UNIVERSITY OF CALIFORNIA, SAN DIEGO

Genetic Manipulation of Fezf2 in Human Embryonic Stem Cells and Its Application in Studying Central Nervous System Development and Repair

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Biomedical Sciences

by

Katherine Marie Ruby

Committee in charge:

Professor Binhai Zheng, Chair
Professor Sylvia Evans
Professor Joseph Gleeson
Professor Lawrence Goldstein
Professor Martin Marsala
Professor Karl Willert

2011

The dissertation of Katherine Marie Ruby is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2011

Dedication

In recognition of the endless support and encouragement throughout my years in graduate school, this dissertation is dedicated to my parents, Timothy and Ruth Ruby, and my sister, Heather Ruby.

Table of Contents

List of Abbreviations

bFGF: basic fibroblast growth factor

CSMN: corticospinal motor neuron

CST: corticospinal tract

DAPI: 4',6-diamidino-2-phenylindole, fluorescent stain for DNA to mark nuclei

DMEM: Dulbecco's Modified Eagle Medium

DMSO: Dimethyl sulfoxide

EB: embryoid body

FACS: Fluorescence activated cell sorting

Fezf2/Fezl: forebrain embryonic zinc finger like-protein 2

FIAU: (2'-deoxy-2'-fluoro-ß-D-arabinofuranosyl)-5-iodouracil, drug used for negative selection against TK

GFRM: growth factor reduced Matrigel

hESC: human embryonic stem cell

HUES-9-h*Fezf2*: HUES-9 cells targeted at the h*Fezf2* locus

HSV-TK: Herpes Simplex Virus Thymidine Kinase, used as a negative selection marker

LDN-193189: a small molecule inhibitor of BMP signaling

MEF: mouse embryonic fibroblast

MCM: MEF conditioned media

NEAA: non-essential amino acids

PBS: phosphate buffered saline

RA: retinoic acid

ROCK: Rho-associated kinase

SB431542: a small molecule inhibitor of Lefty/Activin/TGFβ signaling pathways

SNP: single nucleotide polymorphism

List of Figures

List of Tables

Acknowledgements

Chapter 1, in part, is a reprint of the material as it appears in Ruby and Zheng, Cell Bioengineering and Tissue Microenvironment, in press. The dissertation author is the primary author of this book chapter. All statements were accurate as of the time of publishing.

Chapter 2, in part, is a reprint of the material as it appears in Ruby and Zheng, Stem Cells 2009. The dissertation author is the primary author of this paper. All statements were accurate as of the time of publishing.

Chapter 3, in full, is a reprint of the material as it appears in Ruby and Zheng, Stem Cells 2009. The dissertation author is the primary author of this paper. All statements were accurate as of the time of publishing.

Curriculum Vita

2001-2005	Bachelor of Science, Michigan State University
2006-2007	Teaching Assistant, Department of Cellular and Molecular Medicine University of California, San Diego
2006-2008	Teaching Assistant, Salk Mobile Science Lab
2009-current	Registered Patent Agent, United States Patent and Trademark Office
2005-2011	Graduate Student Researcher, Department of Neurosciences University of California, San Diego
2011	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

Gray RM, Davis MJ, Ruby KM, Voss PG, Patterson RJ, Wang JL. Distinct effects on splicing of two monoclonal antibodies directed against the amino-terminal domain of galectin-3. Arch Biochem Biophys. 2008; 2: 100-8.

Ruby KM and B. Zheng. Gene targeting in a HUES line of human embryonic stem cells via electroporation. Stem Cells. 2009; 7: 1496-506.

Ruby, K. M. and Zheng, B. Genetic engineering of human embryonic stem cells: focusing on targeted modification. In Cell Bioengineering and Tissue Microenvironment, S. Prakash and D. Shum-Tim, eds. World Scientific Publishing Company. In press.

AWARDS AND HONORS

2003	Undergraduate Research Award, Michigan State University, Department of Biochemistry
2005	University Undergraduate Research and Arts Forum Merit Award, Michigan State University
2005	Outstanding Undergraduate Award, Michigan State University, Department of Biochemistry
2006-2008	California Institute of Regenerative Medicine (CIRM) Training Grant Recipient
2009	Travel Award to serve as a delegate at Stem Cells World Congress, Palm Springs, CA
2009	Best Student Presentation Award, Biomedical Sciences Graduate Program Annual Retreat, UCLA Conference Center, Lake Arrowhead, California
2009	Travel Award to Stem Cell Biology Conference, Cold Spring Harbor Laboratory, NY
2010	CIRM Travel Award to International Society for Stem Cell Research (ISSCR) Conference, San Francisco, California

ABSTRACTS

10/20/06	Stem Cells on the Mesa Meeting, La Jolla, California
9/28/07	Annual CIRM Scholars Meeting, Irvine, California
10/3-10/5/07	Biomedical Sciences Annual Retreat, UCLA Conference Center, Lake Arrowhead, California
10/19/07	Stem Cells on the Mesa Meeting, La Jolla, California
9/17/08	Annual CIRM Scholars Meeting, San Francisco, California
5/15-5/18/09	Biomedical Sciences Annual Retreat, UCLA Conference Center, Lake Arrowhead, California
9/22-9/26/09	Stem Cell Biology Meeting, Cold Spring Harbor, New York
6/16-6/19/10	International Society for Stem Cell Research (ISSCR) Annual Meeting, San Francisco, California

Abstract of the Dissertation

Genetic Manipulation of Fezf2 in Human Embryonic Stem Cells and Its Application in Studying Central Nervous System Development and Repair

by

Katherine Marie Ruby

Doctor of Philosophy in Biomedical Sciences

University of California, San Diego, 2011

Professor Binhai Zheng, Chair

One of the major goals in the stem cell biology field is to understand the mechanisms involved in directing cells towards a particular fate. This is an especially complex problem when studying the central nervous system, due to the variety of neuronal cells that comprise the brain and spinal cord. To study any specific lineage or cell type, it is necessary to have tools that can identify the cell type of interest. One such method is to generate a reporter cell line to fluorescently label cells that differentiate towards the lineage of interest.

To study the mechanisms involved in differentiation and specification of corticospinal motor neurons, the neuron type damaged in spinal cord injury and amyotrophic lateral sclerosis, we chose to study *Fezf2*, a gene that has been shown to play an essential role in corticospinal motor neuron

development in mice. To track cells expressing Fezf2, I developed a fluorescent reporter under the control of the endogenous *Fezf2* promoter via homologous recombination. This method is generally applicable to target all genes, regardless of their expression. Analysis of the targeted cell line confirmed functional reporter expression, with fluorescence being detected only after differentiation, consistent with the fact that Fezf2 is not expressed in ESCs.

To study the role of Fezf2 in human cells, we developed a directed differentiation protocol to generate Fezf2+ cells. Analysis of pluripotency and neural markers confirmed our culture system efficiently induces a neural identity. At later time points, Fezf2 expression is detected in a subset of cells. Fezf2+ cells also express β-tubulin, and a subset of Fezf2+ cells co-express CTIP2, a transcription factor downstream of Fezf2. Gene expression analysis of purified Fezf2+ cells revealed conservation of expression of a number of corticospinal markers identified in mice. Together, this data suggests there is some conservation between mouse and human corticospinal development, indicating Fezf2 may be an ideal gene for studying human corticospinal development. This Fezf2 reporter line is a valuable tool that will facilitate further studies into mechanisms involved in human corticospinal development, including in vivo studies looking at development as well as repair strategies for neurological conditions.

Chapter 1

Introduction

1.1 Overview

Millions of people in the United States alone are currently suffering from incurable diseases. A greater understanding of the molecular mechanisms that are misregulated in disease processes is critical to discovering ways to cure, or even prevent, these diseases. While mouse models have proven to be invaluable tools for identifying key genes and pathways involved in a wide variety of diseases, it is essential to remember there are many differences between rodents and humans. Therefore, it is important to develop tools to study disease processes in human cells to increase the chances of developing effective therapeutic strategies for patients while decreasing potential risk and side effects.

One valuable tool that has shown great promise is human embryonic stem cells. Due to their ability to self-renew, there is a limitless supply of cells to work with. Additionally, these cells have the capacity to form every cell in the human body, making them useful tools for studying a variety of diseases involving numerous different cell types. This is especially important in diseases such as spinal cord injury and amyotrophic lateral sclerosis that involve cell types that are not readily accessible in patients. To study the neuronal cells that are affected in these diseases, it is necessary to derive cells from an alternate source, such as human embryonic stem cells.

Although human embryonic stem cells hold great promise to shed valuable insight on the processes underlying disease, the ability to specifically direct these cells to become a particular cell type of interest is a major

obstacle. Great progress has been made in deciphering key pathways and genes involved in cell fate specification during development. As we continue to learn more about the mechanisms controlling cells both during development and in disease, we will increase our ability to develop effective therapeutics for patients suffering from many different diseases.

1.2 Mammalian cortical development

The mammalian neocortex (also referred to simply as the cortex) is the region of the brain responsible for cognitive function, sensory perception, and consciousness. These functions are accomplished through an intricate series of neuronal connections within the cortex, as well as connections between the cortex and other parts of the brain. While the cortex has undergone substantial expansion and development throughout the course of evolution[1], there are many similarities between species that allow scientists to gain a better understanding of the complexities of the human brain.

The cortex is a highly organized and complex structure. It is organized into six layers, which can be identified based on the morphology and density of neurons within each layer. Cortical development occurs in an inside-out pattern[2, 3] **(Figure 1.1)**. The earliest born neurons populate the deep layers (layers V and VI) and later born neurons migrate past the deep layer neurons to form the upper layers (layers II-IV)[4].

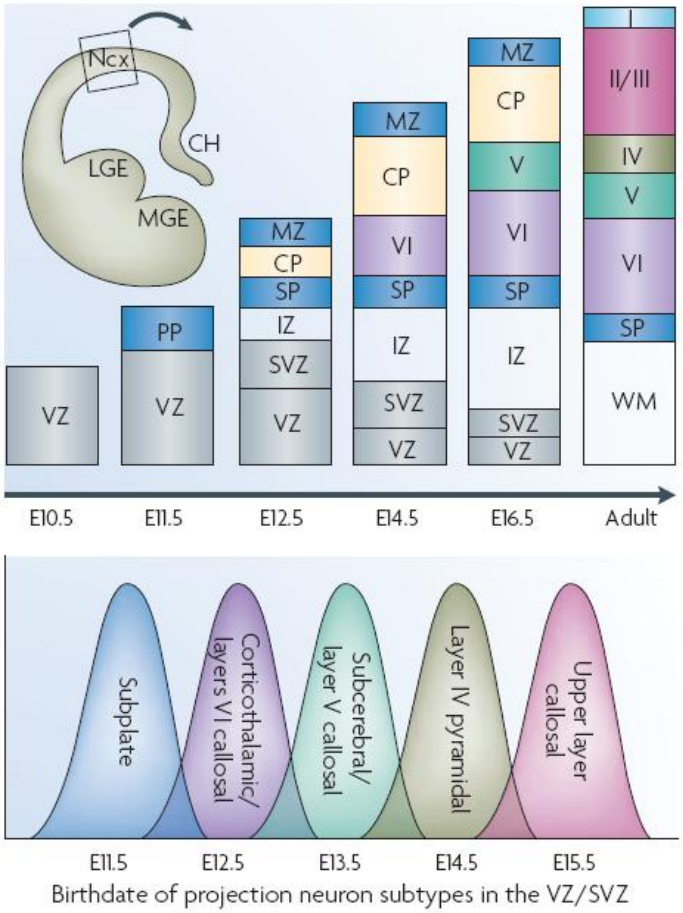


Figure 1.1: Schematic depicting how progenitors residing in the VZ and SVZ in mice produce projection neurons in an "inside-out" fashion. The earliest born neurons form the preplate (PP), which is later split into the more superficial marginal zone (MZ) and the deeply located subplate (SP). The cortical plate (CP), which will give rise to the multilayered neocortex, develops in between these two layers, such that later born neurons arriving at the cortical plate migrate past earlier born neurons. Different classes of projection neuron are born in overlapping temporal waves. All times listed are approximations given the neurogenic gradients that exist across the cortex, where caudomedial neurogenesis lags behind rostrolateral neurogenesis. CH, cortical hem; E, embryonic day; Ncx, neocortex; IZ, intermediate zone; LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence; SVZ, subventricular zone; VZ, ventricular zone; WM, white matter. (Molyneaux, *et al*. 2007)

There are two main classifications of cortical neurons: interneurons and projection neurons. Interneurons are small, inhibitory neurons which make local connections. In contrast, projection neurons are large, excitatory neurons which extend axons to distant targets, including the spinal cord. Projection neurons are characterized by the distinctive pyramidal shape of their cell bodies[5, 6].

