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

Derek K. Smith,1,2 Jianjing Yang,1,2 Meng-Lu Liu,1,2 and Chun-Li Zhang1,2,*

1Department of Molecular Biology
2Hamon Center for Regenerative Science and Medicine
The University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard, Dallas, TX 75390-9148, USA
*Correspondence: chun-li.zhang@utsouthwestern.edu
http://dx.doi.org/10.1016/j.stemcr.2016.09.013

SUMMARY

Pro-neural transcription factors and small molecules can induce the reprogramming of fibroblasts into functional neurons; however, the immediate-early molecular events that catalyze this conversion have not been well defined. We previously demonstrated that neurogenin 2 (NEUROG2), forskolin (F), and dorsomorphin (D) can reprogram fibroblasts into functional neurons with high efficiency. Here, we used this model to define the genetic and epigenetic events that initiate an acquisition of neuronal identity. We demonstrate that NEUROG2 is a pioneer factor, FD enhances chromatin accessibility and H3K27 acetylation, and synergistic transcription activated by these factors is essential to successful reprogramming. CREB1 promotes neuron survival and acts with NEUROG2 to upregulate SOX4, which co-activates NEUROD1 and NEUROD4. In addition, SOX4 targets SWI/SNF subunits and SOX4 knockdown results in extensive loss of open chromatin and abolishes reprogramming. Applying these insights, adult human glioblastoma cell and skin fibroblast reprogramming can be improved using SOX4 or chromatin-modifying chemicals.

INTRODUCTION

Fibroblast-to-neuron reprogramming has been demonstrated using transcription factors (Vierbuchen et al., 2010; Pang et al., 2011) and signaling molecules (Hu et al., 2015; Li et al., 2015). These factors jump-start transcriptional programs that redefine cellular identity; however, the potency of each reprogramming factor is cell age and lineage dependent (Liu et al., 2013; Masserdotti et al., 2015; Mertens et al., 2015). Therefore, mechanistically, the success of neuronal reprogramming depends upon the cell-of-origin genetic and epigenetic environment, as well as the ability of reprogramming factors to navigate that environment and induce reprogramming. For instance, the pioneer activity of ASCL1 and the permissible chromatin structures of mouse embryonic fibroblasts facilitate rapid neuronal reprogramming (Wapinski et al., 2013). Likewise, direct competition for pro-neural genetic elements between endogenous repressor complexes and overexpressed neurogenin 2 (NEUROG2) regulates reprogramming of mouse astroglia in short versus prolonged culture (Masserdotti et al., 2015).

NEUROG2 is a basic helix-loop-helix transcription factor that promotes early neurogenesis (Ma et al., 1996; Fode et al., 1998; Scardigli et al., 2001). NEUROG2 overexpression in mouse embryonic stem cells and cultured mouse cortical astroglia catalyzes conversion into functional, synapse-forming glutamatergic neurons (Berninger et al., 2007; Heinrich et al., 2010; Thoma et al., 2012). While NEUROG2 is sufficient to induce reprogramming in neural lineage cells, this factor is unable to independently reprogram somatic fibroblasts (Liu et al., 2013; Chanda et al., 2014). The high-efficiency reprogramming of human fetal fibroblasts (MRC-5) into functional cholinergic neurons by NEUROG2 requires simultaneous exposure to the small molecules forskolin, a cyclic AMP (cAMP) synthesis activator, and dorsomorphin, an inhibitor of AMP-activated protein kinase and bone morphogenetic protein type 1 receptors (Liu et al., 2013). Furthermore, the inclusion of SOX11 is required to rapidly convert adult human skin fibroblasts into neurons (Liu et al., 2013). These findings demonstrate that age- and lineage-specific genetic and epigenetic factors directly affect the ability of NEUROG2 to catalyze reprogramming.

Here, we perform a systematic analysis of the individual and synergistic actions of NEUROG2 and small molecules during reprogramming. We define critical genetic interactions, hierarchical transcription programs, epigenetic histone signatures, and open chromatin conformations that drive the immediate-early stage of reprogramming, then apply these discoveries to adult human skin and glioblastoma cells refractory to reprogramming.

RESULTS

Small Molecules Synergize with NEUROG2 to Enhance Chromatin Occupancy and Redefine the Transcriptional Profile of Fibroblasts

MRC-5 fibroblasts transduced with NEUROG2-encoding lentivirus exhibit limited capacity for neuronal reprogramming; however, exposure to forskolin and dorsomorphin (FD) triggers adoption of class III β-tubulin (TUBB3)* and microtubule-associated protein 2 (MAP2)* neuronal
identity with greater than 90% efficiency in 14 days (Liu et al., 2013). To investigate the genetic mechanisms that catalyze reprogramming, we performed chromatin immunoprecipitation-sequencing (ChIP-seq) and RNA-sequencing (RNA-seq) analyses at 0.5, 1, and 2 days post FD treatment (DPT) (Figure 1A).

A ChIP-seq time-course analysis was performed to define the ordered genomic targets of NEUROG2 and identify whether FD significantly modify occupancy at these sites. NEUROG2 was immunoprecipitated from MRC-5 in both the absence and presence of FD at 0.5, 1, and 2 DPT. Replicates exhibited strong enrichment of pro-neural genetic targets, as well as the enhancer box motif (5'-CANNTG-3'). A condition-dependent comparison of NEUROG2 occupancy profiles revealed that FD both enhance NEUROG2 binding intensity at shared sites and induce binding at novel targets undetected in the absence of FD (Figure 1B). NEUROG2 was detected at up to 4-fold higher levels for both time-stable and dynamic binding events following FD treatment, indicating that FD affect NEUROG2 chromatin affinity or accessibility (Figure S1A).
The distribution of NEUROG2-bound sites is biased toward intergenic and intronic regions (Figure S1B). Although FD modestly enhanced promoter and transcription start site (TSS) occupancy, NEUROG2 seems to dominantly regulate transcription via distal regulatory elements. To define how occupancy at primary and FD-enhanced sites affects the transcription, we isolated RNA from FD-treated, NEUROG2-transduced, and NEUROG2-transduced FD-treated (NFD) fibroblasts at 0.5, 1, and 2 DPT. Sequence reads were normalized to time-point-specific GFP-transduced MRC-5 for differential expression analysis (Table S1).

NEUROG2 immediately induced an upward shift in the expression of neural genes such as DLL3 and HES6. While FD independently activated relatively few neural-specific genes, these small molecules did significantly upregulate a small subset of diverse genes including C6orf176, C11orf96, DNER, NR4A1, NR4A2, NRG1, and PPARGC1A (Figure 1C). In combination with NEUROG2, FD induced the expression of hundreds of genes unaltered by individual factors (Figure 1D). A temporal analysis of gene expression underscores the rapid rate of change induced by NFD (Figure 1E). Synergistic regulation accounted for 53% of the total change in gene expression at 2 DPT and the total number of genes exhibiting ≥5-fold change increased from 4% to 10% within the 36-hr window from 0.5 to 2 DPT. The synergistic induction of NEUROD1, NEUROD4, and SOX4, as well as ontological classification of immediately repressed and early-induced gene sets, highlight the functional significance of these targets (Figure S2A). Furthermore, upregulation of the basal forebrain markers NTRK1 and SST2 as well as the cholinergic receptor CHRNA3 and vesicular transporter SLC18A3 suggests that these early neurons are initially biased toward a cholinergic fate (Figure S2B).

