Supplementary Information:

Methods:

No statistical methods were used to predetermine sample size and the investigators were not blinded to allocation during experiments and outcome assessment.

Reprogramming and efficiency calculation:

Passage three wild-type, heterozygous TauEGFP knock-in MEFs or human embryonic fibroblasts were infected with indicated lentivirus (available from us through Addgene or specified in Table S4) in MEF media containing 8 µg/ml polybrene (Sigma) for 16-20 h. Cells were switched into fresh MEF media containing 2 µg/ml doxycycline (Sigma) to induce transgene expression for 2 days. At day 3 cells were switched into N3 media (DMEM/F12) containing N2 supplement, B27, 20 µg/ml Insulin, penicillin/streptomycin (all from Invitrogen) and doxycycline to continue reprogramming. The media was changed every 2–3 days for the duration of the reprogramming. To calculate the efficiency of neuronal induction the total number of TauEGFP and/or TUJ1 expressing cells with complex neurite outgrowth (cells having a spherical cell body and at least one thin process three times the size of their cell body), were quantified manually 7 or 14 days after transgene induction by immunofluorescence microscopy. To calculate the efficiency of myocyte-like cells the total number of Desmin and MYH expressing cells was quantified manually 14 days after transgene induction by immunofluorescence microscopy. The quantification was based on the average number of neuronal or myocyte-like cells present in a minimum of 30 randomly selected 20x fields of view from at least three biological replicates. The number of reprogrammed cells was then either normalised to the number of reprogrammed cells in the control condition or to the total number of cells determined by DAPI staining. TauEGFP-expressing cells in Fig. 4a and ED9b were quantified using a LSRFortessa FACS analyser (BD Biosciences) based on the average number of TauEGFP positive cells detected in 60.000 to 120.000 analysed cells from at least three to six biological replicates.

In utero electroporation:

All animal protocols have been approved by Stanford University. For in utero electroporation pregnant CD-1 mice (Charles River) were deeply anesthetized with Isoflurane (Henry Schein), after which the uterine horns were carefully exposed through a midline abdominal incision. pSico constructs encoding both GFP and indicated shRNA-oligos from the same DNA plasmid (1–2 µl of 2 µg/µl) were diluted in PBS containing 0.01% fast green (Sigma) as a tracer and injected in utero into the lateral ventricle of mouse embryos at E13.5 using a micropipette made from G-1 glass capillaries (Narishige). After injection, the embryo in the uterus was placed between a 5 mm platinum tweezertrode and five 50 ms square pulses of 25 V with 950 ms intervals were applied with an ECM 830 electroporation system (both from Harvard Apparatus). We performed randomized electroporations of control or Myt1l-targeting shRNAs in the right and left uterine horns of several pregnant mice to avoid any technical effect on the experimental outcome. Then, uterine horns were placed back into the abdominal cavity, and the abdominal wall of the pregnant mouse was sutured. Embryonic brains were harvested and dissected 2 days (E15.5) after electroporation for subsequent immunofluorescence processing and analysis. Brain sections were then co-stained with indicated antibodies and anti-GFP to identify electroporated cells expressing the indicated knock-down constructs. The distribution of GFP positive cells as well as the number of GFP and respective marker double positive cells was determined by dividing the number of GFP positive cells in each layer or the number of double positive cells by the total number of GFP-positive cells in the entire section, respectively.

Electrophysiology:

