Title
Multilevel regulation of the let-7 miRNAs coordinates human central nervous system developmental maturation

Author
Gaeta, Xavier

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Multilevel regulation of the *let-7* miRNAs coordinates human central nervous system developmental maturation

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular, Cell, and Developmental Biology

by

Xavier Gaeta

2016

ABSTRACT OF THE DISSERTATION

Multilevel regulation of the *let-7* miRNAs coordinates human central nervous system developmental maturation

by

Xavier Gaeta

Doctor of Philosophy in Molecular, Cell, and Developmental Biology

University of California, Los Angeles, 2016

Professor William Edward Lowry, Chair

Pluripotent Stem Cells (PSCs) have the unique ability to divide indefinitely and self-renew *in vitro*, as well as differentiate into all cell types of the body. These cells present intriguing opportunities to study development and biology in novel ways, and can be utilized to produce large amounts of cells with therapeutic applications. In order to safely use PSCs to model difficult to study diseases, or to replace injured or diseased tissues in patients, we need to ensure that the *in vitro* differentiation process is effective. However, the field has encountered great difficulty in making fully differentiated, mature cells *in vitro* from PSCs. While it is currently possible to differentiate PSCs into many types of central nervous system (CNS) cells, we and others have shown that CNS cells derived from PSCs are immature and are most epigenetically and functionally similar to cells

from the CNS at a mid-gestational time point. Since the developing nervous system is tightly temporally regulated in terms of division, differentiation, organization, and connection, this roadblock currently poses a challenge to the utility of PSC-derived CNS cells.

We have identified a molecular circuit consisting of LIN28 proteins and the *let-7* microRNAs that regulates differentiation and maturation in the CNS and other organ systems. In this dissertation I present data to demonstrate that manipulation of the LIN28/*let-7* circuit can induce functional maturation in CNS progenitors by acting through structural epigenetic protein HMGA2 and the Notch signaling pathway. Furthermore, by exploring the regulation of the *let-7* miRNAs at the level of transcription, I propose that dynamic temporal regulation of some *let-7* family members is an important driver of change in the LIN28/*let-7* circuit. Finally, I describe efforts to generate PSC-derived inhibitory interneurons and to characterize their aging and maturation *in vitro* and *in vivo*, with the aim of understanding the mechanisms of post-mitotic neuronal functional maturation. These studies will pave the way for the generation of appropriately mature PSC-derived CNS cells, and may be broadly applicable toward generating mature cells from PSCs for the purposes of biological study or therapeutic utility.

The dissertation of Xavier Gaeta is approved.

Dana Leanne Jones

Kathrin Plath

Alvaro Sagasti

William Edward Lowry, Committee Chair

University of California, Los Angeles

2016

This dissertation is dedicated to

Henry Servín Sr., Lupe G. Servín, Ray R. Gaeta Sr., and Helen R. Gaeta

Table of Contents

Acknowledgements xii

Vita xvi

Chapter 1: Introduction 1

Multicellular organisms require specialized cells 2

Pluripotent Stem Cells 2

Epigenetics: the regulation of gene expression 3

miRNAs: small, regulatory, non-coding RNAs 4

The *let-7* miRNAs 5

The LIN28 proteins regulate *let-7* miRNA biogenesis and form a conserved bistable switch 7

Development of the Central Nervous System 10

Neural cells are generated along a specific time course 11

Persistent differences remain between PSC-derived and *in vivo* differentiated cells 12

Figures 19

References 22

Chapter 2: *let-7* miRNAs can act through Notch to regulate human gliogenesis 28

Introduction 29

Results 30

PSC-NPCs are functionally and transcriptionally distinct from Tissue-derived counterparts……30

let-7 expression and processing across development……32

Manipulation of LIN28A/B shows a modest effect on developmental maturity of NPCs……34

Direct manipulation of *let-7* levels alters cell fate in neural progenitors……34

let-7 acts through HMGA2 to control cell fate……36

The Notch pathway is a key effector of *let-7* in gliogenesis……36

The *let-7* effect on gliogenesis extends to oligodendrocyte formation……38

Discussion……38

Experimental Procedures……40

Cell Culture……40

Transfection……40

Immunostaining……40

Acknowledgements……42

References……43

Supplemental Information……45

Chapter 3: Defining transcriptional regulatory modules for *let-7* pri-miRNAs..... 57

Abstract 58

Introduction 59

Results 61

Discussion 66

Materials and Methods 68

Cell Culture 68

siRNA transfection 68

Measurements of gene expression by RT-PCR 68

Epigenome characterization and candidate TF prediction 69

Figures 70

References 79

Chapter 4*: In vitro* generation of Human PSC-derived Interneurons and progress toward elucidating strategies and mechanisms for their maturation 83

Introduction 84

Results 87

In vitro differentiation of human PSC-derived inhibitory interneurons 87

Electrophysiological characterization of PSC-derived interneurons 90

PSC-derived interneurons injected into the mouse striatum 91

Discussion 91

Materials & Methods 94

In vitro differentiation of interneurons from pluripotent stem cells 94

Fluorescence-Activated Cell Sorting of interneurons 95

Single-cell RNA sequencing 96

Electrophysiology 97

Intracranial injection of human PSC-derived interneurons 97

Figures 98

References 107

Chapter 5: Conclusions 109

Progress toward understanding mature differentiation and the intrinsic clock 110

Mature PSC-derived cells and regenerative medicine 112

References 114

List of Figures

Chapter 1: Introduction

Figure 1-1: let-7 miRNA biogenesis, processing, and LIN28-mediated regulation……..19

Figure 1-2: Phylogenetic conservation and expansion of the let-7 miRNAs……..20

Table 1-1: Consistently differentially expressed genes between PSC-derived progeny and their tissue-derived counterparts……..21

Chapter 2: let-7 miRNAs can act through Notch to regulate human gliogenesis

Figure 2-1: *let-7* activity correlates with human gliogenesis……31

Figure 2-2: Dynamics of *let-7* expression and processing between PSC-NPCs and Tissue-NPCs…..33

Figure 2-3: A role for LIN28/*let-7* in the developmental progression of human Tissue-NPCs…….35

Figure 2-4: Direct introduction of *let-7* miRNAs affects developmental progression of NPCs……37

Figure 2-5: HMGA2 is a critical target of *let-7* in developmental progression……39

Figure 2-6: *let-7*/HMGA2 regulates Notch sensitivity through HES5……41

Figure 2-7: Generation of oligodendrocytes is facilitated by induction of *let-7*……42

Supplemental Figure 2-1: Validation of PSC-NPCs and comparison to Tissue-NPCs……52

Supplemental Figure 2-2: Annotation of *let-7* family members within human genome……53

Supplemental Figure 2-3: Manipulation of LIN28A/B in NPCs……54

Supplemental Figure 2-4: HES family expression across neural development and various human cell types……..55

Supplemental Table 2-2: mRNA changes in PSC NPCs after HMGA2 KD……56

Chapter 3: Defining transcriptional regulatory modules for let-7 pri-miRNAs

Figure 3-1: Dynamic transcriptional regulation of some *pri-let-7* transcripts 70

Figure 3-2: Dynamically and constitutively transcribed *let-7* loci show distinct epigenetic signatures 72

Figure 3-3: Transcription factors predicted to bind to the *let-7a3/b* promoter regulate primary *let-7a3/b* transcription 74

Figure 3-4: FOX proteins are predicted to bind to putative *let-7a3/b* enhancer regions, and affect *let-7* miRNA transcription 77

Chapter 4: In vitro generation of Human PSC-derived Interneurons and progress toward elucidating strategies and mechanisms for their maturation

Figure 4-1: Expression of interneuron markers in the human brain during development 98

Figure 4-2: *In vivo* differentiation of inhibitory interneurons from pluripotent stem cells. 100

Figure 4-3: Single-cell RNA-sequencing reveals heterogeneity of gene expression in human PSC-derived interneurons 102

Figure 4-4: Co-culture of human PSC-derived interneurons with mouse cortical neurons promotes electrophysiological maturation 103

Figure 4-5: Human PSC-derived NPCs can be injected into adult mouse striatum and recovered by Fluorescence-Activated Cell Sorting 105

Acknowledgements

Pursuing a career in research has not been without its challenges, and I have many people to thank for their help and inspiration in getting me to this point.

I would like to thank and acknowledge my PhD advisor and research mentor, Dr. Bill Lowry, for his support and guidance over the past five years. My first inkling that UCLA would be a fulfilling place to conduct my graduate research came during an interview, when I met with Bill for the first time. In that short meeting I was struck by his earnest curiosity, his thoughtful nature, and his ability to always ask just the right question to reach the heart of an issue or to spark a flood of ideas. Through my work with Bill I have grown immensely as a scientist, a communicator, a writer, and as a person. Bill has modeled the type of career I hope to build – deeply investigating crucial but basic scientific questions, following the lead of data with hypothesis-driven inquiry, and strengthening bridges in the scientific community through rich collaboration. The type of lab environment Bill oversees does not develop in a vacuum, and certainly not without a deep commitment to promoting independent thought, creativity, teamwork, equality, and trust.

My thesis committee, Dr. Alvaro Sagasti, Dr. Leanne Jones, and Dr. Kathrin Plath, have also been instrumental in guiding my research progress and my development as a scientist through their questions and their advice.

I would also like to thank my labmates for the boundless support, inspiration, commiseration, and joy they have graciously shared with me. The time we have shared together, whether over experiments and data or impromptu tissue culture karaoke, weekend trips, half-marathon

training, weddings, and mountains of food, has been unparalleled. Thanks to Dr. Michaela Patterson Veldman, Dr. Yuan Xie, Dr. David Chan, Minori Ohashi, Dr. Andrew White, Dr. Peiyee Lee, Dr. Igal Germanguz, Aimee Flores, Matilde Miranda, Jessica Cinkornpumin, Aryeh Solomon, Ying Lin, Soheila Azghadi, Dr. Jim Rheinwald**,** Sarah Gomez, Yan Cui, Eden Maloney, Jeanny Hu, Anqi Liu, Christine Dang, Joan Khuu, Benni Vargas, and Jenny Park.

In particular, I have been incredibly fortunate to work with several talented, driven, and curious undergraduate students over the course of my graduate career. I owe Kimberly Loo, Luat Le, and Namrata Kakade a great debt not only for their help with experiments, but also for their friendship. I have particularly cherished the opportunity to teach and share not just about our research approaches and technical aspects of protocols, but about graduate school and life after college.

My research experience would not have been nearly full, rewarding, or productive without excellent scientific and technical collaborators. Thanks to Dr. Kathrin Plath, Dr. Heather Christofk, Dr. Stephen Smale, Dr. S. Thomas Carmichael, Dr. Istvan Mody, Dr. Amy Rowat, Anna Sahakyan, Dr. Thomas Allison, Dr. Miguel Edwards, Dr. Xiaofei Wei, Dr. Irene Llorente, Wen Gu, Kendra Nyberg, Angelyn Nguyen, Felicia Codrea, Min Zhao, Jeff Calimlim, and Dr. Xinmin Li.

I received constant support from staff and administration within the UCLA MCDB and MBI departments, and the UCLA MSTP including from Dr. Pamela Hurley, Jennifer Miller, and Susie Esquivel. The past and present leadership of the UCLA MSTP, including Dr. Kelsey Martin, Dr. Stephen Smale, Dr. Carlos Portera-Cailliau, and Dr. Leanne Jones, have all been instrumental in

granting me the incredible opportunity to delve deeply into both biomedical science and medicine, and in demonstrating the breadth of what it means to be a physician scientist.

Finally, I would like to extend my deepest gratitude to my family and friends, without whose love and foundational support I would not be here. To my parents, Dr. Raymond and Isaura Gaeta, I am endlessly appreciative. Thanks for nurturing my curiosity at its inception, and for generously providing me with the enthusiastic and caring support to let me strive for my goals. To my grandparents; Henry and Lupe Servín, and Raymond Sr. and Helen Gaeta; thanks for your selfless efforts to provide our family with the tools to succeed and thrive. *Les quiero muchísimo*. To Marcos Gaeta, Andrea Gaeta, and Elisa Nasol, thanks for all you have given me, taught me, and influenced me to be better and work harder through your love, your example, your listening ears, and your open arms.

Chapter 2 was originally published in *Stem Cell Reports*. Patterson M*, Gaeta X**, Loo K, Edwards M, Smale S, Cinkornpumin J, Xie Y, Listgarten J, Azghadi S, Douglass SM, Pellegrini M, Lowry WE. 2014. **let-7 miRNAs can act through Notch to control human gliogenesis.** *Stem Cell Reports.* **3,** 758–773. It is accessible with the DOI:10.1016/j.stemcr.2014.08.015. It is reprinted here under Creative Commons license BY-NC-ND. This work was funded by NIH-NGMS (P01GM99134) and CIRM (RB3-05207). M.P. and Y.X. were supported by a CIRM Training Grant (TG2-01169). X.G. was supported by the UCLA MSTP program and UCLA MBI Whitcome fellowship. W.E.L. was supported by the Maria Rowena Ross Chair in Cell Biology and Biochemistry, and the Eli & Edythe Broad Center of Regenerative Medicine and Stem Cell Research at UCLA Innovation Award

Chapter 3 is unpublished work performed by Xavier Gaeta, Ying Lin, Luat Le, Yuan Xie, and William E. Lowry. This work was funded by NIH-NGMS (P01GM99134). X.G. was supported by the NIH NRSA F31 GM113641-01 NIGMS grant. Y.X was supported by a CIRM Training Grant TG2-01169 and a training grant from the Broad Stem Cell Research Center at UCLA.

Chapter 4 is unpublished work performed by Xavier Gaeta, Thomas Allison, Xiaofei Wei, Luat Le, Istvan Mody, Kathrin Plath, and William E. Lowry. This work was funded by the Paul G. Allen Family Foundation. X.G. was supported by the NIH NRSA F31 GM113641-01 NIGMS grant. Technical details and troubleshooting of the interneuron differentiation protocol were provided by Jennie Close of the Allen Brain Institute. The NKX2.1 GFP line was a gracious gift from A. Elefanty. The LHX6 citrine human ES line was a gracious gift from S.A. Anderson.

Vita

2009	Bachelor of Science, Biological Sciences with Honors Specialization in Cell & Molecular Biology Stanford University, Stanford, CA, USA
2008	Leadership Alliance / SR-EIP Summer Research Fellowship
2009	Stanford University Undergraduate Advising and Research Grant
2009	Stanford University Bio-X Undergraduate Research Fellowship
2010 - Present	UCLA Medical Scientist Training Program (MSTP)
2012 – 2015	UCLA Undergraduate Research Center – Sciences Graduate Student Mentor
2013 – Present	UCLA Broad Stem Cell Research Center Stem Cell Club Coordinator
2013	Poster award, Graduate Division. UCLA Molecular, Cell, & Dev Bio retreat
2014	Outstanding Poster award. UCLA Molecular Biology Institute retreat
2014 - 2015	UCLA Whitcome Pre-doctoral Training Program in Molecular Biology
2015 - 2016	NIH Predoctoral Ruth L. Kirschstein National Research Service Award F31 GM113641-01

PUBLICATIONS:

Patterson M*, **Gaeta X***, Loo K, Edwards M, Smale S, Cinkornpumin J, Xie Y, Listgarten J, Azghadi S, Douglass SM, Pellegrini M, Lowry WE. 2014. **let-7 miRNAs can act through Notch to control human gliogenesis.** *Stem Cell Reports.* **3,** 758–773. doi:10.1016/j.stemcr.2014.08.015

Xie Y, Zhang J, Lin Y, **Gaeta X**, Meng X, Wisidagama DRR, Cinkornpumin J, Koehler CM, Malone CS, Teitell MA, Lowry WE. 2014. **Defining the Role of Oxygen Tension in Human Neural Progenitor Fate.** *Stem Cell Reports*. 2014;3(5):743-757. doi:10.1016/j.stemcr.2014.09.021.

Gaeta X, Xie Y & Lowry WE. 2013. **Sequential addition of reprogramming factors improves efficiency.** *Nature Cell Biology* **15,** 725–727.

Andres RH, Choi R, Pendharkar AV, **Gaeta X,** Wang N, Nathan JK, Chua JY, Lee SW, Palmer TD, Steinberg GK, Guzman R. 2011. **The CCR2/CCL2 Interaction Mediates the Transendothelial**

Recruitment of Intravascularly Delivered Neural Stem Cells to the Ischemic Brain. *Stroke,* 2011:42:2923-2931. doi: 10.1161/STROKEAHA.110.606368.

Chua JY, Pendharkar AV, Wang N, Choi R, Andres RH, **Gaeta X**, Zhang J, Moseley ME, Guzman R. 2011. **Intra-arterial injection of neural stem cells using a microneedle technique does not cause microembolic strokes.** *Journal of Cerebral Blood Flow & Metabolism*, 2011:31:1263-1271. doi:10.1038/jcbfm.2010.213.

Pendharkar AV, Chua JY, Andres RH, Wang N, **Gaeta X**, Wang H, De A, Choi R, Chen S, Rutt BK, Gambhir SS, Guzman R. 2010. **Biodistribution of Neural Stem Cells After Intravascular Therapy for Hypoxic-Ischemia.** *Stroke*, 2010:41: 2064-2070. doi:10.1161/STROKEAHA.109.575993.

PRESENTATIONS

Gaeta X, Choi R, Wang N, Andres RH, Pendharkar A, Guzman R. (5/2009) **Inflammatory Cytokine Pretreatment Alters CCR2 Receptor Externalization in Neural Stem Cells for Therapeutic Delivery after Stroke.** Achauer Biological Sciences Honors Symposium. Stanford University, USA. (poster)

Pendharkar AV, De A. Wang H. **Gaeta X**, Wang N, Andres RH, Chen X, Gambhir SS and Guzman R. (7/2009) **A novel multimodal approach to track neural progenitor cells in vivo.** International Symposium of Cerebral Blood Flow and Metabolism, Chicago, IL, USA. (poster)

Pendharkar AV, Chua JY, Wang N, **Gaeta X**, Choi R, Andres RH, Guzman R. (10/2009**) Intra-arterial injection results in superior delivery of neural stem cells to the ischemic brain in comparison to intravenous injection.** Society for Neuroscience, Chicago IL, USA. (poster)

Gaeta X, Patterson M, Loo K, Azghadi S, Lowry WE. (1/2013) **Tuning the developmental maturity of stem and progenitor cells via the LIN28 *let-7* circuit.** UCLA MCD Research Scholars. UCLA, Los Angeles CA, USA. (invited talk)

Gaeta X, Patterson M, Loo K, Azghadi S, Lowry WE. (6/2013) **Manipulation of the Lin28 let-7 circuit induces changes in developmental maturity of human neural progenitor cells.** New Avenues for Brain Repair: Neural Programming & Reprogramming, Boston MA, USA. (poster)

Gaeta X, Patterson M, Loo K, Edwards M, Azghadi S, Douglass SM, Pellegrini M, Smale S, Lowry WE. (3/2014) **The LIN28/let-7 circuit modulates the developmental maturity of neural progenitor cells through HMGA2 and Notch Signaling.** NIH 5th NIGMS Workshop on Pluripotent Stem Cell Research, Bethesda MD, USA. (poster)

Gaeta X, Patterson M, Loo K, Edwards M, Azghadi S, Douglass SM, Pellegrini M, Smale S, Lowry WE. (4/2014) **The LIN28/*let-7* circuit modulates the developmental maturity of neural progenitor cells through HMGA2 and Notch Signaling.** CIRM Tri-institutional Conference, Asilomar CA, USA. (poster)

Chapter 1: Introduction

Multicellular organisms require specialized cells

For all multicellular, sexually-reproducing organisms on earth, life begins as a single cell. The union of parental gametes leads to the formation of a totipotent zygote, which will make all the cells of the body and extra-embryonic tissue. In mammals, as the newly formed zygote divides to form a multicellular embryo, its individual cells remain broadly unspecialized until, after several rounds of cell division, the packed ball of cells becomes a hollow spherical blastocyst. While the outermost layer of the blastocyst goes on to form the trophoblast and extra-embryonic tissues, an inner cell mass within the blastocyst contains cells which are pluripotent: they give rise to any cells of the embryo proper. The inner cell mass progresses through gastrulation to form an epiblast stratified into the three germ layers: ectoderm, mesoderm, and endoderm. The coordinated, repeated division and specialization of these germ layers into specific cell types -- which go on to comprise tissues and organs -- represent the major task of development.

Pluripotent Stem Cells

Though the cells of the mammalian inner cell mass are only transiently pluripotent during development, advances in stem cell biology have led to cell culture techniques capable of maintaining these cells in their pluripotent state outside of the blastocyst[1]. These pluripotent stem cells (PSCs) are unique in their ability to self-renew indefinitely in culture, and to retain their ability to differentiate into all cells of the body. Because these PSCs were derived from an embryo, they are known as embryonic stem (ES) cells. In 1998 the first human ES cells were made and were maintained with a distinct set of culture conditions compared to mouse ES cells[2].

In addition to ES cells, another type of PSC exists: induced pluripotent stem cells (iPSCs). In 2006, Shinya Yamanaka's team discovered that induced expression of four transcription factors (Oct4, Sox2, Klf4, and cMyc) could revert differentiated somatic cells back to a pluripotent state[3]. This finding was rapidly confirmed in humans, and different cocktails of reprogramming factors have been shown to be sufficient to generate iPS cells[4–6].

Epigenetics: the regulation of gene expression

The Yamanaka lab's ability to restore pluripotency to somatic cells shared a conceptual finding with the work of John Gurdon decades prior. Gurdon's experiments in 1962 consisted of transferring the nucleus of a fully differentiated somatic cell from *Xenopus laevis* into an unfertilized, enucleated egg to generate a functional zygote[7]. This process, somatic cell nuclear transfer, demonstrated not only that the oocyte contains a milieu of factors that can reprogram somatic cells back to totipotency, but also that differentiated somatic cell nuclei retain all the genetic material needed to make a full organism. This realization led to the insight that if certain cellular programs drive different cell fates and behaviors, and if all genetic material is retained during this specialization, there must be a system to regulate which genetic material is active or inactive within a cell.

Epigenetic regulatory mechanisms exist at the level of transcription, splicing, processing and export, translation, trafficking, and post-translation. Some genes are noncoding, exerting their

functional role at one of the steps prior to translation. Regions of the genome that are accessible and are actively transcribed are known as euchromatin, and DNA that is tightly wrapped and inaccessible is known as heterochromatin. A constellation of post-translational modifications of residues along the tails of histone proteins allow for access and recruitment of other components of the epigenetic system, nudging genes toward expression or repression.

miRNAs: small, regulatory, non-coding RNAs

One particular instrument for epigenetic regulation is microRNAs (miRNAs), small 21-22 nt RNAs that can negatively regulate the expression of multiple target messenger RNAs (mRNAs). The multivalent nature of miRNAs – and the large number of validated and predicted miRNAs – has revealed that a substantial portion of the transcriptome is subject to regulation by miRNA-induced silencing[8]. Circulating miRNAs bind to mRNAs at regions of sequence complementarity to the 8 nt miRNA seed sequence, most often at the 3' UTR. The formation of this miRNA-target mRNA duplex prevents the targets from being translated, either by degradation of the dsRNA by the RNA-induced silencing complex (RISC) or by blocking the translation machinery so that no protein may be made from the mRNA[9].

miRNAs are transcribed as large genes (or from within introns of other genes) known as primary miRNAs (pri-miRNAs) and go through splicing and RNA processing. The microprocessor complex, made up of Drosha and Pasha (the protein product encoded by the DGCR8 gene), cleaves the

intermediate RNA molecules down to 60-80 nt RNA hairpins known as precursor miRNAs (pre-miRNAs). These hairpins are then transported to the cytoplasm via Exportin5 and further processed by DICER into mature miRNAs, the functional 21 nt RNAs that can then target cytoplasmic mRNAs for degradation[10].

The *let-7* miRNAs

*let-7*s were among the first miRNAs identified, and were named for the lethal effects observed in *let-7*-deficient *C. elegans*[11,12]. Higher organisms have undergone an expansion of the *let-7* miRNA family but their seed sequence is broadly conserved, so they can all repress the same target mRNAs[12]. In humans, the members of the *let-7* miRNA family include: *let-7a*, *let-7b*, *let-7c*, *let-7d*, *let-7e*, *let-7f*, *let-7g*, *let-7i*, and *mir-98*. Absent are *let-7h* (found in the genomes of bony fish) ; *let-7j* (found in many chordates); and *let-7k* (identified in the mouse and chicken genomes[13]).

The functional role of a miRNA family is inversely related to the functions of its targets. In the case of the *let-7* miRNAs, many of their targets are related to pluripotency[14], embryonic development[15–17], oncogenesis[18], epithelial-to-mesenchymal transition[19], and cell cycle progression[20]. As such, expression of the *let-7* miRNAs is correlated with a progression toward cell differentiation and maturation of tissue from prenatal to postnatal stages of development.

Many *let-7* targets function as oncogenes, or as genes involved in regulating progression of tumors into metastatic disease in humans. For example, HMGA2 is commandeered in cancer to promote migration, invasion, and metastasis of tumor cells[19]. Other targets, including members

of the cyclin family, can promote dysregulated cell division -- a crucial component of malignant transformation[20]. Yet others, including members of the insulin-like growth factors (IGFs) and IGF mRNA binding proteins (IMPs), as well as members of the PI3K pathway, tweak the metabolic balance of cells, which can also become aberrant in tumors[16,21]. Furthermore, increased expression of *let-7* miRNAs correlates with better prognosis in human cancers, and induced expression of *let-7*s in tumor cells negatively impacts their proliferation and survival[22,23]. For these reasons, the *let-7* miRNAs have been designated tumor suppressor miRNAs.

Very few mature *let-7* miRNAs are expressed in the zygote or during early embryonic development, but as specification and differentiation proceed, there is an increase in *let-7* abundance[14]. In the adult organism, many tissues express high levels of the *let-7s*. In recognition of this temporally modulated expression of the *let-7* miRNAs, they have been designated heterochronic – differentially expressed across the lifetime of an organism. Increasing *let-7* abundance occurs at different timescales in each tissue type, but the tight temporal control and unidirectional progression of expression changes suggest that a crucial component of *let-7* miRNA biology is the regulation of this induction. Together, the expression pattern and the direct targets of *let-7* miRNAs describe an epigenetic regulatory mechanism that promotes differentiation and maturation during normal development.

The LIN28 proteins regulate *let-7* miRNA biogenesis and form a conserved bistable switch

A major component of *let-7* miRNA biogenesis and regulation comes at the processing stage, where they are specifically regulated by a well-characterized mechanism. The LIN28 RNA binding proteins, LIN28A and LIN28B, both negatively regulate various stages of *let-7* miRNA processing, summarized in Figure 1. The functions of the LIN28 proteins are numerous and together they regulate multiple aspects of the early embryonic state during development, including self-renewal, metabolism, and mRNA translation[24–27]. In mouse, the role of *Lin28* in coordinating early embryonic gene networks is robust enough that it has been used as a reprogramming factor to convert somatic cells to the pluripotent state alongside OCT4, SOX2, and NANOG[4].

LIN28 proteins are expressed in oocytes, the zygote, and throughout early embryonic development in animals ranging from *C. elegans* to mouse to humans[28,29]. By the mammalian late embryonic period, however, LIN28 expression is markedly reduced except in the progenitor cells of some tissues. Thus, LIN28-mediated suppression of *let-7* miRNA maturation is no longer strong enough to prevent increased mature *let-7* expression[28,30]. Lin-28, the *C. elegans* ortholog of human LIN28, also follows a temporally-regulated shift in expression, leading to its initial description as a heterochronic gene[31]. Worms with aberrant lin-28 expression showed faulty timing and proliferation of the lateral seam cells, a population of epithelial cells whose division is highly regulated and specified during development[11,31,32].

When inappropriately re-expressed during adulthood, the LIN28 proteins also function as potent oncogenes[33]. Retrospective studies based off tumor tissue libraries have demonstrated that 15% of human tumors show re-expression of LIN28A or LIN28B, and this effect does not appear to be restricted to tumors from a specific tissue or cell type[34]. Furthermore, retrospective analysis of the clinical courses of those cancer patients revealed a correlation between high LIN28 expression and poor disease prognosis. Of note is that these studies make no claims about whether LIN28s were the initial drivers of transformation, but *in vitro* experiments subsequently demonstrated that LIN28 overexpression can drive malignant transformation[34].

At a molecular level, the interaction between LIN28 proteins and *let*-7 miRNAs has been thoroughly elucidated. LIN28 protein cold-shock domains initially bind to the preE loop of the *pri-let-7* and *pre-let-7* miRNA hairpins[35], followed by a more stable binding of the LIN28 CCHC-type truncated zinc finger domains to the GGAG sequences in these hairpins[36]. One exception to this relationship exists: the hairpin of *let-7a3* has been shown to evade binding by LIN28A and LIN28B because of its distinct RNA bulge[37].

A subset of *let-7s*, the more recently evolved group II *let-7s*, are monouridylated at their pre-miRNA stage as a LIN28-independent intermediate step in promoting their cleavage (Figure 2)[38]. In contrast to this monouridylation, all pre-*let*-7 miRNAs can be negatively regulated through LIN28-mediated polyuridylation[25]. The LIN28 proteins recruit terminal uridyltransferases (TUT

proteins) ZCCHC11 and ZCCHC6 which then attach many uridines to bound precursor miRNA hairpins[39–41]. Marked polyuridylated *pre-let-7* hairpins are unable to be processed by DICER in the cytoplasm and are instead degraded by the nuclease DIS3L2[25,42]. Biochemical studies show that the LIN28 proteins can substitute for one another[37,39,40] despite showing different affinities for the various TUT proteins used in *let-7* processing, and differential cellular localization (LIN28A in the nucleus).

Together, this highly regulated pathway serves to tightly control when *let-7* miRNAs, once transcribed, can reach maturity and lead to the degradation of their many targets. This regulation is mostly temporal in nature because LIN28s are expressed early in embryonic development and in progenitor cell types. Once expression of the LIN28 proteins slows enough to allow for maturation of some of the transcribed *let-7* miRNAs, a positive feedback loop begins, owing to the fact that the LIN28 proteins are themselves targeted at the mRNA level by the *let-7* miRNAs. Taken together, this relationship describes progression from a LIN28 high immature, embryonic-like state to a *let-7* high mature, differentiated state that is unlikely to reverse. The reinforcing nature of this system has led to the LIN28/*let-7* circuit being designated as a bistable switch – at either of its two states (LIN28 high and *let-7* low, or LIN28 low and *let-7* high) the relationship is stable. It remains unclear what perturbation induces the initial shift between these states during development.

Early observations suggested that *let-7* miRNAs were constitutively transcribed but that their activity was restricted until changes in the LIN28s allowed them to mature. In support of this explanation, another temporally regulated miRNA family, the *mir-125* miRNAs (*lin-4* in *C. elegans*), has been shown to target the *LIN28A* and *LIN28B* mRNAs and reduce their translation[15,32,43]. However, we have demonstrated that at the primary miRNA stage, prior to any splicing or RNA processing, there are differences in transcription of these miRNAs in different cell types. Furthermore, these differences are heterogeneous – some *let-7* miRNAs are very differentially transcribed in the settings we examined, whereas others remained relatively stably transcribed in our experimental setting.

Development of the Central Nervous System

The mammalian nervous system consists of both the central nervous system (brain and spinal cord) and the peripheral nervous system (nerve tissue extending out of the central nervous system to innervate the body with motor and sensory function). Early in development, morphogens from the notochord induce nearby ectoderm to specialize into an intermediate germ layer known as neuroectoderm. This neuroectoderm forms a fold extending in a rostral-caudal direction that generates a closed neural tube as well as neural crests. It is the rostral end of the neural tube that will then continue to expand, thicken, and fold to become the diencephalon, telencephalon, mesencephalon, and rhombencephalon; these structures comprise the embryonic forms of what will become the cerebrum, cerebellum, and brainstem.

In the developing neural tube, cells attached to the central lumen at their apical surfaces are neural stem cells, and are driven and maintained by the activity of the master regulatory transcription factor SOX2[44,45]. At this stage these cells expand through symmetrical division. Signaling molecules (SHH on the ventral side, BMP and Wnt on the dorsal side, Retinoic acid in a rostral-caudal direction) induce a regional specification in a concentration-dependent manner through a diffusion gradient and interpreted via spatially segregated receptors[46–48].

Neural cells are generated along a specific time course

The neural cells of the nervous system consist of three major cell types: neurons, astroglial cells (astrocytes), and oligodendrocytes. The central nervous system also contains cells of non-neural origin, including blood vessels, microglial cells, and at the external borders of many nervous system structures, the meninges. During development, however, these cells are not all generated simultaneously. When the first neural progenitor cells begin to differentiate, they divide and give rise to either 1 neuron by asymmetric division, or 2 neurons by symmetric division[49]. At this point, the neural progenitors have acquired characteristics of glial cells, and a subset of radial glial cells extend outward from the apical surface of the developing ventricles, providing a scaffold for their daughter cells to migrate[49–51]. This neurogenic phase persists until approximately 12 weeks of gestation, at which point neural progenitors can start differentiating into astrocytes as well in multiple neural structures[52]. Astrocytes provide trophic support for nearby neurons, form conduits separating these neurons from the blood (known as the blood-brain barrier), and play a role in regulating the local environment of synapses between neurons. Beginning at 20 weeks of

gestation, a subset of neural progenitors that has specialized to become oligodendrocyte progenitors, marked by SOX10, NKX2-2, and PDFGRα, starts differentiating into oligodendrocytes[53,54]. Oligodendrocytes are responsible for axon myelination in the central nervous system: they wrap layers of a lipid and protein compound myelin around multiple neurons to facilitate salutatory conduction of action potential to speed the transfer of motor and sensory information throughout the body. This myelination continues postnatally, and full myelination of the central nervous system is not complete until the third decade of life.

Each of these major neural cell types has multiple subtypes, distinguishable by region, morphology, neurotransmitters, connectivity, and expression of genes and other markers. Generation of these subtypes is tightly regulated, and is spatially regulated at the progenitor stage, based on the morphogen milieu present during specification[48,49]. Temporal cues, also intrinsic to the progenitor cells, can also drive the generation of distinct daughter cell populations[55,56].

