottom: 5′-AGC TTA AAA AGG CAT ATG CAG CCA GCA AAT CTC TTG AAT TTG CTG CTG CAT ATG CCG.

**Generation and Differentiation of iMN-ESCs.** Generation of inducible motor neuron ES cells (iMN-ESCs) was described previously (6). The A172LoxP ESC line (7) was maintained in an undifferentiated state on 0.1% gelatin-coated dishes in the ESC growth media that consisted of Knockout DMEM, 10% (vol/vol) FBS, 0.1 mM nonessential amino acids, 2 mM L-glutamine, 0.1 mM β-mercaptoethanol, and recombinant leukemia inhibitory factor (LIF, 1000 units/mL, Chemicon). A gene encoding FLAG-tagged Isl1-Lhx3 fusion construct was inserted into Tet-inducible plasmid p2Lox and were cotransfected with pSALK-Cre into ESC lines by electroporation. Stable transfectants were isolated by selection with neomycin (G418, 400 µg/mL) for 7 d. Doxycycline (Dox)-dependent induction of FLAG-Isl1-Lhx3 was confirmed by RT-PCR and Western blotting and immunohistochemical analyses using α-Isl1, α-Lhx3, and α-FLAG antibodies.

The differentiation of iMN-ESCs into MNs was done as previously described (6). Briefly, embryoid bodies of iMN-ESCs were treated with 0.5 µM RA alone for 2 d and then were cultured without or with Dox (2 µg/mL) in the presence of RA for 4 d. The cells were treated with LIF (1,000 µ/mL) for 16 h before harvest.

**ChIP and High-Throughput Sequencing (Seq).** iMN-ESCs were cultured on 0.1% gelatin-coated dishes in the ESC growth media lacking LIF in the presence or absence of Dox (2 µg/mL), which induces the expression of FLAG-tagged Isl1-Lhx3, for 1 d. Approximately $4 \times 10^6$ cells were washed with PBS buffer, fixed by 1% formaldehyde for 10 min at room temperature, and quenched by 125 mM glycine. Cells were washed with Buffer I (0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA, 10 mM Hepes, pH 6.5) and Buffer II (200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10 mM Hepes, pH 6.5) sequentially. Then, cells were lysed with lysis buffer (0.5% SDS, 5 mM EDTA, 50 mM Tris·HCl, pH 8.0, protease inhibitor mixture) and were subjected to sonication for DNA shearing. Next, cell lysates were diluted 1:10 in ChIP buffer (0.5% Triton X-100, 2 mM EDTA, 100 mM NaCl, 50 mM Tris·HCl, pH 8.0, protease inhibitor mixture) and, for immunoclearing, were incubated with IgG and protein A agarose beads for 1 h at 4 °C. Supernatant was collected after quick spin and incubated with α-FLAG antibody (Sigma) and protein A agarose beads to precipitate FLAG-Isl1-Lhx3/chromatin complex overnight at 4 °C. After pull-down of FLAG-Isl1-Lhx3/chromatin/antibody complex with protein A agarose beads, the beads were washed with TSE I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris·HCl, pH 8.0, 150 mM NaCl), TSE II (same components as in TSE I except 500 mM NaCl), and Buffer III (0.25 M LiCl, 1% Nonidet P-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris·HCl, pH 8.0) sequentially for 10 min at each step. Then the beads were washed with Tris-EDTA (TE) buffer three times. FLAG-Isl1-Lhx3/chromatin complexes were eluted in elution buffer (1% SDS, 1 mM EDTA, 0.1 M NaHCO$_3$, 50 mM Tris·HCl, pH 8.0) and decross-linked by incubating at 65 °C overnight. Eluate was incubated at 50 °C for more than 2 h with Proteinase K. Next, DNA was purified with Phenol/chloroform and DNA pellet was precipitated by ethanol and resolved in water. ChIP DNA samples were prepared for sequencing according to the Illumina protocol, and sequenced with the Illumina Genome Analyzer IIx.

**ChIP-Seq Data Analysis.** The sequence reads of 50 or 36 bp generated from Illumina GApipeline were mapped to mouse reference genome (NCBI37, mm9) using BWA software (v0.5.8a) with default parameters (8). To exclude the multiple-hit reads and low-mapping-quality reads, only the reads with mapping quality >25 were kept for further analysis. If several reads had the exact same sequence and direction, only one copy was kept for the analysis to minimize the artifacts introduced by PCR amplification during sample preparation (9). The samples treated with Dox were regarded as Isl1-Lhx3-ChIP, whereas the samples without Dox were regarded as control. To achieve more robust analysis, the sequencing data from two biological replicates were pooled for each dataset. The peak calling was conducted with QuEST v2.4 software (10) using the "transcription factor" option (bandwidth 30 bp, region size 300 bp) and the recommended peak-calling parameters (30-fold Isl1-Lhx3-ChIP to control enrichment for seeding the regions, threefold Isl1-Lhx3-ChIP enrichment for extending the regions, and threefold Isl1-Lhx3-ChIP to control fold enrichment).

