Summer 8-15-2017

Brain Enriched microRNAs Open the Neurogenic Potential of Adult Human Fibroblasts

Daniel Gene Abernathy
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Developmental Biology Commons, and the Neuroscience and Neurobiology Commons

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Developmental Regenerative and Stem Cell Biology

Dissertation Examination Committee:
Andrew S. Yoo, Chair
David H. Gutmann
Bob Mecham
Samantha Morris
Shelly Sakiyama-Elbert

Brain Enriched microRNAs Open the Neurogenic Potential of Adult Human Fibroblasts

by

Daniel G. Abernathy

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2017
St. Louis, Missouri

Table of Contents

iii

List of Figures and Tables

Figures:

<u>Chapter 3</u>

Supplementary Tables:

__Acknowledgments__

This work would not have been possible without the guidance, support and constant encouragement provided by my mentor, Andrew Yoo. I could always walk through the open door of his office and excitedly discuss any aspect of research - the excitement generated in each of these meetings permeated every experiment contained within this body of work.

Thank you to my thesis committee - Dr. Bob Mecham, Dr. Shelly Sakiyama-Elbert, Dr. David Gutmann and Dr. Samantha Morris for the constructive feedback that helped to shape my thesis work.

Thank you to the Department of Developmental, Regenerative and Stem Cell Biology program for the unique friendliness I found among colleagues and co-workers that is not present in many scientific departments.

I would like to thank members of the Yoo lab who have contributed all forms of feedback during my thesis work. The diversity of opinions, approaches, and ideas was integral to my scientific training. Thank you especially to WooKyung Kim and Matthew McCoy for their scientific contributions to this project.

Outside of the laboratory, a collection of friends and family has provided unending encouragement and welcome distractions that I am very grateful for. In particular, I would like to thank Aaron, Zeina, Liam, Larry, Angela, Katherine, Bryan and Greg.

To my grandfather, Barry, your willingness to teach, to share the joys of discovery and to provide a lab where anyone was welcome, has served, and will serve, as my model for what scientific research can and should be.

To my parents, Sandy and Tracy, I cannot thank you enough for providing the most enriching childhood a budding scientist could ask for. Growing up surrounded by books, ideas, honesty, and optimism cultivated a curiosity and steadfast desire to learn that paved the way to my scientific ambitions. I would also like to thank my sister, Amanda, whose seemingly endless drive and creativity has provided a source of joy and inspiration throughout my life.

Finally, to the most important thing I discovered during my thesis research, my wife Caitlin, thank you for your enduring compassion, encouragement, kindness, joy and love. I cannot thank you enough for the clear happiness you bring into my life each and every single day.

<div align="right">Daniel G. Abernathy</div>

Washington University in St. Louis

August 2017

ABSTRACT OF THE DISSERTATION

Brain Enriched microRNAs Open the Neurogenic Potential of Adult Human Fibroblasts

By

Daniel G. Abernathy

Doctor of Philosophy in Biology and Biomedical Sciences

Developmental, Regenerative and Stem Cell Biology

Washington University in St. Louis, 2017

Dr. Andrew S. Yoo, Chair

The seemingly limitless capacities of mammals to sense, respond, and manipulate their environments stems from their structurally and functionally diverse nervous systems. Establishing these complex behaviors requires the integration of many biological phenomena including, morphogenetic gradients, cell-cell signaling, transcriptional networks, cell migration and epigenetic gene regulation. As mammalian development progresses, these pathways coordinate the production of highly specialized neuronal and glial cells, that connect and communicate with another in an even more complex manner. While evolution has shaped a multitude of pathways to produce numerous favorable traits, it has also created an intricate system vulnerable to disease. The loss of different types of neurons, each responsible for specialized biochemical communications within the brain and spinal cord, results in a wide

variety of neurological and neurodegenerative diseases. Unfortunately, many of these diseases are uniquely human and cannot be wholly studied in model organisms such as *Mus musculus* or *Drosophila melanogaster*. Further, their location and absolute necessity precludes isolation directly from patients. Fortunately, advances in our understanding of genetic pathways that specify neuronal cell fates during development have enabled the directed differentiation of embryonic and induced pluripotent stem cells (iPSCs) into specific neuronal subtypes. This knowledge has been further leveraged to directly convert (reprogram) non-neuronal somatic cells into neurons, bypassing the induction of pluripotency. Specifically, ectopically expressing small non-coding microRNAs (miRNAs), miR-9/9* and miR-124 (miR-9/9*-124), with transcription factors in human adult fibroblasts is sufficient to generate functionally mature neuronal subtypes. These direct conversion modalities may prove invaluable in the study of late-onset neurodegenerative diseases, as the original age of human fibroblasts is maintained in converted neurons in contrast to the cellular rejuvenation observed in iPSCs. However, little is known about the epigenetic and molecular events that accompany direct cell-fate conversion limiting the utility of these features. Further, the capacity of miRNAs alone to overcome cell fate barriers has largely been unexplored. Within this thesis I provide mechanistic insights into the cell-fate pioneering activity of miR-9/9*-124. These results demonstrate that miRNAs induce remodeling of chromatin accessibilities, DNA methylation and the transcriptome leading to the generation of functionally excitable neurons. Surprisingly, during neuronal reprogramming, miR-9/9*-124 opens neuronal gene loci embedded in heterochromatic regions while simultaneously repressing fibroblast loci, revealing how miRNAs may overcome the cell-fate barrier that exists in human fibroblasts. These findings led to the discovery of a miRNA-induced permissive neurogenic ground state capable of generating multiple, clinically relevant neuronal subtypes. As such, I

show that the addition of motor neuron factors, ISL1 and LHX3, can function as terminal selectors to specify neuronal conversion to a highly enriched population of human spinal cord motor neurons. Altogether, the work contained within this thesis identifies miRNA-mediated epigenetic remodeling events underlying direct neuronal conversion of human fibroblasts and a modular platform capable of generating multiple, clinically relevant neuronal subtypes directly from patients.

Chapter 1: Genetic and Epigenetic Regulation of Neurogenesis by MicroRNAs

From:

MicroRNA-dependent Genetic Networks During Neural Development

Daniel G. Abernathy and Andrew S. Yoo

Abernathy, D.G., and Yoo, A.S. (2015). MicroRNA-dependent genetic networks during neural development. Cell Tissue Res *359*, 179-185.

Abstract

The specification of multipotent stem cells to defined cell types requires complex integrations of genetic pathways. During neurogenesis - the process of generating functional neurons from neural progenitor cells - many genetic and epigenetic processes play a critical role in regulating genes that control the proliferation and differentiation of neural progenitor cells. MicroRNAs (miRNAs) in particular, play an important role upstream of neurodevelopmental processes by modulating the expression of genes that alter chromatin states. During neural development, miRNAs, particularly brain enriched miR-9/9* and miR-124, regulate neurogenesis by targeting many key epigenetic pathways that affect the transcriptional accessibility of chromatin regions as well as mRNA processing. Importantly, ectopic expression of miR-9/9* and miR-124 in human fibroblasts induces the cells to adopt a neuronal fate. Direct conversion of a non-neuronal somatic cell to a functional neuron by brain-enriched miRNAs not only emphasizes the essential role of miRNAs in neurogenesis, but also provides new strategies for modeling human neurodegenerative diseases and regenerative medicine.

Key words: neurogenesis, epigenetics, miRNA, reprogramming, REST complex, BAF complex, PTBP

Introduction

Development of the mammalian nervous system results from the combinatorial action of morphogenetic gradients, cell-cell signaling, transcriptional networks, and cell migration (Götz and Huttner, 2005). The output of this intricate regulatory network is a collection of interconnected neural cells comprised of glia and neurons. These two cell types can be

subdivided further into distinct, highly specialized cells. In addition to developmental programs governed by transcription factors, non-coding RNAs facilitate neural fate acquisition (Cao et al., 2006; Cochella and Hobert, 2012; Pauli et al., 2011; Sun et al., 2013). The most widely characterized are small 20-24 nucleotide long microRNAs (miRNAs) that coordinate gene expression at the post-transcriptional level (Bartel, 2009; Krol et al., 2010; Pasquinelli, 2012) through translational inhibition and mRNA decay (Bartel, 2009; Bazzini et al., 2012; Djuranovic et al., 2012; Pasquinelli, 2012). These short stretches of RNA can have marked effects on gene networks as a single miRNA can target hundreds of mRNAs (Bernstein et al., 2003; Chi et al., 2009; Peter, 2010). Furthermore, multiple miRNAs can target a single transcript, dramatically increasing the effect on single gene expression (Wu et al., 2010). These properties enable miRNAs to act in both feedforward and feedback loops (reviewed in Ebert and Sharp, 2012) in order to establish developmental transitions, cell fate switches, and to refine gene expression.

miRNAs have been shown to be an integral part of transcriptional networks that drive developmental programs. The importance of miRNAs in many aspects of neural development has been reviewed elsewhere (Cao et al., 2006; Cochella and Hobert, 2012; Pauli et al., 2011; Sun et al., 2013) and in this issue. Here, we discuss studies that have increased our understanding of the activities of miRNAs during neuronal differentiation. We then focus on how genetic programs rely on specific miRNAs to reinforce transcriptional programs during neural development through feedback and feedforward genetic networks. Identification of genetic pathways from neural developmental studies has led to the ability to differentiate neurons from embryonic and induced pluripotent stem cells and recently, reprogram non-neuronal cells into neurons (reviewed in Morris and Daley, 2013), demonstrating the necessity to understand how miRNAs contribute in cell fate decisions during neural development.

3

miRNAs in Development

The major path to a mature miRNA begins in the nucleus where RNA polymerase II transcribes a primary miRNA transcript (pri-miRNA) that ranges in length, but is generally around several hundred base pairs (Lee et al., 2002). The pri-miRNA is processed down to a 60-100 nucleotide precursor hairpin (pre-miRNA) by Drosha and its cofactor DGCR8 (Bartel, 2009; Krol et al., 2010; Pasquinelli, 2012). The intermediate pre-miRNAs are transported into the cytoplasm by Exportin 5 where they are cleaved into their mature 19-25 nucleotide form by Dicer. Mature miRNAs are loaded into the RNA induced silencing complex (RISC) that, in complex with Argonaute (Ago) proteins, typically targets the 3'UTR of mRNAs. This leads to either translational repression or mRNA degradation (Bartel, 2009; Bazzini et al., 2012; Djuranovic et al., 2012; Pasquinelli, 2012).

Initial investigations into the role of miRNAs in development were based on disruption of miRNA biogenesis. With the observation that whole mouse embryo Dicer knockouts die at E7.5, many groups focused on generating tissue-specific Dicer deletion strains using Cre expression (Bernstein et al., 2003). Although the role of specific miRNAs acting through canonical miRNA biogenesis cannot be assessed using this strategy, it has revealed interesting facets of miRNAs in general and their role during neural development. Notably, proper migration, differentiation, and integration of neurons during neurogenesis requires mature miRNAs (Bernstein et al., 2003; Kawase-Koga et al., 2009; Makeyev et al., 2007; Volvert et al., 2012). Deleting Dicer in the telencephalon during neural development using a FoxG1-Cre line resulted in increased apoptosis culminating in a thinner cortical plate (Makeyev et al., 2007). Although there appears to be no effect on differentiation, the stereotyped organization of the cortex is disrupted. Further refining Dicer ablation in the cortex through the use of a Emx1-Cre driver yields analogous results, but

cortical defects appear earlier in development (De Pietri Tonelli et al., 2008). Similarly, deleting Dicer in neural progenitors using Nestin-Cre results in increased apoptosis, reduced cortical thickness, and increased ventricular volume (Kawase-Koga et al., 2009).

FoxG1 and Emx1 are restricted rostrally with Emx1 being the most cortical specific, yet in terms of cortical defects, these two Cre-drivers lead to earlier cell death and premature neuronal differentiation when compared to the Nestin-Cre driver. This is surprising since Nestin is expressed in all neural progenitors. Kawase-Koga and Sun attribute this difference to the timing of Dicer deletion, as opposed to cell types. Delineating the miRNAs responsible for the myriad of phenotypes observed in Dicer knockouts has previously been technically unfeasible. Now, we are able to examine specific miRNA-target networks at a global level, which has facilitated the dissection of miRNA-dependent biological processes.

Identification of miRNAs enriched in the nervous system

Advances in sequencing technology within the past 10 years have greatly increased the detectability of specific miRNAs and their targets within a select population of cells. These techniques have also shifted several commonly accepted notions of miRNA-mediated regulation by demonstrating binding outside the 3'UTR (Lee et al., 2009; Lytle et al., 2007; Ørom et al., 2008), localization of functioning miRNAs in the nucleus (Buckley et al., 2012; Jeffries et al., 2011; Khudayberdiev et al., 2013) and lastly bypassing the requirement of the 5' seed regions (Boudreau et al., 2013; Chi et al., 2012; Helwak et al., 2013). Collectively, these studies reveal a wider network of target genes and further emphasize the importance of miRNAs in gene regulation.

Identifying a collection of miRNAs and their targets has been accomplished primarily through target prediction algorithms, crosslinking techniques such as HITS-CLIP and PAR-CLIP, Ago-HITS-CLIP, and miRNA microarrays (Chi et al., 2009; Hafner et al., 2010; Liu et al., 2004; Ule et al., 2005). Each of these tools generates hundreds of possible miRNAs and targets that must be further validated. Nonetheless, they have revealed changes in miRNA expression as the brain develops and have identified many targets that mediate developmental transitions and neuronal function, including widely characterized miRNAs, miR-124, miR-9, miR-125, miR-132 and the let-7 families (Cochella and Hobert, 2012). In addition, these techniques can be applied to post-mortem human samples thus enabling the identification and validation miRNAs-mRNA interactions within the human brain (Boudreau et al., 2013), allowing comparisons of miRNA expression and target recognitions in different species.

Boudreau et al. observed a marked difference in the conserved binding sites between mouse and human brains (Boudreau et al., 2013). While the authors note the methods used to collect and determine the mouse data could have caused this difference, others have hypothesized that miRNAs are strong contributors to brain evolution and cognition (Bentwich et al., 2005; Hu et al., 2011; Lukiw, 2012) suggesting that the difference between mouse, primate and human miRNA expression may have biological meanings.

Further refinement of miRNA and target expression analysis through the expression of tagged-Ago proteins has enabled miRNA profiling in specific neuronal subtypes. Given the potent effects miRNAs have on both neurogenesis and neuronal function, determining the cell specific miRNA expression profile within neurons is fundamental to understanding their development (Hobert et al., 2010; Sempere et al., 2004). He et al. characterized the miRNA expression profile of five neuronal subtypes within the brain and revealed that glutamatergic,

GABAergic, and subclasses of GABAergic neurons within the neocortex and cerebellum, respectively, contain distinct but overlapping miRNA expression profiles (He et al., 2012). For example, parvalbumin (PV) and somatostatin (SST) expressing GABAergic interneurons overlap in miR-124 and miR-9 expression while SST interneurons express miR-187 and miR-551b at a level 6-fold higher than PV interneurons. Similarly, Camk2α neurons in the neocortex express miR-296 and miR-130b significantly more than Gad2 neurons from the same neuronal compartment. It is important to note that over 100 miRNAs were shown to have significant differential expression between sub-type comparisons. Interestingly, the authors also observed cell-type specific miRNA strand selection from precursors within different neuronal subtypes which could provide alternative target selection. So far, the exact role of miRNAs enriched in different neuronal subtypes largely remains unknown.

Function of miRNAs in Early Neural Development

Once dorsal ectoderm is specified to become neuroectoderm through BMP and activin inhibition and increased Wnt and FGF signaling the central nervous system starts to take shape (Stern, 2005). The appearance of the neural plate and formation of the neural folds dorsal to the neural plate provides the framework for cell specification. Positional information along rostro-caudal and dorso-ventral axis of the neural tube has been shown to demarcate future cell fate largely through the action of ventral Shh, dorsal BMP and rostro-caudal retinoic acid gradients (Jessell, 2000; Stern, 2005). The mechanism that establishes progenitor domains along these axis relies on a combinatorial code of morphogenetic gradients and transcription factor networks (for review see Jessell, 2000). Many components of the transcriptional and signaling components necessary for proper motor neuron pattern formation have been well characterized. Recent evidence suggests that miRNAs also play important roles in regulating the balance between

7

cross-repressive transcription factors within the neural tube. miR-196, miR-17-3p and miR-9 appear to provide boundary information within the developing spinal cord by refining transcription factor networks during motor neuron specification (Asli and Kessel, 2010; Bonev et al., 2011; Chen et al., 2011). Interestingly, the biological role of each of these miRNAs during neurogenesis differs spatially and temporally along the developing embryos' axis, suggesting that miRNAs can have diverse functions within neurons depending on the milieu of mRNAs they encounter. For instance, the combinatorial action of multiple miRNAs can result in stronger gene repression leading to greater changes within genetic networks (Wu et al., 2010). Therefore, regionally specific miRNAs could coordinate specific programs of neural development. Further studies integrating spatially and temporally expressed miRNAs within the nervous system may provide insights into the generation of neuronal diversity.

Despite different biological outputs, miRNAs both inside and outside of the nervous system have been consistently identified has key mediators of both feedforward and feedback loops. Here, the repressive activity of miRNAs can act both to activate or inhibit gene expression. The redundancy built in these systems is thought to provide robustness to genetic networks (Ebert and Sharp, 2012). Below, we review evidence demonstrating the importance of miRNAs in enforcing genetic programs through feedforward and feedback loops during cortical neuron development, one of the most characterized neural fate programs.

Feedback Inhibition During Cortical Neurogenesis

miR-9, one of the most widely studied neuronal miRNAs participates in a wide variety of neuronal differentiation programs (Coolen et al., 2013; Kapsimali et al., 2007). It plays a vital role in the differentiation of neural progenitors into post-mitotic neurons by modulating a host of

targets (reviewed in Coolen et al., 2013). Within the developing forebrain, the orphan nuclear receptor TLX maintains neural stems cells in a proliferative state through activation of Wnt/ß-catenin and recruitment of histone deacetylases that repress the loci of p21 and PTEN (Qu et al., 2010; Sun et al., 2007). Suppression of TLX expression in neural stem cells by miR-9 leads to decreased proliferation, premature differentiation, and migration toward the cortical plate (Zhao et al., 2009). It is thought that during normal development this repressive interaction regulates the differentiation of cortical progenitors. Premature differentiation is prevented at least in part, by the repression of miR-9 by TLX (Zhao et al., 2009). Others have also shown that processes mediated by miR-9 can be attributed to several targets within the cortical progenitor pool (Shibata et al., 2011). This negative feedback loop is further strengthened by additional miRNAs and targets of miR-9. For example, TLX expression is repressed during neurogenesis by let7b (Zhao et al., 2010). miR-137, also highly expressed in the brain, targets LSD1, a co-repressor that acts in conjunction with TLX to mediate transcriptional repression (Sun et al., 2011). This indicates that the effects of these miRNAs, when expressed together, may have a more dramatic phenotype than previously reported. The negative feedback loops present during cortical neurogenesis are depicted in Fig1. It is interesting to note that in each of these studies, over-expression of these miRNAs individually caused remarkably similar phenotypes and the expression of a single target often mitigated these effects. In addition, it is unknown if this regulatory relationship can be extended outside the cortical compartment. While miR-9 is expressed throughout the vertebrate nervous system (Lagos-Quintana et al., 2002), the pool of target mRNAs and co-repressive miRNAs could vary between neuronal populations. The ability of pan neuronal miRNAs to function differently within different neuronal populations remains an open question.

In the case of cortical neuron specification and differentiation, miRNAs clearly integrate into known genetic pathways, providing additional information to the developing nervous system. However, their relationship to one another and the network as a whole remains unclear. If miR-9, miR-137, and let7d have overlapping expression patterns, it is possible that they share common targets and their effect on cell fate is more pronounced than previously reported. Such synergy has been clearly demonstrated with the brain enriched miRNAs miR-9/9* and miR-124 (miR-9/9*-124) in chromatin regulation as described below.

miRNAs and Epigenetic Regulation

During development, ATP-dependent chromatin remodeling complexes acquire new activities by altering their subunit compositions (Ronan et al., 2013). Differentiation of neural progenitors into neurons is accompanied by a switch in the subunit composition of the mammalian SWI/SNF (BAF) chromatin remodeling complex. Specifically, subunits present in neural progenitors are exchanged with the neuron specific BAF45B/C, CREST, and BAF53b subunits (Lessard et al., 2007; Staahl et al., 2013; Wu et al., 2007). miR-9/9* and miR-124, regulate this switch by targeting the 3' UTR of BAF53a and prolonged expression of BAF53a *in vivo* represses BAF53b expression suggesting that miR-9/9*-124 expression initiates a feedforward loop that culminates in the swap of subunits within the BAF complex (Yoo et al., 2009). The onset of this regulatory cascade is instigated through a negative feedback loop between miR-9/9*-124 and the RE1 silencing transcription factor (REST/NRSF) complex (Conaco et al., 2006; Packer et al., 2008).

Figure 1
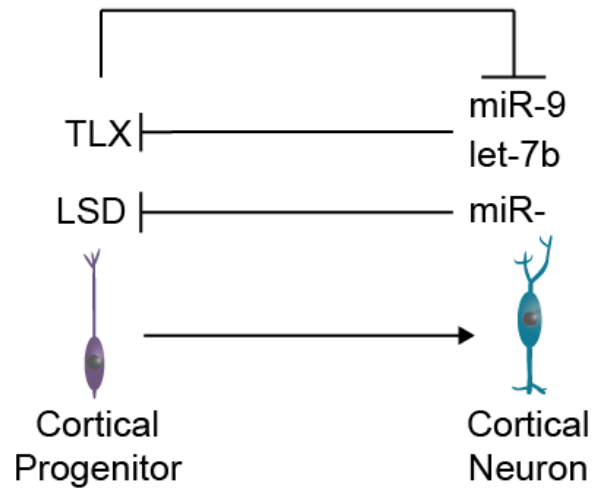
Figure. 1 miRNA-dependent negative feedback loops in cortical neurogenesis. As cortical progenitors differentiate into cortical neurons the expression of miR-9 and miR-137 is induced leading to the direct repression of the orphan nuclear receptor TLX. The expression of miR-9 and miR-137 is normally repressed in cortical progenitors by TLX. This negative feedback loop is further reinforced by the repression of LSD1, a co-repressor that acts in concert with TLX, by miR-137.

REST/NRSF was initially identified as a neuronal repressor that prevents the expression of neuronal genes outside the nervous system (Chong et al., 1995; Schoenherr and Anderson, 1995). This repressive activity is attained by recruiting other co-repressors such as co-REST, SCP1, MeCP2, and histone deacetylases (Andrés et al., 1999; Lunyak, 2002; Roopra et al., 2000; Yeo, 2005). As neural progenitors differentiate into post-mitotic neurons, REST/NRSF-mediated repression is relieved upon the inhibition of REST/NRSF expression (Ballas et al., 2005). Interestingly, REST/NRSF and co-REST have been shown to be regulated by miR-9 and miR-

9*, respectively (Packer et al., 2008). The sharp contrast in REST/NRSF expression between neuronal and non-neuronal cells is maintained through mutual repression between miR-9 and REST/NRSF. This negative feedback loop contributes to the neuronal specific expression of miR-9/9*-124 as all the active loci of miR-9/9* and miR-124 are occupied by REST/NRSF in non-neuronal cells (Conaco et al., 2006). In addition, SCP1, a component of the REST/NRSF silencing complex that facilitates neuronal silencing, was shown to be targeted by miR-124 (Visvanathan et al., 2007).

The loss of REST/NRSF-mediated repression and concomitant upregulation of neuron specific gene expression is also accompanied by neuron specific splicing events (Li et al., 2007). Makayev et al., observed that when miR-124 was overexpressed in two neuroblastoma cell lines they adopted a neuron-like morphology and contained neuron-specific alternatively spliced transcripts (Makeyev et al., 2007). They showed that the downregulation of the splicing repressor PTBP1 by miR-124 leads to neuronal specific splicing events including the production of PTBP2, a neuron-specific splicing regulator that is normally excluded from non-neuronal cells by PTBP1 (Boutz et al., 2007). Here, the reciprocal expression of PTBP1 and PTBP2 is initiated by miR-124 creating both a feedforward and feedback loop. A summary of the miR-9/9*-124 dependent regulatory circuit is shown in Figure 2.