Persistent differences remain between PSC-derived and *in vivo* differentiated cells

While studying PSCs can help us understand early stages of development, the broader promise of PSC technology is predicated on their capacity to generate any cell of the body in large quantities. These cells might be useful for testing new therapeutics for efficacy or off-target effects (*e.g.* treating a panel of ES-derived cardiomyocytes with a novel antiarrhythmic drug to

observe whether it can cause other life-threatening arrhythmias)[57,58], or for modeling diseases by making PSCs with known mutations and generating the cell types known to be most affected (*e.g.* generating iPS cells from a patient with α1 anti-trypsin deficiency, then differentiating those cells into lung and liver cells to study the progression of fibrosis)[59].

A goal for the stem cell biology field is the creation of cell replacement therapies based on PSC-derived cells, as they hold the promise of new therapeutic approaches for currently untreatable ailments. Cell replacement therapies could be used either for replacing specific cells lost in diseases (*e.g.* the focal loss of dopaminergic neurons of the brainstem substantia nigra in Parkinson's Disease, or the depletion of pancreatic β cells in patients with type 1 diabetes mellitus), for replacing diseased cells with healthy cells (*e.g.* introducing muscle satellite cells with functional dystrophin into the muscles of patients with a muscular dystrophy), or for repairing or replacing whole tissues (*e.g.* following traumatic injury or the failure of an entire organ). Crucial for the viability of these novel therapeutic approaches is that the cells generated through PSC differentiation faithfully represent the *bona fide* cell types they are replacing. At present, mounting evidence from a number of labs and in a number of systems suggests that this is not yet the case.

In some cell types, the difference between PSC-derived cells and their *in vivo* counterparts manifests in functional ways. Protocols to generate liver hepatocytes from PSCs, for example, can

create cells that pass through definitive endoderm and foregut differentiation and make cells that morphologically approximate polygonal hepatocytes. However, these cells produce immature versions of a number of crucial liver products: they express the carrier protein α-Fetoprotein (AFP) instead of albumin, they express cytochrome oxidase proteins (CYPs) more consistent with fetal liver (CYP3A7) than adult liver (CYP3A4)[30].

Groups attempting to make PSC-derived pancreatic β cells, or cardiomyocytes, have also encountered similar scenarios: cells can be made that express correct markers (PDX1+ endoderm that eventually express glucokinase and insulin for the pancreas, MHC variants in cardiomyocytes), and that even demonstrate functional characteristics of their target cells (such as propagating contraction and automaticity in cardiomyocytes, or glucose-dependent insulin secretion in β cells)[60–62]. Yet in each case the stem cell-derived target cells remained distinct from adult cells in gene expression and functional outputs, or required further maturation *in vivo*[63].

Stem cell-derived neural progenitors provide another example in which existing approaches to differentiation yield cells that are functionally distinct from those made during development *in vivo.* In addition to differences in gene expression, the differentiation potential of these NPCs varies based on their origin. When differentiated in the absence of epidermal growth factor (EGF) and fibroblast growth factor 2 (FGF-2, or bFGF), PSC-derived NPCs appear stuck at the neurogenic stage of nervous system development, and give rise to a low proportion of glial cells. In contrast, NPCs from neural tissue differentiate into a mix of neurons and glial cells that roughly aligns with

their developmental age, and do not appear to encounter any blocks to glial differentiation when grown and induced to differentiate *in vitro*[30].

The etiology of the disparities between PSC-derived cells and cells generated during development remain unclear, though several hypotheses exist. It remains possible that the culture conditions are still inadequate: perhaps some supplement or coating protein(s) are missing from the milieu that, if included, would permit full specification and maturation of the target cell type. Some other component of the culture system (such as the dimensionality of the culture, or another cell type important for inducing a fully differentiated fate or maintaining those cells once differentiated) might be necessary to overcome the hurdles of fully mature differentiation. However, for many cells that are difficult to generate *in vitro*, existing culture methods are able to maintain their survival and proliferation, suggesting that the major hurdle occurs during specification.

Even if the culture system is complete with all the factors necessary to generate mature differentiated cells from PSCs, the time courses utilized during most *in vitro* culture may be insufficient to drive true maturation of the specified cell types. During human gestational development, organogenesis and initial specification of most cell types occurs over the course of months, and across multiple body systems this maturation at the cellular and organ levels continues, alongside functional changes, past fetal development until the postnatal period[64–66],

early childhood[67], puberty[68,69], or adulthood[70]. Maturation of these cell types *in vitro* might then be occurring along the same timescale as they would during development, implying the existence of a cell-intrinsic clock governing maturation[71–73]. This explanation accounts for the finding that PSC-derived differentiated cells often functionally and transcriptionally resemble cells from the early fetal period of human development[30,74]. Furthermore, the intrinsic clock hypothesis is positioned to explain why *in vitro* differentiation and maturation have been more successful using PSCs generated from animals with shorter gestational periods. In directly comparable PSC-derived cell differentiation protocols in human and mouse, the mouse protocols progress along a shorter timescale, and are more efficient and effective at generating mature cell types[75–77].

This hypothesis was directly tested by comparing PSCs differentiated along ectoderm, mesoderm, and endodermal developmental pathways to the same cell types naturally generated in the developing body[30]. A set of 88 genes was identified as consistently differentially expressed between these two populations, regardless of the germ layer being generated (Table 1). These 88 genes might comprise a conserved genetic mechanism separating the more mature, tissue-derived cells from the less mature PSC-derived cells. Member of the LIN28/*let-7* circuit and their targets were identified within this set of 88 genes, and their expression changes followed a pattern consistent with *let-7* expression promoting maturation. Together, these results raised the intriguing possibility that the LIN28/*let-7* pathway, and the change between the stable states of that pathway, might integrate with the intrinsic developmental clock to compose a major role in the timing of maturation broadly in development.

In the following three chapters, I present work conducted with the aim of addressing this hypothesis in the context of neural development and differentiation.

In Chapter 2 I describe work that directly investigates the functional and mechanistic roles for LIN28 and *let-7* miRNAs in stem cell-derived neural progenitor cells, and directly manipulate the LIN28/*let-7* circuit to induce functional changes in PSC-derived and Tissue-derived NPC differentiation, namely a shift between neurogenesis, astrogliogenesis, and oligodendrogenesis. I then characterize the specific *let-7* targets responsible for these effects, and elucidate a mechanism for the transduction of *let-7* repression into neural-specific signaling pathways and gene expression.

In Chapter 3 I examine the mechanism behind the switch between the bistable states of the LIN28/*let-7* circuit, and determine that transcriptional regulation of several *let-7* loci, independent of later processing, occurs over the course of development. I then take a candidate approach to identifying drivers of this transcriptional shift, particularly at newly annotated promoters and enhancers near important *let-7* loci.

In Chapter 4 I explore maturation of a post-differentiation population of cells, in this case PSC-derived inhibitory interneurons. I describe efforts to generate these cells *in vitro* and to characterize their functional maturation by gene expression and electrophysiology. I then describe efforts to artificially improve the maturation of these cells, using co-culture and injection into a mammalian host, to provide trophic support and a signaling milieu capable of promoting interneuron maturation. Using insights gathered from understanding *let-7* miRNA mediated maturation, we plan to use these model systems to characterize the cell-intrinsic mechanisms of interneuron maturation.

Figures

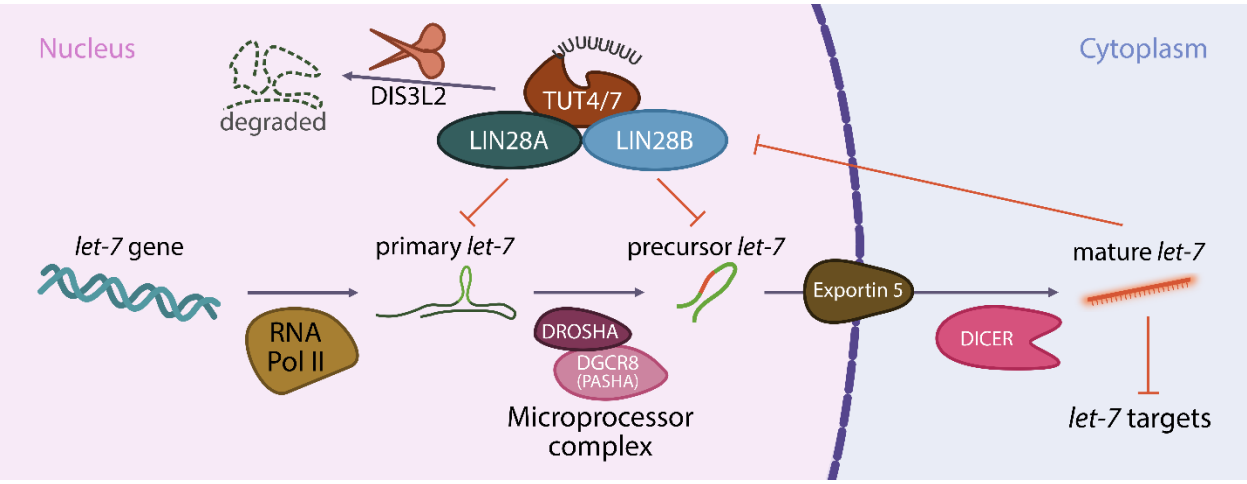


Figure 1-1: let-7 miRNA biogenesis, processing, and LIN28-mediated regulation

Schematic describing the progression from *let-7* gene loci to transcription, processing, and maturation of the *let-7* miRNA. While early stages of *let-7* biogenesis occur in the nucleus, the *pre-let-7* hairpins are transported to the cytoplasm via Exportin 5, and can also be targeted for degradation by LIN28 proteins and their associated TUT proteins in the cytoplasm. It is only in the cytoplasm that the mature *let-7* miRNA are formed and can bind to a variety of mRNA targets, marking them for degradation by the RISC. The mRNAs of the LIN28 proteins are themselves targets of the *let-7* miRNAs, leading to a feedback loop by which *let-7* miRNAs can indirectly promote their own biogenesis.

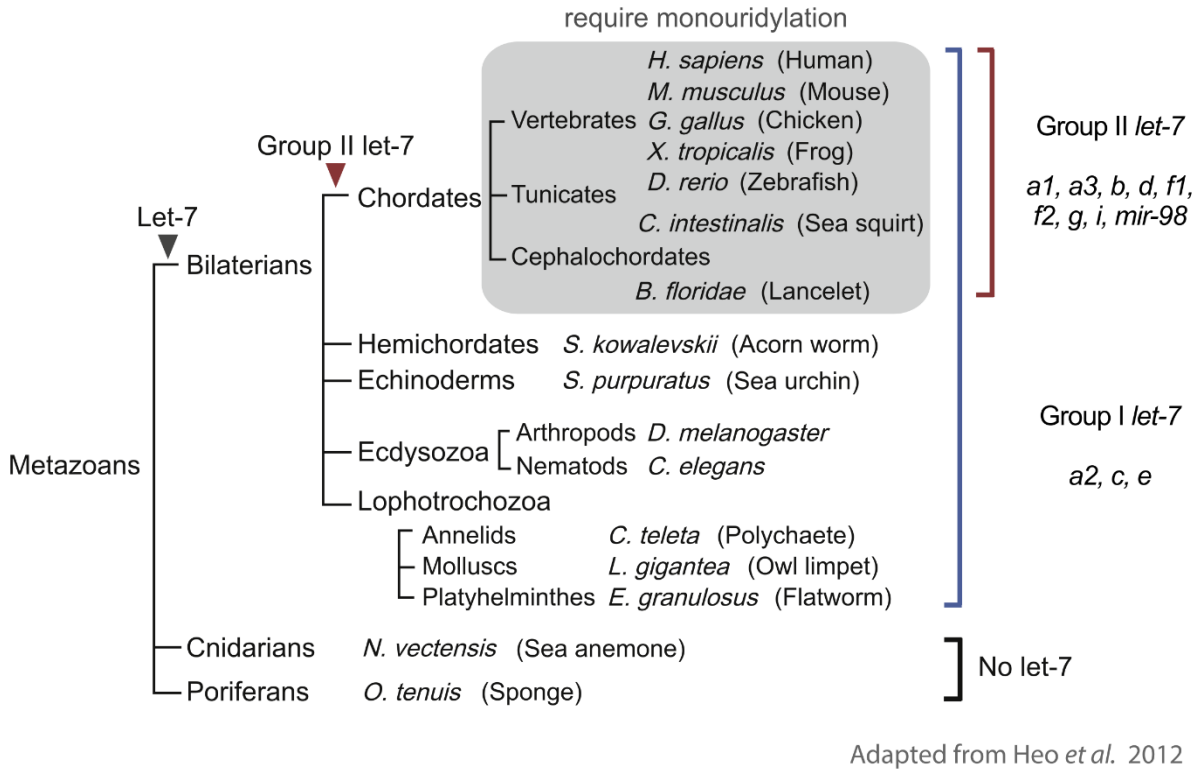


Figure 1-2: Phylogenetic conservation and expansion of the let-7 miRNAs

The *let-7* miRNA family arose as a single member in the precursors of the bilaterian clade of animals, and has expanded over the course of phylogenic evolution. While bilaterian genomes contain *let-7* miRNAs orthologous to the human *let-7a2*, *let-7c*, and *let-7e*, in chordates a second group of *let-7*s has formed and expanded. This *let-7* group II contains human *let-7a1*, *let-7a3*, *let-7b*, *let-7d*, *let-7f1*, *let-7f2*, *let-7g*, *let-7i*, and *mir-98*. Of the group II *let-7*s, only *let-7g* and *let-7i* are not known to be polycistronic, though the human *let-7g* locus is intronic to the gene *WDR82*. The group II *let-7s* are molecularly distinguishable from their group I counterparts by a shorter 3' overhang at the *pre-let-7* hairpin stage, and require LIN28-independent monouridylation in order to be properly processed by DICER.

Higher in Tissue-derived	
ZFP3	PPP1R3C
PRR34-AS1	NFIX
PSMB9	VAMP5
CDKN2C	ACSL1
MGLL	CFLAR
CNTN3	ANKRD35
CASP1	RP11-1216L17.1
ZNF280D	ARNTL
SLC44A1	AC009237.8
SMARCA2	CBR1
RP11-136C24.3	LINC00899
PAX8-AS1	FBXO4
ZNF662	MIRLET7BHG
MAP3K5	RP11-5C23.1
SPESP1	RP11-568K15.1
ZEB1	CREBL2

Higher in PSC-derived		
LIN28B	F11R	ANKRD6
CCDC144NL-AS1	DGKZ	LLGL2
DPPA4	MDK	JUP
CLDN6	ZFHX3	TGIF2
STK26	SPINT1	ANK3
NAALAD2	TMEM88	KLHDC8B
IGF2BP1	KDM6B	ZCCHC3
DSC2	PPM1H	CORO2A
RGS5	ST6GALNAC3	TCF7L1
GPC3	FAR2P1	WDR33
PCDHB2	CYP4F62P	EPHB4
EPPK1	ISYNA1	TMC6
SPINT2	CDH3	PCDHB3
MAB21L2	FZD2	PLEKHG3
HAND1	MYRF	TDG
ALX1	RP3-523K23.2	TLE3
GATA3	PLAGL2	
HIC2	KRBOX1	

Table 1-1: Consistently differentially expressed genes between PSC-derived progeny and their tissue-derived counterparts

Gene expression of PSC-derived endoderm, mesoderm, and ectoderm progeny were compared against their *in vivo* differentiated counterparts by Affymetrix microarray. For cells representative of each germ layer, a list of significantly changed genes (fold change > 1.54, $p < 0.01$ after Benjamini Hochberg FDR correction) was generated, and the intersection of the lists from endoderm, mesoderm, and ectoderm yielded 88 consistently differentially expressed genes. At left are genes consistently more highly expressed in tissue-derived cells, listed in order of absolute fold change. At right are genes consistently more highly expressed in PSC-derived cells, listed in order of absolute fold change. Dark grey boxes denote members of the LIN28/*let-7* circuit and their direct targets.

References

1. Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292,** 154–156 (1981).

2. Thomson, J. A. *et al.* Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* **282,** 1145–1147 (1998).

3. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126,** 663–676 (2006).

4. Yu, J. *et al.* Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science* **318,** 1917–1920 (2007).

5. Takahashi, K. *et al.* Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131,** 861–872 (2007).

6. Lowry, W. E. *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl. Acad. Sci.* **105,** 2883–2888 (2008).

7. Gurdon, J. B. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *Development* **10,** 622–640 (1962).

8. Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell* **120,** 15–20 (2005).

9. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *cell* **116,** 281–297 (2004).

10. Lee, H., Han, S., Kwon, C. S. & Lee, D. Biogenesis and regulation of the let-7 miRNAs and their functional implications. *Protein Cell* (2015). doi:10.1007/s13238-015-0212-y

11. Reinhart, B. *et al.* The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **403,** 901–906 (2000).

12. Pasquinelli, A. E. *et al.* Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408,** 86–89 (2000).

13. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **39,** D152–D157 (2011).

14. Melton, C., Judson, R. L. & Blelloch, R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* **463,** 621–626 (2010).

15. Banerjee, D. & Slack, F. Control of developmental timing by small temporal RNAs: a paradigm for RNA-mediated regulation of gene expression. *BioEssays* **24,** 119–129 (2002).

16. Toledano, H., D'Alterio, C., Czech, B., Levine, E. & Jones, D. L. The let-7–Imp axis regulates ageing of the Drosophila testis stem-cell niche. *Nature* **485,** 605–610 (2012).

17. Gurtan, A. M. *et al.* Let-7 represses Nr6a1 and a mid-gestation developmental program in adult fibroblasts. *Genes Dev.* **27,** 941–954 (2013).

18. Johnson, S. M. *et al.* RAS Is Regulated by the let-7 MicroRNA Family. *Cell* **120,** 635–647 (2005).

19. Watanabe, S. *et al.* HMGA2 maintains oncogenic RAS-induced epithelial-mesenchymal transition in human pancreatic cancer cells. *Am. J. Pathol.* **174,** 854 (2009).

20. Johnson, C. D. *et al.* The let-7 microRNA represses cell proliferation pathways in human cells. *Cancer Res.* **67,** 7713–7722 (2007).

21. Kuppusamy, K. T. *et al.* Let-7 family of microRNA is required for maturation and adult-like metabolism in stem cell-derived cardiomyocytes. *Proc. Natl. Acad. Sci.* **112,** E2785–E2794 (2015).

22. Ma, C. *et al.* H19 promotes pancreatic cancer metastasis by derepressing let-7's suppression on its target HMGA2-mediated EMT. *Tumor Biol.* **35,** 9163–9169 (2014).

23. Takamizawa, J. *et al.* Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res.* **64,** 3753–3756 (2004).

24. Shyh-Chang, N. *et al.* Lin28 Enhances Tissue Repair by Reprogramming Cellular Metabolism. *Cell* **155,** 778–792 (2013).

25. Heo, I. *et al.* Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Mol. Cell* **32,** 276–284 (2008).

26. Polesskaya, A. *et al.* Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency. *Genes Dev.* **21,** 1125–1138 (2007).

27. Shyh-Chang, N. & Daley, G. Q. Lin28: Primal Regulator of Growth and Metabolism in Stem Cells. *Cell Stem Cell* **12,** 395–406 (2013).

28. Yang, D.-H. & Moss, E. G. Temporally regulated expression of Lin-28 in diverse tissues of the developing mouse. *Gene Expr. Patterns* **3,** 719–726 (2003).

29. Moss, E. G. & Tang, L. Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. *Dev. Biol.* **258,** 432–442 (2003).

30. Patterson, M. *et al.* Defining the nature of human pluripotent stem cell progeny. *Cell Res.* **22,** 178–193 (2011).

31. Ambros, V. & Horvitz, H. Heterochronic mutants of the nematode Caenorhabditis elegans. *Science* **226,** 409–416 (1984).

32. Johnson, S. M., Lin, S.-Y. & Slack, F. J. The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev. Biol.* **259,** 364–379 (2003).

33. Viswanathan, S. R. & Daley, G. Q. Lin28: A MicroRNA Regulator with a Macro Role. *Cell* **140,** 445–449 (2010).

34. Viswanathan, S. R. *et al.* Lin28 promotes transformation and is associated with advanced human malignancies. *Nat. Genet.* **41,** 843–848 (2009).

35. Mayr, F., Schutz, A., Doge, N. & Heinemann, U. The Lin28 cold-shock domain remodels pre-let-7 microRNA. *Nucleic Acids Res.* **40,** 7492–7506 (2012).

36. Loughlin, F. E. *et al.* Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. *Nat. Struct. Mol. Biol.* **19,** 84–89 (2011).

37. Triboulet, R., Pirouz, M. & Gregory, R. I. A Single Let-7 MicroRNA Bypasses LIN28-Mediated Repression. *Cell Rep.* (2015). doi:10.1016/j.celrep.2015.08.086

38. Heo, I. *et al.* Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. *Cell* (2012).

39. Piskounova, E. *et al.* Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms. *Cell* **147,** 1066–1079 (2011).

40. Thornton, J. E., Chang, H.-M., Piskounova, E. & Gregory, R. I. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA* **18,** 1875–1885 (2012).

41. Kim, B. *et al.* TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. *EMBO J.* (2015). doi:10.15252/embj.201590931

42. Chang, H.-M., Triboulet, R., Thornton, J. E. & Gregory, R. I. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28–let-7 pathway. *Nature* (2013). doi:10.1038/nature12119

43. Rybak, A. *et al.* A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat. Cell Biol.* **10,** 987–993 (2008).

44. Graham, V., Khudyakov, J., Ellis, P. & Pevny, L. SOX2 functions to maintain neural progenitor identity. *Neuron* **39,** 749–765 (2003).

45. Zappone, M. V. *et al.* Sox2 regulatory sequences direct expression of a (beta)-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells. *Development* **127,** 2367–2382 (2000).

46. Sansom, S. N. & Livesey, F. J. Gradients in the Brain: The Control of the Development of Form and Function in the Cerebral Cortex. *Cold Spring Harb. Perspect. Biol.* **1,** a002519–a002519 (2009).

47. Dupé, V. & Lumsden, A. Hindbrain patterning involves graded responses to retinoic acid signalling. *Development* **128,** 2199–2208 (2001).

48. Tiberi, L., Vanderhaeghen, P. & van den Ameele, J. Cortical neurogenesis and morphogens: diversity of cues, sources and functions. *Curr. Opin. Cell Biol.* **24,** 269–276 (2012).

49. Temple, S. The development of neural stem cells. *Nat.-Lond.*- 112–117 (2001).

50. Fiona Doetsch. The glial identity of neural stem cells. *Nat. Neurosci.* **6,** 1127–1134 (2003).

51. Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. & Kriegstein, A. R. Neurons derived from radial glial cells establish radial units in neocortex. *Nature* **409,** 714–720 (2001).

52. Miller, F. D. & Gauthier, A. S. Timing Is Everything: Making Neurons versus Glia in the Developing Cortex. *Neuron* **54,** 357–369 (2007).

53. Jakovcevski, I., Filipovic, R., Mo, Z., Rakic, S. & Zecevic, N. Oligodendrocyte development and the onset of myelination in the human fetal brain. *Front. Neuroanat.* **3,** (2009).

54. Hsieh, J. IGF-I instructs multipotent adult neural progenitor cells to become oligodendrocytes. *J. Cell Biol.* **164,** 111–122 (2004).

55. Shen, Q. *et al.* The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nat. Neurosci.* **9,** 743–751 (2006).

56. Bayraktar, O. A. & Doe, C. Q. Combinatorial temporal patterning in progenitors expands neural diversity. *Nature* **498,** 449–455 (2013).

57. Braam, S. R. *et al.* Prediction of drug-induced cardiotoxicity using human embryonic stem cell-derived cardiomyocytes. *Stem Cell Res.* **4,** 107–116 (2010).

58. Navarrete, E. G. *et al.* Screening Drug-Induced Arrhythmia Using Human Induced Pluripotent Stem Cell–Derived Cardiomyocytes and Low-Impedance Microelectrode Arrays. *Circulation* **128,** S3–S13 (2013).

59. Wilson, A. A. *et al.* Emergence of a Stage-Dependent Human Liver Disease Signature with Directed Differentiation of Alpha-1 Antitrypsin-Deficient iPS Cells. *Stem Cell Rep.* **4,** 873–885 (2015).

60. Xu, C. *et al.* Efficient generation and cryopreservation of cardiomyocytes derived from human embryonic stem cells. *Regen. Med.* **6,** 53–66 (2011).

61. Chong, J. J. H. *et al.* Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. *Nature* **510,** 273–277 (2014).

62. Pagliuca, F. W. *et al.* Generation of Functional Human Pancreatic β Cells In Vitro. *Cell* **159,** 428–439 (2014).

63. Yang, X., Pabon, L. & Murry, C. E. Engineering Adolescence: Maturation of Human Pluripotent Stem Cell-Derived Cardiomyocytes. *Circ. Res.* **114,** 511–523 (2014).

64. Peters, N. S. *et al.* Spatiotemporal relation between gap junctions and fascia adherens junctions during postnatal development of human ventricular myocardium. *Circulation* **90,** 713–725 (1994).

65. Long, F. & Ornitz, D. M. Development of the Endochondral Skeleton. *Cold Spring Harb. Perspect. Biol.* **5,** a008334–a008334 (2013).

66. Burri, P. H. Structural Aspects of Postnatal Lung Development – Alveolar Formation and Growth. *Biol. Neonate* **89,** 313–322 (2006).

67. Simpson, S. W. & Kunos, C. A. A radiographic study of the development of the human mandibular dentition. *J. Hum. Evol.* **35,** 479–505 (1998).

68. Casey, B. J., Getz, S. & Galvan, A. The adolescent brain. *Dev. Rev.* **28,** 62–77 (2008).

69. Inman, J. L., Robertson, C., Mott, J. D. & Bissell, M. J. Mammary gland development: cell fate specification, stem cells and the microenvironment. *Development* **142,** 1028–1042 (2015).

70. Gogtay, N. *et al.* Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 8174–8179 (2004).

71. Hu, B.-Y. *et al.* Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc. Natl. Acad. Sci.* **107,** 4335–4340 (2010).

72. Gaspard, N. *et al.* An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature* **455,** 351–357 (2008).

73. Cornacchia, D. & Studer, L. Back and forth in time: Directing age in iPSC-derived lineages. *Brain Res.* (2015). doi:10.1016/j.brainres.2015.11.013

74. Fang, H. *et al.* Transcriptome Analysis of Early Organogenesis in Human Embryos. *Dev. Cell* **19,** 174–184 (2010).

75. Maroof, A. M. *et al.* Directed Differentiation and Functional Maturation of Cortical Interneurons from Human Embryonic Stem Cells. *Cell Stem Cell* **12,** 559–572 (2013).

76. Nicholas, C. R. *et al.* Functional Maturation of hPSC-Derived Forebrain Interneurons Requires an Extended Timeline and Mimics Human Neural Development. *Cell Stem Cell* **12,** 573–586 (2013).

77. Shi, Y., Kirwan, P., Smith, J., Robinson, H. P. C. & Livesey, F. J. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nat. Neurosci.* **15,** 477–486 (2012).

Chapter 2: *let-7* miRNAs can act through Notch to regulate human gliogenesis

Stem Cell Reports
Article
OPEN ACCESS

let-7 miRNAs Can Act through Notch to Regulate Human Gliogenesis

M. Patterson,[1,2,5] X. Gaeta,[1,2,3,5] K. Loo,[2] M. Edwards,[1] S. Smale,[1] J. Cinkornpumin,[1,2] Y. Xie,[1,2] J. Listgarten,[4] S. Azghadi,[2] S.M. Douglass,[1,2] M. Pellegrini,[1,2] and W.E. Lowry[1,2,3,*]

[1]Eli and Edythe Broad Center for Regenerative Medicine, UCLA, Box 957357, Los Angeles, CA 90095, USA
[2]Department of Molecular, Cell and Developmental Biology, UCLA, 621 Charles E. Young Drive East, Los Angeles, CA 90095, USA
[3]Molecular Biology Institute, UCLA, 611 Charles E. Young Drive East, Los Angeles, CA 90095, USA
[4]Microsoft Research, 1100 Glendon Avenue Suite PH1, Los Angeles, CA 90024, USA
[5]Co-first author
*Correspondence: blowry@ucla.edu
http://dx.doi.org/10.1016/j.stemcr.2014.08.015

SUMMARY

It is clear that neural differentiation from human pluripotent stem cells generates cells that are developmentally immature. Here, we show that the *let-7* plays a functional role in the developmental decision making of human neural progenitors, controlling whether these cells make neurons or glia. Through gain- and loss-of-function studies on both tissue and pluripotent derived cells, our data show that *let-7* specifically regulates decision making in this context by regulation of a key chromatin-associated protein, HMGA2. Furthermore, we provide evidence that the *let-7*/*HMGA2* circuit acts on *HES5*, a NOTCH effector and well-established node that regulates fate decisions in the nervous system. These data link the *let-7* circuit to NOTCH signaling and suggest that this interaction serves to regulate human developmental progression.

INTRODUCTION

We previously determined that, by both gene expression and functional analyses, the derivatives of human pluripotent stem cells (hPSCs) more closely resembled cell types found prior to 6 weeks of gestation than later time points (Patterson et al., 2012). In fact, this appears to be an emerging theme in hPSC differentiation (Chang et al., 2011; Mariani et al., 2012; Zambidis et al., 2005). This suggests that hPSC derivatives are developmentally immature, which could stem from either inadequate culturing methods or could suggest that developmental timing is somewhat conserved in vitro.

Among the most differentially expressed genes in all PSC derivatives are *LIN28A* and *LIN28B*, RNA binding proteins known to regulate the *let-7* family of miRNAs (Patterson et al., 2012). *LIN28B* seems to function primarily in the nucleus by sequestering *pri-let-7*s to prevent maturation by Microprocessor, whereas *LIN28A* functions in the cytoplasm by recruiting uridylyl transferase to polyuridylate the pre-*let-7*s and prevent their further processing by Dicer (Graf et al., 2013; Hagan et al., 2009; Kim and Nam, 2006; Lee et al., 2014; Piskounova et al., 2011). In lower organisms, *Lin28A* expression is strongly correlated with the differentiation status and self-renewing capacity of cells throughout development. Although there is less evidence for the role of this pathway in human development specifically, many groups have demonstrated that *LIN28A* is reexpressed in a variety of human cancers and is highly correlated with prognosis and disease progression (Viswanathan and Daley, 2010; West et al., 2009). Furthermore, *LIN28A* has also been used to reprogram somatic cells back to the pluripotent state (Yu et al., 2007). All of these known roles are linked to developmental progression and make *LIN28A/B-let-7* an attractive candidate for manipulating the maturity of hPSC-derived cells.

Previous work by other groups in lower organisms has argued that *Lin28A* plays a role in maturation of the nervous system (Balzer et al., 2010), and some have shown that overexpression of *LIN28B* in human adult hematopoietic stem/progenitor cells can reverse their developmental progression to a fetal-like state (Yuan et al., 2012). Downstream of LIN28A/B, however, a role for *let-7* in human gestational maturation in the nervous system has not been established. In fact, one study of a murine model has suggested that the role of *Lin28A* in developmental progression was *let-7* independent (Balzer et al., 2010). Recent work has also suggested that LIN28/*let-7* regulates neurogenesis by controlling the proliferation of progenitors (Cimadamore et al., 2013; Nishino et al., 2013). Here, we explore the role of the *LIN28*/*let-7* pathway in the developmental progression of human neural progenitor cells (NPCs). We demonstrate that *LIN28B* plays a clear role in gestational progression of the developing human nervous system through regulation of *let-7* miRNAs. These miRNAs then go on to regulate *HMGA2*, which has also been implicated in developmental progression (Sanosaka et al., 2008). We also show here that *HMGA2* appears to regulate cell-fate decisions in neural progenitors (NPCs) in this context

through *HES5*, a key node in the NOTCH pathway and previously implicated in neurogenesis.

RESULTS

PSC-NPCs Are Functionally and Transcriptionally Distinct from Tissue-Derived Counterparts

During mammalian development, NPCs progress through several phases: an early expansion phase, characterized by symmetric divisions; a neurogenic phase, characterized by asymmetric divisions and resulting in new neurons; a gliogenic phase where astrocytes are primarily produced; and finally a phase where oligodendrocytes are generated (Figure 1). The result is that neurogenesis precedes gliogenesis on a developmental timescale.

NPCs were derived from either hPSCs or from fetal tissue sources and were validated by immunostaining and judged to be relatively homogenous (Figure S1A; Patterson et al., 2012). We determined that PSC-NPCs across all passages had a higher propensity to differentiate into MAP2/TUJ1$^+$ neurons (~50%) over GFAP/S100/A2B5$^+$ glia (<10%) (Figure 1A). Meanwhile, tissue-derived NPCs isolated from fetal brain or spinal cord samples at 12–19 weeks of gestation (Tissue-NPCs) were more apt to differentiate into glia (~70%) over neurons (<20%) (Figure 1A; Patterson et al., 2012). These data suggest that PSC-NPCs were functionally less mature than tissue-derived counterparts (neurogenesis precedes gliogenesis). Furthermore, we have previously shown that upon subsequent passage of PSC-NPCs the propensity for gliogenesis increased but still did not approach that of tissue-derived cells (Patterson et al., 2012).

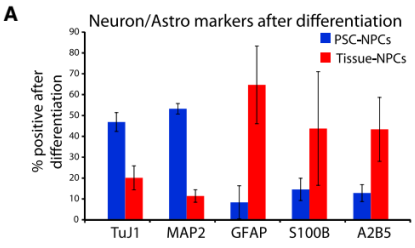
To understand the molecular basis for this observed functional discrepancy, gene expression profiling was performed on PSC derivatives and tissue-derived counterparts (Patterson et al., 2012). Among the most differentially expressed genes were *LIN28A* and *LIN28B* (Figures 1B, top, and S1C), and this was confirmed at the protein level by immunostaining (Figure 1B, bottom). Although continued passaging reduces the levels of *LIN28A* and *LIN28B* in PSC-NPCs, their expression is not decreased to a level found in the Tissue-NPCs within the time points utilized for this study (Figure 1B; Patterson et al., 2012).