While projection neurons populate all six layers of the cortex, distinct populations of projection neurons are located in different cortical layers. Callosal projection neurons that send axons through the corpus callosum into the opposite cortical hemisphere are found in layers II-VI. Neurons that project to subcortical targets are restricted to layers V and VI. Neurons projecting to the thalamus are located in layer VI, whereas those projecting to targets in the midbrain, hindbrain, and spinal cord are found in layer V[7].

1.3 The corticospinal tract

The corticospinal tract (CST), also called the pyramidal tract, is the largest and longest descending axonal pathway in the mammalian CNS. Restricted to mammals, the CST is involved both in motor and sensory control. The corticospinal system is the principal motor system for controlling movements that require the greatest skill and flexibility. It is the last motor system to develop. As the name implies, the CST consists exclusively of motor axons that travel between the cerebral cortex of the brain and the spinal cord[8, 9] **(Figure 1.2)**.

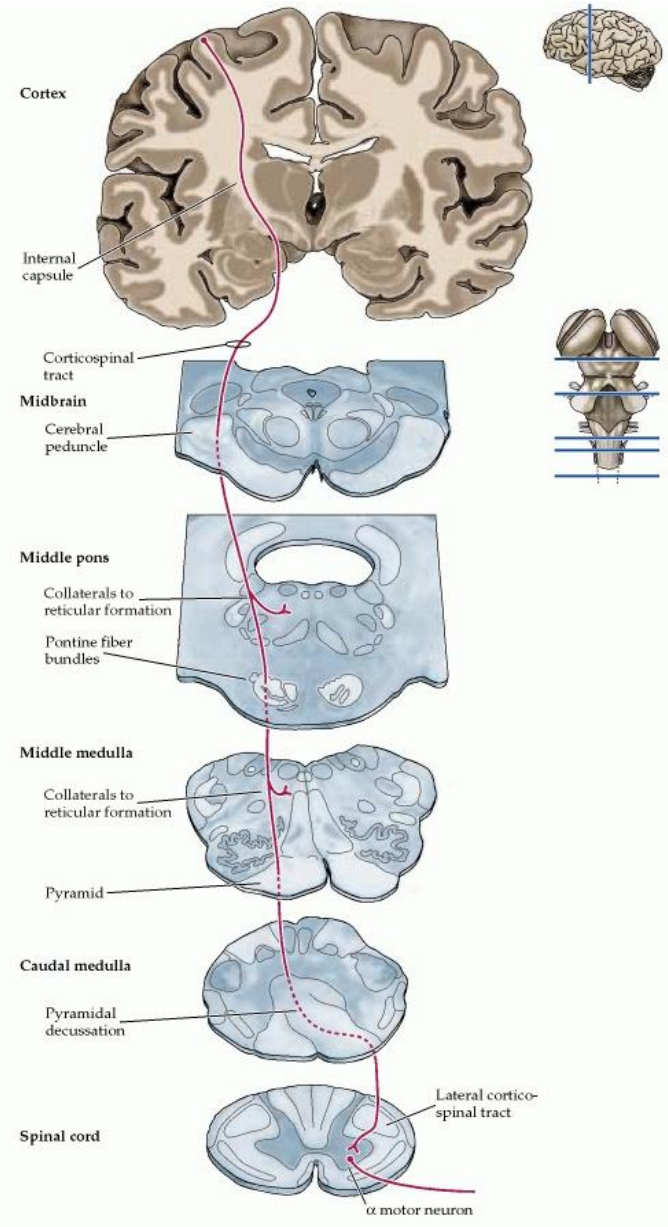


Figure 1.2: The corticospinal tract. Neurons in the motor cortex give rise to axons that travel through the internal capsule and coalesce on the ventral surface of the midbrain, within the cerebral peduncle. These axons continue through the pons and come to lie on the ventral surface of the medulla, giving rise to the pyramids. Most of these pyramidal fibers cross in the caudal part of the medulla to form the lateral corticospinal tract in the spinal cord. Those axons that do not cross (not illustrated) descend on the same side and form the ventral corticospinal tract. The axons that terminate in the reticular formation of the pons and medulla comprise components of the corticobulbar tract. (Purves, *et al*. 2001)

1.3.1 Development of the CST

The timing of motor development is very diverse among different species. The wildebeest is able to maintain a quadrupedal posture and run within minutes after birth. In contrast, many mammals—including mice, cats, and humans—are born only with motor functions necessary for survival, such as respiration and feeding behaviors. Even within these species, there is a marked difference in development after birth. Mice develop mature motor skills within the first month of life, and cats during the first few months. Human development, however, is measured in years[10].

The CST arises exclusively from large pyramidal neurons located in layer V of the cerebral cortex, called corticospinal motor neurons (CSMNs, also called upper motor neurons). The neurons of layer V are born at a distinct and restricted period of development. Consistent with the inside-out development of the cortex, layer V neurons are born relatively early in development, after layer VI neurons but before upper layer neurons[2, 11].

1.3.2 Neurological conditions involving the CST

Neurodegeneration refers to the loss of structure or function of neurons, including neuron death. Neurodegeneration is a common theme in many nervous system disorders, including Parkinson's disease, Huntington's disease, epilepsy, stroke, amyotrophic lateral sclerosis, and spinal cord injury. These neurological conditions are devastating and very costly, and current therapies are lacking. To develop new therapies, it is critical to achieve a

greater understanding of the cell types involved to determine what exactly goes wrong in the disease process, and how the damage might be reversed.

1.3.2.1 Spinal cord injury

There are approximately 250,000 people in the United States currently living with spinal cord injuries (SCI), with about 12,000 new cases each year. The majority of injuries occur at a relatively young age, with the average age at injury being 40.2 years[12]. Spinal cord injury patients suffer permanent functional deficits and paralysis, largely due to the inability to regenerate axons in the adult mammalian central nervous system (CNS).

One of the major neuron types damaged in SCI is the upper motor neurons, or CSMNs. In some cases, lower motor neurons can also be damaged, depending on the location of the injury. In most cases, the neurons do not die, but they are unable to relay signals between the muscles and brain due to damage to the axons that carry the signals. There are currently no effective therapies for SCI which leads to a very large financial burden throughout the life of the patient, costing up to $2.8 million dollars for a single patient.

1.3.2.2 Amyotrophic lateral sclerosis

Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig's disease, is a debilitating disease characterized by progressive muscle weakness and atrophy. ALS affects both upper motor neurons in the brain that project down

the spinal cord and lower motor neurons that relay signals from upper motor neurons to the muscles throughout the body. The progressive degeneration of the motor neurons in ALS eventually leads to their death. When the motor neurons die, the brain is no longer able to control the muscles, leading to muscle atrophy and paralysis.

There are currently about 30,000 Americans living with ALS, with approximately 5,600 new cases diagnosed each year. The average lifetime risk of ALS is 1 in 400[13, 14], despite the fact the prevalence is only about 5 per 100,000. This is because the average life expectancy after diagnosis is only 3-5 years[15]. As the disease progresses, patient care costs increase and can cost up to $200,000 per year. Currently there are no effective therapies to cure ALS. To reverse the paralysis that results from ALS, it will be necessary to develop cell replacement therapy to replace the lost motor neurons. Human ES cells offer great promise as a source of cells, but first it is critical to learn more about the developmental pathways of motor neurons to be able to coax these undifferentiated cells towards the desired fate.

1.4 Overview of stem cells

Stem cells can be found in different tissues of the human body throughout all stages of development. A stem cell is a cell that possesses both the ability to self-renew and the capacity to differentiate into different cell types. There are multiple types of stem cells, each with a different potency, or potential to differentiate into different cell types **(Figure 1.3)**.

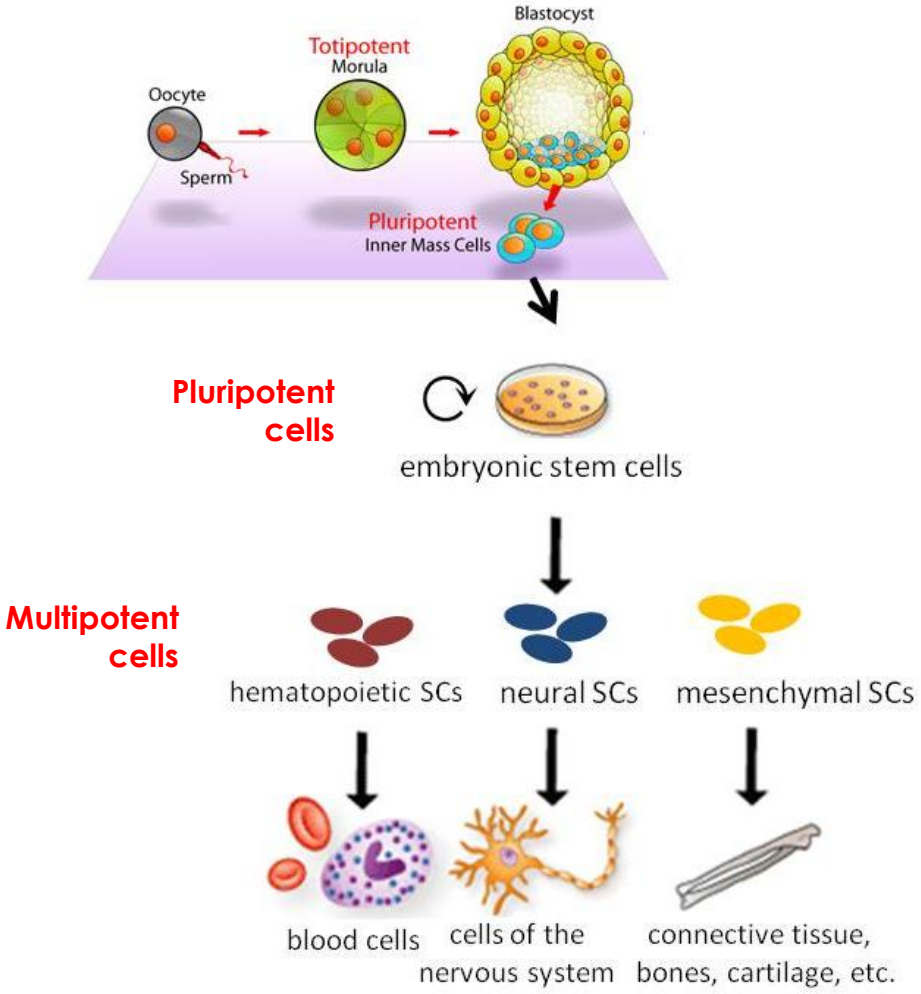


Figure 1.3: Overview of stem cell derivation and differentiation. Human embryonic stem cells are isolated by removing the inner cell mass from a blastocyst. Embryonic stem cells can be maintained in a self-renewing pluripotent state, or can be differentiated into cells from each of the three germ layers.

Totipotency is the ability to differentiate into every cell type of an organism, including extra-embryonic tissues. Totipotent cells are only found at the earliest stages of development, during the first few days after fertilization. As the organism develops, cells lose their totipotency and become pluripotent. Embryonic stem cells are pluripotent cells, meaning they have the ability to produce all the cells of the three germ layers: endoderm (digestive tract, lungs), mesoderm (muscle, bone, blood), and ectoderm (nervous system, epidermis). While pluripotent cells can produce all fetal and adult cell types, they are unable to develop into a fetal or adult animal because they no longer have the ability to produce extra-embryonic tissues. Pluripotent cells undergo further specialization to commit to a particular lineage, at which time the cells are defined as multipotent or progenitor cells. Multipotent cells are capable of becoming the different cell types of their chosen lineage, but are unable to produce cell types of other lineages. An example of a multipotent cell is a neural stem cell, which can make all types of neurons and glia, but cannot form blood or other types of cells. Multipotent cells are present in adult mammalian tissues and are therefore often referred to as adult stem cells.

1.4.1 Embryonic stem cells

Embryonic stem (ES) cells, as the name implies, are isolated from embryos. More specifically, ES cells are isolated from the inner cell mass of a very early stage embryo called a blastocyst. The blastocyst forms prior to

implantation and is comprised of two primary features: the trophoblast and the inner cell mass. The trophoblast is a layer of cells surrounding the inner cell mass that develops into the extra-embryonic support cells, including the placenta. The inner cell mass is a small ball of cells—approximately 100 cells in humans—that develops into the embryo proper. The cells of the inner cell mass are pluripotent and, therefore, the resulting cultured ES cells retain pluripotentcy as long as they receive the appropriate factors to maintain this undifferentiated state. The ability of ES cells to differentiate into all cell types of the body makes them an invaluable tool both for basic biology research, and also for transplantation and regenerative medicine therapies.

