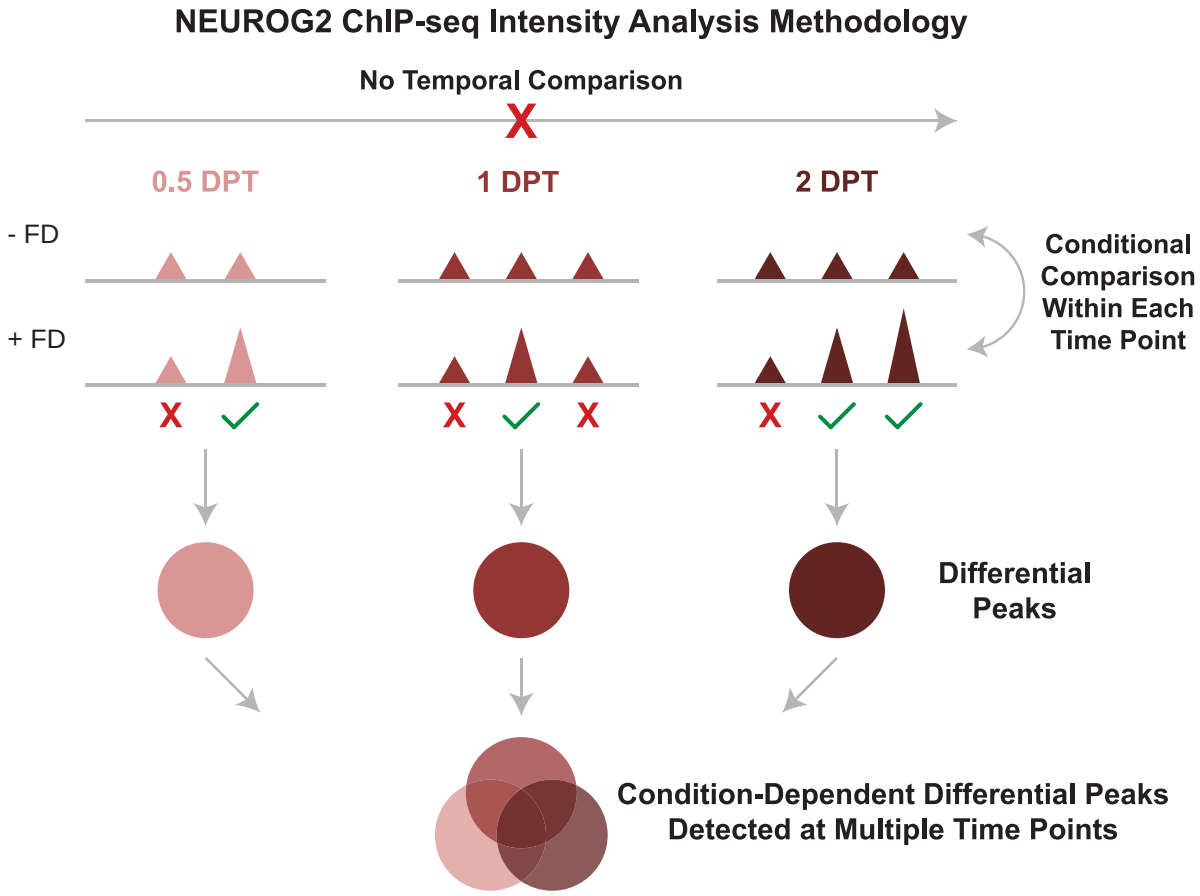
Stem Cell Reports, Volume *7*

Supplemental Information

Small Molecules Modulate Chromatin Accessibility to Promote NEUROG2-Mediated Fibroblast-to-Neuron Reprogramming

Derek K. Smith, Jianjing Yang, Meng-Lu Liu, and Chun-Li Zhang

A



B

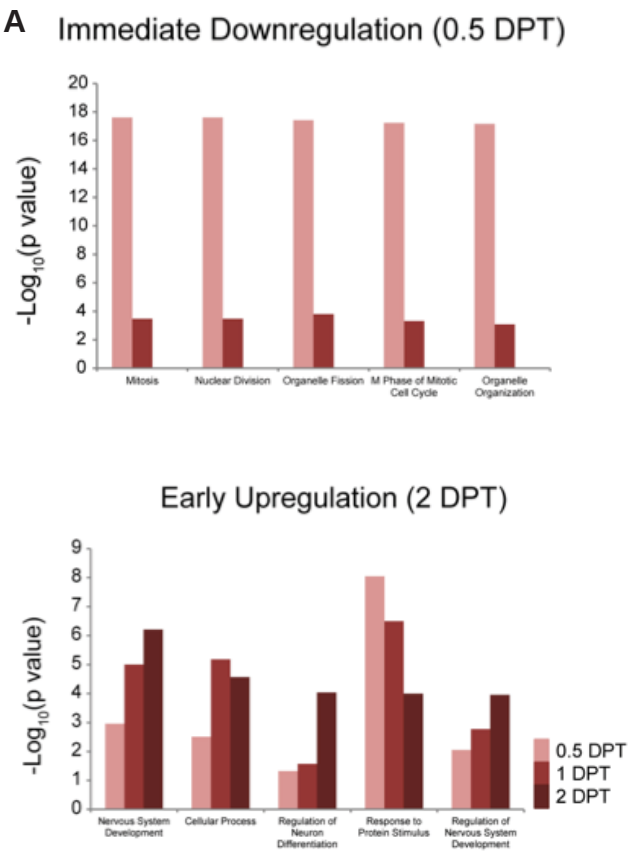
Genomic Distribution of NEUROG2 Binding Events

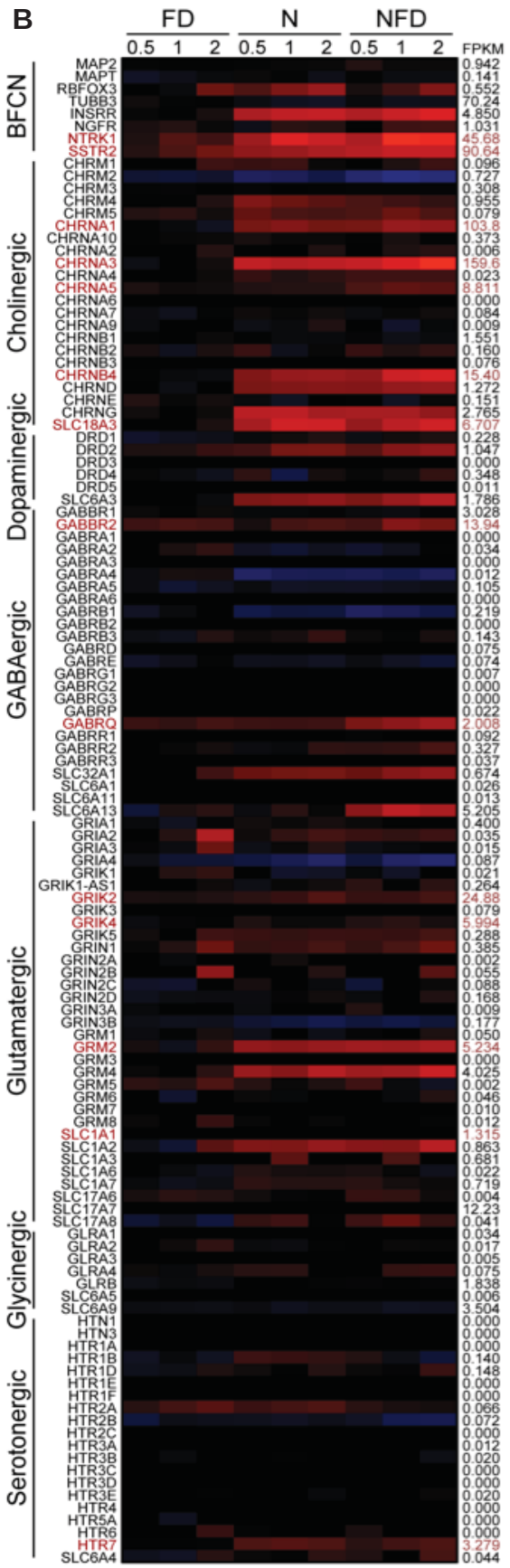
	0.5 DPT			1 DPT			2 DPT		
Location	NFD	Shared	Unique	NFD	Shared	Unique	NFD	Shared	Unique
3' UTR	1.4%	0.2%	1.3%	1.3%	1.4%	1.2%	1.5%	1.5%	1.6%
5' UTR	0.5%	0.3%	0.7%	0.9%	0.6%	1.2%	0.9%	0.6%	1.3%
Exon	2.3%	2.2%	2.3%	3.0%	2.5%	3.6%	3.0%	2.4%	3.8%
Intron	49.7%	49.8%	49.7%	49.1%	49.6%	48.6%	49.6%	49.2%	50.0%
Intergenic	38.1%	40.6%	36.5%	35.9%	38.7%	32.9%	35.8%	39.2%	31.9%
Noncoding	0.6%	0.5%	0.7%	0.7%	0.6%	0.5%	0.7%	0.6%	0.9%
Promoter/TSS	7.4%	5.0%	8.8%	9.0%	6.6%	11.7%	8.4%	6.5%	10.5%

Supplementary Figure 1: Methodological Analyis of ChIP-seq Datasets and Genomic Distribution of NEUROG2 Binding Events , related to Figure 1.

(A) Schematic representation of NEUROG2 ChIP-seq data analysis method.

(B) The genome-wide distribution of NEUROG2 binding events relative to annotate gene structures (TSS, transcription start site; UTR, untranslated region).