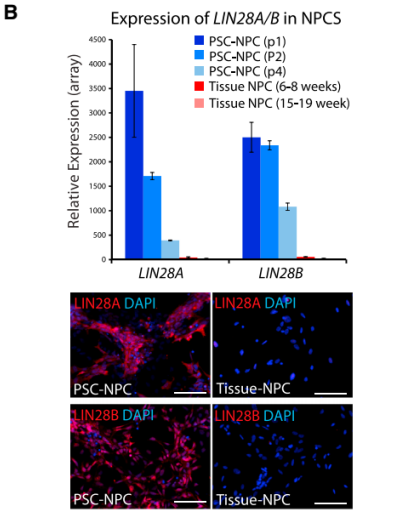
LIN28 homologs are known to negatively regulate the highly conserved *let-7* family of miRNAs. Our previous microarray analyses demonstrated a significantly higher expression of some *let-7* miRNAs in Tissue-NPCs (Figure S1C), and this result was confirmed by direct sequencing of mature miRNA (Figure 1C; Table S1). The latter analysis demonstrated that not only were all *let-7* family members significantly higher in Tissue-NPCs, as we had previously shown by RT-PCR (Patterson et al., 2012), but all nine family members were found among the top 30 differentially expressed miRNAs between Tissue-NPCs and PSC-NPCs (Table S1). Furthermore, the *let-7* family as a whole was the most abundantly expressed miRNA family in 16 week Tissue-NPCs, representing almost 18% of the total miRNA in these cells. In addition, *let-7* family members were expressed at an intermediate level in 6–7 week Tissue-NPCs.

To determine whether *let-7* target genes were among the differentially expressed mRNA distinguishing PSC-NPCs from their tissue-derived counterparts, two lists of *let-7* targets were generated (Figure 1D): one with published *let-7* targets that have been experimentally confirmed (77 genes) and one with predicted *let-7* targets generated by TargetScan 5.2 (751 genes). Of the 77 published *let-7* targets, 20 were differentially expressed between PSC-NPCs and Tissue-NPCs, and 134 of the 751 TargetScan predicted *let-7* targets were differentially expressed. Notably, 95% and 63%, respectively, were specifically lower in Tissue-NPCs, consistent with a negative regulatory activity of *let-7* in tissue-NPCs.

Recent work in mice has suggested that during neurogenesis, the induction of *let-7* leads to exit of the cell cycle and differentiation toward neurons (Cimadamore et al., 2013; Nishino et al., 2013). Analysis of eight tissue-derived NPCs and three PSC-derived NPCs failed to find consistent differences in the percentage of cells in S or G2/M phase of the cell cycle (Figure 1E), suggesting that despite having orders of magnitude more *let-7* (Figure 1C), Tissue-NPCs can still maintain a proliferative state.

To determine when the transition from neurogenesis to gliogenesis occurs endogenously, we probed developmental transcriptome data for glial hallmark genes from the Allen Brain Atlas. GFAP, AQP4, and S100B were all expressed at a low level in all brain regions prior to 12 weeks but surged thereafter suggesting that gliogenesis begins in the human brain at roughly 12 weeks of gestation. A similar examination of oligodendrocyte markers suggests that oligodendrogenesis begins at 24 weeks (Figure 1F, middle). These gene expression data are consistent with pathological data on fetal tissue showing oligodendrocyte progenitors prevalent at 20 weeks of gestation and mature oligodendrocytes at 30 weeks (Craig et al., 2003; Dean et al., 2011; Jakovcevski et al., 2009). Notably, the expression of *LIN28B* in human brain was only detectable in weeks 8–9 of gestation and dropped off significantly thereafter (Figure 1F, middle), whereas *LIN28A* was undetectable in these analyses. *let-7* miRNAs were also detectable in these data sets, and this family showed a striking induction across the human brain in the 9–12 postconception week (PCW) time frame (Figure S1B). Taken together, these data suggested that there is a strong correlation between the transition from high*LIN28*/low*let-7* to low*LIN28*/high*let7* state



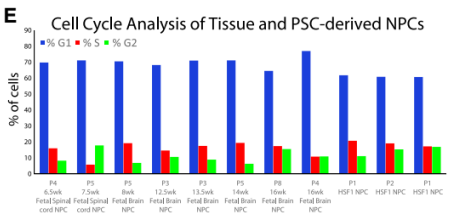


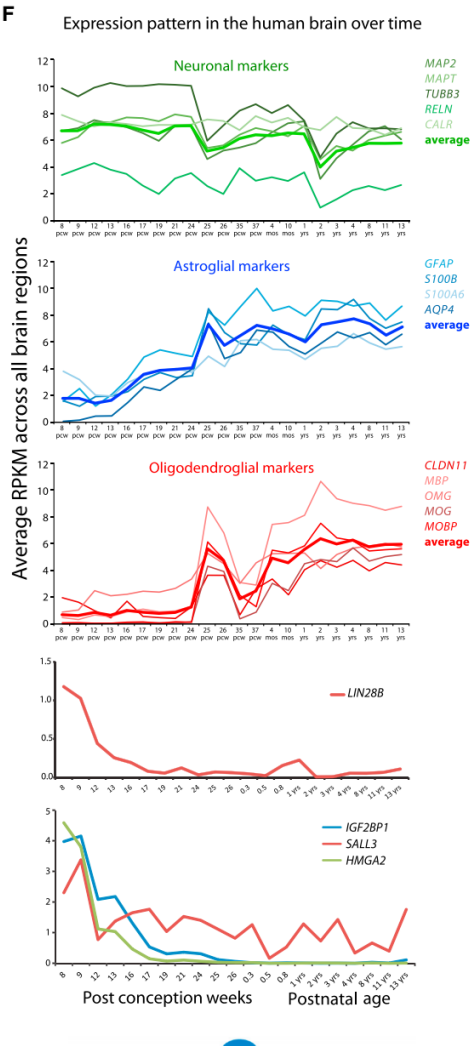
C

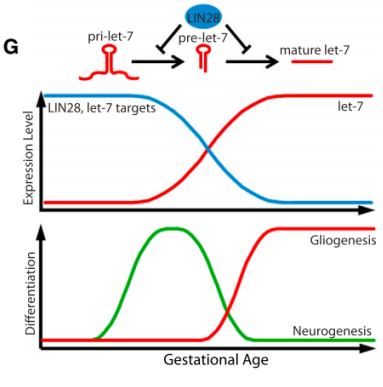
mature miRNA	16wk TissueNPC (average)	6-7wk TissueNPC (average)	PSC-NPC (average)	Fold Change (16wk Tissue v PSC-NPC)
let-7b	4.11%	1.00%	0.01%	460.76
let-7a	6.00%	1.89%	0.11%	52.55
let-7d	0.34%	0.33%	0.01%	43.53
let-7i	1.57%	0.80%	0.04%	43.51
miR-98	0.06%	0.03%	0.00%	24.65
let-7g	0.37%	0.12%	0.02%	23.60
let-7c	3.92%	3.98%	0.20%	19.24
let-7e	1.25%	1.35%	0.10%	12.40
let-7f	0.40%	0.10%	0.03%	12.33
Sum of all let-7 family	18.03%	9.61%	0.52%	34.67

D

	Total		let-7 published (77 genes)		TargetScan 5.2 let-7 (751 genes)	
	genes	% up in PSC-NPC	overlap	% up in PSC-NPC	overlap	% up in PSC-NPC
PSC-NPC vs Tissue-NPC	1885	44.4%	20	95.0%	134	62.7%







(legend on next page)

and the switch from neuro- to gliogenesis. Therefore, we hypothesized that the *let-7* plays a functional role in the decision to make either neurons or glia by neural progenitors (Figure 1G).

let-7 Expression and Processing across Development

LIN28A/B controls the maturation step to generate mature *let-7* (Hagan et al., 2009; Piskounova et al., 2008, 2011; Viswanathan et al., 2008), and *LIN28A* and *B* are also *let-7* targets. A priori, we expected that levels of *let-7* rose during development as a result of *LIN28A/B* downregulation over time. To understand the dynamics of *let-7* processing during development, we assessed the levels of pri, pre, and mature *let-7* miRNAs (Figure 2), we used RT-PCR to detect specifically mature *let-7* family members in NPCs and found that, similar to what was found by miRNA sequencing (miRNA-seq) profiling (Figure 1D), all *let-7* family members were dramatically higher in Tissue-NPCs compared to PSC-NPCs (Figure 2A; Patterson et al., 2012). Conversely, by using a miRNA RT-PCR approach that strictly measures primary *let-7* transcripts, we determined that *pri-let-7b* and *pri-let-7a2* transcript levels are significantly higher in Tissue-NPCs, whereas all the other tested members of the *let-7* family were essentially unchanged (Figure 2C). Similarly, a separate approach that can distinguish pri- and pre-*let-7* messages from mature *let-7* miRNAs also showed that pre-*let-7b* was the only *let-7* family member assayed that was differentially expressed in Tissue- versus PSC-NPCs (Figure 2B). Interestingly, *let-7b* and *let-7a3* are known to be expressed from the same locus (Wang et al., 2012), whereas *let-7a1* and *let-7a2* are transcribed from different loci.

Passaging of PSC-NPCs results in a decrease of both *LIN28A* and *LIN28B* (Figure 1B; Patterson et al., 2012) and increase of mature *let-7*s (Figure 2E; Patterson et al., 2012), suggesting a functional link between *LIN28* expression and levels of mature *let-7* as expected (Hagan et al., 2009; Heo et al., 2008). Interestingly, when assaying for pri-miRNA transcripts, many of the pri-*let-7*s were also induced during passaging as detected by both RT-PCR (Figure 2F) and Chromatin-RNA-seq (discussed more below; Figure 2G), which presumably occurred independently of any direct regulation by LIN28 protein, because they are not thought to play any role in regulation of transcription of pri-*let-7*s. These results point toward the existence of undescribed regulatory mechanisms governing *let-7* transcription independent of *LIN28* during human gestation.

To further verify the apparent transcriptional regulation of the *let-7* family, we employed an approach whereby chromatin-associated RNA is captured and sequenced (Bhatt et al., 2012). This approach generates data on relative amounts of nascent transcript, and therefore a bona fide measure of transcription as opposed to steady-state levels of RNA. Consistent with results from RT-PCR, transcription at the *let-7a3/let-7b* locus was significantly different between PSC-NPCs and Tissue-NPCs, demonstrating a *LIN28A/B* independent regulation of *let-7* (Figures 2D and 2G). These data also suggest that the *let-7* family of miRNAs is each subject to unique modes of transcriptional regulation that occurs prior to the actions of *LIN28A/B* on *let-7* maturation. Also shown is the expression of several unrelated miRNAs (miR-10, 15A, 15B) to demonstrate that the *let-7* effect is specific to this family and not indicative of imbalanced analysis (Figure 2G). This method allowed for a complete annotation of human *let-7* transcriptional loci based on analysis of transcribed RNA, as

Figure 1. *let-7* Activity Correlates with Human Gliogenesis

(A) Differentiation of human NPCs by growth factor withdrawal drives the generation of neurons and astrocytes. Percentage of positive PSC-NPCs and Tissue-NPCs (7–19 weeks of gestation) undergoing neuronal (TUJ1, MAP2) versus glial (GFAP, S100B, A2B5) differentiation. Error bars represent standard error of the mean (SEM) over three at least three biological replicates.

(B) Average expression of *LIN28A* and *LIN28B* probe sets from (Patterson et al., 2012). PSC-NPCs p1 (n = 7), PSC-NPCs p2 (n = 2), PSC-NPCs p4 (n = 2), 6–8 week tissue-NPCs (n = 6), 15–19 week fetal NPCs (n = 5). Error bars represent SEM over biological replicate lines. Scale bars, 100 μm.

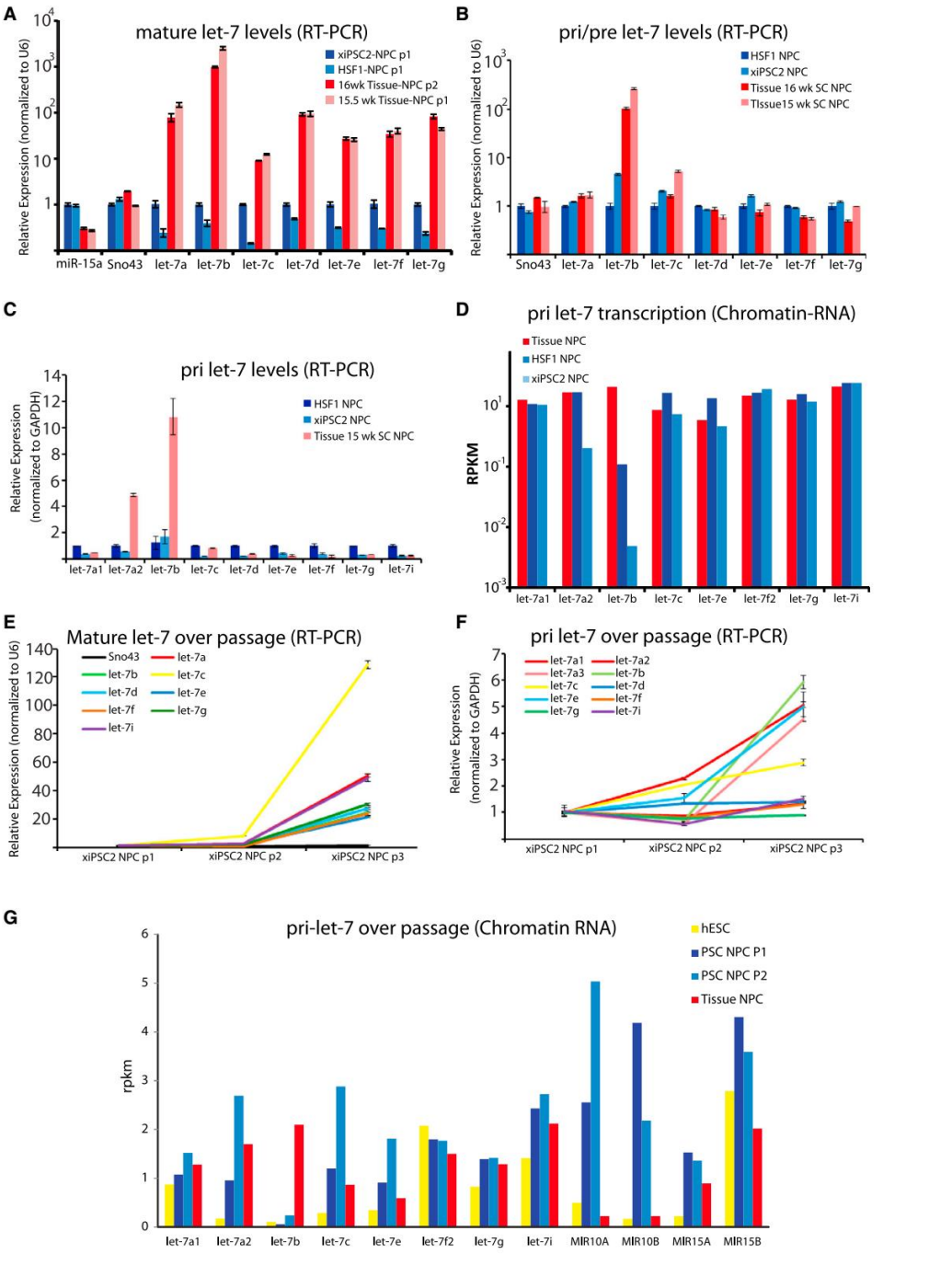
(C) Percentage of contribution of *let-7* miRNA family members to all miRNA present in cells derived from 16 week Tissue-NPCs (n = 2), 6–7 week Tissue-NPCs (n = 2), and PSC-NPCs (n = 3) by miRNA-seq. Fold change of normalized expression between 16 week Tissue-NPCs and PSC-NPCs for each family member is shown at right. Data for all miRNAs can be found in Table S1.

(D) Genes differentially expressed between PSC-NPCs (n = 7) and Tissue-NPCs (n = 5) as identified (Patterson et al., 2012) and the number of *let-7* targets represented in the data. *let-7* target lists were generated from published articles and TargetScan 5.2.

(E) Despite having significantly different levels of *let-7*, PSC- and Tissue-NPCs show similar proliferative potential as measured by cell-cycle analysis based on propidium iodide staining and flow cytometry.

(F) RNA-seq data from the Allen Institute's Brainspan developmental transcriptome database displayed as log-scale reads per kilobase measured ($\log_2$ RPKM) across the developing human brain. Top, transcription of neuronal, glial and oligodendrocyte marker genes (the average of all markers indicated by boldface line); middle, transcription of *LIN28B*; bottom, transcription of selected *let-7* target genes.

(G) Top, schematic of *LIN28*'s role in the progression of *let-7* microRNAs and the pathway's temporal correlation with development. Bottom, differentiation propensity of NPCs correlated with gestational age.



(legend on next page)

opposed to in silico prediction, and showed a large degree of polycistronic expression of groups of *let-7* family members (Table S2).

Manipulation of LIN28A/B Shows a Modest Effect on Developmental Maturity of NPCs

To determine what role *LIN28* had on the developmental maturity of NPCs, we utilized two strategies. First, we used a small interfering RNA (siRNA) approach to knockdown both homologs. When si*LIN28A* and si*LIN28B* were introduced by transfection, mRNA for these genes was suppressed by 70%–75% as measured by RT-PCR (Figure S3A). This was also confirmed by western blot (Figure S3B). As a result of this knockdown, mature *let-7* family miRNAs accumulated, as measured by RT-PCR, relative to transfection of a nontarget (siNT) control (Figure S3C). Also of note, siRNAs against both homologs were necessary to observe an induction of the *let-7* family members, demonstrating their known semiredundant roles in *let-7* maturation (Nam et al., 2011).

Microarray profiling identified many genes differentially expressed as a result of si*LIN28A/B* dual knockdown. However, these differentially expressed genes did not significantly overlap with the original list of probe sets differentially expressed between PSC-NPCs and their tissue-derived counterparts (Figure S3H; Patterson et al., 2012), nor were *let-7* downstream targets, besides *LIN28A* and *LIN28B*, selectively knocked down (Figure S3F). In addition, the dual knockdown had no effect on the propensity to differentiate into neurons or glia (Figure S3D). These findings suggested that RNAi based knockdowns can lower *LIN28* expression and induce *let-7* maturation, but the resulting subtle *let-7* induction exerted little effect functional effect. The second strategy to assess the functional relevance of *LIN28B* in NPC maturation utilized re-expression in Tissue-NPCs. When *LIN28B* was overexpressed in Tissue-NPCs by viral infection, mature *let-7* levels decreased (Figure S3E), a portion of the differentially expressed genes between Tissue-NPCs and PSC-NPCs were corrected (Figure S3H), including some *let-7* target genes (Figure S3F). Importantly, neurogenesis was promoted at the expense of gliogenesis upon induction of *LIN28B* in Tissue-NPCs (Figure S3G).

Direct Manipulation of *let-7* Levels Alters Cell Fate in Neural Progenitors

Previous studies either did not directly address the role of *let-7* in *LIN28*-mediated developmental progression or ruled it out altogether (Balzer et al., 2010; Yuan et al., 2012). To assay whether *let-7* miRNAs play a role in Tissue-NPCs, we introduced antagomirs against *let-7b* and *let-7g*. Microarray analysis on two independent experiments demonstrated that *HMGA2* specifically was strongly induced by antagomirs compared to nonspecific controls at both the RNA and protein levels (Figures 3A and 3B). Looking at global gene expression, 649 probe sets were differentially expressed between *let-7b/g* antagomir-treated and nontarget controls (antag CTRL). Of these probe sets, 181 overlapped with previous comparisons between PSC- and Tissue-NPCs ($p = 3.72 \times 10^{-55}$, Figure 3C), and *let-7* targets were specifically increased in these cells as expected (Figure 3D). Furthermore, Tissue-NPCs transfected with *let-7b/g* antagomirs were significantly more similar to PSC-NPCs than antag-CTRL-transfected cells on a global transcriptome level as measured by Pearson correlation (Figure 3E). Together, these data indicated that levels of *let-7* in PSC- and Tissue-NPCs play a role in the transcriptional disparity between these sources of NPCs, and that manipulation of *let-7* levels can shrink the dissimilarity between them. Finally, when antagomirs against *let-7b/g* were introduced into Tissue-NPCs, they became significantly more neurogenic and less gliogenic (Figure 3F; $p < 0.005$).

Figure 2. Dynamics of *let-7* Expression and Processing between PSC-NPCs and Tissue-NPCs

(A) RT-PCR for mature *let-7* family members normalized against small nucleolar RNA U6 in two PSC-NPCs and two Tissue-NPCs. miR-15a and Sno43 were assayed as controls.

(B) Real-time RT-PCR for pre and pri *let-7* family members normalized against small nucleolar RNA U6.

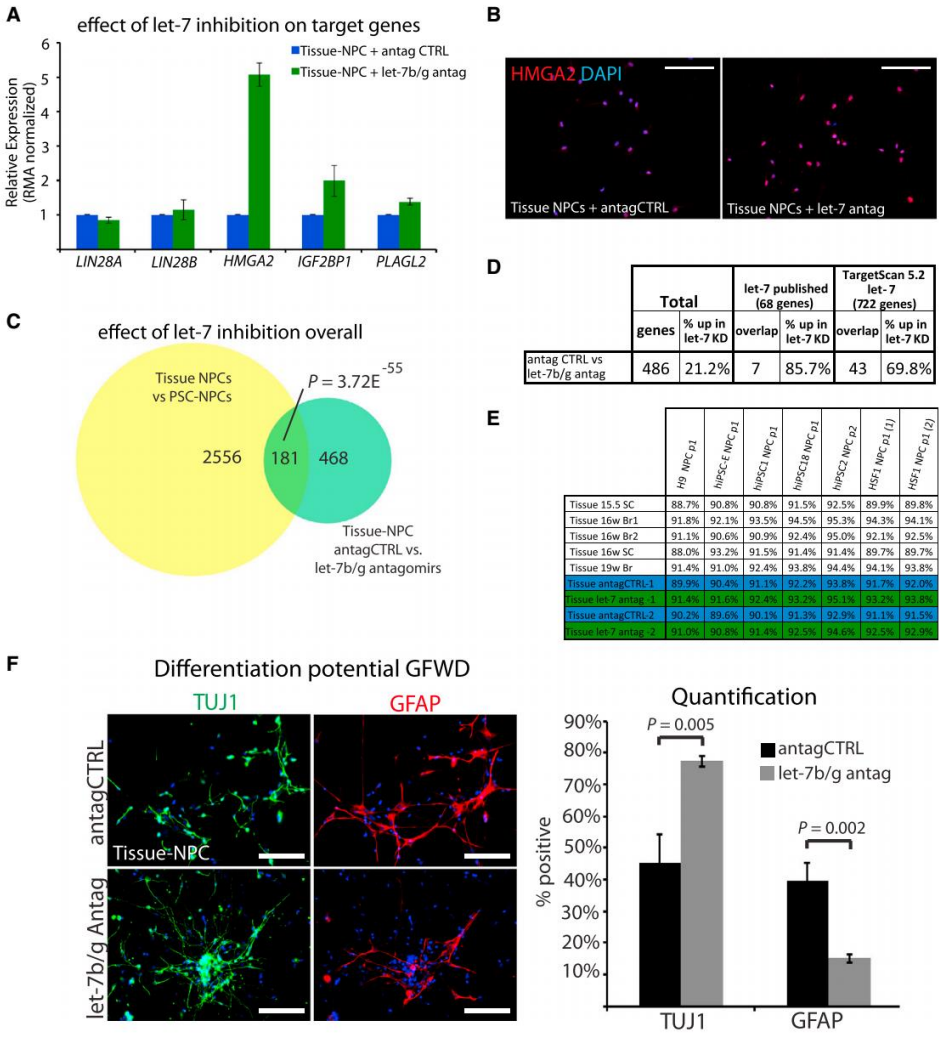
(C) Real-time RT-PCR for pri-*let-7* miRNAs normalized against the relative levels of GAPDH in each cell type.

(D) miRNA-seq of chromatin-associated primary *let-7* family miRNA transcripts between PSC-NPCs and Tissue-NPCs displayed as reads per kilobase measured (RPKM).

(E) Real-time RT-PCR for mature *let-7* family miRNAs in PSC-NPCs over three passages normalized against U6. Sno43 was run as a negative control and was unchanged over passage.

(F) Real-time RT-PCR for primary *let-7* family miRNA transcripts in PSC-NPCs over three passages normalized against U6.

(G) miRNA-seq of chromatin-associated primary *let-7* family miRNA transcripts in PSC-NPCs over two passages (n = 2 for each group) displayed as reads per kilobase measured (RPKM). Data for miRs-10a, 10b, 15a, and 15b are shown to indicate that *let-7* transcriptional regulation is unique in PSC versus Tissue-NPCs, even over passage, where *let-7a1*, *a2*, *b*, *c*, and *e* were induced. Error bars in all RT-PCR graphs represent SEM over three to four technical replicates, and results shown are representative of at least three independent experiments.



	Total		let-7 published (68 genes)		TargetScan 5.2 let-7 (722 genes)	
	genes	% up in let-7 KD	overlap	% up in let-7 KD	overlap	% up in let-7 KD
antag CTRL vs let-7b/g antag	486	21.2%	7	85.7%	43	69.8%

	H9 NPC p1	hiPSC-F NPC p1	hiPSC1 NPC p1	hiPSC18 NPC p1	hiPSC2 NPC p2	HSF1 NPC p1 (1)	HSF1 NPC p1 (2)
Tissue 15.5 SC	88.7%	90.8%	90.8%	91.5%	92.5%	89.9%	89.8%
Tissue 16w Br1	91.8%	92.1%	93.5%	94.5%	95.3%	94.3%	94.1%
Tissue 16w Br2	91.1%	90.6%	90.9%	92.4%	95.0%	92.1%	92.5%
Tissue 16w SC	88.0%	93.2%	91.5%	91.4%	91.4%	89.7%	89.7%
Tissue 19w Br	91.4%	91.0%	92.4%	93.8%	94.4%	94.1%	93.8%
Tissue antagCTRL-1	89.9%	90.4%	91.1%	92.2%	93.8%	91.7%	92.0%
Tissue let-7 antag -1	91.4%	91.6%	92.4%	93.2%	95.1%	93.2%	93.8%
Tissue antagCTRL-2	90.2%	89.6%	90.1%	91.3%	92.9%	91.1%	91.5%
Tissue let-7 antag -2	91.0%	90.8%	91.4%	92.5%	94.6%	92.5%	92.9%

Figure 3. A Role for LIN28/*let-7* in the Developmental Progression of Human Tissue-NPCs

(A) Averaged robust multiarray average (RMA) normalized expression of selected *let-7* target genes in Tissue-NPCs transfected with *let-7b/g* antagomiRs or nontargeted control antagomiRs by gene expression microarray. Error bars represent SEM across two biological replicates.

(B) Immunostaining for *HMGA2* in Tissue-NPCs showed that *let-7* antagomirs induce *HMGA2* at the protein level.

(C) Venn diagram demonstrating the original differences identified by Patterson et al. (2012) (yellow) and the overlap with gene expression differences (>1.54-fold change) between *let-7b/g* antagomirs versus nonspecific control (antagCTRL) in Tissue-NPCs (green; n = 2).

(D) Overlap of published and predicted *let-7* targets with genes changed after transfection of Tissue-NPCs with *let-7b/g* antagomiRs or nontargeted control antagomiRs measured by gene expression microarray.

(E) Pearson correlations of global gene expression similarity between Tissue-NPCs, Tissue-NPCs transfected with *let-7b/g* antagomiRs or nontargeted control antagomiRs, and PSC-NPCs.

(F) Immunofluorescence (LEFT) and quantification (RIGHT) for TUJ1 (neurons) and GFAP (glia) on Tissue-NPCs differentiated for 3 weeks in growth factor withdrawal following transfection with *let-7b/g* antagomirs or antag CTRL. p value was calculated with Student's t test for at least 1,200 cells across eight to ten fields of view. Error bars represent SEM over fields of view, and results shown are representative of at least three independent experiments. Scale bar, 100 μm.

To determine whether gliogenesis in PSC-NPCs could be regulated by *let-7*, mature oligonucleotides (mimics) for *let-7b* and *let-7g* were transfected into PSC-NPCs. Several *let-7* targets (*LIN28A*, *LIN28B*, *HMGA2*, *IGF2BP1*, and *PLAGL2)* were suppressed as measured by microarray profiling (Figure 4A). Global expression analysis of the of *let-7* mimic-treated NPCs over two independent experiments identified 331 probe sets differentially expressed between *let-7b/g*-transfected PSC-NPCs over respective nonspecific mimic control (miCTRL). Furthermore, among the transcriptional differences observed between these two conditions a significant number overlapped with the original list of genes differentially expressed between PSC- and Tissue-NPCs ($p = 3.76 \times 10^{-46}$, Figure 4B). Focusing on direct *let-7* downstream targets, expression profiling determined that not only were *let-7* target genes overrepresented in the data, but they were specifically downregulated in *let-7b/g*-transfected NPCs over miCTRLs (Figure 4C). Furthermore, *let-7b/g*-transfected PSC-NPCs were significantly more similar to 16–19 week tissue-NPCs than miCTRL-transfected PSC-NPCs on a transcriptome level as measured by Pearson correlation (Figure 4D). Following *let-7* induction, the PSC-NPCs were differentiated for 3 weeks in growth factor withdrawal media. miCTRL-transfected NPCs overwhelmingly produced $TUJ1^+$ neurons, whereas only rare $GFAP^+$ cells were found. When *let-7b/g* mimics were transfected prior to differentiation, the PSC-NPCs were slightly less neurogenic and significantly more gliogenic (Figure 4E; $p = 3.58E^{-05}$).

Taken together, experiments that employed direct regulation of *let-7* levels showed significant effects in both PSC- and Tissue-NPCs at both the molecular and functional levels, whereas attempts to manipulate *LIN28A/B* yielded more subtle effects. This observation is summarized in Figure S3, where manipulation of *LIN28A/B* only showed modest effects on both gene expression and differentiation in NPCs (Figures S3G and S3H). Instead, direct manipulation of *let-7* clearly shows that this miRNA family can control the fate of NPCs (Figures 3 and 4), in contrast to what was shown in a murine model (Balzer et al., 2010).

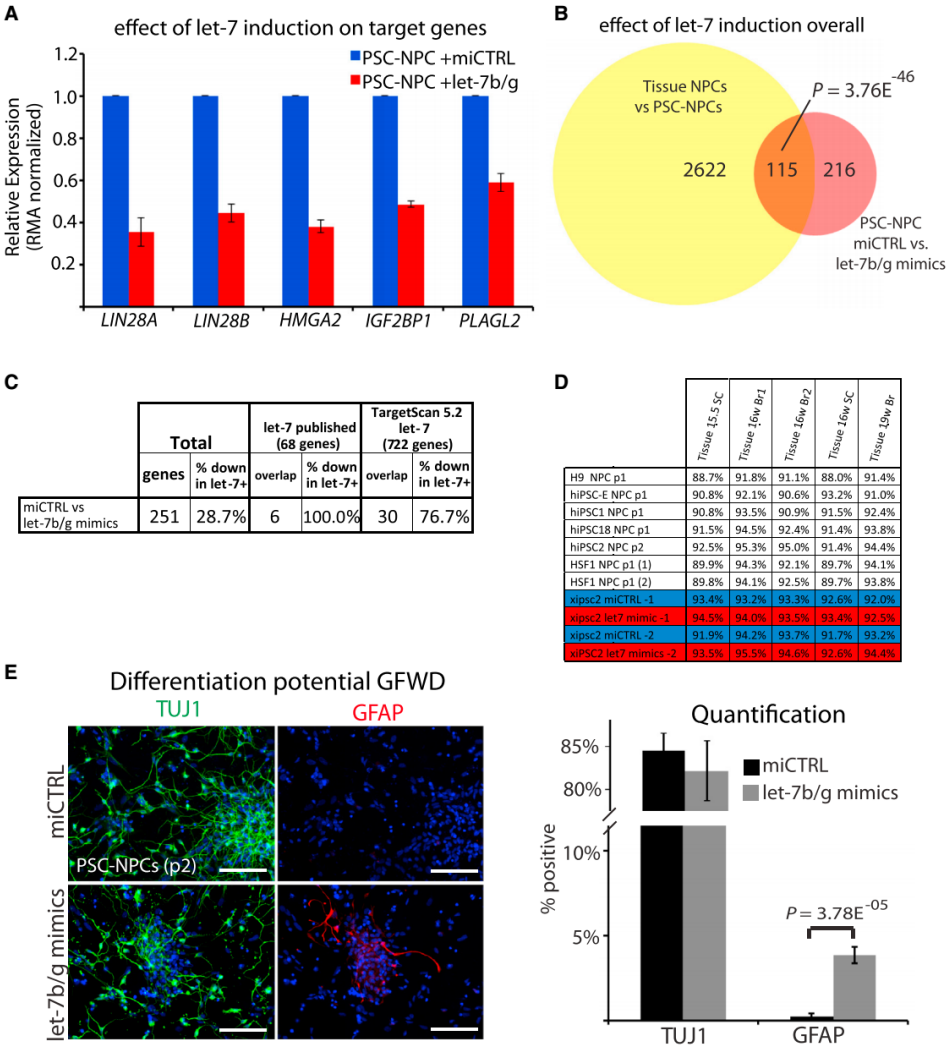
let-7 Acts through HMGA2 to Control Cell Fate

We hypothesized that the key *let-7* target genes responsible for cell fate in our model should show differences in gene expression levels in experimental settings presented here where NPC fate was affected (Figures 1A, 4E, and 5F). Taking the genes changed in every experimental model we produced that correlated with changes in neurogenesis versus gliogenesis, we narrowed the list of candidate genes. We furthered narrowed the list by taking gene expression data from the Allen Brain Atlas to identify genes changed from the pregliogenic state to the gliogenic stage (8–9 weeks PCW versus 16–18 weeks postconception) (Figure 6A). Just six genes appeared to have expression patterns that correlate with the switch from neurogenesis to gliogenesis in all these settings (Figure 5A). Of these six genes, two were predicted *let-7* target genes (*HMGA2* and *GABBR2*), but only *HMGA2* has been implicated in neurogenesis (Sanosaka et al., 2008). *HMGA2* was shown to affect cell fate in the murine brain (Nishino et al., 2008; Sanosaka et al., 2008) as well as various tissues (Monzen et al., 2008; Yu et al., 2013) and is thought to be mostly expressed prenatally (Ayoubi et al., 1999; Gattas et al., 1999). We found that *HMGA2* was among the genes whose expression was limited to fetal brain prior to 10 weeks of gestation, just prior to the surge of gliogenesis predicted by the Allen Database (Figure 1F).