For the analysis of peak location, we used mm9 RefSeq gene annotation. We assigned the peaks to the exons that exclude untranslated regions, 5′ UTRs, 3′ UTRs, and introns, up to 20 kb upstream from the 5′ end of a transcription start site (TSS), up to 20 kb downstream from the 3′ transcription end site (TES), and intergenic regions (with a distance greater than 20 kb upstream from any TSS or 20 kb downstream from any TES).

To associate each peak with genes, the potential target RefSeq genes were identified with the following criterion: at least one ChIP-seq peak located from 20 kb upstream or downstream of a gene. If no gene is found within the selected range from a peak, the nearest genes to the peak were chosen as potential targets.

Analysis of gene ontology (GO) terms was performed using DAVID (11). Functional annotation charts were generated for the GO term GOTERM_BP_ALL using a mouse background.

For the analyses to combine ChIP-seq and RNA-seq. (6) data, the predicted target genes from ChIP-seq are associated with the differentially expressed genes from RNA-seq only if the ChIP-seq peak is found within 20 kb from the gene body and the fold change is ≥1.5 in normalized RNA-seq expression levels.

**ChIP-Seq Motif Analysis.** MEME-ChIP Suite (12) was used for motif analysis with some modifications based on its standard protocol. The sequences of each peak were defined by extending 250 bp from the peak summit to both sides, and the 500-bp sequences were extracted from the UCSC genome browser mm9 assembly. By default, the web-based Multiple Expectation Maximization for Motif Elicitation (MEME)-ChIP Suite limits MEME (13), a probabilistic model-based algorithm, to the data size of 600 sequences because of the demanding running time. If more than 600 sequences are provided, the program randomly samples 600 sequences without duplicates and trims the center 100 bp for de novo MEME motif search. However, because we wanted to search significant motifs from the total of 2,455 peaks, we removed this limit by running individual commands in our own servers for de novo MEME search. MEME searched both the provided sequences and their reverse complements for any number of nonoverlapping occurrences of each motif.

Although MEME is powerful and does not require any prior knowledge, it still has limitations to detect very short or low complexity motifs of interest. Thus, MEME-ChIP Suite provides a complementary algorithm, named Discriminative Regular Expression Motif Elicitation (DREME), which is a simpler and nonprobabilistic model using regular expressions to search for enriched motifs (12). DREME is capable of detecting very short motifs. The motifs identified by MEME and DREME were compared with a database of known transcription factors (JASPAR_CORE_2009.meme) using a motif comparison tool TOMTOM algorithm (14) that is also included in the MEME-ChIP Suite. In addition, Motif Alignment and Search Tool (MAST) (15) and Average Motif Affinity (AMA) (16) were used for motif visualization and binding strength analysis. Finally, to detect any additional known transcription factor binding sites with higher sensitivity, Analysis of Motif Enrichment (AME) algorithm (17) was used for motif enrichment analysis.

To visualize the frequency of each motif in relation to the summit of the Isl1-Lhx3–bound peaks, a graph of positional distribution was plotted. The number of sites for each motif was binned in size of 10 or 20 according to its relative distance to the summit of the peaks, which was set as 0. The curves were smoothened by the smooth.spline function in splines, an R package, with a smoothing parameter of 0.45. We noticed that some hexamer response element (HxRE)-long and HxRE-short motifs shared the same sequences. To determine whether a motif contains HxRE-long motif alone, HxRE-short motif alone, or both HxRE motifs, the motifs are considered to detect the same sequences if the distance between the centers of HxRE-long and HxRE-short motifs is <5 bases. This strategy was used to ensure that the HxRE-long

and HxRE-short overlapping sequences were counted only once in retrieving the total number of HxRE sites. If one HxRE-long motif contains two HxRE-short motifs because of sequence similarity, this region was considered to have two HxRE-short motifs in calculating the total number of HxREs. To compare the positional distribution patterns of different motifs in a pairwise manner, the peaks containing both motifs were identified and their relative positions were plotted.