Electrophysiological recordings were performed from mouse primary hippocampal neurons 14 days after *in vitro* culture and 11 days upon shRNA-infection or MEF-derived iN cells 21 days after transgene induction and 14 days after co-culturing on glia following pervious protocols [9]. In brief, action potentials were recorded using current-clamp configuration with pipette solution containing (in mM): 130 $KMeSO_3$, 10 NaCl, 2 $MgCl_2$, EGTA 0.5 for MEF-iN 1 for primary cultures, $CaCl_2$ 0.16 for MEF-iN 0.1 for primary cultures, 4 Na_2ATP, 0.4 NaGTP, 14 Tris-creatine phosphate, and 10 HEPES-KOH (pH adjusted to 7.3, 310 mOsm). The bath solution contained (in mM): 140 NaCl, 5 KCl, $CaCl_2$ 2 for MEF-iN 3 for primary cultures, 1 $MgCl_2$, 10 glucose, and 10 HEPES-NaOH (pH 7.4). Membrane potentials were kept around -60 mV using small holding currents, and step currents were injected to elicit action potentials. Recordings of the intrinsic and active membrane properties were performed in the presence of 50 µM picrotoxin, 10 µM CNQX and 50 µM D-AP5 in the bath solution (all from Tocris). The synaptic current recordings were performed in voltage-clamp mode with internal solution containing (in mM): 135 $CsCl_2$, 1 EGTA, 4 Na_2ATP, 1 Na-GTP, 1 QX-314, and 10 HEPES-NaOH (pH adjusted to 7.4, 310 mOsm). AMPA-receptor EPSCs were pharmacologically isolated by application of 50 µM picrotoxin, and were subsequently blocked by addition of 50 µM CNQX. Evoked synaptic responses were triggered by 1-ms, 1 mA current injection through a local extracellular electrode (FHC concentric bipolar electrode) with a Model 2100 Isolated Pulse Stimulator (A-M Systems). All recordings were performed in whole-cell configuration using a Multiclamp 700B amplifier (Molecular Devices). All average data were analyzed from three or more biological replicates using Clampfit 10.4 (Axon Instruments).

Cell line generation and maintenance:

Mouse embryonic fibroblasts (MEFs) were isolated from wild-type or heterozygous TauEGFP knock-in mouse embryos at E13.5 (Jackson Laboratories) after removal of all neural cell containing tissues. Mouse glial cells were isolated from forebrains of wild-type mice (Jackson Laboratories) at postnatal day two. Both cell types were maintained in MEF media (DMEM; Invitrogen) containing 10% cosmic calf serum (CCS; Hyclone), beta-mercaptoethanol (Sigma), non-essential amino acids, sodium pyruvate and penicillin/streptomycin (all from Invitrogen) and passaged three times before experiments. Mouse neural stem cells (NSCs) were isolated from forebrains at E13.5 or cortex at E14.5 of wild-type mouse embryos. The cells were maintained on PO-laminin coated tissue culture dishes or as neurospheres in proliferation medium (DMEM/F12) containing N2 or B27 supplement, penicillin/streptomycin (all from Invitrogen), and 20 ng/ml EGF, 10 ng/ml FGF (both from Peprotech). Cells were infected and kept in proliferation medium for 7 days after transgene induction using doxycycline or changed one day after infection to N3 media (DMEM/F12) containing N2 supplement, B27, 20 µg/ml Insulin, penicillin/streptomycin (all from Invitrogen) 1 % fetal bovine serum (FBS; Hyclone), and doxycycline to induce differentiation for 7 days. For primary neuronal cultures the hippocampus of P0 wild-type mouse pups was isolated and cultured on Matrigel-coated plates (Corning) in MEM supplemented with B27, glucose, transferrin and 5% FBS. 2 days after plating the medium was supplemented with 2 mM Ara-C (Sigma) as published previously [29]. Primary myoblasts were isolated and cultured on Collagen-coated plates (Corning) in DMEM supplemented with 20% FBS and penicillin/streptomycin (Invitrogen) as described earlier [30]. One day after infection and transgene induction using doxycycline primary myoblasts were changed to DMEM containing penicillin/streptomycin and 5 % horse serum (all from Invitrogen) to induce muscle differentiation for 4 days. The media for all cells was changed every 2–3 days for the duration of the experiment and all cells were grown at 37 °C and 5% CO_2. Cells were tested negative for mycoplasma contamination.

Virus production and infection:

Lentivirus was produced by transfection of lentiviral backbones containing the indicated transgenes together with third-generation packaging plasmids into HEK293 cells following the Trono laboratory protocol [31]. Virus was concentrated from culture supernatant by ultra-centrifugation (23000 rpm, 2 h, 4 °C) and cells were infected with three different viral titers. Infected cell populations that were used for reprogramming and genomic analysis were verified by immunofluorescence to contain approximately 70%-90% transgene-positive cells two days after induction based on immunofluorescence microscopy.