1.4.2 Using embryonic stem cells to study gene function

Embryonic stem cells were first isolated from the laboratory mouse in the early 1980s.[16, 17] The possibility to transfer any genetic mutations acquired in mouse ES cells back into a living organism provided unlimited potential for studying gene function in development and physiology and for generating mouse models of human diseases.[18] In particular, gene targeting, the process to modify an endogenous gene via homologous recombination in mouse ES cells,[19, 20] has been one of the most widely used tools in biomedical research since the late 1980s.

The isolation of human ES cells in the late 1990s provided exciting new possibilities for both basic research of human biology and for development of novel therapies for human diseases through transplantation strategies.[21] On

the one hand, just as mouse ES cells, human ES cells can serve as a tool for basic research. Much we know about human gene function has come from one of several ways: 1) identifying mutations in genes that cause human genetic disorders or predisposition to a disease; 2) extrapolating data from model organisms such as the laboratory mouse; 3) investigating the *in vitro* function of human gene products either by studying their structure or biochemical function in a test tube or by studying their role in a cell culture system. It is often difficult, if not impossible, to experimentally manipulate endogenous human genes and to study their role *in vitro* in a highly relevant cell type that can be isolated in sufficient quantity. It is even more challenging to do so to study their role *in vivo*. Since human ES cells can potentially turn into virtually any type of cell *in vitro* or *in vivo* (other than cells of the extra-embryonic origin) under appropriate conditions, the effect of any induced genetic changes in human ES cells can then be assessed in many different cell types. For example, gain or loss of function mutations can be introduced into a human ES cell line, and any effect of such genetic alterations can be assessed after transplanting the cells or their differentiated derivatives into an experimental animal.

On the other hand, unlike mouse ES cells, human ES cells can be directly used as a tool to develop cell transplantation-based therapies for a variety of human conditions. The excitement generated by human ES cells is in large part due to the potential of such cells for therapeutic development. A spectrum of human diseases or conditions where cells have been lost,

damaged or degenerating may benefit from cellular transplant with derivatives of human ES cells: type I diabetes, Parkinson's disease, and spinal cord injury, to name a few. Current challenges to this approach include the knowledge and ability to differentiate human ES cells into the desired cell types *in vitro* and/or *in vivo*, ways to deliver these cells or their derivatives into the body, the functional integration and long term survival of differentiated derivatives *in vivo*, and the problem of immune rejection and tumorigenicity associated, respectively, with allograft transplant of cells in general and that of ES cells in particular. The development of patient-specific induced pluripotent stem (iPS) cells may bypass the issue of immune rejection in the future.[22-24] However, whether iPS cells are identical to their human ES cell counterparts, especially in terms of their full differentiation potential, remains to be investigated systematically.

1.4.3 Stem cells as cellular therapy

Because embryonic stem cells are capable of forming any cell type in the body, there is great promise that scientists can direct stem cells specifically towards a desired cell type. The hope is that one day these cells will be used in therapies for a wide variety of diseases ranging from diabetes to spinal cord injuries. However, there are a number of challenges that must be overcome before stem cell therapies can efficiently and safely be translated to the clinic. One of the greatest challenges is the need to generate large quantities of the desired cell type at high purity.

1.4.4 Directed differentiation of human embryonic stem cells

To generate large quantities of a particular cell type, it is critical to know the pathways and molecules involved in directing an undifferentiated stem cell towards the differentiated cell type of interest. To develop cellular therapies for patients suffering from spinal cord injuries or ALS, one such desired cell type would be CSMNs. Currently, little is known about the molecular pathways involved in specifying CSMN fate. However, studies performed in rodent models have begun to reveal some factors that are critical for CSMN development. One of these factors is the transcription factor *Fezf2*.

1.5 Fezf2 in CSMN development

Forebrain embryonic zinc finger-like 2 (Fezf2, also called Fezl) is a six zinc finger domain-containing transcription factor[25-27]. In humans, the gene is located on chromosome 3 and encodes a 459 amino acid protein comprised of five exons.

Fezf2 was initially identified as a possible CSMN identity gene in a microarray study comparing the expression pattern of CSMNs to two other populations of cortical projection neurons in mouse. Fezf2 was shown to be highly expressed in CSMNs at all stages of development studied (E18.5-P14), indicating its potential as a key regulator of CSMN development[28]. Additionally, between E12.5 and E14.5, the time period when subcerebral

projection neurons are born, *Fezf2* is detected in the ventricular zone in a pattern consistent with the location of neural progenitors as well as in the developing cortical plate where newly born subcerebral projection neurons are located[29, 30] **(Figure 1.4)**. Taken together, this information suggests *Fezf2* may play a critical role in the specification of CSMN and other subcerebral projection neurons.

1.5.1 Fezf2 loss of function studies

Examination of *Fezf2* null mutant mice revealed a complete absence of subcerebral projection neurons upon staining with CTIP2, a transcription factor that labels layer V subcerebral projection neurons[28, 31]. Histological analysis of the *Fezf2*-/- neocortex confirmed all large neurons with pyramidal morphology were missing from layer V. However, this absence of CSMN is not due to cell death or abnormal neuronal migration. Interestingly, despite the absence of projection neurons, which results in a 44% reduction in layer V thickness, the total cortical thickness remains the same in *Fezf2*-/- mice. The decrease in layer V thickness in *Fezf2*-/- mice is compensated by an increase in layer VI thickness, suggesting the possibility that lack of *Fezf2* might result in neural progenitors continuing to generate layer VI neurons instead of switching to produce layer V subcerebral projection neurons, as they normally do[31].

The most striking evidence confirming the complete absence of CSMNs is the lack of long-distance axonal projections from the sensorimotor cortex to

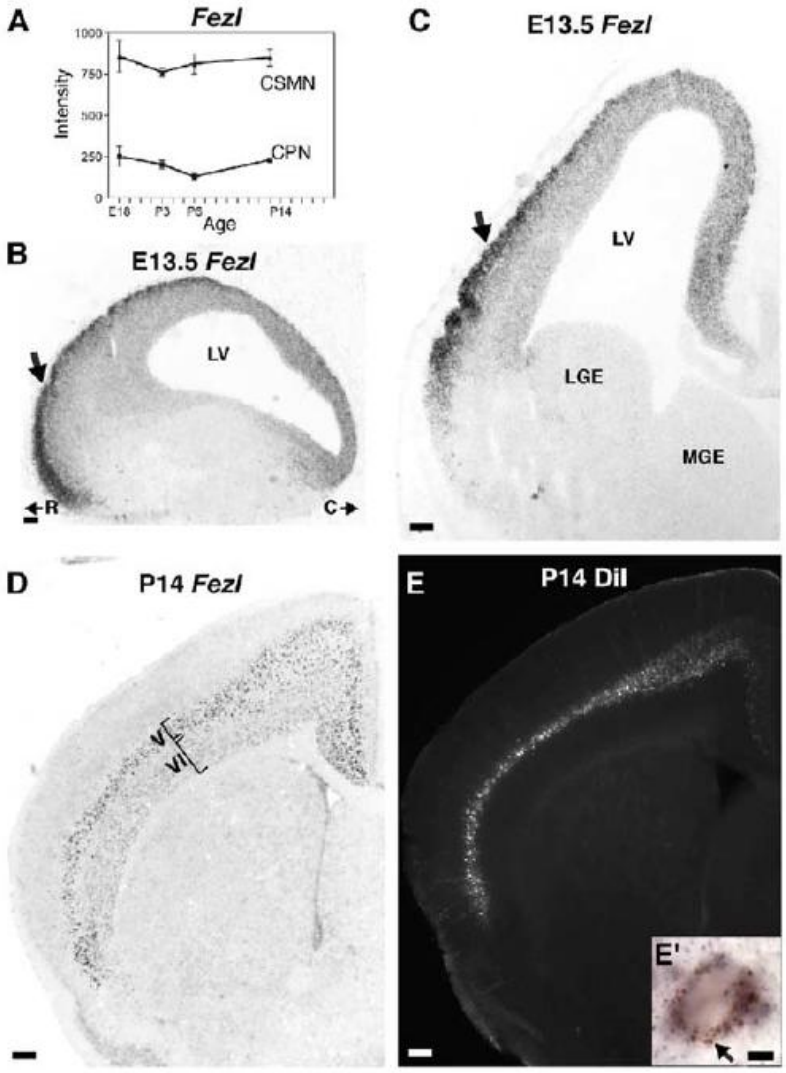


Figure 1.4: *Fezl* Is Expressed in Subcerebral Projection Neurons Throughout Development (A) Microarray expression profiling reveals that *Fezl* is expressed at stable levels in CSMN throughout development (modified from Arlotta et al., 2005).(B and C) In situ hybridization for *Fezl* at E13.5 in sagittal (B) and coronal (C) section, showing highest expression in the cortical plate (arrow).(D and E) In situ hybridization for *Fezl* at P14 (D) indicates that *Fezl* is expressed at high levels in layer V in the same distribution as subcerebral projection neurons, identified by DiI retrograde labeling (E). (E') In situ hybridization for *Fezl* (purple in situ signal) after retrograde labeling of subcerebral projection neurons with DiI (brown precipitate; arrow) confirms expression in subcerebral projection neurons. R, rostral, to C, caudal, indicated in (B). LV, lateral ventricle; MGE, medial ganglionic eminence; LGE, lateral ganglionic eminence.Scale bars, 100 μm (B and C), 250 μm (D and E), 5 μm (E'). (Molyneaux, *et al*. 2005)

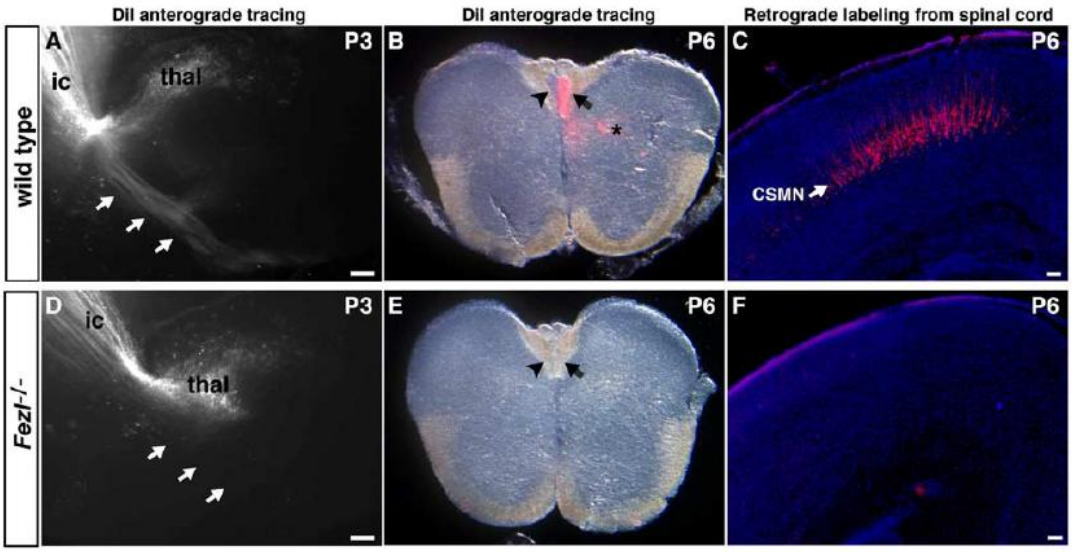


Figure 1.5: Absence of the Corticospinal Tract and Other Subcerebral Projections in *Fezl*$^{-/-}$ Mice (A and D) Anterograde DiI-labeled projections showing absence of any subcerebral axonal tract in *Fezl*$^{-/-}$ brain ([D], arrows), compared to well-defined axon tracts in wild-type brain ([A], arrows). Both thalamic projections via the internal capsule (ic) and retrogradely labeled neurons in the thalamus (thal) are present in *Fezl*$^{-/-}$ mice.(B and E) Cross-sections of P6 cervical spinal cord, showing the corticospinal tract in wild-type dorsal funiculus (B) by both DiI tracing of CSMN axons from the cortex (arrow in [B] points at DiI-labeled corticospinal tract; asterisk indicates fibers growing into dorsal horn), and oblique coherent contrast imaging (arrowhead in [B]), while no corticospinal fibers are detected in *Fezl*$^{-/-}$ mice ([E]; arrow, arrowhead).(C and F) Retrograde labeling of CSMN reveals normal location and numbers in wild-type P6 motor cortex (C), while no neurons are labeled in *Fezl*$^{-/-}$ cortex (F). DAPI staining in blue shows underlying cellular structure. Scale bars, 100 μm (A, C, D, and F). (Molyneaux, *et al*. 2005)

the spinal cord **(Figure 1.5)**. Due to the possibility the lack of *Fezf2* might simply result in projection neurons acquiring an altered form or ectopic location, it is critical to examine if any neurons send subcerebral projections. Anterograde tracing into the sensorimotor cortex revealed intact thalamic projections, but no projections beyond the thalamus to the brainstem or spinal cord in *Fezf2*-/- mice. Retrograde tracing from the spinal cord was also performed to further verify the lack of projections to the spinal cord. No retrogradely labeled neurons were detected in the cortex of *Fezf2*-/- mice, definitively establishing the complete absence of CSMN and subcerebral projection neurons in *Fezf2*-/- mice[31, 32].