Figure 2
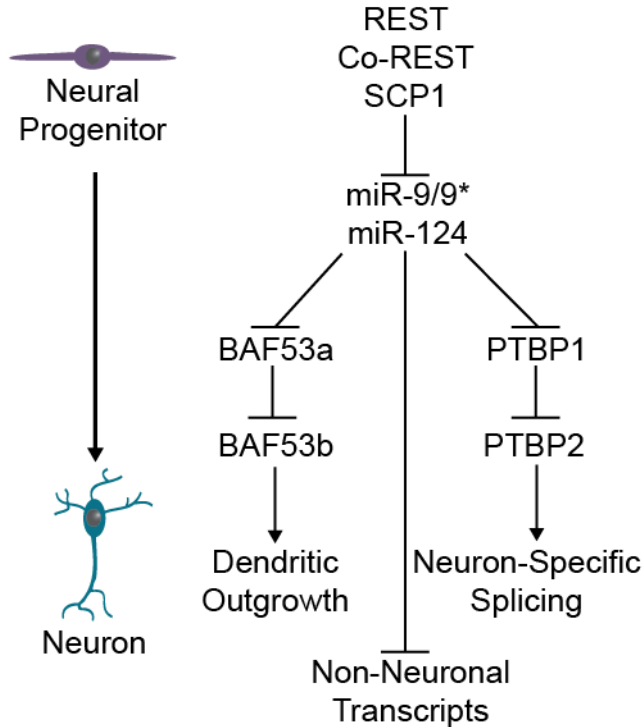
Figure 2. miR-9/9*, miR-124 and REST coordinate neurogenesis through multiple positive and negative feedback loops. REST silences many neuronal genes including miR-9/9* and miR-124. At the onset of neural differentiation, miR-9/9* and miR-124 are upregulated leading to the repression of their targets, REST, BAF53a and PTBP1. REST, co-REST, and SCP1 repress miR-9/9* and miR-124 which creates a feedback loop allowing for distinct, reciprocal expression patterns. BAF53a normally inhibits BAF53b, the relief of this repression by miR-9/9* and miR-124 enables BAF53b expression and formation of neuron specific BAF chromatin remodeling complex which is important for proper dendritic outgrowth. This feedforward loop is distinct from PTBP1's repression by miR-124. Here, neuron specific splicing is accomplished by repression of one splicing repressor, PTBP1, which directly represses the expression of another splicing regulator PTBP2. During neurogenesis, the expression of miR-9/9* and miR-124 that

13

are activated at least in part by dismissal of REST binding leads to a cascade of genetic events that strongly re-enforces the neurogenic program

Negative feedback loops are continually being characterized between miRNAs and their repressors, and the sum of these feedback loops can have dramatic effects when the balance is sharply tipped in one direction. For instance, ectopic expression of miR-9/9*-124 directly converts human fibroblasts into post-mitotic neurons (Figure 3) (Yoo et al., 2011). While neither miRNA alone is capable of reprogramming the human fibroblasts, the neuronal conversion can be enhanced with the addition of neuronal transcription factors, revealing a synergy between multiple miRNAs and pro-neural transcription factors. The observation that miRNAs alone are capable of switching cell fate demonstrates their potency in regulating neuronal gene networks. Others have also taken advantage of manipulating the genetic circuit involved in neuronal specific splicing to achieve neuronal reprogramming (Xue et al., 2013).

Figure 3
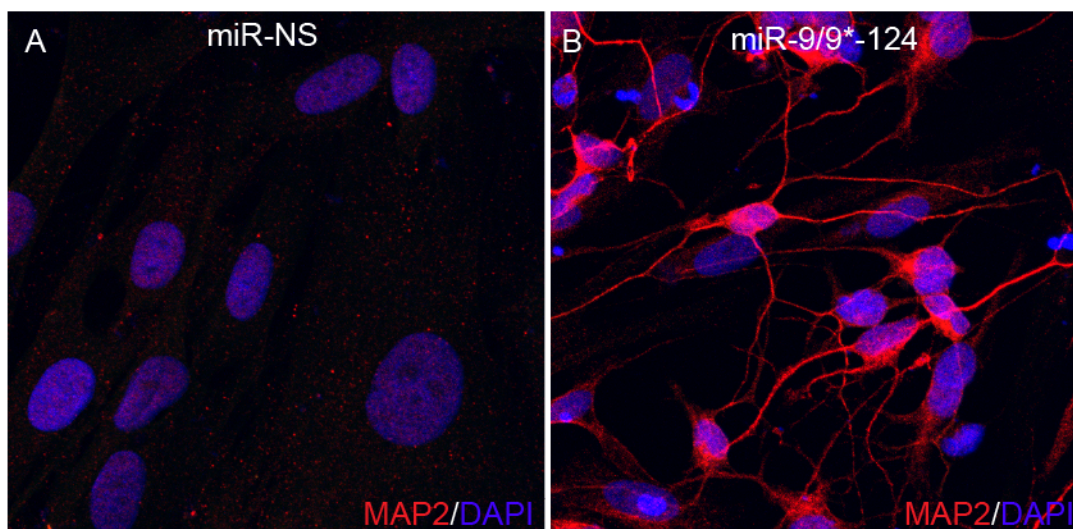
Figure 3. Neuronal conversion of primary human dermal fibroblasts mediated by miR-9/9*-124. Pictures depict (A) human neonatal fibroblasts and (B) converted cells immunostained with a pan-neuronal marker, MAP2 (adapted from Yoo et al., 2011).

Asymmetric Cell Division and Neuron Specific miRNA Activity

One of the large outstanding questions in mammalian neurogenesis is how asymmetric division of neural stem cells is regulated. Recently Schwamborn et al, uncovered that TRIM32, a TRIM NHL protein, is asymmetrically inherited during neural progenitor division (Schwamborn et al., 2009). The daughter cell that does not receive TRIM32 remains a neural progenitor while the daughter inheriting the TRIM32 protein differentiates into a neuron. The mechanism behind this differentiation is due to two separate activities performed by TRIM32. As a ubiquitin ligase TRIM32 tags c-Myc for degradation. Repression of c-Myc leads to cell cycle exit and further participates in neural induction. Secondly, TRIM32 complexes with Ago1 leading to enhanced activity of several pro-neural miRNAs. One of these miRNAs, let7a, contributes to neural differentiation although not to the extent of TRIM32. Others have also shown let7a has both anti-proliferative effects and is upregulated during neurogenesis (Peng et al., 2008; Rybak et al., 2008). The effect(s) of TRIM-32 on other pro-neural miRNAs remains unexplored.

Conclusion

Much like transcription factors, miRNAs can exist in feedback loops with their repressors or activators enabling binary switches and signal amplification, respectively. The cumulative effect of each regulatory node described thus far results in the vast array of highly specialized cells within the nervous system that connect and communicate in an even more complex manner.

Our understanding of the development of specific types of neurons has been limited to the transcription factors necessary for the development of these populations. The advent of new technologies enables mechanistic studies focusing on the function of specific miRNAs during the development, and subsequent maintenance of neuronal subtypes. Further understanding of the genetic networks that connect each of these pathways within a single cell–type will greatly facilitate our understanding of neural development. To date, much progress has been made in understanding miRNAs and target interactions in neurogenesis. The results described here are a group of select studies that demonstrate both the importance of miRNA-dependent genetic networks and the need to connect this accumulated knowledge.

Author Contributions

DGA is responsible for the design and content of this review. ASY contributed to the overall organization and content of this work.

Chapter 2: MiR-9/9*-124 Induce Widespread Transcriptomic and Epigenetic Changes During Neuronal Transdifferentiation

Adapted from:

MicroRNA-induced Epigenetic Remodeling During Direct Cell-Fate Conversion of Adult Human Fibroblasts

Daniel G. Abernathy, Wookyung Kim, Matthew J. McCoy, Allie Lake, Rebecca Ouwenga, Xiaoyun Xing, Daofeng Li, Hyung Joo Lee, Robert O. Heuckeroth, Joseph D. Dougherty, Ting Wang, Andrew S. Yoo

Abstract

Neuronal microRNAs, miR-9/9* and miR-124 (miR-9/9*-124) direct cell-fate conversion of adult human fibroblasts to post-mitotic neurons and with additional transcription factors enable the generation of discrete neuronal subtypes. Previously, the molecular events underlying the neurogenic switch mediated by microRNAs during neuronal reprogramming were unknown. Here, we systematically dissected the neurogenic state induced by miR-9/9*-124 expression alone. We found that miR-9/9*-124 surprisingly stimulate reconfiguration of chromatin accessibility, DNA methylation and mRNA levels, leading to the generation of functionally excitable miRNA-induced neurons (miNs) that are yet uncommitted towards a particular subtype-lineage. Further subtype identity can be programmed through additional transcription factors. As such, we show ISL1 and LHX3 selectively commit conversion to a highly homogenous population of human spinal cord motor neurons (Moto-miNs). Taken together, our study reveals a modular synergism between microRNAs and transcription factors that allows lineage-specific neuronal reprogramming, providing a platform for generating distinct subtypes of human neurons.

INTRODUCTION

Understanding genetic pathways that specify neuronal cell fates during development has enabled directed differentiation of pluripotent stem cells to specific neuronal subtypes (Perrier et al., 2004; Wichterle et al., 2002). This knowledge has been further leveraged to directly convert (or reprogram) non-neuronal somatic cells into neurons via ectopic expression of pro-neural transcription factors (TFs) or neurogenic miRNAs with TFs (Mertens et al., 2016). These direct conversion modalities may prove invaluable in the study of late-onset neurodegenerative diseases, as the original age of human fibroblasts is maintained in converted neurons (Huh et al., 2016; Mertens et al., 2015) in contrast to the cellular rejuvenation observed in induced

18

pluripotent stem cells (Horvath, 2013; Miller et al., 2013). Interestingly, the miRNA-mediated reprogramming approach boasts high conversion efficiency in adult human fibroblasts, which may provide unique opportunities in modeling neurological disorders using patient derived neurons (Victor et al., 2014). However, despite the advantages of direct reprogramming, little is known about the epigenetic and molecular events that accompany direct cell-fate conversion.

MiRNAs regulate genetic pathways by binding to their target transcripts and repressing their expression (Pasquinelli, 2012). Target specificity is governed largely through short sequence complementarity within the 5' end of a miRNA enabling a single miRNA to target hundreds of mRNA transcripts (Boudreau et al., 2014; Chi et al., 2009). Moreover, a single mRNA can be targeted by multiple miRNAs, markedly enlarging the effect on single gene repression (Wu et al., 2010). The convergence of genetic controls by miRNAs towards a specific biological process is exemplified by miR-9/9* and miR-124 miRNAs activated at the onset of neurogenesis (Conaco et al., 2006; Makeyev et al., 2007). For example, miR-9* and miR-124 synergistically initiate subunit switching within BAF chromatin remodeling complexes (Staahl et al., 2013; Yoo et al., 2009) while separately repressing the neuronal cell-fate inhibitors REST, Co-REST and SCP1 (Packer et al., 2008; Visvanathan et al., 2007). These examples suggest that miR-9/9* and miR-124 target components of genetic pathways that antagonize neurogenesis to promote a neuronal identity during development.

Co-expressing miR-9/9* and miR-124 (miR-9/9*-124), with TFs enriched in the cortex and striatum directly converts primary adult human fibroblasts to cortical and striatal medium spiny neurons, respectively (Victor et al., 2014; Yoo et al., 2011). However, the same TFs without miR-9/9*-124 fail to trigger neuronal conversion (Victor et al., 2014; Yoo et al., 2011), suggesting that the miRNA-induced neuronal state is permissive to terminal selector TFs which, upon determination of a neuronal fate, initiate and advance mature subtype-identities (Stefanakis et al., 2015).

Here, we systematically investigated a miRNA-induced neuronal state in adult human cells. Longitudinal analyses of the transcriptome, genome-wide DNA-methylation and chromatin accessibilities revealed that miR-9/9*-124 induced extensive remodeling of the epigenome, including simultaneous activation of a pan-neuronal program and the reconfiguration of chromatin accessibilities. These changes precede the emergence of differentially methylated genomic regions. Because miR-9/9*-124 also led to the opening of genomic loci for multiple subtype-specific genes including established motor neuron markers, we postulated that motor neuron-enriched transcription factors would cooperate with miR-9/9*-124 to specify a motor neuron lineage. As such, we demonstrate that co-expressing TFs ISL1 and LHX3 along with miR-9/9*-124 generated a highly pure population of human spinal cord motor neurons. Taken together, these results demonstrate that miR-9/9*-124 opens the neurogenic potential of adult human fibroblasts and provides a platform for subtype-specific neuronal conversion of human cells.

RESULTS

Neuronal Conversion of Human Adult Fibroblasts With miR-9/9*-124 Alone

To dissect how miRNAs alone contribute to neuronal conversion, we first tested the ability of miR-9/9*-124 to convert primary human fibroblasts collected from multiple adult individuals from ages 22 to 68 into microRNAs-induced neurons (miNs). We transduced fibroblasts with lentivirus containing a doxycycline (Dox)-inducible promoter driving miR-9/9*-124 and BCL-XL expression (Victor et al., 2014) (Figures 1A and S1A). At 35 days post-transduction, we evaluated cell morphology and the expression of neuronal markers MAP2, TUBB3 and NEUN by immunocytochemistry (Figure 1B). Strikingly, miR-9/9*-124 alone converted 80% of the fibroblasts to neuronal cells displaying complex neurite outgrowth and neuronal marker expression (Figures 1B and 1C). The converted cells stained positive for

20

voltage-gated sodium channels SCN1A and Ankyrin G, which localized at axonal initial

segments with a characteristic polarized staining pattern (Figure 1D). The synaptic vesicle

marker SV2 displayed defined puncta along neurites, consistent with the adoption of a neuronal

fate (Figure 1D).

Figure 1
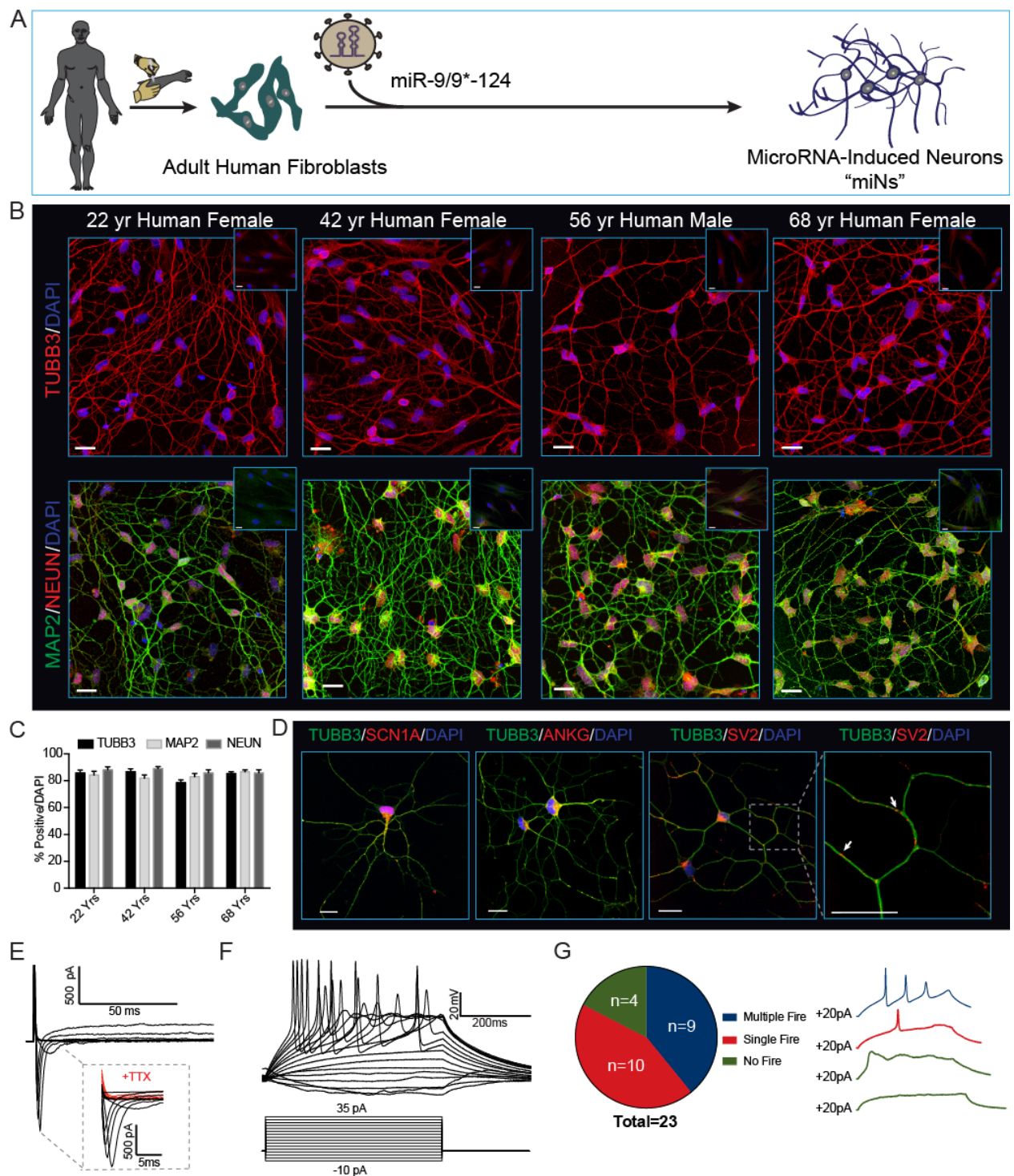
Figure 1. Direct Conversion of Young and Old Primary Adult Human Fibroblasts into Neurons via miRNA Overexpression.

(A) Experimental schema for miR-9/9*-124 mediated direct neuronal conversion.

(B) Adult human fibroblasts ectopically expressing miR-9/9*-124 for 35 days immunostained for the pan-neuronal markers TUBB3, MAP2 and NEUN. Insets represent starting fibroblasts co-stained as negative controls. Scale bars = 20μm.

(C) Quantification of TUBB3, MAP2 and NEUN positive cells over total number of cells (DAPI). For TUBB3 and MAP2, only cells with processes at least three times the length of the soma were counted. For NEUN, only cells with proper nuclear localization were counted. Data are represented as mean ± SEM. 22 Yr Female n = 238 cells, 42 Yr Female, n = 100 cells, 56 Yr Male n = 171 cells, and 68 Yr Female n = 216.

(D) Converted neurons display hallmark sodium channel (SCN1A), axonal initial segment (ANKG) (left) and synaptic vesicle (SV2) (right) staining patterns. Scale bars = 20μm.

(E) Representative traces of TTX-sensitive inward and potassium whole-cell currents.

(F) Repetitive AP waveforms in response to 500 ms current injections recorded from neurons converted in monoculture.

(G) Summary of firing patterns observed in 23 neurons recorded in current-clamp mode (left) and representative waveforms within each firing pattern recorded (right).

See also Figure S1.

Functional Properties and Stability of MiRNA-induced Neurons

To determine if miNs displayed membrane excitability, we performed whole-cell recording on miNs without glial co-culture. All cells recorded (23/23 cells) exhibited fast TTX-sensitive inward currents upon depolarization (Figure 1E, while 19/23 cells fired action potentials (APs) during current injections (9 cells fired multiple APs, 10 fired single APs)

(Figures 1F and 1G). Similar current (I)-voltage (V) curve relationships were observed between cells that fired multiple or single action potentials (Figure S1B). All miNs had a stable hyperpolarized resting membrane potential ranging from -52.7 to -84.5 mV with a mean value of -69.63 mV ± 2.15 mV (S.E.M.) (Figures S1C and S1D). There was no correlation between capacitance and firing patterns suggesting this membrane property is not as an accurate measure of neuronal maturation (Figure S1E). Together, this data indicates that miNs exhibit the membrane functionality of neurons.

The minimum duration of miRNA expression required for neuronal conversion was determined by inactivating the doxycycline-inducible promoter at 3-day intervals by Dox removal beginning at day 9 until reprogramming day 30 (Figure S1F). Loss of fibroblast identity and gain of neuronal identity was assayed by analyzing fibroblast-specific protein (FSP1) and MAP2 expression, respectively. Surprisingly, we saw a reduction in the number of FSP1-positive cells and the appearance of MAP2-positive cells after only 9 days of miRNA expression, however; efficient switching of a cell population from FSP1- to MAP2+ required 30 days of miR-9/9*-124 expression (Figure S1F). These data reveal the non-synchronous process of the neuronal conversion and temporal requirements for highly efficient neuronal conversion. The stability of miRNA-induced neuronal conversion was determined by following miNs for an additional 30 days after removing miR-9/9*-124 exposure (Figure S1G, top). The majority of miNs remained as post-mitotic neurons (marked by the absence of Ki67, a cell proliferation marker) expressing MAP2, TUBB3, NeuN and NCAM, in contrast to non-converted fibroblasts (Figure S1G, left panel), indicating that the morphological and protein expression changes that accompany miR-9/9*-124-mediated conversion of adult human fibroblasts are stable after 30 days of miR-9/9*-124 expression.

Transcriptional Profiling of miNs

To further explore the miR-9/9*-124-mediated neuronal output, we profiled the transcriptome of starting human adult fibroblasts and miNs after 30 days of neuronal conversion by RNA-Seq. We identified 2,692 differentially expressed genes (DEGs) in miNs representing 1,251 up-regulated and 1,441 down-regulated genes in comparison to fibroblasts (log fold change ≥ 2 ; adj.P-value <0.01) (Figure 2A). A robust down-regulation of fibroblast-specific genes (for instance, *S100A4*, *VIM*, *FBN1* and, *COL13A1*) was accompanied an enrichment of pan-neuronal genes including *MAP2*, *SCN1A*, *SNAP25*, *NRCAM,* and *NEFM* (Figure 2A, 2B top two traces). Analysis of top 10 gene ontology (GO) terms revealed that upregulated genes in miNs are primarily enriched with terms related to neuronal development and functionality (Figure 2C) while down-regulated genes associate with fibroblast functions (Figure 2C). Down-regulated genes also included key cell-cycle components (data not shown), consistent with the previous finding that miR-9/9*-124 expression in human fibroblasts caused rapid cell cycle exit without transitioning through a neural stem cell-like state (Yoo et al., 2011). Interestingly, neuronal subtype-specific genes such as *TH* (dopaminergic neurons), *GABBR2*, *GABR1*, and *GAD2* (GABAergic neurons), *CHAT* (cholinergic neurons), or *DARPP-32* (striatal medium spiny neurons), were not significantly enriched in miNs (Figure 2B, bottom two traces as examples). Overall, our transcriptome analyses show that miR-9/9*-124 induce a neuronal state characterized by the loss of fibroblast identity and the presence of a pan-neuronal gene expression program without commitment to a particular subtype.

Figure 2
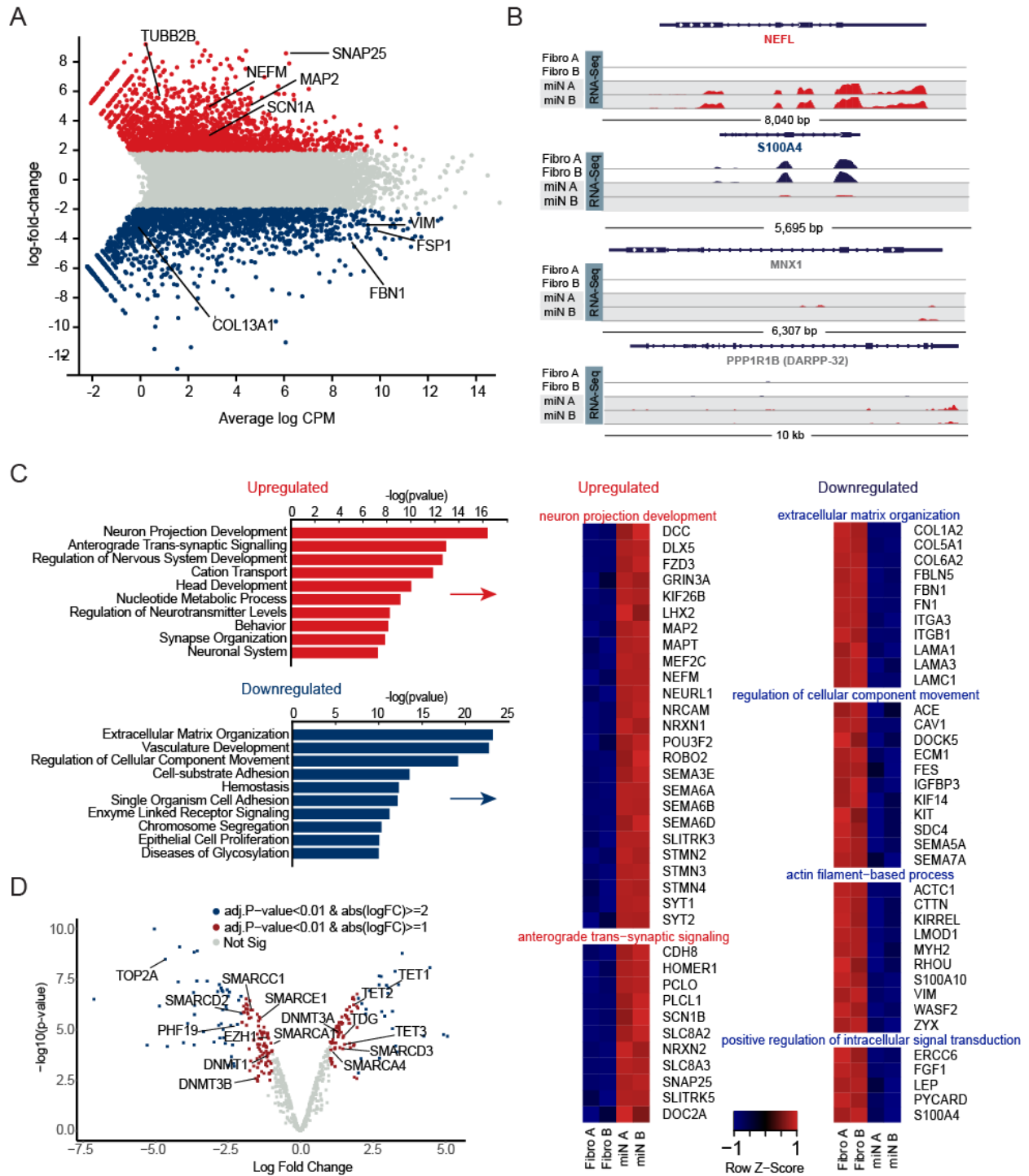
Figure 2. Gene Expression Profiling Reveals Pan-Neuronal Identity Induced by miRNAs Alone

(A) Genome wide expression analysis of miNs and starting fibroblasts by RNA-seq. Plot shows the relationship between average gene expression (logCPM) and log fold-change of miNs compared to fibroblasts. A selection of pan-neuronal and fibroblast-specific genes are highlighted in black text. Blue = fold change <-2 $\log_2$ p<.01 (more abundant in fibroblasts), grey = fold change >-2 $\log_2$<2 $\log_2$ p>.01, and red = fold change>2 $\log_2$ p<.01 (more abundant in miNs). FDR <.01.