A closer look at *HMGA2* transcription shows that at least eight transcripts are produced from this locus. One of these transcripts contains a large 3′ UTR that contains six *let-7* binding sites, whereas the other transcripts are predicted to be much less sensitive to suppression by *let-7*. Reanalysis of expression analysis by array with various probe sets that each recognize distinct transcripts showed that the dominant transcript expressed contains multiple *let-7* sites (identified by 208025_s_at), and that this is the only one affected by *let-7* manipulation (data not shown; Figure 5B). Furthermore, this particular transcript is the only one differentially expressed between PSC- and Tissue- NPCs (Figure 5B). The difference in expression of *HMGA2* between Tissue- and PSC- NPCs was also obvious at the protein level (Figure 5C). Taken together, these data suggested that suppression of *HMGA2* levels by induction of *let-7* correlated with the switch from neurogenesis to gliogenesis in the nervous system. To directly test the role of *HMGA2* in the developmental progression of PSC-NPCs, we silenced this gene by RNAi prior to terminal differentiation. Three distinct siRNAs were used (a–c), and just one of them (siHMGA2b) was predicted to target the long isoform with *let-7* sites. *HMGA2* message was suppressed by this method at both the RNA (Figure 5D, left) and protein levels (Figure 5D, right). As with induction of *let-7* mimics, suppression of *HMGA2* significantly induced gliogenesis in PSC-NPCs (Figure 5E).

The Notch Pathway Is a Key Effector of *let-7* in Gliogenesis

Profiling RNA expression after HMGA2 KD uncovered a modest number of genes changed. Intersection of these genes with those that were changed during human neurogenesis (Allen Database) or in Tissue-NPCs where *let-7* was inhibited by antagomirs found just two genes consistently altered (*HES5* and *USP44*)(Figure 6A). Significantly, *HES5* is a key target gene and effector of the NOTCH pathway, and perhaps the best known regulator of gliogenesis in mice (Hojo et al., 2000; Kageyama and Ohtsuka, 1999; Kageyama et al., 2008; Ohtsuka et al., 1999, 2001). RT-PCR



	Total		let-7 published (68 genes)		TargetScan 5.2 let-7 (722 genes)	
	genes	% down in let-7+	overlap	% down in let-7+	overlap	% down in let-7+
miCTRL vs let-7b/g mimics	251	28.7%	6	100.0%	30	76.7%

	Tissue 15.5 SC	Tissue 16w Br1	Tissue 16w Br2	Tissue 16w SC	Tissue 19w Br
H9 NPC p1	88.7%	91.8%	91.1%	88.0%	91.4%
hiPSC-E NPC p1	90.8%	92.1%	90.6%	93.2%	91.0%
hiPSC1 NPC p1	90.8%	93.5%	90.9%	91.5%	92.4%
hiPSC18 NPC p1	91.5%	94.5%	92.4%	91.4%	93.8%
hiPSC2 NPC p2	92.5%	95.3%	95.0%	91.4%	94.4%
HSF1 NPC p1 (1)	89.9%	94.3%	92.1%	89.7%	94.1%
HSF1 NPC p1 (2)	89.8%	94.1%	92.5%	89.7%	93.8%
xipsc2 miCTRL -1	93.4%	93.2%	93.3%	92.6%	92.0%
xipsc2 let7 mimic -1	94.5%	94.0%	93.5%	93.4%	92.5%
xipsc2 miCTRL -2	91.9%	94.2%	93.7%	91.7%	93.2%
xiPSC2 let7 mimics -2	93.5%	95.5%	94.6%	92.6%	94.4%

Figure 4. Direct Introduction of *let-7* miRNAs Affects Developmental Progression of PSC-NPCs

(A) Averaged RMA normalized expression of selected *let-7* target genes in PSC-NPCs transfected with *let-7b/g* mimics or nontargeted control miRNAs by gene expression microarray. Error bars represent SEM across two biological replicates.

(B) Venn diagram demonstrating the original differences identified by Patterson et al. (2012) (yellow), and the overlap with gene expression differences (>1.54-fold change) between *let-7b/g* mimics versus nonspecific control (miCTRL) in PSC-NPCs (red; n = 2).

(C) Overlap of published and predicted *let-7* targets with genes changed after transfection of PSC-NPCs with *let-7b/g* mimics or nontargeted control miRNAs measured by gene expression microarray.

(D) Pearson correlations of global gene expression similarity between PSC-NPCs, PSC-NPCs transfected with *let-7b/g* mimics or nontargeted control miRNAs, and Tissue-NPCs.

(E) Immunofluorescence (left) and quantification (right) for TUJ1 (neurons) and GFAP (glia) on PSC-NPCs differentiated for 3 weeks in growth factor withdrawal following transfection with *let-7b/g* mimics or miCTRL. p value indicated reflects Student's t test for at least 800 cells in multiple wells and across six fields of view. Error bars represent SEM over fields of view, and results shown are representative of at least three independent experiments. Scale bar, 100 μm.

of PSC-NPCs with knockdown of *HMGA2* clearly showed decreased *HES5* levels, suggesting a functional correlation between the two (Figure 6B). This correlation did not appear to be due to a difference in the activity of NOTCH, as both PSC and Tissue-NPCs showed similar activity (Figure 6C). Although NOTCH signaling has been shown to either promote or inhibit gliogenesis depending on the context, in our human model abrogation of all NOTCH signaling with gamma-secretase inhibitor (DAPT) stimulated gliogenesis and suppressed neuronal differentiation (Figure 6D).

Using the Allen database, it is clear that *HES5* is coexpressed with HMGA2 at 8–11 weeks of gestation just prior onset of gliogenesis, and expression of both genes drop significantly thereafter (Figure 6E). On the other hand, most other NOTCH effectors did not significantly change over the same time (Figure S4A). Expression patterns for *HES5* across nine different human cell types representing derivatives of all three germ layers suggested that *HES5* is restricted to the nervous system, whereas the other *HES/HEY* family members had a more widespread distribution throughout various organs (Figures 6F and S4B). This analysis also showed that PSC-NPCs expressed the highest levels of *HES5*, followed by lower expression in Tissue-NPCs, which were more gliogenic (Figure 1A; Patterson et al., 2012).

In order to functionally probe for a link between *HES5* and *let-7*, this miRNA was experimentally abrogated in Tissue-NPCs with antagomirs while suppressing *HES5* induction with siRNA. As shown above (Figure 3F), suppression of *let-7*s with antagomirs increased neurogenesis and decreased gliogenesis in Tissue-NPCs, but when *HES5* induction was simultaneously blocked by siRNA (Figure 6G), this effect was lost (Figure 6H). This result was consistent with the described role for *HES5* in murine neural development in the absence of LIF, whereas, in the presence of LIF, *HES5* has a progliogenic role (Chambers et al., 2001; Chenn, 2009; Hirabayashi and Gotoh, 2005).

As an AT-hook binding protein, *HMGA2* is well known to be broadly associated with chromatin. To begin to understand why *HMGA2* expression strongly correlated with *HES5* levels in NPCs, we probed whether *HMGA2* can regulate access of NICD/RBPj transcriptional complexes to the *HES5* promoter in response to NOTCH activation. Chromatin immunoprecipitation (ChIP) for NICD uncovered significant binding to two RBP binding sites within the *HES5* regulator region (Figure 6I), as expected in cells with strong NOTCH activation (Figure 6C). On the other hand, when *HMGA2* expression was abrogated by siRNA, NICD did not appear to significantly associate with these same sites, consistent with the notion that *HMGA2* plays a role in regulating access of NICD to the *HES5* promoter.

The *let-7* Effect on Gliogenesis Extends to Oligodendrocyte Formation

The fact that induction of *let-7* drove PSC-NPCs toward a more astrocytic fate suggested that *let-7* could play an important role in developmental switches. In the same vein, established protocols to generate oligodendrocytes with PSC-NPCs normally take at least 17 weeks (Stacpoole et al., 2013; Wang et al., 2013), which is akin to the time it takes to produce them in utero (Figure 1F). Using one of these protocols, we observed oligodendrocyte progenitors (OPCs) and mature oligodendrocytes after just 6–7 weeks of culture, but only in PSC-NPCs cultures following *let-7* induction at the NPC stage (Figure 7). These cells were judged to be bona fide oligodendrocytes on the basis of coexpression of O4 and Myelin Basic Protein (MBP) in confocal imaging. This result again suggested that *let-7* can play a role in the developmental maturity of PSC-NPCs and provides an approach to speed the generation of astrocytes and oligodendrocytes from human PSCs.

DISCUSSION

The critical advance of this study is definitively showing that *let-7* levels not only correlate with developmental progression in the human nervous system, but also serve to functionally drive the transition. Furthermore, whereas the *let-7* target gene *HMGA2* had previously been shown to promote neurogenesis in murine NPCs, we showed here that this is also the case for humans as well. Finally, we provide evidence that the mechanism by which *let-7* miRNAs drives progression in the nervous system is by altering expression of a key NOTCH effector gene *HES5*. This NOTCH effector, and NOTCH signaling in general, has been shown to be a promoter of gliogenesis in murine models, whereas our data suggest a proneurogenic role. This discrepancy could be due to several obvious differences in context: murine versus human models, the presence or absence of LIF signaling, and/or in vitro versus in vivo settings. Nevertheless, the effect in this human setting was clear: inhibiting NOTCH activity generated NPCs that went from negligible numbers to a state where over 10% of differentiated cells were astrocytes. We did not try the converse experiment, inducing NOTCH activity, because it appeared as though all NPCs tested, whether neurogenic or gliogenic, showed considerable NOTCH activity. Interestingly, data presented here suggest that no other NOTCH effectors were affected by *let-7* or *HMGA2* manipulation, only *HES5*. These data suggest that the Notch pathway can be regulated by *HES/HEY* effectors by fine-tuning access to the binding sites of its effectors.

However, this study provides only a partial understanding of the relation between HMGA2 and regulation of

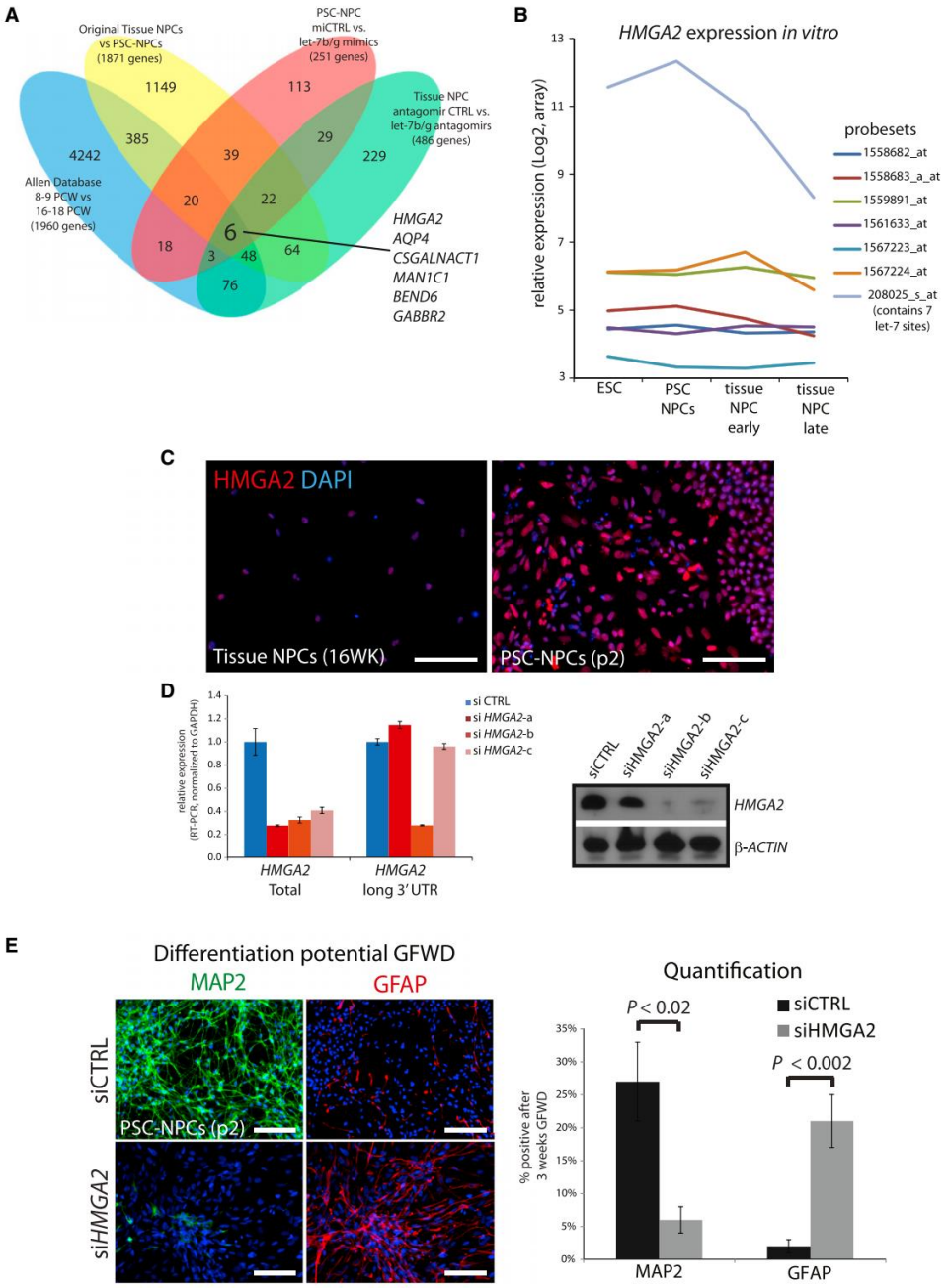


Figure 5. HMGA2 Is a Critical Target of *let-7* in Developmental Progression

(A) Venn diagram demonstrating the overlap of genes changed in comparisons between neurogenic and gliogenic NPCs in the indicated contexts. Original list (yellow): n = 7 PSC-NPC lines and n = 5 Tissue-NPC lines; mimic experiment (red): n = 2 for both mimic- and control-treated lines; antagomir experiment (green): n = 2 for both antag- and control-treated lines.

(legend continued on next page)

Notch signaling. This is due in part to a poorly defined binding pattern of HMGA2 throughout the genome. It is known to be abundant when expressed during development and, as an AT-binding protein, can in theory bind nearly the entire genome. Thus, whereas we show here that the expression of HMGA2 appears to affect access of NICD to the HES5 promoter, it is unclear whether this is due to direct binding at this locus, or a general effect on chromatin compaction.

This study also highlights a role for transcriptional control of let-7 miRNAs in development. It was thought that primary *let-7* transcription occurred at a constant level, and that levels of mature *let-7* were strictly a matter of *LIN28* activity. However, data provided here clearly demonstrate that *let-7b* is regulated at the transcriptional level even in cells that do not express *LIN28A/B*. Our data suggest that, developmentally, one *let-7* family member (pri-*let-7b*) is induced at the transcriptional level prior to the loss of *LIN28A/B* protein and may be a key driver of *LIN28A/B* suppression as gliogenesis is initiated. It will be important to determine whether the methods developed here to facilitate developmental progression also apply to cells that have been transplanted into tissue.

EXPERIMENTAL PROCEDURES

Cell Culture

Undifferentiated hPSC lines HSF1, H9, and xiPSC2 were maintained as previously described (Lowry et al., 2008) and in accordance with UCLA ESCRO. Neural progenitors were derived through formation of neural rosettes and maintained and further differentiated as described (Karumbayaram et al., 2009). NPCs were derived from fetal tissue (either brain or spinal cord) as described. All NPCs were judged to be pure by immunostaining as described (Patterson et al., 2012) (Figure S1).

Transfection

siRNAs (Thermo Dharmacon), *let-7* mimics (Thermo Dharmacon), and *let-7* antagomirs (Thermo Dharmacon "inhibitors") were transfected with Lipofectamine RNAiMAX (Invitrogen) at a ratio of 5 μl lipofectamine:20 nM siRNA or 5 μl lipofectamine:40 nM mimic or antagomirs for each well of a 6-well plate. A reverse transfection method was used. Briefly, Lipofectamine and oligos were premixed in 1 ml of OptiMEM (Gibco) for 25 min in the precoated receiving plate. Cells were then passaged with TrypLE (Gibco), resuspended in 1 ml of NPC media without antibiotics, and plated on top of transfection media. Transfections were incubated overnight at 37°C after which time media was replaced with standard NPC culturing media with antibiotics for the indicated lengths of time. Lentivirus for CMV-LIN28B-SV40-mCherry (Genecopoeia) or UBC-mCherry (Kohn lab) was made by the UCLA vector core. NPCs were reverse infected overnight in NPC media without antibiotic.

Immunostaining

PSC-NPCs or Tissue-NPCs differentiated for 3 weeks were passaged with TrypLE (GIBCO) to glass coverslips 2–7 days before fixing. Coverslips were fixed and stained using standard protocols as described (Patterson et al., 2012). Primary antibodies included rabbit × TUJ1 (1:2,000; Covance), rabbit × GFAP (1:1,000; Dako), rabbit × S100 (1:1,000; Dako), chicken × GFAP (1:2,000, Abcam), mouse × MAP2 (1:500; Abcam), rabbit × Cleaved Notch1 (Val1744) (1:500; Cell Signaling Technologies), rabbit × LIN28A (1:300; Cell Signal), rabbit × LIN28B (1:300; Cell Signal), rabbit × HMGI-C (HMGA2) (1:200, Santa Cruz Biotechnology), mouse × O4 (1:300; R&D Systems), rat × MBP (1:50; Abcam), mouse × A2B5 (1:1,000; Abcam). All images were captured on a Zeiss microscope using Axiovision software (Zeiss) for image capture. GFAP or TUJ1 positive cells were counted using ImageJ software as a percentage of the total DAPI-labeled nuclei or as a percentage of the mCherry$^+$ cells within the field. Average percentage positive cells and SEM was calculated over at least six fields of view. Western blot analysis was performed using standard procedures as described (Lowry et al., 2005). Primary western blot antibodies include rabbit × HMGI-C (1:500; Santa Cruz), rabbit × LIN28A (1:500; Cell Signal), rabbit × LIN28B (1:500; Cell Signal), and mouse × actin (1:1,000).

Additional materials and methods can be found in the Supplemental Information.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2014.08.015.

AUTHOR CONTRIBUTIONS

M.P. and X.G.: conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript; K.L., M.E., J.C., Y.X., J.L., S.A., S.M.D.: collection and assembly of data; S.S. and M.P.: data analysis and interpretation; W.E.L.: conception and design, data analysis and

(B) Normalized expression of all probe sets recognizing *HMGA2* in cell types representing developing NPCs. Probe set 208025_s_at is the only one that recognizes the *HMGA2* isoform containing multiple *let-7* target sites.
(C) Immunostaining for *HMGA2* in PSC and Tissue-derived NPCs.
(D) Real-time RT-PCR and western blot for *HMGA2* expression in PSC-NPCs transfected with either control siRNA or each of three different siRNAs against *HMGA2*.
(E) Percentage of positive PSC-NPCs transfected with either control siRNA or siRNA against HMGA2 undergoing neuronal (MAP2) versus glial (GFAP) differentiation following 3 weeks growth factor withdrawal. Images and quantification are representative of three independent experiments. At least 1,900 cells were analyzed across at least eight fields of view. Error bars represent SEM over fields of view, and these data are representative of at least three independent experiments. Scale bar, 100 μm.

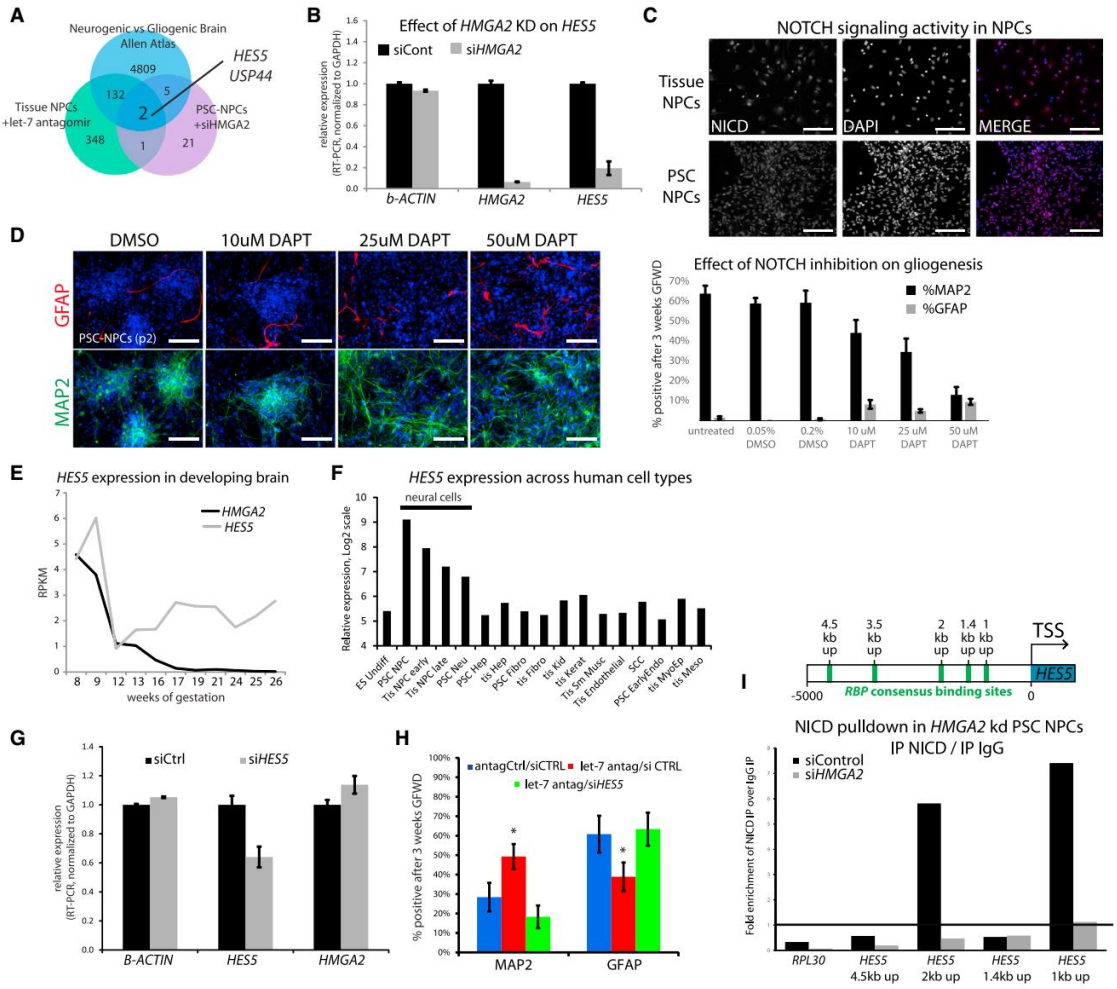


Figure 6. *let-7*/HMGA2 Regulates Notch Sensitivity through HES5

(A) Venn diagram demonstrating overlap of genes changed in comparisons between neurogenic and gliogenic NPCs in the indicated contexts.

(B) RT-PCR for relative expression after siRNA transfection to suppress *HMGA2* expression.

(C) PSC-NPCs and Tissue-NPCs show similar levels of nuclear NICD accumulation suggesting that both have similar activation of the NOTCH pathway. Scale bar, 100 μm.

(D) Using a gamma-secretase inhibitor to block NOTCH activation decreases neuronal differentiation and induces gliogenesis as measured by MAP2 and GFAP staining and quantification. Scale bar, 100 μm.

(E) Transcription of *HMGA2* and *HES5* quantified by RNA-seq across the developing human brain.

(F) Gene expression data taken from an in-house database of a variety of human cell types suggests that *HES5* is mostly expressed in the nervous system, whereas all the other family members are scattered throughout various types of cells from different organs (Log2 scale, RMA transformation).

(G) siRNA silencing of *HES5* shows a 40% knockdown of message compared to scramble control.

(H) As shown previously in Figure 5E, inhibition of *let-7* with antagomirs increased neurogenesis at the expense of gliogenesis. Silencing *HES5* blocks this effect demonstrating a clear link between *let-7* and HES5 activity in neurogenesis. *Student's t test $p < 0.05$ for $n > 450$ cells in multiple wells. Paired t tests were performed for each manipulation, and only those with significant changes were indicated by *. Error bars represent SEM over technical replicates. Results shown are representative of at least three independent experiments.

(legend continued on next page)

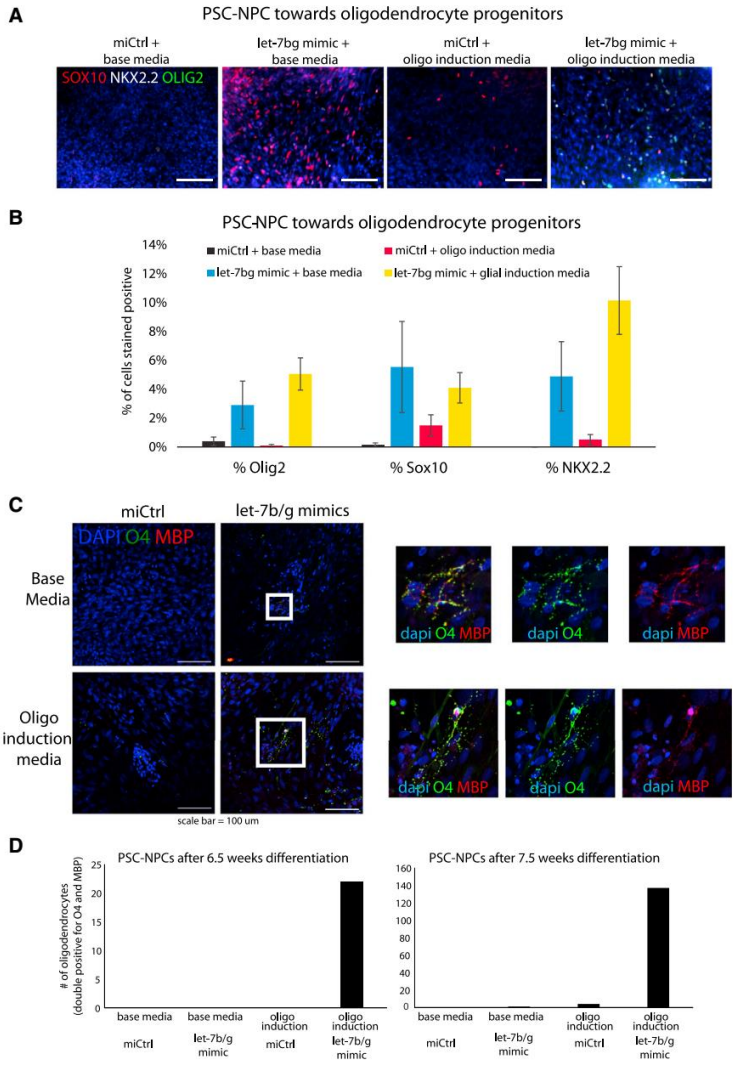


Figure 7. Generation of Oligodendrocytes Is Facilitated by Induction of *let-7*

(A) At 6 weeks of differentiation toward the oligodendrocyte lineage, OPC markers were identified by immunostaining and quantified in (B).

(B) Quantification of results presented in (A).

(C) Following at least 6 weeks of differentiation in either base or glial induction medium, mature oligodendrocytes were identified by immunostaining and quantified in (D). The bottom panels followed a glial induction protocol (Wang et al., 2013) and produced oligodendrocytes after just 6.5 weeks, again, only with *let-7* induction at the NPC stage prior to directed differentiation. At 7.5 weeks, the number of oligodendrocytes produced was considerably larger. Scale bar, 100 mm.

(D) A quantification of oligodendrocytes in each condition at two time points in each well. Cells were judged as bona fide oligodendrocytes if positive for both O4 and MBP.

interpretation, manuscript writing, final approval of manuscript, and financial support.

ACKNOWLEDGMENTS

We would like to thank the core facilities at UCLA (Flow Cytometry Core, EEBRCM; Clinical Genomics, Pathology). We would also like to thank Wen Gu for help with cell-cycle analysis and Gerry Weinmaster for useful suggestions on the Notch pathway. M.P. and Y.X. were supported by a CIRM Training Grant (TG2-01169). X.G. was supported by the MSTP program, David Geffen School of Medicine at UCLA. W.E.L. was supported by the Maria Rowena Ross Chair in Cell Biology and Biochemistry, and this work was supported by grants from CIRM (RB3-05207), NIH (P01GM99134-NIGMS), and

(I) Chromatin immunoprecipitation with antibody against the active NOTCH1 product, NICD, shows strong enrichment for NICD binding in at two predicted RBPj binding sites in the *HES5* upstream region (2 kp up and 1 kb up, see schematic). This enrichment was lost in cells with siHMGA2. Binding is represented as enrichment of pull-down over input and calculated as a function of enrichment of IgG over input. *RPL30* was employed as a negative control locus for a constitutively expressed gene not sensitive to NOTCH1 signaling. Shown is the average of two experiments.

an Eli & Edythe Broad Center of Regenerative Medicine and Stem Cell Research at UCLA Innovation Award.

Received: January 19, 2014
Revised: August 27, 2014
Accepted: August 28, 2014
Published: October 2, 2014

REFERENCES

Ayoubi, T.A., Jansen, E., Meulemans, S.M., and Van de Ven, W.J. (1999). Regulation of HMGIC expression: an architectural transcription factor involved in growth control and development. Oncogene *18*, 5076–5087.

Balzer, E., Heine, C., Jiang, Q., Lee, V.M., and Moss, E.G. (2010). LIN28 alters cell fate succession and acts independently of the let-7 microRNA during neurogliogenesis in vitro. Development *137*, 891–900.

Bhatt, D.M., Pandya-Jones, A., Tong, A.J., Barozzi, I., Lissner, M.M., Natoli, G., Black, D.L., and Smale, S.T. (2012). Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. Cell *150*, 279–290.

Chambers, C.B., Peng, Y., Nguyen, H., Gaiano, N., Fishell, G., and Nye, J.S. (2001). Spatiotemporal selectivity of response to Notch1 signals in mammalian forebrain precursors. Development *128*, 689–702.

Chang, C.J., Mitra, K., Koya, M., Velho, M., Desprat, R., Lenz, J., and Bouhassira, E.E. (2011). Production of embryonic and fetal-like red blood cells from human induced pluripotent stem cells. PLoS ONE *6*, e25761.

Chenn, A. (2009). A top-NOTCH way to make astrocytes. Dev. Cell *16*, 158–159.

Cimadamore, F., Amador-Arjona, A., Chen, C., Huang, C.T., and Terskikh, A.V. (2013). SOX2-LIN28/let-7 pathway regulates proliferation and neurogenesis in neural precursors. Proc. Natl. Acad. Sci. USA *110*, E3017–E3026.

Craig, A., Ling Luo, N., Beardsley, D.J., Wingate-Pearse, N., Walker, D.W., Hohimer, A.R., and Back, S.A. (2003). Quantitative analysis of perinatal rodent oligodendrocyte lineage progression and its correlation with human. Exp. Neurol. *181*, 231–240.

Dean, J.M., Moravec, M.D., Grafe, M., Abend, N., Ren, J., Gong, X., Volpe, J.J., Jensen, F.E., Hohimer, A.R., and Back, S.A. (2011). Strain-specific differences in perinatal rodent oligodendrocyte lineage progression and its correlation with human. Dev. Neurosci. *33*, 251–260.

Gattas, G.J., Quade, B.J., Nowak, R.A., and Morton, C.C. (1999). HMGIC expression in human adult and fetal tissues and in uterine leiomyomata. Genes Chromosomes Cancer *25*, 316–322.

Graf, R., Munschauer, M., Mastrobuoni, G., Mayr, F., Heinemann, U., Kempa, S., Rajewsky, N., and Landthaler, M. (2013). Identification of LIN28B-bound mRNAs reveals features of target recognition and regulation. RNA Biol. *10*, 1146–1159.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat. Struct. Mol. Biol. *16*, 1021–1025.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol. Cell *32*, 276–284.

Hirabayashi, Y., and Gotoh, Y. (2005). Stage-dependent fate determination of neural precursor cells in mouse forebrain. Neurosci. Res. *51*, 331–336.

Hojo, M., Ohtsuka, T., Hashimoto, N., Gradwohl, G., Guillemot, F., and Kageyama, R. (2000). Glial cell fate specification modulated by the bHLH gene Hes5 in mouse retina. Development *127*, 2515–2522.

Jakovcevski, I., Filipovic, R., Mo, Z., Rakic, S., and Zecevic, N. (2009). Oligodendrocyte development and the onset of myelination in the human fetal brain. Front Neuroanat. *3*, 5.

Kageyama, R., and Ohtsuka, T. (1999). The Notch-Hes pathway in mammalian neural development. Cell Res. *9*, 179–188.

Kageyama, R., Ohtsuka, T., and Kobayashi, T. (2008). Roles of Hes genes in neural development. Dev. Growth Differ. *50* (*Suppl 1*), S97–S103.

Karumbayaram, S., Novitch, B.G., Patterson, M., Umbach, J.A., Richter, L., Lindgren, A., Conway, A.E., Clark, A.T., Goldman, S.A., Plath, K., et al. (2009). Directed differentiation of human-induced pluripotent stem cells generates active motor neurons. Stem Cells *27*, 806–811.