**Conservation Analysis of the ChIP-Seq Peaks.** To examine whether the sequences under identified peaks are evolutionarily conserved, we analyzed the conservation scores (18) under each 500-bp peak with the randomly picked 500-bp region. For each particular chromosome, we randomly generated the same number of regions as that of peaks. The sum of the conservation scores under the peaks or randomly picked regions was compared with each other using the Wilcoxon test.

**Software Availability.** Analysis tools other than the published packages were implemented in Perl, R, and Linux shell scripts. Source codes and scripts are available upon request.

**ChIP Assays.** To perform the ChIP assays with mouse embryonic tissues, we dissected E12.5 spinal cords. The microdissected spinal cords from five to 12 E12.5 embryos were combined for each ChIP reaction with a specific antibody. The tissues were dissociated completely before the ChIP process. For the ChIP assays, P19 cells were transfected using Lipofectamine 2000 (Invitrogen), and iMN-ESCs were cultured with or without Dox for 2 d to induce Isl1-Lhx3 expression. For STAT3 knockdown studies, P19 cells were transfected with both shRNA-STAT3 #1 and shRNA-STAT3 #2 constructs. The cells were fixed by 1% formaldehyde for 10 min at room temperature, and quenched by 125 mM glycine. Next, cells were washed with Buffer I (0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA, 10 mM Hepes, pH 6.5) and Buffer II (200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10 mM Hepes, pH 6.5) sequentially. Then, cells were lysed with lysis buffer (0.5% SDS, 5 mM EDTA, 50 mM Tris·HCl, pH 8.0, protease inhibitor mixture) and were subjected to sonication for DNA shearing. Next, cell lysates were diluted 1:10 in ChIP buffer (0.5% Triton X-100, 2 mM EDTA, 100 mM NaCl, 50 mM Tris·HCl, pH 8.0, protease inhibitor mixture) and, for immunoclearing, were incubated with IgG and protein A agarose beads for 1 h at 4 °C. Supernatant was collected after quick spin and incubated with α-FLAG antibody (Sigma) and protein A agarose beads to precipitate FLAG-Isl1-Lhx3/chromatin complex overnight at 4 °C. After pull-down of FLAG-Isl1-Lhx3/chromatin/ antibody complex with protein A agarose beads, the beads were washed with TSE I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris·HCl, pH 8.0, 150 mM NaCl), TSE II (same components as in TSE I except 500 mM NaCl) and Buffer III (0.25 M LiCl, 1% Nonidet P-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris·HCl, pH 8.0) sequentially for 10 min at each step. Then the beads were washed with TE buffer three times. FLAG-Isl1-Lhx3/ chromatin complexes were eluted in elution buffer (1% SDS, 1 mM EDTA, 0.1 M NaHCO₃, 50 mM Tris·HCl, pH 8.0) and decross-linked by incubating at 65 °C overnight. Eluate was incubated at 50 °C for more than 2 h with Proteinase K. Next, DNA was purified with Phenol/chloroform and DNA pellet was precipitated by ethanol and resolved in water. The purified final DNA samples were subjected to quantitative PCR reactions using the SYBR green kit (11762-500, Invitrogen) and Mx3000P (Stratagene). The total input was used for normalization. All ChIP experiments were repeated independently at least three times. Data are represented as the mean of duplicate or triplicate values obtained from representative experiments; error bars represent SD. The following primers were used for ChIP-PCR; *Hb9*-MNe (P-85-1), forward: 5′-TAC TCT CCC TAC AGT CTC

TGG GGT, reverse: 5′-TGT CCA GAA ATC CAC AGG CCT GCG; Fgf9 (P-35-1), forward: 5′-CAA TCC TTC CAG ATG GTC CAA AAT, reverse: 5′-AAG TTC CAG GGG CAG GCT TTT CCA; 9530008L14Rik (P-87-1), forward: 5′-CTC CCT CAG GGG CTG GTG GGA ACT, reverse 5′-ATT CCT GTG CCC TAT CTT GAT CCC; Arhgap36, forward: 5′-ACT GCC TAT TCG CAT CGG CCT TTG A, reverse 5′-TTC TGC GGA GCC ATT AGT GCG ATT G; Chodl (P-538-1), forward: 5′-TTA ATC ACC AAG TGC ATT AAG CCT G, reverse 5′-TGC CTT CCT TCT TGC AAT CAG CAT; Hecw2 (P-504-2), forward: 5′-ACT CAT TTC CCG ACC TGG AGA AAT, reverse 5′-ATA ATT AAT GTG CTT TCA GTG TGA; Fgf9 (P-674-1), forward: 5′-AAA GCC ATG TTT ACT AAG GCA TTC, reverse 5′-AAT GGC AGC TCC TAA GCG CAT CCT; Hs3st5, forward: 5′-GAG CTA AGT CAA GAT TCC TGA TCA, reverse 5′-GGG CTG GCT GCT TAA GCA CCT GCT; Megf11 (P-1522-1), forward: 5′-GAT CTG CTA CAG GAA ACT GTT CAG, reverse 5′-AAG TAA TTA TAC AGC ATG CTT TGT; Ppargc1a (P651-1), forward: 5′-AAG GCT CAA TAA TTG TTAA TCT CA, reverse 5′-TTC ATC GTG GCA GGG AAG TCA GGG; and Fam19a4, forward: 5′-TTT CCA GTT TGA GAG AAA GTA AAG, reverse 5′-CCA TTT GTC AAG TTA CCT CAT TAG.