(B) Representative genome browser snapshots demonstrating increased expression for a pan-neuronal gene (*NEFL*), loss of fibroblast marker gene expression (*S100A4*), and absence of neuronal subtype marker gene expression (*MNX1*, motor neuron marker; *DARPP-32*, medium spiny neuron marker).

(C) Gene ontology (GO) terms associated with genes upregulated in miNs (red) and GO terms associated with genes downregulated in miNs (blue). Right, genes that fall within top GO categories listed (top to bottom) in order of lowest to highest p-value.

(D) Volcano plot representing chromatin remodeling genes differentially expressed between fibroblasts and miNs. Blue dot, abs(logFC) > 2 and p < 0.01, red dot, abs(logFC) > 1 and p < 0.01, grey dot, no significant difference.

See also Figure S2.

Transcriptional Changes in Epigenetic Machinery

Epigenetic modifications can markedly affect gene expression and developmental programs (Cantone and Fisher, 2013). Our gene expression studies showed that when compared to fibroblasts, miNs had markedly altered expression of genes involved in DNA methylation, histone modifications, chromatin remodeling and chromatin compaction (Figures 2D and S2). For instance, the TET family of proteins, key mediators of DNA-demethylation (Wu

and Zhang, 2011) were upregulated along with the brain-enriched *de novo* DNA-methyltransferase DNTM3A (Lister et al., 2013), while DNMT3B (Okano et al., 1999) mRNA levels were reduced in miNs compared to fibroblasts (Figures 2D and S2). Transcripts encoding histones and histone variants were altered (Figure S2) suggesting that changes in histone composition may accompany neuronal conversion of human fibroblasts. Genes encoding chromatin remodelers important for neurogenesis like CHD5, CHD7 and components of the BAF chromatin remodeling complex were expressed at higher levels in miNs than in fibroblasts (Egan et al., 2013; Feng et al., 2013; Lessard et al., 2007) (Figures 2D and S2). Additionally, the main DNA topoisomerase 2 family member expressed in miNs is TOP2B, which replaces the non-neuronal TOP2A, a switch that has been observed during normal neuronal differentiation (Tiwari et al., 2012) (Figures 2D and S2). In sum, dynamic changes and switches within diverse epigenetic modifiers coincide with neuronal differentiation and appear to be recapitulated in direct neuronal conversion of fibroblasts by miR-9/9*-124.

Dynamic Regulatory Events During Neuronal Reprogramming

Because transcriptome profiling at day 30 only provided a snapshot of the functional output of neuronal reprogramming, we explored transcriptome dynamics by profiling intermediary timepoints (days 3, 6, 10, and 20) by RNA-seq. The Dynamic Regulatory Events Miner (DREM) (Schulz et al., 2012) reports 13 paths of co-regulated, differentially expressed genes during the first 20 days of neuronal conversion (Figure 3A). Combining DREM with predicted TF-gene binding interactions (Ernst et al., 2010) revealed several potential TFs associated with major regulatory events (bifurcations in each path; Figure 3A). Altogether, major regulatory events were observed before day 10, suggesting genetic networks are established within 10 days of miR-9/9*-124 expression. Thereafter, the directionality of gene expression stays the same but transcript levels markedly change. This transcriptional maturation over time may explain why the acquisition of functional neuronal characteristics requires 30 days of

culture. GO analyses of each path revealed enrichment of neuronal terms in the most

upregulated path, whereas down-regulated paths were enriched for cell cycle and extracellular

matrix-related terms (Figure 3B), consistent with results for day 30 transcriptome profiling

(Figure 2C). The transcriptome switch from a fibroblast to a neuronal program was seen as

early as day 10 of miR-9/9*-124 expression (Figure 3C), while genes associated with synaptic

functionality (e.g. HOOK1) are activated at a later time point (day 20, Figure 3C).

Finally, there were no significant changes in ASCL1 or SOX2, TFs that have been used

to reprogram somatic cells into neurons (Niu et al., 2013; Pang et al., 2011). suggesting miR-

9/9*-124-induced neuronal conversion activates a neuronal program through mechanisms

distinct from those previously reported.

Figure 3
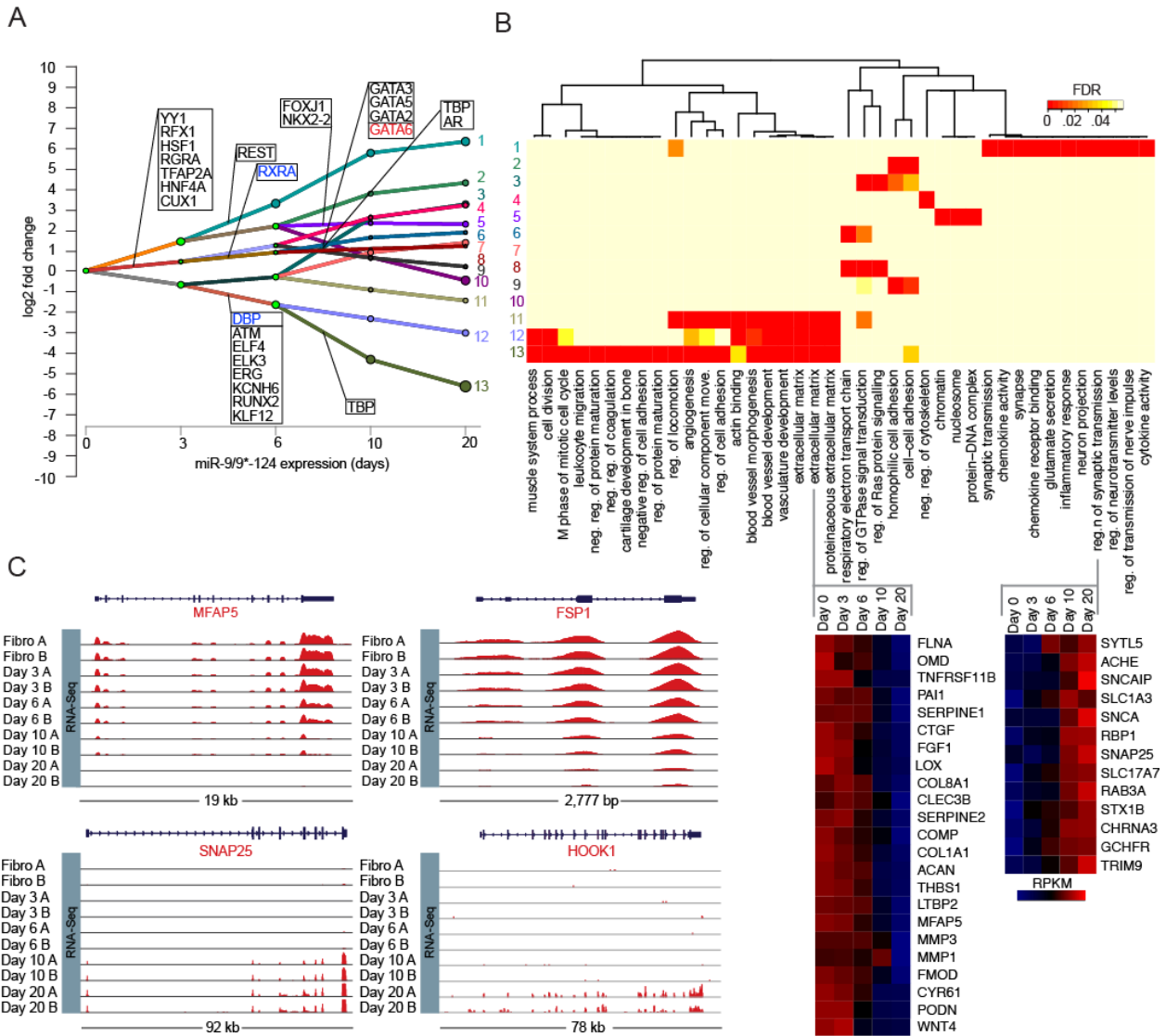
Figure 3. Time Series Transcriptome Analysis Reveals Early Dynamic Gene Expression Changes Followed by a Stable Transcriptome Switch

(A) Dynamic Regulatory Event Miner (DREM) analysis reveals multiple paths of co-regulated genes emerge over time (days).

(B) GO terms enriched in each path (top) and heat maps of genes within the extracellular matrix and regulation of synaptic transmission terms (bottom). Z-score normalized RPKM.

(C) Representative genome browser snapshots demonstrating the time-dependent loss of fibroblast gene expression (top), emergence of pan-neuronal marker gene expression, and synaptic component expression (bottom).

DNA Methylation Profiling of miNs

After observing numerous changes in DNA methylation machinery we assessed genome-wide DNA methylation at an early (day 10), intermediate (day 20), and late stage (day 30) of neuronal reprogramming by combining methylated DNA immunoprecipitation sequencing (MeDIP-seq) and methylation sensitive restriction enzyme sequencing (MRE-seq; Figure 4A) (Zhang et al., 2013). No significant changes in DNA methylation were detected at day 10 between cells that were exposed to miRNAs and non-treated control (ctrl) cells. In contrast, we identified 1,540 differentially methylated regions (DMRs) at day 20 (miN day 20 vs. Ctrl day 20) that overlap with DMRs at day 30 (miN day 30 vs. Ctrl day 30; Figures 4B and 4C). The difference in DNA methylation at day 30 between treated and control cells was more dramatic than changes observed at day 20 with most DNA regions undergoing demethylation. Gene ontology (GO) analysis of Mouse Genome Informatics (MGI) expression using the GREAT (McLean et al., 2010) tool to characterize the top overlapping DMRs in miNs, enriched only for neuronal tissue developmental processes (Figure 4D). Interestingly, these differentially methylated regions undergo changes at TS15 in mouse development (equivalent to mouse E9/10), which is when *miR-9* expression is first detected in developing mouse telencephalon (Shibata et al., 2008). Two examples of top DMRs within genes important for neuronal development and function are shown in Figure 4E (Gehrke et al., 2010; Vadhvani et al., 2013). Nearly 64% of DMRs were located in introns and 28% of DMRs are in intergenic regions (Figure 4F). Comparison of data from day 30 demethylated and methylated DMRs and RNA-seq revealed 882 differentially expressed genes in total. GO analysis of top demethylated DMRs associated with upregulated genes (> 2.5 logFC) was enriched for neuronal terms. In contrast,

top DMRs associated with down-regulated genes (< 2.5 logFC) did not match GO terms involved in neural development (Figure 4G; data not shown). Furthermore, since miR-9/9*-124 expression quickly induces cell cycle exit, these changes in DNA methylation must necessarily occur via active processes, such as those catalyzed by TET and TDG enzymes (Kohli and Zhang, 2013), rather than via failure to remethylate DNA after DNA replication.

Figure 4
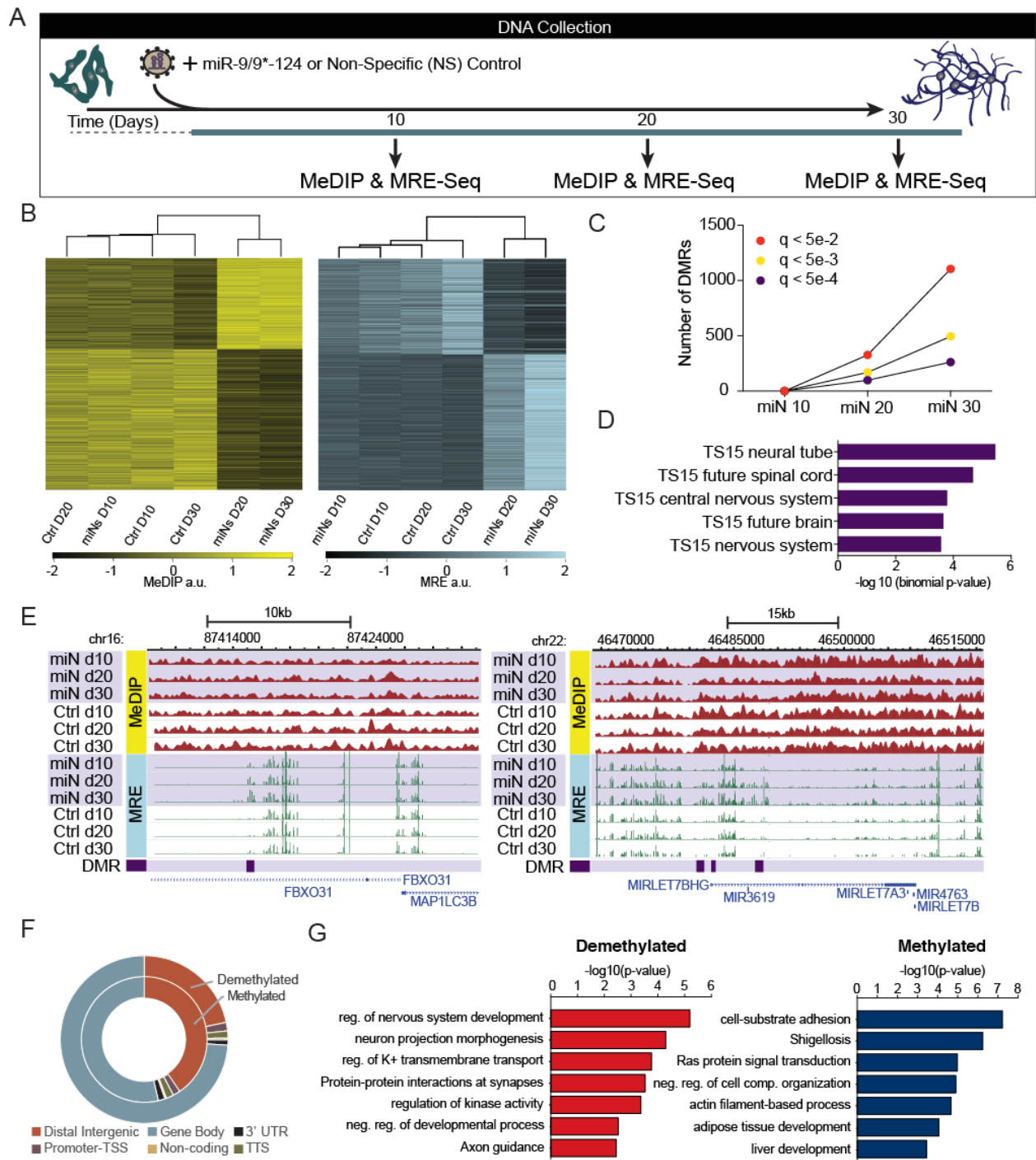
Figure 4. MiR-9/9*-124 Alter DNA Methylation at Neuronal Loci

(A) Schematic of sample collection during miR-9/9*-124-mediated neuronal reprogramming for

DNA methylation studies. Human fibroblasts were transduced with virus expressing miR-9/9*-

33

124 or a non-specific (NS) control (Ctrl) virus at day 0. Samples were collected at day 10, day 20, and day 30.

(B) Biclustering analysis of DMRs. Heatmaps based on MeDIP-seq RPKM (left) and MRE-seq RPKM (right) showing overlapping DMRs at days 20 and day 30.

(C) Quantification of DMRs at multiple q-value cutoffs (q < 5e-2 in red; q < 5e-3 in yellow; q < 5e-4 in purple) across all time points: miN 10 (miN day 10 vs. Ctrl day 10), miN 20 (miN day 20 vs. Ctrl day 20), and miN 30 (miN day 30 vs. Ctrl day 30).

(D) Tissue development enrichment of top overlapping DMRs at day 20 and day 30 show neuronal tissue terms at TS15 (~E9/10 in mouse development).

(E) WashU Epigenome Browser screenshots of two DMRs: *FBXO31* (left) and *MIRLET7BHG* (right) loci are shown with MeDIP-seq tracks (red; top), MRE-seq tracks (green; middle), and DMR positions (purple; bottom).

(F) Genomic distribution of differentially methylated and demethylated regions.

(G) Functional enrichment of top demethylated and upregulated (red; left) or methylated and downregulated (blue; right) DMRs overlapping at day 20 and day 30 compared with RNA-seq at day 30.

Chromatin Remodeling in miNs

Extensive expression changes in diverse chromatin remodeling genes during neuronal reprogramming (Figures 2D and S2) suggest that miR-9/9*-124 may also alter chromatin accessibility. We therefore performed Assay for Transpose-Accessible Chromatin followed by high throughput sequencing (ATAC-seq) 10 and 20 days after initiating miR-9/9*-124 induced

neuronal conversion. A high correlation between replicates confirmed the reproducibility of our

analyses (Figure 5A). We obtained 154,406 total peaks across all samples and identified 59,200

differential peaks (Figure 5B). Of the total peaks detected, 20,712 became more accessible

(open) and 38,882 peaks became inaccessible (closed). Most differential peaks in miNs at day

10 overlapped with the peaks at day 20 and the signal intensity of peaks gradually increased or

decreased during the conversion (Figure 5C), suggesting gradual transition of chromatin

accessibility during reprogramming (Figure 5C). Interestingly, the ratio of open intragenic to

distal intergenic regions increased during miR-9/9*-124-mediated conversion (Figure 5D).

Erasure of Fibroblast Epigenetic Identity and Gain of Neuronal Chromatin Architecture

To gain insight into biological relevance of changes in chromatin accessibility after miR-

9/9*-124 expression, we performed gene enrichment analysis on genes with differential ATAC

signals around the transcription start site ± 2Kb (TSS). We identified 4,915 genes with gradual

increases and 1,763 genes with gradual decreases in ATAC signals during conversion. Top GO

terms associated with genes with increased ATAC signals are enriched in neuronal terms

(Figure 5E), while closed regions were not (Figure S3A). The number of genes in open regions

associated with neuronal terms gradually increased from day 10 to 20 consistent with our

transcriptional profiling and coinciding with *bona fide* neuronal commitment of miNs (Figure 5E).

Consistent with this hypothesis, chromatin accessibility for fibroblast marker genes like *S100A4*,

S100A10, *VIM* and *COL13A1* gradually decreased between 10 and 20 days into neuronal

conversion. Collectively, our ATAC-seq analyses demonstrate miR-9/9*-124-induced chromatin

remodeling events are characterized by concurrent closing of fibroblast-related genomic loci and

opening of neuronal gene loci.

Next, we examined whether miR-9/9*-124-induced chromatin accessibility correlated

with changes in mRNA levels. We compared DEGs (logFC > 2 or < -2 adjusted P-value < 0.01)

to genes with altered chromatin accessibility around the TSS and identified 501 upregulated and 184 down-regulated genes that coincide with open and close regions, respectively (Figure 5F). GO enrichment analysis revealed that upregulated genes from open regions associate with neurogenic terms, while downregulated genes from closed regions are connected with terms important for fibroblast function (Figure 5G, see S3E for example tracks). This demonstrates the concordant regulation of transcription and chromatin accessibility during reprogramming. Interestingly, some opened genomic regions that displayed no changes in gene expression contained genes that are uniquely expressed in neuronal subtypes, including those enriched in: dopaminergic neuron markers (*TH and SLC6A3*), serotonergic neuron markers (*FEV, LMX1B and SLC6A4*), GABAergic neuron markers (*SLC6A1, SLC32A1, GAD2*), striatal medium spiny neuron marker (*PPP1R1B*), glutamatergic neuron markers (*GLUL and SLC1A6*), and cholinergic or motor neuron markers (*MNX1, CHAT, and SLC5A7*) (Figure S3G for example tracks). These results suggest that miR-9/9*-124 poise chromatin to accept additional inputs from subtype-specifying determinants without activating subtype-specific programs. In all, the chromatin dynamics observed during miRNA mediated neuronal conversion are consistent with time-dependent suppression of fibroblast identity concurrent with the opening of neuronal loci and activation of pan-neuronal gene expression.

Figure 5
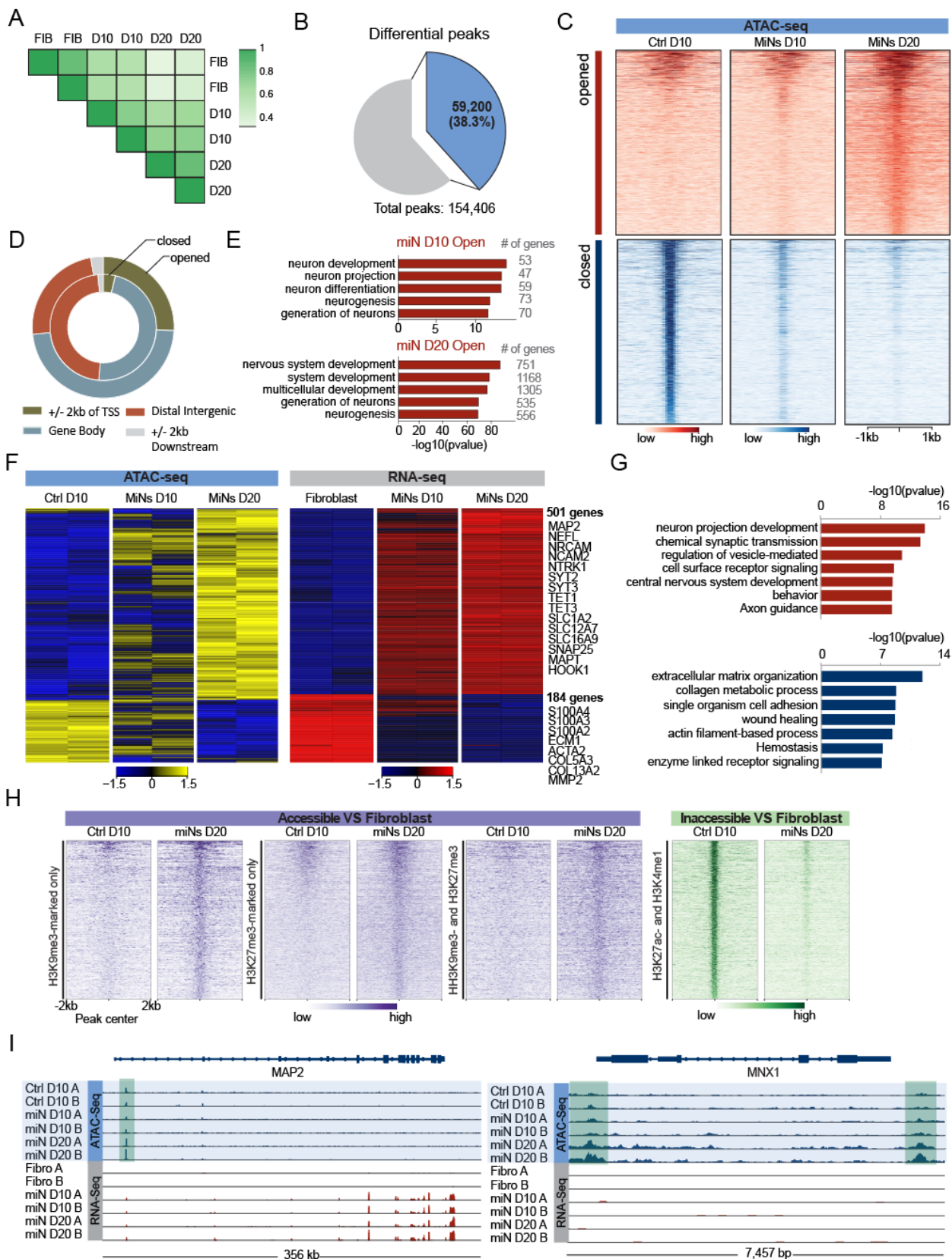
Figure 5. MiR-9/9*-124 Globally Changes Chromatin Accessibility

(A) Two dimensional correlation plot of samples. Pearson's correlation coefficient: 0.90 for Ctrl D10 (or FIB); 0.83 for miNs D10 (D10); 0.90 for miNs D20 (D20).