To define how NEUROG2 binding correlates to the changes detected by RNA-seq, we used proximity-based peak-to-gene annotation to identify differentially expressed genes targeted by NEUROG2 (Figure 1F). These results revealed direct overlap between ChiP-seq and RNA-seq datasets ranging from high-level pan-neuronal genes to single genes such as DLL3 and SOX4. This analysis demonstrates that NEUROG2 is the primary activator of pro-neural gene expression, and FD complement this neurogenic program by enhancing transcription in a NEUROG2-sensitive manner.

Small Molecules Induce Genome-wide Chromatin Remodeling

The relative accessibility of genetic elements is cell lineage dependent and directly regulates the reprogramming potency of transcription factors. To define the binding properties of NEUROG2 in MRC-5 fibroblasts, we used the assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) to generate an atlas of accessible and modified chromatin sites.

Initially, chromatin conformation was probed at locations corresponding to the 1,000 highest score peaks from 2-DPT NEUROG2 ChIP-seq datasets. Enrichment at these sites was evaluated using ATAC-seq datasets collected from MRC-5, NEUROG2-transduced MRC-5, and NFD-treated MRC-5. The majority of these binding events occurred in the absence of ATAC-seq signal, indicating that NEUROG2 functions effectively as a pioneer factor to access closed chromatin (Figure 2A). While only a modest increase in open chromatin is detected upon NEUROG2 expression, FD significantly enhances accessible chromatin at NEUROG2 binding sites and other pro-neural genetic regulatory elements. For instance, NEUROG2 binds within closed chromatin in the INSR-NTRK1, LLRC1OB-SYT7, and GADD45S loci, but a significant increase in relaxed chromatin is observed only in the presence of FD (Figure 2B). This demonstrates a role for FD in modulating accessibility at targets relevant to neuronal reprogramming.

Lineage-determinative transcription factors can bind within localized regions of chromatin containing high densities of histone modifications such as H3K27 acetylation, known as super-enhancers, to regulate gene expression (Hnisz et al., 2013). An analysis of H3K27AC super-enhancers in NFD-treated fibroblasts identified 454 potential H3K27AC super-enhancers containing a NEUROG2 bound site, of which 277 NEUROG2 binding sites were localized to regions with closed chromatin signatures (Figure 2C). While FD promoted an increase in accessibility from 50 to 118 sites, the vast majority of NEUROG2 binding sites in super-enhancers remained localized to regions of transposase-inaccessible chromatin (Figure 2C). As histone acetylation is generally considered a marker of active transcription, these dichotomous acetylation and chromatin accessibility results suggest that these features might not be easily correlated during the earliest stage of reprogramming. This also highlights the functional importance of NEUROG2 pioneer activity as a complement to FD-induced chromatin remodeling.

An analysis of accessible chromatin sites uniquely detected upon NFD treatment revealed that newly open chromatin was primarily localized to intergenic and intronic regions (Figure 2D). A functional classification of the genes annotated to these newly detected sites revealed strong enrichment of differentiation and neurogenesis-related factors (Figure 2E). A scan for transcription factor motifs that might respond to FD identified 50.1% of peaks as containing the NEUROG2-targeted enhancer box motif, as well as 4,672 peaks and 2,075 peaks containing cAMP response element (CRE) half-site and high-mobility group (HMG)-box motifs, respectively (Figure 2E).
A pairwise analysis of RNA-seq datasets was performed for genes annotated to ATAC-seq peaks uniquely detected in the NFD condition to identify how changes in chromatin structure affect gene expression (Figure 2F). This integrated analysis identified pro-neural genes involved in neurogenic transcription, Notch and Wnt signaling, and chromatin remodeling. Collectively, these datasets demonstrate that NEUROG2 functions as a pioneer transcription factor, while FD promote chromatin remodeling at pro-neural genes upregulated during reprogramming.

PRKACA Can Phosphorylate NEUROG2 Heterodimers to Modify DNA Affinity

Forskolin is a well-established activator of cAMP synthesis and PRKACA (protein kinase cAMP-activated catalytic

Figure 2. NEUROG2 Acts as a Pioneer Transcription Factor and Small Molecules Promote Chromatin Accessibility at Pro-neural Genetic Elements

(A) Clustered heatmaps representing open chromatin events detected at NEUROG2 binding sites (2-DPT ChIP-seq, leftmost black) in fibroblasts (MRC-5, black), NEUROG2-transduced fibroblasts (N, blue), and NFD-treated fibroblasts (NFD, rightmost red). Input-normalized tag densities are plotted in log2 scale and centered in a 4-kb window around each peak.

(B) NEUROG2 ChIP-seq (2 DPT) and ATAC-seq tracks for INSR, NTRK1, LRRC10B, SYT7, and GADD45G genomic loci. Gray shading highlights regions of NEUROG2 pioneer activity or FD-induced chromatin remodeling.

(C) Super-enhancers identified using H3K27AC and 2-DPT NEUROG2 ChIP-seq datasets with a Venn diagram depicting the distribution of super-enhancers in open versus closed chromatin.

(D) Genomic distribution of ATAC-seq peaks newly detected upon NFD treatment is biased toward intergenic and intronic genetic elements.

(E) Gene ontology analysis of open chromatin peaks newly detected following NFD treatment reveals an enrichment of neurogenesis- and differentiation-related classifications. Additionally the enhancer box, CRE half-site, and HMG-box motifs are diversely represented within these sites.

(F) Pairwise scatter plot representing the change in expression of genes annotated to ATAC-seq open chromatin peaks newly detected upon NFD treatment. Upregulated genes (red) were enriched for pro-neuronal identity, while downregulated genes (blue) had primarily non-neural function. Black dots indicate genes without significant changes.
subunit α2) kinase activity. To demonstrate that increased intracellular cAMP is fundamental to reprogramming, we successfully replaced forskolin with cAMP or dibutyryl cAMP during reprogramming (Figure S3A). Increased intracellular cAMP catalyzes the release of PRKACA from its regulatory subunit and stimulates protein kinase A activity. Remarkably, fibroblasts co-transduced with NEUROG2 and PRKACA rapidly outgrow TUBB3* and MAP2* neurites within 12 days (Figure S3B).

From these findings, we hypothesized four mechanisms by which FD might enhance NEUROG2 binding and promote chromatin remodeling (Figure 3A). The simplest mechanism is an FD-induced post-translational phosphorylation event that modulates NEUROG2 function. Cyclin-dependent kinases and glycogen synthase kinase 3 have been shown to regulate NEUROG2 DNA affinity, neuron differentiation, and motor neuron specification through site-specific phosphorylation (Ali et al., 2011; Ma et al., 2008). To identify whether increased NEUROG2 phosphorylation by PRKACA or another kinase is a primary function of FD during reprogramming, we performed introduced phosphodeficient and phosphomimetic amino acids at serine residues with a high probability for phosphorylation. Surprisingly, the replacement of S24, S91, S193, S207, S209, S219, S232, S239, or S242 had no effect on reprogramming efficiency, speed of conversion, or neuron morphology (Figures S3C and S3D). Even a NEUROG2 construct carrying substitutions at seven of these residues exhibited no effect on reprogramming. Although not an exhaustive analysis, this screen suggests that FD might act via another mechanism to drive the changes observed in ChIP-seq and ATAC-seq datasets.