The following antibodies were used for ChIP-PCR; mouse/rabbit IgG (Santa Cruz), rabbit anti-Isl1 (19), rabbit anti-Lhx3 (20), rabbit anti-NLI (21), mouse anti-FLAG (Sigma), rabbit anti-STAT3 (#9132, Cell Signaling), and rabbit anti–Phosph-STAT3 (Tyr705) (#9145, Cell Signaling).

**In Ovo Electroporation and Immunohistochemistry.** These assays were performed as described previously (1, 22). In chick electroporation assays, DNAs were injected into the equivalent of a Hamburger–Hamilton stage 13 chick neural tube. The embryos were harvested 3 d postelectroporation and fixed in 4% paraformaldehyde, embedded in Optimal Cutting Temperature compound (OCT), and cryosectioned in 12 μm thickness for immunohistochemistry assays or 18 μm thickness for in situ hybridization with digoxigenin-labeled probes. Each set of chick electroporation experiments was repeated independently three to six times with at least three embryos injected with the same combination of plasmids for each experimental set. Representative sets of images from reproducible results were presented. Mouse embryos were collected at indicated developmental stages and processed similarly to chicken embryos as described previously.

For immunohistochemistry assays, the following antibodies were used; rabbit anti-STAT3 (#9132, Cell Signaling), rabbit anti-Hb9 (23), mouse anti-Mnr2/Hb9 (5C10, DSHB), rabbit anti-Isl1/2 (19), rabbit anti–Tyr705-phospho-STAT3 (#9145, Cell Signaling), rabbit anti-STAT1 (sc-591, Santa Cruz), guinea pig anti-Chx10 (23), rabbit anti-Lhx3 (20), rabbit anti-Lim1 (19), rabbit anti-Olig2, rabbit phsopho-histone3 (sc-8656-R, Santa Cruz), and rabbit anti-GFP (A6455, Molecular Probes) antibodies. Rabbit anti-Olig2 antibodies were generated using bacterially purified GST-Olig2 recombinant proteins. TUNEL assays were performed using In Situ Cell Death Detection Kit (Roche).

**Luciferase Assays.** P19 embryonic carcinoma cells were cultured in α-minimal essential media supplemented with 2.5% FBS and 7.5% bovine calf serum. For luciferase assays, P19 cells were seeded and incubated for 24 h, and transient transfections were performed using Lipofectamine 2000 (Invitrogen). An actin promoter-β-galactosidase plasmid was cotransfected for normalization of transfection efficiency, and empty vectors were used to equalize the total amount of DNA. Cells were harvested 36–40 h after transfection. For some experiments, cells were treated with LIF for 16 h before harvest for the subsequent luciferase and β-galactosidase assays. Cell extracts were assayed for luciferase activity and the values were normalized with β-galactosidase activity. Data are represented as means of triplicate or duplicate values obtained from represented experiments. All transfections were repeated independently at least four times. Luciferase reporter data are shown in relative luciferase units (mean ± SD).

**CoIP and Western Blotting Assays.** HEK293T cells were cultured in DMEM media supplemented with 10% FBS. For coimmunoprecipitation (CoIP), cells were seeded in 100-mm dishes, incubated for 24 h, and transfected using SuperFect (QIAGEN). For in vivo CoIP experiments, spinal cords from E12.5 WT mice were microdissected, dissociated, and pooled. Cells were treated with LIF (1,000 μ/mL) for 16 h before harvest. Cells were lysed in IP buffer (20 mM Tris·HCl, pH 8.0, 0.5% Nonidet P-40, 1 mM EDTA, 150 mM NaCl, 2 mM PMSF, 10% Glycerol, 4 mM $Na_3VO_4$, 200 mM NaF, 20 mM Na-pyrophosphate, and proteinase inhibitor mixture), and subject to IP overnight at 4 °C followed by Western blotting assays. For CoIP and Western blotting assays, the following antibodies were used: mouse anti-HA (Covance), rabbit anti-STAT3 (#9132, Cell Signaling), rabbit anti-phosph-STAT3 (Tyr705) (#9145, Cell Signaling), rabbit anti-Isl1, rabbit anti-Lhx3, mouse anti-FLAG (Sigma), and rabbit anti-Tubulin antibodies. Blotted membranes were scanned using Odyssey infrared imaging system (Li-COR).