(B) Pie chart showing the proportion of differential peaks and total peaks. Differential peaks were obtained by combining all significant peaks (Ctrl D10 vs miNs D10, miNs D10 vs miNs D20).

(C) Heatmaps showing signal intensity in open and close chromatin peaks across all time points. All open and closed chromatin regions were ranked according to maximum intensity across all samples.

(D) The genomic distribution of open and closed chromatin regions.

(E) Comparison of GO terms for genes with open chromatin regions at promoters in miNs at day 10 and 20, but closed chromatin regions in fibroblasts.

(F) Heatmaps showing gene expression levels for DEGs positively correlated with ATAC-seq signal intensity in their promoter regions. Signal intensity is based on normalized counts per million (CPM) values.

(G) Top GO terms associated with DEGs which correlate with ATAC-seq signal intensity in promoter regions.

(H) Heatmaps showing signal intensity in the open (accessible) and closed (inaccessible) chromatin regions that overlapped with histone-marked regions of fibroblasts.

 I) Integrated genomics viewer (IGV) screenshots showing two different examples of ATAC-seq and RNA-seq integration. An example of ATAC-seq and RNA-seq peaks within a pan-neuronal

gene: *MAP2* (left) and an example of ATAC-seq peaks within a subtype-specific locus without gene expression changes: *MNX1* (right).

See also Figures S3 and S4.

MicroRNA-induced Chromatin Remodeling at Heterochromatin Regions in Fibroblasts

To gain a more complete understanding of the epigenetic architecture within opened and closed chromatin sites we examined the relationship of these regions to pre-existing histone marks present in fibroblasts. We hypothesized that regions that close during reprogramming would overlap with the active enhancer/euchromatin marks H3K27ac and H3K4me1 in fibroblasts. Conversely, regions that open during neuronal reprogramming may overlap with heterochromatic H3K9me3 and H3K27me3 signatures pre-existing in fibroblasts.

We selected 70,661 regions commonly marked by H3K27ac and H3K4me1 and 5,843 regions commonly marked by H3K9me3 and H3K27me3 in human fibroblasts based on the Roadmap Epigenome database (Roadmap Epigenomics et al., 2015). We confirmed regions with altered chromatin accessibility overlapped with these histone marks. Strikingly, we found that 1,128 ATAC signal peaks present in day 20 miNs that display increased accessibility during conversion overlapped with regions of fibroblasts marked by H3K9me3/H3K27me3, which were heterochromatic regions in fibroblasts. Whereas H3K27ac/H3K4me1 marked euchromatic regions in fibroblasts overlapped with 16,207 peaks that were closed in miNs (Figure 5H). GO enrichment analysis of genes associated with open chromatin regions in miNs that were marked by H3K9me3 or H3K27me3 in fibroblasts resulted in neuronal differentiation and function-related terms (Figure S3B) In contrast, regions that lose chromatin accessibility in miNs were enriched with non-neuronal terms (Figure S3D). These results demonstrate the surprisingly potency of miR-9/9*-124 to open heterochromatin regions needed for neuronal development and to close enhancer regions that preexist in fibroblasts and are not active in neurons. Our

results collectively provide an unprecedented demonstration that microRNAs can change chromatin architecture to promote neuronal and repress fibroblast fates during the direct conversion of human fibroblasts to neurons.

Chromatin Remodeling is Required for Direct Conversion

To determine if chromatin changes were necessary for cell fate conversion we knocked down the expression BRG1, a core component of BAF chromatin remodeling complex whose reduced function has been shown to collapse the overall chromatin architecture (Kadoch et al., 2017). After 20 days of neuronal conversion, loss of BRG1 markedly decreased the amount of MAP2 positive cells when compared to a control shRNA (Figure S4A). ATAC-seq revealed regions which failed to open in response to miR-9/9*-124 in the absence of BRG1 (Figure S4A). These regions were associated with neuronal GO terms in contrast to the fibroblast-related GO terms associated with regions that failed to close (Figure S4D). These data demonstrate the requirement of chromatin remodeling in cell fate conversion. Collectively, our transcriptome and ATAC-seq results provide mechanistic insight into the possibility of deriving additional clinically relevant neuronal subtypes through miRNA-mediated conversion in addition to cortical and striatal medium spiny neurons (Victor et al., 2014; Yoo et al., 2011).

Supplementary Figures
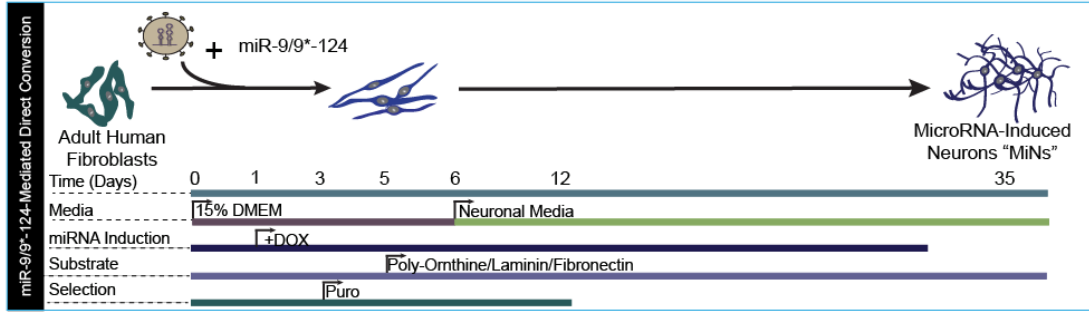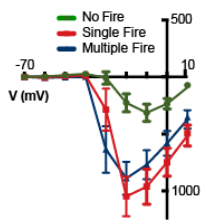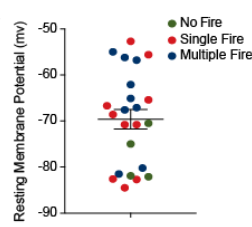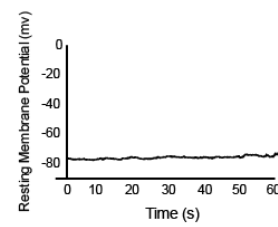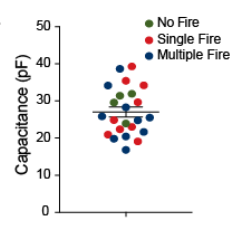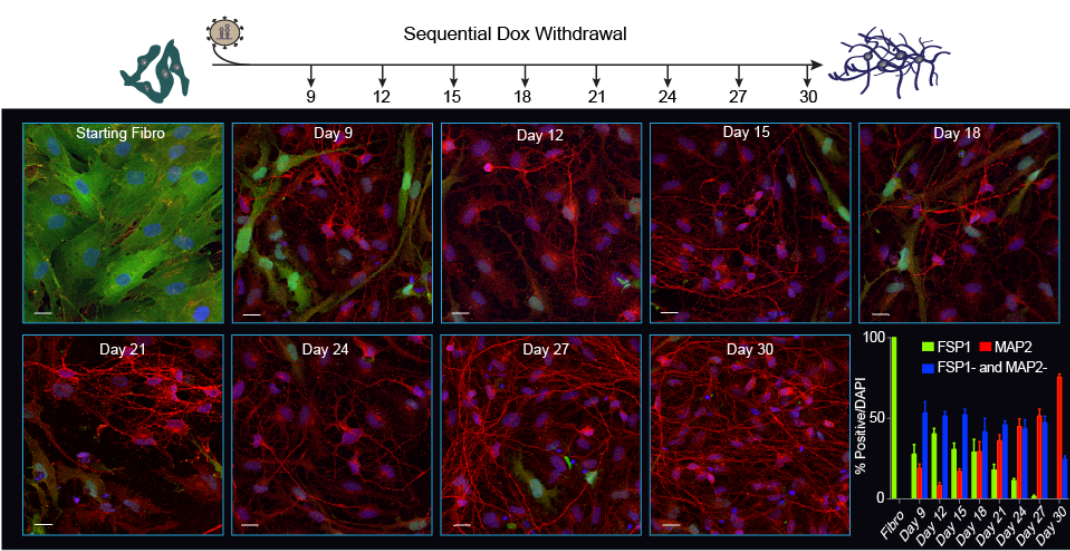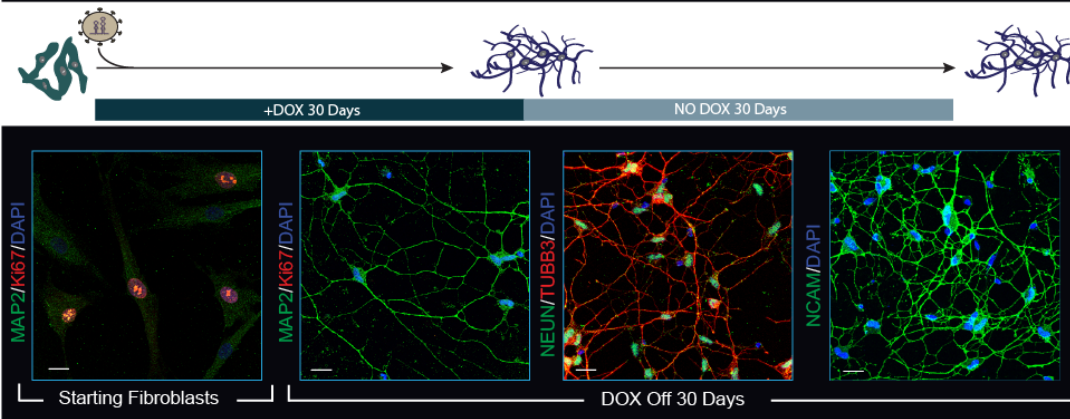
Figure S1

Figure S1, Related to Figure 1

MiRNA-Mediated Conversion into Neuronal Fate is Stable

(A) Detailed schematic of miR-9/9*-124 direct conversion protocol.

(B) Combined plot of the current (I) - voltage (V) relationship for every neuron recorded.

(C) Tabulated values of miN resting membrane potentials.

(D) Gap-free recording of miN resting membrane potential.

(E) Tabulated values of miN capacitance values

(F) MiR-9/9*-124 was ectopically expressed under a doxycycline (DOX) inducible promoter in 22 year old human fibroblasts for 30 days then DOX was removed from the media and cells were cultured for an additional 30 days. Stable adoption of the neuronal fate was assayed by immunostaining for the pan-neuronal markers MAP2, NeuN, TUBB3 and NCAM. To assay if cells remained post-mitotic cells were stained with the proliferative marker ki-67. Scale bar = 20μm.

All data are represented as mean ± SEM.

Figure S2
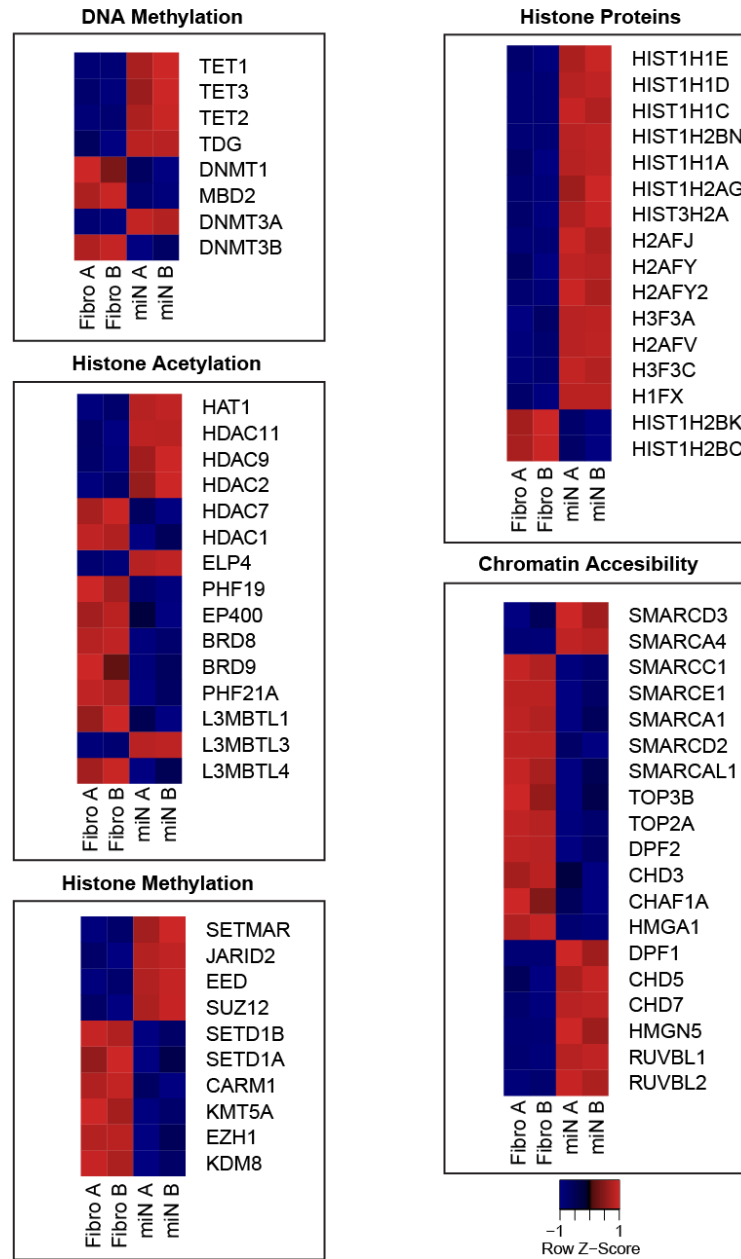
Figure S2, Related to Figure 2

Widespread Expression Changes in Epigenetic Modifiers During Reprogramming

Heatmaps show expression changes observed between miNs and fibroblasts within a subset of

proteins that recognize or modify distinct parts of the epigenome

Figure S3

A

miN d10 Closed

	# of genes
Cell Communication	133
Single Organism Signaling	130
Signaling	130
Signal Transduction	121
Reg. of Biological Quality	88
Cellular Response to Stimulus	146
Single-organism Process	256
Positive Reg. of Inflammatory Response	9
Reg. of Cell Shape	10
Single-organism Cellular Process	232

-Log (p-value)

miN d20 Closed

	# of genes
Positive Reg. of Biological Process	427
Signal Transduction	391
Positive Reg. of Cellular Process	369
Positive Reg. of Metabolic Process	291
Reg. of Cellular Component Organization	191
Reg. of Multicellular Organismal Process	214
Single Organism Signaling	412
Cell Communication	418
Signaling	412
Response to Stimulus	568

-Log (p-value)

B — H3K9me3

- Neuron development
- Synaptic signaling
- Cell Differentiation in spinal cord
- Cell fate commitment
- Ventral nervous System Neuron Differentiation
- Spinal cord development
- Neuron projection development
- Axogenesis

C — H3K27me3

- Neurogenesis
- Synaptic transmission
- Central nervous system development
- Neuron projection development
- Spinal cord development
- Motor neuron axon guiadance
- Spinal cord patterning

D — H3k27ac and H3K4me1

- Positive reg. of biological process
- Reg. of cell motility
- Reg. of intracellular signal transduction
- Reg. of phosphate metabolic process
- Vasculature development

E — Fibroblast Genes

ECM1 — 6,748 bp; MFAP5 — 19 kb; VIM — 10 kb

F — Neural Stem Cell and Neuronal Progenitor Genes

SOX2-OT SOX2 — 150 kb; OLIG2 — 4,261 bp; ASCL1 — 3,268 bp

G — Neuronal Subtype Genes

GAD2 — 92 kb; GABRA2 — 155 kb; TH — 8,673 bp

44

Figure S3, Related to Figure 5

Pre-existing Heterochromatic Neuronal Loci Open in Response to miR-9/9*-124 Expression.

(A) Top GO terms associated term with promoter regions that close during reprogramming from fibroblasts to miNs.

(B) Closed regions in fibroblasts marked by H3K9me3 that open during neuronal reprogramming are enriched for neuronal GO terms.

(C) Closed regions in fibroblasts marked by H3K27me3 that open during reprogramming are also enriched for neuronal GO terms.

(D) Pre-existing distal H3K27ac and H3K4me1 marks within fibroblasts that close during neuronal reprogramming show GO terms related to general biological processes.

(E) Genome browser snapshots demonstrating closing and loss of fibroblast gene expression (*ECM1*, *MFAP5*, and *VIM*).

(F) Genome browser snapshots demonstrating neither opening or activation of gain of progenitor genes (*SOX2*, *OLIG2* and *ASCL1*).

(G) Genome browser snapshots demonstrating neuronal subtype gene loci that open, but do not show gene expression changes (*GAD2* and *GABRA2*, GABAergic markers; *TH*, dopaminergic neuron marker).

Figure S4
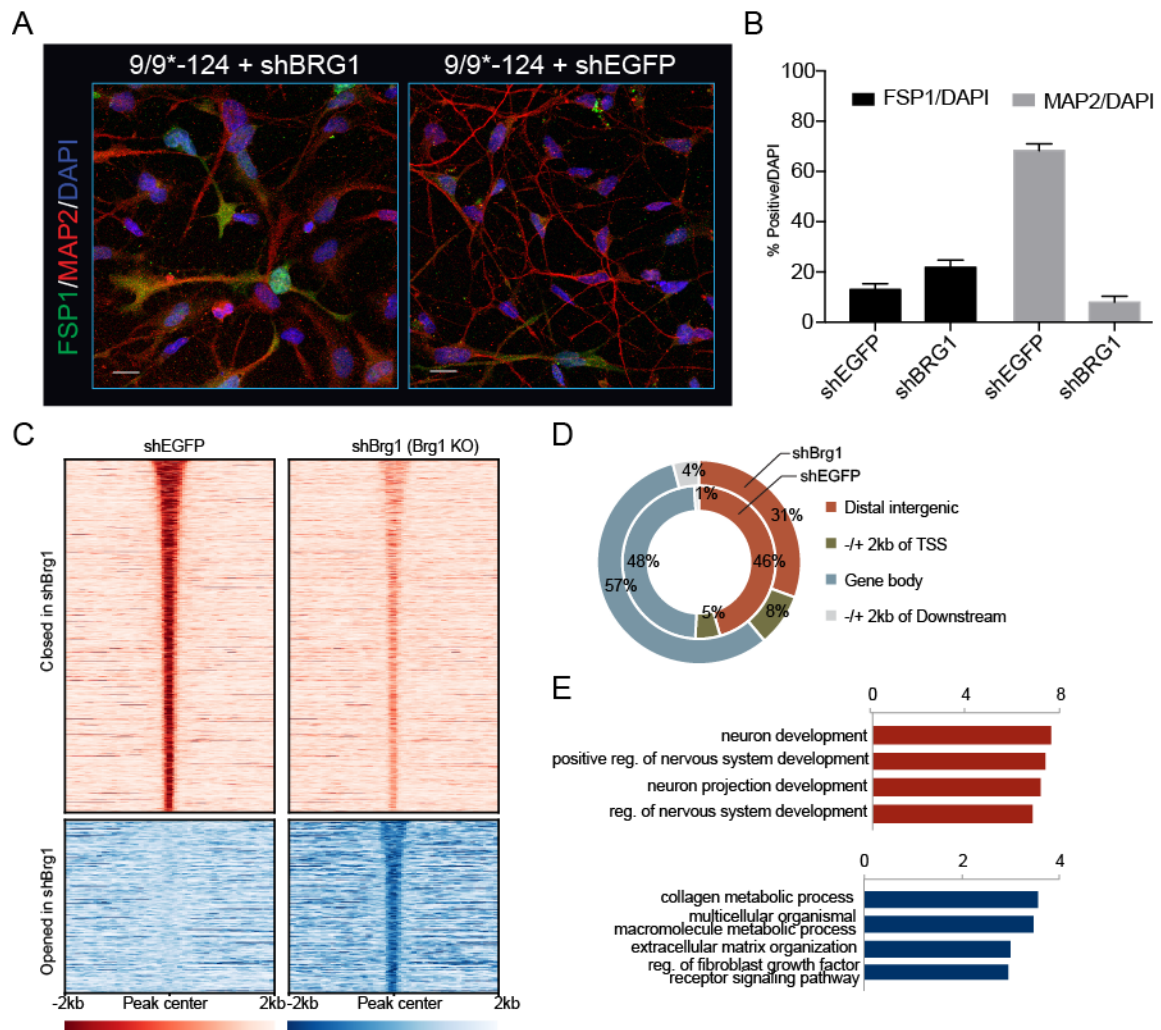
Figure S4, Related to Figure 5

Loss of BRG1 Prevents Neuronal Fate Acquisition

(A) Adult human fibroblasts ectopically expressing miR-9/9*-124 and shBRG1 or shEGFP

control for 20 days immunostained for the pan-neuronal marker MAP2 and fibroblast marker

FSP1. Scale bars = 20μm.

(B) Quantification of (A) MAP2 and FSP1 positive cells over total number of cells (DAPI). For MAP2, only cells with processes at least three times the length of the soma were counted. Data are represented as mean ± SEM. shBRG1 n = 155 cells, shEGFP, n = 173 cells,

(C) Heatmaps showing signal intensity in open and close chromatin peaks in shBRG1 and shEGFP. All open and closed chromatin regions were ranked according to maximum intensity across all samples.

(D) The genomic distribution of open and closed chromatin regions.

(E) GO terms showing biological function of closed (lost in shBRG1, red) and open (retained in shBRG1, blue) regions.

Materials and Methods

Cell Culture

Adult and neonatal human fibroblasts were obtained from commercial sources and maintained in fibroblast media comprised of Dulbecco's Modified Eagle Medium (Invitrogen) supplemented with 15% fetal bovine serum (Life Technologies), 0.01% β- mercaptoethanol (Life Technologies), 1% non-essential amino acids, 1% sodium pyruvate, 1% GlutaMAX, 1% 1M HEPES buffer solution, and 1% penicillin/streptomycin solution (all from Invitrogen). Cells were never passaged more than 15 times.

Plasmid construction and virus production.

Complementary cDNA was generated from adult human spinal cord (Clontech) from which individual motor neuron transcription factors (Figure S5A) were subcloned into the N174 and N106 lentiviral vectors using standard techniques. Lentivirus was produced in 293le cells plated in 10cm dishes (6.5×10^6 cells per dish) via polyethylineimine (48 µL of 2 mg/mL, Polysciences) assisted transfection of 3rd generation packaging vectors (1.5 µg pMD2.G and 4.5 µg psPAX2), and 6 µg of lentiviral backbone plasmid (e.g. pT-BCL-9/9*-124) 16 hours after initial plating. Media was changed the next day. After 2 days, media was harvested, filtered through a 0.45 µm polyethersulfone (PES) syringe filter and then concentrated by centrifugation at 70,000xG for 2 hours at 4°C. Virus collected from a single 10 cm dish was resuspended in 1 mL of sterile PBS then aliquoted and stored at -80°C. Before each transduction, virus aliquots were spun at 5,000xG for 5 minutes at 4°C to remove debris. Control vector expressing non-specific (NS) miRNA and BCL-XL was generated previously (Victor et al., 2014).