Plasmid constructs:

DNA constructs were generated by PCR amplification with Phusion polymerase followed by restriction digest using indicated enzymes and ligation with T4 DNA ligase (all from NEB). Site directed mutations of *Myt1l ZF2-3* were generated using the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent Technologies) following the manufacturer's instructions. A complete list of all constructs and primers generated in this study can be found in Table S4&6.

Protein expression and purification:

Proteins for immunisation and GST pull downs were expressed in *E. coli* Rosetta DE3 (Novagen). All His- and His-GST-tagged Myt1l fragments were purified under native conditions using Ni-NTA Agarose (Qiagen) and eluted in bacterial lysis buffer containing (in mM): 20 Tris pH 7.5, 500 NaCl, 1 $MgCl_2$ supplemented with 400 imidazole (all from Sigma). Eluted proteins were dialysed overnight against phosphate buffered saline (PBS) and supplemented with 10% glycerol (both from Sigma) prior to use or storage at -80 °C.

Antibodies:

Polyclonal antibodies against Myt1l were raised in rabbits using recombinant His-tagged mmMyt1l aa170-240. Immune sera were affinity purified against the antigen immobilised on Affigel (Bio-rad) and eluted in buffer containing (in mM): 100 glycine pH 2.7, 150 NaCl followed by 100 glycine pH 2.2, 150 NaCl. 50 ml anti-Ascl1 hybridoma supernatant, 2.5 mg affinity purified anti-Myt1l antibody, and mouse or rabbit control IgG (Sigma) were coupled to 250 µl protein A Sepharose (GE Healthcare) using buffer containing (in mM): 20 dimethylpimelimidate in 200 $Na_2B_4O_7$ pH 9, blocked with 1000 Tris pH 8, followed by 200 ethanolamine pH 8, washed with 200 $Na_2B_4O_7$ pH 9, washed with 200 glycine pH 2.2, 150 NaCl, and finally washed extensively using PBS (all from Sigma). A complete list of all primary antibodies used in this study can be found in Table S7. Secondary Alexa-conjugated antibodies were used at 1:2000 (all from Invitrogen), secondary HRP-conjugated antibodies were used at 1:5000 (all from Jackson Immunoresearch) and secondary IRDye-conjugated antibodies were used at 1:5000 (all from LI-COR).

Immunofluorescence:

Cultured cells were washed with PBS and fixed using 4% paraformaldehyde (PFA: Affymetrix) for 10 min. Cells were then permeabilised in 0.5% Triton X-100 (Sigma) in PBS for 5 min, blocked in a solution of PBS containing 2% bovine serum albumin (BSA; Sigma) for 1 h, followed by incubation with primary antibody for 1 h at room temperature or overnight at 4°C. Cells were washed three times for 15 min using blocking solution prior to incubation for 30 min with secondary antibodies. For DNA staining 100 ng/ml DAPI was added to the last washing step (Invitrogen). Embryonic mouse brains were

dissected and immersion fixed in 4% PFA overnight, while adult mice were transcardially perfused first with PBS followed by ice cold 4% PFA, then the brains were isolated and post fixed in 4% PFA overnight. After fixation embryonic and adult brains were cryoprotected overnight in 30% sucrose in PBS. Embryonic brains were embedded in OCT compound (Sakura), frozen in dry ice, cryostat sectioned at 20 µm and mounted on glass slides. Adult brains were sectioned at 40 µm using a freezing sliding microtome. For immunofluorescence, mounted (embryonic) or free-floating (adult) sections were washed in potassium phosphate buffer (KPBS). Then incubated in KPBS with 0.25% Triton-X100 and 5% normal donkey and goat serum (pre-incubation solution; both from Jackson Immunoresearch) for 1 h at room temperature followed by primary antibody incubation in pre-incubation solution overnight at 4°C. Following washes in KPBS with 0.25% Triton-X100, sections were incubated with secondary antibodies and Hoechst (Invitrogen) in pre-incubation solution for 2 h at room temperature. After final washes in KPBS free-floating sections were mounted on gelatin-coated slides and sealed using glycerol based mounting media. Microscopy images were obtained using a DM6000 B microscope equipped with a 20x HCX PL air objective (NA 0.4) and a DFC365 FX digital camera (all from Leica).