**Statistical Analysis of the Data.** Data are presented as mean ± SD or mean ± SEM as indicated. Statistical comparisons were conducted by two-tailed Student $t$ test. The $P$ values in motif analyses were calculated by the corresponding programs in MEME suite (13).

1. Thaler JP, Lee SK, Jurata LW, Gill GN, Pfaff SL (2002) LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions. *Cell* 110(2):237–249.
2. Lee SK, Pfaff SL (2003) Synchronization of neurogenesis and motor neuron specification by direct coupling of bHLH and homeodomain transcription factors. *Neuron* 38(5):731–745.
3. Lee SK, Jurata LW, Funahashi J, Ruiz EC, Pfaff SL (2004) Analysis of embryonic motoneuron gene regulation: Derepression of general activators function in concert with enhancer factors. *Development* 131(14):3295–3306.
4. Bromberg JF, et al. (1999) Stat3 as an oncogene. *Cell* 98(3):295–303.
5. Bousquet C, Melmed S (1999) Critical role for STAT3 in murine pituitary adrenocorticotropin hormone leukemia inhibitory factor signaling. *J Biol Chem* 274(16):10723–10730.
6. Lee S, et al. (2012) Fusion protein Isl1-Lhx3 specifies motor neuron fate by inducing motor neuron genes and concomitantly suppressing the interneuron programs. *Proc Natl Acad Sci USA* 109(9):3383–3388.
7. Iacovino M, et al. (2011) Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells. *Stem Cells* 29(10):1580–1588.
8. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
9. Cao Y, et al. (2010) Genome-wide MyoD binding in skeletal muscle cells: A potential for broad cellular reprogramming. *Dev Cell* 18(4):662–674.
10. Valouev A, et al. (2008) Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat Methods* 5(9):829–834.
11. Dennis G, Jr., et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4(5):3.
12. Machanick P, Bailey TL (2011) MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics* 27(12):1696–1697.
13. Bailey TL, et al. (2009) MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res* 37(Web Server issue):W202–208.
14. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS (2007) Quantifying similarity between motifs. *Genome Biol* 8(2):R24.
15. Bailey TL, Gribskov M (1998) Combining evidence using p-values: Application to sequence homology searches. *Bioinformatics* 14(1):48–54.
16. Buske FA, Bodén M, Bauer DC, Bailey TL (2010) Assigning roles to DNA regulatory motifs using comparative genomics. *Bioinformatics* 26(7):860–866.
17. McLeay RC, Bailey TL (2010) Motif enrichment analysis: A unified framework and an evaluation on ChIP data. *BMC Bioinformatics* 11:165.
18. Siepel A, et al. (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 15(8):1034–1050.
19. Tsuchida T, et al. (1994) Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. *Cell* 79(6):957–970.
20. Sharma K, et al. (1998) LIM homeodomain factors Lhx3 and Lhx4 assign subtype identities for motor neurons. *Cell* 95(6):817–828.

21. Jurata LW, Kenny DA, Gill GN (1996) Nuclear LIM interactor, a rhombotin and LIM homeodomain interacting protein, is expressed early in neuronal development. *Proc Natl Acad Sci USA* 93(21):11693–11698.

22. Joshi K, Lee S, Lee B, Lee JW, Lee SK (2009) LMO4 controls the balance between excitatory and inhibitory spinal V2 interneurons. *Neuron* 61(6):839–851.