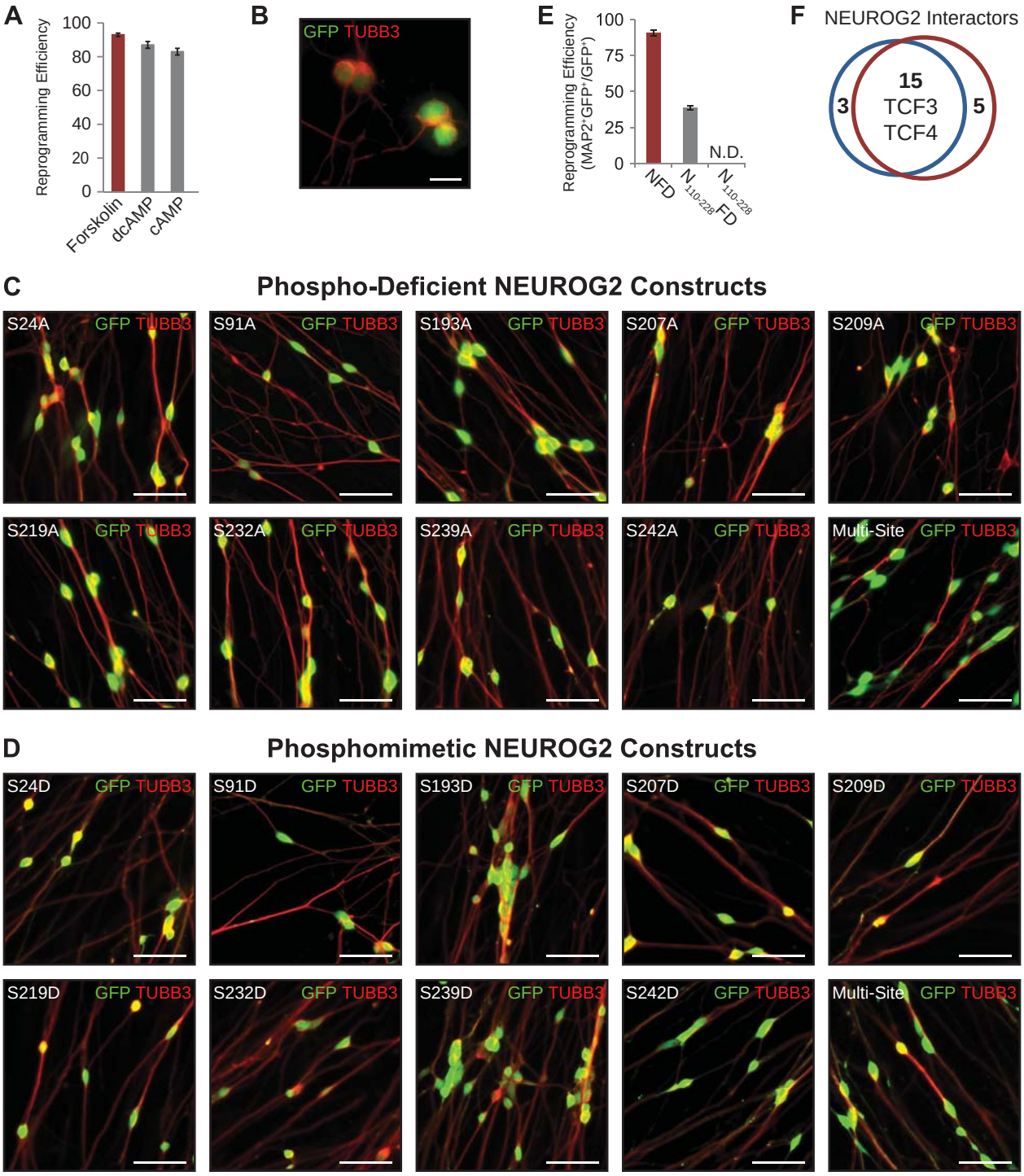

A
Immediate Downregulation (0.5 DPT)
-Log10(p value)
Mitosis
Nuclear Division
Organelle Fission
M Phase of Mitotic Cell Cycle
Organelle Organization
Early Upregulation (2 DPT)
-Log10(p value)
Nervous System Development
Cellular Process
Regulation of Neuron Differentiation
Response to Protein Stimulus
Regulation of Nervous System Development
0.5 DPT
1 DPT
2 DPT


B
FD
N
NFD
0.5 1 2 0.5 1 2 0.5 1 2
FPKM
BFCN
Cholinergic
Dopaminergic
GABAergic
Glutamatergic
Glycinergic
Serotonergic
MAP2 0.942
MAPT 0.141
RBFOX3 0.552
TUBB3 70.24
INSRR 4.850
NGFR 1.031
NTRK1 45.68
SSTR2 90.64
CHRM1 0.096
CHRM2 0.727
CHRM3 0.308
CHRM4 0.955
CHRM5 0.079
CHRNA1 103.8
CHRNA10 0.373
CHRNA2 0.006
CHRNA3 159.6
CHRNA4 0.023
CHRNA5 8.811
CHRNA6 0.000
CHRNA7 0.084
CHRNA9 0.009
CHRNB1 1.551
CHRNB2 0.160
CHRNB3 0.076
CHRNB4 15.40
CHRND 1.272
CHRNE 0.151
CHRNG 2.765
SLC18A3 6.707
DRD1 0.228
DRD2 1.047
DRD3 0.000
DRD4 0.348
DRD5 0.011
SLC6A3 1.786
GABBR1 3.028
GABBR2 13.94
GABRA1 0.000
GABRA2 0.034
GABRA3 0.000
GABRA4 0.012
GABRA5 0.105
GABRA6 0.000
GABRB1 0.219
GABRB2 0.000
GABRB3 0.143
GABRD 0.075
GABRE 0.074
GABRG1 0.007
GABRG2 0.000
GABRG3 0.000
GABRP 0.022
GABRQ 2.008
GABRR1 0.092
GABRR2 0.327
GABRR3 0.037
SLC32A1 0.674
SLC6A1 0.026
SLC6A11 0.013
SLC6A13 5.205
GRIA1 0.400
GRIA2 0.035
GRIA3 0.015
GRIA4 0.087
GRIK1 0.021
GRIK1-AS1 0.264
GRIK2 24.88
GRIK3 0.079
GRIK4 5.994
GRIK5 0.288
GRIN1 0.385
GRIN2A 0.002
GRIN2B 0.055
GRIN2C 0.088
GRIN2D 0.168
GRIN3A 0.009
GRIN3B 0.177
GRM1 0.050
GRM2 5.234
GRM3 0.000
GRM4 4.025
GRM5 0.002
GRM6 0.046
GRM7 0.010
GRM8 0.012
SLC1A1 1.315
SLC1A2 0.863
SLC1A3 0.681
SLC1A6 0.022
SLC1A7 0.719
SLC17A6 0.004
SLC17A7 12.23
SLC17A8 0.041
GLRA1 0.034
GLRA2 0.017
GLRA3 0.005
GLRA4 0.075
GLRB 1.838
SLC6A5 0.006
SLC6A9 3.504
HTN1 0.000
HTN3 0.000
HTR1A 0.000
HTR1B 0.140
HTR1D 0.148
HTR1E 0.000
HTR1F 0.000
HTR2A 0.066
HTR2B 0.072
HTR2C 0.000
HTR3A 0.012
HTR3B 0.020
HTR3C 0.000
HTR3D 0.000
HTR3E 0.020
HTR4 0.000
HTR5A 0.000
HTR6 0.000
HTR7 3.279
SLC6A4 0.044

Supplementary Figure 2: Gene Ontology and Specific Neuron Identity, related to Figure 2

(A) The five most significant gene ontology terms that represent the genes sets immediately repressed by NFD treatment at 0.5 DPT. The significances of each functional classification at 0.5, 1, and 2 DPT are compared to represent the progressive or transient enrichment of individual ontologies. The five most significant gene ontology terms representing the genes sets upregulated by NFD treatment at 2 DPT.

(B) Heatmap representing the change in expression of subtype-specific neuron genes detected by RNA-seq relative to GFP-transduced control fibroblasts. Genes exhibiting at least 2-fold enrichment, $\log_2$(NFD FPKM) ≥ 1, and p value ≤ 0.05 are labeled red. (BFCN, basal forebrain cholinergic neuron; FD, fibroblasts exposed to FD; FPKM, fragments per kilobase per million reads for NFD 2 DPT RNA-seq replicates; N, NEUROG2-transduced fibroblasts; NFD, *NEUROG2*-transduced fibroblasts exposed to FD).


A
Reprogramming Efficiency
100
80
60
40
20
0
Forskolin
dcAMP
cAMP
B
GFP TUBB3
E
Reprogramming Efficiency
(MAP2+GFP+/GFP+)
100
75
50
25
0
N.D.
NFD
N110-228
N110-228FD
F
NEUROG2 Interactors
3
15
TCF3
TCF4
5
C
Phospho-Deficient NEUROG2 Constructs
S24A GFP TUBB3
S91A GFP TUBB3
S193A GFP TUBB3
S207A GFP TUBB3
S209A GFP TUBB3
S219A GFP TUBB3
S232A GFP TUBB3
S239A GFP TUBB3
S242A GFP TUBB3
Multi-Site GFP TUBB3
D
Phosphomimetic NEUROG2 Constructs
S24D GFP TUBB3
S91D GFP TUBB3
S193D GFP TUBB3
S207D GFP TUBB3
S209D GFP TUBB3
S219D GFP TUBB3
S232D GFP TUBB3
S239D GFP TUBB3
S242D GFP TUBB3
Multi-Site GFP TUBB3
G
PRKACA Phosphorylation
TCF3 (E12)
TCF3 (E47)
TCF4
TCF12
OE CIP PKA
OE CIP PKA
OE CIP PKA
OE CIP PKA

Supplementary Figure 3: PRKACA Phosphorylates Co-Factors Essential to NEUROG2 Function, related to Figure 3.

(A) Reprogramming efficiency for *NEUROG2*-transduced cells treated with dorsomorphin and cAMP synthesis activators. Efficiency was calculated as (GFP-TUBB3-positive cells) / (total GFP-positive cell population).

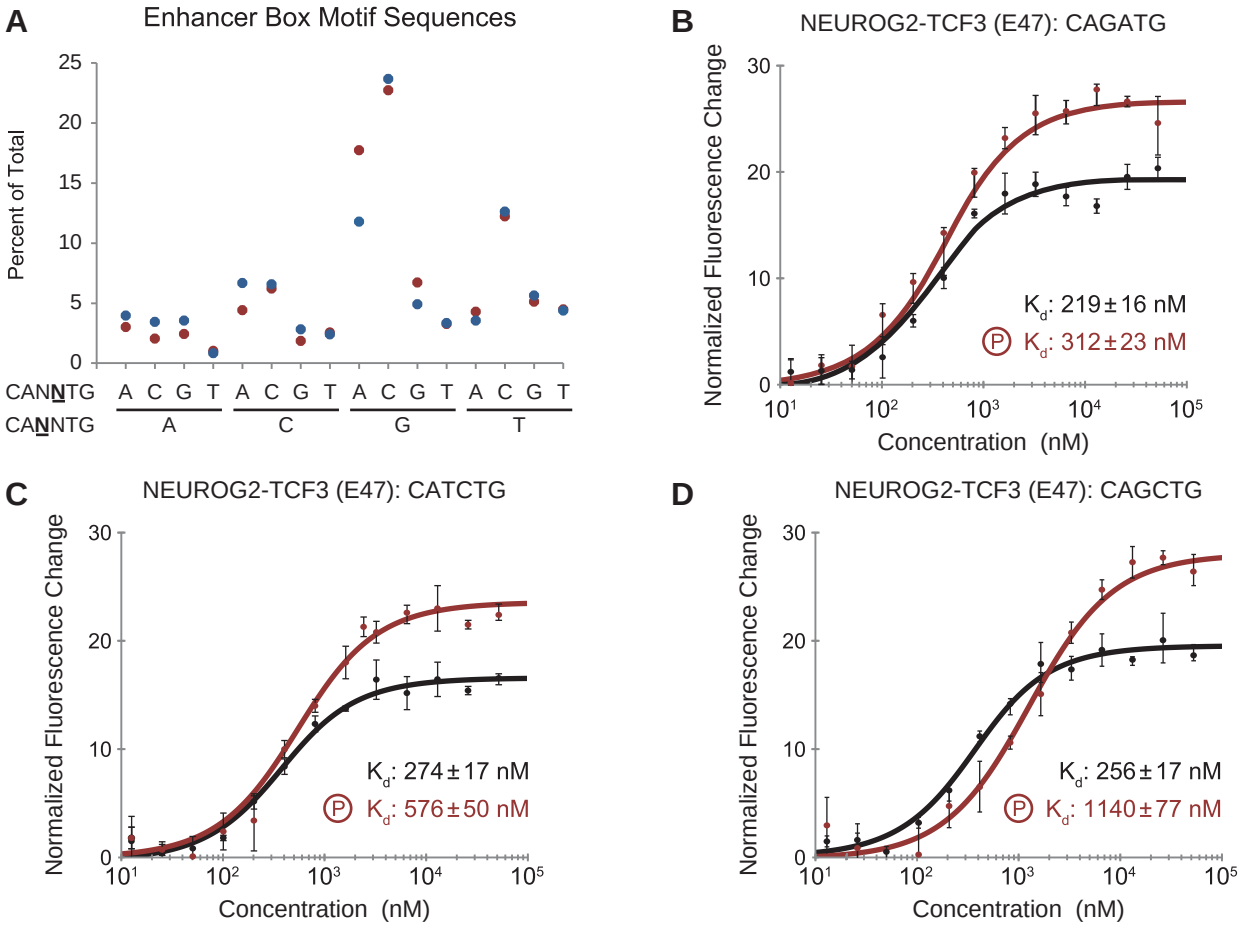
(B) Representative TUBB3 staining of fibroblast-derived neurons 12 days after lentiviral delivery of NEUROG2 and caPRKACA, scale: 25 μm.