Immunoprecipitation experiments (IP):

For each immunoprecipitation $20x10^6$ MEF cells infected with the indicted transgenes two days after induction were lysed in 1 ml cell lysis buffer containing (in mM): 0.5% Tween-20, 50 Tris pH 7.5, 2 EDTA, 1 DTT, 1 PMSF, 5 NaF (all from Sigma), and complete protease inhibitor (Roche) for 15 min at 4 °C. Nuclei were pelleted by centrifugation (3200 rpm, 1 min, 4 °C) and resuspended in 1 ml NP-40 lysis buffer containing (in mM): 0.5% NP-40, 50 Tris pH 8, 150 NaCl, 2 EDTA, 1 DTT, 1 PMSF, 5 NaF, 5 µg/ml cytochalasin B (all from Sigma), complete protease inhibitor (Roche), and benzonase (Merck). Debris was removed by centrifugation (14000 rpm, 10 min, 4 °C) and nuclear lysate was pre-cleared by incubation with 15 µl uncoupled protein A beads prior to addition with 15 µl antibody coupled beads for 2 h at 4 °C. After binding the beads were washed extensively in NP-40 lysis buffer and bound proteins were eluted with SDS-PAGE sample buffer.

GST pull-down experiments (GST-PD):

For each pull-down $20x10^6$ MEF cells were lysed in 1 ml cell lysis buffer containing (in mM): 0.5% Tween-20, 50 Tris pH 7.5, 2 EDTA, 1 DTT, 1 PMSF, 5 NaF (all from Sigma), and complete protease inhibitor (Roche) for 15 min at 4 °C. Nuclei were pelleted by centrifugation (3200 rpm, 1 min, 4 °C) and resuspended in 0.5 ml cell lysis buffer containing 500 mM NaCl, 5 µg/ml cytochalasin B (Sigma), and benzonase (Merk). Debris was removed by centrifugation (14000 rpm, 10 min, 4 °C) and nuclear lysate was diluted with cell lysis buffer 1:1.25 (200 mM NaCl c.f.). 50 µg of GST alone or GST-tagged Myt1l fragments were incubated with 20 µl glutathione Sepharose 4B beads (GE Healthcare) in PBS supplemented with 2 µg/ml BSA (Sigma) and complete protease inhibitor (Roche) for 1 h at 4 °C. Beads were then washed extensively in cell lysis buffer containing 200 mM NaCl prior to addition of diluted nuclear lysate and incubation for 2 h at 4 °C. After binding the beads were washed extensively in cell lysis buffer containing 200 mM NaCl and bound proteins were eluted with SDS-PAGE sample buffer.

Chromatin immunoprecipitation sequencing (ChIP-seq) and computational analysis:

For each ChIP-seq experiment either 4 brains from wild-type E13.5 mouse embryos or $20\text{-}40x10^6$ primary cells two days after transgene induction were cross-linked using 1% formaldehyde (EMS) for