23. Thaler J, et al. (1999) Active suppression of interneuron programs within developing motor neurons revealed by analysis of homeodomain factor HB9. *Neuron* 23(4): 675–687.
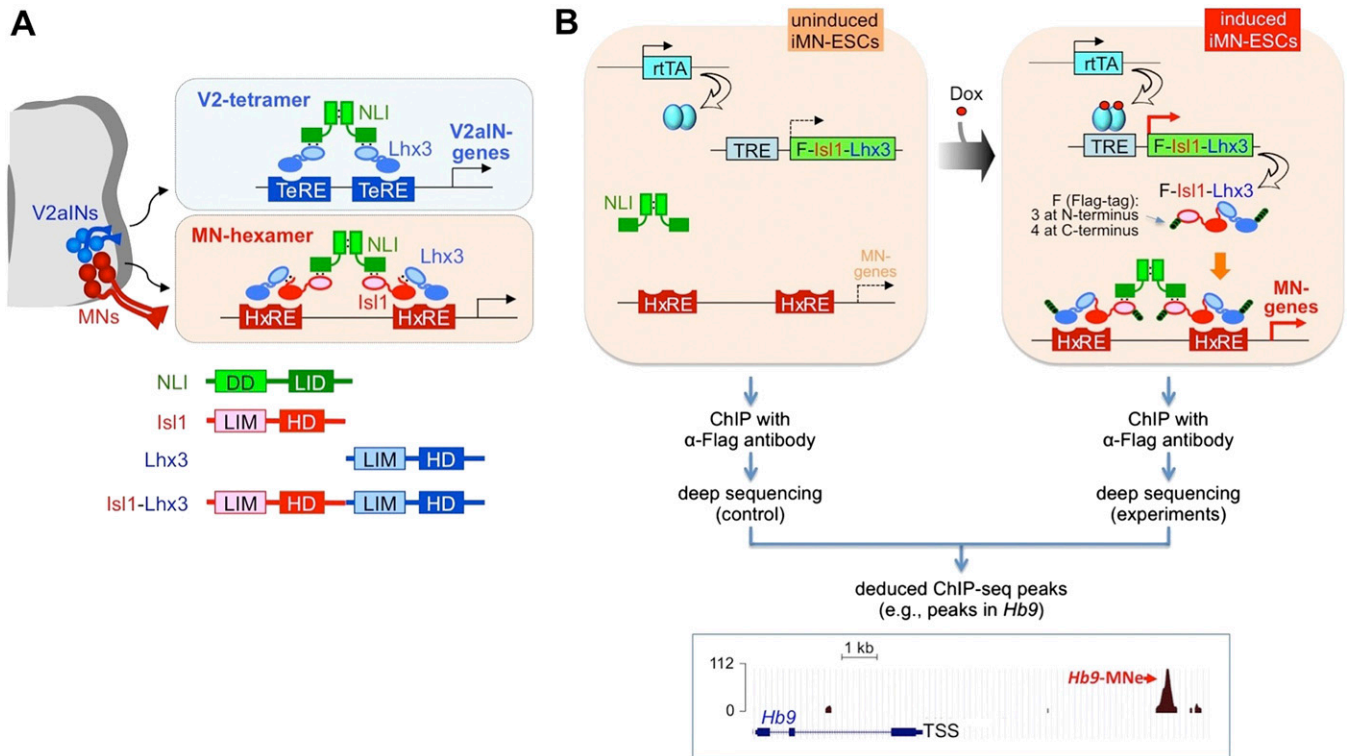
**Fig. S1.** Schematic representation of the cell type-specific LIM complexes and ChIP-seq experiments. (*A*) Schematic representation of the LIM complexes and their target genes. The MN–hexamer and V2–tetramer complexes direct the specification of MNs and V2a interneurons (V2aINs), respectively, in the developing spinal cord. In differentiating MNs, the MN–hexamer consisting of two Lhx3, two Isl1, and two nuclear LIM interactor (NLI) molecules binds to the HxRE in MN genes and induce their expression. During V2-interneuron specification, the V2–tetramer composed of two NLI and two Lhx3 molecules is recruited to TeRE (V2-tetramer-response elements) to active V2a-interneuron genes. (*B*) In iMN-ESCs, the expression of a FLAG-tagged Isl1-Lhx3 fusion protein is controlled by tetracycline response element (TRE), whereas reverse tetracycline transactivator (rtTA) is constitutively expressed. In the absence of Dox, FLAG-Isl1-Lhx3 is not expressed and therefore the MN genes are not activated. Upon addition of Dox, rtTA binds to TRE and activates the expression of FLAG-Isl1-Lhx3, which forms the MN–hexamer complex with endogenous NLI. Then, the MN–hexamer is recruited to the HxRE motif in MN genes. ChIP with anti-FLAG antibody was performed with uninduced and induced iMN-ESCs, and immunopurified DNA samples were subjected to deep sequencing. The ChIP samples from uninduced iMN-ESCs served as a control to deduce Isl1-Lhx3–binding peak profiles, as exemplified by the *Hb9* gene locus shown in the box.
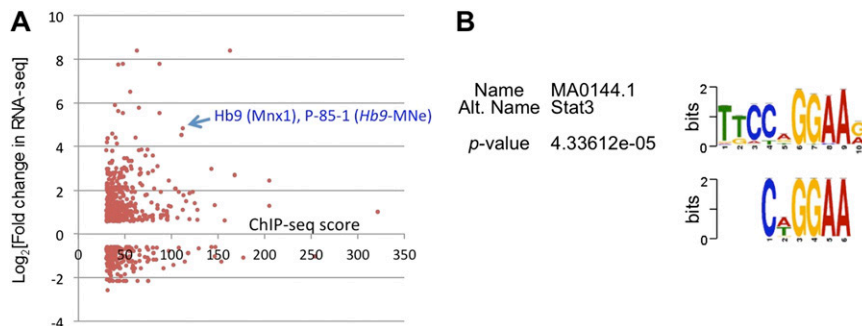