To evaluate the second hypothetical mechanism underlying enhanced NEUROG2 chromatin occupancy, we investigated whether FD promote the phosphorylation of heterodimeric NEUROG2 complexes to enhance chromatin affinity. NEUROG2 heterodimers have been shown to exhibit variable chromatin affinity and functional stability in developing and mature neurons (Bronicki et al., 2012), GFP fusion does not bind DNA (Bronicki et al., 2012), and GFP fusion does not affect NEUROG2 association with TCF3, and NEUROG2 efficiently binds a ChIP-identified DLL3 enhancer box in vitro (Figure S4E). Thermoporetic traces were collected using differentially phosphorylated NEUROG2-TCF3 heterodimers and a DLL3 5'-CAGCTG-3', 5'-CAGATG-3', and 5'-CATCTG-3', were strongly enriched within this subset of binding sites. Phosphorylation-dependent binding of NEUROG2 heterodimers to each of these enhancer box sequences was assayed using microscale thermophoresis (Figures S4B–S4D) (Wienken et al., 2010).

DNA binding-induced changes in the thermophoresis of recombinant GFP-labeled NEUROG2 enabled us to quantitatively define the effects of differential phosphorylation on heterodimer affinity for sequence-specific enhancer box motifs. A preliminary electrophoretic mobility shift assay demonstrated that NEUROG2 homodimers do not bind DNA (Bronicki et al., 2012), GFP fusion does not affect NEUROG2 association with TCF3, and NEUROG2 efficiently binds a ChIP-identified DLL3 enhancer box in vitro (Figure S4E). Thermoporetic traces were collected using differentially phosphorylated NEUROG2-TCF3 heterodimers and a DLL3 5'-CAGCTG-3', native enhancer box or mutated internal dinucleotide pair to encode 5'-CAGATG-3' and 5'-CATCTG-3'. Surprisingly, phosphorylation caused sequence-specific reductions in enhancer box affinity (Figures S4B–S4D). Among several possibilities, this result suggests that TCF3 phosphorylation might be a mechanism for enhancing the specificity and resolution of NEUROG2 binding in a sequence-dependent manner, but is not likely to be the primary mechanism underlying enhanced chromatin accessibility or occupancy.

**CREB1 Promotes Neuron Survival**

In contrast to direct post-translational modifications, FD might mediate synergistic changes in gene expression and chromatin remodeling through PRKACA-targeted transcription factors that complement the NEUROG2 transcription program. To investigate our third proposed mechanism, we performed sequence-based discovery of regulons using RNA-seq datasets for upregulated genes without NEUROG2 annotation (Janky et al., 2014). This enabled the identification of 79 FD-induced genes containing at least one CRE motif within 5 kb of the target TSS (Table S2). CRE binding protein 1 (CREB1) directly interacts with this enriched motif and is an established target of

To investigate whether PRKACA phosphorylates TCF3 and related co-factors that heterodimerize with NEUROG2, we co-expressed PRKACA with TCF3 isoform E12, TCF3 isoform E47, TCF4, and TCF12. A denaturing electrophoretic analysis revealed that TCF3, TCF4, and TCF12 are phosphorylated by PRKACA (Figure S3G). Next, we assessed the functional roles for this phosphorylation. Genomic sequences extracted from NEUROG2 ChIP-seq occupancy sites annotated to the 100 most upregulated and downregulated genes identified by RNA-seq were directionally searched to identify the distribution of enhancer box sequences at these sites (Figure S4A). Three unique sequences, 5'-CAGCTG-3', 5'-CAGATG-3', and 5'-CATCTG-3', were strongly enriched within this subset of binding sites. To identify NEUROG2 co-factors (Figure S3F). This revealed direct interaction with transcription factor 3 (TCF3), which has been shown to modulate NEUROG2 chromatin binding dynamics in a phosphorylation-dependent manner (Li et al., 2012).
PRKACA phosphorylation (Gonzalez and Montminy, 1989). To investigate whether FD-activated CREB1 is essential to reprogramming, we performed CREB1 knockdown using a small hairpin RNA (shRNA) co-expressed with NEUROG2 (Figure 3B). CREB1 knockdown significantly reduced the survival of NFD-treated fibroblasts, indicating...
that CREB1 is critical to induced neuron survival (Figure 3C). This survival effect during reprogramming is consistent with its known role in endogenous neurons (Finkbeiner, 2000).

Next, an antibody specific for PRKACA-mediated CREB1 phosphorylation was used to perform ChIP-qPCR (Figure 3D). ChIP specificity was validated using established target genes (NR4A1 and SIK1), which both exhibited significant FD-induced upregulation in RNA-seq datasets. Interestingly, FD-enhanced phospho-CREB1 binding at SOX4 compared with binding in the absence of FD (Figure 3D). A genomic search revealed that several consensus CRE sites are located adjacent to NEUROG2 occupancy sites in the SOX4 enhancer (Figure 3E). Supporting a cooperative role for active CREB1 and NEUROG2 in SOX4 regulation, upregulation of SOX4 by FD could be further enhanced by NEUROG2 (Figure 3F). To evaluate whether co-binding might occur at numerous locations, we performed co-immunoprecipitation (coIP) on NEUROG2 and CREB1 chromatin (Figure 3G). NEUROG2-bound chromatin exposed positive for CREB1 and CREB1-bound chromatin validated this result, with NEUROG2 detected in NFD-treated cells but not control cells. These results demonstrate that FD enhances CREB1 enrichment at chromatin locations co-bound by NEUROG2.

SOX4 Promotes Chromatin Remodeling during Early Reprogramming

SOX4 is an HMG transcription factor with roles in pan-neuronal gene expression networks and the maturation of differentiating neurons (Bergsland et al., 2006; Shim et al., 2012). To investigate whether SOX4 upregulation is critical during early reprogramming, we performed SOX4 knockdown in NFD-treated fibroblasts (Figure 4A). Remarkably, disrupted SOX4 expression resulted in morphological defects and nearly abolished reprogramming (Figure 4B). ChIP-seq was next used to define the genetic targets of this factor. Alignment of SOX4 and NEUROG2 profiles revealed significant overlap with 413 of 1,119 total SOX4 binding events directly shared with NEUROG2 (Figure 4C). Peak-to-gene annotation next enabled an analysis of how co-binding effects gene expression. Relative to NEUROG2 alone, NFD-treated cells exhibited enhanced transcription for 81% of co-bound genes including NEUROD1 and NEUROD4 (Figure 4C). In addition, these datasets demonstrated that SOX4 also targets components of the SWI/SNF
nucleosome remodeling complex including ARID1A, ARID1B, SMARCA2, SMARCAD1, SMARCC1, SMARCC2, SMARCD1, SMARCD2, and SMARCE1 (Kwon et al., 1994; Son and Crabtree, 2014). This suggests that SOX4 not only functions to regulate transcription via direct chromatin binding but might indirectly promote chromatin remodeling to increase accessibility.

A comprehensive analysis of this indirect chromatin remodeling activity was performed using SOX4 knockdown coupled with ATAC-seq in NFD-treated fibroblasts. FD-induced chromatin decondensation at NEUROG2 binding sites and adjacent genetic elements in the NEUROD1 and NEUROD4 loci is significantly reduced following SOX4 depletion (Figure 4D). As both of these genes are significantly upregulated only in the NFD condition, this analysis demonstrates how the synergistic co-binding of NEUROG2 and SOX4 accompanied by a transition to open chromatin patterns is fundamental to the expression of genes that promote reprogramming. Broadening this to a genome-wide analysis, SOX4 knockdown results in an extensive loss of approximately 59% of total open chromatin sites in NFD-treated fibroblasts (Figure 4E). A simultaneous decrease in ATAC score is also observed for 22% of the remaining peaks following knockdown.