Kim, V.N., and Nam, J.W. (2006). Genomics of microRNA. Trends Genet. *22*, 165–173.

Lee, S.H., Cho, S., Sun Kim, M., Choi, K., Cho, J.Y., Gwak, H.S., Kim, Y.J., Yoo, H., Lee, S.H., Park, J.B., and Kim, J.H. (2014). The ubiquitin ligase human TRIM71 regulates let-7 microRNA biogenesis via modulation of Lin28B protein. Biochim. Biophys. Acta *1839*, 374–386.

Lowry, W.E., Blanpain, C., Nowak, J.A., Guasch, G., Lewis, L., and Fuchs, E. (2005). Defining the impact of beta-catenin/Tcf transactivation on epithelial stem cells. Genes Dev. *19*, 1596–1611.

Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc. Natl. Acad. Sci. USA *105*, 2883–2888.

Mariani, J., Simonini, M.V., Palejev, D., Tomasini, L., Coppola, G., Szekely, A.M., Horvath, T.L., and Vaccarino, F.M. (2012). Modeling human cortical development in vitro using induced pluripotent stem cells. Proc. Natl. Acad. Sci. USA *109*, 12770–12775.

Monzen, K., Ito, Y., Naito, A.T., Kasai, H., Hiroi, Y., Hayashi, D., Shiojima, I., Yamazaki, T., Miyazono, K., Asashima, M., et al. (2008). A crucial role of a high mobility group protein HMGA2 in cardiogenesis. Nat. Cell Biol. *10*, 567–574.

Nam, Y., Chen, C., Gregory, R.I., Chou, J.J., and Sliz, P. (2011). Molecular basis for interaction of let-7 microRNAs with Lin28. Cell *147*, 1080–1091.

Nishino, J., Kim, I., Chada, K., and Morrison, S.J. (2008). Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression. Cell *135*, 227–239.

Nishino, J., Kim, S., Zhu, Y., Zhu, H., and Morrison, S.J. (2013). A network of heterochronic genes including Imp1 regulates temporal changes in stem cell properties. eLife *2*, e00924.

Ohtsuka, T., Ishibashi, M., Gradwohl, G., Nakanishi, S., Guillemot, F., and Kageyama, R. (1999). Hes1 and Hes5 as notch effectors in mammalian neuronal differentiation. EMBO J. *18*, 2196–2207.

Ohtsuka, T., Sakamoto, M., Guillemot, F., and Kageyama, R. (2001). Roles of the basic helix-loop-helix genes Hes1 and Hes5 in expansion of neural stem cells of the developing brain. J. Biol. Chem. *276*, 30467–30474.

Patterson, M., Chan, D.N., Ha, I., Case, D., Cui, Y., Van Handel, B., Mikkola, H.K., and Lowry, W.E. (2012). Defining the nature of human pluripotent stem cell progeny. Cell Res. *22*, 178–193.

Piskounova, E., Viswanathan, S.R., Janas, M., LaPierre, R.J., Daley, G.Q., Sliz, P., and Gregory, R.I. (2008). Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. J. Biol. Chem. *283*, 21310–21314.

Piskounova, E., Polytarchou, C., Thornton, J.E., LaPierre, R.J., Pothoulakis, C., Hagan, J.P., Iliopoulos, D., and Gregory, R.I. (2011). Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. Cell *147*, 1066–1079.

Sanosaka, T., Namihira, M., Asano, H., Kohyama, J., Aisaki, K., Igarashi, K., Kanno, J., and Nakashima, K. (2008). Identification of genes that restrict astrocyte differentiation of midgestational neural precursor cells. Neuroscience *155*, 780–788.

Stacpoole, S.R., Spitzer, S., Bilican, B., Compston, A., Karadottir, R., Chandran, S., and Franklin, R.J. (2013). High yields of oligodendrocyte lineage cells from human embryonic stem cells at physiological oxygen tensions for evaluation of translational biology. Stem Cell Reports *1*, 437–450.

Viswanathan, S.R., and Daley, G.Q. (2010). Lin28: a microRNA regulator with a macro role. Cell *140*, 445–449.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320*, 97–100.

Wang, D.J., Legesse-Miller, A., Johnson, E.L., and Coller, H.A. (2012). Regulation of the let-7a-3 promoter by NF-κB. PLoS ONE *7*, e31240.

Wang, S., Bates, J., Li, X., Schanz, S., Chandler-Militello, D., Levine, C., Maherali, N., Studer, L., Hochedlinger, K., Windrem, M., and Goldman, S.A. (2013). Human iPSC-derived oligodendrocyte progenitor cells can myelinate and rescue a mouse model of congenital hypomyelination. Cell Stem Cell *12*, 252–264.

West, J.A., Viswanathan, S.R., Yabuuchi, A., Cunniff, K., Takeuchi, A., Park, I.H., Sero, J.E., Zhu, H., Perez-Atayde, A., Frazier, A.L., et al. (2009). A role for Lin28 in primordial germ-cell development and germ-cell malignancy. Nature *460*, 909–913.

Yu, J., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917–1920.

Yu, K.R., Park, S.B., Jung, J.W., Seo, M.S., Hong, I.S., Kim, H.S., Seo, Y., Kang, T.W., Lee, J.Y., Kurtz, A., and Kang, K.S. (2013). HMGA2 regulates the in vitro aging and proliferation of human umbilical cord blood-derived stromal cells through the mTOR/p70S6K signaling pathway. Stem Cell Res. *10*, 156–165.

Yuan, J., et al. (2012). Lin28b reprograms adult bone marrow hematopoietic progenitors to mediate fetal-like lymphopoiesis. Science *335*, 1195–1200.

Zambidis, E.T., Peault, B., Park, T.S., Bunz, F., and Civin, C.I. (2005). Hematopoietic differentiation of human embryonic stem cells progresses through sequential hematoendothelial, primitive, and definitive stages resembling human yolk sac development. Blood *106*, 860–870.

Stem Cell Reports, Volume *3*

Supplemental Information

let-7 miRNAs Can Act through Notch to Regulate Human Gliogenesis

M. Patterson, X. Gaeta, K. Loo, M. Edwards, S. Smale, J. Cinkornpumin, Y. Xie, J. Listgarten, S. Azghadi, S.M. Douglass, M. Pellegrini, and W.E. Lowry

Supplemental Figure Legends

Supplemental Table 1. miRNA sequencing on PSC- and Fetal-derived NPCs Reads from miRNA sequencing for sequences (18-25bp in length) with homology to known miRNAs in 16 week Fetal-NPCs (n=2); 6-7 week Fetal-NPCs (n=2) and PSC-NPCs (n=3). Expression of each miRNA represented by the percent of the total RNA. Data sorted by fold change between 16 week Fetal-NPCs and PSC-NPCs.

Supplemental Table 2. mRNA changes in PSC-NPCs after HMGA2 KD In two independent experiments, *HMGA2* was silenced by siRNA-b. 5 days after transfection, the cells were lysed and array analysis was performed. This list of genes represent those that satisfied at student's t-test ($p < 0.02$) and were changed by at least 1.5 fold.

Supplemental Figure 1. Validation of PSC-NPCs and comparison to Tissue-NPCs

(A) Validation of identity of NPCs used in this manuscript. Shown is a quantification of an immunostaining typical of PSC and Tissue-derived NPCs. The vast majority of the cells are positive for at least SOX1, SOX2 and NESTIN, as shown here. (B) RNA-seq data from the Allen Institute's Brainspan developmental transcriptome database showing the relative expression of *let-7* miRNAs displayed as log-scale reads per kilobase measured (log_2 RPKM). miR-98 was not detectable in this dataset at this timescale. (C) Microarray gene expression data from PSC-derived NPCs (n=30 samples) and Tissue-derived NPCs ranging from 6.5 weeks of gestation to 19 weeks of gestation (n=23 samples) were grouped by origin, RMA normalized, and compared using Genespring GX software. Probsesets identified as significantly changed between the two cell types ($p<0.01$ by unmodified t-test after Benjamini-Hochberg FDR correction) and demonstrating >1.54 fold change in either direction were included. A heatmap was generated for the 5479 microarray probesets satisfying these conditions, and these

probesets were ordered by an unbiased hierarchical clustering algorithm. Probesets were curated manually into 6 cluster. Selected gene symbols are listed right of the heatmap for *let-7* miRNA family members (in black) and a subset of predicted *let-7* targets with the most conserved target sites as generated by Targetscan 6.2 (in blue). Table at right shows Gene Ontology terms significantly enriched in each group ($p<0.01$ after Benjamini-Hochberg correction) and filtered for redundancy.

Supplemental Figure 2. Annotation of *let-7* family members within human genome

Shown is a table to annotate all members of the *let-7* family of miRNAs based on their transcription as measured by Chromatin-RNAseq, followed by mapping onto the human genome (Bhatt et al., 2012).

Supplemental Figure 3. Manipulation of *LIN28A/B* in NPCs

(A) Real-time RT-PCR for *LIN28A* and *LIN28B* normalized to *GAPDH* following transfection of si*LIN28B* alone, si*LIN28A* and si*LIN28B* together, or of non-target control (siNT) 2 and 4 days post transfection. Error bars represent SEM over technical replicates. (B) Western blot against LIN28A, LIN28B, and β-ACTIN proteins following transfection of indicated siRNAs. (C) Real-time RT-PCR of mature let-7 family members following transfection of indicated siRNAs. Error bars represent SEM over technical replicates. (D) results from GFWD assay following siRNA for *LIN28A/B*, indicated no significant effect of silencing *LIN28* on neurogenesis or gliogenesis. (E) RT-PCR from experiments where a retrovirus was used to overexpress *LIN28B* (+ IRES Cherry) in Tissue-NPCs. Levels of mature *let-7* family members demonstrated that LIN28B overexpression suppressed all *let-7* miRNAs. Expression normalized to U6. (F) Genes differentially expressed between Tissue-NPCs overexpressing *LIN28B* or Cherry alone and the number of *let-7* targets represented in the data. *Let-7* target lists were generated from published articles and TargetScan 5.2. (G) Immunostaining and quantification to identify the percent

positive Tissue-NPCs infected with either *LIN28B* or Cherry alone undergoing neuronal (TUJ1) vs. glial (GFAP) differentiation following 3 weeks growth factor withdrawal. Note that in these experiments, only Cherry positive cells were quantified for the indicated differentiation marker. P values for student's t-test are indicated for N > 90 Cherry+ cells in multiple wells. Results shown are representative of three independent experiments. (H) Venns comparing gene expression changes in PSC-NPCs after si*LIN28A/B* treatment, Tissue-NPCs after *LIN28B* overexpression, PSC-NPCs after *let-7b/g* mimic treatment, and Tissue-NPCs after *let-7b/g* antagomiR treatment compared against gene expression differences between PSC-NPCs and Tissue-derived NPCs. Direct manipulations of *let-7* miRNAs are more effective than *LIN28* manipulations at changing the expression of those genes as measured by overlaps in the venn diagrams. (I) Quantification of the number of genes changed and overlap in the experiments listed in H.

Supplemental Figure 4. HES family expression across neural development and various human cell types. (A) Data taken from the Allen Brain Atlas suggests that the expression of most HES family members do not change dramatically. (B) Gene expression data taken from an in-house database of a variety of human cell types suggests that *HES5* is mostly expressed in the nervous system, whereas all the other family members are scattered throughout various types of cells from different organs (Log2 scale, RMA transformation).

Supplementary Materials and Methods

Expression Analysis: mRNA

RNA isolation, mRNA reverse transcription, and real-time PCR were performed as described (Lowry et al., 2008). RNA Microarray profiling was performed with Affymetrix Human HG-U133 2.0 Plus arrays as described (Chin et al., 2009; Lowry et al., 2008). Data were normalized with Robust Multichip Algorithm in Genespring separately from previous analyses performed(Patterson et al., 2011). Probe sets that were not expressed at a raw value of > 50 in

at least 10% of samples were eliminated from further analysis. Only those probe sets that made it past the filtering after both normalizations were included in this analysis. Gene expression differences for the *let-7* manipulations (n=2) and for siRNA experiments (n=2) were identified by a fold change of ≥1.54 in pairwise samples, where both biological replicates had a >1.54 fold change compared to their respective control. For *LIN28B* overexpression experiments (n=3), gene expression differences were identified by a paired student t-test ($p<0.05$) and a fold change of >1.54. Gene expression differences for the original list of genes differentially expressed between PSC-NPCs and Tissue-NPCs were determined as previously described(Patterson et al., 2011). Further statistical analysis for hypergeometric distribution and three-way simulation was performed with R, package 2.9.2 as described(Chin et al., 2009). Overlap with *let-7* target genes was performed in Excel using lists generated either from the literature or from TargetScan 5.2. 77 published targets and 751 Target scan targets were originally identified; however some targets were excluded in later analyses because they were not represented in the normalized, filtered data. Pearson correlations were performed in Excel using all normalized, filtered probe sets. All of the microarray data are included in NIHGEO Dataset 51635 and 47796.

Expression analysis: miRNA and mRNA

miRNA was isolated by stratagene RNA isolation kit. Mature and pre/pri miRNA reverse transcription was performed with miScript reverse transcription kit (Qiagen) and real-time PCR was performed using miScript Sybr Green PCR kit and miScript Primer and Precursor Assays (Qiagen). Samples were normalized against small nucleolar RNAs, U6 and Sno43. pri miRNA alone was treated like mRNA and normalized against GAPDH. miRNA seq was performed with Illumina TruSeq Small RNA Sample Prep Kit - Set A (cat. #: RS-200-0012) for constructing the small RNA library; TruSeq PE Cluster Kit v3 - cBot - HS (cat#: PE-401-3001) for template amplification and TruSeq SBS Kit v3 - HS (200-cycles) (cat#: FC-401-3001) for sequencing.

The sequencing reactions were run on HiSeq 2000 as a pair-end 100bp run. Data were normalized as a percent of the total RNA. Individual reads were aligned to known miRNAs with no more than two mismatches. Individual reads for each miRNA were summed together and summed totals were averaged across cohorts: 16wk Tissue-NPC (n=2); 8 wk Tissue-NPC (n=2); PSC-NPC (n=3). Chromatin isolation and RNA sequencing was performed as described previously (Bhatt et al., 2012). The primers used for RT-PCR on mature miRNAs are proprietary but available from Qiagen as a kit for detection.

Primers used for RT-PCR:

Target	**Forward primer**	**Reverse Primer**
Primers for RT-PCR for relative mRNA levels		
HMGA2 total	ACTTCAGCCCAGGGACAAC	TTGAGCTGCTTTAGAGGGAC
HMGA2 long 3' UTR	GGCCAGCTCATAAAATGGAA	TACTGTTCCATTGGCCACAA
LIN28A	CACAGCCCTACCCTGCTC	CTAGCCCCAATGCACCCTAT
LIN28B	CACCAAGCTGGCTTCAATTA	CGCATTCCAAACATCTACCC
GAPDH	CGACCACTTTGTCAAGCTCA	AGGGGTCTACATGGCAACTG
B-ACTIN	GATCATTGCTCCTCCTGAGC	AAAGCCATGCCAATCTCATC
pri-*let-7a1*	GATTCCTTTTCACCATTCACCCTGGATGTT	TTTCTATCAGACCGCCTGGATGCAGACTTT
pri-*let-7a2*	TGTTTAGTGCAAGACCCAAGG	GACTGCATGCTCCCAGGT
pri-*let-7a3*	CCCTTTGGGGTGAGGTAGTAG	CAGATATTACAGCCACTTCAGGAAA
pri-*let-7b*	GCCCCTCGGAAGATAACTATACA	CAAGTTCATGGTCAGAACAGCTT
pri-*let-7c*	TGAAGCAACATTGGAAGCTG	GCCCAAATCAATGATCCAAG
pri-*let-7d*	TGCCAAGTAGAAGACCAGCA	CACCAAAGCAAAGTAGCAAGG
pri-*let-7e*	CCTGGTCCCTGTCTGTCTGT	CAGCCCAGTGGGGCTAAG
pri-*let-7f*	TTGCTTCTTGTCTTATTTCTCTGTG	TCCTCAGGGAAGGCAATAGA
pri-*let-7g*	CCATTACCTGGTTTCCCAGAGA	GTTCCTCCAGCGCTCCGTT
pri-*let-7i*	CGAGGAAGGACGGAGGAG	ACCAGCACTAGCAAGGCAGT

ChIP-PCR

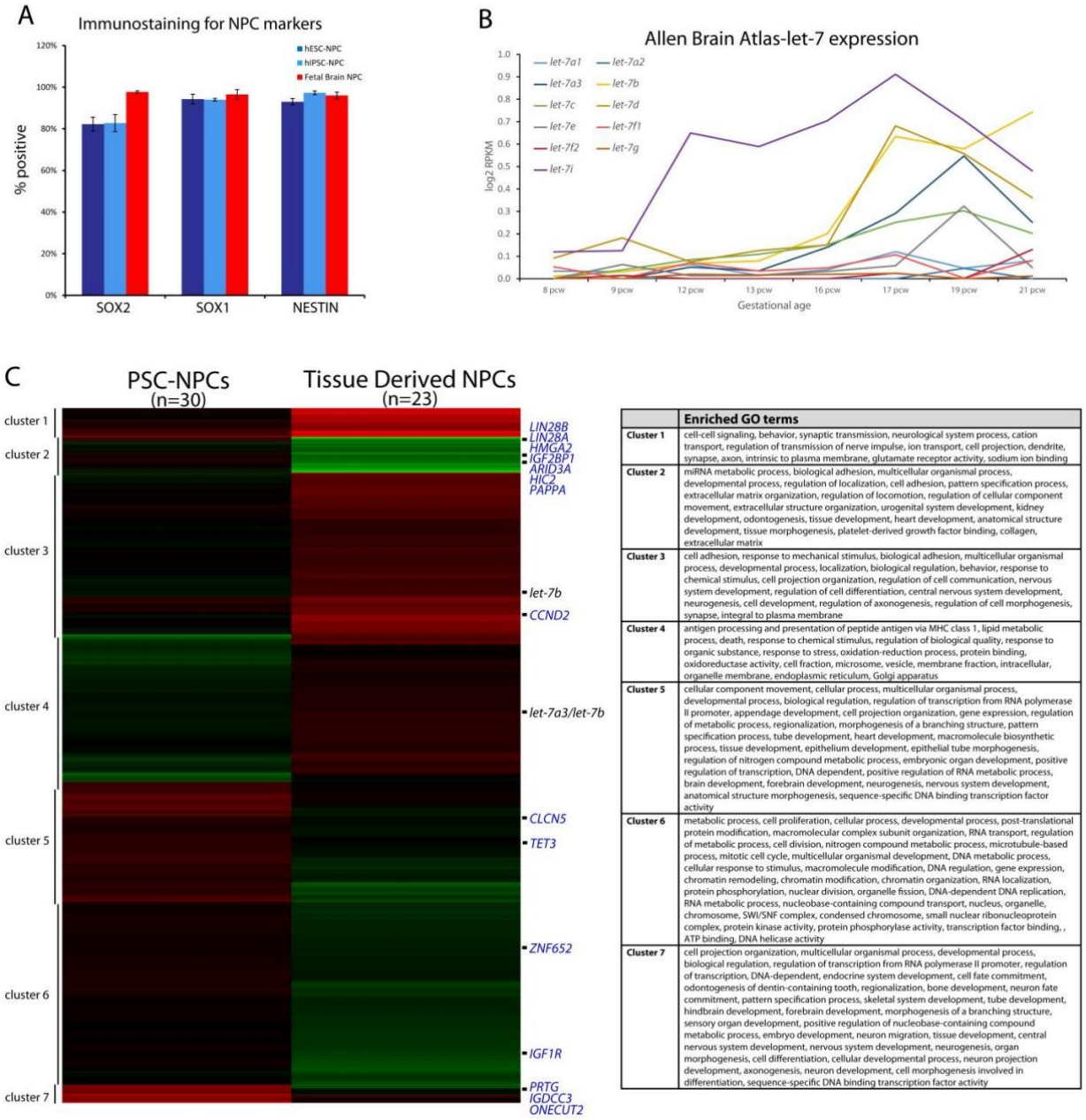
PSC derived NPCs were crosslinked with 1% PFA and chromatin was extracted and purified using a SimpleChIP Enzymatic Chromatin IP kit with magnetic beads (Cell Signaling). IP was performed with normal Rabbit IgG, Rabbit Histone H3, and Rabbit Cleaved Notch1 (Val1744) (D3B8) antibodies (Cell Signaling). DNA was eluted from the beads and used as template for PCR with Real-Time Roche Lightcycler. Signal was quantified first as a function of the

appropriate input sample with the same primers, and then expressed as a function of enrichment over IgG control.

Primers used:

RPL30 exon 3	AAGTCGCTGGAGTCGATCAA	CGATTACCTCAAAGCTGGGC
HES5 4.5kb upstream	GCCGCTGTCTCTGAAATCTG	CCTCTGGGTGGAAACTCTACT
HES5 2kb upstream	CCGCCATCAATGCCCAGA	CTTGATGGATGCAGGAGGG
HES5 1.4kb upstream	CTGCCTAACCAGCCCTGAT	GCTCCTAGAGACAGGTTGGG
HES5 1kb upstream	AAAGCAAGCCCTACAAGTGC	GCCACTGTTTATGCATGCCA

Supplemental Figure 1

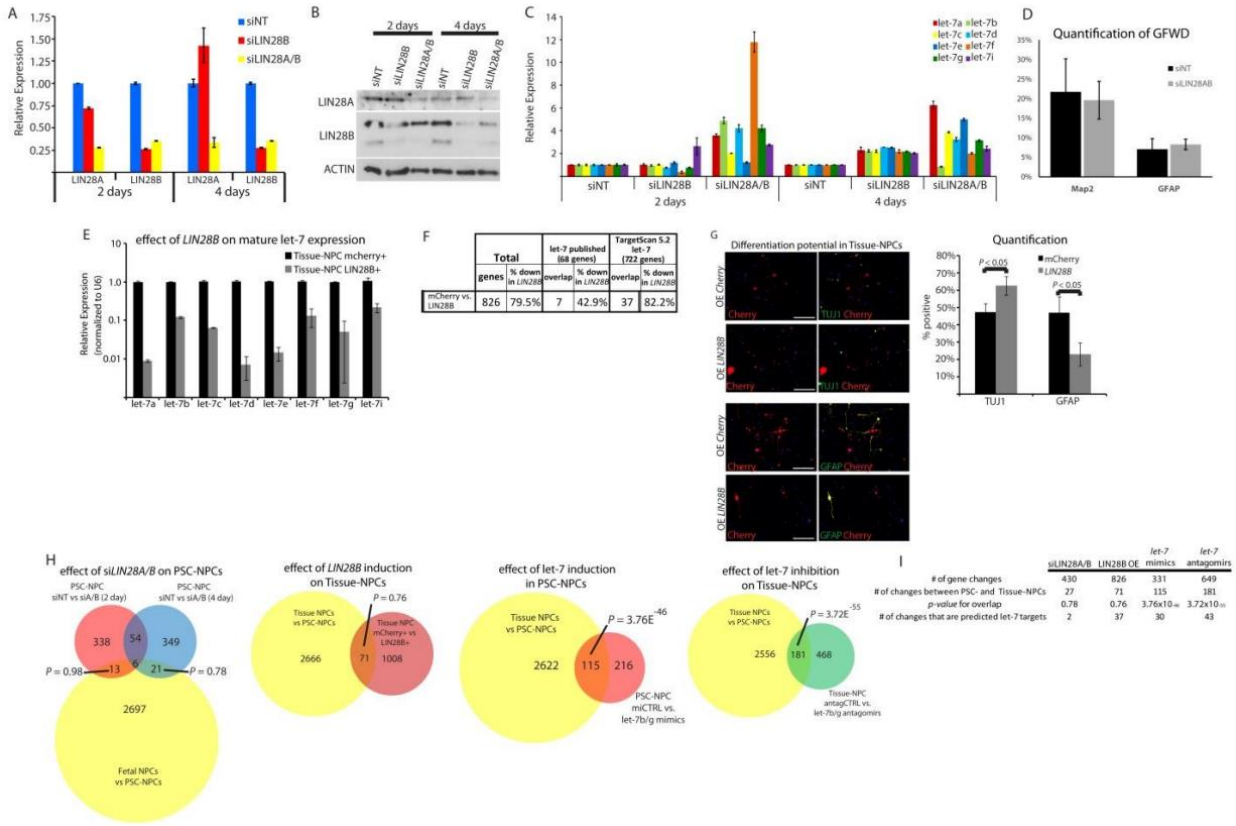


	Enriched GO terms
Cluster 1	cell-cell signaling, behavior, synaptic transmission, neurological system process, cation transport, regulation of transmission of nerve impulse, ion transport, cell projection, dendrite, synapse, axon, intrinsic to plasma membrane, glutamate receptor activity, sodium ion binding
Cluster 2	miRNA metabolic process, biological adhesion, multicellular organismal process, developmental process, regulation of localization, cell adhesion, pattern specification process, extracellular matrix organization, regulation of locomotion, regulation of cellular component movement, extracellular structure organization, urogenital system development, kidney development, odontogenesis, tissue development, heart development, anatomical structure development, tissue morphogenesis, platelet-derived growth factor binding, collagen, extracellular matrix
Cluster 3	cell adhesion, response to mechanical stimulus, biological adhesion, multicellular organismal process, developmental process, localization, biological regulation, behavior, response to chemical stimulus, cell projection organization, regulation of cell communication, nervous system development, regulation of cell differentiation, central nervous system development, neurogenesis, cell development, regulation of axonogenesis, regulation of cell morphogenesis, synapse, integral to plasma membrane
Cluster 4	antigen processing and presentation of peptide antigen via MHC class 1, lipid metabolic process, death, response to chemical stimulus, regulation of biological quality, response to organic substance, response to stress, oxidation-reduction process, protein binding, oxidoreductase activity, cell fraction, microsome, vesicle, membrane fraction, intracellular, organelle membrane, endoplasmic reticulum, Golgi apparatus
Cluster 5	cellular component movement, cellular process, multicellular organismal process, developmental process, biological regulation, regulation of transcription from RNA polymerase II promoter, appendage development, cell projection organization, gene expression, regulation of metabolic process, regionalization, morphogenesis of a branching structure, pattern specification process, tube development, heart development, macromolecule biosynthetic process, tissue development, epithelium development, epithelial tube morphogenesis, regulation of nitrogen compound metabolic process, embryonic organ development, positive regulation of transcription, DNA dependent, positive regulation of RNA metabolic process, brain development, forebrain development, neurogenesis, nervous system development, anatomical structure morphogenesis, sequence-specific DNA binding transcription factor activity
Cluster 6	metabolic process, cell proliferation, cellular process, developmental process, post-translational protein modification, macromolecular complex subunit organization, RNA transport, regulation of metabolic process, cell division, nitrogen compound metabolic process, microtubule-based process, mitotic cell cycle, multicellular organismal development, DNA metabolic process, cellular response to stimulus, macromolecule modification, DNA regulation, gene expression, chromatin remodeling, chromatin modification, chromatin organization, RNA localization, protein phosphorylation, nuclear division, organelle fission, DNA-dependent DNA replication, RNA metabolic process, nucleobase-containing compound transport, nucleus, organelle, chromosome, SWI/SNF complex, condensed chromosome, small nuclear ribonucleoprotein complex, protein kinase activity, protein phosphorylase activity, transcription factor binding, , ATP binding, DNA helicase activity
Cluster 7	cell projection organization, multicellular organismal process, developmental process, biological regulation, regulation of transcription from RNA polymerase II promoter, regulation of transcription, DNA-dependent, endocrine system development, cell fate commitment, odontogenesis of dentin-containing tooth, regionalization, bone development, neuron fate commitment, pattern specification process, skeletal system development, tube development, hindbrain development, forebrain development, morphogenesis of a branching structure, sensory organ development, positive regulation of nucleobase-containing compound metabolic process, embryo development, neuron migration, tissue development, central nervous system development, nervous system development, neurogenesis, organ morphogenesis, cell differentiation, cellular developmental process, neuron projection development, axonogenesis, neuron development, cell morphogenesis involved in differentiation, sequence-specific DNA binding transcription factor activity

Supplemental Figure 2

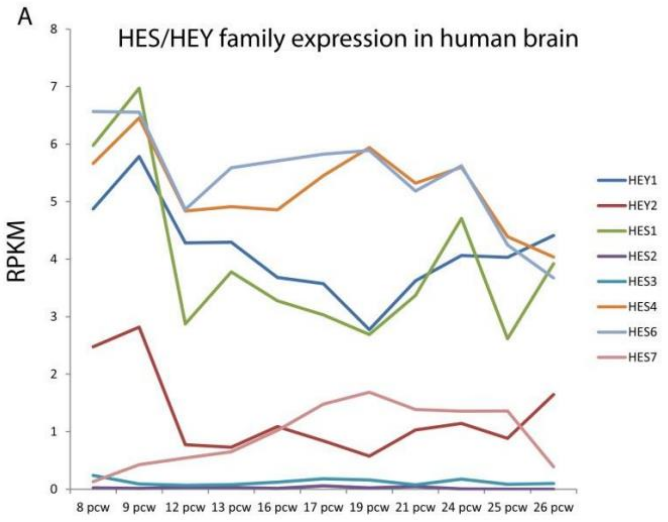
Let7 family member	Other miRs	Chr	Strand	Start	End	Length (bp)	miRStart Annotation
let7b	4763, 3619	22	+	46,467,707	46,514,431	46,724	Based on one existing MIRLET7B host gene. One of the predicted start sites corellates with chromatin data **TSS 46467598: 109bp difference** **http://mirstart.mbc.nctu.edu.tw/mirna.php?no=50548**
let7a3		22	+	46,467,707	46,514,431	46,724	
let7a2	100, 125B1	11	-	121,961,330	122,076,912	115,582	Longest start site prediction is ~40kb upstream of mature miRNA **http://mirstart.mbc.nctu.edu.tw/mirna.php?no=50546**
let7a1		9	+	96,929,333	96,967,519	38,186	One of the predicted start sites overlaps well with chromatin data and has a block of conservation **TSS: 96929288 45bp difference**
let-7d		9	+	96,929,333	96,967,519	38,186	**http://mirstart.mbc.nctu.edu.tw/mirna.php?no=50545**
let-7f1		9	+	96,929,333	96,967,519	38,186	
let-7c	99A, 125B2	21	+	17,442,842	18,024,259	581,417	**http://mirstart.mbc.nctu.edu.tw/mirna.php?no=50549**
let-7e	99B, 125A	19	+	52,193,049	52,214,933	21,884	One of the predicted start sites overlaps well with chromatin data and has a block of conservation **TSS: 52193273 224bp difference** **http://mirstart.mbc.nctu.edu.tw/mirna.php?no=50551**
let-7f2	98	X	-	53,559,057	53,713,697	154,640	**http://mirstart.mbc.nctu.edu.tw/mirna.php?no=50553**
let-7g		3	-	52,288,438	52,312,659	24,221	**http://mirstart.mbc.nctu.edu.tw/mirna.php?no=50896**
let-7i		12	+	62,993,322	63,024,823	31,501	No predicted site with 3kb start site predicted by Chromatin data. All the other predicted start sites are upstream within another coding gene MON2 **http://mirstart.mbc.nctu.edu.tw/mirna.php?no=50897**

Supplemental Figure 3



Supplemental Figure 4

A



B

Gene of Interest	ES Undiff	PSC NPC	Tis NPC early	Tis NPC late	PSC Neu	PSC Hep	tis Hep	PSC Fibro	tis Fibro	tis Kid	tis Kerat	Tis Sm Musc	Tis Endothelial	SCC	PSC EarlyEndo	tis MyoEp	tis Meso
HMGA2	11.56	12.33	10.87	8.32	10.33	10.78	6.06	11.18	9.13	10.63	11.89	10.37	9.88	11.12	11.42	10.63	9.27
Hes1	8.33	10.68	10.42	9.88	10.84	9.90	8.04	9.57	6.22	7.23	9.77	7.31	9.70	8.04	8.73	7.99	7.64
Hes2	6.71	6.84	6.79	6.59	6.90	6.95	7.88	6.75	6.70	6.72	9.92	6.66	8.51	7.27	6.91	7.81	6.81
Hes4	6.44	10.14	9.10	8.96	9.86	9.49	6.93	8.01	6.25	7.86	8.00	5.91	8.29	8.21	7.21	5.78	6.11
Hes5	5.41	9.11	7.95	7.20	6.79	5.24	5.74	5.39	5.24	5.84	6.06	5.29	5.33	5.78	5.07	5.90	5.52
Hey1	7.49	10.12	10.08	10.68	10.10	9.78	6.43	7.30	6.52	6.64	7.11	6.96	10.24	8.14	8.70	5.77	6.58
Hey2	8.71	7.26	10.26	9.09	9.04	7.65	6.28	5.06	4.80	4.92	4.59	5.78	8.03	5.32	8.69	4.41	4.79
Notch1	8.51	9.78	10.97	10.44	9.95	8.35	8.04	8.13	7.59	8.34	9.97	8.09	10.19	7.77	8.65	8.95	8.13

Supplemental Table 2

Expression in PSC HMGA2 kd v. control PSC NPCs

Gene Name	Fold Change
HES5	4.83
ARHGEF4	2.00
MBTD1	1.98
GAD1	1.97
LUC7L3	1.86
BEND5	1.79
REM2	1.77
SFRP1	1.73
USP44	1.71
DAAM1	1.70
WNT4	1.67
C21orf71	1.61
GABBR2	1.61
DSC3	1.61
LOC147670	1.60
PREX1	1.60
SOX2	1.58
HOXD3	1.57
IKZF4	1.54
FERMT2	-1.57
OXTR	-1.58
ATP11B	-1.62
EYA4	-1.62
TBC1D8B	-1.67
TGFB2	-1.67
CXCL12	-1.68
LIMA1	-1.69
CALD1	-1.70
CCPG1	-1.76
MAN2A1	-1.82
FAR2	-1.87
C12orf39	-1.95
IL1RAP	-2.01
ENPP1	-2.18

Chapter 3: Defining transcriptional regulatory modules for *let-7* pri-miRNAs

Abstract

The *let-7* family of miRNAs have been shown to control developmental timing in organisms from *C. elegans* to humans; their function in several essential cell processes throughout development is also well conserved. Numerous studies have defined several steps of post-transcriptional regulation of *let-7* production; from pri-miRNA through pre-miRNA, to the mature miRNA that targets endogenous mRNAs for degradation or translational inhibition. Less-well defined are modes of transcriptional regulation of the pri-miRNAs for *let-7*. *let-7* pri-miRNAs are expressed in polycistronic fashion, in long transcripts newly annotated based on chromatin-associated RNA-sequencing. Upon differentiation, we found some *let-7* pri-miRNAs are regulated at the transcriptional level, while others appear to be constitutively transcribed. Annotation of the regulatory elements of each polycistron identified putative promoters and enhancers. Probing these regulatory elements for transcription factor binding sites identified factors that regulate transcription of *let-7* in both promoter and enhancer regions.