**Fig. S2.** ChIP-seq data analyses. (*A*) The analyses to combine the genomic Isl1-Lhx3–binding site profile (ChIP-seq) and Isl1-Lhx3–induced gene transcriptome changes (RNA-seq). A total of 460 ChIP-seq peaks, which are associated with 218 genes whose expression changes upon Isl1-Lhx3 expression, were plotted in this graph. A total of 378 peaks are associated with 167 genes, whose expression is induced by Isl1-Lhx3. The arrow indicates the Hb9 (Mnx1) gene associated with the ChIP-seq peak P-85-1 (also previously identified as *Hb9*-MNe). x axis, ChIP-seq score; y axis, fold changes in RNA-seq. (*B*) STAT3 site is highly enriched in Isl1-Lhx3–bound ChIP-seq peaks. One of highly enriched motifs in Isl1-Lhx3–bound ChIP-seq peaks was identified to be a STAT3-binding motif compared with a JASPAR CORE 2009 database of known transcription factor–binding sites using the TOMTOM algorithm. The figure shows the outcome of search by the TOMTOM algorithm.

**A**

| Gene Name | Peak_id | ChIP-seq enric. fold | RNA-seq fold change | HxRE-long | HxRE-short | STAT3 site |
|---|---|---|---|---|---|---|
| Fgf9 | P-35-1 | 142.9 | 7.9 | Yes | Yes | Yes |
| Mnx1 (Hb9) | P-85-1 (Hb9-MNe) | 111.6 | 28.5 | Yes | Yes | Yes |
| 9530008L14Rik | P-87-1 | 110.8 | 22.8 | Yes | Yes | Yes |
| Arhgap36 | P-355-1 | 65.0 | 54.9 | | Yes | Yes |
| Chodl | P-538-1 | 55.7 | 91.4 | Yes | Yes | Yes |
| Hecw2 | P-504-2 | 54.1 | 10.1 | Yes | Yes | Yes |
| Fgf9 | P-674-1 | 51.4 | 7.9 | Yes | Yes | Yes |
| Hs3st5 | P-2078-1 | 31.6 | 20.9 | | Yes | Yes |
| Megf11 | P-1522-1 | 30.8 | 13.9 | | Yes | Yes |
| Ppargc1a | P-651-1 | 28.9 | 10.9 | Yes | Yes | Yes |
| Fam19a4 | P-222-1 | 15.3 | 12.0 | Yes | Yes | Yes |