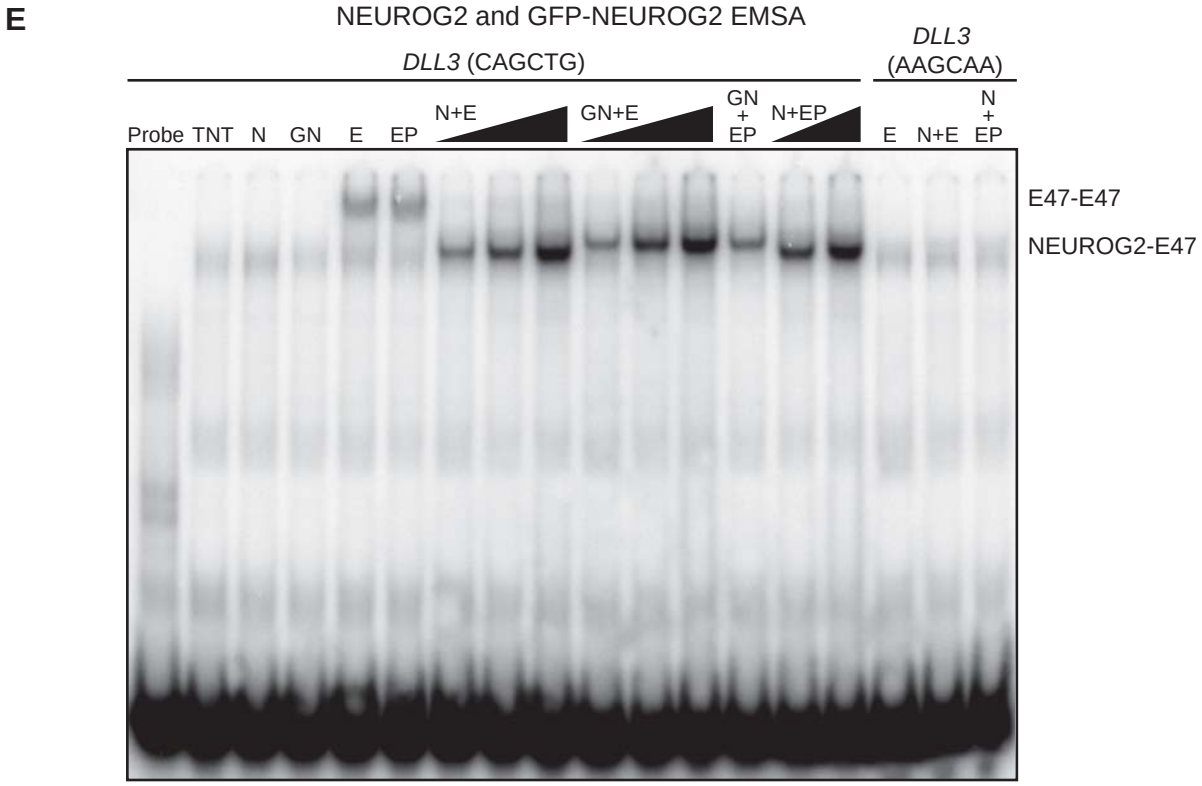
(C,D) Site-directed mutagenesis of NEUROG2 serine residues computationally identified as potential phosphorylation sites. (C) Alanine-substitution mimicked permanent dephosphorylation and (D) aspartate-substitution mimicked constitutive phosphorylation. Multi-site designates a NEUROG2 construct containing substitutions at S193, S207, S209, S219, S232, S239, and S242. None of the substitutions repressed the ability of NEUROG2 to induce TUBB3- and MAP2-positive neurons from fibroblasts at 14 DPI, scale: 50 μm.

(E) Reprogramming efficiency for $NEUROG2_{110\text{-}228}$-transduced fibroblasts. Efficiency was calculated as (GFP-MAP2-positive cells) / (total GFP-positive cell population).

(F) NEUROG2-interacting proteins identified in the absence (blue) and presence (red) of FD by mass spectrometry.

(G) Analysis of PRKACA-mediated E protein phosphorylation by denaturing polyacrylamide gel electrophoresis. TCF3 (E12 isoform), TCF3 (E47 isoform), TCF4, and TCF12 were overexpressed (OE), co-overexpressed with PRKACA (PKA), and co-overexpressed with PRKACA then incubated with calf intestinal alkaline phosphatase (CIP).


A
Enhancer Box Motif Sequences
Percent of Total
25
20
15
10
5
0
CANNTG A C G T A C G T A C G T A C G T
CANNTG A C G T
B
NEUROG2-TCF3 (E47): CAGATG
Normalized Fluorescence Change
30
20
10
0
Kd: 219±16 nM
Ⓟ Kd: 312±23 nM
10¹ 10² 10³ 10⁴ 10⁵
Concentration (nM)
C
NEUROG2-TCF3 (E47): CATCTG
Normalized Fluorescence Change
Kd: 274±17 nM
Ⓟ Kd: 576±50 nM
Concentration (nM)
D
NEUROG2-TCF3 (E47): CAGCTG
Normalized Fluorescence Change
Kd: 256±17 nM
Ⓟ Kd: 1140±77 nM
Concentration (nM)


E
NEUROG2 and GFP-NEUROG2 EMSA
DLL3 (CAGCTG)
DLL3 (AAGCAA)
Probe TNT N GN E EP N+E GN+E GN + EP N+EP E N+E N + EP
E47-E47
NEUROG2-E47

Supplementary Figure 4: NEUROG2 Heterodimers Exhibit Sequence-Specific Chromatin Affinity, related to Figure 4

(A) The distribution of enhancer box motif sequences identified within NEUROG2-target sequences for the 100 most significantly upregulated genes (red, 284 binding events) and 100 most significantly downregulated genes (blue, 192 binding events).

(B-D) Thermophoretic binding curves that depict the effect of phosphorylation (red) on the affinity of GFP-NEUROG2 heterodimers for sequence-specific enhancer box motifs.

(E) EMSA demonstrating NEUROG2 and GFP-NEUROG2 binding at an enhancer box upstream of *DLL3* (chr19:39,988,668-39,988,693). Probe, ^{32}P-labeled double-stranded probe only; TNT, T7 reticulocyte lysate expressing control vector; N, recombinant NEUROG2; GN, recombinant GFP-NEUROG2; E, recombinant TCF3 isoform E47; EP, recombinant TCF3 isoform E47 incubated with PRKACA. Filled ramps represent increased recombinant protein per reaction.

Supplementary Table S2. List of FD-induced genes containing at least one cAMP responsive element within 5-kilobases of the target TSS, related to Figure 3.

Gene Names
ABTB2
ADAMTS6
AHI1
AKAP12
ATP1B1
AVPI1
C6orf176
CD36
CENPW
CGA
CHMP1B
CLASP2
COL8A1
CREB3L2
CREM
DGKD
DIRAS3
DUSP1
EDNRA
ESRRG
FAM196A
FAM5C
FNDC3A
FOS
FYN
GAB2
GEM
GPCPD1
HES1
HES4
HS3ST3A1
ID3
IL11
ITPRIP
JUP
KCNF1
KIAA1217
KISS1

KLF9
KLHL13
LTBP1
MAOA
MLF1
MYOCD
NEAT1
NHS
NR4A1
NR4A2
NR4A3
NRG1
OLFML2B
P4HA3
PAPPA
PDE4B
PDE7B
PON2
PPARGC1A
PRMT10
RAB3A
RELL1
RNF122
SCG2
SDCBP2
SGIP1
SGK1
SIK1
SIK2
SLC46A3
SLC6A15
SNAP25
SOX4
SPAG4
SSTR2
SYNM
TMEM100
TMEM198
TMTC4
TUBA4A
WSB1

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Abbreviations

°C	degree(s) Celsius
× g	relative centrifugal force
µg	microgram(s)
µl	microliter(s)
µm	micrometer(s)
µM	micromolar
5-azacitidine	4-amino-1-β-D-ribofuranosyl-1,3,5-triazin-2(1*H*)-one
ATAC-seq	assay for transposase-accessible chromatin using sequencing
BCL2L1	B-cell CLL/lymphoma 2 like 1
BDNF	brain-derived neurotrophic factor
BSA	bovine serum albumin
caPRKACA	constitutively active protein kinase A catalytic subunit
cm	centimeter(s)
CREB1	cyclic AMP responsive element binding protein 1
decitabine	4-amino-1-(2-deoxy-β-D-*erythro*-pentofuranosyl)-1,3,5-triazin-2(1*H*)-one
deoxycholic acid	(3α,5β,12α,20*R*)-3,12-Dihydroxycholan-24-oic acid
DLL3	delta-like 3
DMEM	Dulbecco's modified eagle medium
DNA	deoxyribonucleic acid
dorsomorphin	6-[4-(2-Piperidin-1-ylethoxy)phenyl]-3-pyridin-4-ylpyrazolo[1,5-a]pyrimidine
DTT	dithiothreitol
DZNep	(1*S*,2*R*,5*R*)-5-(4-Amino-1*H*-imidazo[4,5-*c*]pyridin-1-yl)-3-(hydroxymethyl)-3-cyclopentene-1,2-diol hydrochloride
EDTA	ethylenediaminetetraacetic acid
EPZ5676	9H-purin-6-amine,9-[5-deoxy-5-[[cis-3-[2-[6-(1,1-dimethylethyl)-1H-benzimidazol-2-yl]ethyl]cyclobutyl](1-methylethyl)amino]-β-D-ribofuranosyl]
FGF2	fibroblast growth factor 2
FK228	cyclo[(2*Z*)-2-amino-2-butenoyl-L-valyl-(3*S*,4*E*)-3-hydroxy-7-mercapto-4-heptenoyl-D-valyl-D-cysteinyl], cyclic (3-5) disulfide