10 min followed by lysis in 1 ml swelling buffer containing (in mM): 0.5% NP-40, 5 HEPES pH 7.9, 85 KCl, 1 DTT, 1 PMSF (all from Sigma), and complete protease inhibitors (Roche) for 20 min on ice. Nuclei were pelleted by centrifugation (3200 rpm, 1 min, 4 °C) and lysed in 1 ml lysis buffer containing (in mM): 1% SDS, 50 Tris pH 8, 10 EDTA, 1 DTT, 1 PMSF (all from Sigma), and complete protease inhibitors (Roche) for 10 min on ice. Chromatin was sheared using a Bioruptor sonicator (Diagenode) until DNA was fragmented to 200-500 bp followed by 1:4 dilution with buffer containing (in mM): 1% Triton X-100, 20 Tris pH 8, 2 EDTA, 150 NaCl, 1 PMSF (all from Sigma), and complete protease inhibitors (Roche). Diluted lysate was pre-cleared by incubation with 15 µl Staph A cells or 50 µl protein A beads prior to addition with antibody (20 µg Sin3b or 50 µg HDAC1) or 50 µl antibody coupled beads (Myt1l) overnight at 4 °C, respectively. After binding the antibody coupled beads were washed extensively in wash buffer containing (in mM): 1% NP-40, 0.05% SDS, 20 Tris pH 8, 250 NaCl, 2 EDTA (all from Sigma), and complete protease inhibitors (Roche). The reactions with uncoupled antibody were supplemented with 15 µl Staph A cells washed extensively in the following three wash buffers containing (in mM): 1% Triton X-100, 0.1% SDS, 20 Tris pH 8, 150 NaCl, 2 EDTA; 1% Triton X-100, 0.1% SDS, 20 Tris pH 8, 500 NaCl, 2 EDTA; and 1% NP-40, 1% DOC, 10 Tris pH 8, 250 LiCl, 1 EDTA. After washing, the bound fraction was eluted in 100 µl elution buffer containing (in mM): 1% SDS, 50 Tris pH 8, 10 EDTA, 100 PMSF and reversed cross-linked by overnight incubation at 65 °C. The isolated DNA was RNase (NEB) treated and purified using QIAGEN columns. Libraries were generated using the NEBNext ChIP-seq Library Prep Master Mix Set for Illumina (NEB) and single-end sequencing reads (50bp) were generated on Hi-Seq 2000 platforms (Illumina). ChIP-sequence reads are available on NCBI GEO accession GSE72121. Raw reads were mapped to mouse reference genome GRCm38/mm10 using bowtie2 (version 2.2.3) allowing maximum one mismatch [32]. Peaks for each sample were called using MACS2 algorithm (version 2.0.10.20131216) [33] using the shift size values calculated from the run_spp.R script from the SPP peak caller [34]. High quality peaks were identified using IDR2 (https://github.com/nboley/idr version 2.0.2). A negative control ChIP was performed as described above using the Myt1l antibody on rtTA-only infected. Any rtTA peak that overlapped a peak in any other condition by greater than 50% was excluded from the final analyses. Heatmaps of ChIP-seq signals were generated around peak summits (± 2kb region) using 25 bp sliding windows with the HOMER suite (version 4.7) [35]. Read histograms, motif analysis, peak annotation, and GO analysis was also performed using HOMER and PHANTER [28]. Both Bedtools and SAMtools were used for file processing and format conversions [36,37]. Differential Myt1l localisation was determined using the DiffBind R package [38]. All ChIP-sequence data was analysed from two or three biological replicates.

MNase-seq analysis:

To assess nucleosome occupancy at candidate regions, we utilised a previously reported MEF MNase-seq dataset (GSE40896) [25]. The raw reads were mapped to GRCm38/mm10 using bowtie (version 1.1.1) allowing for one mismatch as described [39]. Histograms of read densities for the Myt1l peaks and Ascl1 peaks [8](lifted over to mm10) were compiled using HOMER.

Transcript analysis using RNA-seq and quantitative PCR:

RNA was isolated using Trizol (Invitrogen) and treated with DNase (NEB). For RNA-sequencing, libraries were prepared following the dUTP protocol [41] and paired-end sequencing reads (100bp) were generated on Hi-Seq 2000 platforms (Illumina). RNA-sequence reads are available on NCBI GEO

accession GSE72121. Raw reads were mapped to mouse reference genome GRCm38/mm10 using TopHat2 (version 2.0.10) [42] and transcript assembly and differential expression determined using cufflinks (version 2.1.1) according to the cuffquant pipeline [43]. CummeRbund was used to generate expression scatterplots and cluster3/treeview were used to generate heatmaps [44-46]. GSEA was performed using all genesets in the MsigDB database (GSE14012) [47,48], including a MEF gene signature that was compiled by taking the genes that were enriched by a factor of 10 in the MEF condition compared to the iN condition from the Wapinski et al. raw RNA-seq data (Table S3) [8]. Significance for GSEA profiles were determined by an FDR < 0.25 as described [48]. Cell type specific gene signatures were derived from the top 10% of the genes unique to the particular cell type [49-52], excluding any genes common amongst 4 out of 5 sets to exclude housekeeping genes (Table S3). Odds ratio analysis was performed using the GeneOverlap R package (version 1.6.0) [27]. For quantitative PCR, RNA was reverse transcribed using Superscript II (Invitrogen) following the manufacturer's instructions. Quantitative real time PCR was performed from cDNA templates using the SYBER Green PCR master mix and a 7900HT Fast Real-Time PCR System (Applied Biosystems). Transcript levels were determined after normalisation against GAPDH using the Relative Expression Software Tool (REST) [26]. A complete list of real time PCR primers used in this study can be found in Table S6. All RNA-sequence data was analysed from two biological replicates.