**B**



**C**
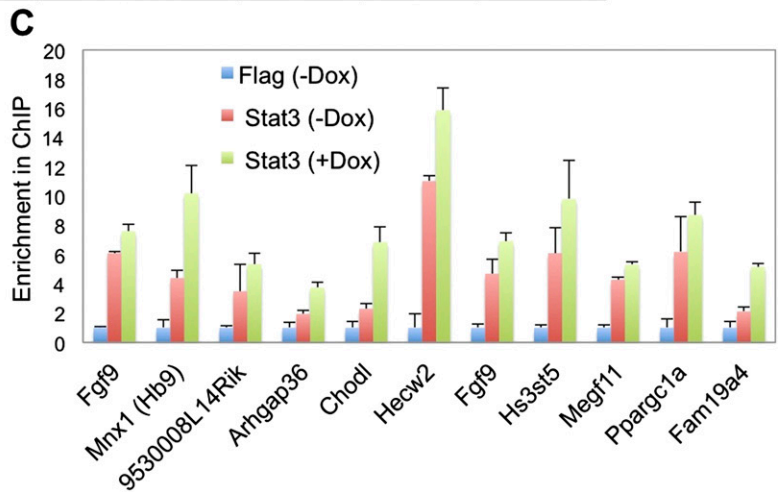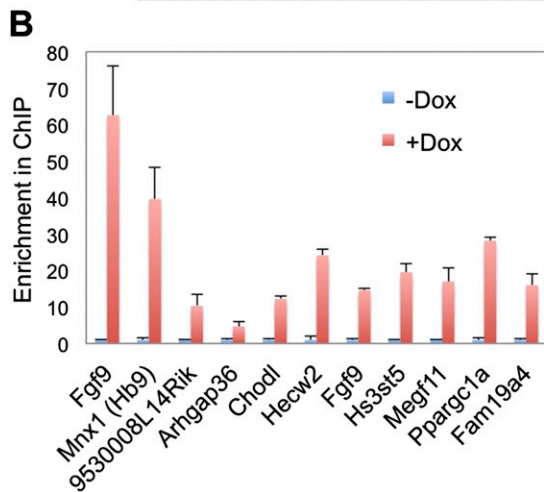


**Fig. S3.** The MN–hexamer and STAT3 are recruited to hexamer target genes that have both HxRE and STAT3 sites. (*A*) The list of selected MN–hexamer target genes and their associated Isl1-Lhx3–bound peaks that have both HxRE and STAT3 sites. ChIP-seq enric. fold, fold enrichment in ChIP-seq analyses for each peak; RNA-seq fold change, induction fold of each gene after Dox treatment (i.e., Isl1-Lhx3 expression). (*B* and *C*) ChIP assays with α-FLAG or α-STAT3 antibodies in iMN-ESCs with or without Dox treatment (i.e., expression of FLAG-tagged Isl1-Lhx3). The cells were treated with LIF in all conditions. (*B*) Isl1-Lhx3 bound to each peak region. (*C*) The recruitment of STAT3 to each peak region did not require Isl1-Lhx3, and was slightly enhanced upon the expression of Isl1-Lhx3. Error bars represent the SD.

**Fig. S4.** STAT3 facilitates the MN–hexamer–binding to target enhancers. (*A*) Western blotting assays in P19 cells transfected with sh-RNA vectors as listed at the top. shRNA-STAT3 #1 and shRNA-STAT3 #2 efficiently down-regulated endogenous STAT3 protein. (*B*) ChIP assays in P19 cells transfected FLAG-Isl1-Lhx3 and shRNA constructs. The cells were treated with LIF in all conditions. Isl1-Lhx3 was recruited to the ChIP-seq peak regions, and the Isl1-Lhx3–binding to each ChIP-seq peak was reduced when STAT3 was knocked down. scr., scrambled sequences. Error bars represent the SD.



**Fig. S5.** STAT1 and STAT3 are needed for MN differentiation. Analyses of chick spinal cord electroporated with either STAT3-DNs (*A*) or shRNA constructs (*B*). Only the ventral portion of the neural tube is shown. +, electroporated side. GFP marks the electroporated cells. (*A*) Inhibition of STAT3 activity using STAT3-DNs resulted in reduction of the number of Hb9$^+$ MNs, but it did not change the number of Chx10$^+$ V2-interneurons. (*B*) Knockdown analyses in chicken embryos electroporated with shRNA constructs.

**Dataset S1.   A list of ChIP-seq peaks that recruit Isl1-Lhx3 in iMN-ESCs**

[Dataset S1](#)

This dataset shows the genome coordinates for each ChIP-seq peak. The Isl1-Lhx3–bound chromatin fragments immunopurified with α-FLAG antibodies from Dox-induced iMN-ESCs are marked as ChIP. α-FLAG–immunoprecipitated chromatin fragments from non-Dox–treated iMN-ESCs are marked as background. The cutoff of ChIP score is set at 30.

**Dataset S2.   A list of RefSeq genes that have at least one Isl-Lhx3–bound ChIP-seq peak within 20 kb from a RefSeq transcription unit or the RefSeq genes located nearby to intergenic Isl1-Lhx3–bound ChIP-seq peaks**

[Dataset S2](#)

    This dataset shows the genome coordinates for the genes associated with ChIP-seq peaks (peak_id) and the location of the ChIP-seq peaks relative to the genes.

**Dataset S3.   The combined analyses of ChIP-seq and RNA-seq data**

[Dataset S3](#)

    The predicted target genes from ChIP-seq are associated with the differentially expressed genes from RNA-seq only if the ChIP-seq peak is found within 20 kb from the gene body; the induction fold is ≥1.5 in normalized RNA-seq expression levels.

**Dataset S4.   Lists of ChIP-seq peak ID and the sequences of identified motifs in each peak**

[Dataset S4](#)

    These datasets show ChIP-seq peak ID and the sequences of HxRE-long, HxRE-short, and Stat3 binding motifs, which were used to generate the motif logos in Figs. 1*E* and 2*A*.