formaldehyde	methanal
forskolin	7β-Acetoxy-8,13-epoxy-1α,6β,9α-trihydroxylabd-14-en-11-one
FPKM	fragments per kilobase of exon per million mapped fragments
G	gauge
GDNF	glial cell-derived neurotrophic factor
GFP	green fluorescent protein
GSK126	1-(S)-sec-butyl-N-((4,6-dimethyl-2-oxo-1,2-dihydropyridin-3-yl)methyl)-3-methyl-6-(6-(piperazin-1-yl)pyridin-3-yl)-1H-indole-4-carboxamide
GSK3	glycogen synthase kinase 3
HA	hemagglutinin tag
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
I-BET151	7-(3,5-dimethyl-4-isoxazolyl)-8-(methyloxy)-1-[(1R)-1-(2-pyridinyl)ethyl]-1,3-dihydro-2H-imidazo[4,5-c]quinolin-2-one
I-BET762	4H-[1,2,4]Triazolo[4,3-a][1,4]benzodiazepine-4-acetamide, 6-(4-chlorophenyl)-N-ethyl-8-methoxy-1-methyl-, (4S)
IGEPAL CA-630	octylphenoxypolyethoxyethanol
JQ1	(6*S*)-4-(4-Chlorophenyl)-2,3,9-trimethyl-6*H*-thieno[3,2-*f*][1,2,4]triazolo[4,3-*a*][1,4]diazepine-6-acetic acid 1,1-dimethylethyl ester
liproxstatin-1	N-[(3-chlorophenyl)methyl]-spiro[piperidine-4,2′(1′H)-quinoxalin]-3′-amine
ln	natural log
M	molar
MAP2	microtubule-associated protein 2
ml	milliliter(s)
MOPS	3-(*N*-morpholino)-propanesulfonic acid
NEUROD1	neuronal differentiation 1
NEUROD4	neuronal differentiation 4
NEUROG2	neurogenin 2
ng	nanogram(s)
NT-3	neurotrophin 3
PBS	phosphate buffered saline

PCR	polymerase chain reaction
PRKACA	protein kinase A catalytic subunit
RepSox	2-[3-(6-methyl-2-pyridinyl)-1H-pyrazol-4-yl]-1,5-naphthyridine
RG108	*N*-phthalyl-L-tryptophan
RNA	ribonucleic acid
rpm	rotations per minute
SAHA	*N*-hydroxy-*N*'-phenyloctanediamide
SDS	sodium lauryl sulfate
SGI1027	*N*-[4-[(2-Amino-6-methyl-4-pyrimidinyl)amino]phenyl]-4-(4-quinolinylamino)benzamide
shRNA	microRNA30-based short hairpin RNA
SMARCA4	SWI/SNF related actin dependent regulator of chromatin, subfamily a member 4
SMARCB1	SWI/SNF related actin dependent regulator of chromatin, subfamily b member 1
SMARCC2	SWI/SNF related actin dependent regulator of chromatin, subfamily c member 2
SOX4	sex determining region Y-box 4
SWI/SNF	switch/sucrose non-fermentable
TBST	tris-buffered saline and tween 20
TCF12	transcription factor 12
TCF3	transcription factor 3
TCF4	transcription factor 4
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
Triton X-100	polyethylene glycol p-(1,1,3,3-tetramethylbutyl)-phenyl ether
TUBB3	tubulin, beta 3 class III
Tween 20	polyoxyethylene (20) sorbitan monolaurate
UNC669	(2-(phenylamino)-1,4-phenylene)bis((4-(pyrrolidin-1-yl)piperidin-1-yl)methanone)
v/v	volume per volume
valproic acid	2-propylpentanoic acid
w/v	weight per volume
zebularine	1-β-D-ribofuranosyl-2(1*H*)-pyrimidinone

Cell culture, lentivirus production, and neuron induction

Cell culture

293T/17, human fibroblast, and human glioblastoma cell lines (Supplemental Table 2) were cultured in Dulbecco's modified eagle medium (DMEM, GE Healthcare, SH30243.01) supplemented with 10% (v/v) (293T/17 and glioblastoma cell lines) or 15% (v/v) (fibroblast cell lines) fetal bovine serum (Corning, 35-010-CV) and 1% (v/v) penicillin-streptomycin (GE Healthcare, SV30010). Cells were maintained at 37°C in a humidified incubator with 5% CO_2.

Lentivirus production

293T/17 cells were seeded at a density of 3×10^6 cells in a 10 cm polystyrene dish 16 hours prior to transient transfection. 293T/17 cultures were treated with fresh medium 45 minutes prior to transfection. Polyethylenimine (Polysciences, 23966) was used to transfect cells with third-generation lentiviral packaging, envelope, transfer, and expression vectors (Supplemental Table 2). Cultures were treated with fresh medium 16 hours post transfection and replication-deficient lentivirus was collected 24 and 48 hours following this medium change. Lentivirus-containing medium was syringe filtered through a 0.22 μm polyvinylidene fluoride membrane (EMD Millipore, SLGV033RS) and stored at 4°C.

MRC-5 transdifferentiation: NEUROG2, forskolin, and dorsomorphin

MRC-5 fibroblasts were seeded at a density of 0.8×10^6 cells in a 10 cm polystyrene dish or 2×10^4 cells on standard 24-well glass coverslips sequentially treated with

0.1% (w/v) gelatin (Bio-Rad Laboratories, 170-6537) for 10 minutes at 24°C and Matrigel Basement Membrane Matrix (BD Biosciences, 356234) diluted 200-fold in DMEM for 16 hours at 37°C. Fibroblasts were transduced with *NEUROG2*-encoding lentivirus 24 hours after plating. The optimal lentivirus titer was empirically determined as 4-fold dilution for 10 ml-capacity 10 cm dishes and 10-fold dilution for 1 ml-capacity 24-well plates. These dilutions consistently yielded ≥90% TUBB3- and MAP2-positive neurons. Transduced-fibroblast cultures were treated with fresh medium 24 hours post infection. Cultures were transitioned to neuron induction medium (DMEM, Ham's F12 nutrient mixture (GE Healthcare, SH30026.02), neurobasal medium (Life Technologies, 21103-049), N-2 supplement (Life Technologies, 17502-048), B-27 supplement (Life Technologies, 17504-044), and penicillin-streptomycin at 1 : 1 : 0.5 : 0.02 : 0.01 : 0.025) supplemented with 10 μM forskolin (Sigma Aldrich, F6886) and 1 μM dorsomorphin (Millipore, 171260) 48 hours post infection. Four days post infection half of the total volume of medium was removed and replaced by one full volume of medium (for example, 500 μl removed and 1 ml added). Six days post infection half of the total volume of medium was removed and replaced with half of the original volume of neuron induction medium supplemented with 10 μM forskolin (for example, 750 μl removed and 500 μl added). Neuron induction medium supplemented with 10 μM forskolin was then half changed every two days until replating (Liu et al. 2013). Immunocytochemical staining, gene expression profiling, and electrophysiological characterization of induced neurons was previously demonstrated (Liu et al. 2013). Neuron induction efficiency was calculated as the total number of cells expressing a neuron-specific marker (TUBB3 or

MAP2) with characteristic neuron morphology (rounded soma and neurites at least five times the soma length) relative to the total number of viable GFP-expressing cells.

MRC-5 transdifferentiation: forskolin replacement screen

MRC-5 fibroblasts were cultured, transduced with *NEUROG2*-encoding lentivirus, and treated with neuron induction medium along the above-described time course. Neuron induction medium was supplemented with 1 μM dorsomorphin and either 10 μM forskolin, 10 μM cyclic AMP (cAMP, Sigma Aldrich, A6885) or 10 μM dibutyryl cAMP (Sigma Aldrich, D-0627). TUBB3 immunocytochemical staining was performed in triplicate 12 days post infection and reprogramming efficiency quantified using 10 random fields from each replicate.

MRC-5 transdifferentiation: NEUROG2, caPRKACA, and dorsomorphin

A constitutively active form of the human protein kinase A catalytic subunit (*caPRKACA*) was generated by site-directed mutagenesis of the wild-type catalytic subunit (Addgene, 23495) to introduce H88Q and W197R amino acid substitutions (Orellana et al. 1992) (Supplemental Table 2). MRC-5 fibroblasts were cultured, transduced with a *NEUROG2*+*caPRKACA*-encoding lentivirus, and treated with neuron induction medium supplemented with 1 μM dorsomorphin along the above-described time course. Immunocytochemical staining for TUBB3 and MAP2 was performed to quantify reprogramming efficiency.

MRC-5 transdifferentiation: NEUROG2 phosphomutant screen

A multi-species alignment of the *NEUROG2* gene identified nine potential GSK3- or PRKACA-targeted phosphorylation sites. Phosphomimetic and phospho-deficient NEUROG2 constructs were generated for these sites by consecutive rounds of site-directed mutagenesis of wild-type *NEUROG2* to introduce phosphomimetic (S24D, S91D, S193D, S207D, S209D, S219D, S232D, S239D, S242D) and phospho-deficient (S24A, S91A, S193A, S207A, S209A, S219A, S232A, S239A, S242A) mutations (Supplemental Table 2). MRC-5 fibroblasts were cultured, transduced with phosphomutant *NEUROG2*-encoding lentivirus, and treated with neuron induction medium or neuron induction medium supplemented with 10 μM forskolin and 1 μM dorsomorphin along the above-described time course. Immunocytochemical staining for TUBB3 and MAP2 was performed to quantify reprogramming efficiency.

MRC-5 transdifferentiation: NEUROG2 deletion construct

A NEUROG2 deletion construct (*NEUROG2* coding sequence positions 328 to 651 or NEUROG2 residue positions 110 to 217) was cloned into pCSC-SP-PW-IRES/GFP using AgeI and PstI restriction sites. MRC-5 fibroblasts were cultured, transduced with *NEUROG2* deletion-encoding lentivirus (10-fold dilution for 1 ml-capacity 24-well plates), and treated with neuron induction medium with or without 10 μM forskolin and 1 μM dorsomorphin along the above-described time course. MAP2 immunocytochemical staining was performed in triplicate 14 days post infection and reprogramming efficiency quantified using three random fields per replicate.

Glioblastoma cell transdifferentiation: NEUROG2 and SOX4

Human U-251 glioblastoma cells were seeded at a density of 2×10^5 cells (U-251) on standard 24-well glass coverslips treated with 500 µl Matrigel Basement Membrane Matrix diluted 200-fold in DMEM for 24 hours at 37°C. Glioblastoma cells were transduced with *NEUROG2*- (5-fold dilution for 1 ml-capacity 24-well plates) or *NEUROG2*+*SOX4*-encoding lentivirus (3.3-fold dilution for 1 ml-capacity 24-well plates) 24 hours after plating. Transduced cultures were treated with fresh medium 24 hours post infection. Cultures were transitioned to neuron induction medium supplemented with 10 µM forskolin and 1 µM dorsomorphin 48 hours post infection. Four days post infection half of the total volume of medium was removed and replaced by one full volume of medium (for example, 500 µl removed and 1 ml added). Six days post infection half of the total volume of medium was removed and replaced with half of the original volume of medium (for example, 750 µl removed and 500 µl added). Neuron induction medium with supplements was then half changed every two days until analysis. Neuron induction efficiency was calculated as the total number of cells expressing a neuron-specific marker (TUBB3 or MAP2) with characteristic neuron morphology relative to the total number of viable GFP-expressing cells.