NEUROD1 is required to promote human fibroblast-to-neuron reprogramming (Pang et al., 2011). As they represent direct downstream targets of NEUROG2 and SOX4, ChIP for NEUROD1 and NEUROD4 was used to investigate the occupancy of these factors within regions of newly open chromatin (Figure 4F). Publicly available ChIP-seq datasets from mouse cortical tissue were analyzed for potential target sites that directly overlap with human NEUROG2 ChIP and ATAC-seq sites (Pataskar et al., 2016). NEUROD1 was enriched at NEUROG2 bound sites in regions of DLL3, HES6, and NHLH1, while NEUROD4 was significantly less enriched within these loci in reprogrammed human fibroblasts. Furthermore, NEUROD1 exhibited modest enrichment at the NEUROD4 promoter within an open chromatin site lost following SOX4 depletion. These findings collectively implicate FD-induced upregulation of SOX4 in pro-neural chromatin remodeling activity.

NEUROG2 and SOX4 Induce H3K27 Acetylation at Pro-neural Genetic Elements

We next performed a condition-dependent analysis of H3K27AC and H3K27 trimethylation (H3K27ME3) to define the native epigenetic landscape of fibroblasts and factor-specific changes triggered during reprogramming. First, a quantitative analysis of H3K27AC ChIP-seq read densities around NEUROG2 ChIP-seq peaks revealed that FD only modestly enhanced acetylation at these sites. In contrast, FD significantly enhanced H3K27AC enrichment surrounding SOX4-targeted genomic sites (Figure 5A). Secondly, this FD-enhanced acetylation is even more pronounced at occupancy sites annotated to genes targeted by both NEUROG2 and SOX4 (Figure 5B). This further demonstrates that the synergistic genetic and epigenetic reprogramming induced by NEUROG2 and FD is, at least in part, mediated by SOX4.

In-depth analysis of transcription factor binding profiles and histone marks at DLL3 and SMARCC2 provided insights into both the mechanism of reprogramming and the epigenetic environment of MRC-5 fibroblasts. For instance, binding sites for NEUROG2 and SOX4 directly overlap upstream of DLL3, and H3K27AC enrichment significantly increases at this site upon FD treatment (Figure 5C). A simultaneous loss of H3K27ME3 is observed throughout this region. In contrast, the SWI/SNF subunit SMARCC2 is targeted only by SOX4, and the H3K27ME3 signal is relatively low throughout this locus (Figure 5D). This indicates that MRC-5 fibroblasts might be more epigenetically poised for reprogramming than other lineage-committed adult skin fibroblasts.

Induced Neuronal Identity in Cell Types Refractory to NEUROG2 Reprogramming

The success of neuronal reprogramming depends upon the ability of exogenous reprogramming factors and downstream induced transcription factors to traverse the cell-of-origin epigenetic environment. Unlike the high-efficiency conversion of fetal MRC-5 fibroblasts, NEUROG2 and FD do not functionally or morphologically reprogram adult skin fibroblast lines without additional SOX11 overexpression (Liu et al., 2013). Likewise, the addition of FGF2 and kenpaullone to this reprogramming cocktail promote the survival of reprogrammed neurons (Liu et al., 2013, 2015). As SOX4 and SOX11 exhibit redundant functions in the developing nervous system (Cheung et al., 2000; Miller et al., 2013), we investigated whether an inability to activate this factor hinders adult cell reprogramming efficiency.

An initial assessment of SOX4 expression in adult skin fibroblast lines revealed that adult fibroblasts express SOX4 at comparatively lower levels than MRC-5 (Figure 6A). Next, we evaluated whether SOX4 upregulation is beneficial to reprogramming by transducing adult fibroblasts with NEUROG2- and SOX4-encoding lentivirus in combination with FD and FGF2. These factors induced TUBB3 expression and modest neurite outgrowth in all four lines (Figure 6B). However, the progressive maturation of these neurons is significantly slower and less complex than neurons derived from fetal cells (Figure 6C).

A SOX4 ChIP time course using AG05811 adult fibroblasts provided insight into the varied genetic landscapes of adult and fetal fibroblast lines. SOX4 enrichment at sites identified by ChIP-seq in 1-DPT MRC-5 cells exhibited
significantly delayed occupancy in adult cells through 4 DPT (Figure 6D). This suggests that condensed chromatin might occlude the binding sites of SOX4 and other downstream NEUROG2-driven transcription factors to inhibit the neurogenic transcription programs critical to adult fibroblast reprogramming.

Under conditions of chromatin accessibility, NEUROG2, FD, and the endogenous upregulation of SOX4 should be sufficient to drive reprogramming of adult fibroblasts. Since SOX4 targets numerous subunits of the SWI/SNF complex (Figure 4C), SMARCA4 and NEUROG2 were co-expressed to determine whether chromatin remodeling could potentiate reprogramming (Figure 6E). A fraction of transduced fibroblasts rounded, outgrew neurites, and exhibited modest TUBB3 expression but, similar to SOX4 overexpression, these fibroblasts did not adopt highly complex neuronal morphology. The co-expression of SMARCB1 and SMARCC2 with NEUROG2 and SMARCA4 failed to further improve conversion. A chemical screen targeting chromatin modifiers was then performed to identify whether compounds that enhance chromatin accessibility could improve adult fibroblast reprogramming (Figure 6E). We identified several chemicals sufficient to induce rounding and modest neurite outgrowth, but only the histone deacetylase inhibitor FK228 was sufficient to rapidly induce a large cell population with neuron-like morphology (Figure 6F). Remarkably, this morphological conversion occurred within 20–24 hr of exposure; however, cells treated with 0.5 μM FK228 failed to survive long term in culture.

In addition to histone deacetylase inhibition, FK228 suppresses the RAS-MAP kinase signaling pathway to induce apoptosis (Kobayashi et al., 2006). To abrogate these pro-apoptotic functions while retaining histone deacetylase inhibition, we pre-treated adult skin fibroblasts with the pan-caspase inhibitor Z-VAD-FMK (Wang et al., 2012), the necrosis inhibitor necrosulfonamide, and the ferroptosis inhibitor liproxstatin-1. Fibroblasts were continuously

**Figure 5. Small Molecules Enhance H3K27 Acetylation at NEUROG2- and SOX4-Targeted Genetic Elements**

(A and B) Distribution of H3K27 acetylation (H3K27AC) ChIP-seq reads from untreated fibroblasts (black), NEUROG2-transduced fibroblasts (blue), and NFD-treated fibroblasts (red) centered around (A) SOX4 ChIP-seq peaks and (B) NEUROG2 and SOX4 co-bound peaks within a 4-kb window.

(C and D) ChIP-seq tracks for SOX4, NEUROG2, H3K27 acetylation, and H3K27 trimethylation generated from fibroblasts (MRC-5), NEUROG2-transduced fibroblasts (N), and NFD-treated fibroblasts (NFD) around the (C) DLL3 and (D) SMARCC2 loci.
Figure 6. Increased Chromatin Accessibility Promotes Human Adult Fibroblast and Glioblastoma Cell Reprogramming

(A) Expression of SOX4 RNA and protein in adult fibroblast lines relative to fetal MRC-5 fibroblasts (mean ± SD; biological quadruple samples for qPCR).