Although *Fezf2*-deficient layer V neurons were unable to extend projections to subcortical targets, they were able to adopt an alternate trajectory across the midline through the anterior commissure[32]. This trajectory towards the contralateral hemisphere suggests mutant subcortical projection neurons might adopt a callosal projection neuron fate. Although callosal projection neurons normally reach the opposite hemisphere via the corpus callosum, this commissure is missing in *Fezf2* mutants[32], preventing axons from passing through that route. When the corpus callosum is restored, *Fezf2*-/- cells extend projections across the corpus callosum into the contralateral hemisphere[32]. Additionally, *Fezf2* mutant neurons display the electrophysiological properties of callosal projection neurons[33], indicating *Fezf2* plays a key role in fate determination for subcerebral and callosal projection neurons. Expression of *Fezf2* determines a subcerebral projection

neuron identity, while absence of *Fezf2* results in a callosal projection neuron fate.

1.5.2 Fezf2 gain of function studies

The loss of function experiments detailed above demonstrate that *Fezf2* is necessary for the specification of CSMN and other subcerebral projection neurons in mice. The next important question to address is if *Fezf2* is sufficient to specify the CSMN lineage. To investigate this question, *Fezf2* was overexpressed through *in vivo* electroporation into the neocortical ventricular zone at E13.5, the time when layer VI neurons are born. Electroporation with a control plasmid containing only GFP confirmed cells at this time point give rise to neurons of layer VI. In contrast, cells overexpressing *Fezf2* formed a new, heterotopic layer of neurons beneath the corpus callosum and failed to migrate to layer VI[6] **(Figure 1.6)**. Only cells expressing *Fezf2* at much lower levels were able to migrate into the cortex. While these results suggest overexpression of *Fezf2* results in an inability of cells to migrate, it is consistent with the role of *Fezf2* during normal development. *Fezf2* expression is dramatically upregulated postmigrationally during normal cortical development. Therefore, the ability of cells with low levels of expression to migrate while higher expressing cells fail to migrate indicates that the normal developmental role of *Fezf2* might be recapitulated in these overexpression studies. Although the *Fezf2* overexpressing neurons are ectopically located and therefore presumably would not receive the normal guidance cues from

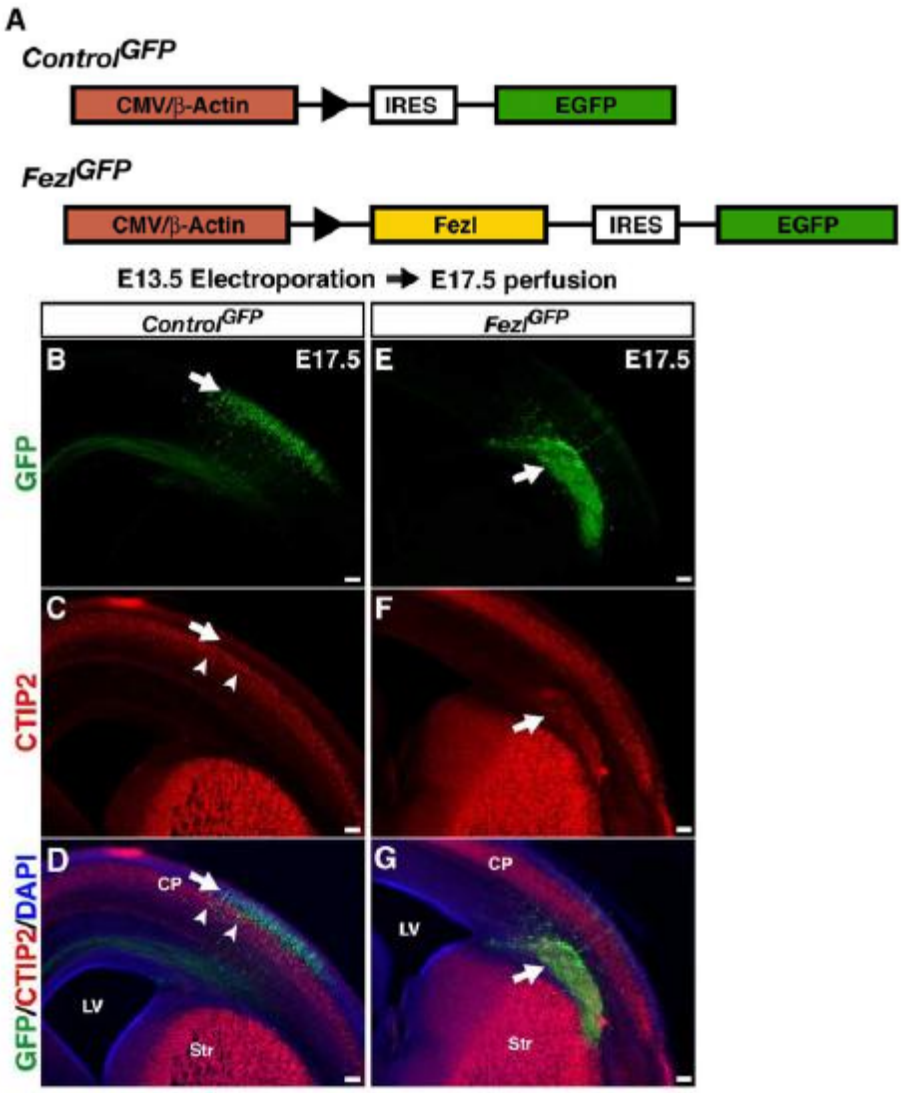


Figure 1.6: Overexpression of *Fezl* Produces CTIP2-Expressing Subcerebral Projection Neurons (A) *Control*^{GFP} and *Fezl*^{GFP} constructs. (B–D) Coronal sections from brains after electroporation of *Control*^{GFP} at E13.5 and perfusion at E17.5. GFP-expressing neurons from the controls are primarily located in layer IV (B, arrows), superficial to CTIP2-labeled subcerebral projection neurons ([C and D], arrowheads). (E–G) Coronal sections after electroporation of *Fezl*^{GFP} at E13.5 and perfusion at E17.5 (E–G). *Fezl*- and GFP-expressing neurons accumulate overwhelmingly in the germinal zone (arrow) beneath the corpus callosum, and virtually all of them express CTIP2 (arrows, [F and G]). (Molyneaux, *et al*. 2005)

the cortex, they are still able to extend axons through the internal capsule toward subcerebral targets[31].

Even more striking is the ability of ectopic *Fezf2* expression in upper layer neurons (layers II/III) to specify projections to locations characteristic of deeper layer neurons, including the CST[33] **(Figure 1.7)**. In the normal developing cortex, progenitors for earlier born neurons have the capacity to form any of the later born neurons. For example, layer VI progenitors have the ability to form not only layer VI neurons, but also neurons of layers II-V. However, progenitors that give rise to later born neurons have lost the capacity to generate earlier born neurons. Therefore, layer II/III progenitors do not have the intrinsic ability to generate layer V neurons that project subcortically to the CST.

The ability of *Fezf2* to induce a corticospinal fate at different stages of cortical development suggests it may also be sufficient to specify a projection neuron identity to progenitors of non-cortical cell fate. To test this, Rouaux and Arlotta expressed *Fezf2* in progenitors of the lateral ganglionic eminence. These progenitors normally give rise to GABAergic medium spiny neurons of the striatum and interneurons of the olfactory bulb, but do not produce glutamatergic cortical projection neurons. Despite their location in the striatum, cells ectopically expressing *Fezf2* expressed proteins specific to cortical neurons, adopted a pyramidal morphology, and were able to extend axons to targets appropriate for projection neurons[34]. In addition, Fezf2-expressing cells did not express markers specific to medium spiny neurons,

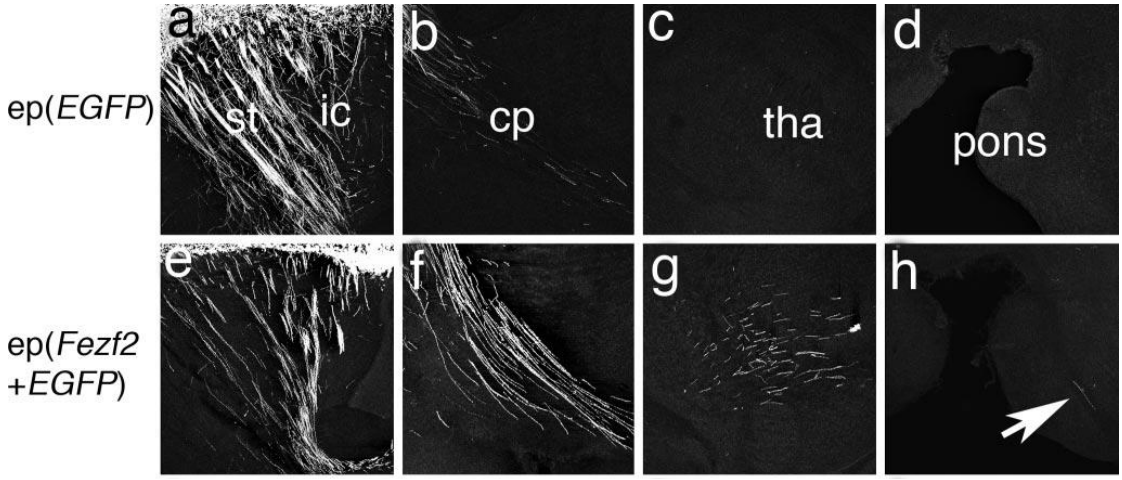


Figure 1.7: Ectopic expression of *Fezf2* or *Ctip2* in wild-type mice is sufficient to alter the axon trajectories of upper-layer neurons, which normally form corticocortical connections, at P5. (*a–d*) Axonal projections of layer 2/3 neurons electroporated with *pCA-EGFP* reveal that some labeled axons occupy the striatum (*a*) and a few descend through the internal capsule and cerebral peduncle (*a* and *b*). No EGFP-labeled axons were observed in the thalamus (*c*) or the vicinity of the pons (*d*). (*e–h*) Layer 2/3 neurons coelectroporated with *pCA-Fezf2* and *pCA-EGFP* extend EGFP-labeled axons through the internal capsule (*e*) and cerebral peduncle (*f*). Labeled axons were abundant in the thalamus (*g*), and a few were present in the CST above the pons (*h*, arrow) (Scale bar, 200 μm.) (Chen, *et al.* 2008)

indicating expression of Fezf2 caused a complete fate switch from GABAergic medium spiny neurons to glutamatergic cortical projection neurons **(Figure 1.8)**. Together with the loss of function analyses discussed above, these data indicate that *Fezf2* is not only necessary, but also sufficient to induce a subcerebral projection neuron identity.

1.6 Fezf2 interacts with other transcription factors to establish cortical projection neuron identities

The loss and gain-of-function studies described above indicate the essential role of Fezf2 in determining subcerebral projection neuron identity. However, Fezf2 does not function alone. Fezf2 positively regulates CTIP2, another zinc finger transcription factor, which then directs the differentiation and axonal projections of these neurons[28, 31-33, 35]. Fezf2 and CTIP2 expression levels correlate in different neuronal types to determine alternative cell fates: high *Fezf2* gene (and thus presumably Fezf2 protein) expression in layer V pyramidal neurons leads to high CTIP2 expression levels, which directs cells to become CSMNs. In contrast, low Fezf2 and CTIP2 expression in layer VI cells results in corticothalamic neuron fate.