Adult fibroblast transdifferentiation: NEUROG2 and SOX4

Adult fibroblast lines (AG05811, AG09969, ND29563, and ND39027) were seeded at a density of 0.3×10^6 cells in a 6 cm polystyrene dish or 1×10^4 cells in a 48-well plate treated with 100 µl DMEM containing 500-fold diluted Matrigel Basement Membrane Matrix for 16 hours at 37°C. Fibroblasts were transduced with *NEUROG2*- or

NEUROG2+SOX4-encoding lentivirus 24 hours after plating. The optimal lentivirus titer was empirically determined for each cell line. Transduced-fibroblast cultures were treated with fresh medium 24 hours post infection. Cultures were transitioned to neuron induction medium supplemented with 10 μM forskolin, 1 μM dorsomorphin, and 20 ng/ml FGF2 48 hours post infection. Four days post infection half of the total volume of medium was removed and replaced by one full volume of medium (for example, 250 μl removed and 500 μl added). Six days post infection half of the total volume of medium was removed and replaced with half of the original volume of neuron induction medium containing supplements (for example, 375 μl removed and 250 μl added). Neuron induction medium with supplements was then half changed every two days until analysis of neuron induction efficiency. Efficiency was calculated as the total number of GFP- and TUBB3-expressing cells with rounded morphology and at least one neurite relative to the total number of viable GFP-expressing cells. Alternatively, neurons were replated 14 days post infection. Replated neurons were washed twice with 2 ml phosphate-buffered saline (PBS, GE Healthcare, SH30028.02), incubated in 1 ml of 10-fold diluted trypsin (GE Healthcare, SH30042.01) for 10 minutes at 37°C, and suspended in 6 ml 15% (v/v) fetal bovine serum-containing medium. Cells were transferred to a 10 cm polystyrene dish pre-treated with 6 ml of 0.1% gelatin for 10 minutes at room temperature and incubated for 30 minutes at 37°C. Medium was collected without disturbing adherent fibroblasts and neurons were collected by centrifugation (500 × g, 3 minutes, 23°C). Neurons were suspended in 100 μl neuron maturation medium (DMEM, Ham's F12 nutrient mixture, neurobasal medium, N-2 supplement, B-27 supplement, and penicillin-streptomycin at

1 : 1 : 0.5 : 0.02 : 0.01 : 0.025) supplemented with 5 µM forskolin, 20 ng/ml BDNF (PeproTech, 450-02), 20 ng/ml GDNF (PeproTech, 450-10), and 20 ng/ml NT-3 (PeproTech, 450-03). Neuron suspensions were seeded in 1 ml neuron maturation medium on standard 24-well glass coverslips treated with 500 µl coating solution (5 µg fibronectin (ThermoFisher Scientific, CB-40008A), 5 µg laminin (ThermoFisher Scientific, CB-40232), and Matrigel Basement Membrane Matrix diluted 500-fold in DMEM for 48 hours at 37°C then washed once with PBS before plating). Medium was half changed every 48 hours for the remainder of neuron culture.

Adult Fibroblast Transdifferentiation: NEUROG2 and SWI/SNF Factors

AG05811 adult skin fibroblasts were plated at a density of 1×10^4 cells in a 48-well plate treated with 100 µl DMEM containing 500-fold diluted Matrigel Basement Membrane Matrix, 1 µg fibronectin, and 1 µg laminin for 48 hours at 37°C. Fibroblasts were transduced with *NEUROG2*- and *SMARCA4*- or *NEUROG2*-, *SMARCA4*-*SMARCB1*-, and *SMARCC2*-encoding lentiviruses 24 hours after plating. Transduced-fibroblast cultures were treated with fresh medium 24 hours post infection then transitioned to neuron induction medium supplemented with 10 µM forskolin, 1 µM dorsomorphin, and 20 ng/ml FGF2 48 hours post infection. Reprogramming was performed in both the presence and absence of 20 ng/ml BDNF, 20 ng/ml GDNF, and 20 ng/ml NT-3. Neuron induction medium was changed along the above-described time course until TUBB3 immunocytochemical staining 10 days post infection.

Adult Fibroblast Transdifferentiation: NEUROG2 and Chemicals

AG05811 adult fibroblasts were seeded at a density of 1×10^4 cells in a 48-well plate treated with 100 µl DMEM containing 500-fold diluted Matrigel Basement Membrane Matrix, 1 µg fibronectin, and 1 µg laminin for 48 hours at 37°C. Fibroblasts were transduced with 25 µl *NEUROG2*-encoding lentivirus 24 hours after plating. Transduced-fibroblast cultures were treated with fresh medium 24 hours post infection then transitioned to neuron induction medium supplemented with 10 µM forskolin, 1 µM dorsomorphin, 20 ng/ml FGF2, and one or a combination of the following small molecules: 15 µM 5-azacitidine (Selleck Chemicals, L1700), 32 µM decitabine (Selleck Chemicals, L1700), 0.5 µM DZNep (Selleck Chemicals, L1700), 1 µM EPZ5676 (Selleck Chemicals, L1700), 1 µM FK228 (Selleck Chemicals, L1700), 2 µM GSK126 (EMD Millipore, 500580), 0.5 µM I-BET151 (Selleck Chemicals, L1700), 0.5 µM I-BET762 (Selleck Chemicals, L1700), 0.5 µM JQ1 (Selleck Chemicals, L1700), 1 µM liproxstatin-1 (Sigma Aldrich, SML1414), 25 µM RepSox (Selleck Chemicals, L1700), 50 µM RG108 (Selleck Chemicals, L1700), 2.5 µM SAHA (Selleck Chemicals, L1700), 2.5 µM SGI1027 (Selleck Chemicals, L1700), 10 µM UNC669 (Selleck Chemicals, L1700), 1 mM valproic acid (Sigma Aldrich, P4543) or 10 µM zebularine (Selleck Chemicals, L1700) 48 hours post infection. Four days post infection half of the total volume of medium was removed and replaced by one full volume of medium containing supplements. Six and eight days post infection half of the total volume of medium was removed and replaced with half of the original volume of medium containing only 10 µM forskolin, 1 µM dorsomorphin, and 20 ng/ml FGF2. Small molecule-enhanced reprogramming was performed in both the presence and absence of 20 ng/ml BDNF,

20 ng/ml GDNF, and 20 ng/ml NT-3. Neuron induction medium was changed along the above-described time course until immunocytochemical staining 10 days post infection. FK228 treatment was performed using 1 μM, 0.5 μM, 0.1 μM, and 0.01 μM concentrations for 7 hour and 24 hour exposure periods; however, none of these conditions induced neuron-like cells with both high-efficiency and high survival rates. High-efficiency conversion was achieved by co-expression of NEUROG2 and BCL2L1 in AG05811 fibroblasts prior to treatment with 0.5 μM FK228. Fibroblasts were exposed to FK228 for a 24 hour period ranging from the 2-3, 3-4 or 4-5 day window following initial infection. Neuron reprogramming was performed in the presence of 10 μM forskolin, 1 μM dorsomorphin, and 20 ng/ml FGF2, 20 ng/ml BDNF, 20 ng/ml GDNF, and 20 ng/ml NT-3.

Chromatin immunoprecipitation and next generation sequencing

Assay for transposase-accessible chromatin using sequencing

MRC-5 fibroblasts were cultured, transduced with *GFP*-encoding, *NEUROG2*-encoding or *NEUROG2*+*SOX4*-shRNA-encoding lentivirus, and treated with DMEM containing 15% fetal bovine serum, neuron induction medium or neuron induction medium supplemented with 10 μM forskolin and 1 μM dorsomorphin for 4 days, respectively. Approximately 5×10^6 cells were treated with Accutase cell detachment solution (Innovative Cell Technologies) for 3 minutes and collected in ice-cold PBS by centrifugation (525 × g, 5 minutes, 4°C). Cells were resuspended by repeated gentle pipetting in 2 ml ice-cold resuspension buffer (PBS containing 25 mM HEPES (pH 7.0) and 5 mM EDTA). GFP-expressing viable cells were isolated using fluorescence-based

sorting on a MoFlo platform (Beckman Coulter). Exactly 50,000 cells were sorted into lysis buffer (10 mM Tris (pH 7.5), 10 mM NaCl, 3 mM $MgCl_2$, and 0.1% (v/v) IGEPAL CA-630) at 4°C then collected by centrifugation (21,130 × g, 2 minutes, 4°C). The pellet was resuspended in 50 µl transposition mix (25 µl 2X TD buffer, 22.5 µl nuclease-free H_2O, and 2.5 µl Nextera Tn5 transposase (Illumina, FC-121-1030)) and incubated exactly 30 minutes at 37°C with gentle agitation (400 rpm). Chromatin was purified with a MinElute PCR purification kit (Qiagen, 28004) and twice eluted in 10 µl elution buffer. Amplification reactions of 20 µl eluted chromatin, 2.5 µl universal ATAC primer (25 µM, Supplemental Table), 2.5 µl barcoded ATAC primer (25 µM, Supplemental Table), and 25 µl NEBNext High-Fidelity 2X PCR Master Mix (New England Biolabs, M0541S) were amplified for 1 cycle (5 minutes, 72°C) followed by 5 cycles (10 seconds, 98°C; 30 seconds, 63°C; 1 minute, 72°C). To determine the number of additional amplification cycles required, quantitative real-time PCR using 5 µl ATAC library reaction, 2.5 µl NEBNext High-Fidelity 2X PCR Master Mix, 0.125 µl universal ATAC primer (25 µM), 0.125 µl barcoded ATAC primer (25 µM), and 2 µl 5X SYBR Green I (ThermoFisher Scientific, S7563) was performed for 1 cycle (30 seconds, 98°C) followed by 20 cycles (10 seconds, 98°C; 30 seconds, 63°C; 1 minute, 72°C). The appropriate number of additional amplification cycles was determined for each library (11 total cycles) and amplified libraries were purified with a MinElute PCR purification kit then eluted in 20 µl elution buffer. Right side size selection with 0.4X Agencourt AMPure XP beads (Beckman Coulter, A63881) was used to reduce fragments larger than 1,000 nucleotides and left side size selection with 1X Agencourt AMPure XP beads was used to eliminate fragments smaller than 150 nucleotides. Fragment length and size selection

were evaluated after library amplification using a high-sensitivity DNA analysis kit (Agilent Technologies, 5067-4626). Library quantification prior to flow cell loading was performed using bioanalyzer traces and a Quant-iT PicoGreen dsDNA assay kit (ThermoFisher Scientific, P11496).