(B) Percentage of rounded adult fibroblasts expressing TUBB3 when treated with NFD, FGF2, and mouse Sox4 (red) or human SOX4 (gray) (mean ± SD; n = 10 random image fields from biological triplicates).

(C) Examples of progressive neuronal conversion using adult fibroblasts treated with reprogramming factors. Fibroblasts adopt a rounded morphology (5–10 DPT), outgrow low-complexity neurites (10–20 DPT), and infrequently adopt complex neuronal morphology (≥21 DPT). Scale bars, 20 μm.

(D) SOX4 ChIP-qPCR time course demonstrating SOX4 chromatin occupancy at 1, 2, and 4 DPT in adult fibroblasts treated with reprogramming factors (mean ± SD; biological triplicates for each ChIP).

(legend continued on next page)
treated starting at 1 day before FK228 treatment and 2 days following FK228 exposure. Due to the rapid adoption of neuronal morphology, we also included the brain-derived neurotrophic factor BDNF, the glial cell line-derived neurotrophic factor GDNF, and neurotrophin-3 at the outset of reprogramming. Fibroblasts treated with FK228 for 24 hr starting 2 days after infection exhibited modest improvements in cell survival upon Z-VAD-FMK treatment. Liproxstatin-1 and necrosulfonamide exhibited no effect on cell survival or reprogramming efficiency. While Z-VAD-FMK increased the survival of induced neurons compared with FK228 alone, a significant fraction of cells still died 5–7 days after initial transduction.

FK228-induced apoptosis is highly sensitive to BCL2L1 activity, but remains unaffected by BCL2 (Newbold et al., 2008). We then co-expressed BCL2L1 with NEUROG2 to determine whether cell survival could be enhanced following FK228 treatment. Interestingly, adult fibroblasts co-expressing NEUROG2 and BCL2L1 for 4 days in neuron induction medium followed by a brief 24-hr exposure to FK228 from days 4–5 generated morphologically complex neurons that stained TUBB3+, MAP2+, and RBFOX3+ after 12 days in culture (Figures 6G and 6H). These results show the neuronal reprogramming capacity of FK228 when apoptotic mechanisms are controlled during conversion. This chemical-based approach to chromatin modification further demonstrates that NEUROG2 and small molecules are sufficient to induce early neurogenic programs in fibroblasts refractory to reprogramming in the absence of epigenetic barriers to activity. The toxicity of FK228 to fibroblasts not transduced with BCL2L1 resulted in fluorescent cell debris that occluded an extensive immunocytochemical analysis of neuronal maturation. Further optimization is needed for long-term survival, maturation, and subtype identity analyses.

Distinct cell types exhibit disparate heterochromatin condensation patterns. To investigate whether a more permissive chromatin state would facilitate rapid reprogramming, we reprogrammed adult human U-251 glioblastoma cells. In the absence of SOX4 overexpression, NFD-treated cells adopt TUBB3+, and MAP2+ neuron-like morphology with 21% ± 3% efficiency (Figure 6I). However, inclusion of SOX4 endows these cells with complex morphology and conversion efficiency of greater than 90% (Figure 6I). This alternative cellular context provides independent confirmation of the fundamental roles played by SOX4 and SOX4-targeted downstream actors in neuronal reprogramming.

Collectively, these findings represent a stepwise molecular mechanism for the synergistic events underlying early fibroblast-to-neuron reprogramming (Figure 7). The hierarchical activation of NEUROG2, CREB1, and SOX4 transcriptional programs, as well as, FD-induced chromatin remodeling, can potentiate reprogramming in a cell age- and lineage-dependent manner.

DISCUSSION

Here we performed genome-wide ATAC-seq, ChIP-seq, and RNA-seq profiling to define the immediate-early molecular events that catalyze fibroblast-to-neuron reprogramming by NEUROG2 and small molecules. While NEUROG2 is sufficient to reprogram glial cells into neurons, this transcription factor fails to reprogram somatic fibroblasts independent of the small molecules F and D (Liu et al., 2013). This synergistic reprogramming activity is the result of extensive SOX4-mediated chromatin remodeling that enhances NEUROG2 occupancy at pro-neural genetic elements and, in coordination with downstream factors, redefines the fibroblast transcriptome. These findings enabled us to reprogram adult fibroblasts and glioblastoma cells refractory to conversion by NEUROG2 alone.

NEUROG2 Functions as a Pioneer Factor during Early Reprogramming

Transcription factors in both pluripotent and direct reprogramming systems utilize either off-target cooperative chromatin binding (Soufi et al., 2012) or on-target pioneer binding mechanisms to activate gene expression (Wapinski et al., 2013). Similar to ASCL1 function in
mouse embryonic fibroblasts, NEUROG2 exhibits on-target pioneer binding in heterochromatin regions when expressed in human fetus. Unexpectedly, however, FD-induced chromatin remodeling dramatically enhances NEUROG2 chromatin occupancy. This enhanced occupancy demonstrates that FD act synergistically with NEUROG2 to promote reprogramming rather than function as complementary, but independent transcriptional activators. We investigated four mechanisms by which FD might enhance NEUROG2 activity in a cell lineage-dependent manner. This revealed that FD enhance chromatin accessibility to promote hierarchical associations with CREB1 and SOX4 to initiate secondary pro-neural transcription programs. These observations validated an earlier computational model indicating that CREB1 acts as a co-activator of NEUROG2 during mouse telencephalon development (Gohlke et al., 2008), as well as roles for SOX4 in establishing a cross-regulatory neurogenic network involving SWI/SNF complex factors (Ninkovic et al., 2013). The genome-wide interplay of CREB1, NEUROG2, and SOX4 during reprogramming highlights how tightly interwoven genetic regulatory networks require coordinated multi-factor occupancy to effectively enhance the expression of pro-neural downstream actors previously implicated in reprogramming such as NEUROD1 (Pang et al., 2011, Guo et al., 2014) and NEUROD4 (Masserdotti et al., 2015).

**SOX4 Indirectly Enhances Neuronal Identity through Chromatin Remodeling**

Sox4 and Sox11 exhibit similar expression patterns in the developing nervous system, and simultaneous, but not independent, knockout of these factors results in reduced neuron survival (Cheung et al., 2000; Miller et al., 2013). CHD7 activates a neuronal differentiation program in neural stem cells by remodeling the Sox4 and Sox11 promoters to upregulate expression (Feng et al., 2013). Of these transcription factors, only SOX4 is significantly expressed in human MRC-5 fibroblasts, and transient knockdown of SOX4 abolishes reprogramming. Importantly, NEUROG2 outcompetes REST to drive Sox4 and promote neuronal differentiation in post-mitotic neural precursor cells (Bergsland et al., 2006). In parallel with this mechanism, NEUROG2 co-binds with CREB1 to upregulate SOX4 in MRC-5 fibroblasts, demonstrating similar regulatory mechanisms during development and reprogramming.

An analysis of SOX4 targets revealed extensive coordination with NEUROG2 at factors with FD-enhanced or synergistically induced expression such as HES6, NEUROD1, NEUROD4, and WNT10A. In addition to
synergistic regulation, SOX4 targets epigenetic remodeling factors such as EZH2 (Tiwari et al., 2013) and numerous SWI/SNF subunits.