Fezf2 also participates in suppressing alternative neuron fate by working in a regulatory network with other transcription factors including Satb2, Tbr1, and Sox5 **(Figure 1.9)**. Fezf2 regulates a fate switch between subcerebral and callosal projection neurons by suppressing Satb2, a chromatin remodeling protein required for the development of callosal projection neurons[33, 36]. In the

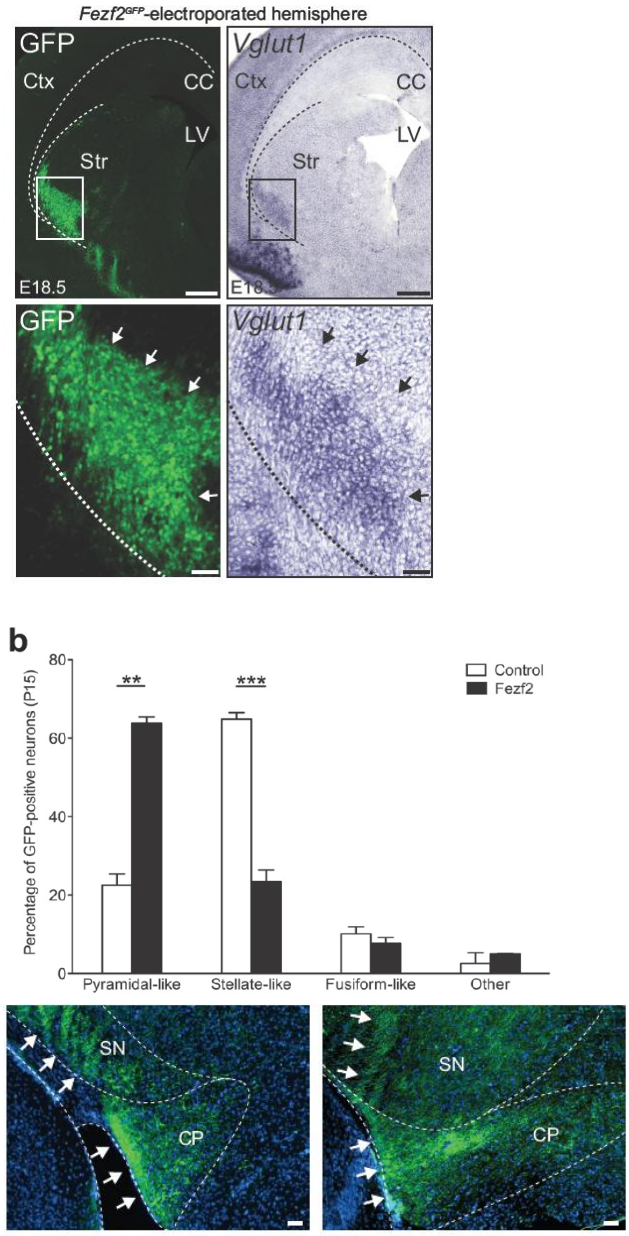


Figure 1.8: *Fezf2* instructs the generation of glutamatergic neurons from LGE progenitors. (**a**) Coronal sections at E18.5 from electroporated brains showing that *Fezf2GFP*-expressing neurons (arrows) also express *Vglut1*. Scale bars, 500 μm and 100 μm in boxed areas. (**b**) Quantification at P15 of GFP-positive neurons that show pyramidal-like, stellatelike and fusiform-like morphology after electroporation of either *ControlGFP* or *Fezf2GFP*. (**c**) Immunocytochemistry for GFP and DAPI at P15 shows *Fezf2GFP*-axonal projections in the the cerebral peduncle and the substantia nigra. SN, Substantia Nigra; CP, Cerebral Peduncle. Scale bar, 100 μm. (Rouaux and Arlotta, 2010)

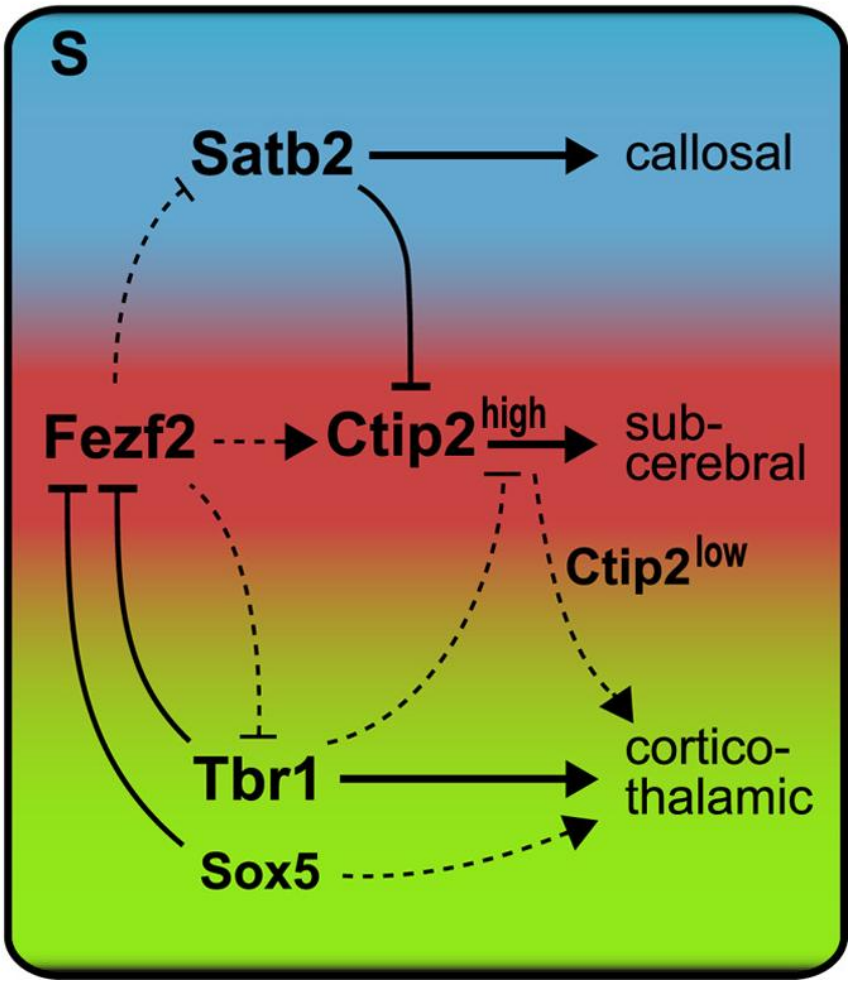


Figure 1.9: Proposed mechanisms for establishing cortical projection neuron identities. The dashed arrow between *Fezf2* and *Ctip2*: an indirect positive regulation. The dashed, bar-end lines from *Fezf2* to *Satb2*, *Fezf2* to *Tbr1*, and *Tbr1* to *Ctip2*: direct inhibition has not been established. The dashed arrows from Sox5 and Ctip2 low demonstrate that they are important for corticothalamic axon development, although corticothalamic axons are still present in *Sox5*$^{-/-}$ and *Ctip2*$^{-/-}$ mice. (McKenna, *et al.* 2011)

absence of Fezf2 expression, cells express Satb2 and adopt a callosal fate. Satb2 then represses subcerebral fates by downregulating CTIP2 expression via direct binding to regulatory regions of the CTIP2 locus, leading to chromatin remodeling[37, 38].

To repress corticothalamic neuron fate, Fezf2 represses Tbr1, a T-box transcription factor that is highly expressed in layer VI neurons and regulates their development[36]. In Fezf2$^{-/-}$ mice, subcerebral neurons ectopically express Tbr1 and send their projections to the thalamus. In return, Tbr1 expression promotes corticothalamic identity by repressing Fezf2 and CTIP2. Both Fezf2 and CTIP2 are upregulated in early-born neurons of Tbr1-/- mice, suggesting Tbr1 represses expression of these genes, either directly or indirectly[39]. Specifically, Tbr1 binds to highly conserved noncoding regions at the 3' end of *Fezf2*, suggesting a possible mechanism of direct inhibition of *Fezf2*[36].

Similar to Tbr1, Sox5 also represses *Fezf2* by binding to an enhancer element downstream of the *Fezf2* gene[40]. Sox5, a member of the SRY-box containing gene family, postmitotically controls the molecular identity, laminar positioning, and axonal connectivity of early born projection neurons. Fezf2 and CTIP2 both show broader expression in early postmigratory subplate, layer VI, and layer V neurons. This broad expression is later refined by Sox5 expression in subplate and layer VI neurons, which results in the layer V-specific enrichment of Fezf2 and CTIP2 expression[40, 41].

1.7 Using developmental cues to generate Fezf2+ CSMN from human embryonic stem cells

Much has been learned about some of the key molecules and pathways that are involved in early neuronal development. Studies conducted in multiple different species have begun to shed light on some basic developmental programs that have been maintained throughout evolution. By applying this knowledge to human cells, the goal is to be able to specifically and efficiently direct cells towards a particular cell fate.

1.7.1 Neural induction during development

The first step towards generating CSMNs from human embryonic stem cells is to provide cues to instruct the cells to adopt a cortical neuronal fate. The first step of neural induction is initiated by molecules that inhibit BMP signaling to form anterior neuroectoderm. One of the key molecules involved in this step is noggin. In the absence of any caudalizing factors, the induced neuroectoderm retains its initial forebrain identity[42].

Following neural induction, the forebrain undergoes patterning along the dorso-ventral axis. Ventral identity is induced by sonic hedgehog (SHH), which is secreted from the ventral neural tube[43]. SHH signaling is important for specification of the two major cortical neuronal populations, pyramidal neurons and interneurons, which are generated from the dorsal and ventral telencephalon, respectively[42, 44-46].

To generate the complex and organized layers of the cortex, further

specification takes place to create the different subtypes of neurons that populate the different layers. This specification occurs in an organized temporal manner with subtypes that populate each layer being generated sequentially[5]. The mechanisms that control this process, however, remain poorly understood.

1.7.2 Directing human embryonic stem cells towards a cortical fate

Although our understanding of the complex processes involved in the specification of the numerous subtypes of cortical neurons remains incomplete, we can apply our current understanding of the key developmental pathways to develop more efficient methods of directing human embryonic stem cells towards a cortical fate. A wide variety of neural induction protocols have been developed, the majority of which yield very low numbers of neural cells, making these protocols insufficient for regenerative medicine. Several neural induction protocols include treatment with noggin to inhibit BMP signaling and increase the efficiency of neural induction[47, 48]. Additionally, the drug SB431542 has been used to inhibit the Lefty/Activin/TGFβ pathways[49]. Used separately, each compound alone improves the efficiency of neural induction; however, the efficiency still remains only approximately 10%. In contrast, when used together, the efficiency of neural induction increases dramatically to greater than 90%[50] **(Figure 1.10)**.

While such a dramatic increase seems quite surprising, a closer look at

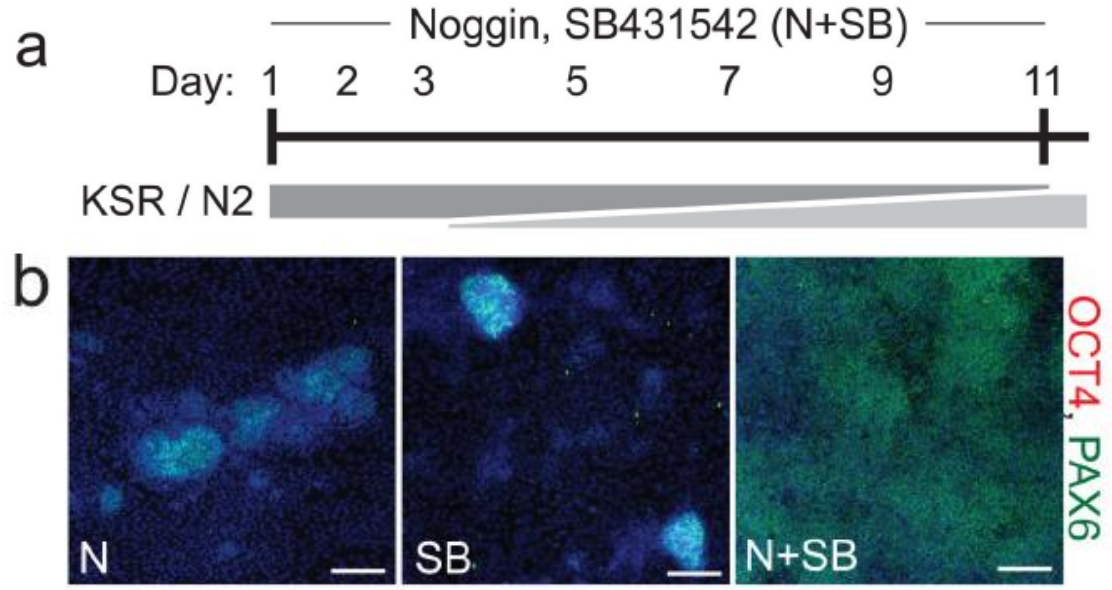


Figure 1.10: Dual SMAD inhibition allows for a highly efficient feeder-free neural induction in adherent cultures in seven days. (a) Differentiation scheme used for achieving neural induction can be achieved with the combination of SB431542, an ALK inhibitor, and Noggin, a BMP inhibitor. (b) The dual SMAD inhibition greatly improves neural differentiation (PAX6 expression, green) to greater than 80%. Infrequent neural differentiation (< 10% $PAX6^+$ cells) can be observed when the single factors are used. (Chambers, *et al*. 2009)

the pathways involved reveals the synergistic action of these two compounds inhibits the major non-neural pathways, allowing the cells essentially only one option: adopt a neural fate. Inhibition of BMP signaling by noggin suppresses differentiation towards the trophoblast[51] and endodermal[50, 52] lineages and promotes neuralization of primitive ectoderm[53]. SB431542-mediated inhibition of activin signaling suppresses mesodermal differentiation[50, 54]. Once a neural identity has been specified, additional patterning factors can then be added to activate or repress certain pathways involved in specification of a particular cell type.