Crosslinking chromatin immunoprecipitation

MRC-5 fibroblasts were cultured, transduced with lentivirus (*HA-NEUROG2*-encoding or *NEUROG2+HA-SOX4*-encoding), and treated with neuron induction medium or neuron induction medium supplemented with 10 μM forskolin and 1 μM dorsomorphin along the above-described time course. AG05811 fibroblasts were cultured, transduced with *NEUROG2+HA-SOX4*-encoding lentivirus, and treated with neuron induction medium supplemented with 10 μM forskolin, 1 μM dorsomorphin, and 20 ng/ml FGF2 along the above-described time course. Approximately 1×10^7 cells were treated with 16% (w/v) methanol-free formaldehyde (Thermo Scientific, 28908) at a final concentration of 1% (v/v) for 8 minutes at room temperature with gentle rotation following 0.5, 1, 2 or 4 days in neuron induction medium. Crosslinking was quenched with 1.375 M glycine at a final concentration of 0.125 M for 5 minutes at room temperature with gentle rotation. Cells were immediately placed on ice, washed twice with 10 ml ice-cold PBS, scraped from the dish surface, and collected by centrifugation (525 × g, 5 minutes, 4°C). Pelleted cells were resuspended in 5 ml lysis buffer (100 mM HEPES (pH 8.0), 85 mM KCl, 1% (v/v) IGEPAL CA-630, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail (Roche, 11873580001)) and incubated on ice for 20 minutes. Cells were Dounce homogenized (15 repetitions) on ice and nuclei

collected by centrifugation (525 × g, 5 minutes, 4°C). Pelleted nuclei were resuspended in 260 µl briefly-chilled shearing buffer (50 mM HEPES (pH 8.0), 10 mM EDTA (pH 8.0), 1% (w/v) SDS, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail) and transferred to a TPX microtube (Diagenode, C-30010010-50) on ice. Chromatin was sheared for 45 minutes (sonication for 30 seconds followed by a 30 second pause for 45 cycles, power setting: high) using a Bioruptor (Diagenode, B01010002) equipped with a 4°C refrigerated water bath. Sheared chromatin was purified by centrifugation (21,130 × g, 10 minutes, 4°C) to remove precipitated detergent and insoluble debris. Chromatin concentration was measured using a microvolume spectrophotometer (DeNovix, DS-11+) and 150 µg (MRC-5) or 200 µg (AG05811) was 5-fold diluted in ice-cold immunoprecipitation buffer (50 mM HEPES (pH 8.0), 20 mM NaCl, 1 mM EDTA (pH 8.0), 0.1% (v/v) Triton X-100, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail). CREB1-bound DNA fragments were immunoprecipitated with 5 µg rabbit anti-phospho-CREB1 monoclonal antibody (Cell Signaling Technology, 9198S), HA-NEUROG2-bound and HA-SOX4-bound DNA fragments were immunoprecipitated with 5 µg rabbit anti-HA polyclonal antibody (Abcam, AB9110), NEUROD1-bound DNA fragments were immunoprecipitated with 5 µg rabbit anti-NEUROD1 monoclonal antibody (Abcam, AB109224), and NEUROD4-bound DNA fragments were immunoprecipitated with 5 µg rabbit anti-NEUROD4 polyclonal antibody (Abcam, AB90484) for 18 hours at 4°C with gentle nutation (Labnet, S0500). One aliquot of 10 µg sheared chromatin was stored at -20°C as input. Another aliquot of 10 µg sheared chromatin was sequentially treated with 10 µg RNase A (Roche, 10109142001) for 1 hour at 37°C, 40 µg Proteinase K (ThermoFisher Scientific, BP1700-50) for

2 hours at 55°C, and 2 µl 5 M NaCl for 16 hours at 67°C. This sample was loaded to a 1% (w/v) agarose gel, electrophoresed at 100 volts for approximately 2 hours, and stained with ethidium bromide to confirm fragmentation within a range of 100-500 nucleotides. The following day, 100 µl Magnetic Protein G Dynabeads (Life Technologies, 10003D) were washed three times with 1 ml ice-cold PBS containing 0.2% (v/v) Tween 20 then transferred to a 1.5 ml microtube on a MagneSphere separation stand (Promega, Z5332) at 4°C. The wash solution was removed and the beads were resuspended in immunoprecipitation samples for 2 hours at 4°C with gentle nutation. The bead complexes were magnetically isolated, twice washed by resuspension in ice-cold immunoprecipitation buffer for 2 minutes, twice washed by resuspension in ice-cold wash buffer (100 mM Tris-HCl (pH 9.0), 500 mM LiCl, 1% (v/v) IGEPAL CA-630, 1% (w/v) deoxycholic acid, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail) for 1 minute, and finally resuspended in ice-cold high-salt wash buffer (wash buffer containing 150 mM NaCl) for transfer to a 1.5 ml microtube on ice followed by bead collection. Chromatin was eluted from collected beads by addition of 100 µl elution buffer (1% (w/v) SDS and 50 mM $NaHCO_3$) for 30 minutes at 27°C with vigorous agitation (1,400 rpm, Eppendorf Thermomixer). Beads were magnetically collected and eluted chromatin was transferred to a sterile 1.5 ml microtube. Input chromatin was thawed on ice and treated with 85 µl elution buffer for 30 minutes at 27°C. Input and immunoprecipitated chromatin were sequentially treated with 10 µg RNase A for 1 hour at 37°C, 80 µg Proteinase K for 2 hours at 55°C, and 15 µl 5 M NaCl for 18 hours at 67°C. Chromatin was purified with a QIAquick PCR purification kit (Qiagen, 28104). Each sample was mixed with 1.4 ml PB binding buffer, isolated on a

QIAquick column matrix by centrifugation (15,800 × g, 1 minute), twice washed with 750 µl PE wash buffer, and eluted with 50 µl elution buffer into 1.5 ml DNA LoBind microtubes (Eppendorf, 022431021). Enrichment of CREB1-, NEUROD1-, NEUROD4-, NEUROG2-, and SOX4-targeted chromatin sites was evaluated by quantitative real-time PCR. Chromatin concentration was determined using a Qubit fluorometer and Qubit dsDNA HS assay kit (Life Technologies, Q32850).

Crosslinking chromatin co-immunoprecipitation

MRC-5 fibroblasts were cultured, transduced with *NEUROG2*-encoding lentivirus, and treated with neuron induction medium supplemented with 10 µM forskolin and 1 µM dorsomorphin along the above-described time course. Chromatin was isolated from approximately 1×10^7 cells, sheared into 100-500 nucleotide fragments, immunoprecipitated with appropriate antibodies (Supplemental Table 2), captured, washed, and eluted as described for crosslinking chromatin immunoprecipitation experiments. Input (25 µg) and immunoprecipitation (20% of eluted volume) samples were treated with Laemmli denaturing buffer, electrophoresed on 4-15% Mini-PROTEAN TGX precast gels, transferred to Immobilon-P polyvinylidene difluoride membrane (EMD Millipore, IPVH00010), incubated with appropriate antibodies (Supplemental Table 2), and treated with enhanced chemiluminescence horseradish peroxidase substrate for detection (Life Technologies, 32106).

Flow cytometry and low-cell native chromatin immunoprecipitation

MRC-5 fibroblasts were cultured, transduced with *NEUROG2*-encoding lentivirus,

and treated with neuron induction medium or neuron induction medium supplemented with 10 μM forskolin and 1 μM dorsomorphin for 2 days. Approximately 1×10^7 cells were treated with Accutase cell detachment solution (Innovative Cell Technologies) for 3 minutes and collected in ice-cold PBS by centrifugation (525 × g, 5 minutes, 4°C). Cells were resuspended by repeated gentle pipetting in 4 ml ice-cold resuspension buffer (PBS containing 25 mM HEPES (pH 7.0) and 5 mM EDTA). GFP-expressing viable cells were isolated using fluorescence-based sorting on a MoFlo platform (Beckman Coulter). Approximately 500,000 cells were collected by centrifugation (525 × g, 5 minutes, 4°C) then washed with 1 ml ice-cold PBS for chromatin digestion (Gilfillan et al. 2012). Cells were resuspended in 95 μl Micrococcal Nuclease digestion buffer (New England Biolabs, B0247S) supplemented with 0.01 μg BSA (New England Biolabs, B9001S), 0.2% (v/v) Triton X-100, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail. Resuspended cells were treated with 100 gel units Micrococcal Nuclease (New England Biolabs, M0247S) at 37°C for 5 minutes. Cells were immediately transferred to ice, treated with 10 μl ice-cold quench buffer (100 mM HEPES (pH 8.0) and 55 mM EDTA), and sonicated in a 1.5 ml TPX microtube for 60 seconds (power setting: high) using a Bioruptor equipped with a 4°C refrigerated water bath. Chromatin was diluted with 110 μl ice-cold immunoprecipitation buffer (50 mM HEPES (pH 8.0), 40 mM NaCl, 5 mM EDTA (pH 8.0), 0.2% (v/v) Triton X-100, 0.2% (w/v) SDS, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail), cell debris removed by centrifugation (21,130 × g, 10 minutes, 4°C), and 100 μl of supernatant transferred into two 0.2 ml microtubes (VWR, 732-0547). Chromatin was incubated with 1 μg rabbit polyclonal histone H3 acetyl K27 (Abcam, ab4729) or mouse

monoclonal histone H3 tri-methyl K27 (Abcam, ab6002) antibody for 18 hours at 4°C with gentle nutation. The remaining 20 µl of supernatant was stored at -20°C as input. The following day, 20 µl Magnetic Protein G Dynabeads were washed three times with 200 µl ice-cold PBS containing 0.2% (v/v) Tween 20 on a MagneSphere separation stand at 4°C. The wash solution was removed and the beads were resuspended in 10 µl immunoprecipitation buffer then added to immunopreciptation samples for 2 hours at 4°C with gentle nutation. The bead complexes were magnetically isolated, twice washed by resuspension in 150 µl ice-cold immunoprecipitation buffer for 2 minutes, twice washed by resuspension in 150 µl ice-cold wash buffer (100 mM Tris-HCl (pH 9.0), 500 mM LiCl, 1% (v/v) IGEPAL CA-630, 1% (w/v) deoxycholic acid, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail) for 1 minute, and finally resuspended in 150 µl ice-cold high-salt wash buffer (wash buffer containing 150 mM NaCl) for transfer to a 1.5 ml microtube on ice followed by bead collection. Chromatin was eluted from collected beads in 50 µl elution buffer (1% (w/v) SDS and 50 mM $NaHCO_3$) by two sequential additions of 25 µl elution buffer for 30 minutes at 27°C with vigorous agitation (1,400 rpm). Beads were magnetically collected and eluted chromatin was transferred to a sterile 1.5 ml microtube. Input chromatin was thawed on ice and treated with 50 µl elution buffer for 30 minutes at 27°C. Input and immunoprecipitated chromatin were sequentially treated with 10 µg RNase A for 15 minutes at 37°C and 80 µg Proteinase K for 3 hours at 55°C. Chromatin was purified with a QIAquick PCR purification kit and concentration was determined using a Qubit fluorometer and Qubit dsDNA HS assay kit as described above.