In parallel with this epigenetic function, the inclusion of either SOX4 or SOX11 is required to induce adult fibroblasts and glioblastoma cells to adopt TUBB3+ identity (Liu et al., 2013; Su et al., 2014). The morphological complexity of these neurons can be significantly enhanced with ISL1 and LHX3 overexpression (Liu et al., 2015). As adult fibroblast lineages exhibit highly defined epigenetic signatures that make these cells less poised for reprogramming, mechanistic studies have predominantly relied on embryonic or fetal fibroblasts (Masserotti et al., 2015; Wapinski et al., 2013; Gascon et al., 2016). Our proposed mechanism suggests that SOX4 reduces epigenetic barriers to reprogramming to facilitate fate conversion of adult cell lineages.

Mature mammalian neurons exhibit distinct epigenomic signatures that reflect subtype-specific functions. Based on our ATAC-seq and SOX4 ChIP-seq datasets, we hypothesized that modulating the epigenome of fibroblasts with chromatin remodeling complexes might enhance reprogramming efficiency as observed with induced pluripotent stem cells (Singhal et al., 2010). While a subset of adult fibroblasts transduced with both SMARCA4 and NEUROG2 expressed TUBB3 and low-complexity neurites, the addition of other SWI/SNF subunits such as SMARCB1 and SMARCC2 failed to improve reprogramming, likely due to imbalances in subunit stoichiometric required for a minimally functional SWI/SNF complex.

Targeting Epigenetic Modifiers Enhances Adult Human Fibroblast Reprogramming

We next turned to a small-molecule-based strategy for epigenetic remodeling inspired by recently described chemical reprogramming systems (Hu et al., 2015; Li et al., 2015; Zhang et al., 2015). This led to the replacement of SOX4 by FK228, which rapidly induces neuron morphology but confers poor survival. Upon co-expression of NEUROG2 and BCL2L1 with FK228 we observed morphologically complex neurons. While we were not able to directly image MAP-tau expression due to high background, these cells did become TUBB3+, MAP2+, and RBFOX3+. Collectively, these results highlight multiple roles for SOX4 in early reprogramming as well as the functional impact of epigenetic remodeling on reprogramming efficiency in adult cells.

Chromatin Accessibility and Histone Signatures Regulate Reprogramming Efficiency

Distinct histone acetylation and methylation signatures regulate transcription factor activity and gene expression. Patterned post-translational modifications direct transcription factors to specific genomic regulatory regions. H3K27AC enrichment was observed at both NEUROG2 and SOX4 bound sites in reprogrammed fibroblasts. Furthermore, FD significantly increased this enrichment at cooperatively bound sites.

CREB1 phosphorylation promotes the recruitment of histone acetyltransferases that induce local nucleosome remodeling (Michael et al., 2000). This activity likely mediates, in part, the FD-enhanced acetylation observed at NEUROG2 occupancy sites. Histone acetyltransferase inhibitors such as valproic acid (Niu et al., 2013) and FK228 drive the adoption of neuronal identity, underscoring the significance of persistent acetylation at neurogenic promoters during early conversion.

Collectively, this work presents a stepwise mechanism detailing the early events critical to fibroblast-to-neuron reprogramming and highlights how changes to the cell-of-origin chromatin landscape measurably affect conversion efficiency.

EXPERIMENTAL PROCEDURES

ATAC-Seq

Fluorescent-sorted cells were lysed, resuspended in transposition mix, incubated for 30 min at 37°C, and amplified with barcoded primers. Single-end 50-base length reads were generated and aligned to GRCh37/hg19 for analysis.

ChIP, ChIP ColP, and ChIP-Seq

ChIP and ChIP-seq used fibroblasts transduced with HA-NEUROG2 or NEUROG2-HA-SOX4. Cells were fixed and lysed, and nuclei suspended in shearing buffer. Chromatin was sheared, quantified, and combined with target-specific antibodies. Chromatin was captured with magnetic beads, washed, eluted, and reverse crosslinked. Single-end 50-base length reads were generated and aligned to GRCh37/hg19 for analysis. ChIP ColP followed the same protocol using denaturing buffer elution. Detailed protocols are provided in Supplemental Experimental Procedures and Table S3.

Neuron Induction

MRC-5 fibroblasts were seeded on standard 24-well glass coverslips, transduced with NEUROG2-encoding lentivirus, and supplemented with 10 μM forskolin and 1 μM dorsomorphin 48 hr after infection. Medium was changed every 2 days for the duration of culture.

RNA-Seq

Total RNA was isolated from control or treated MRC-5 fibroblasts and quantified for library synthesis. Single-end 50-base length sequencing reads were aligned to GRCh37/hg19 for differential expression analysis.

shRNA-Mediated Gene Knockdown

PCR was used to generate a lentiviral construct containing miR30 regulatory sequences and shRNAs (Table S3; Fellmann et al., 2013).
ACCESSION NUMBERS
The accession number for bioinformatics datasets is GEO: GSE75912.

SUPPLEMENTAL INFORMATION
Supplemental Information includes Supplemental Experimental Procedures, four figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2016.09.013.

AUTHOR CONTRIBUTIONS
D.K.S., conceptualization, data curation, formal analysis, investigation (Figures 1, 2, 3, 4, 5, 6A–6H, 7, and S1–S4), methodology, software, supervision, validation, visualization, writing – original draft, and writing – review and editing; J.Y., investigation (Figures 6G–6I); M.-L.L., supervision; C.-L.Z., conceptualization, funding acquisition, resources, and supervision.

ACKNOWLEDGMENTS
We thank Mark Borromeo, Bradford Casey, Jane E. Johnson, Stephen Johnson, Taekyung Kim, Zhigao Wang, and members of the Zhang laboratory. This work was supported by the Decherr Foundation, the Mobility Foundation, the NIH (NS070981, NS092616, NS093502, and NS088095), the Texas Institute for Brain Injury and Repair, and the Welch Foundation (I-1724).

Received: June 6, 2016
Revised: September 28, 2016
Accepted: September 29, 2016
Published: October 27, 2016

REFERENCES
Ali, F., Hindley, C., McDowell, G., Deibler, R., Jones, A., Kirschnerr, M., Guillemot, F., and Philpott, A. (2011). Cell cycle-regulated multi-site phosphorylation of Neurogenin 2 coordinates cell cycling with differentiation during neurogenesis. Development 138, 4267–4277.

Bergsland, M., Werme, M., Malewicz, M., Perlmann, T., and Muhr, J. (2006). The establishment of neuronal properties is controlled by Sox4 and Sox11. Genes Dev. 20, 3475–3486.

Berninger, B., Costa, M.R., Koch, U., Schroeder, T., Sutor, B., Grothe, B., and Götz, M. (2007). Functional properties of neurons derived from in vitro reprogrammed postnatal astroglia. J. Neurosci. 27, 8654–8664.

Bronicki, L.M., Bélanger, G., and Jasmin, B.J. (2012). Characterization of multiple exon 1 variants in mammalian HuD miRNA and neuron-specific transcriptional control via neurogenin 2. J. Neurosci. 32, 11164–11175.