1.8 Conclusion

Human embryonic stem cells offer great promise both as a tool for studying developmental processes and as a limitless source of cells for regenerative medicine. One major hurdle that must be overcome before stem cells can be widely used for regenerative medicine is the ability to efficiently direct them from the embryonic state towards a particular cell type. As our understanding of the mechanisms controlling development and cell type specification increases, we will improve our ability to develop effective therapeutic treatments for patients suffering from many different diseases.

Chapter 1, in part, is a reprint of the material as it appears in Ruby and Zheng, Cell Bioengineering and Tissue Microenvironment, in press. The dissertation author is the primary author of this book chapter. All statements were accurate as of the time of publishing.

1.9 References

1. Finlay BL, Darlington RB. Linked regularities in the development and evolution of mammalian brains. Science. 1995;268:1578-1584.

2. Angevine JB, Jr., Sidman RL. Autoradiographic study of cell migration during histogenesis of cerebral cortex in the mouse. Nature. 1961;192:766-768.

3. Rakic P. Mode of cell migration to the superficial layers of fetal monkey neocortex. J Comp Neurol. 1972;145:61-83.

4. Rakic P. Neurons in rhesus monkey visual cortex: systematic relation between time of origin and eventual disposition. Science. 1974;183:425-427.

5. McConnell SK. Constructing the cerebral cortex: neurogenesis and fate determination. Neuron. 1995;15:761-768.

6. O'Leary DD, Koester SE. Development of projection neuron types, axon pathways, and patterned connections of the mammalian cortex. Neuron. 1993;10:991-1006.

7. Bayer S, & Altman, J. Neocortical Development: Raven Press; 1991.
8. Wonders CP, Anderson SA. The origin and specification of cortical interneurons. Nat Rev Neurosci. 2006;7:687-696.

9. Jones EG. Laminar distributions of cortical efferent cells. In: A. Peters and Jones EG, ed. Cerebral Cortex: Plenum Publishing Corp; 1984:521-553.

10. The Primary Motor Cortex: Upper Motor Neurons That Initiate Complex Voluntary Movements. In: Dale Purves GJA, David Fitzpatrick, Lawrence C Katz, Anthony-Samuel LaMantia, James O McNamara, and S Mark Williams, ed. Neuroscience, 2 ed: Sinauer Associates, Inc.; 2001.

11. Martin JH. The corticospinal system: from development to motor control. Neuroscientist. 2005;11:161-173.

12. Muir GD. Early ontogeny of locomotor behaviour: a comparison between altricial and precocial animals. Brain Res Bull. 2000;53:719-726.

13. Spinal cord injury facts and figures at a glance. J Spinal Cord Med. 2010; 33:439-440.

14. Phukan J, Hardiman O. The management of amyotrophic lateral sclerosis. J Neurol. 2009;256:176-186.

15. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature. 1981;292:154-156.

16. Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A. 1981;78:7634-7638.

17. Robertson E, Bradley A, Kuehn M, et al. Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. Nature. 1986;323:445-448.

18. Doetschman T, Gregg RG, Maeda N, et al. Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature. 1987;330:576-578.

19. Thomas KR, Capecchi MR. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell. 1987;51:503-512.

20. Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. Science. 1998;282:1145-1147.

21. Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007;131:861-872.

22. Yu J, Vodyanik MA, Smuga-Otto K, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007;318:1917-1920.

23. Park IH, Arora N, Huo H, et al. Disease-Specific Induced Pluripotent Stem Cells. Cell. 2008.

24. Hashimoto H, Yabe T, Hirata T, et al. Expression of the zinc finger gene fez-like in zebrafish forebrain. Mech Dev. 2000;97:191-195.

25. Matsuo-Takasaki M, Lim JH, Beanan MJ, et al. Cloning and expression of a novel zinc finger gene, Fez, transcribed in the forebrain of Xenopus and mouse embryos. Mech Dev. 2000;93:201-204.

26. Yang Z, Liu N, Lin S. A zebrafish forebrain-specific zinc finger gene can induce ectopic dlx2 and dlx6 expression. Dev Biol. 2001;231:138-148.

27. Arlotta P, Molyneaux BJ, Chen J, et al. Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. Neuron. 2005;45:207-221.

28. Hirata T, Suda Y, Nakao K, et al. Zinc finger gene fez-like functions in the formation of subplate neurons and thalamocortical axons. Dev Dyn. 2004;230:546-556.

29. Inoue K, Terashima T, Nishikawa T, et al. Fez1 is layer-specifically expressed in the adult mouse neocortex. Eur J Neurosci. 2004;20:2909-2916.

30. Molyneaux BJ, Arlotta P, Hirata T, et al. Fezl is required for the birth and specification of corticospinal motor neurons. Neuron. 2005;47:817-831.

31. Chen B, Schaevitz LR, McConnell SK. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. Proc Natl Acad Sci U S A. 2005;102:17184-17189.

32. Chen B, Wang SS, Hattox AM, et al. The Fezf2-Ctip2 genetic pathway regulates the fate choice of subcortical projection neurons in the developing cerebral cortex. Proc Natl Acad Sci U S A. 2008;105:11382-11387.

33. Rouaux C, Arlotta P. Fezf2 directs the differentiation of corticofugal neurons from striatal progenitors in vivo. Nat Neurosci.13:1345-1347.

34. Chen JG, Rasin MR, Kwan KY, et al. Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. Proc Natl Acad Sci U S A. 2005;102:17792-17797.

35. McKenna WL, Betancourt J, Larkin KA, et al. Tbr1 and Fezf2 Regulate Alternate Corticofugal Neuronal Identities during Neocortical Development. J Neurosci.31:549-564.

36. Alcamo EA, Chirivella L, Dautzenberg M, et al. Satb2 regulates callosal projection neuron identity in the developing cerebral cortex. Neuron. 2008;57:364-377.

37. Britanova O, de Juan Romero C, Cheung A, et al. Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex. Neuron. 2008;57:378-392.

38. Bedogni F, Hodge RD, Elsen GE, et al. Tbr1 regulates regional and laminar identity of postmitotic neurons in developing neocortex. Proc Natl Acad Sci U S A.107:13129-13134.

39. Kwan KY, Lam MM, Krsnik Z, et al. SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. Proc Natl Acad Sci U S A. 2008;105:16021-16026.

40. Lai T, Jabaudon D, Molyneaux BJ, et al. SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. Neuron. 2008;57:232-247.

41. Wilson SW, Houart C. Early steps in the development of the forebrain. Dev Cell. 2004;6:167-181.

42. Gunhaga L, Jessell TM, Edlund T. Sonic hedgehog signaling at gastrula stages specifies ventral telencephalic cells in the chick embryo. Development. 2000;127:3283-3293.

43. Gotz M, Sommer L. Cortical development: the art of generating cell diversity. Development. 2005;132:3327-3332.

44. Schuurmans C, Guillemot F. Molecular mechanisms underlying cell fate specification in the developing telencephalon. Curr Opin Neurobiol. 2002;12:26-34.

45. Sur M, Rubenstein JL. Patterning and plasticity of the cerebral cortex. Science. 2005;310:805-810.

46. Elkabetz Y, Panagiotakos G, Al Shamy G, et al. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes Dev. 2008;22:152-165.

47. Lee H, Shamy GA, Elkabetz Y, et al. Directed differentiation and transplantation of human embryonic stem cell-derived motoneurons. Stem Cells. 2007;25:1931-1939.

48. Smith JR, Vallier L, Lupo G, et al. Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. Dev Biol. 2008;313:107-117.

49. Chambers SM, Fasano CA, Papapetrou EP, et al. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol. 2009;27:275-280.

50. Xu RH, Chen X, Li DS, et al. BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat Biotechnol. 2002;20:1261-1264.

51. D'Amour KA, Agulnick AD, Eliazer S, et al. Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat Biotechnol. 2005;23:1534-1541.

52. Munoz-Sanjuan I, Brivanlou AH. Neural induction, the default model and embryonic stem cells. Nat Rev Neurosci. 2002;3:271-280.

53. Laflamme MA, Chen KY, Naumova AV, et al. Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. Nat Biotechnol. 2007;25:1015-1024.

Chapter 2

Experimental Procedures

2.1 Cell culture of undifferentiated human ES cells

Human ES cells (HUES-9, obtained from Harvard HUES Cell Facility) were cultured on a layer of irradiated mouse embryonic fibroblasts (MEFs, Globalstem) in HUES media as described [55] with minor modifications: Knockout DMEM (Invitrogen/GIBCO) supplemented with 10% Knockout Serum Replacement (GIBCO), 10% Plasmanate (Bayer), 1× (0.1 mM) Non-essential amino acids (GIBCO), 1× (2 mM) Glutamax (Invitrogen), 55 μM β-mercaptoethanol (GIBCO) and 15 ng/mL Human Fibroblast Growth Factor-basic (bFGF, GIBCO). bFGF was added fresh daily to the media. Cells were passaged as needed by trypsinization with 0.05% Trypsin/EDTA (GIBCO). For some applications (e.g. to make DNA or prior to electroporation), cells were cultured feeder-free on Matrigel (BD Biosciences) in MEF-Conditioned Media (MCM), supplemented with 15 ng/mL bFGF. MCM was made by adding HUES media to plated irradiated MEFs for 20-24 hours. All experimental procedures were approved by the UCSD Embryonic Stem Cell Research Oversight Committee and Institutional Review Board.

2.2 Targeting vector construction

The configuration of the targeting vector is described in Results (Fig. 1-2A). HUES-9 genomic DNA was used as the template to PCR amplify the homology arms using the Expand High Fidelity PCR kit (Roche). Genomic PCR products and multiple clones of the homology arms were sequenced. The clones that showed the least amount of discrepancy with the corresponding

genomic PCR product were used to construct the targeting vector. Out of six single nucleotide polymorphisms (SNPs) mapped to the h*Fezf2* homology arms by the International HapMap Project (www.hapmap.org), two showed variations between the two HUES-9 alleles. There are two additional nucleotides that showed allelic variations in HUES-9 that are not mapped in HapMap. Thus, there are a total of four confirmed SNPs between the two HUES-9 alleles within the total region of homology. In addition, there are 10 single nucleotides that showed discrepancy between the targeting vector and the sequenced human genome that currently we cannot exclude as SNPs but are more likely due to PCR errors by the Expand High Fidelity PCR kit. Detailed sequence information will be provided upon request.

2.3 Transfection and establishment of targeted clones

In a typical experiment, HUES-9 cells were cultured feeder-free on Matrigel to 80-90% confluence. Cells were trypsinized and passed through a 40 μm cell strainer (BD Falcon) to create a single cell suspension, which was routinely verified under a microscope. Cells were centrifuged at 1000 rpm for 4 min at room temperature and a small sample was counted using a hemocytometer. Typically ~80% of trypsinized cells were single cells. Cells were resuspended in 900 μL of HUES-9 media and incubated with the appropriate amount of DNA at 37°C for 5 minutes. The cell suspension was then transferred to a Gene Pulser cuvette (0.4 cm gap) and electroporated with Gene Pulser II (Bio-Rad) at 320V, 200 μF. The cuvette was incubated at 37°C for 5-10 minutes

to allow cell recovery, and then plated on neomycin-resistant MEFs. Geneticin (GIBCO) selection was started at 24 hours at 50 µg/mL. FIAU ((2'-deoxy-2'-fluoro-ß-D-arabinofuranosyl)-5-iodouracil) selection, when used, was started at 48 hours at 200 nM. Selection was applied for 8-10 days until cell death ended. Selection was removed 1 day prior to colony picking to allow cells to recover. Drug-resistant colonies were picked and, after one passage, grown in duplicate: on gelatin or Matrigel plates to harvest DNA and on MEFs to maintain cells in culture. For transient Cre expression, 1×10^7 cells were electroporated with 15 µg of Cre plasmid POG231 [56].

2.4 Detection of targeted clones by PCR analysis

Clones were grown feeder-free to confluence to make DNA. Cells were lysed in DNA lysis buffer (10mM Tris pH 7.5, 10 mM EDTA pH 8.0, 10mM NaCl, 0.5% Sarcosyl (N-Lauroylsarcosine), and 1 mg/mL Proteinase K) overnight at 60°C in a humidified chamber as described [57]. DNA was precipitated using 75 mM NaCl in cold ethanol for 15-30 minutes, washed with 70% ethanol, and resuspended in 0.1× TE. PCR was performed using the Expand High Fidelity PCR kit supplemented with 5% DMSO. PCR products from positive clones were sequenced to verify correct targeting.