Massively parallel DNA sequencing

ChIP-seq libraries were synthesized from 20 ng purified input chromatin and 3.5-10 ng purified immunoprecipitated chromatin using a NEBNext ChIP-Seq Library Prep Master Mix Set for Illumina (New England Biolabs, E6240S) with NEBNext Multiplex Oligos for Illumina (New England Biolabs, E7335S). Replicate libraries for each experimental condition and time point were prepared from independent immunoprecipitations. Condition-dependent replicates were not performed for ATAC-seq libraries. Single-end 50-base length sequencing reads were generated on an Illumina HiSeq 2500 System. Reads were aligned to the human reference sequence GRCh37/hg19 with the Bowtie algorithm (version 1.0.0) (Langmead et al. 2009). Peak calling, peak-gene annotation, motif discovery (parameter: 200 nucleotide window from peak center), and generation of heatmap matrices were performed using HOMER (version 4.7) (Heinz et al. 2010). Heatmaps were generated by hierarchical clustering using Cluster (version 3.0) (Eisen et al. 1998) and Java TreeView (version 1.1.6) (Saldanha et al. 2004). Gene ontology classification was performed using HOMER and Genomic Regions Enrichment of Annotations Tool (version 2.0.2) (McLean et al. 2010). The UCSC Genome Browser was used to visualize tag densities and multi-experiment datasets (Kent et al. 2002). Sequences under peaks annotated to the 100 greatest upregulated and downregulated genes were extracted from the UCSC Genome Browser for enhancer box motif compositional analysis (NEUROG2 peak height ≥4 in the NFD condition, genomic sequence was defined between two points where the peak height equals 2 units, peak assignment required a maximum distance of 100 kilobases from the annotated gene and a minimum of 15 kilobases from any neighboring gene

with exception for peaks within 10 kilobases of the annotated gene). Super-enhancer analysis was performed with HOMER using a 12.5 kilobase stitching window and NEUROG2 ChIP-seq and ATAC-seq normalized tag count cutoff of 10.1 (Whyte et al. 2013). Genome-wide positional analysis of histone 3 lysine 27 acetylation and tr-methylation reads was performed with NGSplot (Shen et al. 2014).

Quantitative real-time PCR

Enrichment of transcription factor-targeted chromatin sites was validated for immunoprecipitation experiments on a 7900HT Fast Real-Time PCR System (Applied Biosystems, 4329002) using FastStart Universal SYBR Green Master Mix (ROX) (Roche, 04913922001) and primers specific to published target genes containing enhancer box, high mobility group box or cAMP response element motifs, as well as, controls for non-targeted chromatin and chromatin shearing efficiency (Supplemental Table 2). DNA amplification was performed for 2 minutes at 50°C, 10 minutes at 95°C, 40 cycles of 14 seconds at 95°C, and 60 seconds at 58°C. Amplification specificity was confirmed using a heat dissociation protocol during the final cycle.

Electrophoretic mobility shift assay

PRKACA, *NEUROG2*, and N-terminal HA-tagged *TCF3* (E12 isoform, Harvard Medical School Plasmid Repository), *TCF3* (E47 isoform, Addgene, 16059), *TCF4* (Harvard Medical School Plasmid Repository), and *TCF12* (Harvard Medical School Plasmid Repository) were cloned into bacterial expression vectors. An in-frame GFP-flexible-linker-NEUROG2 fusion protein was generated in a bacterial expression

vector using a 22 amino acid linker peptide (Neuhold et al. 1993). Recombinant protein was generated using a TNT T7 Quick Coupled Transcription/Translation System (Promega, L1170). Single stranded oligos corresponding to *DLL3* (chr19:39,988,668-39,988,693) or *DLL3* with destroyed enhancer box (AAGCAA) were end-labeled using Titanium Taq DNA polymerase (Takara, 639242) with ^{32}P-dCTP (PerkinElmer, BLU513A250UC) and annealed for 5 minutes at 95°C then 68°C for 15 minutes. Double stranded oligos were purified using a 1 ml hand-packed column containing Sephadex G-25 resin (Sigma Aldrich, S5772) pre-soaked in TE buffer (10 mM Tris (pH 8.0) and 1 mM EDTA) for 24 hours at 4°C. Labeling efficiency was determined using a scintillation counter (Beckman). Recombinant proteins were incubated with 12,500 units PRKACA (New England Biolabs, P6000S) supplemented with protein kinases buffer (New England Biolabs, B6022S) and 200 µM adenosine 5'-triphosphate magnesium salt (Sigma Aldrich, A9187) for 45 minutes at 30°C. Binding reactions containing recombinant protein (1, 2 or 3 µl TCF3 isoform E47 and GFP-NEUROG2 or NEUROG2), 2 µl reaction buffer (100 mM Tris (pH 7.5), 500 mM NaCl, 10 mM DTT, 10 mM EDTA, 50% glycerol), 0.1 nM ^{32}P-labeled double stranded oligo, and 2 µg Poly(dI-dC) (Sigma Aldrich, P4929) were mixed, incubated at 30°C for 60 minutes, and electrophoresed on a 1.5-hour pre-run 5% native polyacrylamide gel at 4°C for 2 hours at 200 volts. Each polyacrylamide gel was transferred to 3 mm chromatography paper (GE Healthcare, 3030-6461) and vacuum dried for 1.5 hours at 80°C. Protein-DNA interactions were visualized following 30 minute to 12 hour phosphorscreen exposure using a STORM 820 phosphorimager (GE Healthcare).

Mass spectrometry

MRC-5 fibroblasts were cultured, transduced with *NEUROG2*-encoding lentivirus, and treated with neuron induction medium or neuron induction medium supplemented with 10 μM forskolin and 1 μM dorsomorphin along the above-described time course. Approximately 1.25×10^7 cells were placed on ice, washed twice with 10 ml ice-cold PBS, scraped from the dish surface into ice-cold PBS, and collected by centrifugation (525 × g, 5 minutes, 4°C). Cells were suspended in 500 μl ice-cold lysis buffer (100 mM Tris (pH 7.5), 100 mM KCl, 5 mM $MgCl_2$, 0.2% (v/v) Triton X-100, 0.1% (v/v) Tween 20, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail) and homogenized for 10 repetitions using a 25 G × 5/8 inch sterile PrecisionGlide needle (BD Biosciences, 305122). Cell lysate was treated with 1,000 units of Pierce Universal Nuclease (ThermoFisher Scientific, 88701) and incubated on ice for 30 minutes with 3-second bursts of vortexing every 10 minutes. Cell debris was removed by centrifugation (21,130 × g, 10 minutes, 4°C). In parallel, 100 μl Pierce Anti-HA magnetic beads (ThermoFisher Scientific, 88836) were washed with 1 ml ice-cold lysis buffer then transferred to a 1.5 ml microtube on a MagneSphere separation stand at 4°C. The wash solution was removed and the beads were resuspended in 500 μl cell lysate for 16 hours at 4°C with gentle nutation. The following day, bead complexes were magnetically isolated, washed three times by resuspension in ice-cold lysis buffer for 2 minutes, and protein eluted with 20 μl glycine (pH 2) at 27°C and 1,000 rpm shaking for 7 minutes. Eluate was transferred to a 1.5 ml microtube, neutralized with 4 μl 1 M Tris (pH 8), mixed with 24 μl Laemmli denaturing buffer (Sigma Aldrich, S3401), and incubated at 100°C for 5 minutes. Protein was electrophoresed 1 cm into a

10% Mini-PROTEAN TGX precast gel (Bio-Rad Laboratories, 456-1033) using NuPAGE MOPS SDS Running Buffer (Life Technologies, NP0001). Protein was visualized with blue silver stain (10% (v/v) phosphoric acid, 10% (w/v) ammonium sulfate, 0.12% (w/v) Brilliant blue G 250 (Sigma Aldrich, 27815), and 20% (v/v) anhydrous methanol) for 12 hours at room temperature with gentle rotation (Candiano et al. 2004). The gel was washed for 2 hours in HyPure Molecular Biology Grade Water (GE Healthcare, SH30538.02) then gel bands were removed by sterile excision and transferred into a 1.5 ml microfuge tube rinsed one time with 50% ethanol. Samples were reduced and alkylated with DTT (Sigma Aldrich, 9779) and iodoacetamide (Sigma Aldrich, I6125) then digested overnight with trypsin (Promega, V5280). Protein was purified by solid-phase extraction using an Oasis HLB plate (Waters, WAT058951) then analyzed by liquid chromatography-mass spectrometry-mass spectrometry using an Ultimate 3000 RSLC-Nano liquid chromatography system (Dionex) coupled to a Q Exactive mass spectrometer (Thermo Electron). Raw data files were converted to peak-list format and analyzed using the central proteomics facilities pipeline (version 2.0.3) (Trudgian et al. 2010, Trudgian et al. 2012). Peptide identification was performed using the X!Tandem (Craig et al. 2004) and open mass spectrometry search algorithm engines (Geer et al. 2004) against the Uniprot human protein database with common contaminants and reversed decoy sequences flagged (Elias et al. 2007). Fragment and precursor tolerances of 20 parts per million and 0.1 Daltons were specified with three missed cleavages permitted. Label-free quantitation of proteins was performed using SINQ normalized spectral index software (Trudgian et al. 2011). Three replicates from NEUROG2-transduced fibroblasts and two replicates from NFD-treated fibroblasts were

used for analysis. Classification as a NEUROG2 interactor protein required spectral counts ≥4 in at least two replicates of one condition with the ratio of normalized spectral counts indicating no detection in the untreated fibroblast control sample. Interactors were assigned as unique if undetected in any replicate of the opposite condition.

Microscale thermophoresis

GFP-NEUROG2 fusion and N-terminal HA-tagged TCF3 (E12 isoform), TCF3 (E47 isoform), TCF4, and TCF12 recombinant proteins were prepared as described above (see electrophoretic mobility shift assay). Single stranded oligos corresponding to *DLL3* (chr19:39,988,668-39,988,693) or *DLL3* with mutated enhancer box (CAGATG or CATCTG) were annealed in thermophoresis buffer (25 mM HEPES (pH 7.3), 50 mM NaCl, 2.5 mM $MgCl_2$, 0.025% IGEPAL CA-630, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail) using a 98°C to 4°C gradient at a cooling rate of 0.5°C per 20 seconds. GFP-NEUROG2 was mixed at a 1:1 (v/v) ratio with each E protein and incubated at 30°C for 30 minutes. In parallel, 1:1 (v/v) NEUROG2-E protein mixtures were incubated with 12,500 units PRKACA in the presence of 200 μM adenosine 5'-triphosphate magnesium salt for 30 minutes at 30°C. Recombinant protein heterodimer mixtures were added to double stranded oligos (1:7 (v/v) dilution) in a titration series from 208 μM to 6 nM. Binding reactions were loaded into standard treatment capillaries (NanoTemper Technologies, MO-K002) and thermophoretic mobility quantified using a Monolith NT.115 Platform (NanoTemper Technologies). Thermophoretic traces were collected from three positions in each capillary using 100% light emitting diode power, 40% infrared laser power, and standard thermophoresis

program (fluorescence before: 5 seconds, laser on: 30 seconds, fluorescence after: 5 seconds, delay: 25 seconds). NTControl (version 2.2.1) and PALMIST (Chad Brautigam, Ph.D., The University of Texas Southwestern Medical Center) were used to derive average binding curves and apparent K_d values.