Chanda, S., Ang, C.E., Davila, J., Pak, C., Mall, M., Lee, Q.Y., Athienitis, H., Jung, S.W., Stühlfot, T.C., and Wernig, M. (2014). Generation of induced neuronal cells by the single reprogramming factor ASCL1. Stem Cell Rep. 3, 282–296.

Cheung, M., Abu-Elmagd, M., Clevers, H., and Scotting, P.J. (2000). Roles of Sox4 in central nervous system development. Mol. Brain Res. 79, 180–191.

Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhr, 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.

Feng, W., Khari, M.A., Bellvis, P., Zhu, Z., Bernhardt, O., Herold-Mende, C., and Liu, H.K. (2013). The chromatin remodeler CHD7 regulate adult neurogenesis via activation of SoxC transcription factors. Cell Stem Cell 13, 62–72.

Finkbeiner, S. (2000). CREB couples neurotrophin signals to survival messages. Neuron 25, 11–14.

Fode, C., Gradwohlg, G., Morin, X., Dierich, A., LeMeur, M., Goridis, C., and Guillemtot, F. (1998). The bHLH protein Neurogenin 2 is a determinant factor for epibranchial placode-derived sensory neurons. Neuron 20, 483–494.

Gascon, S., Murenu, E., Masserdottt, G., Ortega, F., Russo, G.L., Petrik, D., Deshpande, A., Heinrich, C., Karow, M., Robertson, S.P., et al. (2016). Identification and successful negotiation of a metabolic checkpoint in direct neuronal reprogramming. Cell Stem Cell 18, 396–409.

Gohlke, J.M., Armant, O., Parham, F.M., Smith, M.V., Zimmer, C., Castro, D.S., Nguyen, L., Parker, J.S., Gradwohlg, G., Portier, C.J., and Guillemtot, F. (2008). Characterization of the proneural gene regulatory network during mouse telencephalon development. BMC Biol. 6, 15.

Gonzalez, G.A., and Montminy, M.R. (1989). Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. Cell 59, 675–680.

Guo, Z., Zhang, L., Wu, Z., Chen, Y., Wang, F., and Chen, G. (2014). In vivo direct reprogramming of reactive glial cells into functional neurons after brain injury and in an Alzheimer's disease model. Cell Stem Cell 14, 188–202.

Heinrich, C., Blum, R., Gascon, S., Masserdottt, G., Tripathi, P., Sanchez, R., Tiedt, S., Schroeder, T., Götz, M., and Berninger, B. (2010). Directing astroglia from the cerebral cortex into subtype specific functional neurons. PLoS Biol. 8, e1000373.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell 155, 934–947.

Hu, W., Qiu, B., Guan, W., Wang, Q., Wang, M., Li, W., Gao, L., Shen, L., Huang, Y., Xie, G., et al. (2015). Direct conversion of normal and Alzheimer’s disease human fibroblasts into neuronal cells by small molecules. Cell Stem Cell 17, 204–212.

Janky, R., Verfaillie, A., Imrichová, H., Van de Sande, B., Standaert, L., Christians, V., Hulsel mans, G., Herten, K., Sanchez, M.N., Potter, D., et al. (2014). PRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput. Biol. 10, e1003731.

Kobayashi, Y., Ohtsuki, M., Murakami, T., Kobayashi, T., Sutheesophon, K., Kitayama, H., Kano, Y., Kusano, E., Nakagawa, H., and Furukawa, Y. (2006). Histone deacetylase inhibitor FK228 suppresses the Ras-MAP kinase signaling pathway by upregulating Rap1 and induces apoptosis in maqlignant melanoma. Oncogene 25, 512–524.
Kwon, H., Imbalzano, A.N., Khavari, P.A., Kingston, R.E., and Green, M.R. (1994). Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex. Nature 370, 477–481.

Li, S., Mattar, P., Zinyk, D., Singh, K., Chaturvedi, C.-P., Kovach, C., Dixit, R., Kurrasch, D.M., Ma, Y.-C., Chan, J.A., et al. (2012). GSK3 temporally regulates Neurogenin 2 proneural activity in the neocortex. J. Neurosci. 32, 7791–7805.

Li, X., Zuo, X., Jing, J., Ma, Y., Wang, J., Liu, D., Zhu, J., Du, X., Xiong, L., Du, Y., et al. (2015). Small-molecule-derived direct reprogramming of mouse fibroblasts into functional neurons. Cell Stem Cell 17, 195–203.

Liu, M.-L., Zhang, 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.

Liu, M.-L., Zhang, T., and Zhang, C.-L. (2015). Direct lineage reprogramming reveals disease-specific phenotypes of motor neurons from human ALS patients. Cell Rep. 14, 115–128.

Ma, Q., Kittner, C., and Anderson, D.J. (1996). Identification of neurogenin, a vertebrate neuronal determination gene. Cell 87, 43–52.

Ma, Y.-C., Song, M.-R., Park, J.P., Ho, H.-Y.H., Hu, L., Kurtev, M.V., Zieg, J., Ma, Q., Pfaff, S.L., and Greenberg, M.E. (2008). Regulation of motor neuron specification by phosphorylation of Neurogenin 2. Neuron 58, 65–77.

Masserotti, G., Gillotin, S., Sutor, B., Drechsel, D., Irmler, M., Jørgensen, H.F., Sass, S., Theis, E.J., Beckers, J., Berninger, B., et al. (2015). Transcriptional mechanisms of proneural factors and REST in regulating neuronal reprogramming of astrocytes. Cell Stem Cell 17, 74–88.

Mertens, J., Paquola, A.C.M., Ku, M., Hatch, E., Böhne, L., Ladjevardi, S., McGrath, S., Campbell, B., Lee, H., Herdy, J.R., et al. (2015). Directly reprogrammed human neurons retain aging-associated transcriptomic signatures and reveal age-related nucleocytoplasmic defects. Cell Stem Cell 17, 705–718.

Michael, L.F., Ashara, H., Shulman, A.I., Kraus, W.L., and Montminy, M. (2000). The phosphorylation status of a cyclic AMP-responsive activator is modulated via a chromatin-dependent mechanism. Mol. Cell Biol. 20, 1596–1603.

Miller, J.A., Nathanson, J., Franjic, D., Shim, S., Dalley, R.A., Shapouri, S., Smith, K.A., Sunkin, S.M., Bernard, A., Bennett, J.L., et al. (2013). Conserved molecular signatures of neurogenesis in the hippocampal subgranular zone of rodents and primates. Development 140, 4633–4644.

Newbold, A., Lindermann, R.K., Cluse, L.A., Whitecross, K.F., Dear, A.E., and Johnstone, R.W. (2008). Characterisation of the novel apoptotic and therapeutic activities of the histone deacetylase inhibitor romidepsin. Mol. Cancer Ther. 7, 1066–1079.

Ninkovic, J., Steiner-Mexadri, A., Jawerka, M., Akinci, U., Masserdotti, G., Petricca, S., Fischer, J., von Holst, A., Beckers, J., Lie, C.D., et al. (2013). Essential role for BAF complex interacting with Pax6 in establishment of a core cross-regulatory neurogenic network. Cell Stem Cell 13, 403–418.

Niu, W., Zang, T., Zou, Y., Fang, S., Smith, D.K., Bachoo, R., and Zhang, C.-L. (2013). In vivo reprogramming of astrocytes to neuroblasts in the adult brain. Nat. Cell Biol. 15, 1164–1175.

Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., Citri, A., Sebastiani, V., Marro, S., Südhof, T.C., et al. (2011). Induction of human neuronal cells by defined transcription factors. Nature 476, 220–223.

Pataskar, A., Jung, J., Smialowski, P., Noack, F., Calegari, F., Straub, T., and Tiwari, V.K. (2016). NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. EMBO J. 35, 24–45.

Shim, S., Kwan, K.Y., Li, M., Lefèvre, V., and Sestan, N. (2012). Cis-regulatory control of corticospinal system development and evolution. Nature 486, 74–79.

Scardigli, R., Schuurmans, C., Gradwohl, G., and Guillemot, F. (2001). Crossregulation between neurogenin 2 and pathways specifying neuronal identity in the spinal cord. Neuron 31, 203–217.

Singhal, N., Graumann, J., Wu, G., Araúzo-Bravo, M.J., Han, D.W., Greber, B., Gentile, L., Mann, M., and Schönler, H.R. (2010). Chromatin-remodeling components of the BAF complex facilitate reprogramming. Cell 141, 943–955.

Son, E.Y., and Crabtree, G.R. (2014). The role of BAF (mSWI/SNF) complexes in mammalian neural development. Am. J. Med. Genet. 166C, 333–349.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors’ initial engagement with the genome. Cell 15, 994–1004.

Su, Z., Zang, T., Liu, M.L., Wang, L.L., Niu, W., and Zhang, C.-L. (2014). Reprogramming the fate of human glioma cells to impede brain tumor development. Cell Death Dis. 5, e1463.

Thoma, E.C., Wischmeyer, E., Offen, N., Maurus, K., Sirén, A.-L., Schartl, M., and Wgner, T.U. (2012).ECTopic expression of neurogenin 2 alone is sufficient to induce differentiation of embryonic stem cells into mature neurons. PLoS One 7, e38651.

Tiwari, N., Tiwari, V.K., Waldmeier, L., Bialwierz, P.J., Arnold, P., Pachkov, M., Meyer-Schaller, N., Schübeler, D., van Nimwegen, E., and Christofori, G. (2013). Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming. Cancer Cell 23, 768–783.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature 463, 1035–1041.

Wang, Z., Jiang, H., Chen, S., Du, E., and Wang, X. (2012). The mitochondrial phosphatase PGAMS functions at the convergence point of multiple necrotic death pathways. Cell 148, 228–243.

Wapinski, O.L., Vierbuchen, T., Qu, K., Lee, Q.Y., Chanda, S., Fuentes, D.R., Giresi, P.G., Ng, Y.H., Marro, S., Neff, N.E., et al. (2013). Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. Cell 155, 621–635.

Wienken, C.J., Baaske, P., Rothbauer, U., Braun, D., and Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. Nat. Commun. 1, 100.

Zhang, L., Yin, J.C., Yeh, H., Ma, N.X., Lee, G., Chen, X.A., Wang, Y., Lin, L., Chen, L., Jin, P., et al. (2015). Small molecules efficiently reprogram human astroglial cells into functional neurons. Cell Stem Cell 17, 735–747.
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  NEUROG2 ChIP-seq Intensity Analysis Methodology

No Temporal Comparison

- FD

+ FD

0.5 DPT  1 DPT  2 DPT

Conditional Comparison Within Each Time Point

Differential Peaks

Condition-Dependent Differential Peaks Detected at Multiple Time Points

B  Genomic Distribution of NEUROG2 Binding Events

| Location  | 0.5 DPT   | 1 DPT   | 2 DPT   |
|-----------|-----------|---------|---------|
|           | 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 Analysis 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).
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, \( \log_2(\text{NFD FPKM}) \geq 1 \), and p value \( \leq 0.05 \) 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, \textit{NEUROG2}-transduced fibroblasts exposed to FD).
A

Reprogramming Efficiency

B

GFP TUBB3

C

Phospho-Deficient NEUROG2 Constructs

D

Phosphomimetic NEUROG2 Constructs

E

Reprogramming Efficiency

F

NEUROG2 Interactors

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*<sub>110-228</sub>-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).
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**

| 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

Cell culture, lentivirus production, and neuron induction

Cell culture

Lentivirus production

MRC-5 transdifferentiation: NEUROG2, forskolin, and dorsomorphin
MRC-5 transdifferentiation: forskolin replacement screen
MRC-5 transdifferentiation: NEUROG2, caPRKACA, and dorsomorphin
MRC-5 transdifferentiation: NEUROG2 phosphomutant screen
MRC-5 transdifferentiation: NEUROG2 deletion construct
Glioblastoma cell transdifferentiation: NEUROG2 and SOX4
Adult fibroblast transdifferentiation: NEUROG2 and SOX4
Adult fibroblast transdifferentiation: NEUROG2 and SWI/SNF factors
Adult fibroblast transdifferentiation: NEUROG2, small molecules, and Bcl2/l1

Chromatin immunoprecipitation and next-generation sequencing

Assay for transposase-accessible chromatin
Crosslinking chromatin immunoprecipitation
Crosslinking chromatin co-immunoprecipitation
Flow cytometry and low-cell native chromatin immunoprecipitation
Massively parallel DNA sequencing
Quantitative real-time PCR

Electrophoretic mobility shift assay

Mass spectrometry

Microscale thermophoresis

PRKACA phosphorylation assay

RNA sequencing

Reverse transcription and quantitative real-time PCR

shRNA-mediated gene knockdown

References
Abbreviations

°C     degree(s) Celsius
\times g relative centrifugal force
\mu g  microgram(s)
\mu l  microliter(s)
\mu m  micrometer(s)
\mu M  micromolar
5-azacitidine 4-amino-1-\beta-D-ribofuranosyl-1,3,5-triazin-2(1H)-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-\beta-D-erythro-pentofuranosyl)-1,3,5-triazin-2(1H)-one
dehydrocholic acid (3\alpha,5\beta,12\alpha,20R)-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 (1S,2R,5R)-5-(4-Amino-1H-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]-\beta-D-ribofuranosyl]
FGF2 fibroblast growth factor 2
FK228 cyclo[(2Z)-2-amino-2-butenoyl-L-valyl-(3S,4E)-3-hydroxy-7-mercapto-4-heptenoyl-
D-valyl-D-cysteiny], cyclic (3-5) disulfide
formaldehyde  methanal
forskolin  7β-Acetoxy-8,13-epoxy-1α,6β,9α-trihydroxyabd-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  (6S)-4-(4-Chlorophenyl)-2,3,9-trimethyl-6H-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(1H)-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% CO₂.

Lentivirus production

293T/17 cells were seeded at a density of 3×10⁶ 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 0.8×10⁶ cells in a 10 cm polystyrene dish or 2×10⁴ 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 $2 \times 10^5$ 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 $0.3 \times 10^6$ cells in a 6 cm polystyrene dish or $1 \times 10^4$ 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
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 $1 \times 10^4$ 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 MgCl₂, 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 H₂O, 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 NaHCO₃) 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 immunoprecipitation 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 NaHCO₃) 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 MgCl₂, 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 \( \geq 4 \) 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 \( \mu \)M 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 \( \mu \)M 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 MnCl₂ 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 \( \ln(\text{FPKM}) \geq 1 \) 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 $6 \times 10^5$ transduced cells were scraped from the dish surface in ice-cold PBS and collected by centrifugation ($525 \times g$, 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 ($21,130 \times g$, 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, 450-1083) 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.