2.5 Detection of targeted clones by Southern blot analysis

5 µg DNA was digested overnight at 37°C with the appropriate restriction enzyme, separated on a 0.7% agarose gel and transferred to a

Hybond-XL nylon membrane (Amersham). Membranes were prehybridized in hybridization buffer (1.5× SSPE, 1% SDS, 0.5% fat-free milk, and 40 µg/mL herring sperm DNA (Promega)) for 1 hour rotating at 65ºC. Probes were labeled with the Prime-It II Random Primer Labeling Kit (Stratagene) and EasyTides a-P^{32}dCTP (Perkin Elmer), purified on Micro Bio-Spin columns (BioRad), and then incubated with the membrane with rotation overnight at 65ºC. Membranes were washed briefly three times with low stringency washing buffer (2× SSC, 1% SDS), followed by one 20 minute wash in low stringency buffer and 2 washes in high stringency buffer (0.2× SSC, 0.1% SDS) rotating at 65ºC. Membranes were placed in heat-sealed pouches and exposed to Kodak BioMax MS Film at -80ºC.

2.6 Karyotype analysis of targeted clones

Clones were grown feeder-free to 70% confluence and submitted to the UCSD Medical Genetics facility for karyotype analysis. Clone G6 was electroporated at passage 35 and karyotyped at passage 47. Clones B4 and E11 were electroporated at passage 26 and karyotyped at passage 32. Clone B4 after Cre excision was karyotyped at passage 39.

2.7 Embryoid body (EB) formation and initial neural differentiation

HUES-9-h*Fezf2* cells were grown for one passage on Matrigel, fed with MEF conditioned media and 15 ng/mL bFGF. One day prior to EB formation, feeder-free cells were pre-treated with 10 µM Rho-associated kinase (ROCK)

inhibitor Y-27632 (EMD Chemicals). EBs were made as described [58] with minor modifications. Cells were plated in a V-bottom 96-well plate at a final concentration of 5,000 cells/well in EB media (Knockout DMEM supplemented with 1× N2, 2% B27, 1× Glutamax, 1 μM non-essential amino acids (NEAA), and 1 μM β-mercaptoethanol) supplemented with 10 μM Y- 27632. After 2 days, EBs were transferred to an ultra-low binding 6-well plate in EB media.

After another 4 days, EBs were transferred to a new ultra-low binding 6-well plate in neural induction media (DMEM/F12 with 1× N2, 1 μM NEAA, and 1 μg/mL heparin) supplemented with 20 ng/mL bFGF. Two days later, EBs were transferred to 6-well plates pre-treated with either laminin or Matrigel and plated at 20 aggregates/well. Once adhered, cells were maintained in neural induction media supplemented with 100 ng/mL bFGF. Rosette-like structures formed within 5 days. At 7 days post-plating, rosettes were isolated with dispase as described [59] and plated onto an untreated 6-well plate in neural precursor media (Neurobasal media supplemented with 1× Glutamax, 1× Pen/Strep, 1× B27 (-RA) and 20 ng/mL bFGF). After 6 days, neural precursors were lifted from the plate with Accutase (Sigma), plated on coverslips coated with growth factor reduced Matrigel (GFRM), and maintained in neural differentiation media (Neurobasal media supplemented with 1× Glutamax, 1× Pen/Strep, 2% B27 (-RA) and 1× N2). After one week of differentiation coverslips were fixed and stained.

2.8 Immunocytochemistry

Cells were cultured feeder-free. All staining steps were carried out at room temperature except where indicated otherwise. Cells went through the following steps: 1) fixed with 4% paraformaldehyde for 10 minutes; 2) washed 3× with 1× PBS (phosphate-buffered saline); 3) incubated in blocking solution (5% normal goat serum with 1% Triton-X 100 in PBS) for 30 minutes, 4) incubated with the primary antibody diluted in blocking solution overnight at 4°C; 5) washed 3× with PBS; 6) incubated with the secondary antibody for 90 minutes; 7) washed 3× with PBS, incubated with DAPI for 5 minutes and washed 3× with PBS and 1× with milli-Q H_2O; 8) cells on coverslips were mounted on slides using Fluoromount-G (Southern Biotech). Cells in culture plates were stored in Fluoromount G. Images were captured and analyzed with an Axio Imager M1 fluorescent microscope (Zeiss) equipped with an Axiocam MRm digital imaging system.

2.9 Dual SMAD inhibition neural differentiation

Cells were differentiated as described[50] with modification. HUES-9-h*Fezf2* cells were grown for one passage on Matrigel, fed with MEF conditioned media and 15 ng/mL bFGF until ~80% confluence. Cells were trypsinized until completely dissociated and plated onto GFRM at a density of 20,000 cells/cm^2 in N2B27 media supplemented with 15ng/mL bFGF and 10μM ROCK inhibitor. Once cells reached >90% confluence, bFGF was removed and patterning factors added. Cells were patterned as follows: dual SMAD

inhibitors SB431542 (5μM) and LDN-193189 (100μM, Stemgent) were added from days 1-10; cyclopamine (2μM, Calbiochem) was added on days 2-10.

2.10 First-strand synthesis for RT-PCR

Total RNA was harvested from cells using the RNeasy Plus Mini Kit (Invitrogen). In cases where the starting sample was very small (i.e. after FACS), total RNA was harvested using the NucleoSpin RNA XS kit (Macherey-Nagel). In both cases, RNA was eluted in RNase-free water and stored at -20°C. First-strand cDNA was synthesized using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen).

2.11 Fluorescence activated cell sorting (FACS)

Differentiated HUES-9-h*Fezf2* cells were pre-treated one hour with 10μM ROCK inhibitor. Cells were treated with Accutase for a total of 40 minutes. For the last 10 minutes of Accutase treatment, 100U/mL DNase was added. Cells were collected by centrifugation and resuspended at 2-3M cells/mL in sort buffer (Ca/Mg free HBSS with 20% FBS). To create a single cell suspension, cells were passed through a 40μM cell strainer. To remove dead cells during sorting, 25ug/mL propidium iodide (Roche) was added to the cells. Cells were put on ice and kept at 4°C throughout the entire sorting process. Wild type differentiated HUES-9 cells were used as a negative control to set sorting gates. After sorting, RNA was harvested from cells using the NuceloSpin RNA XS kit (Macherey-Nagel) or cultured further on GFRM.

2.12 Whole genome expression microarray analysis

Total RNA was harvested from FACS sorted HUES-9-*hFezf2* cells on day 19 of differentiation. RNA was submitted to the Biomedical Genomics laboratory (BIOGEM) at UCSD. Concentration and quality of the RNA was analyzed using the Agilent BioAnalyzer. Total RNA was labeled according to the Illumina labeling protocol. Prior to hybridization, the quality of the cRNA was analyzed. After hybridization and washing, the bead chip was scanned according to Illumina's specifications and the resulting image files and data analysis files were deposited on the BIOGEM server. The quality of the hybridizations and overall experimental performance was determined by a visual inspection of the raw scanned data.

Chapter 2, in part, is a reprint of the material as it appears in Ruby and Zheng, Stem Cells 2009. The dissertation author is the primary author of this paper. All statements were accurate as of the time of publishing.

2.13 References

1. Cowan CA, Klimanskaya I, McMahon J, et al. Derivation of embryonic stem-cell lines from human blastocysts. N Engl J Med. 2004;350:1353-1356.

2. O'Gorman S, Dagenais NA, Qian M, et al. Protamine-Cre recombinase transgenes efficiently recombine target sequences in the male germ line of mice, but not in embryonic stem cells. Proc Natl Acad Sci U S A. 1997;94:14602-14607.

3. Ramirez-Solis R, Rivera-Perez J, Wallace JD, et al. Genomic DNA microextraction: a method to screen numerous samples. Anal Biochem. 1992;201:331-335.

4. Burridge PW, Anderson D, Priddle H, et al. Improved human embryonic stem cell embryoid body homogeneity and cardiomyocyte differentiation from a novel V-96 plate aggregation system highlights interline variability. Stem Cells. 2007;25:929-938.

5. Zhang SC, Wernig M, Duncan ID, et al. In vitro differentiation of transplantable neural precursors from human embryonic stem cells. Nat Biotechnol. 2001;19:1129-1133.

6. Chambers SM, Fasano CA, Papapetrou EP, et al. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol. 2009;27:275-280.

Chapter 3

Targeting human *Fezf2* in HUES-9 cells

3.1 Abstract

Genetic modification is critical for achieving the full potential of human embryonic stem (ES) cells as a tool for therapeutic development and for basic research. Targeted modifications in human ES cells have met with limited success due to the unique culture conditions for many human ES cell lines. The HUES lines of human ES cells were developed for ease of manipulation and are gaining increased utility in stem cell research. We tested conditions for gene targeting via electroporation in the HUES-9 human ES cell line and demonstrate here successful gene targeting at the gene encoding Fezf2 (also known as Fezl), a transcription factor involved in corticospinal neuron development. With a targeting strategy involving positive and negative selection that is applicable to all genes, we observed a gene targeting frequency of ~1.5% at *Fezf2*, a gene not expressed in human ES cells. We found that conditions developed for gene targeting in mouse ES cells can be readily adapted to HUES cells with few key modifications. HUES-9 cells exhibit an intrinsically high efficiency of clonal expansion and sustain electroporation-based gene targeting procedures without any significant loss of pluripotency marker expression or karyotypic stability. Thus, human ES cell lines adapted for enzymatic passage and efficient clonal expansion can be highly amenable to genetic modifications, which will facilitate their application in basic science and clinical development.

3.2 Introduction

The derivation of mouse embryonic stem (ES) cells and the development of gene targeting technology in mouse ES cells via homologous recombination have revolutionized ways of deciphering mammalian gene function *in vivo* [60]. The advent of human ES cells, on the other hand, has offered enormous potential for transplantation therapies and regenerative medicine as well as for advancing basic understanding of human biology [21]. Cell transplantations with human ES cells or their derivatives hold great promise for treating conditions ranging from diabetes, Parkinson's disease, amyotrophic lateral sclerosis to spinal cord injury. As a basic research tool, human ES cells allow for *in vitro* and *in vivo* characterization of human gene function in development, physiology and disease biology.

To achieve the full potential of human ES cells as a tool for both therapeutic development and basic research, a key element lies in our ability to genetically modify these cells [61]. In particular, gene targeting via homologous recombination, which has been well established in mouse ES cells and represents one of the most powerful ways to assess gene function [62], will play an important role in human ES cell research for several reasons. First, as in mouse ES cells, gene targeting can be used to assess human gene function either *in vitro* or *in vivo*, with the latter in animal transplantation models. Second, gene targeting can be used to repair faulty genes in patient-specific human ES cells or induced pluripotent stem (iPS) cells [22-24] prior to transplantation therapies. Finally, gene targeting can be used to track cells

(e.g. neurons), cellular processes (e.g. axons) or an endogenous protein with a knockin reporter (e.g. a gene encoding a fluorescent protein such as GFP) or protein tag (e.g. FLAG) [61, 63]. This homologous recombination-based strategy has distinct advantages over conventional or virally mediated transgenics via random integration in that the knockin reporter or tagged protein expression will be more likely to faithfully recapitulate that of the endogenous gene or protein. This is a particularly critical issue for *human* ES cells because unlike in *mouse* ES cells, the expression pattern of the reporter gene cannot be readily verified *in vivo* with *human* ES cells. Knockin cellular reporters would greatly facilitate the monitoring of human ES cell differentiation *in vitro* and in animal transplantation models; in addition, knockin protein tags would allow for the study of endogenous gene products using advanced genomic and proteomic tools.

In spite of these important utilities, our ability to perform gene targeting in human ES cells is still immature compared to that in mouse ES cells [61]. Success with homologous recombination in human ES cells with the time-tested, electroporation-based method has only been reported in a limited number of cases [64-68]. Among these reports, all except one either targeted the *Hprt* locus in a male human ES cell line [65, 67, 68] or targeted a gene that is expressed in human ES cells [66, 68], using strategies that are not readily transferable to other genes.

The HUES lines of human ES cells were developed under conditions for ease of manipulation and well defined culture media [55], lending to their

increased utility [69]. The amenability of these cells to enzymatic passage with trypsin makes them ideal vehicles to test gene targeting methods that were refined and time tested in mouse ES cells. We chose the human *Fezf2* gene (also known as *Fezl*, referred to as h*Fezf2* from here on) to test gene targeting in a HUES line of human ES cells. The mouse orthologue of hFezf2, a forebrain embryonic zinc-finger-like transcription factor, is expressed throughout corticospinal neuron development and plays a critical role in this process [31, 32]. A fluorescent reporter targeted to the h*Fezf2* locus will be a useful tool in guiding the development of a differentiation protocol for corticospinal neurons, a major therapeutic target for neurological conditions such as amyotrophic lateral sclerosis and spinal cord injury. It will also be useful for *in vivo* tracking of HUES cell derivatives that have taken the fate of corticospinal and related neurons in animal transplantation models. We document here successful gene targeting at the h*Fezf2* gene in HUES-9 cells. We report that HUES-9 cells can be readily modified genetically via homologous recombination with a robust protocol similar to that for mouse ES cells without compromising pluripotency marker expression or karyotypic stability.