PRKACA phosphorylation assay

Escherichia coli Rosetta (DE3) cells (EMD Millipore, 70954) were transformed with *GFP-NEUROG2*, *NEUROG2*, *HA-TCF3* (isoform E12), *HA-TCF3* (isoform E47), *HA-TCF4*, and *HA-TCF12* constructs independently or co-transformed with *PRKACA*. Single colonies were used to inoculate 2 ml Luria broth containing 100 µg/ml ampicillin (Fisher Scientific, BP1760-25). Co-transfected cells were dual-selected with 25 µg/ml kanamycin (Fisher Scientific, BP906-5). Cultures were grown 16 hours at 37°C with shaking (225 rpm) then 500 µl was transferred to 50 ml antibiotic-containing Luria broth and grown to an optical density$_{600}$ of 0.6. Protein overexpression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (Fisher Scientific, BP1620-10) for 16 hours at 30°C with shaking (225 rpm). Protein-expressing cells were collected by centrifugation (525 × g, 5 minutes, 4°C), lysed in 250 µl ice-cold lysis buffer (50 mM Tris (pH 7.4), 150 mM NaCl, 1% (v/v) Triton X-100, 1% (w/v) deoxycholic acid, 0.1% (w/v) SDS, and 1% (v/v) cOmplete EDTA-free Protease Inhibitor Cocktail), and incubated on ice for 30 minutes with 3 second vortex bursts every 10 minutes. Lysed cell debris was removed by centrifugation (21,130 × g, 10 minutes, 4°C) and protein concentration quantified by colorimetric assay (Bio-Rad Laboratories, 500-0006). Protein extracts were treated with calf intestinal alkaline phosphatase (New England BioLabs, M0290S)

for 2 hours at 37°C. Protein extracts were treated with Laemmli denaturing buffer for 5 minutes at 100°C then electrophoresed on an 8% denaturing polyacrylamide gel with and without 50 μM Phos-tag reagent (NARD Institute, 300-93523) and 100 μM $MnCl_2$ in NuPAGE MOPS SDS Running Buffer. Each gel was rinsed two times with 10 ml transfer buffer (250 mM Tris, 2 M glycine, 20% (v/v) methanol) then consecutively incubated two times in 15 ml transfer buffer containing 10 mM EDTA for 10 minutes with gentle rotation followed by one 15 minute incubation in transfer buffer without EDTA. Protein was transferred to Immobilon-P polyvinylidene difluoride membrane, blocked with 5% (w/v) non-fat dry milk solution (5% (w/v) non-fat dry milk dissolved in TBST (100 mM Tris (pH 7.4), 150 mM NaCl, and 0.1% (v/v) Tween 20) for 1 hour at room temperature, incubated with anti-HA polyclonal antibody (Supplemental Table 2) for 18 hours at 4°C with gentle rotation, washed five times with TBST for 5 minutes per wash, incubated with secondary antibody for 1 hour at room temperature with gentle rotation, washed three times with TBST for 5 minutes per wash, and treated with enhanced chemiluminescence horseradish peroxidase substrate for detection.

Reverse transcription and quantitative real-time PCR

Total RNA was isolated from cells using TRIzol (Life Technologies, 15596-018) followed by chloroform extraction and column-based purification (Zymo Research, R1016). Reserve transcription was performed using 1.5 μg purified RNA and SuperScript III Reverse Transcriptase (Invitrogen, 18080-093) with random primers. Amplification was performed for 2 hours at 42 °C then 15 minutes at 72°C. Samples

were cooled to 4°C then analyzed by quantitative real-time PCR using target-specific primers (Supplemental Table 2) as described above.

RNA sequencing

Total RNA was isolated in triplicate from fibroblasts transduced with *GFP*-encoding lentivirus (control), fibroblasts exposed to forskolin and dorsomorphin, fibroblasts transduced with *NEUROG2*-encoding lentivirus, and fibroblasts both transduced with *NEUROG2*-encoding lentivirus and exposed to forskolin and dorsomorphin. RNA isolation was performed 0.5, 1, and 2 days following conversion to neuron induction medium using TRIzol (Life Technologies, 15596-018) followed by chloroform extraction and column-based purification (Zymo Research, R1016). Libraries were synthesized from 4 µg purified RNA using a TruSeq Stranded Total RNA Sample Preparation Kit (Illumina, RS-122-2201). Single-end 50-base length sequencing reads were generated on an Illumina HiSeq 2500 System. Reads were aligned to the human reference sequence GRCh37/hg19 with the TopHat (version 1.4.1) algorithm and transcript assembly performed with Cufflinks (version 2.1.0) (Trapnell et al. 2012). Expression levels of RefSeq-annotated genes were calculated in units of fragments per kilobase of exon per million mapped fragments (FPKM) and differential expression analysis was performed with Cuffdiff (version 2.1.0). Genes were defined as significant if the following three criteria were satisfied: Student's t-test p value ≥0.05, at least 2-fold gene expression change relative to the GFP-transduced control, and the triplicate average of ln(FPKM) ≥1 in either the experimental condition or fibroblast control. Gene ontology

analysis was performed using the DAVID Functional Annotation Tool (version 6.7) (Huang et al. 2009a, Huang et al. 2009b).

shRNA-mediated gene knockdown

Multiple rounds of PCR were used to introduce miR30 regulatory sequences at the 3' end of a *NEUROG2-IRES-GFP* lentiviral construct. Two restriction sites, XhoI and PstI, were engineered into these regulatory sequences at optimized processing locations to permit efficient generation of shRNAs targeting *CREB1* and *SOX4* (Supplemental Table 2) (Fellmann et al. 2013). MRC-5 fibroblasts were cultured, transduced with *NEUROG2*+shRNA-encoding lentivirus, and treated with neuron induction medium supplemented with 10 μM forskolin and 1 μM dorsomorphin along the above-described time course. Approximately 6×10^5 transduced cells were scraped from the dish surface in ice-cold PBS and collected by centrifugation (525 × g, 5 minutes, 4°C). Cells were suspended in 75 μl ice-cold lysis buffer (50 mM Tris (pH 7.4), 150 mM NaCl, 1% (v/v) Triton X-100, 1% (w/v) deoxycholic acid, and 0.1% (w/v) SDS) and incubated on ice for 30 minutes with 3 second vortex bursts every 10 minutes. Lysed cell debris was removed by centrifugation (21,130 × g, 10 minutes, 4°C) and protein concentration quantified by colorimetric assay. Protein extracts (25 μg) were treated with Laemmli denaturing buffer for 5 minutes at 100°C then electrophoresed on 4-15% Mini-PROTEAN TGX precast gels (Bio-Rad Laboratories, 456-1083) and transferred to Immobilon-P polyvinylidene difluoride membrane. Membranes were blocked with 5% (w/v) non-fat dry milk or 5% (w/v) bovine serum albumin dissolved in TBST for 1 hour at room temperature, treated with target-specific antibodies (Supplemental Table 2) for 18

hours at 4°C with gentle rotation, washed five times with TBST for 5 minutes per wash, treated with corresponding secondary antibodies (Supplemental Table 2) for 1.5 hours at room temperature with gentle rotation, washed three times with TBST for 5 minutes per wash, and treated with enhanced chemiluminescence horseradish peroxidase substrate for detection. In addition, MAP2 immunocytochemical staining was performed in triplicate 12 days post infection and reprogramming efficiency quantified using 10 random fields from each replicate.

References

Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G.M., Carnemolla, B., Orecchia, P., Zardi, L., and Righetti, P.G. (2004) Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. Electrophoresis *25*, 1327-1333.

Craig, R., and Beavis, R.C. (2004) TANDEM: matching proteins with tandem mass spectra. Bioinformatics *20*, 1466-1467.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. U.S.A. *95*, 14863-14868.

Elias, J.E., and Gygi, S.P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods *4*, 207-214.

Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhar, M., Roth, M., Lai, D.Y., Barbosa, I.A.M., Kwon, J.S., Guan, Y., *et al.* (2013) An optimized microRNA backbone for effective single-copy RNAi. Cell Rep. *5*, 1704-1713.

Geer, L.Y., Markey, S.P., Kowalak, J.A., Wagner, L., Xu, M., Maynard, D.M., Yang, X., Shi, W., and Bryant, S.H. (2004) Open mass spectrometry search algorithm. J. Proteome Res. *3*, 958-964.

Gilfillan, G.D., Hughes, T., Sheng, Y., Hjorthaug, H.S., Straub, T., Gervin, K., Harris, J.R., Undlien, D.E., and Lyle, R. (2012) Limitations and possibilities of low cell number ChIP-seq. BMC Genomics *12*, 645.

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.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nat. Protoc. *4*, 44-57.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene list. Nucleic Acids Res. *37*, 1-13.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002) The human genome browser at UCSC. Genome Res. *12*, 996-1006.

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.

Liu, M.-L., Zang, T., Zou, Y., Chang, J.C., Gibson, J.R., Huber, K.M., and Zhang, C.-L. (2013) Small molecules enable Neurogenin 2 to efficiently convert human fibroblasts into cholinergic neurons. Nat. Commun. *4*, 2183.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010) GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495-501.

Neuhold, L.A., and Wold, B. (1993) HLH forced dimers: tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. Cell *74*, 1033-1042.

Orellana, S.A., and McKnight, G.S. (1992) Mutations in the catalytic subunit of cAMP-dependent protein kinase result in unregulated biological activity. Proc. Natl. Acad. Sci. U.S.A. *89*, 4726-4730.

Saldanha, A.J. (2004) Java Treeview-extensible visualization of microarray data. Bioinformatics *20*, 3246-3248.

Shen, L., Shao, N., Liu, X., and Nestler, E. (2014) ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics *15*, 284.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. *7*, 562-578.

Trudgian, D.C., Thomas, B., McGowan, S.J., Kessler, B.M., Salek, M., Acuto, O. (2010) CPFP: a central proteomics facilities pipeline. Bioinformatics *26*, 1131-1132.

Trudgian, D.C., Ridlova, G., Fischer, R., Mackeen, M.M., Ternette, N., Acuto, O., Kessler, B.M., and Thomas, B. (2011) Comparative evaluation of label-free SNIQ

normalized spectral index quantitation in the central proteomics facilities pipeline. Proteomics *11*, 2790-2797.

Trudgian, D.C., and Mirzaei, H. (2012) Cloud CPFP: a shotgun proteomics data analysis pipeline using cloud and high performance computing. J. Proteome Res. *11*, 6282-6290.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307-319.