SELEX:

Constructs encoding Myt1l DNA-binding domains were generated by gene synthesis (codon optimised, Genscript) as indicated in Table S5. Clones were transferred to N-terminal thioredoxin hexahistidine bacterial expression vector (pETG-20A-SBP; [53]) by a Gateway LR reaction (Invitrogen). Proteins were produced in the *E. coli* Rosetta DE3 pLysS (Novagen), and purified, and HT-SELEX analyses were performed as described previously [54]. After each SELEX cycle, the selection ligands were sequenced using Illumina HiSeq 2000 sequencer, and PWM models generated using the seed, cycle and multinomial model reported in Table S5 as described here [54].

Data availability statement:

ChIP and RNA-sequencing data were deposited on NCBI GEO accession number GSE72121.

Supplementary Information References:

26. Pfaffl, M. W., Horgan, G. W. & Dempfle, L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* **30,** e36 (2002).
27. Shen, L. GeneOverlap: An R package to test and visualize gene overlaps. (2014).
28. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nature protocols* **8,** 1551–1566 (2013).
29. Maximov, A., Pang, Z. P., Tervo, D. G. R. & Südhof, T. C. Monitoring synaptic transmission in primary neuronal cultures using local extracellular stimulation. *Journal of Neuroscience Methods* **161,** 75–87 (2007).
30. Springer, M. L., Rando, T. A. & Blau, H. M. Gene delivery to muscle. *Curr Protoc Hum Genet* **Chapter 13,** Unit13.4 (2002).
31. Dull, T. *et al.* A third-generation lentivirus vector with a conditional packaging system. *Journal of Virology* **72,** 8463–8471 (1998).
32. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9,** 357–359 (2012).
33. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biology* **9,** R137 (2008).
34. Kharchenko, P. V., Tolstorukov, M. Y. & Park, P. J. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat Biotechnol* **26,** 1351–1359 (2008).
35. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. **38,** 576–589 (2010).
36. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
37. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).
38. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481,** 389–393 (2012).
39. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10,** R25 (2009).
40. Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Meth* **9,** 215–216 (2012).
41. Levin, J. Z. *et al.* Comprehensive comparative analysis of strand-specific rnA sequencing methods. *Nat Meth* **7,** 709–715 (2010).
42. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes inthe presence of insertions, deletions and genefusions. *Genome Biology* **14,** R36 (2013).
43. Trapnell, C. *et al.* Differential analysis of gene regulation at transcript resolution with rNA-seq. *Nat Biotechnol* **31,** 46–53 (2012).
44. de Hoon, M. J. L., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20,** 1453–1454 (2004).
45. Goff, L. A., Trapnell, C. & Kelley, D. CummeRbund: Visualization and Exploration of Cufflinks High-throughput Sequencing Data. R package version 2.8.2. 1–45 (2014).
46. Saldanha, A. J. Java Treeview--extensible visualization of microarray data. *Bioinformatics* **20,** 3246–3248 (2004).
47. Mootha, V. K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34,** 267–273 (2003).
48. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102,** 15545–15550

(2005).
49. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28,** 511–515 (2010).
50. Bao, X. *et al.* A novel ATAC-seq approach reveals lineage-specific reinforcement of the open chromatin landscape via cooperation between BAF and p63. *Genome Biology* **16,** 284 (2015).
51. Huang, L. *et al.* Partial Hepatectomy Induced Long Noncoding RNA Inhibits Hepatocyte Proliferation during Liver Regeneration. *PLoS ONE* **10,** e0132798 (2015).
52. Zhang, Y. *et al.* An RNA-Sequencing Transcriptome and Splicing Database of Glia, Neurons, and Vascular Cells of the Cerebral Cortex. *Journal of Neuroscience* **34,** 11929–11947 (2014).
53. Jolma, A. *et al.* DNA-Binding Specificitiesof Human Transcription Factors. *Cell* **152,** 327–339 (2013).
54. Nitta, K. R. *et al.* Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. *eLife* **4,** (2015).