Direct Conversion

To initiate direct conversion, 1.8×10^6 cells were seeded onto Costar 6 well cell culture

48

vessels (Corning; 300,000 cells/well). The following day, each plate was transduced with the following reprogramming cocktail: 750 uL of concentrated lentivirus containing the reverse tetracycline-controlled transactivator (rtTA; Addgene, 66810) and 500 μL of virus containing pT-BCL-9/9*-124 or pT-BCL-9/9*-124 plus 500 μl of each individual TF driven by the EF1α promoter in the presence of polybrene (8 μg/mL ; Sigma-Aldrich) all diluted up to 18 mL (3 mL per well) then spinfected at 37°C for 30 minutes at 1,000xG using a swinging bucket rotor. The following day media was changed to fresh fibroblast media (2 mL per well) supplemented with doxycycline (Dox; Sigma Aldrich, 1 μg/mL). After 2 days, fresh fibroblast media was changed and supplemented with Dox and antibiotics for respective vectors (Puromycin, 3 μg/ml; Blasticidin 5 μg/ml; Geneticin, 400 ug/mL; all from Invitrogen). Five days post-transduction cells were replated on to poly-ornithine/laminin/fibrobnectin (PLF) coated glass coverslips. Before PLF coating, glass coverslips were acid treated according to (Richner et al., 2015). To transfer cells, for each well of a 6 well plate, cells were first washed 2x with 1 mL sterile PBS. Then 320 μL of 0.25% Trypsin (Gibco) was added to each well then placed in an incubator. Cells were monitored every 2 minutes, as soon as cells began to detach (no more than 6 minutes) 1 ml of MEF media supplemented with 1 μg/mL Dox was added to each well. One by one, each well was gently triturated three times to remove remaining attached cells then transferred to a sterile 1.5 mL Eppendorf tube. Cells were then spun at 200xG for 5 minutes at 37°C. The supernatant was aspirated and cells were gently resuspended in 300 μL MEF media supplemented with Dox. Cells were then drop-plated onto either 18 mm (150 μL per/c.s.; placed in 12 well plate) or 12mm (60 μL per c.s.; placed in 24 well plate) coverslips. Cells were left to settle for 15 minutes in an incubator then each well was flooded with MEF media supplemented with 1 μg/mL Dox. The following day media was then changed to Neuronal Media (Sciencell) supplemented with

49

Dox, valproic acid (1 mM; EMD Millipore) dibutyryl cAMP (200 μM; Sigma-Aldrich), BDNF, NT-3, CNTF, GDNF (all 10 ng/ml, Peprotech), and Retinoic Acid (1 uM; Sigma-Aldrich) and antibiotics for each vector. Dox was replenished every two days and half the media was changed every 4 days. Drug selection was halted 14 days into conversion. A diagram of the reprogramming protocol is available in Figure S1.

For time-course studies, the above reprogramming protocol was utilized with the following exception. In 3 day intervals starting on reprogramming day 9, a full media change excluding DOX was used to effectively remove DOX from the well at each timepoint.

For DNA methylation profiling, 1.8x10^6 human neonatal fibroblasts were seeded onto 10 cm plates (Corning). To achieve the cell number required for meDIP and MRE-seq human neonatal fibroblasts were utilized as adult fibroblasts were unable to expand to sufficient quantities. The following day, each plate was transduced with 10 ml of un-concentrated lentivirus containing media with a doxycycline inducible miR-9/9*-124 vector (Victor et al., 2014) and polybrene (8 μg/mL ; Sigma-Aldrich). The following day media was changed to fresh fibroblast media (2 mL per well) supplemented with Dox. After 2 days, fresh fibroblast media was changed and supplemented with Dox and antibiotics for respective vectors (see Table 1). Seven days post-transduction, cells were first washed 2x with 3 mL sterile PBS. Then 1 ml of 0.25% Trypsin (Gibco) was added to each plate then placed in a 37°C tissue culture incubator. Cells were monitored every 2 minutes and as soon as cells began to detach (no more than 6 minutes), 4 ml of MEF media supplemented with 1 μg/mL Dox was added to each plate. Cells were transferred to Primaria modified 10 cm plates (Corning) and 5 ml fresh fibroblast media supplemented with 1 μg/mL Dox was added to a final volume of 10 ml. The following day media was changed to Neuronal Media (Sciencell) supplemented with Dox, valproic acid (1 mM; EMD

50

Millipore) dibutyryl cAMP (200 μM; Sigma-Aldrich), BDNF, and NT-3 (all 10 ng/ml,

Peprotech), Retinoic Acid (1 uM; Sigma-Aldrich), and 4% FBS. Dox was replenished every two

days and half the media was changed every 4 days.

Immunocytochemistry and cell counting

Cells were fixed using 4% formaldehyde for 18 minutes at room temperature (RT) then

blocked and permeabilized for one hour at room temperature in PBS containing 0.3% Triton-

X100, 5% bovine serum albumin (Sigma-Aldrich), and 2% of either goat or donkey serum

(Sigma-Aldrich). Primary antibodies were incubated overnight at 4°C in blocking buffer. Cells

were then washed 3x and incubated with secondary antibodies conjugated to either Alexa-488, -

594 or -647, for one hour at room temperature. a-Bungarotoxin was incubated with secondary

antibodies at a concentration of 1:200. Images were obtained on a Leica SP-2 Confocal

Microscope. Quantifications were performed on at least 6 random fields of view in duplicate

experiments.

Electrophysiology

Whole-cell patch-clamp recordings were performed 35-40 days post-transduction. Data

was acquired using pCLAMP 10 software with multiclamp 700B amplifier and Digidata 1550

digitizer (Molecular Devices). Electrode pipettes were pulled from borosilicate glass (World

Precision Instruments) and typically ranged between 5–8 MΩ resistance. Intrinsic neuronal

properties were studied using the following solutions (in mM): Extracellular: 140 NaCl, 3 KCl,

10 Glucose, 10 HEPES, 2 CaCl$_2$and 1 MgCl$_2$ (pH adjusted to 7.25 with NaOH). Intracellular:

130 K- Gluconate, 4 NaCl, 2 MgCl$_2$, 1 EGTA, 10 HEPES, 2 Na-ATP, 0.3 Na-GTP, 5 Creatine

phosphate (pH adjusted to 7.5 with KOH). Membrane potentials were typically kept at -65 mV.

In voltage-clamp mode, currents were recorded with voltage steps ranging from -20 mV to +90

mV. In current-clamp mode, action potentials were elicited by injection of step currents that modulated membrane potential from -10 mV to +35 mV. Data was collected in Clampex and initially analyzed in Clampfit (Molecular Devices). Further analysis was done in GraphPad Prism 7 (GraphPad Software). Liquid junction potential was calculated to be 15.0 mV and corrected in calculating resting membrane potential according to previously published methods (Barry, 1994).

RNA-Seq library preparation and sequencing

Day 30 miNs and starting human adult fibroblasts (22 yr old) were extracted by RNeasy plus micro kit (Qiagen). The RNA samples with > 9.5 of RIN based on a 2100 Bioanalyzer were used for RNA-Seq library preparation. Library preparation and sequencing were performed by Genome Technology Access Center in Washington University School in St. Louis. Briefly mRNA was isolated by using SMARTer Ultra Low RNA Kit for Illumina sequencing (Clontech). All cDNA libraries, based on two biological replicates for each condition, were sequenced on Illumina Hi-Seq 2500 with single-end 50 bp read length.

For RNA-seq time series data in Figure 3, total RNA was extracted from starting human adult fibroblasts (22 yr old) and day 3, 6, 10, and 20 miNs by RNeasy plus micro kit (Qiagen). The RNA samples with > 9.5 of RIN based on a 2100 Bioanalyzer were used for RNA-Seq library preparation. Total RNA was processed for library construction by Cofactor Genomics (http://cofactorgenomics.com, St. Louis, MO) according to the following procedure: Briefly, total RNA was reverse-transcribed using an Oligo(dT) primer, and limited cDNA amplification was performed using the SMARTer® Ultra® Low Input RNA Kit for Sequencing – v4 (Takara Bio USA, Inc., Mountain View, CA). The resulting full-length cDNA was fragmented and tagged, followed by limited PCR enrichment to generate the final cDNA sequencing library

(Nextera® XT DNA Library Prep, Illumina, San Diego, CA). Libraries were sequenced as single-end 75 base pair reads on an Illumina NextSeq500 following the manufacturer's instructions by Cofactor Genomics.

RNA-Seq data analysis

More than 35 million reads of each RNA-seq data were aligned to human genome assembly GRCh 37. For differential expression analysis, edgeR and limma were used. Genes with low read counts, regarded as genes not expressed at a biologically meaningful level were filtered out before read normalization. The cut-off for low read count was counts per million (CPM) <1 in at least any two samples across the experiment. Reads for each sample were normalized by the edgeR method of trimmed mean of M-values (TMM). The quantitative difference of read counts between miNs and starting fibroblast samples were evaluated by carrying out limma and graphically represented by Glimma. Gene enrichment analysis for differentially expressed genes was performed using Metascape Gene Annotation and Analysis Resource tool.

For RNA-seq time series data in Figure 3, quality control, alignment, clustering, normalization, and expression comparison were performed by Cofactor Genomics (http://cofactorgenomics.com, Saint Louis, Missouri, USA). Raw sequence data in FASTQ format were assessed for quality (FastQC, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and ribosomal RNA content (sortmeRNA, http://bioinfo.lifl.fr/RNA/sortmerna/). NovoAlign (Novocraft, http://novocraft.com) was used to align reads to a set of transcript sequences. NovoAlign parameters were set to allow multiple alignments to the transcriptome set to allow for isoforms. Alignments to the genome were performed using STAR (https://github.com/alexdobin/STAR).

Only unique alignments to the genome were allowed. The genome alignment loci from all samples were combined and clustered to generate genomic loci ("patches") with contiguous read coverage. Patches overlapping reference genome annotation loci were labelled as such. For each transcript or patch, the RPKM expression value was calculated for each sample. These RPKM values are the basis for expression comparison and statistics generation. For each replicate group, the mean and coefficient of variation for each transcript or patch were calculated across the expression values for the samples in that group. These means were considered to be the expression values for the replicate group. P-values were calculated between the means of each pair of replicate groups using a Welch's t-test corrected for FDR by the method of Benjamini-Hochberg

DREM Analysis

The Dynamic Regulatory Events Miner (DREM) and cDREM were used to integrate time series gene expression data with predicted TF-gene binding interactions (top 100 genes per PWM from Ernst et al., 2010 to identify patterns of temporal gene expression patterns and the associated regulators. The log fold change of 7310 DE genes (fold-change > 10; q-value < 0.1) was used to generate all paths in Figure 3.

Methylated DNA immunoprecipitation sequencing

MeDIP-seq was performed as in Maunakea *et al.* (Maunakea et al., 2010). Five micrograms of genomic DNA was sonicated to a fragment size of ~100-400 bp using the Bioruptor sonicator (Diagenode). End-repair, addition of 3'-A bases and PE adapter ligation with 2 μg of sonicated DNA was performed according to the Illumina Genomic DNA Sample Prep Kit protocol. Adapter-ligated DNA fragments were size selected to 166-366 bp and purified by gel electrophoresis. DNA was heat denatured and then immunoprecipitated with 5-

54

methylcytidine antibody (Eurogentec; 1 μg of antibody per 1 μg of DNA) in 500 μl of

immunoprecipitation buffer (10 μM sodium phosphate, pH 7.0, 140 mM sodium chloride and

0.05% Triton X-100) overnight at 4 °C. Antibody/DNA complexes were isolated by addition of

1 μl of rabbit anti-mouse IgG secondary antibody (2.4 mg ml^{-1}, Jackson Immunoresearch) and

100 μl protein A/G agarose beads (Pierce Biotechnology) for 2 h at 4 °C. Beads were washed

nine times with immunoprecipitation buffer and then DNA was eluted in TE buffer with 0.25%

SDS and 0.25 mg ml^{-1} of proteinase K for 2 h at 50 °C. DNA was then purified with the Qiagen

QIAquick kit and eluted in 30 μl EB buffer. Ten microliters of DNA was used for a PCR-

enrichment reaction with PCR PE Primers 1.0 and 2.0. PCR products were size selected (220-

420 bp) and purified by gel electrophoresis. Methylated DNA enrichment was confirmed by PCR

on known methylated (SNRPN and MAGEA1 promoters) and unmethylated (a CpG-less

sequence on chromosome 15 and glyceraldehyde 3-phosphate dehydrogenase promoter)

sequences. DNA libraries were checked for quality by Nanodrop (Thermo Scientific) and

Agilent DNA Bioanalyzer (Agilent). Reads were aligned to hg19 using BWA and pre-processed

using methylQA (an unpublished C program; available at http://methylqa.sourceforge.net/).

Detailed library construction protocols for MRE-seq and MeDIP-seq are publically available at

the NIH Roadmap Epigenomics project website

(http://www.roadmapepigenomics.org/protocols/type/experimental/.

Methylation-sensitive restriction enzyme sequencing

     MRE-seq was performed as in Maunakea *et al.* (Maunakea et al., 2010), with

modifications as detailed below. Five parallel restriction enzyme digestions (*Hpa*II, *Bsh*1236I,

*Ssi*I(*Aci*I) and *Hin*6I (Fermentas), and *Hpy*CH4IV (NEB)) were performed, each using 1 μg of

DNA per digest for each of the samples. Five units of enzyme were initially incubated with DNA

for 3 h and then an additional five units of enzyme were added to the digestion for a total of 6 h of digestion time. DNA was purified by phenol/chloroform/isoamyl alcohol extraction, followed by chloroform extraction using phase lock gels. Digested DNA from the different reactions was combined and precipitated with one-tenth volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ethanol. The purified DNA was size selected and purified (50-300 bp) by gel electrophoresis and Qiagen MinElute extraction. Library construction was performed as per the Illumina Genomic DNA Sample Prep Kit protocol with the following modifications. During the end-repair reaction, T4 DNA polymerase and T4 PNK were excluded and 1 μl of 1:5 diluted Klenow DNA polymerase was used. For the adapter ligation reaction, 1 μl of 1:10 diluted PE adapter oligo mix was used. Ten microliters from the 30 μl of purified adapter ligated DNA was used for the PCR enrichment reaction with PCR PE Primers 1.0 and 2.0. PCR products were size selected and purified (170-420 bp) by gel electrophoresis and Qiagen QIAquick extraction. DNA libraries were checked for quality by Nanodrop (Thermo Scientific) and Agilent DNA Bioanalyzer (Agilent). Reads were aligned to hg19 using BWA and pre-processed using methylQA. MRE reads were normalized to account for differing enzyme efficiencies and methylation values were determined by counting reads with CpGs at fragment ends (Maunakea et al., 2010).

Differential DNA-methylated region analysis

The M&M statistical model (Zhang et al., 2013), which integrates MeDIP-seq and MRE-seq data to identify differentially methylated regions between two samples was implemented with a window size of 500 bp and a q-value (false discovery rate (FDR)-corrected P-value) cutoff of 5e-2. This cutoff was determined from Figure 3C, where only 1 DMR was detected at day 10 (miN day 10 vs. Ctrl day 10). For Figure 3B, only regions that were considered DMRs (q-

value < 1e-5) at both day 20 (miN day 20 vs. Ctrl day 20) and day 30 (miN day 30 vs. Ctrl day 30) are displayed.

GO enrichment analyses

DNA methylation GO analyses of MGI (Mouse Genome Informatics) expression (Smith et al., 2014) presented in Figure 3D were performed using the GREAT package (McLean et al., 2010). Gene regulatory domains were defined by default as the regions spanning 5 kb upstream and 1 kb downstream of the TSS (regardless of other nearby genes). Gene regulatory domains were extended in both directions to the nearest gene's basal domain, but no more than a maximum extension in one direction. The top 100 most significant overlapping DMRs from day 20 (miN day 20 vs. Ctrl day 20) and day 30 (miN day 30 vs. Ctrl day 30) were used as input. For Figure 3G, GO analyses were performed using Metascape (Tripathi et al., 2015) with a minimum enrichment of 1.5, a minimum overlap of 3, and a *p*-value cutoff of 0.01 using all demethylated or methylated DMRs at day 30 (miN day 30 vs. Ctrl day 30) that were either up- or down-regulated, respectively, by RNA-seq at day 35 using a cutoff of 2.5 logFC.

Genomic features

DMRs from day 30 (miN day 30 vs. Ctrl day 30) were segregated into exons, introns, intergenic regions, 3' UTRs, 5' UTRs, non-coding regions, promoter-TSSs, and TTSs by using the annotatePeaks program provided by HOMER (Heinz et al., 2010).

ATAC-sequencing library preparation and data processing

ATAC-seq was performed as previously described (Buenrostro et al., 2013). Briefly, 50,000 cells were collected for ATAC-seq library preparation at Ctrl day 10, miNs day 10, miNs day 20, shBRG1 Day20 and shEGFP day 20. Transposition reaction was carried out with Nextera Tn5 Transposase for 30 min at 37 °C. Library fragments were amplified for optimal

57

amplification condition. Final libraries were purified using AMPure XP beads (Ampure) and sequenced with 50 bp paired-end reads on Illumina HiSeq 2500.

More than 50 millions of ATAC-seq reads were trimmed for Nextera adapter sequences using TrimGalore and aligned to hg19 human genome assembly using bowtie2 with parameters --very-sensitive --maxins 2000 --no-discordant –-no-mixed. Duplicate reads were discarded with Picard and uniquely mapped reads were used for downstream analysis. Peaks were called using Homer with parameters findPeaks -region -size 150 -minDist 300. Peaks called from all the samples were combined together and raw reads mapped on the combined peaks were counted using HTSeq count. Differential peaks between any two different samples were identified using edgeR with a cut-off: a fold-change threshold of 1.5 and FDR < 0.01. Differential peaks were regarded as peaks that are gained or lost at each time point.

Gained peaks at miNs D10 and D20 were combined together and defined as open chromatin regions. Conversely, all lost peaks at miNs D10 and D20 were defined as close chromatin regions. The genomic features in the differential open and close chromatin regions were distributed by the CEAS software (Shin et al., 2009). We annotated Ref-seq genes that are most nearest located from differential peaks with Homer annotatePeaks command. Based on those genomic distribution and peak annotation, we defined the promotor regions (-/+ 2Kb of TSS) and distal regions (all peak positions except the promoter regions). GO enrichment analysis were performed by Metascape or the Gene Ontology. All heatmaps were made based on normalized signal intensity values (i.e. $\log_2$CPM) of each sample on relevant specific regions.

All histone mark ChIP-seq data were obtained from Roadmap Epigenome database of human fibroblasts (Roadmap Epigenomics et al., 2015). To identify histone mark-occupied chromatin

accessibility during reprogramming, we compared each histone ChIP-seq data with open and

close chromatin regions based on ATAC-seq. We confirmed that most open and close chromatin

peaks overlapped with histone mark-ChIP peaks were found in regions outside of promoter

regions (+/- 2Kb of TSS). Open and close chromatin regions excluding the promoter regions

were then used to compare with histone ChIP-seq data and perform further GO enrichment

analysis.

Chapter 3: Guiding the miRNA-Induced Neuronal Ground State to Motor Neurons

Adapted from:

MicroRNA-induced Epigenetic Remodeling During Direct Cell-Fate Conversion of Adult Human Fibroblasts

Daniel G. Abernathy, Wookyung Kim, Matthew J. McCoy, Allie Lake, Rebecca Ouwenga, Xiaoyun Xing, Daofeng Li, Hyung Joo Lee, Robert O. Heuckeroth, Joseph D. Dougherty, Ting Wang, Andrew S. Yoo

Abstract

Motor neurons provide the electrical bridge between the brain and musculature enabling vertebrates to interact with their environment. The inherent complexity in motor behaviors requires a robust developmental program to ensure proper specification, differentiation, and connection of motor neurons. A large amount of progress has been made in the study of each of these aspects of motor neuron biology. This work has been inspired by the disorders that arise when this intricate system is perturbed by disease or injury. Most well known of the motor neuron disorders are Amyotrophic Lateral Sclerosis (ALS), Spinal Muscular Atrophy (SMA) and Spinal Cord Injuries (SCI). Despite our increased understanding of motor neuron biology, a cure for any of these debilitating diseases has not been found. One of the greatest hurdles in studying neurodegenerative diseases is access the affected cell types. The location and absolute necessity of neurons within the human body prohibits their extraction for further studies. Furthermore, post-mortem samples cannot be expanded in culture since neurons are post-mitotic, thus preventing the accumulation of cell numbers needed for proper experimentation. In the same vein, neurons lost in these diseases cannot be replaced by the host as there is little adult neurogenesis (Ming and Song, 2011). If it were possible to generate the neurons affected by these diseases *in vitro*, a valuable resource would open up both for neurodegenerative disease modeling and prospective regenerative therapies. Our lab has previously shown that the ectopic expression of the brain enriched miRNAs, miR-9/9* and miR-124 (miR-9/9*-124) is sufficient to convert human fibroblasts into a heterogeneous population of both excitatory and inhibitory neurons (Yoo et al., 2011). Importantly, this method robustly converts cells originating from adult humans, a cell type historically difficult to reprogram efficiently (Caiazzo et al., 2011; Ring et al., 2012; Vierbuchen et al., 2010).

With the observation that neurodegenerative diseases often affect specific cell types we have sought to guide this conversion into distinct neuronal subtypes. We have found that ectopic co-expression of transcription factors enriched in specific regions of the brain and miR-9/9*-124, reprograms fibroblasts into the neuronal subtype that resides within the region of interest (Victor et al., 2014). For example, the introduction of transcription factors enriched in the striatum to the miRNA-mediated reprogramming process guides conversion into striatal medium spiny neurons (MSNs), the principle cell-type affected in Huntington's disease. The ability of miR-9/9*-124 to open loci important in the acquisition and maintenance of multiple neuronal subtypes, including but not limited motor neuron loci, suggests that motor neurons could be derived through the addition of motor neurons transcription factors to miR-9/9*-124. In pursuit of this hypothesis I screened a pool of developmentally informed motor neurons transcription factors in the background of ectopic miR-9/9*-124 expression and assayed for the acquisition of motor neuron identity. I identified two transcription factors, ISL1 and LHX3, that when co-expressed with miR-9/9*-124 were sufficient to directly convert adult human fibroblasts into functional motor neurons. Further, these motor neurons share significant genetic networks with their i*n vivo* mouse counterparts, suggesting that this reprogramming modality is capable of generating *bona fide motor* neurons.

Main Text

Instructing the miRNA-induced Neurogenic State to a Motor Neuron Fate

The deposition and removal of nucleosomes along regulatory elements within DNA inhibits or enables the binding of TFs, simultaneously facilitating and reinforcing cell-type specific gene expression programs (Jiang and Pugh, 2009). In our previous analyses we noted chromatin regions in miNs with enhanced accessibility were proximal to *MNX1* and choline

acetyl transferase *CHAT,* two of the hallmark genes expressed by motor neurons (Fonnum, 1973; Tanabe et al., 1998) (Figure 5I). Neither *MNX1* nor *CHAT* mRNA levels were elevated in day 30 miNs after miR-9/9*-124 directed neuronal conversion, but we hypothesized the open chromatin would facilitate expression of these genes in response to motor neuron-specific TFs. To test this hypothesis we expressed a panel of TFs known to promote motor neuron identity (Figure S5A) in combination with ectopic miR-9/9*-124 expression, and assayed for MAP2 and MNX1 expression by immunostaining (data not shown). While all combinations resulted in MAP2 positive cells, only ISL1 and LHX3 robustly led to the generation of MNX1-positive cells. Co-expressing LHX3 and ISL1 with miR-9/9*-124 in adult human fibroblasts from 22-, 42-, 56- and 68-year old donors resulted in MAP2, TUBB3 and NCAM-positive cells with complex neuronal morphologies (Figures 6A and 6B). Approximately 80% of cells positive for DAPI staining were positive for TUBB3, MAP2 and NCAM expression (Figure 6C). The majority of converted cells displayed nuclear staining of MNX1 (~85% of TUBB3 positive cells in each line, Figures 6D and 6E). Similarly, cytoplasmic CHAT protein and SMI-32 (a neurofilament protein found in motor neurons) were detected in ~80% of TUBB3 positive cells within each age group (Figure 6D and 6E).

Interestingly, ISL1 and LHX3 alone were not sufficient to induce neuronal conversion when co-expressed with a non-specific miRNA (miR-NS) (Figure S5B), supporting the notion that miR-9/9*-124 is necessary for the subtype-specifying activities of ISL1 and LHX3. Motor neuron conversion induced by miR-9/9*-124 plus ISL1 and LHX3 was stable, displaying neuronal morphologies, cell cycle exit, and motor neuron marker expression for 30 days after doxycycline removal (Figure 6F).

Figure 6

A — Guiding miRNA-Induced Neurogenic State to Motor Neurons

Adult Human Fibroblasts + Isl1 Lhx3 → Motor Neurons

B

MAP2/DAPI   TUBB3/DAPI   NCAM/DAPI

C

TUBB3/DAPI
MAP2/DAPI
NCAM/DAPI

% Positive/DAPI

22 yr old   42 yr old   56 yr old   68 yr old

D

22 Yrs Female   42 Yrs Female   56 Yrs Male   68 Yrs Female

TUBB3/MNX1/DAPI

CHAT/TUBB3/DAPI

SMI-32/TUBB3/DAPI

E

MNX1/TUBB3
CHAT/TUBB3
SMI-32/TUBB3

% Positive/TUBB3

22 yr old   42 yr old   56 yr old   68 yr old

F

+DOX 30 Days   NO DOX 30 Days

TUBB3/Ki67/DAPI   TUBB3/Ki67/DAPI   NEUN/MAP2/DAPI   TUBB3/MNX1/DAPI   CHAT/TUBB3/DAPI

Starting Fibroblasts   DOX Off 30 Days

Figure 6. MiRNA-Induced Neuronal Competence Enables Motor Neuron Transcription Factors, ISL1 and LHX3, to Determine Motor Neuron Identity.