Introduction

The *let-7* family of miRNAs were first identified in *C. elegans* as a single heterochronic factor controlling developmental timing[1]. Since then, this family of miRNAs has been shown to play somewhat equivalent roles in all bilaterian organisms, and the *let-7*s were the first miRNAs identified in humans[2]. The *let-7*s have now been implicated in differentiation and maturation of many tissues during development *in vivo* and *in vitro*[3–8]. As with other miRNAs, the initial pri-*let-7* transcripts are first transcribed by RNA polymerase II, then processed via the canonical pathway through the pre-miRNA stage generated by the action of Drosha/DGCR8[9–12]. The pre-miRNA is then processed in the cytoplasm by Dicer to generate the mature version of the miRNA[13]. In addition, in the case of *let-7* miRNAs, other processes such as uridylation are used to stabilize or destabilize miRNAs[14–16]. LIN28A and LIN28B are RNA binding proteins that regulate several of these processing steps to control levels of mature *let-7* transcripts[13,16–18]. Over evolution, *let-7* isoforms have expanded such that the human genome contains 9 isoforms. The study of regulation of the *let-7* family of miRNAs has focused on these processing steps, but less is understood about how the pri-*let-7* transcripts are regulated by transcription prior to any processing.

Studies in *C. elegans*, where the activity and expression of *let-7* is regionally and temporally constrained, have attempted to clarify transcriptional regulation from the single *let-7* locus. Two regulatory regions upstream of the locus were identified as the temporally regulated expression binding site (TREB) and the *let-7* transcription element (LTE), and many studies have tested the binding and transcriptional control exerted by several TFs including elt-1 and daf-12[19–23]. These sequences are not present upstream of mammalian *let-7* gene, and there are not similarly consistently present sequences near all the different *let-7* loci. In higher organisms, a different system for regulating *let-7* miRNA transcription must have been established.

The study of mammalian pri-*let-7* transcription is hampered by the relative scarcity of the transcript which is typically processed immediately and therefore difficult to detect. We previously took advantage of a method that allows for the capture of nascent RNA transcripts, which are still associated with the chromatin from which they are transcribed, to carefully annotate pri-*let-7* transcripts[6,24]. Another group later induced *pri-let*-7 accumulation in the context of DGCR8 knockout, and validated with RACE PCR that primary *let*-7 transcripts have multiple isoforms, some of which aligned nearly identically to our observed annotation patterns and varied in different cellular contexts[25]. From these annotations, it is clear that many *let-7* family members are transcribed within very long (up to 200KB), often polycistronic transcripts[25]. While some studies have identified transcriptional models of pri-miRNAs in higher organisms, the lack of proper annotation left the precise regulatory motifs for human *let-7* transcripts undefined[23,26–28]. Here, using precise annotation, we attempt to define regulatory motifs for the

let-7 family of miRNAs by taking advantage of Chromatin-associated RNA-seq and the latest genomic descriptions of chromatin states within *let-7* loci. We model *let-7* transcription in distinct neural paradigms to reveal subsets of *let-7* family members that are transcribed constitutively versus dynamically regulated in particular contexts. Finally, we identify transcription factors that appear to regulate *let-7* transcription by acting at either promoter or enhancer elements enriched in dynamically regulated *let-7* polycistrons.

Results

As a first step to determine how *let-7* miRNAs are transcriptionally regulated, we attempted to define developmental models that display dynamism of transcription. We previously identified dynamic transcriptional regulation of some *let-7* family members between neural progenitors that represent distinct developmental stages[6]. We also show here that as human pluripotent stem cells are specified to neural progenitors, and subsequently into neurons, some primary *let-7* transcripts are strongly induced as measured by RT-PCR. In both developmental scenarios, we observed that a subset of *let-7* family members showed transcriptional induction over developmental time, while other members appeared to be constitutively transcribed (Fig 1A and B). On the other hand, the levels of all mature *let-7* family members were strongly induced across development (Fig 1C). Using Chromatin-associated RNA-seq data, we not only observed the same dynamism of *let-7* transcription, but also mapping reads to the *let-7* loci highlighted the fact that some *let-7* transcripts are long and polycistronic (Fig 1D).

As further evidence for the long length of these transcripts, RT-PCR was performed using primers that recognize different regions of the predicted transcript from the *let-7a3/b* locus (Figure 3A). In addition, we posited that this transcript would accumulate in abundance if downstream processing by DGCR8 was inhibited. siRNA-mediated silencing of DGCR8 increased levels of the *let-7a3/b* transcript as measured by all the primers across the entire predicted polycistron. Silencing of DICER, necessary for the final step of miRNA processing, did not change the level of any portion of the *let-7a3/b* transcript (Figure 3B). As further evidence that *let-7* transcripts are polycistronic, the data in Fig 1A and B on dynamic versus constitutive indeed showed a shared pattern for those *let-7*s that are in the same polycistron. For instance, the pattern of *let-7a* and *let-7b* was conserved and dynamic in both contexts, while *let-7a1*, *let-7d* and *let-7f1,* which are also polycistronically transcribed, were constitutively expressed in both contexts.

We then sought to determine whether the dynamic versus constitutive *let-7* polycistrons display distinct regulatory schemes. Using data from the Epigenetic Roadmap, we annotated the chromatin states across each polycistronic *let-7* locus. The Roadmap database includes data from dozens of human cell types, including several of the neural lineage and pluripotent stem cells, both highly relevant to our current study[29]. Using these data and the imputed chromatin state model in tamed, we clearly identified transcriptional start sites, promoters (active and poised), enhancers, and actively transcribed regions for each of the *let-7* polycistrons (Fig 2A). As further evidence for their polycistronic nature, these epigenetic data again predicted single, long

transcripts across entire loci that encompass multiple *let-7* family members. Importantly, some of the genomic state models predicted variation of states in distinct cell types. Notably, the predicted promoter of *let-7a3/b* was shown to be poised in hPSCs and hPSC-derived NPCs, and active and transcribed in later neural derivatives and in brain[30]. This pattern is highly consistent with our own transcriptional data whereby the *let-7a3/b* polycistron was not transcribed significantly until hPSC-derived NPCs were driven further to neurons (Fig 1A).

Globally, the utility of these analyses was to define more precisely the location of promoters and enhancers for each of the *let-7* polycistrons. Taking advantage of the annotation of promoters, we attempted to identify mechanisms of transcriptional regulation of the dynamic versus constitutively regulated *let-7* polycistrons. With a focus on *let-7a3/b*, we searched for transcription factors that could regulate this polycistron through interactions at the promoter. We first used transcription factor ChIP-Seq data from the ENCODE and Roadmap datasets to detect transcription factor binding sites enriched in this promoter (Fig 3C)[29,31]. We then narrowed the list of candidates to include just those whose expression changes in contexts where *let-7a3/b* transcription also changes. This led to the identification of 10 TFs. To functionally determine whether any of these TFs can affect *let-7a3/b* transcription, we silenced some of them in tissue-NPCs (where transcription of *let-7a3/b* is high) and performed RT-PCR. Silencing of N-MYC, AP2a, or EGR1 all appeared to lead to an increase in *let-7a3/b* transcription after just two days, indicating a role for these TFs in transcriptional regulation of this polycistron.

To focus on potential regulatory mechanisms at enhancers, we next looked for putative enhancers by looking for regions of enriched DNAse-hypersensitivity, peaks of H3K27ac, and peaks associated with p300 binding in the *let-7a3/b* locus (Fig 4A). Several predicted enhancers, outlined in red rectangles, were highly DNAse sensitive region in Fetal and Adult brain samples but not in PSCs or PSC-derived NPCs. This pattern correlates with the timing of increased *pri-let7a3/b* transcription. A different site, 10kb upstream of the newly annotated TSS, is outlined in a green rectangle, and instead showed DNAse sensitivity only in PSC-derived NPCs, and not in either the undifferentiated or fully differentiated cells in the database. All of these identified regions showed P300 binding, were surrounded by ChIP-seq peaks for acetylated H3K27 in neural samples, and were significant for a specific depletion of histone-associated ChIP-seq binding peaks right at the site of DNAse sensitivity.

We used a similar approach to identify predicted and validated TF binding sites in the enhancer regions as on promoter sequences. This analysis yielded a strong enrichment of binding by the forkhead box transcription factors (FOX proteins), all of which can bind the same motif: 5'-[AC]A[AT]T[AG]TT[GT][AG][CT]T[CT]-3' [32]. Note the increased intensity of FOX protein ChIP-seq signal within the putative enhancers in Figure 4A. The furthest upstream such region, outlined in green, is shown in more detail, with both predicted and experimental FOX protein binding localizing to one highly conserved area (Figure 4B). The forkhead box TFs contain winged helix domains, which contribute to the pioneer transcription factor activity of the entire family and could alone be responsible for observed changes in chromatin accessibility[33]. We compared this

finding with expression data to subsequently determine which candidate forkhead box TFs could potentially be acting at the *let-7a3/b* enhancer during neural development (Fig 4C).

The forkhead box proteins FOXP2, FOXP1, FOXP4, FOXN2, FOXN3, FOXN4, and FOXG1 showed both high baseline expression and dynamic changes in expression over the course of nervous system development (Figure 4C). FOXP2's role in brain development is linked closely to its involvement in diseases of speech and language[34,35]. In the murine developing spinal cord and cortex, Foxp2 and Foxp4 are expressed in neural progenitor cells and increase in abundance during neuronal differentiation[36]. In Foxp4$^{-/-}$ mice, these NPCs fail to exit the progenitor stage and cause major disruptions in the developing neural tube. Foxg1 has been shown to act as a transcriptional repressor, and suppresses differentiation into an early neuron subtype in the cortex[37]. While no function in the nervous system has been ascribed to either FOXN2 or FOXN3, murine Foxn4 is expressed in the brain and retina, and is necessary for the specification of retinal amacrine cells[38]. Taken together, forkhead box proteins have the molecular components necessary to induce reorganizations of the epigenetic state, and some are expressed at anatomic locations and times that correlate with *let-7* expression.

siRNA-mediated silencing of FOXG1, which is normally induced over the same time course as *let-7a3/b,* showed an increase in the expression of several primary *let-7* miRNAs including *pri-let-7b*, but not pri-*let-7a3* (Figure 4D). Conversely, silencing FOXP2, which is normally suppressed over

the same developmental period, had little effect on transcription of the *let-7a3/b* polycistron, and showed induction of *pri-let-7a1* and *pri-let7f*, but a decrease in *pri-let-7g* and *pri-let-7i*.

Discussion

Together, these analyses define contexts in which particular *let-7* polycistrons are transcriptionally regulated, and identify TFs that play roles in this dynamism. This study is not the first to identify transcriptional mechanisms for *let-7* family members, but previous studies did not take advantage of genome-wide analyses to systematically define regulatory modules or transcription factors that regulate them[23,26]. The fact that *let-7* miRNAs can be dynamically regulated at the transcriptional level has only recently been appreciated, but the relative contribution of this regulation to levels of mature *let-7*s remains undefined. This is potentially an important issue to resolve as recent evidence suggests that not all *let-7* miRNAs are processed by the same machinery[39], and therefore, the level of mature *let-7* might not simply be DICER dependent.

These issues bring to light an interesting question, why have mammals evolved to have so many *let-7* isoforms in their genomes, and why do so in polycistronic fashion? Because all the *let-7* family members have the same seed sequence, it seems redundant to express so many. Even in the early neural lineage where mature *let-7*s are scarce, some of the *let-7* polycistrons are not transcribed, whereas others appear to be constitutively expressed. While we can only speculate,

it is possible that both dynamic and constitutive *let-7* transcription is a function of feed-back activity of *let-7*-target interactions. It is worth pointing out that some *let-7* targets also regulate *let-7* maturation, such as LIN28A, LIN28B and LIN41. Furthermore, it has been proposed that some *let-7* target RNAs can act as ceRNA or sponges of mature *let-7* to regulate their activity[8,40]. In addition, some of the TFs shown here and elsewhere to regulate *let-7* transcription (e.g. N-MYC) are also *let-7* target genes[41]. Perhaps, the constitutive transcription and maturation of small amounts of *let-7* serves as something of a rheostat of developmental timing that is tuned as cells become more specified, leading to changes in *let-7* targeted TFs that can then in turn regulate *let-7* transcription, leading to even more mature *let-7* through an additional feed-forward mechanism.

In *C. elegans*, where *let-7*s were first discovered, there is evidence for both transcriptional and maturation control despite the fact that all *let-7* is transcribed from a single locus. In fact, there are two distinct transcriptional start sites and these are distinctly regulated by both cis and trans mechanisms. There is further evidence that *let-7*s play a more general role in miRNA biogenesis through an interaction with Argonaute[23,42]. Therefore, sophisticated mechanisms for *let-7* regulation have been preserved and expanded across evolution, perhaps pointing to their critical roles in both developmental timing and tumorigenesis.

Materials and Methods

Cell Culture

Pluripotent stem cell culture and differentiation into NPCs and neurons was performed as previously described[6]. Briefly, PSCs were induced to differentiate along the neuroepithelial lineage by treatment with dual inhibitors of the SMAD signaling pathway, SB431542 (Sigma, 5 µM) and LDN193189 (Sigma, 50 nM). Neuroepithelial rosettes were manually picked and replated onto plates coated with ornithine and laminin. Cells were maintained and expanded in NPC media, containing DMEM/F-12 (Gibco), B27 supplement (Gibco), N2 (Gibco), EGF and bFGF. To induce differentiation, cells were fed with media lacking EGF and bFGF for 3 weeks. Tissue-derived NPCs were cultured and differentiated with the same reagents[6].

siRNA transfection

Gene knockdowns were performed by transfecting cells with double stranded 27mer RNAs (OriGene) using the lipofectamine RNAiMAX reagent (Thermo Fisher) according to the protocol provided.

Measurements of gene expression by RT-PCR

Cells were lysed in Trizol lysis reagent (Thermo Fisher), and total RNA was purified from lysates using the QIAgen miRNeasy kit. cDNA was made by reverse transcription from mRNAs with the SuperScript III First Strand Synthesis system (Thermo Fisher), or from miRNAs with the miScript

II RT Kit (Qiagen). Realtime PCR was performed on a Roche Lightcycler 480 instrument. For mRNA-derived cDNA, Roche 480 SYBR green I was used. For miRNA-derived cDNA, miScript SYBR (Qiagen) was used.

Epigenome characterization and candidate TF prediction

ENCODE and Roadmap gene expression data, ChIP-Seq mapping data, and ChromHMM chromatin state prediction were accessed and visualized using the UCSC genome browser and the WashU Epigenome Browser[43,44]. These tools were also used to import and visualize Chromatin-associated RNA-seq reads from Patterson *et al.* and miRNA gene transcripts in cells lacking DGCR8 from Chang *et al*[6,25].

Transcription factor binding site predictions were performed with the ORCA Toolkit web server, and with the MEME suite of motif analysis applications[45,46].

Figures

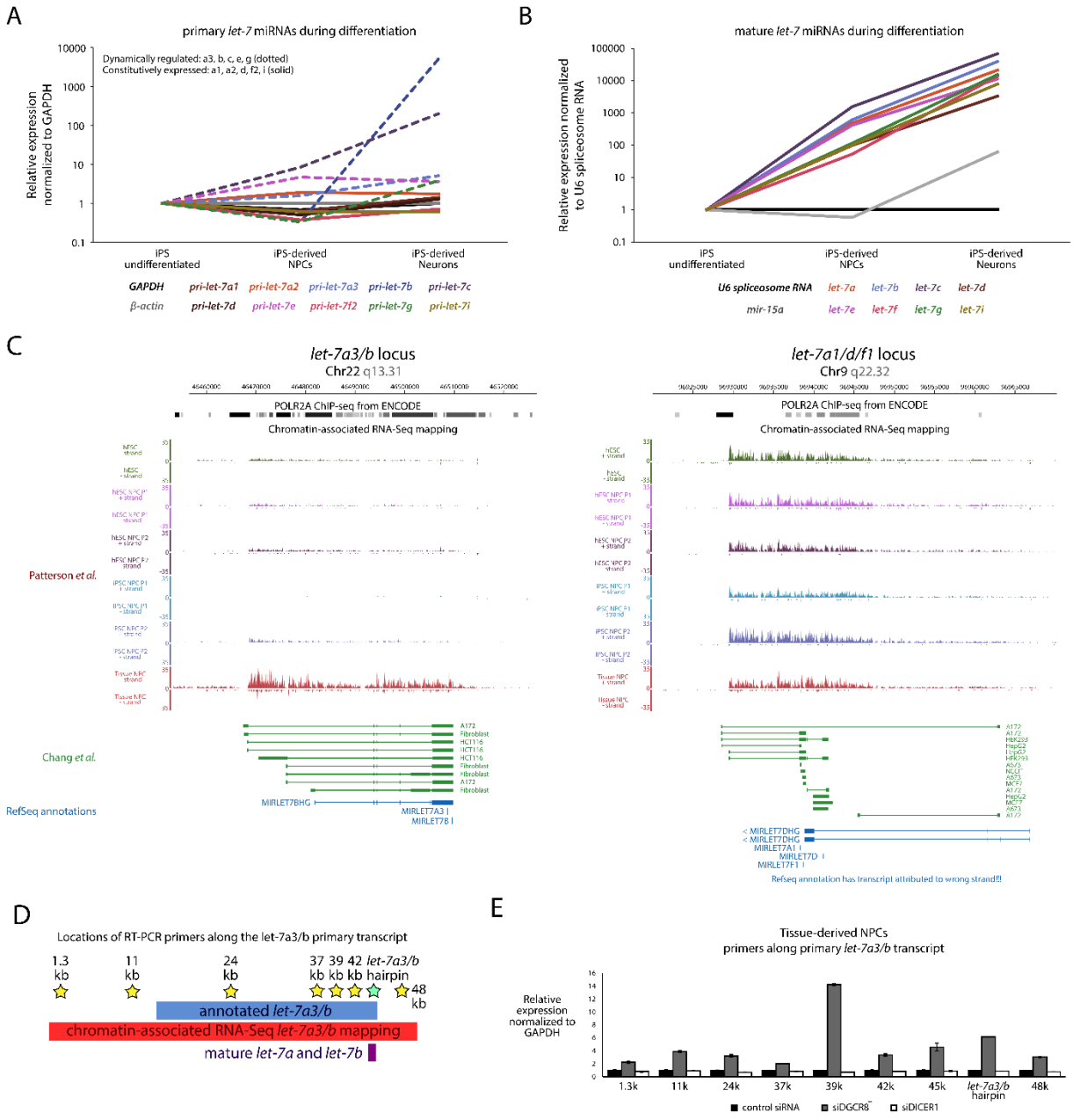


Figure 3-1: Dynamic transcriptional regulation of some *pri-let-7* transcripts

Pluripotent stem cells were differentiated through the neural lineage to neural progenitor cells (NPCs) and then to neurons. Using RT-PCR with primers specific to the *let-7* miRNAs at different

stages of processing, we tested changes in expression of the *pri-let-7*s **(A)** and their mature forms **(B)**. While all mature miRNAs increased over the course of differentiation, only a subset (marked with dotted lines), the dynamically regulated *let-7*s, also increased before processing, at the primary *let-7* stage. RT-PCRs were also performed beginning with ES cells. **(C)** Chromatin-associated RNA-seq reads were mapped onto two distinct polycistronic *let-7* loci. At left, the *let-7a3/b* locus, is dynamic, while at right, the *let-7a1/d/f1* locus, is constitutively expressed. Reads are shown for ESC, iPSC, PSC-derived NPC, and neural tissue-derived NPC stages. These reads are aligned with validated primary miRNA transcripts from RACE PCR experiments in green and RefSeq annotated genes in blue[25]. Note that Chromatin-associated RNA-seq and RACE PCR annotated transcripts demonstrate the existence of longer transcripts from different transcriptional start sites than suggested by the RefSeq annotation. In the case of *let-7a1/d/f1*, this discrepancy extends to the strand from which initial transcription occurs. **(D)** Graphic comparing the length of the RefSeq annotated *let-7a3/b* with our predicted transcript. Stars mark primer pairs for RT-PCR along the full transcript. **(E)** RT-PCR of *pri-let-7a3/b* transcript in tissue-derived NPCs, in which transcription is abundant. In control, siDGCR8 (to block Microprocessor function and pri-to-pre conversion), and siDICER (to block pre-to-mature conversion) conditions. When Microprocessor is disabled, the entire *let-7a3/b* transcript accumulates.

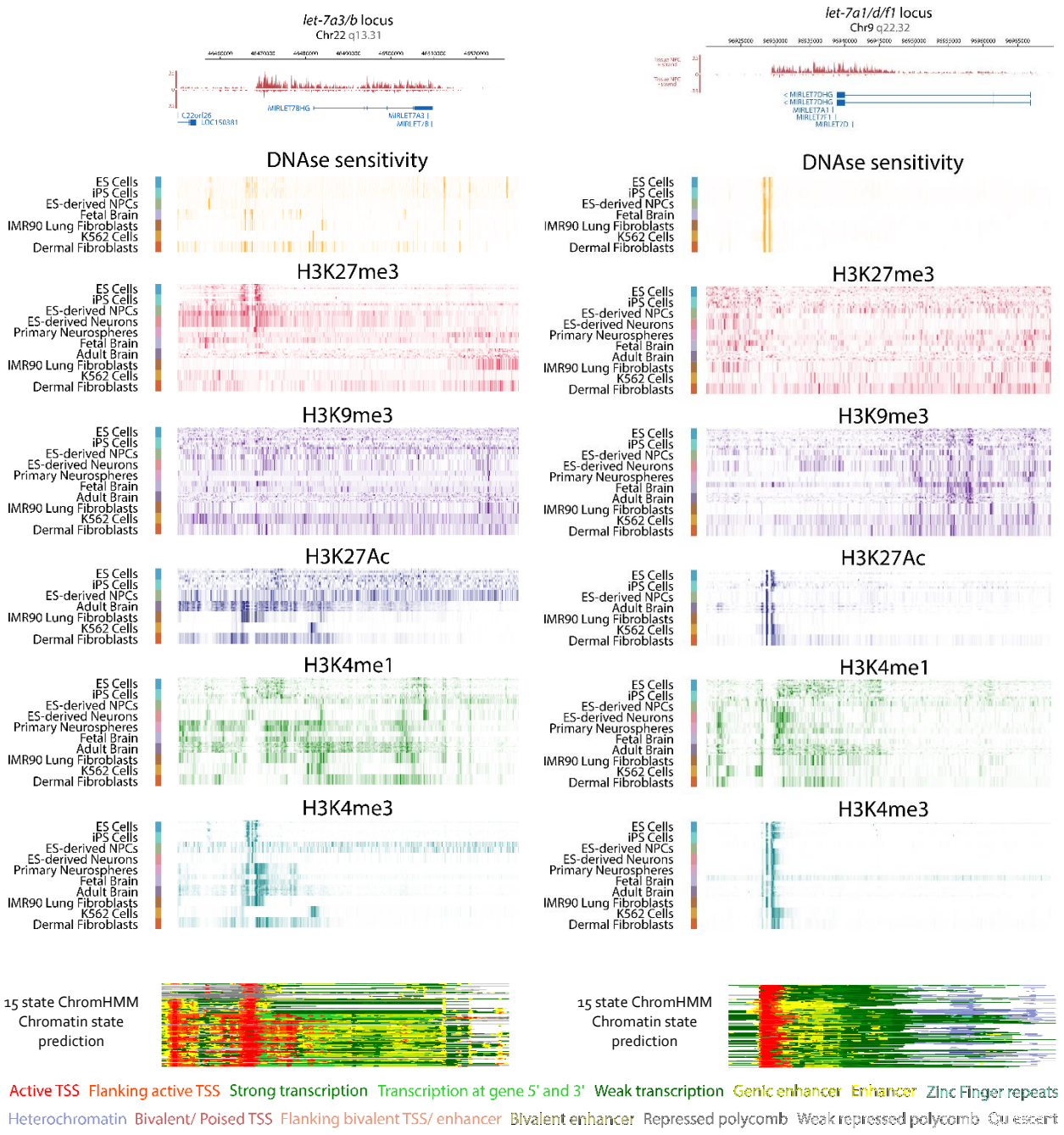


Figure 3-2: Dynamically and constitutively transcribed *let-7* loci show distinct epigenetic signatures

(A) Epigenetic marks from the Roadmap Epigenomics project at the dynamic and constitutive polycistronic loci. At top are the Chromatin-associated RNA-Seq peaks and RefSeq annotations

of the primary *let-7* transcripts, and below are the relative intensities of DNAse sensitivity or histone modification ChIP-Seq peaks at those loci. **(B)** Computationally imputed chromatin states generated by the ChromHMM algorithm at the same *let-7* loci. Each row represents one biological sample. These states show active transcriptional marks at the predicted TSS for *let-7a1/d/f1* in multiple cell types. At the *let-7a3/b* locus, ES cells, iPS cells, and PSC-derived NPCs have marks consistent with poised promoters, but later in differentiation active TSS marks appear at the same sites, reflecting changes in epigenetic state during neural differentiation. Epigenetic marks in K562 leukemia cells show active transcription at the RefSeq annotated *let-7a3/b* locus.

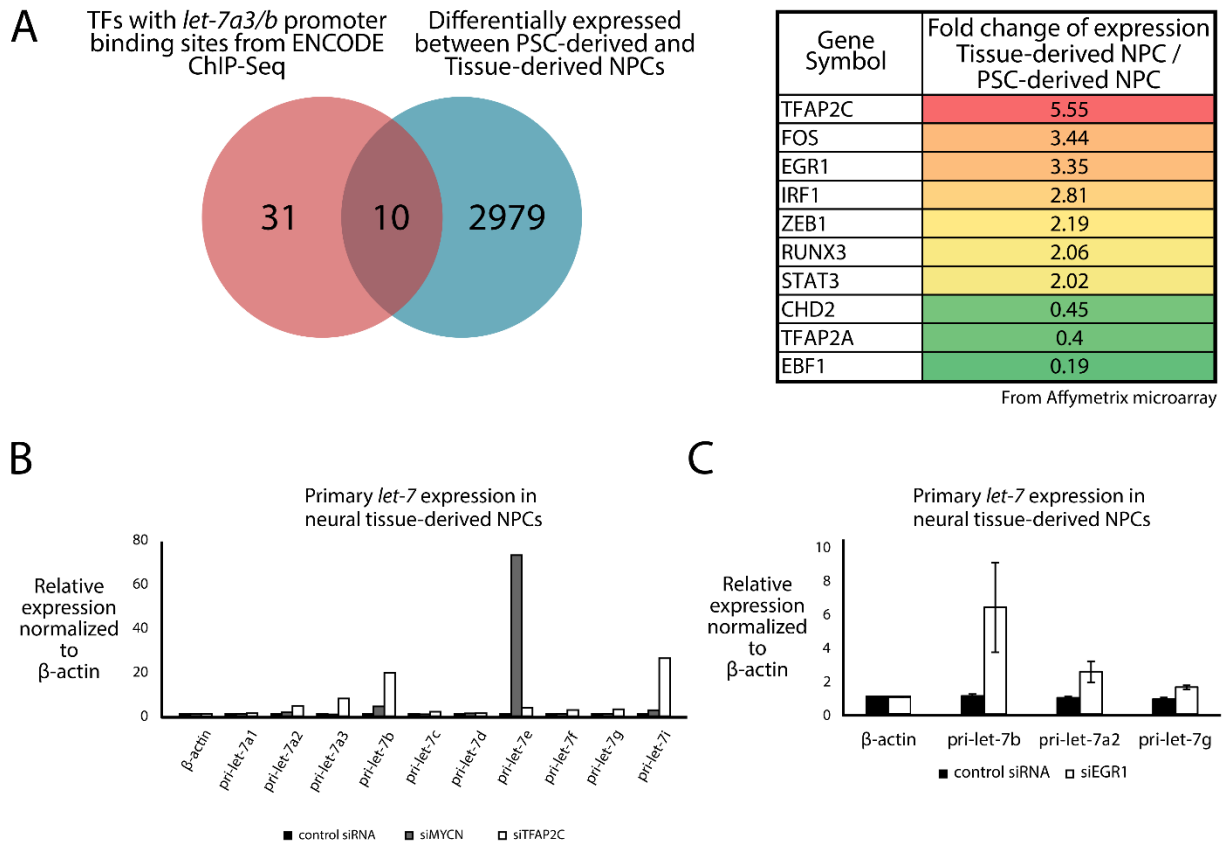
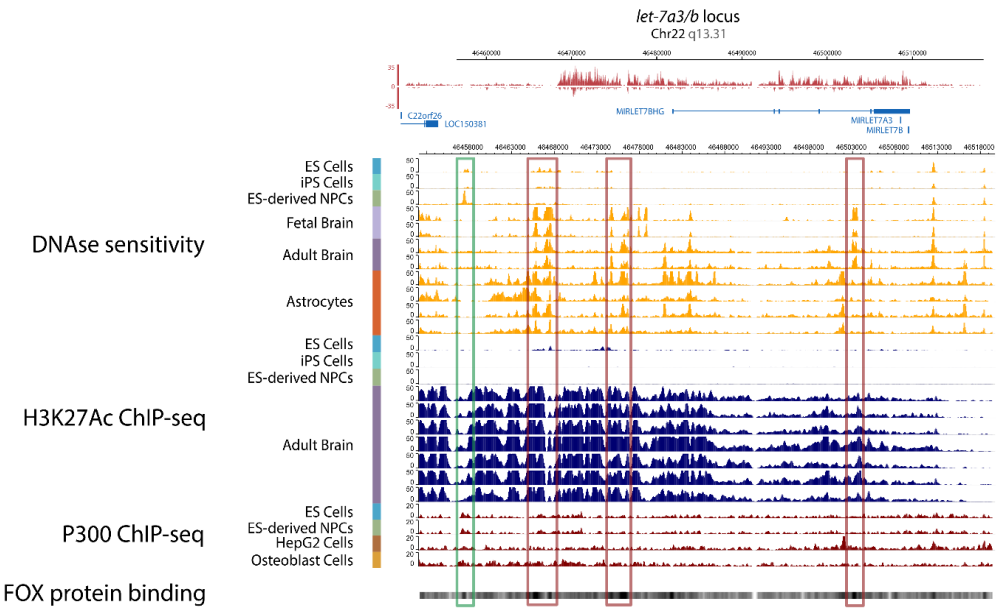


Figure 3-3: Transcription factors predicted to bind to the *let-7a3/b* promoter regulate primary *let-7a3/b* transcription

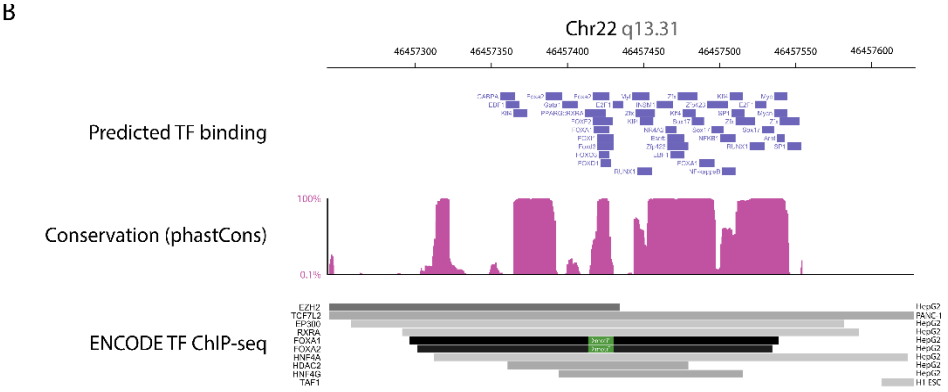
(A) Comparison of transcription factors with experimentally determined binding sites to the *bona fide let-7a3/b* promoter from the ENCODE database with genes differentially expressed between tissue-derived NPCs (in which *let-7a3/b* is abundantly transcribed) and PSC-derived NPCs (in which it is not). 10 genes were present in both sets, and are shown at right, ranked by their fold change of expression between tissue-derived and PSC-derived NPCs from microarray based gene expression measurements. We knocked down several of these candidate *let-7* regulator transcription factors in tissue-derived NPCs. **(B)** Knockdown of the *TFAP2C* gene encoding the AP-2γ protein, and of the *MYCN* gene increase transcription of several *let-7* genes. Data shown are

representative of 3 independent experiments. **(C)** Knockdown of *EGR1* increases transcription of primary *let-7b* and other *let-7* genes. Error bars are ± SEM from n=3 biological replicates.