(A) Schematic of neuronal induction paradigm using miR-9/9*-124 plus *ISL1* and *LHX3*.

(B) Representative immunohistochemistry for pan-neuronal markers in neurons generated from fibroblasts through 35 days of ectopic miR-9/9*-124, ISL1, and LHX3 co-expression. Fibroblasts were isolated from a 22 year old female donor. Scale bars = 20μm.

(C) Quantification of 4 independent primary human fibroblast lines from both male and female donors stained with TUBB3, MAP2 and NCAM. Percentages represent total number of positive cells over all cells (DAPI) and are represented as mean ± SEM. Cells analyzed: 22 yr old N=TUBB3 325, MAP2 219, NCAM 275; 42 yr old N=TUBB3 304, MAP2 236, NCAM 129; 56 yr old N=TUBB3 275, MAP2 279, NCAM 213; 68 yr old N=TUBB3 282, MAP2 234, NCAM 190.

(D) Expression and correct localization of motor neuron markers in neurons converted by miR-9/9*-124 and ISL1/LHX3 as demonstrated by immunohistochemistry. MNX1, (top), CHAT (middle) and SMI-32 (bottom). Scale bars = 20μm.

(E) Quantification of (D) represents the total percentage of MNX1, CHAT and SMI-32-positive cells over TUBB3-positive cells. Data are represented as mean ± SEM. Cells analyzed: 22 yr old N=MNX1 256, CHAT 256, SMI-32 113; 42 yr old N= MNX1 151, CHAT 151, SMI-32 283; 56 yr old N= MNX1 207, CHAT 207, SMI-32 174; 68 yr old N= MNX1 151, CHAT 151, SMI-32 96.

(F) After 30 days of neuronal conversion by ectopic miR-9/9*-124 expression, doxycycline was removed and cells were cultured for an additional 30 days. Immunocytochemistry showing motor neurons produced by miR-9/9*-124 plus ISL1 and LHX3 (Moto-miNs) remain Ki-67 negative (2nd panel), retain expression and localization of the neuronal proteins TUBB3, NEUN,

and MAP2 (2nd and 3rd panel), and express the motor neuron proteins MNX1 and CHAT (4th and 5th panel). Scale bars = 20μm.

See also Figure S5.

Electrophysiological properties of Moto-miNs

Motor neurons produced from fibroblasts by co-expression of miR-9/9*-124, ISL1 and LHX3 (Moto-miNs) demonstrated robust inward and outward currents in response to depolarizing steps (Figure 7A) and displayed action potential trains through the injection of step-wise depolarizations (Figures 7B and S5C). The visualization of single traces recorded at individual current steps revealed the characteristic hyperpolarization following each action potential seen in mature neurons (Figure 7C). To assess the percentage of functionally mature Moto-miNs, we patched 45 randomly chosen Moto-miNs from 22-year-old and 68-year-old donors. All Moto-miNs fired APs. Most cells fired multiple APs (80% n=20 and 74% n=25) while single APs were observed in ~20-25% of the patched cells (Figure 7D). Gap-free recordings in current clamp mode revealed cells capable of firing spontaneous action potentials (Figure 7E), demonstrating the excitability of these cells. Similar I-V curve relationships were observed between donors and firing patterns (Figure 7F). Peak inward currents were substantially higher than those observed in miNs (Figure S6A). Lastly, resting membrane potentials in all Moto-miNs tested were hyperpolarized (Figure 7G; 22 yr old, -67.2 mV ± 3.3 mV; 68 yr old, -72.8 mV ± 2.0 mV, S.E.M.). Coupled with the increased proportion of cells that fire multiple APs these data suggest that the addition of ISL1 and LHX3 to miR-9/9*-124 produced more mature neurons than exposure to miR-9/9*-124 alone.

The ability of motor neurons to control voluntary muscle movement stems from their ability to form neuromuscular junctions (NMJs), unique synapses formed between motor neurons and muscle cells. The formation of NMJs was visualized through the co-localization of

EGFP-labeled Moto-miNs, Alexa-fluor-594 Bungarotoxin (BTX, a toxin that binds to the nicotinic

acetylcholine receptor (AChR) of NMJs), and myosin heavy chain. BTX puncta were not

observed in the absence of Moto-miNs (Figure 7H, left inset). In contrast, Moto-miNs were able

to induce characteristic BTX-clustering in close apposition with EGFP labeled neurons and

myotubes (Figure 7H) indicating the formation of putative NMJs.

Transcriptional Profiling of Moto-miNs

To fully characterize the acquisition of a motor neuron fate and assess the contribution

of ISL1 and LHX3, we profiled the transcriptome of 22-yr-old starting fibroblasts, miNs and

Moto-miNs by microarray. We again observed the loss of fibroblast gene expression (for

example, *S100A4*, *VIM* and *COL13A1*), and the gain of a pan-neuronal identity (for example,

MAP2, NEFL, SNAP25 and *SCN1A*) 35 days after the expression of miR-9/9*-124 (Figure 7I).

Expression of motor neuron-specific genes was not significantly different between starting

fibroblasts and miNs. While the addition of ISL1 and LHX3 to miR-9/9*-124 did not significantly

change the expression of pan-neuronal genes when compared to miNs (Figure 7J), ISL1 and

LHX3 selectively activated key motor neuron genes including *MNX1*, *CHAT*, *VACHT*, *LMO1* and

LMO4 (Figure 7J). We further validated the loss of fibroblast identity and gain of motor-neuron

identity in Moto-miNs derived from 42-, 56- and 68-year-old donors by qRT-PCR using RNA

from human spinal cord as a positive control (Figure 7K). In addition, we observed dramatic

upregulation of miR-218, a recently identified motor neuron-specific microRNA (Amin et al.,

2015; Thiebes et al., 2015) in Moto-miNs (Figure 7L).

Figure 7

68
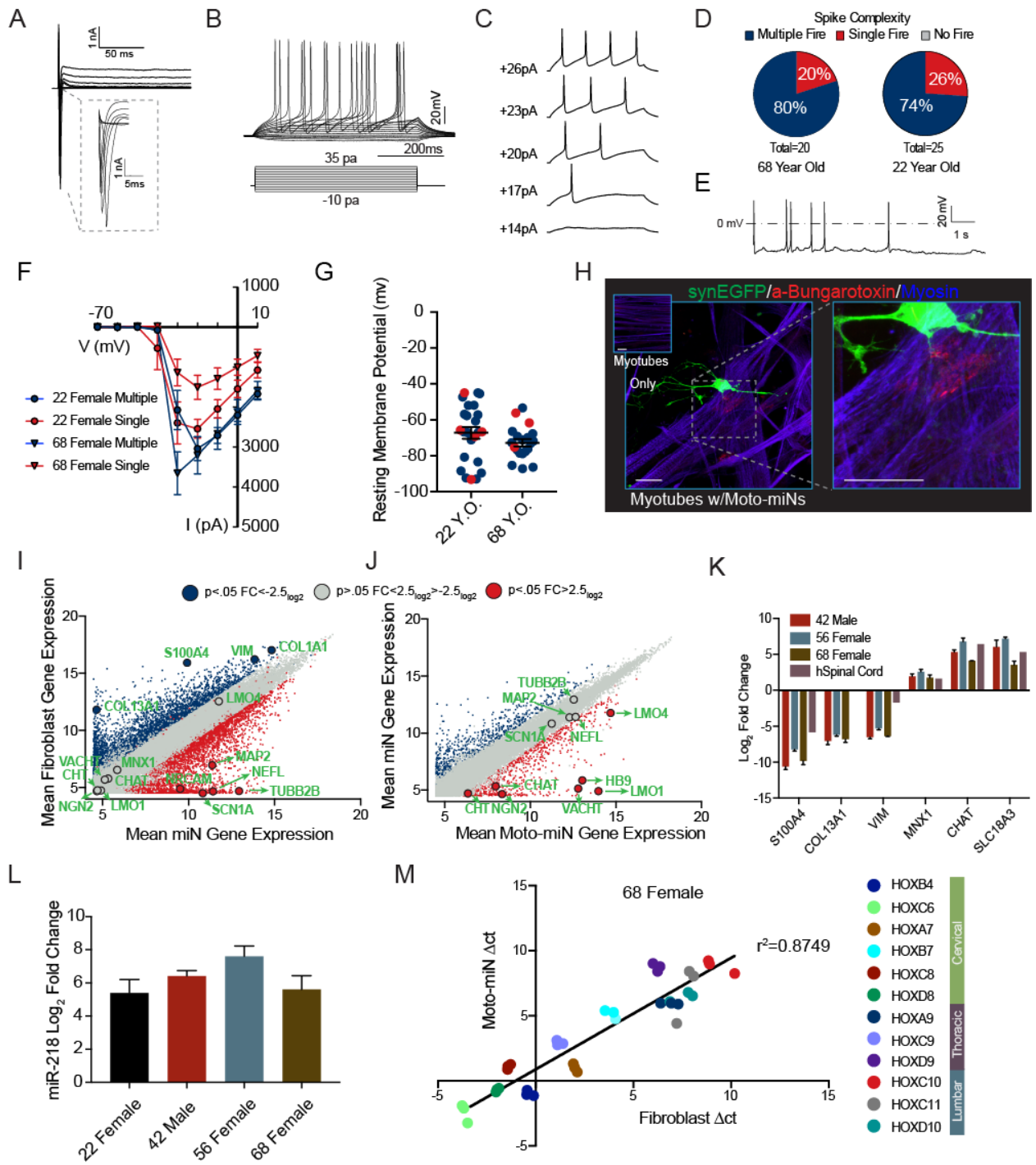
Figure 7. Functional Properties and Gene Expression Profile of Moto-miNs

(A) Representative traces of inward sodium and outward potassium whole-cell currents.

(B) Repetitive AP waveforms in response to 500 ms current injections recorded from Moto-miNs converted in monoculture.

(C) Representative waveforms of a single Moto-miN at increasing current injections.

(D) Summary of firing patterns observed in Moto-miNs converted from both old and young donors. 68 yr old 80% multiple (N=20), 22 yr old 74% multiple (N=25).

(E) Spontaneous firing activity was observed in a small percentage of Moto-miNs (3/20).

(F) Combined plot of the current (I) - voltage (V) relationship for every Moto-miN recorded. Data are represented as mean ± SEM.

(G) Moto-miNs converted from both young and old donors are hyperpolarized, demonstrating mean resting membrane potentials of -67.2mV and -72.8mV, respectively. Data are represented as mean ± SEM.

(H) Staining of Moto-miNs cultured with differentiated human myotubes. Moto-miNs were labeled with synapsin-EGFP via viral transduction and then plated onto human myotubes. Myotube only cultures did not have α-Bungarotoxin-594 (red) puncta (top left inset). Scale bar = 20μm.

(I) Scatterplot comparing the mean gene expression of starting fibroblasts from a 22 year old donor (y-axis) and miNs generated from the same individual (x-axis). Plot highlights a selection of pan-neuronal and fibroblast-specific genes in green text. Blue = $\log_2$FC < -2.5 and p < 0.05, (more abundant in fibroblasts) grey = $\log_2$FC >-2.5 and < 2.5 p>.05 (no significant difference), and red = $\log_2$FC > 2.5 and p < 0.05 (more abundant in miNs).

(J) Scatterplot comparing the mean gene expression of miNs from a 22-year old donor (y-axis) and Moto-miNs generated from the same individual (x-axis). Plot highlights a selection of pan-neuronal and motor neuron-specific genes in green text. Blue = $log_2FC < -2.5$ and $p < 0.05$, (more abundant in miNs) grey = $log_2FC > -2.5$ and < 2.5 $p > .05$ (no significant difference), and red = $log_2FC > 2.5$ and $p < 0.05$ (more abundant in Moto-miNs).

(K) Moto-miNs generated from multiple donors have lower mRNA levels for fibroblast genes and higher mRNA levels for motor neuron-specific genes. Moto-miNs were analyzed by qRT-PCR 35 days post-transduction. Human spinal cord RNA served as a positive control (normalized to 42 yr fibroblasts, ΔΔct method). Data are represented as mean ± SEM

(L) MiR-9/9*-124 and ISL1/LHX3 activate the expression of the motor neuron specific miRNA, miR-218. RNA was isolated from fibroblasts and Moto-miNs 35 days post-transduction and analyzed by qRT-PCR. Data are represented as mean ± SEM

(M) Moto-miNs retain donor fibroblast HOX gene expression pattern as demonstrated by qRT-PCR. Δct method. Data represent Δct values for each biological replicate (3 separate Moto-miN conversions).

See also Figure S6 and S7.

Next, we used the cell type-specific enrichment analysis tool (CSEA) (Xu et al., 2014) to test whether the gene expression profile within each population of miNs and Moto-miNs would be associated with distinct subtypes of *in vivo* neurons. When queried with a gene list, CSEA identifies neuronal subtypes that show significant enrichment in genes from the input list through curated transcriptomic data. Our CSEA analysis of the 100 most enriched genes within Moto-miNs identified two subtypes, brainstem motor neurons and spinal motor neurons, to be significantly associated. No subtype specificity was assigned to the miN transcriptome (Figure

S6B). This unbiased bioinformatics approach further supports the motor neuron identity of converted Moto-miNs.

Lastly, HOX gene expression patterns were compared by qRT-PCR between starting fibroblasts and Moto-miNs. Interestingly, we observed a high correlation (R^2 = 0.88) between the expression levels of HOX genes before and after conversion within each of the defined spinal cord regions tested, indicating Moto-miNs retain the positional identity that existed in original fibroblasts (Figure 7M, S6C).

Transcriptional Activation of ISL1 and LHX3 Genomic Targets

An alternative approach for generating motor neurons is forced expression of NGN2, ISL1, and LHX3 in human embryonic stem cells (ESCs) (Mazzoni et al., 2013). We compared the genomic targets of ISL1 and LHX3 identified by Mazzoni et al. through ChIP-seq to genes whose expression increases in Moto-miNs compared to miNs. Surprisingly, we identified a large cohort of overlapping genes (323) that include numerous hallmark motor neuron markers (Figure S7A), suggesting that ISL1/LHX3 activate expression of a core gene regulatory network that underlies the specification of diverse cellular states towards motor neurons.

Comparison of Moto-miNs to Endogenous Spinal Cord Motor Neurons

The inability to obtain a pure population of motor neurons from humans prevents direct transcriptional comparisons between Moto-miNs and their human *in vivo* counterparts. Therefore, we directly compared Moto-miNs to fully differentiated *in vivo* mouse motor neurons. To interrogate the gene expression of motor neurons within the large heterogeneity of cell-types present in the spinal cord, we performed Translating Ribosomal Affinity Purification (TRAP) followed by RNA-seq (Figure S7B). The use of two mouse lines expressing EGFP tagged ribosomes (one line under the pan-neuronal SNAP25 promoter, and the other through the CHAT promoter), enabled the enrichment and subsequent sequencing of actively transcribed

mRNA in all neurons and motor neurons within the spinal cord. We also profiled the transcriptome of the entire spinal cord as an additional 'pre-IP' control. Comparisons between CHAT pre-IP controls and CHAT-TRAP transcripts confirmed significant enrichment of motor neuron markers (Figure S7C, left) by the TRAP procedure. To further separate motor neuron enriched genes from common neuronal genes we first normalized CHAT-TRAP and SNAP-25 datasets to their pre-IP controls, then directly compared the normalized gene expression values. This analysis revealed the co-expression of pan-neuronal genes such as *MAP2*, *NRCAM*, *SCN1A* and *TUBB3* and highly enriched for motor neurons transcripts such as *SLC18A3 (VACHT)*, *CHAT* and *MNX1* (Figure S7C, right). This dataset also serves as a resource for genes enriched in spinal cord neurons over all SNAP25 expressing neurons within the spinal cord. To that end, we compared differentially expressed mouse spinal cord motor neuron genes to genes enriched in Moto-miNs over miNs. We observed a significantly larger overlap between Moto-miNs and mouse motor neurons than expected by chance (Fisher's exact test p = 3.522e-06), indicating that Moto-miNs and endogenous mouse motor neurons utilize similar genetic networks. This includes expression of canonical motor neuron markers such as, SLIT2 and SLIT3, host genes for the motor neuron specific miRNA, miR-218 (Amin et al., 2015) (Figure S7D). Altogether, our in depth analyses of Moto-miNs at cellular, functional and transcriptomic levels confirm that co-expression of miR-9/9*-124 and ISL1/LHX3 can directly convert adult fibroblasts into spinal cord motor neurons.

DISCUSSION

MiRNA-mediated neuronal conversion appears to be distinct from current models of cell fate reprogramming. Two models of lineage reprogramming have been proposed: one based on transcription factor cooperativity and positive feedback loops (Jaenisch and Young, 2008; Soufi et al., 2012; Vierbuchen and Wernig, 2012), and the other proposes that the "on-target" pioneer activity of a TF initiates and enables additional TFs to assist in cellular conversion (Wapinski et

72

al., 2013). In stark contrast, canonical gene regulation by miRNAs requires the removal of information through translational repression and transcript degradation. This mode of repression in conjunction with the multitude of anti-neurogenic genes targeted by miR-9/9*-124 suggests miRNA-mediated reprogramming acts through an alternative mechanism. We propose miR-9/9*-124 expression in non-neuronal somatic cells initiates gradual, yet active changes in the activities of multiple chromatin modifiers while simultaneously repressing anti-neuronal genes and activating neuronal genes culminating in a binary cell-fate switch. This model is supported by the rapid cell cycle exit observed upon ectopic miR-9/9*-124 expression, the subsequent neuronal switching within chromatin modifiers, steady increase in epigenetic and transcriptional changes, and the time scale in which conversion takes place.

Chromatin Remodeling Accompanies Cell Fate Conversion

In this study, we reveal the surprising potency of miR-9/9*-124 for remodeling chromatin and altering DNA methylation. Surprisingly, preexisting neuronal loci within the heterochromatic regions in human fibroblasts opened up in response to miR-9/9*-124. These data suggest the robustness of miRNA-mediated reprogramming observed in human cells could stem from their ability to induce epigenetic changes. Cellular processes and identity are governed by the cumulative action of multiple levels of genome regulation and it is unlikely a single genetic component downstream of miR-9/9*-124 mediates these changes and ultimately cell fate conversion. For example, almost every level of epigenetic remodeling participates in the induction of pluripotency (Takahashi and Yamanaka, 2016). Instructions operating through multiple levels of genetic and epigenetic regulation are likely required for true cell-fate conversion. Our thorough characterization of miR-9/9*-124 induced transdifferentiation of human fibroblasts into functional neurons highlights molecular processes that are critical to cell fate conversion.

A Modular Neuronal State

The plastic neuronal platform presented here affords modularity to direct cell fate conversion. Numerous studies in developmental neuroscience have identified subtype-specific TFs or terminal selector genes that could be incorporated in neuronal reprogramming technology. Yet, identifying molecules capable of overcoming the cell-fate barrier present in human somatic cells and eliciting a permissive environment in which terminal selector genes can act has proven to be challenging. We demonstrate this property by generating a neuronal population highly enriched in spinal cord motor neurons from human adult fibroblasts through the coexpression of miR-9/9*-124, ISL1 and LHX3. Because motor neurons are the major neuronal subtype affected in Amyotrophic Lateral Sclerosis (ALS) and Spinal Muscular Atrophy (SMA), the robustness and specificity of neuronal conversion employing miRNAs and motor neuron TFs may pave the way towards generating patient-specific MNs for disease modeling.

AUTHOR CONTRIBUTIONS

ASY, DGA, MJM and WK conceived the concept for this study. ASY and DGA designed and performed experiments determining the neurogenic potential of miR-9/9*-124. DGA and WK designed and performed gene expression studies. ASY, TW, HL and MJM designed DNA-methylation studies, and MJM generated and analyzed DNA-methylation data with assistance from XX and DL. ASY, DGA and WK designed and performed ATAC-Seq experiments and analyses. ASY and DGA designed and performed motor neuron experiments. JDD, ROH, AL and RO designed and analyzed TRAP-Seq experiments. ASY and DGA wrote the manuscript with input from all authors.

Supplementary Figures

Figure S5

A

TF's Tested with miR-9/9*-124

Isl1	Ngn2 Olig2 RXRG RARβ
Lhx3a	Ngn2 RXRG RARβ
Isl1 Lhx3a	Nkx 2.2
Nd2	Nkx 6.2
Ngn2	Nkx 6.2 Isl1
Ngn2 Isl1	Nkx 6.2 Lhx3a
Ngn2 Isl1 RXRG RARβ	Olig2
Ngn2 Lhx3a	Olig2 Isl1
Ngn2 Lhx3a Isl1	Olig2 Lhx3a
Ngn2 Lhx3a Isl1 RXRG RARβ	Olig2 Nkx 6.2
Ngn2 Lhx3a RXRG RARβ	Olig2 Nkx 6.2 RXRG RARβ
Ngn2 Olig2	Olig2 Nkx6.2 Isl1
Ngn2 Olig2 Isl1	RXRG RARβ
Ngn2 Olig2 Isl1 RXRG RARβ	

B

miR-NS+Isl1/Lhx3   miR-9124+Isl/Lhx3

MNX1/TUBB3/DAPI

22 Yrs Female
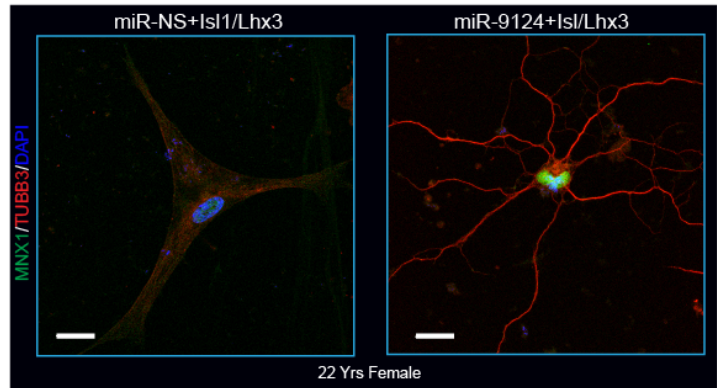
C
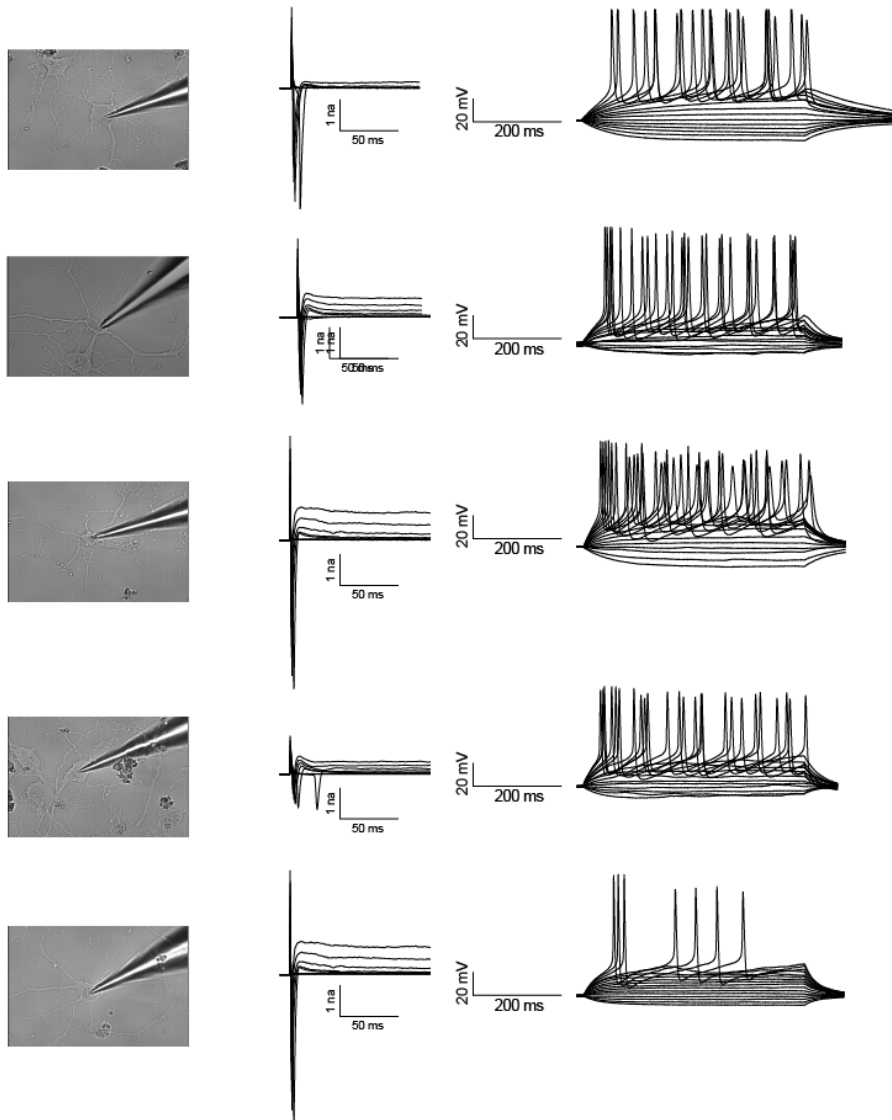
Figure S5, Related to Figure 6

Identification of Transcription Factors For Defining Motor Neuron Specific Conversion

(A) Candidate motor neuron transcription factors and factor combinations that co-expressed with miR-9/9*-124 in human fibroblasts.

(B) Immunocytochemistry of adult human fibroblasts overexpressing a non-specific miRNA (miR-NS) and ISL1/LHX3 (left) or miR-9/9*-124 and ISL1/LHX3 (right) for 35 days. Images demonstrate the necessity of miR-9/9*-124 for opening the neurogenic potential of human fibroblasts. Scale bar = 20μm.

(C) Additional representative inward/outward whole-cell currents and repetitive AP waveforms generated from whole cell patch clamp recordings of Moto-miNs. Images show the patch clamped cells.

Figure S6

A

Peak Inward Current

B

22 Yr miNs

22 Yr Moto-miNs

C

42 Female

56 Male
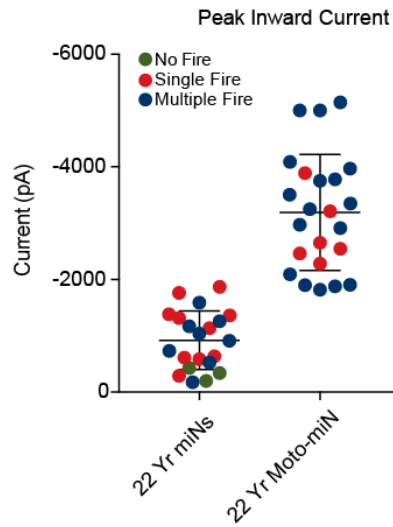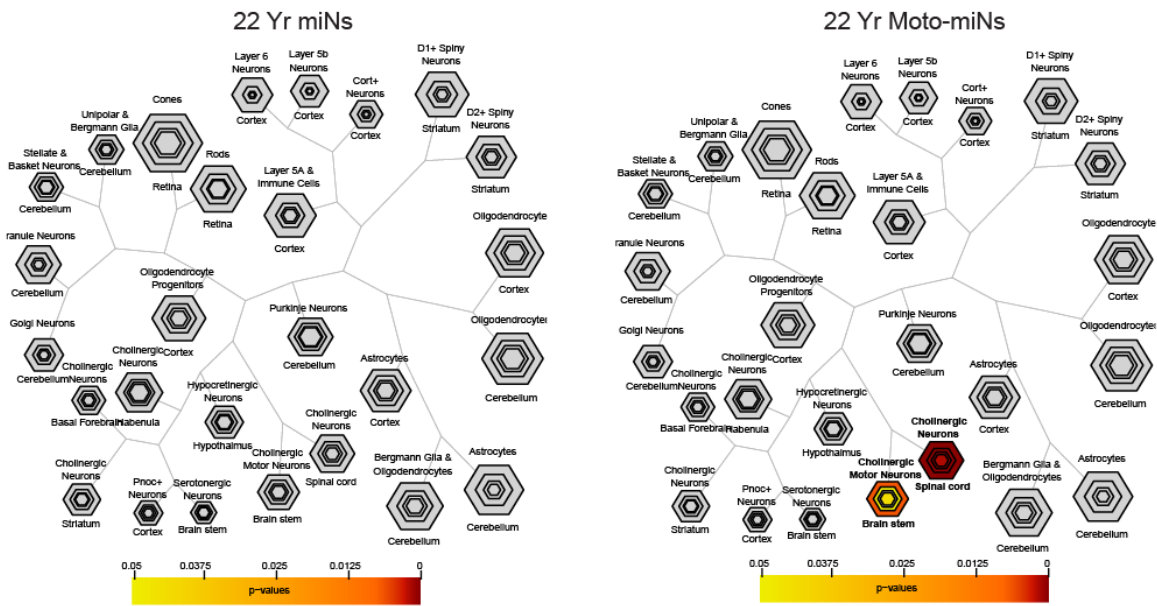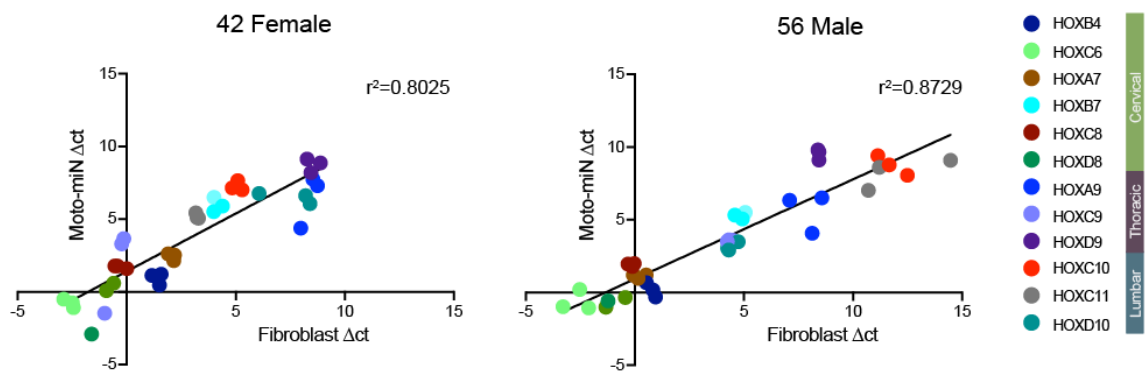
Figure S6, Related to Figure 7

Addition of ISL1/LHX3 Increases Functional Maturity and Generates Motor Neuron Transcriptional Network

(A) Peak inward current measured during voltage clamp mode of miNs and Moto-miNs reveals increased peak inward current in Moto-miNs (-3,189 pA ± 214 pA) compared to miNs (-919 pA ±113 pA). Data are represented as mean ± SEM.

(B) Cell Type-specific Enrichment Analysis (CSEA) tool reveals the top 100 most significantly expressed genes in miNs do not enrich for defined neuronal subtypes (left), while the top 100 most significantly expressed genes in Moto-miNs are enriched in cholinergic motor neurons in the brain stem and spinal cord.

(C) HOX gene expression analysis by qRT-PCR in 42 year old female and 56 year old male donor fibroblasts before and after conversion confirms that Moto-miNs retain donor fibroblast HOX gene expression patterns. Data represent Δct values for each biological replicate (3 separate Moto-miN conversions).

Figure S7

80
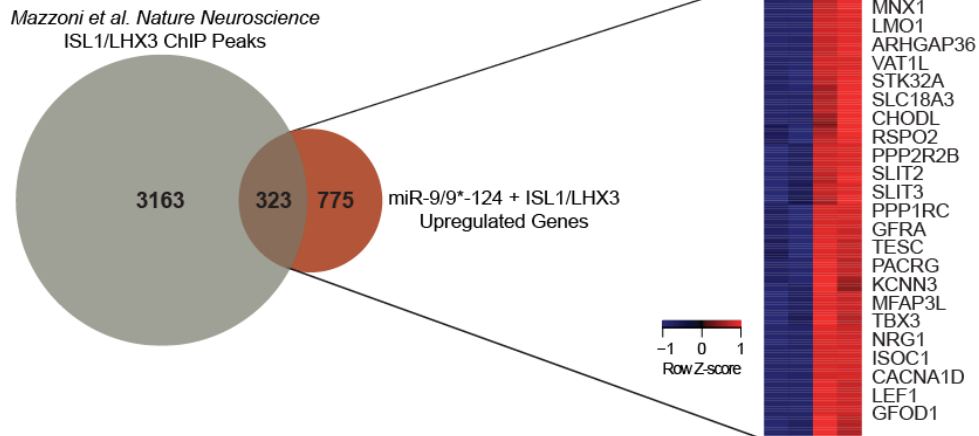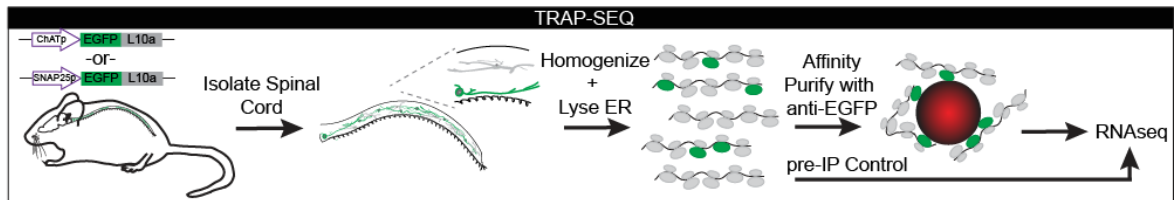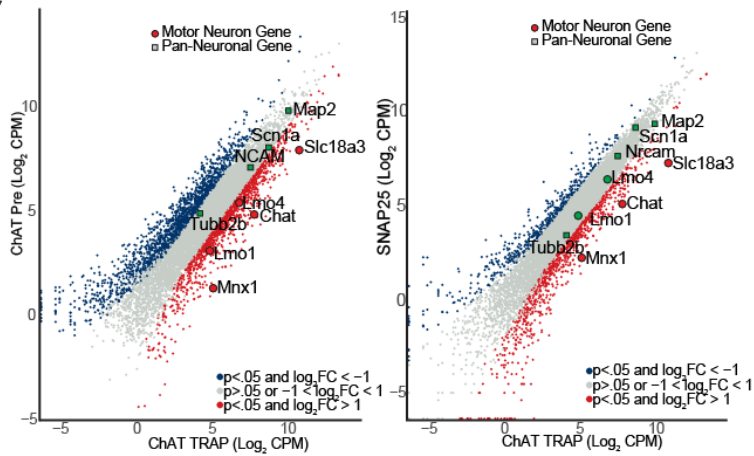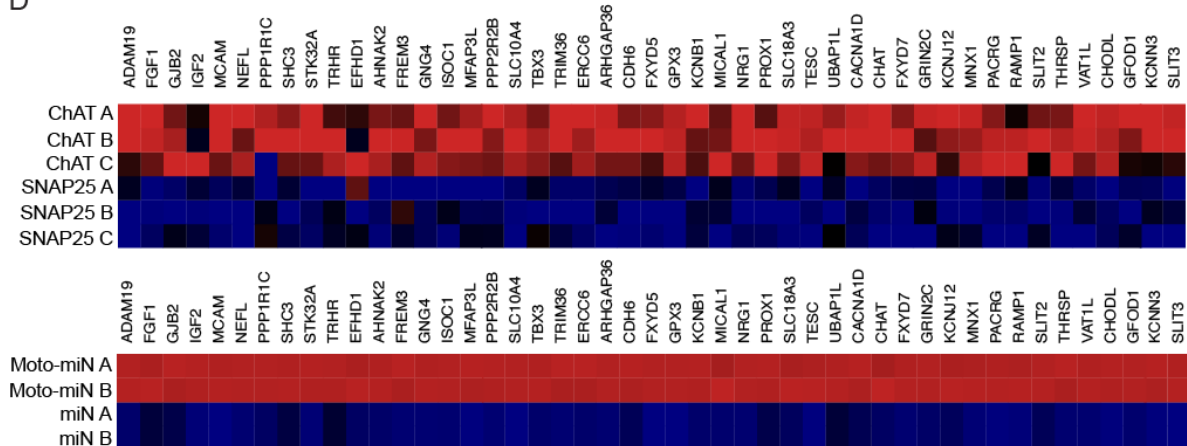
Figure S7, Related to Figure 7

Direct Comparison of Moto-miN Transcriptome to *in vivo* Mouse Motor Neurons by

Translating Ribosomal Affinity Purification (TRAP) Sequencing.

(A) Venn Diagram depicting the number of ISL1/LHX3 ChiP-seq peaks identified by *Mazzoni et al.* during ISL1/LHX3 directed ES to motor neuron differentiation (3,486) and genes enriched in Moto-miN transcriptome (775). Heatmap shows overlapping activated genes (323) include hallmark motor neuron markers.

(B) Schematic of TRAP-Seq strategy used to identify transcripts in all neurons (SNAP-25 genetic driver) and motor neurons (CHAT genetic driver) in mouse spinal cord. TRAP is a method to precipitate actively translated mRNA bound to ribosomes using an antibody to EGFP-L10A.

Materials and Methods

Cell Culture

Cell culture and differentiation was performed as described in Chapter 3.

Myotube Differentiation

Human myotubes were generated by differentiating human myoblasts using defined culture conditions (Steinbeck et al., 2016). Briefly, human skeletal myoblasts were cultured according to manufacturers recommendations (HSMM; CC-2580, Lonza) then were plated on matrigel (.1mg/mL) coated 12mm glass coverslips at a density of 80,000 cells/well. The following day HSMM's were differentiated by switching media to skeletal muscle differentiation media comprised of a 1:1 mixture of DMEM F12 (Gibco) and Complete Neuronal Media + 2% Horse Serum (Gibco). Every 2 days 75% of the media was replaced with fresh differentiation media. After 10-14 days of differentiation, Moto-miNs 14 days into conversion and labeled with synapsin-eGFP via lentiviral transduction were replated onto the established myotubes at a 1:1 ratio (i.e. one 12mm Moto-miN coverslip was replated on top of a 12mm myotube coverslip). The following day media was changed to complete neuronal media and cells were cultures for 2 weeks. Dox was replenished every two days and half the media was changed every 4 days. After two weeks cells were fixed with 4% paraformaldehyde and processed for immunocytochemistry and a-Bungarotoxin staining.

Electrophysiology

Electrophysiology was performed as described in Chapter 3 of this thesis.

MicroArray analyses

Total RNA was extracted from miNs and Moto-miNs derived from 22 yr old donor fibroblasts alongside corresponding starting fibroblast controls using TRIzol (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's instruction followed by extraction using chloroform and then ethanol precipitation. RNA quality was determined by the ratio of absorbance at 260 nm and 280 nm to be approximately 2.0. Samples for RNA microarray were then standardly prepped and labeled with Illumina TotalPrep kits (Thermo Fisher Scientific, Waltham, MA) for Agilent Human 4x44Kv1. Standard hybridization and imagine scanning procedure were performed according to the manufacturer's protocol at Genome Technology Access Center at Washington University School of Medicine, St. Louis. The intensity of the probes was imported into Partek and quantile normalized. Differentially expressed genes were identified using Partek with a cut-off of adjusted p-value <0.05 and over 2.5 $\log_2$ fold expression change.

Translating ribosome affinity purification (TRAP)

Translating ribosome affinity purification (Heiman et al., 2014) was performed on spinal cord dissections pooled from 3-4 mice 21 days post birth that were positive for the eGFP-L10A fusion ribosomal marker protein under either the Chat promoter (Tg(Chat-EGFP/Rpl10a)DW167Htz) or the Snap25 promoter (Tg(Snap25-EGFP/Rpl10a)JD362Jdd). TRAP samples underwent immunopurification for four hours at 4°C. Both TRAP and pre-immunopurification control RNA samples were extracted through TRIzol purification, DNase treatment, and Qiagen RNeasy Mini columns (74104). Quality and quantity of RNA was assessed using a Bioanalyzer 2100 RNA Pico Chip. Sequencing libraries were amplified using Nugen Amplification Kit Ovation® RNA Seq System V2 (7102). Genome Technology Access Center at Washington University in St. Louis performed adapter ligation and sequencing of the

libraries on the Illumina Hiseq2500. Three replicates of this procedure were analyzed.

Analysis of TRAP RNA-Seq data

RNA-Seq reads were mapped to Ensembl release 76 using STAR (analysis performed by Genome Technology Access Center at Washington University in St. Louis). For downstream analyses, only those genes with >1 CPM in at least 3 samples, with an Ensembl gene biotype of "protein_coding," were retained. For gene symbols mapping to multiple Ensembl gene IDs, only the ID with the highest number of mapped reads was retained, resulting in a total of 14,009 genes used for downstream analyses. Using edgeR, read counts were fit to a negative binomial generalized log-linear model, and a likelihood ratio test was done to determine differential expression.

Comparative analysis of RNA-Seq and microarray data

For comparative analysis, only probes with a detected call in at least 1 of 6 samples was retained, resulting in 23,775 probes mapping to a gene symbol. Expression level was then averaged over all probes for each gene, resulting in a total of 15,333 genes that were used for comparative analysis, 10,736 of which were also present in the gene set retained from the RNA-seq dataset (described above) after CPM filtering. Within the genes retained in both datasets, the top differentially expressed genes between motor neurons and controls - CHAT IP vs. SNAP25 IP (logFC > 1 and $p < 0.05$) in the RNA-seq dataset, and Moto-miN vs. miN (logFC >2.5 and $p < .05$) in the microarray dataset - were assessed for significant contingency using a one-tailed Fisher's exact test.

Quantitative PCR

Total RNA was extracted using TRIzol (Invitrogen, USA) according to the manufacturer's instruction. Reverse-transcribed complementary DNA (cDNA) was synthesized

from 500ng of RNA with SuperScript III First-Strand Synthesis SuperMix (Invitrogen, USA) or

from 10 ng of RNA for microRNAs expression analyses using specific stem-loop primer probes

from TaqMan MicroRNA Assays (Invitrogen, USA). Subsequently, the cDNA was analyzed on

a StepOnePlus Real-Time PCR System (AB Applied Biosystems, Germany). Expression data

were normalized to housekeeping genes *HPRT1* and *RNU44* for coding genes and microRNAs,

respectively, and analyzed using the $2^{-\Delta\Delta CT}$ relative quantification method. The following primers

were utilized:

miRNA qRT-PCR primers

hsa-miR-218 (Thermo Fisher Cat. # 4427975)

RNU-44 (Thermo Fisher Cat. # 4427975)

For primers utilized for mRNA qRT PCR see Table S2.

Overlap with ISL1/LHX3 ChIP Seq

To identify the genes regulated by LHX3 and ISL1 in motor neurons, ISL1- and LHX3-

ChIP sequencing data (Mazzoni et al., 2013) were used. We selected the regions co-occupied by

ISL and LHX3 during ES to motor neuron differentiation, accounting for 84.2% of peak regions

called in each ChIP-seq data. Based on the peaks co-occupied by ISL1 and LHX3, we annotated

3,486 closest genes with peaks located within 5Kb upstream of TSS and intragenic regions.

Comparing those annotated genes with genes selectively enriched in Moto-miNs vs miNs ($\log_2$

fold change=>2.5, p<0.01), identified 323 genes co-occupied by ISL1 and LHX3 that are also

upregulated when miR-9/9*-124 is co-expressed with ISL1/LHX3.

Data Availability

The RNA-seq, ATAC-seq, MeDIP and MRE-seq, and microarray data generated in this study

are: 14 samples, single-ended RNA-seq libraries from starting fibroblasts and miNs throughout

direct conversion (Days 3, 6 ,10, 20 and 30), 2 replicates per timepoint; 10 pair-ended ATAC-seq libraries from day 10 Ctrl, day 10 and day 20 miNs, day 20 miNs with shBRG1 and day 20 miNs with an shCtrl, 2 replicates for each timepoint; 6 microarrays from starting fibroblasts, day 35 miNs, and day 35 moto-miNs, 2 replicates per condition; 6 samples of pair-ended MeDIP libraries and 6 samples of MRE single-ended libraries from day 10, 20, 30 miNs and corresponding Ctrls. The accession number for the RNA-seq, ATAC-seq, MeDIP and MRE-seq, and microarray data data reported in this paper is currently being processed and will be available in the paper cited at the beginning of this chapter. The accession number for the TRAP-seq data is GEO:GSE93412.

Chapter 4: Conclusions and Future Directions

Significance

One of the greatest barriers in the study and treatment of neuronal disorders and neurodegenerative diseases has been a lack of patient-specific neuronal subtypes that retain their age-associated phenotypes. The desire to procure these cells stems from the observation that many of these neuronal diseases are uniquely human and may require a cellular platform that is truly representative of the affected species. My thesis has been aimed at meeting this goal through miRNA-mediated direct conversion technology. While examples of this conversion modality had been demonstrated at and near the onset of this thesis, that is, the generation of cortical neurons (Yoo et al., 2011) and of medium spiny neurons (Victor et al., 2014), the general applicability of this reprogramming paradigm and it's molecular underpinnings remained unknown. To this end, I set out to determine whether miRNAs alone, i.e. in the absence of transcription factors, were capable of trans-differentiating human fibroblasts into functional neurons. Surprisingly, I found the cellular activities of miR-9/9* and miR-124 were sufficient to elicit direct neuronal conversion of adult human fibroblasts. More surprisingly, while probing the identity of converted neurons using gene expression profiling, I discovered these miRNA-induced neurons (miNs) do not possess neuronal subtype identity, suggesting that the activities of subtype-specific transcription factors utilized in previous studies requires a permissive environment brought about by the ectopic expression of miR-9/9*-124.

Exploring this permissive state provided molecular insight into the phenomena of direct conversion and cell-fate acquisition. In particular, the observation that many neuronal loci encoding genes important in diverse subtypes open during direct conversion by miR-9/9*-124 but are not expressed provided an epigenetic blueprint for generating other clinically relevant neuronal subtypes. The further derivation of motor neurons, highlighted in chapter 3, explores this hypothesis and opens the door to generating other neuronal subtypes.

Altogether, the significance of this work can be separated into two broad findings. First, despite the short sequence length of miRNAs, they are capable of carrying enough regulatory information to reprogram cellular identity at epigenetic, transcriptomic, and functional levels. This is a significant finding as the predominant view of miRNAs revolves around the notion of "fine-tuning" gene expression. My findings provide a counterpoint to this hypothesis, suggesting that miRNAs are indeed capable of radically altering gene expression.

The second significant finding, the ectopic expression of miR-9/9*-124 and the motor neuron transcription factors, ISL1 and LHX3, is sufficient to directly convert adult human fibroblasts from young and old individuals into motor neurons, provides a platform for motor neuron disease modeling.

Future Directions

What targets of miR-9/9*-124 are responsible for chromatin remodeling?

To date, miRNAs have not been shown to directly alter local chromatin structure. Thus, our current understanding of miRNA biology suggests that miR-9/9* and miR-124 likely repress the expression of chromatin remodeling genes. Indeed, during neuronal development miR-9* and miR-124 repress the expression of the chromatin regulating BAF-complex subunit, BAF53a.

This allows for the replacement of BAF53a with BAF53b as neuronal progenitors differentiate (Yoo et al., 2009). This switch is also observed during miRNA-mediated direct conversion (unpublished data). However, preventing this switch by overexpressing BAF53A during conversion does not inhibit neuronal conversion of fibroblasts. Therefore other chromatin remodeling genes likely play a role in the epigenetic changes observed during conversion. These genes may not be direct targets of miR-9/9*-124, but may be part of a regulatory cascade initiated by miR-9/9*-124 overexpression. To identify direct targets of miR-9/9*-124 a technique such as HITS-CLIP followed by high throughout sequencing could be employed. Coupled with RNA-SEQ, these two techniques would generate predicted miRNA-mRNA binding and correlative gene expression changes. To confirm predicted miRNA-target interactions a luciferase essay could be employed consisting of a luciferase-target-3'UTR fusion that is co-transfected with the predicted miRNA binding partner. Decreases in luciferase activity compared to a non-target control 3'UTR would validate an interaction between predicted miRNAs and their targets. To further demonstrate binding the target sequence within the 3'UTR should be mutated. If this restores luciferase activity it can be inferred that the miRNA of interest binds.

The gene expression studies described in chapter two of this thesis revealed widespread changes in genes responsible for diverse epigenetic regulatory programs. Many of these genes are not predicted to be targets of miR-9/9* or miR-124 suggesting their expression changes are the results of secondary or greater events downstream of these miRNAs. Nonetheless, their role in reprogramming could be systematically addressed using overexpression and knockdown studies. For example, I knocked down the expression of BRG1, the catalytic unit of the BAF chromatin remodeling complex, and performed both ICC and ATAC-seq to test the role of chromatin remodeling in conversion. Using a scheme such as this, it would be interesting to test

the individual roles of DNA-methylation writers and readers, histone proteins, histone modifiers, and chromatin remodeling genes whose expression changes during conversion. Paring down the list of differentially expressed epigenetic modifiers would represent a key step in understanding the contribution of each component in reprogramming and would begin to define pathways required for cell fate conversion. Ultimately, uncovering the genes that create or remove barriers in reprogramming may offer avenues for increasing the speed and efficiency of direct cell fate conversion.

Guiding Conversion to other Clinically Relevant Neuronal Subtypes.