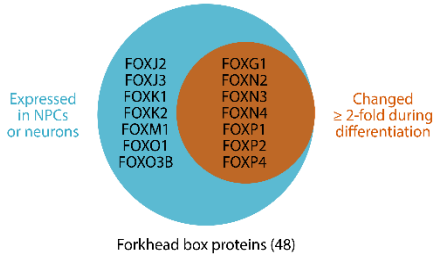
A


let-7a3/b locus
Chr22 q13.31
C22orf26
LOC150381
MIRLET7BHG
MIRLET7A3
MIRLET7B
DNAse sensitivity
ES Cells
iPS Cells
ES-derived NPCs
Fetal Brain
Adult Brain
Astrocytes
H3K27Ac ChIP-seq
P300 ChIP-seq
HepG2 Cells
Osteoblast Cells
FOX protein binding

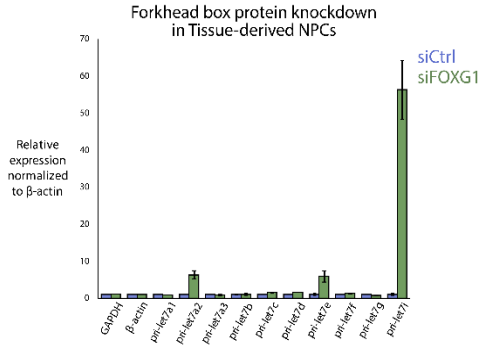
B


Chr22 q13.31
Predicted TF binding
Conservation (phastCons)
ENCODE TF ChIP-seq

C


Expressed in NPCs or neurons
FOXJ2
FOXJ3
FOXK1
FOXK2
FOXM1
FOXO1
FOXO3B
FOXG1
FOXN2
FOXN3
FOXN4
FOXP1
FOXP2
FOXP4
Changed ≥ 2-fold during differentiation
Forkhead box proteins (48)

D


Forkhead box protein knockdown in Tissue-derived NPCs
siCtrl
siFOXG1
Relative expression normalized to β-actin
GAPDH
β-actin
pri-let7a1
pri-let7a2
pri-let7a3
pri-let7b
pri-let7c
pri-let7d
pri-let7e
pri-let7f
pri-let7g
pri-let7i

E

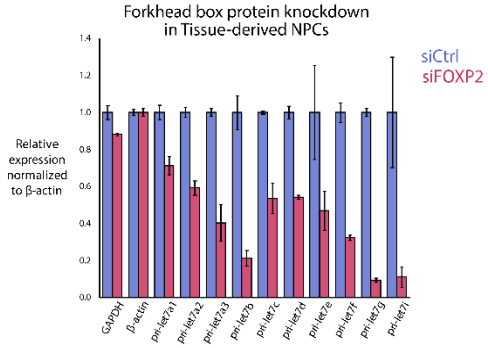

Forkhead box protein knockdown in Tissue-derived NPCs
siCtrl
siFOXP2
Relative expression normalized to β-actin
GAPDH
β-actin
pri-let7a1
pri-let7a2
pri-let7a3
pri-let7b
pri-let7c
pri-let7d
pri-let7e
pri-let7f
pri-let7g
pri-let7i

Figure 3-4: FOX proteins are predicted to bind to putative *let-7a3/b* enhancer regions, and affect *let-7* miRNA transcription

(A) The predicted existence of an upstream enhancer for the *let-7a3/b* locus was based on the epigenetic state at a region 10kb upstream of the TSS, outlined in green. In addition to being marked by H3K27Ac ChIP-Seq peaks with a localized dip in signal intensity, and peaks for the enhancer-associated histone acetyltransferase protein P300, this region showed dynamic changes in DNAse sensitivity. Note that a large DNAse sensitivity peak appears only in ES-derived NPCs, suggesting a differentiation state-specific chromatin opening at this region. At bottom, relative intensity of forkhead box protein ChIP-Seq from multiple cell types are pooled, with the darkest regions indicating intense FOX protein binding. Outlined in red are similar regions that show DNAse sensitivity beginning at the fetal brain stage that also colocalize with FOX protein binding. **(B)** A zoomed in view of the green region of increased DNAse sensitivity in PSC-derived NPCs. In blue are computationally predicted transcription factor binding sites from the ORCAtk database. The degree of genomic conservation along this region from the PhastCons64 database is shown in purple. At bottom are transcription factor ChIP-seq mapped peaks from the ENCODE database. The regions in green mark forkhead box transcription factor conserved motifs. Note that the forkhead box motifs co-localize with a region of highly conserved sequence, and the redundant binding of the forkhead box motif by many family members predicts that many such proteins can bind there. **(C)** There are 48 human forkhead box proteins, most of which share binding identity to the motif identified above. By filtering for FOX genes that are actively transcribed in our neural cells and differentially expressed between PSC-derived NPCs and Tissue-

derived NPCs, we generated a list of candidate proteins that might mediate changes in *let-7a3/b* transcription. **(D,E)** Knockdown of two candidate FOX proteins, FOXG1 **(D)** and FOXG1 **(E)**, in Tissue-derived NPCs show mild effects on primary *let-7* transcription, particularly *pri-let-7a1*, *pri-let-7e*, and *pri-let-7f*.

References

1. Reinhart, B. *et al.* The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **403,** 901–906 (2000).

2. Pasquinelli, A. E. *et al.* Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408,** 86–89 (2000).

3. Patterson, M. *et al.* Defining the nature of human pluripotent stem cell progeny. *Cell Res.* **22,** 178–193 (2011).

4. Gurtan, A. M. *et al.* Let-7 represses Nr6a1 and a mid-gestation developmental program in adult fibroblasts. *Genes Dev.* **27,** 941–954 (2013).

5. Madison, B. B. *et al.* Let-7 Represses Carcinogenesis and a Stem Cell Phenotype in the Intestine via Regulation of Hmga2. *PLOS Genet.* **11,** e1005408 (2015).

6. Patterson, M. *et al.* let-7 miRNAs can act through Notch to regulate human gliogenesis. *Stem Cell Rep.* **3,** 758–773 (2014).

7. Xia, X. & Ahmad, I. let-7 microRNA regulates neurogliogenesis in the mammalian retina through Hmga2. *Dev. Biol.* **410,** 70–85 (2016).

8. Ma, C. *et al.* H19 promotes pancreatic cancer metastasis by derepressing let-7's suppression on its target HMGA2-mediated EMT. *Tumor Biol.* **35,** 9163–9169 (2014).

9. Lee, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23,** 4051–4060 (2004).

10. Bracht, J. Trans-splicing and polyadenylation of let-7 microRNA primary transcripts. *RNA* **10,** 1586–1594 (2004).

11. Gregory, R. I. *et al.* The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432,** 235–240 (2004).

12. Lee, H., Han, S., Kwon, C. S. & Lee, D. Biogenesis and regulation of the let-7 miRNAs and their functional implications. *Protein Cell* (2015). doi:10.1007/s13238-015-0212-y

13. Piskounova, E. *et al.* Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms. *Cell* **147,** 1066–1079 (2011).

14. Heo, I. *et al.* Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. *Cell* (2012).

15. Thornton, J. E., Chang, H.-M., Piskounova, E. & Gregory, R. I. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA* **18,** 1875–1885 (2012).

16. Heo, I. *et al.* Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Mol. Cell* **32,** 276–284 (2008).

17. Viswanathan, S. R., Daley, G. Q. & Gregory, R. I. Selective blockade of microRNA processing by Lin28. *Science* **320,** 97–100 (2008).

18. Chang, H.-M., Triboulet, R., Thornton, J. E. & Gregory, R. I. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28–let-7 pathway. *Nature* (2013). doi:10.1038/nature12119

19. Johnson, S. M., Lin, S.-Y. & Slack, F. J. The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev. Biol.* **259,** 364–379 (2003).

20. Bethke, A., Fielenbach, N., Wang, Z., Mangelsdorf, D. J. & Antebi, A. Nuclear hormone receptor regulation of microRNAs controls developmental progression. *Science* **324,** 95–98 (2009).

21. Hammell, C. M., Karp, X. & Ambros, V. A feedback circuit involving let-7-family miRNAs and DAF-12 integrates environmental signals and developmental timing in Caenorhabditis elegans. *Proc. Natl. Acad. Sci.* **106,** 18668–18673 (2009).

22. Cohen, M. L., Kim, S., Morita, K., Kim, S. H. & Han, M. The GATA Factor elt-1 Regulates C. elegans Developmental Timing by Promoting Expression of the let-7 Family MicroRNAs. *PLoS Genet.* **11,** e1005099–e1005099 (2015).

23. Kai, Z. S., Finnegan, E. F., Huang, S. & Pasquinelli, A. E. Multiple cis-elements and trans-acting factors regulate dynamic spatio-temporal transcription of let-7 in Caenorhabditis elegans. *Dev. Biol.* **374,** 223–233 (2013).

24. Bhatt, D. M. *et al.* Transcript Dynamics of Proinflammatory Genes Revealed by Sequence Analysis of Subcellular RNA Fractions. *Cell* **150,** 279–290 (2012).

25. Chang, T.-C., Pertea, M., Lee, S., Salzberg, S. L. & Mendell, J. T. Genome-wide annotation of microRNA primary transcript structures reveals novel regulatory mechanisms. *Genome Res.* **25,** 1401–1409 (2015).

26. Wang, D. J., Legesse-Miller, A., Johnson, E. L. & Coller, H. A. Regulation of the let-7a-3 promoter by NF-κB. *PloS One* **7,** e31240 (2012).

27. Van Wynsberghe, P. M. & Pasquinelli, A. E. Period homolog LIN-42 regulates miRNA transcription to impact developmental timing. *Worm* **3,** e974453 (2014).

28. Van Wynsberghe, P. M. *et al.* LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans. *Nat. Struct. Mol. Biol.* **18,** 302–308 (2011).

29. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518,** 317–330 (2015).

30. Ernst, J. & Kellis, M. Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nat. Biotechnol.* **33,** 364–376 (2015).

31. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2012).

32. Pierrou, S., Hellqvist, M., Samuelsson, L., Enerbäck, S. & Carlsson, P. Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending. *EMBO J.* **13,** 5002 (1994).

33. Cirillo, L. A. *et al.* Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9,** 279–289 (2002).

34. Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413,** 519–523 (2001).

35. Enard, W. *et al.* Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* **418,** 869–872 (2002).

36. Rousso, D. L. *et al.* Foxp-Mediated Suppression of N-Cadherin Regulates Neuroepithelial Character and Progenitor Maintenance in the CNS. *Neuron* **74,** 314–330 (2012).

37. Hanashima, C. Foxg1 Suppresses Early Cortical Cell Fate. *Science* **303,** 56–59 (2004).

38. Li, S. *et al.* Foxn4 controls the genesis of amacrine and horizontal cells by retinal progenitors. *Neuron* **43,** 795–807 (2004).

39. Triboulet, R., Pirouz, M. & Gregory, R. I. A Single Let-7 MicroRNA Bypasses LIN28-Mediated Repression. *Cell Rep.* (2015). doi:10.1016/j.celrep.2015.08.086

40. Kallen, A. N. *et al.* The Imprinted H19 LncRNA Antagonizes Let-7 MicroRNAs. *Mol. Cell* **52,** 101–112 (2013).

41. Chang, T.-C. *et al.* Lin-28B transactivation is necessary for Myc-mediated let-7 repression and proliferation. *Proc. Natl. Acad. Sci.* **106,** 3384–3389 (2009).

42. Zisoulis, D. G., Kai, Z. S., Chang, R. K. & Pasquinelli, A. E. Autoregulation of microRNA biogenesis by let-7 and Argonaute. *Nature* (2012). doi:10.1038/nature11134

43. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).

44. Zhou, X. *et al.* The Human Epigenome Browser at Washington University. *Nat Meth* **8,** 989–990 (2011).

45. Portales-Casamar, E. *et al.* The PAZAR database of gene regulatory information coupled to the ORCA toolkit for the study of regulatory sequences. *Nucleic Acids Res.* **37,** D54–D60 (2009).

46. Bailey, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37,** W202–W208 (2009).

Chapter 4*: In vitro* generation of Human PSC-derived Interneurons and progress toward elucidating strategies and mechanisms for their maturation

Introduction

Our previous efforts to understand the mechanisms of progenitor cell maturity yielded changes in gene expression and differentiation, but we were unable to discern functional differences between the differentiated cells, in this case neurons and glia, produced from those experiments. This is a key issue as neurons and glia do most of the work in the brain, whereas progenitors just give rise to these cells. The neurons described in earlier chapters were differentiated by growth factor withdrawal which created mostly inhibitory interneurons (Ohashi *et al.*, in prep). However, despite prolonged maintenance in *in vitro* culture, these interneurons never attained electrophysiological maturity. These findings led us to question how cell culture approaches might be changed to drive the maturation of fully specified and differentiated cell types. While other groups have used prolonged culture and co-culture to attempt to clear this maturation hurdle, the mechanisms behind interneuron maturation remains unclear[1,2].

Interneurons were initially described and recognized for their short projections in the early 20th century, and were later functionally characterized as providing inhibitory input on neural circuits[3–5]. More than a century ago it was also recognized that interneurons comprise a much higher percentage of CNS neurons in humans and primates than in lower animals[3]. The major neurotransmitter secreted by interneurons is γ-amino butyric acid (GABA), which exerts primarily inhibitory effects on postsynaptic neurons by triggering the influx of Cl^- ions through ligand-gated GABA A receptors[5]. Chloride anion influx hyperpolarizes postsynaptic neurons when extracellular

Cl^- concentrations are higher than intracellular Cl^- concentrations, as is the case in the adult mammalian brain, but not during early brain development[6,7].

The vast majority of the inhibitory interneurons that will eventually populate the adult brain are generated from telencephalic structures known as the ganglionic eminences during development. The medial ganglionic eminence (MGE) generates a population of NKX2.1 expressing progenitor cells, whereas the caudal ganglionic eminence (CGE) generates progenitors expressing DLX1/2 and GSH2 [4,8,9]. Each of these populations is biased toward generating subsets of interneurons that will eventually migrate and populate the cortex, comprising up to 20% of the neurons in the adult human cortex[3,10].

In the developing human brain, the initial signatures of interneuron differentiation begin to appear soon after the onset of neurogenesis. Figure 1 shows RNA sequencing-based gene expression data from the developing human brain, filtered for markers of the specification, differentiation, and specialization of CNS interneurons (Allen Brain Atlas). By grouping mapped RNA-Seq reads by brain region and condensing data from all time points together, we can pinpoint the brain locations at which interneuron markers are overrepresented (Fig 1A). For example, transcription factors of the DLX, NKX, and LHX families, which mark progenitors of CNS interneurons, were expressed more highly in the lateral ganglionic eminence (LGE), medial ganglionic eminence (MGE), caudal ganglionic eminence (CGE), and striatum relative to in the rest of the brain. Figure 1B reorganizes these data to show the depletion of these progenitor

stage transcription factors during the second and third trimester specifically in those brain regions (marked with asterisks in 1A). Lastly, Figure 1C summarizes expression of interneuron subtype markers throughout the entire brain across human development and adult life. While some subtype markers are already expressed by the initial 8 week gestation time point captured in this database, the genes *GAD1*, *PVALB*, *VIP*, and *CCK* all show increased transcription during the third trimester of fetal development and beyond. Together, these data describe a time course of CNS interneuron specification and subtype commitment that may serve as a guide for *in vitro* differentiation of interneurons from PSCs.

Crucially, the increased abundance of interneuron subtype markers in aggregate in the developing brain does not imply that their expression is co-regulated or that they are rising across the entire population. In fact, several of these markers are mutually exclusive, not co-expressed. Approaches to cataloguing the true diversity of interneurons necessitate a more sensitive method for detecting rare cells and distinguishing them from other, similar cells. One such approach has been measurements of gene expression at the single cell level[11,12]. While so far this approach has primarily been used to identify different types of cells present in a population, the same approach might also be used to distinguish functional subpopulations by expression of non-marker genes.

In this chapter, I discuss preliminary efforts toward generating human PSC-derived interneurons *in vitro* from several distinct pluripotent stem cell lines, and describe efforts to drive those cells

toward functional maturation *in vitro* and *in vivo*. I then present initial characterization of those changes at a molecular level, with the eventual goal of improving the differentiation potential and electrophysiological maturity of human PSC-derived NPCs

Results

In vitro differentiation of human PSC-derived inhibitory interneurons

Three different pluripotent stem cell lines were used over the course of these studies, all of which were transgenically modified to express fluorescent reporter proteins. The xfiPSC GFP line, an induced pluripotent stem cell line generated free of xenobiotics, constitutively expressed a green fluorescent protein[13]. The NKX2.1::GFP line was modified from the HES3 human ES cell line to express GFP under the control of the NKX2.1 promoter[14]. The LHX6::citrine line was modified from the H9 human ES cell line to express the citrine fluorescent protein under control of the LHX6 promoter.

Mammalian MGE-derived interneuron progenitors are specifically marked by NKX2.1 expression, and inhibitory interneurons generated from this region and the CGE express LHX6[5,14]. For these reasons, repeating the interneuron differentiation protocol with these three pluripotent stem cell lines demonstrated the fidelity of the differentiation protocol and additionally validated progress toward generating *bona fide* interneurons at intermediate steps of *in vitro* differentiation.

Figure 2A briefly describes the *in vitro* differentiation protocol used in this study to generate interneurons from PSC. Initial specification of neuroepithelial differentiation relies on culture with 3 small molecule inhibitors of developmentally relevant signaling pathways, followed by stimulation of the Sonic Hedgehog (SHH) signaling pathway with recombinant SHH and an agonist of SMO, the human Smoothened receptor. These growth factors are briefly withdrawn from the culture media, and then trophic growth factors are added to the final differentiation cocktail. The reporter lines were tested for epifluorescence over the course of this differentiation protocol, which validated that the newly generated neural progenitors drove activity of the *NKX2.1* promoter and that some cells retained this activity after more than a month of culture in differentiation media (Figure 2B). By 8 weeks of culture, staining for MAP2, a microtubule-associated protein specifically expressed in neurons, demonstrated that many of the cells in culture had terminally differentiated, and a staining for NKX2.1 validated activation of the promoter-driven reporter alongside expression of the *bona fide* transcription factor (Figure 2C).

The citrine fluorescent reporter driven by the *LHX6* promoter, though less bright than the *NKX2.1* reporter, also demonstrated activation during the differentiation phase of the protocol (Figure 2D). LHX6+ cells were also positive for MAP2, and subsequent staining with NKX2.1 also confirmed some prolonged expression of the progenitor-state transcription factor even after successful interneuron differentiation (Figure 2E). Together, activity of these reporters *in vitro*

suggested that our differentiation protocol was able to induce PSCs to differentiate into functional MGE interneuron progenitors and CNS interneurons.

A more specific characterization of the interneurons generated from PSCs revealed subtype heterogeneity among NKX2.1::GFP interneurons (Figure 2F). By 8 weeks of differentiation, the cultured neurons were broadly positive for GABA and components of the GABA synthesis pathway, GAD65/67. Direct quantification of these cells revealed that 30% of cells are GABA+, 41% of cells are MAP2+, and 93% of cells expressed GAD65/67. Many GAD65/67+ cells were observed that lacked NKX2.1 driven GFP expression. Software-based estimates were performed to quantify the interneuron subtype populations (Figure 2H). FOXG1+ progenitors were undetectable, and only rare RELN+, SST+, or NPY+ interneurons could be identified. Of particular note is that several RELN+ but GFP- cells were detected, suggesting that our protocol yields some non-MGE specific interneurons. A subset of neurons expressed Calretinin (CALB2), a calcium-binding protein common in cortical interneurons, but only rare cells expressed Parvalbumin (PVALB), which characterizes the fast-spiking interneurons most commonly generated from progenitors in the MGE[4,5]. Similar characterization was performed with the xfiPSC GFP line, though without detailed quantification.

In such a heterogeneous population of cells, measuring gene expression of markers specific for certain subtypes of interneurons would be impossible at the population level. In order to faithfully represent the variability of the interneurons generated by this protocol, we performed

single-cell RNA sequencing on dissociated human PSC-derived interneurons (Fig 3). Using this method, we were able to capture and identify different cells within a population of interneurons (*e.g.* a single SST+ cell, the co-expression of SOX2 and DCX).

Electrophysiological characterization of PSC-derived interneurons

Multiple attempts to measure the electrophysiological properties of our human PSC-derived interneurons consistently failed to demonstrate any spontaneously occurring or evoked excitatory potentials within those cells. We hypothesized this might be an effect of the developing interneurons not receiving signals from excitatory cells, so we performed *in vitro co*-culture experiments using murine cortical excitatory neurons and human PSC-derived interneurons. Figure 4A shows images and voltage tracings from two distinct PSC-derived interneurons after 10 days of co-culture. Although the cells demonstrate different degrees of excitability in response to injected current, both of them show evoked depolarizations. One neuron in particular was able to fire at a rate of 20 Hz, which is still slower than predicted for mature interneurons.

As a positive control, we tested the electrophysiological properties of the mouse striatum, in which interneurons are plentiful. Many interneurons demonstrated spontaneous and evoked action potentials at a rate near 50 Hz. This fast spiking is a characteristic of parvalbumin-expressing interneurons, and suggests a target for electrophysiological maturity in PSC-derived interneurons[4].

PSC-derived interneurons injected into the mouse striatum

If *in vitro* co-culture with a specified population of neurons induced maturation in PSC-derived interneurons, embedding those cells in the milieu of a brain already containing a broad diversity of cell types might also drive them toward maturity. We injected LHX6::citrine PSC-derived interneurons at one month of differentiation bilaterally into the striata of immunodeficient adult mice (Figure 5A). One month later, we examined the brains of these mice for the presence of human cells. One hemisphere of the brain was frozen, sliced, and stained using immunofluorescence, which allowed us to detect human cells migrating locally into brain regions nearby the injection site (Figure 5B). The other hemisphere was broken up and enzymatically dissociated, and then individual cells expressing citrine were isolated with FACS. We set our fluorescence intensity thresholds based on citrine-expressing cells cultured *in vitro* (Figure 5C), and were able to identify and collect live cells from the mouse brain in the area directly surrounding the original injection site (Figure D). We were unable to detect any citrine+ cells from the cerebellum, a region distant from the injection site and which does not receive migratory interneurons from the MGE (Figure 5E).

Discussion

By leveraging multiple lines of human pluripotent stem cells expressing markers of interneurons during development, we have been able to generate inhibitory interneurons *in vitro*. We characterized these interneurons by marker expression and gene expression to identify and

quantify the different subtypes of interneurons generated by this protocol. The types of interneurons we made were heterogeneous, more so than would be predicted to developmentally arise from just the MGE[4,15]. This suggests that even in the controlled environment of the tissue culture dish, different local concentrations of morphogens and growth factors can give rise to variations in differentiated cells. Characterizing these cells at later time points of differentiation will improve our understanding of the temporal dynamics of interneuron specification.

Despite expressing markers of committed interneuron differentiation, the PSC-derived interneurons we generate *in vitro* are not able to normally respond to electrical signals. Only through co-culture with excitatory cells were we able to induce some electrophysiological maturity, as other groups have also observed with PSC-derived interneurons[1,2]. Even so, these cells still do not display a signature consistent with mature interneuron differentiation. Further exploration of the co-culture effects, and supplementation with media made for the purpose of electrophysiologically maturing cells *in vitro* will help us explore and improve this process[16]. With this approach set up and our ability to high-throughput perform single cell RNA sequencing, we will then be able to profile the changes in gene expression that occur within individual cells as they reach increasing electrophysiological maturity over a wider range of time points and culture conditions.

Early insights from our injection of human PSC-derived interneurons into the mouse brain include findings that cells survive, migrate, and can be recovered from within a pool of mouse cells. We are continuing to gather data from later time points after injection from these experiments. We also plan to leverage these experiments to test the electrophysiological maturation of injected cells *in situ* once they have been integrated into the mouse brain. Furthermore, we can also look for changes in gene expression in PSC-derived interneurons over a prolonged course of engraftment into the mouse brain by sorting them and using single-cell RNA sequencing. By querying which gene expression changes occur in this setting, we can hypothesize about ways to more reliably induce maturation during *in vitro* differentiation.

The maturation of PSC-derived interneurons is different in many respects from the maturation of neural progenitor cells we have characterized before. While each cell type functionally changes, whether in differentiation potential or in terms of electrical properties, during maturation, we know that even in immature human interneurons there is already high expression of miRNA factors responsible for the maturation of progenitor cells[17,18]. Our experiments described here suggest that receiving signals from other types of neurons nearby, and perhaps the subsequent changes in gene expression in the cells receiving the signals, might be the driving mechanism behind functional interneuron maturation. We now have established tools to identify these changes, and can test their effects on interneuron gene expression, electrophysiology, and marker expression in a candidate driven approach. Together, we can broaden our understanding of neural mechanisms of cellular maturation.

Materials & Methods

In vitro differentiation of interneurons from pluripotent stem cells

On day 0, Pluripotent stem cells were dissociated with collagenase and triturated gently, then plated onto a matrigel (BD) coated plate in mTESR1 ES media containing 10 µM Y-27632, a rho kinase inhibitor (Stem Cell Technologies). Cells were plated to an 80% confluence density as observed on day 1.

Beginning on day 1 and continuing for 2 more days, cells were fed daily with NIMX media, containing DMEM/F-12 media (Thermo Fisher), B27 supplement with RA (Thermo Fisher), N2 supplement (Thermo Fisher), 2 mM Glutamax (Thermo Fisher), MEM NEAA (Thermo Fisher), 110 µM 2-mercaptoethanol, 300 pg/mL BSA, 5 µM SB431542 (Stem Cell Technologies), 2 µM XAV939 (Stem Cell Technologies), 50 nM LDN193189 (Stem Cell Technologies), Penicillin/Streptomycin (Thermo Fisher), and Primocin (Invivogen).

From Day 4 onward, NIMX media was tapered down with N2 media, containing DMEM/F-12 media with N2 supplement, 55 µM 2-mercaptoethanol, 1.5 mg/mL Dextrose, Penicillin/Streptomycin, and Primocin. Cells received 75% NIMX and 25% N2 on day 5, 50% each on day 5, 25% NIMX and 75% N2 on day 6.

From day 7-12, cells were fed every other day with P/S media, containing DMEM/F-12 media with B27 supplement with RA, N2 supplement, 55 µM 2-mercaptoethanol, 1.5 mg/mL Dextrose, 100 ng/mL recombinant mouse Shh (Stem Cell Technologies), 1 µM purmorphamine (Stem Cell Technologies), Penicillin/Streptomycin, and Primocin. On Day 10 cells were dissociated with 20%

TrypLE and replated at a ratio of 1:1onto freshly matrigel coated plates in P/S media containing 10 μM Y-27632.

From day 13-18, cells were fed daily with N2/B27 media containing DMEM/F-12 media with B27 supplement with RA, N2 supplement, 55 μM 2-mercaptoethanol, 1.5 mg/mL Dextrose, Penicillin/Streptomycin, and Primocin. On Day 18 cells were dissociated with 30% TrypLE at a ratio of 1:1.5 or 1:2 onto new matrigel coated plates in B27/N2 media containing 10 μM Y-27632.

Beginning on day 20 and continuing thereafter cells were fed with NBND media containing Neurobasal media (Thermo Fisher), B27 supplement with RA, N2 supplement, 2 mM Glutamax, 100 μM dibutyryl cAMP (Sigma), 55 μM 2-mercaptoethanol, 300 pg/mL BSA,200 μM Ascorbic Acid (Sigma), 10 ng/mL BDNF (Stem Cell Technologies), 10 ng/mL GDNF (Stem Cell Technologies), 10 ng/mL Neurotrophin-3 (Stem Cell Technologies), Penicillin/Streptomycin, and Primocin. Half of the media was changed every other day.

For co-culture experiments, frozen mouse cortical neurons (Thermo Fisher) were thawed and plated onto matrigel coated plates at an estimated confluence of 40% and fed with NBND media. The next day, dissociated human PSC-derived interneurons were plated atop the mouse neurons, and the mixed population continued to receive NBND media.

Fluorescence-Activated Cell Sorting of interneurons

Interneurons were dissociated using 20% TrypLE (Thermo Fisher) at 37° C for 5 minutes, were resuspended in PBS with 1% Bovine Calf Serum, and were filtered through a 40 μm mesh filter.

0.1 µg/mL DAPI (Sigma) was added to the resuspended cells at least 15 minutes prior to sorting in order to stain dead cells. Resuspended cells were stored on ice until sorting.

Cells were sorted on a BD FACSAria II sorter using a 70 µm aperture. Selected populations (usually $DAPI^{-}GFP^{+}$ or $DAPI^{-}GFP^{-}$) were sorted either into collection tubes containing NBND media or into individual wells of a cooled 60-well terasaki plate filled with RNA lysis buffer.

In experiments in which cells from the mouse brain were sorted, mouse brains were extracted and dissociated. Briefly, mice were anesthetized with isoflurane and decapitated following UCLA Chancellor's Animal Research Committee protocol. The mouse brain was coarsely chopped using a sterile razor, then collected in Hibernate-A media (Thermo Fisher) and incubated at 37° C for 1 hour with Papain and DNAse I to dissociate cells from the brain. Cells were then more fully dissociated with mild trituration, filtered through a 40 µm mesh filter, spun down, and resuspended in 0.1% BCS with 0.1 µg/mL DAPI for sorting.

Single-cell RNA sequencing

Dissociated interneurons were either sorted into individual wells of lysis buffer by FACS or were automatically separated and lysed using the Fluidigidm C_1 instrument. RNA was converted to cDNA using the SMARTer low input kit (Clontech), and sequencing libraries were generated using the Nextera kit (Illumina). During this library step, DNA from individual cells was barcoded and combined. Sequencing was performed using an Illumina HiSeq system, and reads were demultiplexed, processed, mapped, and quantified.

Electrophysiology

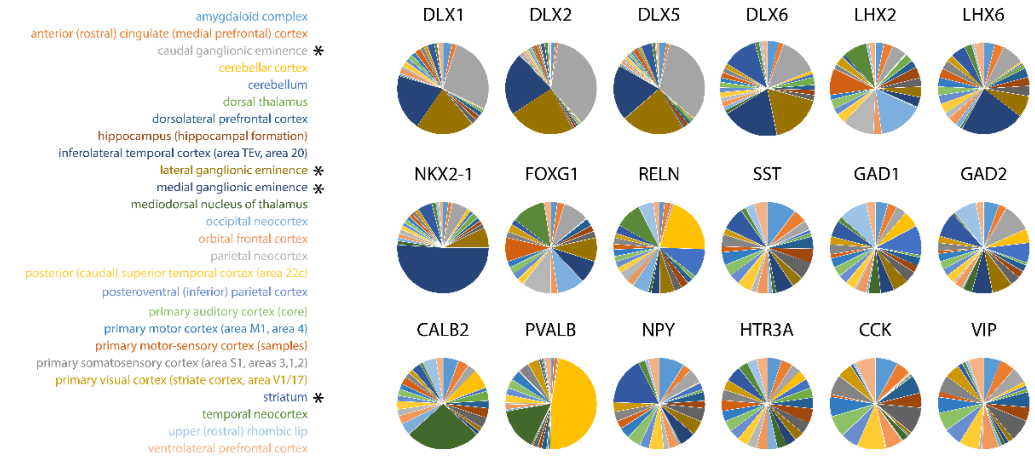
Interneurons were cultures on matrigel-coated glass coverslips, which were bathed in artificial CSF containing 126 mM NaCl, 10 mM D-glucose, 2.5 mM KCl, 1.25 mM NaH_2PO_4, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 26 mM $NaHCO_3$, 1.5 mM Na pyruvate, and 1 mM L-Glutamine as previously described[19]. Recordings were performed as previously described, with a patch pipette under control of a motorized stage, connected to an electrophysiology amplifier setup. Current clamp was performed, maintaining current at 0 with 500 ms pulses of current beginning with 200 pA and increasing 200 pA per step, and cellular voltage was measured.

Intracranial injection of human PSC-derived interneurons

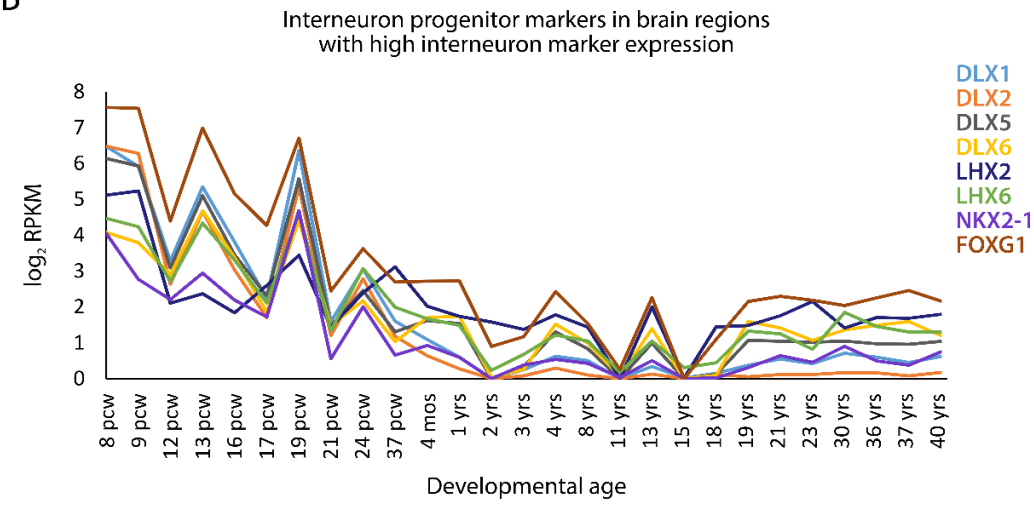
Dissociated interneurons were resuspended in Isolyte media at a concentration of 1 million cells per 30 µL. Under isoflurane anesthesia, Adult NSG mice received bilateral injections of 100,000 cells via stereotactic injection at the coordinates ± 3mm lateral, 0.95 mm posterior, and 2.6 mm inferior from the bregma at a controlled flow rate.