Chromatin profiling by ATAC-seq revealed the opening of key subtype-specific loci in response to miR-9/9*-124 expression. Specifically, dopaminergic neuron markers (*TH and SLC6A3*), serotonergic neuron markers (*FEV, LMX1B and SLC6A4*), GABAergic neuron markers (*SLC6A1, SLC32A1, GAD2*), striatal medium spiny neuron marker (*PPP1R1B*), glutamatergic neuron markers (*GLUL and SLC1A6*), and cholinergic or motor neuron markers (*MNX1, CHAT, and SLC5A7*) opened but were not expressed. To date, three neuronal subtypes have been derived using miRNA-mediated direct conversion therefore the probability of generating neuronal subtypes marked by regions that open in response to miR-9/9*-124 is high. Towards this goal, I have cloned genes important in the development and specification of dopaminergic neurons into lentiviral vectors, for example, NURR1, IRX3, OTX2, and LMX1A. These genes will be expressed individually and in combination with one another in the background of miR-9/9*-124 and assayed for their ability to activate key dopaminergic markers such as TH. The acquisition of the dopaminergic fate will be further determined through gene expression and functional studies. As dopaminergic neurons are the major neuronal subtype affected in Parkinson's Disease (PD), the ability generate these cells from patients may provide a valuable tool in the study of PD. Additionally, establishing multiple cultures of Dopaminergic

neurons from multiples patients with pathological PD may provide a powerful cellular platform for drug screening. The output of such a screen would be the production of dopamine.

Does Trans-differentiation Transit Through Analog or Digital Cellular States?

Our time-course RNA and ATAC-seq studies revealed steady increases and decreases of signal intensity over time suggesting cell fate conversion occurs gradually until terminal neuronal identity has been acquired. However, at the time these experiments were conducted only entire populations of cells could be analyzed. RNA-seq in particular required approximately 200,000 cells per condition to yield sufficient quantities of RNA for sequencing. Thus, an alternative hypothesis that would explain the data is: direct conversion consists of multiple binary cell fate switches that occur over time non-synchronously towards a final neuronal identity. Such events would be masked at the population level. To determine whether trans-differentiation is binary or digital, single-cell RNA-seq could be employed at multiple timepoints during conversion. In collaboration with Dr. Samantha Morris at Washington University in St. Louis, we have begun working towards this goal. To begin, I collected 10,000 miNs and Moto-miNs at 5, 10 15, 20 and 30 days into reprogramming in addition to starting adult human fibroblasts. After all samples were collected and slowly frozen they were transferred to the Morris Lab where single-cell libraries were generated. At the time of writing these cells are in the process of being sequenced. To date, little is known about direct neuronal cell-fate conversion at the single cell level and the results of this experiment may reveal more definitive insights into the dynamics of direct cell-fate conversion by removing the mask of population effects. Further, specific genes that associate with early or failed commitment to the neuronal fate may constitute accelerants or barriers to the conversion process, respectively. Exploring the

roles of these genes further in conversion through activation or inhibition studies may provide methods to increase the efficiency and speed of conversion.

Can Local Cues Guide Subtype Acquisition?

All of our efforts have been aimed at identifying transcription factors capable of guiding neuronal conversion. The rationale behind this approach is that it is consistent and does not require complex growth media changes where cocktails of patterning molecules are added at defined timepoints. However this approach is also limiting in that it relies on known developmental pathways. It remains unknown if miNs can be terminally differentiated *in vivo* using local patterning cues provided by the developing neuroectoderm. By transplanting fibroblasts into different regions of the developing mouse brain we could simultaneously test the plasticity of the miRNA-induced neuronal state and identify spatio-temporal cues important in the establishment of neuronal subtypes. Two different approaches could be used to explore this question. First, 30 days after injecting miNs into defined regions, mouse brains could be harvested, sectioned and stained for human specific proteins, subtype specific markers, pan-neuronal markers, and fibroblast specific proteins. The combination of overlaps between these markers would reveal if conversion occurred and if subtype identity was acquired.

Can miRNAs illicit *in vivo* reprogramming?

After optimizing the stereotaxic injection technique, the possibility of converting non-neuronal cells to neurons *in vivo* could also be examined. The development of such *in vivo* reprogramming techniques could provide a therapeutic intervention for neuronal trauma or disease. The ideal cell to reprogram would be resident astrocytes due to their location and

tendency to proliferate upon injury. To determine if trans-differentiation occurs, lineage tracing

could be employed whereby a GFAP-Cre mouse line is crossed with a mouse strain that contains

a LoxP flanked STOP codon preceding RFP at the ROSA26 locus. Progeny containing both the

GFAP-Cre and LoxP-STOP-LoxP-RFP alleles would also contain RFP-positive astrocytes.

These mice would then be injected with a GFP-miR-9/9*-124 containing lentivirus at different

locations in the brain. After 2, 4 and 8 weeks mice would be sacrificed and their brains processed

for ICC. Cells that were transduced with the miRNA-containing construct would be identified by

GFP fluorescence. Transduction of astrocytes would be ascertained by presence of GFP and RFP

fluorescence. Co-staining with neuronal markers such as MAP2 or TUJ1 would test whether

GFP/RFP double positive cells converted into neurons. Obtaining proof of principle experiments

using miRNAs would open the gateway to guiding *in vivo* reprogramming of astrocytes to

defined neuronal subtypes through the addition of transcription factors.

Evaluate the capability of reprogrammed motor neurons to function properly *in vivo*

If reprogrammed cells are to be used for regenerative therapies they must be able to

function correctly *in vivo*. In particular, functional recovery mediated by cell replacement

strategies requires that transplanted cells both survive and restore lost connectivity. Several

studies have demonstrated efficient engraftment and functional recovery when using ES and iPS

derived neural progenitors in rat and mouse models of SCI (Lu et al., 2014a). The ability of *ex

vivo* derived post-mitotic neurons to integrate into the mammalian spinal cord has not been

explored. By supplying all of the required information for conversion genetically, our

reprogramming scheme may be well-suited for delivery of motor neurons, and motor neurons

alone, to sites of injury. Testing this hypothesis would require transplanting moto-miNs into the

spinal cord of immunodeficient mice at 1, 2, 3, and 4 weeks post-transduction to ascertain the most favorable timepoint for effective survival. After determining the time of transplantation that yields the highest amount of neuronal engraftment I would test if transplanted cells can integrate into the local circuitry by recording from spinal cord slices.

Upon verifying that induced neurons are capable of engrafting into the murine spinal cord, I would transplant cells into a mouse model of spinal cord injury. Spinal cord injury could be induced by a dorsal hemisection of the spinal cord in immunodeficient mice, an injury model that generates stereotyped locomotor deficits such as impaired paw placement during walking (Zorner et al., 2010). This phenotype has the potential to be rescued by cell transplantation and is less debilitating than other models of spinal cord injury (Zorner et al., 2010). I would then examine the extent of cell survival, axon outgrowth, myelination by the host and formation of synaptic structures 2 and 4 weeks post transplantation. Functional recovery could be assessed by a battery of locomotor tests and compared to results obtained from controls injected with post-mitotic fibroblasts and mock injections.

Approaches in transplanting neural progenitors into the mammalian nervous system have been met with mixed success (Lu et al., 2014a). Recently, the injection of cells contained within biomaterial scaffolds and supplemented with neurotrophic factors showed greatly enhanced survival and growth of transplanted cells (Lu et al., 2014a; Lu et al., 2014b). If poor survival and/or engraftment of transplanted cells is observed, the use of biomaterial scaffolds and growth factor supplementation to cell suspensions during injection may need to be explored. Restoring locomotor activity requires functional integration into the site of injury and surrounding circuit therefore, if the environment created by dorsal hemisection is too hostile for transplantation, or the injury too severe, additional models of spinal cord injury may need to be utilized.

94

Elucidate the mechanism of guiding reprogramming into motor neurons.

The process by which a fully differentiated fibroblast of mesodermal origin is directly converted into a functional motor neuron – a highly specialized cell normally arising from neuroectoderm - remains largely enigmatic. Exploring the etiology of cell fate conversion presents a unique opportunity in the study of lineage commitment. Simultaneously, understanding the molecular pathways activated during neural conversion is necessary for the interpretation of experiments utilizing reprogrammed cells. In addition to determining the manner in which transcription factors guide reprogramming towards motor neurons, identifying the logic and function of pro-neural transcription factors within the neurogenic state primed by miR-9/9*-124 may provide a molecular foundation for choosing transcription factors that may guide conversion towards additional clinically relevant neural subtypes.

To understand how miR-9/9*-124 and ISL1 and LHX3 are able to reprogram human fibroblasts into motor neurons, I would conduct a series of ChIP-Seq experiments to determine if genome binding requires miRNAs and determine the binding locations of each TF. From these data, I will examine the association of transcription factors with genes important in neurogenesis and motor neuron development. To explore the possibility that transcription factors are cooperating with one another, I would ask if transcription factors co-occupy distinct sites. Direct associations between reprogramming factors could then be tested using Co-IP assays. The relative contribution of endogenous and exogenous transcription factors within complexes may be resolved through the use of tagged exogenous genes. I have generated FLAG-HA tagged

versions of both LHX3 and ISL1 and shown they do not interfere with motor neuron fate acquisition.

The molecular underpinnings of direct conversion remain largely unexplored. This set of experiments is designed to gain insights into DNA accessibility, the activation of defined genetic networks, and the role of transcription factor cooperation in miRNA-mediated reprogramming. Since little is known about each of these processes in the context of cell fate conversion, these experiments could shed light on how one cell type is able to change into another without going through a pluripotent intermediate.

Impact and Closing Remarks

During the course of this dissertation I was able to induce and characterize a miRNA-induced neurogenic state in adult human fibroblasts. The insights provided through this in depth characterization are critical to further reprogramming efforts in the laboratory and clinic. This notion is exemplified by the generation of motor neurons directly from adult human fibroblasts. By further exploring the neurogenic ground state elicited by miRNAs alone, an epigenetic basis for subtype-specification emerged. Pursuing the hypothesis that miRNAs open the neurogenic potential of fibroblasts by enabling access to TF binding sites necessary for neuronal subtype development, led to the derivation of functional motor neurons directly from adult human fibroblasts.

The impact of this work revolves around the pioneering activities of miR-9/9*-124. Before this work, it was broadly accepted that miRNAs only "fine tune gene expression." Contrary to this idea, I have shown that despite their short sequence length, miRNAs contain

enough genetic information to radically reprogram the epigenome and transciptome. These activities lead to the generation of functional neurons, a property that I exploited to derive a clinically relevant neuronal subtype. Going forward, the method of generating motor neurons directly from adult human fibroblasts described within this thesis may pave the way to a greater understanding of motor neuron diseases such as ALS and SMA. Further, the epigenetic blueprint formed by miRNA overexpression that revealed by ATAC-seq could lead to the generation of additional neuronal subtypes.

References

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Caiazzo, M., Dell'Anno, M.T., Dvoretskova, E., Lazarevic, D., Taverna, S., Leo, D., Sotnikova, T.D., Menegon, A., Roncaglia, P., Colciago, G.*, et al.* (2011). Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. Nature *476*, 224-227.

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

Heiman, M., Kulicke, R., Fenster, R.J., Greengard, P., and Heintz, N. (2014). Cell type-specific mRNA purification by translating ribosome affinity purification (TRAP). Nat Protoc *9*, 1282-1291.

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

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Lu, P., Kadoya, K., and Tuszynski, M.H. (2014a). Axonal growth and connectivity from neural stem cell grafts in models of spinal cord injury. Curr Opin Neurobiol *27*, 103-109.

Lu, P., Woodruff, G., Wang, Y., Graham, L., Hunt, M., Wu, D., Boehle, E., Ahmad, R., Poplawski, G., Brock, J*., et al.* (2014b). Long-distance axonal growth from human induced pluripotent stem cells after spinal cord injury. Neuron *83*, 789-796.

Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y*., et al.* (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature *466*, 253-257.

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

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.

Ming, G.L., and Song, H. (2011). Adult neurogenesis in the mammalian brain: significant answers and significant questions. Neuron *70*, 687-702.

Richner, M., Victor, M.B., Liu, Y., Abernathy, D., and Yoo, A.S. (2015). MicroRNA-based conversion of human fibroblasts into striatal medium spiny neurons. Nat Protoc *10*, 1543-1555.

Ring, K.L., Tong, L.M., Balestra, M.E., Javier, R., Andrews-Zwilling, Y., Li, G., Walker, D., Zhang, W.R., Kreitzer, A.C., and Huang, Y. (2012). Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor. Cell Stem Cell *11*, 100-109.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res *43*, e47.

Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J.*, et al.* (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317-330.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Schulz, M.H., Devanny, W.E., Gitter, A., Zhong, S., Ernst, J., and Bar-Joseph, Z. (2012). DREM 2.0: Improved reconstruction of dynamic regulatory networks from time-series expression data. BMC Syst Biol *6*, 104.

Shin, H., Liu, T., Manrai, A.K., and Liu, X.S. (2009). CEAS: cis-regulatory element annotation system. Bioinformatics *25*, 2605-2606.

Smith, C.M., Finger, J.H., Hayamizu, T.F., McCright, I.J., Xu, J., Berghout, J., Campbell, J., Corbani, L.E., Forthofer, K.L., Frost, P.J.*, et al.* (2014). The mouse Gene Expression Database (GXD): 2014 update. Nucleic Acids Res *42*, D818-824.

Steinbeck, J.A., Jaiswal, M.K., Calder, E.L., Kishinevsky, S., Weishaupt, A., Toyka, K.V., Goldstein, P.A., and Studer, L. (2016). Functional Connectivity under Optogenetic Control Allows Modeling of Human Neuromuscular Disease. Cell Stem Cell *18*, 134-143.

Tripathi, S., Pohl, M.O., Zhou, Y., Rodriguez-Frandsen, A., Wang, G., Stein, D.A., Moulton, H.M., DeJesus, P., Che, J., Mulder, L.C.*, et al.* (2015). Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding. Cell Host Microbe *18*, 723-735.

Victor, M.B., Richner, M., Hermanstyne, T.O., Ransdell, J.L., Sobieski, C., Deng, P.Y., Klyachko, V.A., Nerbonne, J.M., and Yoo, A.S. (2014). Generation of human striatal neurons by microRNA-dependent direct conversion of fibroblasts. Neuron *84*, 311-323.

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

Wise, A., and Bar-Joseph, Z. (2015). cDREM: inferring dynamic combinatorial gene regulation. J Comput Biol *22*, 324-333.

Yoo, A.S., Staahl, B.T., Chen, L., and Crabtree, G.R. (2009). MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. Nature *460*, 642-646.

Yoo, A.S., Sun, A.X., Li, L., Shcheglovitov, A., Portmann, T., Li, Y., Lee-Messer, C., Dolmetsch, R.E., Tsien, R.W., and Crabtree, G.R. (2011). MicroRNA-mediated conversion of human fibroblasts to neurons. Nature *476*, 228-231.

Zhang, B., Zhou, Y., Lin, N., Lowdon, R.F., Hong, C., Nagarajan, R.P., Cheng, J.B., Li, D., Stevens, M., Lee, H.J.*, et al.* (2013). Functional DNA methylation differences between tissues, cell types, and across individuals discovered using the M&M algorithm. Genome Res *23*, 1522-1540.

Zorner, B., Filli, L., Starkey, M.L., Gonzenbach, R., Kasper, H., Rothlisberger, M., Bolliger, M., and Schwab, M.E. (2010). Profiling locomotor recovery: comprehensive quantification of impairments after CNS damage in rodents. Nat Methods *7*, 701-708.

Supplemental Table 1, Key Resources Table

REAGENT or	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-MAP2	Sigma-Aldrich	Cat# M4403; RRID: AB_477193
Rabbit polyclonal anti-TUBB3B	Covance Research Products Inc	Cat# PRB-435P-100; RRID: AB_291637
Chicken polyclonal anti- NeuN	Aves Labs	Cat# NUN, RRID:AB_2313556
Mouse monoclonal anti-SCN1A	Sigma-Aldrich	Cat# S8809, RRID:AB_477552
Mouse monoclonal anti-ANKG	UC Davis/NIH NeuroMab Facility	Cat# 73-146, RRID:AB_10697718
Mouse monoclonal anti-SV2	DSHB	Cat# SV2, RRID:AB_2315386
Mouse monoclonal anti-MNX1	DSHB	Cat# 81.5C10, RRID:AB_2145209
Goat polyclonal anti-CHAT	Millipore	Cat# AB144P, RRID:AB_2079751
Mouse monoclonal anti-SMI-32	BioLegend	Cat# 801701, RRID:AB_2564642
Rabbit polyclonal anti-Ki-67	Abcam	Cat# ab15580, RRID:AB_443209
Mouse monoclonal anti-Myosin	DSHB	Cat# A4.1025, RRID:AB_528356
Mouse monoclonal anti-NCAM	Santa Cruz Biotechnology	Cat# sc-106, RRID:AB_627128
Mouse monoclonal anti-5-Methylcytidine	Eurogentec	Cat# BI-MECY-0100, RRID:AB_2616058
Rabbit polyclonal anti-Mouse IgG	Jackson ImmunoResearch Labs	Cat# 315-001-003, RRID:AB_2340031

Monoclonal anti-EGFP	Memorial Sloan-Kettering Monoclonal Antibody Facility	Clone 19C8
Monoclonal anti-EGFP	Memorial Sloan-Kettering Monoclonal Antibody Facility	Clone 19F7
Chemicals, Peptides, and Recombinant Proteins		
Fetal bovine serum (FBS), qualified, US Department of Agriculture (USDA)-approved regions	Life Technologies	Cat# 10437028
Polybrene	Sigma-Aldrich	Cat# H9268
Neuronal Media	ScienCell Research Laboratories	Cat# 1521
Doxycycline hyclate (Dox)	Sigma-Aldrich	Cat# D9891
Poly-L-ornithine solution	Sigma-Aldrich	Cat # P4957
Laminin	Sigma-Aldrich	Cat# L2020
Fibronectin,	Sigma-Aldrich	Cat# F4759
Valproic acid (VPA), sodium salt,	Sigma-Aldrich	Cat# 676380
Dibutyryl-cAMP sodium salt,	Sigma-Aldrich	Cat# D0627
Retinoic acid (RA)	Sigma-Aldrich	Cat# R2625
Pierce Protein A/G Agarose	ThermoFisher Scientific	Cat# 20421
Ampure XP beads	Ampure	Cat# A63880
Critical Commercial Assays		
RNeasy plus micro kit	Qiagen	Cat# 74034
SMARTer Ultra Low RNA Kit for Illumina sequencing	Clontech	Cat# 635029
Illumina Genomic DNA Sample Prep Kit	Illumina	
Illumina TotalPrep kit	ThermoFisher Scientific	Cat# AMIL1791
Whole Human Genome Microarray Kit, 4x44K	Agilent Technologies	Cat# G4112F
QIAquick PCR Purification Kit	Qiagen	Cat# 28104
MinElute Gel Extraction Kit	Qiagen	Cat# 28604
RNeasy Mini Kit	Qiagen	Cat# 74104

Nugen Amplification Kit Ovation® RNA-Seq System V2	NuGEN	Cat# 7102
Deposited Data		
Experimental Models: Mice		
Tg(Chat-EGFP/Rpl10a)DW167Htz	PMID: 19013282	N/A
Tg(Snap25-EGFP/Rpl10a)JD362Jdd	PMID: 22865458	N/A
Experimental Models: Cell Lines		
Primary 22 year old human adult fibroblasts	NIGMS Human Genetic Cell Repository at the Coriell Institute for Medical Research	GM02171
Primary 42 year old human adult fibroblasts	Washington University in St. Louis School of Medicine iPSC core facility	F09-238
Primary 56 year old human adult fibroblasts	NIA Aging Cell Repository at the Coriell Institute for Medical Research	AG04148
Primary 68 year old human adult fibroblasts	NINDS Cell Line Repository at the Coriell Institute for Medical Research	ND34769
Primary neonatal human dermal fibroblasts	ATCC	PCS-201-010
Lenti-X 293LE cell line	Clontech	Cat# 632180
Human skeletal myoblasts (HSMM)	Lonza	Cat# CC-2580
Recombinant DNA		
N174	Addgene	60859
N106	Addgene	66808
pMD2.G	Addgene	12259
psPAX2	Addgene	12260
pT-BCL-9/9*-124	Addgene	60859
pT-BCL-NS	(Victor et al., 2014)	N/A
LHX3-N174	This paper	
ISL1-N174	This paper	
pLKO-shBRG1	This Paper	

pLKO shEGFP	This Paper	
synEGFP	(Victor et al., 2014)	N/A
Software and Algorithms		
pCLAMP 10 Software	Molecular Devices	https://www.moleculardevices.com/systems/conventional-patch-clamp/pclamp-10-software
Graphpad Prism 7	GraphPad Software Inc	http://www.graphpad.com/
edgeR 3.4	(Robinson et al., 2010)	https://bioconductor.org/packages/release/bioc/html/edgeR.html
limma	(Ritchie et al., 2015)	https://bioconductor.org/packages/release/bioc/html/limma.html
Metascape	(Tripathi et al., 2015)	http://metascape.org/gp/index.html#/main/step1
methylQA	N/A	http://methylqa.sourceforge.net/
M&M statistical model	(Zhang et al., 2013)	http://epigenome.wustl.edu/MnM/methylMnM.pdf
GREAT package	(McLean et al., 2010)	http://bejerano.stanford.edu/great/public/html/
HOMER software	(Heinz et al., 2010)	http://homer.salk.edu
TrimGalore, version 0.4.2	Felix Krueger, Babraham Bioinformatics	http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
bowtie2	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Picard Tools	Broad Institute	https://broadinstitute.github.io/picard/
HTSeq count	(Anders et al., 2015)	http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html
CEAS software	(Shin et al., 2009)	http://liulab.dfci.harvard.edu/CEAS/
STAR	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
DREM	(Schulz et al., 2012)	http://www.sb.cs.cmu.edu/drem/
cDREM	(Wise and Bar-Joseph, 2015)	http://sb.cs.cmu.edu/cdrem/

Supplemental Table 2, Primer sequences

used for qRT PCR analysis

Primer Name	Sequence (5'-3')
S100A4 FWD	GATGAGCAACTTGGACAGCAA
S100A4 REV	CTGGGCTGCTTATCTGGGAAG
VIM FWD	AGTCCACTGAGTACCGGAGAC
VIM REV	CATTTCACGCATCTGGCGTTC
COL13A1 FWD	GGAGACGGCTATTTTGGGACG
COL13A1 REV	TCCTTGAGTGGAGCTTCCATT
ChAT FWD	TCAATCATGTCCAGCGAGTC
ChAT REV	AACGAGGACGAGCGTTTG
HB9 FWD	CTCCTACTCGTACCCGCAG
HB9 REV	TTGAAGTCGGGCATCTTAGGC
SLC18A3 FWD	TTCGCCTCTACAGTCCTGTTC
SLC18A3 REV	GCTCCTCCGGGTACTTATCG
HOXB4 FWD	CGTGAGCACGGTAAACCCC
HOXB4 REV	CGAGCGGATCTTGGTGTTG
HOXC6 FWD	ACAGACCTCAATCGCTCAGGA
HOXC6 REV	AGGGGTAAATCTGGATACTGGC
HOXA7 FWD	CGTTCCGGGCTTATACAATGT
HOXA7 REV	CTCGTCCGTCTTGTCGCAG
HOXB7 FWD	TTCCCAGAACAAACTTCTTGTGC
HOXB7 REV	GCATGTTGAAGGAACTCGGCT
HOXC8 FWD	ACCGGCCTATTACGACTGC
HOXC8 REV	TGCTGGTAGCCTGAGTTGGA
HOXD8 FWD	GGAAGACAAACCTACAGTCGC
HOXD8 REV	TCCTGGTCAGATAGGGGTTAAAA
HOXA9 FWD	TACGTGGACTCGTTCCTGCT

HOXA9 REV	CGTCGCCTTGGACTGGAAG
HOXC9 FWD	ACTCGCTCATCTCTCACGACA
HOXC9 REV	GACGGAAAATCGCTACAGTCC
HOXD9 FWD	GGACTCGCTTATAGGCCATGA
HOXD9 REV	GCAAAACTACACGAGGCGAA
HOXC10 FWD	ACATGCCCTCGCAATGTAACT
HOXC10 REV	GAGAGGTAGGACGGATAGGTG
HOXC11 FWD	ATGTTTAACTCGGTCAACCTGG
HOXC11 REV	GCATGTAGTAAGTGCAACTGGG
HOXD11 FWD	TCTCCGAGTCCTCGTGGGGA
HOXD11 REV	GCAAAACACCAGCGCCTTCTA
HPRT FWD	TCCTTGGTCAGGCAGTATAATCC
HPRT REV	GTCAAGGGCATATCCTACAACAAA