Figures

A


Relative expression of Interneuron markers across brain regions
(all timepoints averaged)
amygdaloid complex
anterior (rostral) cingulate (medial prefrontal) cortex
caudal ganglionic eminence *
cerebellar cortex
cerebellum
dorsal thalamus
dorsolateral prefrontal cortex
hippocampus (hippocampal formation)
inferolateral temporal cortex (area TEv, area 20)
lateral ganglionic eminence *
medial ganglionic eminence *
mediodorsal nucleus of thalamus
occipital neocortex
orbital frontal cortex
parietal neocortex
posterior (caudal) superior temporal cortex (area 22c)
posteroventral (inferior) parietal cortex
primary auditory cortex (core)
primary motor cortex (area M1, area 4)
primary motor-sensory cortex (samples)
primary somatosensory cortex (area S1, areas 3,1,2)
primary visual cortex (striate cortex, area V1/17)
striatum *
temporal neocortex
upper (rostral) rhombic lip
ventrolateral prefrontal cortex
DLX1
DLX2
DLX5
DLX6
LHX2
LHX6
NKX2-1
FOXG1
RELN
SST
GAD1
GAD2
CALB2
PVALB
NPY
HTR3A
CCK
VIP

B


Interneuron progenitor markers in brain regions
with high interneuron marker expression
log2 RPKM
0
1
2
3
4
5
6
7
8
8 pcw
9 pcw
12 pcw
13 pcw
16 pcw
17 pcw
19 pcw
21 pcw
24 pcw
37 pcw
4 mos
1 yrs
2 yrs
3 yrs
4 yrs
8 yrs
11 yrs
13 yrs
15 yrs
18 yrs
19 yrs
21 yrs
23 yrs
30 yrs
36 yrs
37 yrs
40 yrs
Developmental age
DLX1
DLX2
DLX5
DLX6
LHX2
LHX6
NKX2-1
FOXG1

C

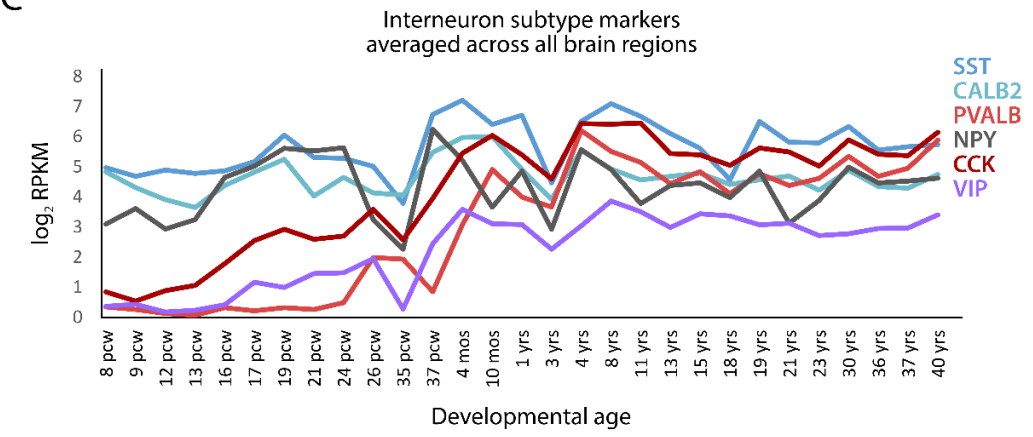

Interneuron subtype markers
averaged across all brain regions
log2 RPKM
0
1
2
3
4
5
6
7
8
8 pcw
9 pcw
12 pcw
13 pcw
16 pcw
17 pcw
19 pcw
21 pcw
24 pcw
26 pcw
35 pcw
37 pcw
4 mos
10 mos
1 yrs
3 yrs
4 yrs
8 yrs
11 yrs
13 yrs
15 yrs
18 yrs
19 yrs
21 yrs
23 yrs
30 yrs
36 yrs
37 yrs
40 yrs
Developmental age
SST
CALB2
PVALB
NPY
CCK
VIP

Figure 4-1: Expression of interneuron markers in the human brain during development

Gene expression data queried from Allen Brain Atlas Developmental Transcriptome and filtered for markers of interneuron progenitors and interneurons reveal spatial distribution of interneurons during development. **(A)** Regional expression of interneuron progenitors. Gene expression, measured in reads per kilobase per million reads (RPKM) from all time points for which data is available and reliable, were averaged together and grouped by brain region, and visualized as a proportion of the total RPKM for all regions. The brain regions with the highest representation of interneuron markers (MGE, CGE, LGE, striatum) are identified with asterisks. **(B)** Gene expression, graphed as $\log_2$ RPKM, of interneuron progenitor transcription factors graphed over developmental time only using data from the regions identified in figure 4A. Interneuron progenitor TF expression decreases during the second trimester of gestation. **(C)** Gene expression, graphed as $\log_2$ RPKM, of interneuron markers graphed over developmental time. Data from all brain regions were averaged together. Parvalbumin, Cholecystokinin, and Vasoactive Intestinal Peptide increased during the last trimester of gestation, while Somatostatin, Calretinin, and Neuropeptide Y are already expressed at the earliest time points of this dataset.

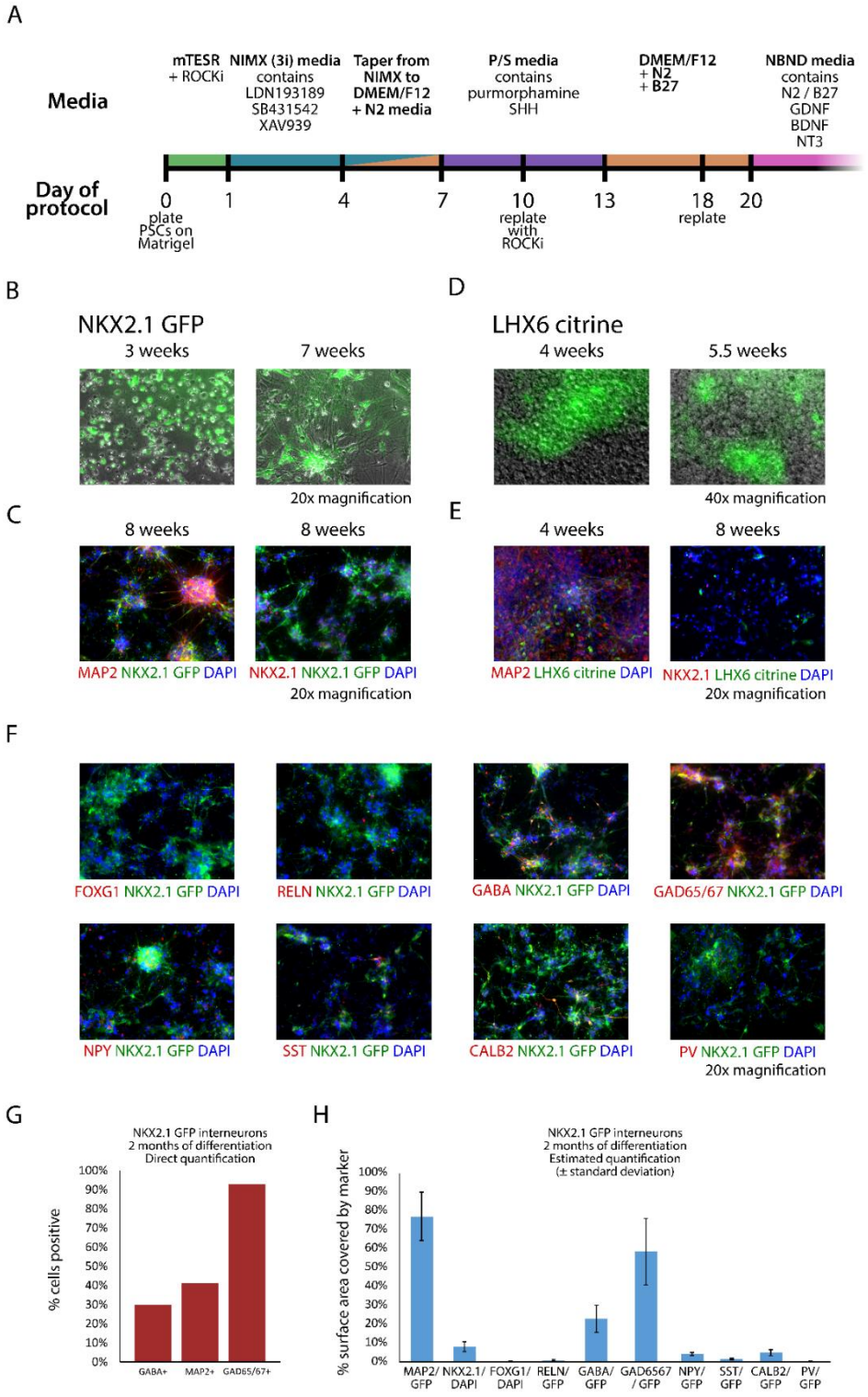


Figure 4-2: *In vivo* differentiation of inhibitory interneurons from pluripotent stem cells.

(A) Schematic of interneuron differentiation protocol beginning with PSCs grown on Matrigel. Differentiation progresses through three distinct phases: neuroepithelial differentiation via culture with NIMX media containing SMAD and Wnt pathway inhibitors, regional specification via culture with P/S media containing purmorphamine and recombinant Shh, and differentiation and maintenance via culture with NBND media containing trophic growth factors. **(B)** Interneuron progenitor reporter driven fluorescent protein expression activates during interneuron differentiation. GFP expressed under the NKX2.1 reporter is broadly present as early as 3 weeks into interneuron differentiation, and remains present following prolonged differentiation. **(C)** Immunostaining reveals that NKX2.1::GFP cells are MAP2+ neurons, and expression of the NKX2.1 protein in GFP+ cells confirms the validity of the reporter. **(D)** Interneuron reporter driven fluorescent protein expression activates during interneuron differentiation. Citrine expressed under the LHX6 reporter is present in a subset of differentiated neurons as early as 4 weeks into interneuron differentiation, and persists thereafter. **(E)** Immunostaining reveals that LHX6::Citrine cells are MAP2+ neurons, and that they also continue to express NKX2.1 despite terminal differentiation. **(F)** NKX2.1::GFP interneurons at 8 weeks of differentiation immunostaining demonstrates expression of interneuron and interneuron subtype markers. The percentages of cells that are GABA+, MAP2+, and GAD65/67+ are quantified in **(G)**. A computer-assisted estimated quantification of relative percentages of cells expression other markers (± standard deviation between fields of view) is presented in **(H)**.

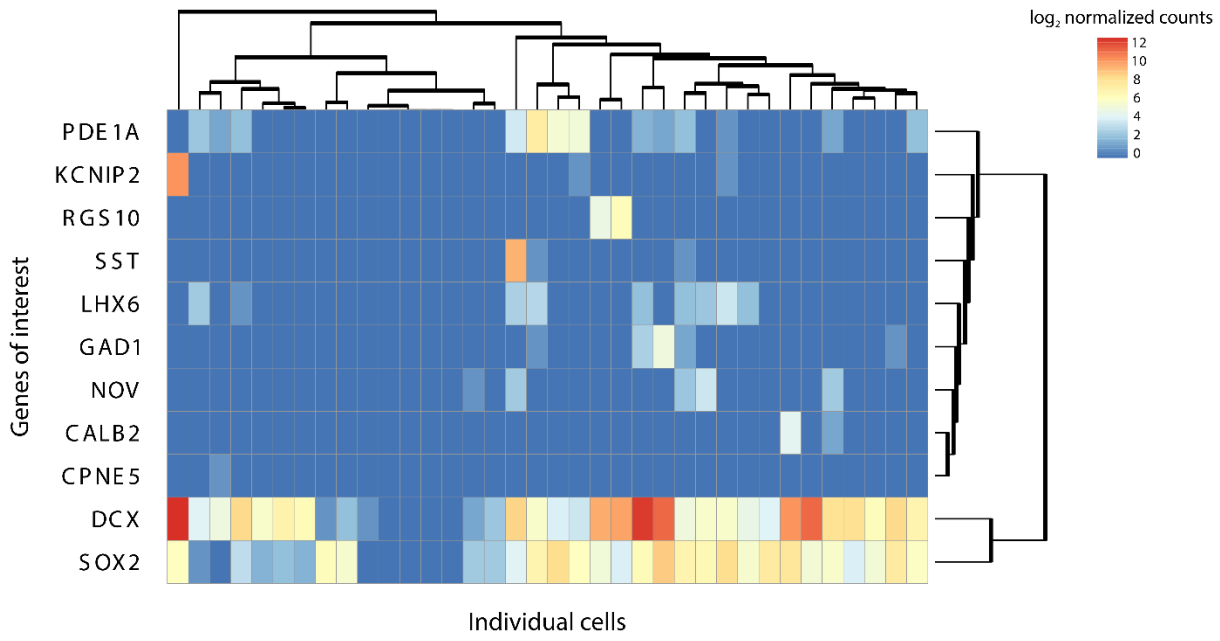


Figure 4-3: Single-cell RNA-sequencing reveals heterogeneity of gene expression in human PSC-derived interneurons

36 single cells were isolated, lysed, and used to purify RNA and generate cDNA using the Fluidigm C_1 instrument. RNA-sequencing was performed for each of these single cells. The relative, log-normalized expression of markers of neural differentiation and interneuron specification are graphed in grid form. Bidirectional hierarchical clustering was performed in order to sort cells with those most similar in gene expression, and to group expressed markers by their degree of correlation with one another. This method was able to detect rare cells reflective of interneuron subtypes, such as the single cell strongly positive for Somatostatin, or the more commonly represented SOX2+ DCX+ LHX6+ cells.

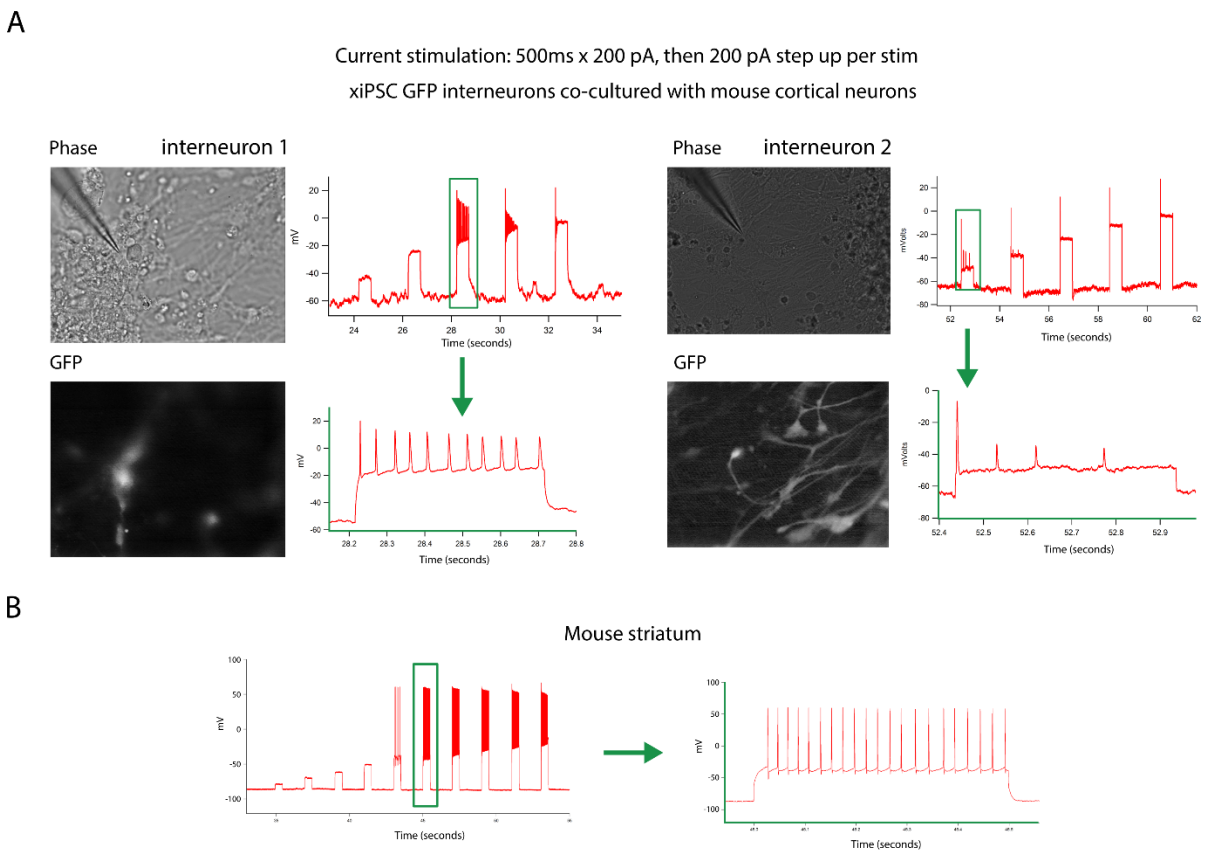


Figure 4-4: Co-culture of human PSC-derived interneurons with mouse cortical neurons promotes electrophysiological maturation

Images and electrophysiological tracings of neurons recorded by current clamp. **(A)** Interneurons differentiated from human PSCs constitutively expressing GFP were co-cultured with mouse cortical neurons for 10 days, then their membrane potential was tested while inducing progressively larger stimulatory currents. Phase images and epifluorescent images confirm that patched cells demonstrate neuronal morphology and reporter expression. The top electrophysiological tracing for each cell demonstrates changes in membrane potential representing either spontaneous or evoked excitation potentials in response to current injected in steps of 200 pA. Bottom tracing for each cell shows an inset of one set of evoked potentials.

The left interneuron showed a 20 Hz firing rate, consistent with activity of an immature interneuron. The right interneuron showed less robust, less frequent evoked potentials, representative of a very immature interneuron. No evoked or spontaneous potentials were observed for PSC-derived interneurons without co-culture conditions. **(B)** Current clamp tracing and inset of a mouse striatal interneuron shows fast-spiking activity at 50 Hz, which is expected for mature parvalbumin-expressing interneurons.

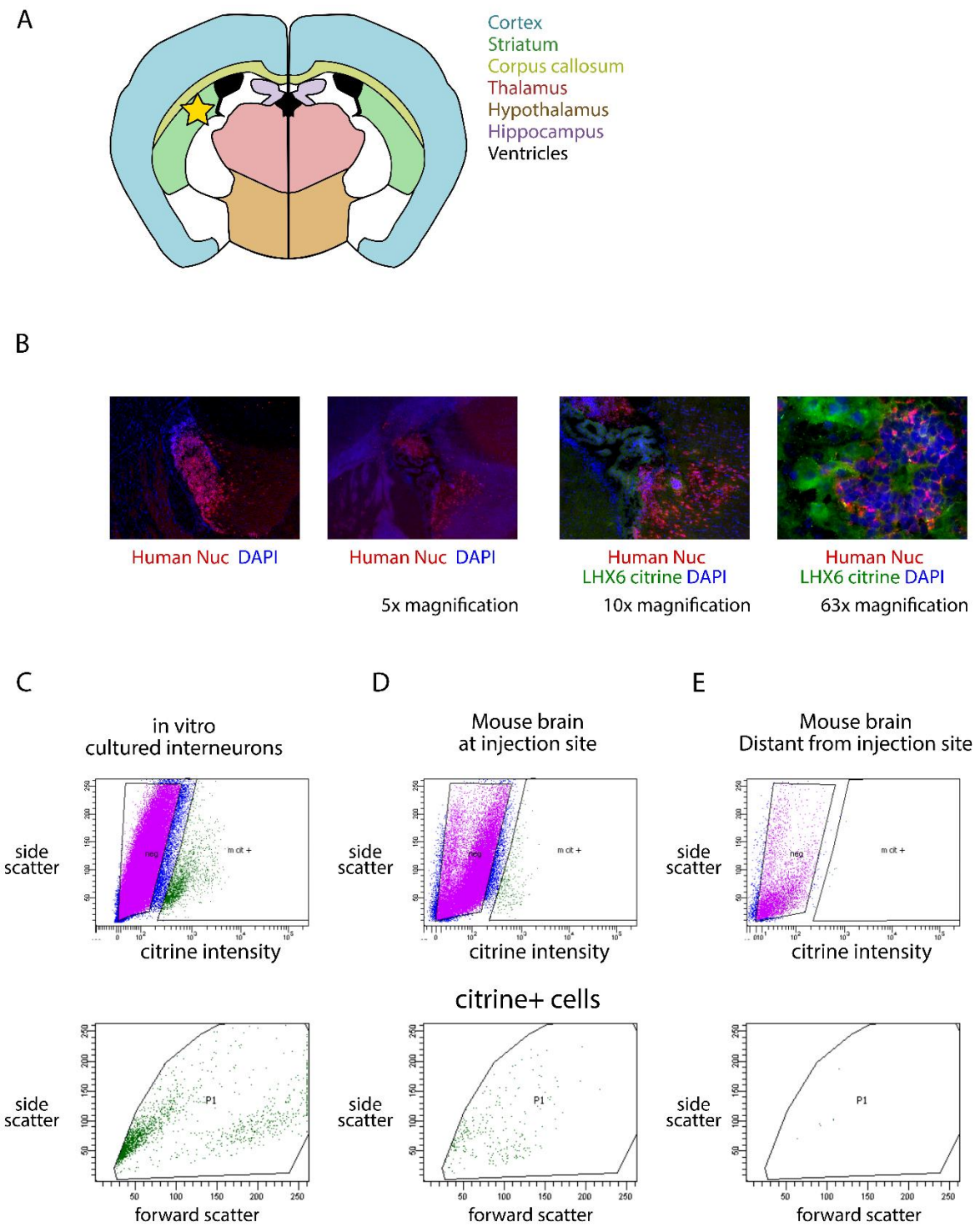


Figure 4-5: Human PSC-derived NPCs can be injected into adult mouse striatum and recovered by Fluorescence-Activated Cell Sorting

(A) Diagram of mouse brain, coronal view at .95 mm posterior of bregma. Star marks injection site, one on each hemisphere of the brain. Within this view, endogenous Interneurons are found most prominently in the cortex, striatum, and thalamus. **(B)** Human LHX6::citrine PSC-derived interneuron cells, marked by human nuclear antigen (red), were detected in the mouse brain 1 month after injection. While the majority of human cells were present in a large bolus, there was some migration of cells into the nearby brain parenchyma. **(C)** FACS plots demonstrating the ability to sort out LHX6::citrine^{+} cells from dissociated mouse brain one month after injection. Left: human LHX6::citrine cells grown for 1 month *in vitro* as a positive control. Middle: cells sorted from the dissociated mouse brain from slices surrounding the injection site. Right: citrine+ cells were absent from sections of the mouse cerebellum, far from the injection site. For each FACS experiment, the top plot shows all cells measured by their side scatter and their fluorescence intensity. The bottom plot shows only the citrine+ cells gated from the top plot, graphed according to their forward and side scatter.

References

1. Nicholas, C. R. *et al.* Functional Maturation of hPSC-Derived Forebrain Interneurons Requires an Extended Timeline and Mimics Human Neural Development. *Cell Stem Cell* **12,** 573–586 (2013).

2. Maroof, A. M. *et al.* Directed Differentiation and Functional Maturation of Cortical Interneurons from Human Embryonic Stem Cells. *Cell Stem Cell* **12,** 559–572 (2013).

3. Ramón y Cajal, S. *Recuerdos de mi vida: historia de mi labor científica*. (Imprenta y Libreria de N. Moya, 1917).

4. DeFelipe, J. *et al.* New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nat. Rev. Neurosci.* **14,** 202–216 (2013).

5. Kelsom, C. & Lu, W. Development and specification of GABAergic cortical interneurons. *Cell Biosci.* **3,** 19 (2013).

6. Livesey, M. R. *et al.* Maturation of AMPAR Composition and the GABAAR Reversal Potential in hPSC-Derived Cortical Neurons. *J. Neurosci.* **34,** 4070–4075 (2014).

7. Ben-Ari, Y. Excitatory actions of gaba during development: the nature of the nurture. *Nat. Rev. Neurosci.* **3,** 728–739 (2002).

8. DeBoer, E. M. & Anderson, S. A. Fate determination of Cerebral Cortical GABAergic Interneurons and their Derivation from Stem Cells. *Brain Res.* (2015).

9. Butt, S. J. B. *et al.* The Requirement of Nkx2-1 in the Temporal Specification of Cortical Interneuron Subtypes. *Neuron* **59,** 722–732 (2008).

10. Xu, Q. Origins of Cortical Interneuron Subtypes. *J. Neurosci.* **24,** 2612–2622 (2004).

11. Zeisel, A. *et al.* Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347,** 1138–1142 (2015).

12. Pollen, A. A. *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32,** 1053–1058 (2014).

13. Karumbayaram, S. *et al.* From Skin Biopsy to Neurons Through a Pluripotent Intermediate Under Good Manufacturing Practice Protocols. *Stem Cells Transl. Med.* **1,** 36–43 (2012).

14. Goulburn, A. L. *et al.* A Targeted NKX2.1 Human Embryonic Stem Cell Reporter Line Enables Identification of Human Basal Forebrain Derivatives. *STEM CELLS* **29,** 462–473 (2011).

15. Butt, S. J. B. *et al.* The Temporal and Spatial Origins of Cortical Interneurons Predict Their Physiological Subtype. *Neuron* **48,** 591–604 (2005).

16. Bardy, C. *et al.* Neuronal medium that supports basic synaptic functions and activity of human neurons in vitro. *Proc. Natl. Acad. Sci.* **112,** E2725–E2734 (2015).

17. Patterson, M. *et al.* let-7 miRNAs can act through Notch to regulate human gliogenesis. *Stem Cell Rep.* **3,** 758–773 (2014).

18. Tuncdemir, S. N., Fishell, G. & Batista-Brito, R. miRNAs are Essential for the Survival and Maturation of Cortical Interneurons. *Cereb. Cortex* **25,** 1842–1857 (2015).

19. Ferando, I. & Mody, I. In vitro gamma oscillations following partial and complete ablation of δ subunit-containing GABAA receptors from parvalbumin interneurons. *Neuropharmacology* **88,** 91–98 (2015).

Chapter 5: Conclusions

Pluripotent stem cells represent a source of huge potential for the biomedical sciences. They can be used to study and model normal and aberrant development, allowing a direct approach to previously unanswerable questions of development in a tractable experimental system. Their potential uses also include repairing and replacing cells and tissues lost or rendered nonfunctional through disease or injury. Together, these opportunities represent a paradigm-shifting development that despite decades of study we are still attempting to understand and leverage. While a great deal of progress has been made toward differentiating varied, specialized cells from pluripotent stem cells, groups focusing on generating many distinct cells types have encountered similar roadblocks that prevent the generation of fully mature cells *in vitro*[1–6]. The work presented in this thesis represents our efforts to understand and overcome these limitations by discovering their mechanistic underpinnings.

Progress toward understanding mature differentiation and the intrinsic clock

Studies from our lab led to the hypothesis that one of the major hurdles toward differentiation from pluripotent stem cells was a switch from a LIN28-high *let-7*-low state to a LIN28-low *let-7*-high state[3]. While this hypothesis was generalized using data from 3 distinct cell types each representing a different germ layer, the maturation of many cell types not used in that study have now been shown to correlate with *let-7* expression[1,7–11]. In Chapter 2 we validate this hypothesis by direct manipulation of the LIN28/*let-7* circuit, and show that changes in this pathway have functional consequences in neural progenitor cells. Gene expression analysis confirmed that NPCs spiked with artificial *let-7*s miRNAs more closely resemble NPCs from later

ages during fetal development. Together, these studies suggest that manipulating the *let-7* miRNAs might be a generally applicable tactic for pushing cells toward maturity *in vitro*.

Based on the findings from Chapter 2, we know that changes in the expression of *let-7* correlate with and are sufficient to drive neural progenitor cells toward maturation. Yet progress toward understanding the genesis of the switch to a high let-7 state requires elucidating the mechanism behind that rise in mature *let-7*s. Chapter 3 summarizes our efforts to understand the regulatory environment driving changes in the transcription of primary *let-7*s, which is possible now because of efforts to better characterize the true boundaries of the *let-7* genes[12,13]. We make the case that while the LIN28 proteins and their cofactors are crucial for regulating the biogenesis and maturation of *let-7* miRNAs, the change between states in the bistable LIN28/*let-7* switch actually comes from rapid increases in transcription of just a few *let-7* loci. New developments in miRNA biology raise the intriguing possibility that these specific *let-7s* are also able to bypass LIN28-mediated regulation, meaning that they could be sufficient for both increasing *let-7* expression and decreasing LIN28 expression on their own[14].

While we and others have shown that *let-7* is important for maturation along the time scale of mid-gestation, other cell types in which *let-7* miRNA expression does not change also experience maturation. In Chapter 4 we describe preliminary efforts to drive maturation in one such system – inhibitory interneurons of the striatum and cortex. While our *in vitro* system is capable of generating large, diverse populations of interneurons, we have only detected signatures of neuronal maturity in cells that have been co-cultured with mature excitatory cells, or that have

been grown within the mouse brain for an extended time. Others have reported similar approaches to inducing maturation in PSC-derived neurons, but our ability to profile these cells after driving them toward maturity represents a possible way to discover the genetic and molecular mechanisms regulating maturation on this scale[4,5]. Such a mechanism would help fill in the gaps of our understanding of aging and maturation, as existing ways of cataloguing cellular age apply either at very short, early time scales or at a lifespan-wide time scale[15,16].

Mature PSC-derived cells and regenerative medicine

A better understanding of cellular maturation and aging is important for our understanding of development, organogenesis, growth, and epigenetics. These aspects of biology are no less important for applications more directly applicable to human health. Using pluripotent stem cells to model disease is only possible if we can faithfully recapitulate the setting in which disease arises[17]. And making sure pluripotent stem cell derived cell products are functional and appropriately mature is of central importance if they are to be used clinically. Especially as clinical trials have now begin using stem cell-derived progeny, research that improves the fidelity and safety of those cells is a critical priority[18,19].

Though our focus has been improving the functional resemblance of PSC-derived cells to their counterparts from normal development, ensuring that signatures of embryonic age and immaturity are appropriately silenced during differentiation is also a concern. Cells inappropriately aged *in vitro* might retain expression of genes that can also predispose cells toward tumorigenesis[20–22]. Finding the correct balance, and taking our cues from development

as it occurs, are essential preconditions for safe, effective cell replacement therapies. Eventually clinical trials will be a vital step in determining the safety and utility of PSC-derivatives driven *in vitro* toward different cell types and cellular ages. Our work here describes hypothesis-driven approaches to elucidating and manipulating the cellular and molecular systems behind developmental maturation, and may be fruitful for future biomedical uses of stem cell-derived progeny.

References

1. Kuppusamy, K. T. *et al.* Let-7 family of microRNA is required for maturation and adult-like metabolism in stem cell-derived cardiomyocytes. *Proc. Natl. Acad. Sci.* **112,** E2785–E2794 (2015).

2. Pagliuca, F. W. *et al.* Generation of Functional Human Pancreatic β Cells In Vitro. *Cell* **159,** 428–439 (2014).

3. Patterson, M. *et al.* Defining the nature of human pluripotent stem cell progeny. *Cell Res.* **22,** 178–193 (2011).

4. Nicholas, C. R. *et al.* Functional Maturation of hPSC-Derived Forebrain Interneurons Requires an Extended Timeline and Mimics Human Neural Development. *Cell Stem Cell* **12,** 573–586 (2013).

5. Maroof, A. M. *et al.* Directed Differentiation and Functional Maturation of Cortical Interneurons from Human Embryonic Stem Cells. *Cell Stem Cell* **12,** 559–572 (2013).

6. Daniel, M. G., Pereira, C.-F., Lemischka, I. R. & Moore, K. A. Making a Hematopoietic Stem Cell. *Trends Cell Biol.* **26,** 202–214 (2016).

7. Copley, M. R. *et al.* The Lin28b–let-7–Hmga2 axis determines the higher self-renewal potential of fetal haematopoietic stem cells. *Nat. Cell Biol.* **15,** 916–925 (2013).

8. Gurtan, A. M. *et al.* Let-7 represses Nr6a1 and a mid-gestation developmental program in adult fibroblasts. *Genes Dev.* **27,** 941–954 (2013).

9. Cimadamore, F., Amador-Arjona, A., Chen, C., Huang, C.-T. & Terskikh, A. V. SOX2-LIN28/let-7 pathway regulates proliferation and neurogenesis in neural precursors. *Proc. Natl. Acad. Sci.* **110,** E3017–E3026 (2013).

10. Madison, B. B. *et al.* Let-7 Represses Carcinogenesis and a Stem Cell Phenotype in the Intestine via Regulation of Hmga2. *PLOS Genet.* **11,** e1005408 (2015).

11. Xia, X. & Ahmad, I. let-7 microRNA regulates neurogliogenesis in the mammalian retina through Hmga2. *Dev. Biol.* **410,** 70–85 (2016).

12. Patterson, M. *et al.* let-7 miRNAs can act through Notch to regulate human gliogenesis. *Stem Cell Rep.* **3,** 758–773 (2014).

13. Chang, T.-C., Pertea, M., Lee, S., Salzberg, S. L. & Mendell, J. T. Genome-wide annotation of microRNA primary transcript structures reveals novel regulatory mechanisms. *Genome Res.* **25,** 1401–1409 (2015).

14. Triboulet, R., Pirouz, M. & Gregory, R. I. A Single Let-7 MicroRNA Bypasses LIN28-Mediated Repression. *Cell Rep.* (2015). doi:10.1016/j.celrep.2015.08.086

15. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14,** R115 (2013).

16. Melton, C., Judson, R. L. & Blelloch, R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* **463,** 621–626 (2010).

17. Miller, J. D. *et al.* Human iPSC-Based Modeling of Late-Onset Disease via Progerin-Induced Aging. *Cell Stem Cell* **13,** 691–705 (2013).

18. Trounson, A., Thakar, R. G., Lomax, G. & Gibbons, D. Clinical trials for stem cell therapies. *BMC Med.* **9,** 1 (2011).

19. Tabar, V. & Studer, L. Pluripotent stem cells in regenerative medicine: challenges and recent progress. *Nat. Rev. Genet.* **15,** 82–92 (2014).

20. Viswanathan, S. R. *et al.* Lin28 promotes transformation and is associated with advanced human malignancies. *Nat. Genet.* **41,** 843–848 (2009).

21. Busch, B. *et al.* The oncogenic triangle of HMGA2, LIN28B and IGF2BP1 antagonizes tumor-suppressive actions of the let-7 family. *Nucleic Acids Res.* gkw099 (2016). doi:10.1093/nar/gkw099

22. Ma, C. *et al.* H19 promotes pancreatic cancer metastasis by derepressing let-7's suppression on its target HMGA2-mediated EMT. *Tumor Biol.* **35,** 9163–9169 (2014).