3.3 Results

To achieve robust gene targeting in HUES-9 cells, there are three key technical measures that need to be achieved: 1) a reasonably high efficiency of stable transfection with appropriate positive selection; 2) an acceptable *relative* gene targeting frequency (the ratio of the number of targeted clones

over that of stable transfectants); 3) a streamlined protocol to screen a large number of human ES cell clones for targeted events. Our strategy was to adapt a well-established method for AB2.2 mouse ES cells [70] to the HUES-9 cells since they share a high tolerance for enzymatic passage with trypsin [55]. We have tested each of the three key elements, which we describe below, followed by a description of gene targeting at h*Fezf2*.

3.3.1 Stable transfection via electroporation in HUES-9 cells

The robustness of an ES cell line for stable transfection relies on its tolerance for clonal expansion. Following electroporation of a targeting vector, drug resistant colonies are clonally expanded before a molecular screen for targeted clones can be conducted. Following electroporation of a Cre plasmid to excise a *loxP*-flanked selection marker, a commonly performed procedure during more complex gene targeting schemes, cells are plated at a low density for clonal expansion in the absence of drug selection. We first tested the efficiency of clonal expansion for HUES-9 cells following enzymatic passage with trypsin but without electroporation. When plated at a low density, ~24% of individual HUES-9 cells (469, 497, and 492 out of 2000 cells plated on a 10-cm plate in three experiments) grew into single colonies, indicating a robust capability for HUES-9 cells to undergo clonal expansion. It should be noted that this high clonal expansion efficiency was observed under standard culture and passage conditions for HUES-9 (e.g. with standard HUES media), without the application of any additional factors such as the

ROCK inhibitor Y-27632 or a neurotrophin cocktail [71, 72].

The Neomycin resistance gene (*Neo*), which confers resistance to Geneticin (also known as G418), is the most commonly used positive selection marker in mouse ES cells. Guided by the standard Geneticin concentrations for mouse ES cells at 125 – 200 μg/mL, we tested the killing efficacy of Geneticin on *wild type* HUES-9 cells ranging between 50-150 μg/mL. Cell death was seen as early as 24 hours, whereas it typically takes 2-3 days for Geneticin selection to become effective in mouse ES cells. After 8-10 days of drug selection, 50 μg/mL of Geneticin was sufficient to kill all wild type HUES-9 cells.

We next tested electroporation conditions for stable transfection based on our mouse ES cell protocol [70] and a protocol developed for the H1.1 human ES cells [68]. Our mouse ES cell protocol (230 V, 500 μF in 1× PBS) routinely gives 500 – 2000 Geneticin resistant colonies per 10^7 electroporated AB2.2 cells. Interestingly, differing voltage (230 vs. 320 V), capacitance (200 vs. 500 μF) or electroporation medium (1× PBS vs. 1× HUES culture media) did not lead to an appreciable difference in the efficiency of stable transfection for HUES-9 cells (data not shown). Typically, 200 - 500 Geneticin resistant colonies per 1-1.5 × 10^7 electroporated HUES-9 cells were obtained with 10 μg of linearized DNA carrying a Neomycin resistance gene (*Neo*). Although this number is about 3-4 fold lower than that for AB2.2 mouse ES cells, it is sufficient for gene targeting experiments because only 100-400 colonies are routinely screened for targeted events. Thus, our protocol produces a sufficient number of stable

transfectants for gene targeting.

3.3.2 Negative selection in HUES-9 cells to increase the relative gene targeting frequency

The second key measure towards successful gene targeting is the *relative* gene targeting frequency, expressed as the percentage of the number of targeted clones among the total number of drug resistant colonies (i.e. stable transfectants). The *relative* gene targeting frequency can be enhanced with negative selection [73]. By selectively killing random integrants, which, unlike the targeted clones, tend to retain the negative selection marker, negative selection enriches targeted clones among stable transfectants. The Herpes Simplex Virus Thymidine Kinase (HSV-TK) gene is a commonly used negative selection marker. Mouse ES cells that express HSV-TK can be selected against with FIAU [1-(2'-deoxy-2'-fluoro-ß-D-arabinofuranosyl)-5-iodouracil] at 200 nM.

We first tested FIAU toxicity at 200 nM – 1 μM on wild type HUES-9 cells, which lack the HSV-TK gene and should not be sensitive to FIAU. Starting at 400 nM, FIAU became toxic to HUES-9 cells, with increasing toxicity at higher concentrations (**Figure 3.1A**). We then tested the efficacy of FIAU negative selection in a mixed population of HUES-9 cells stably transfected with a vector carrying both Neo and HSV-TK, as would be performed during gene targeting. At 200 nM, ~60% of the colonies were killed by FIAU (**Figure 3.1B**). Thus, ~60% of the random integrants are expected to be killed at this

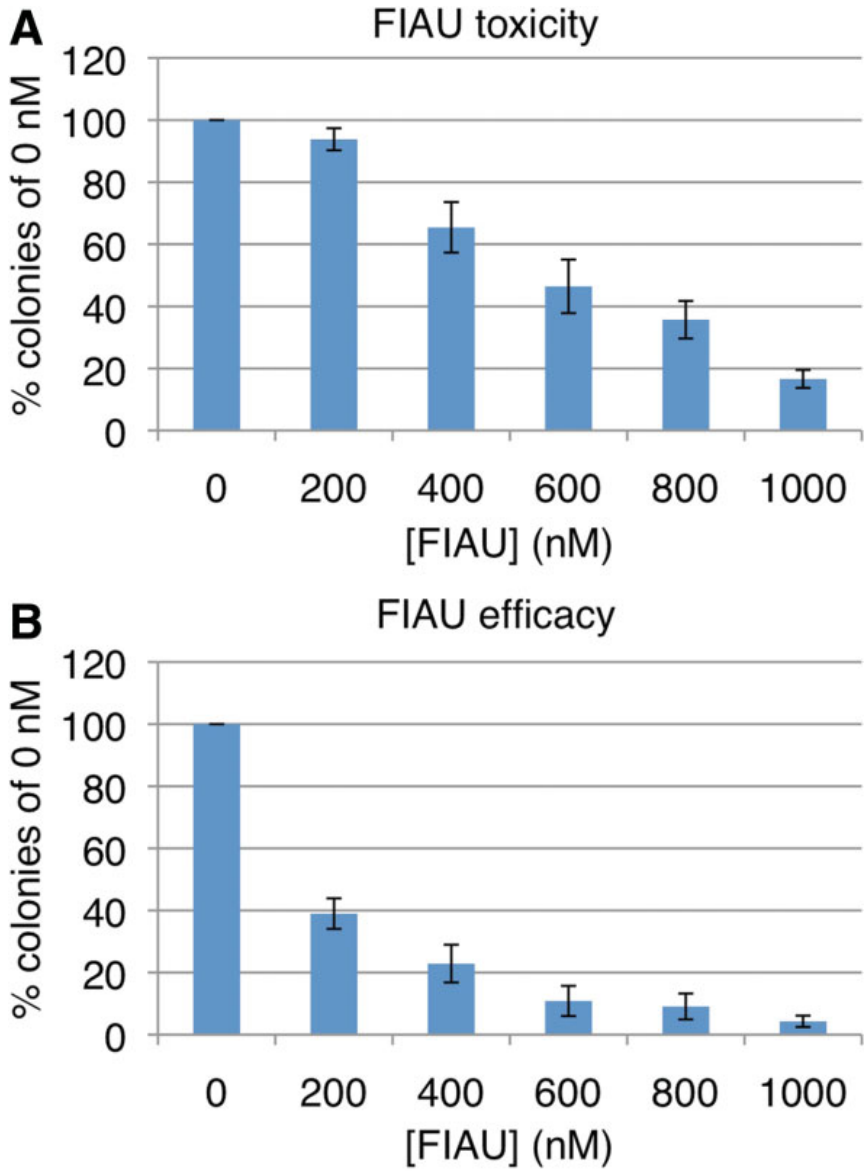


Figure 3.1. Testing of negative selection drug FIAU. HUES-9 cells were treated with various concentrations of FIAU for 8-10 days. Error bars, SEM (n = 3). (**A**) Toxicity of FIAU was tested on wild type cells. (**B**) Efficacy of FIAU was tested on a mixed population of cells stably transfected with a *Neo-HSV-TK* cassette.

concentration in a gene targeting experiment, which translates to ~2.5 fold enrichment for targeted clones. This is in line with the typical 3-5 fold enrichment associated with FIAU negative selection in mouse ES cells [74]. Thus, negative selection with FIAU at 200 nM on HUES-9 cells would provide effective enrichment without apparent toxicity.

3.3.3 A streamlined protocol to detect targeted clones in HUES-9 cells

The third key element in a robust gene targeting method is efficient screening of individual colonies for targeted events. We experimented with screening colonies in a 96-well format with either PCR or Southern blot. Both strategies worked for HUES-9 cells (data not shown). However, PCR is considerably more time efficient (< 1 day vs. typically ≥ 3-5 days for Southern blot) and is now our method of choice for initial screens. Positive clones identified by PCR are then expanded and frozen for long-term storage before verification with Southern blot analysis. Rapid initial screens proved crucial for HUES-9 cells because in our hands freezing and thawing HUES-9 cells in 96-wells significantly reduced their viability and consequently individual clones were kept in culture while the initial screen was being conducted.

3.3.4 Gene targeting at the hFezf2 locus in HUES-9 cells

We constructed a replacement gene targeting vector for h*Fezf2* with the following features (**Figure 3.2A**): 1) 4.2 and 4.1 kb of homology for the left and right arm respectively; 2) homology arms amplified by high-fidelity PCR

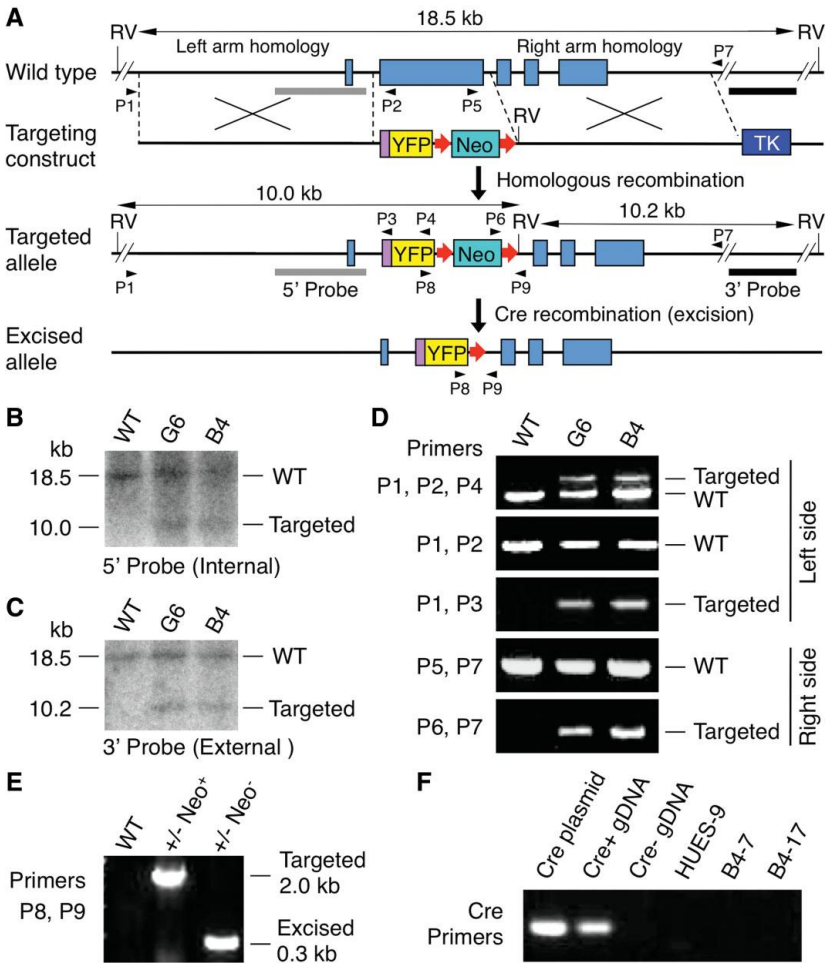


Figure 3.2. Gene targeting via homologous recombination at the h*Fezf2* locus in HUES-9 human ES cells. (**A**) Gene targeting and Cre excision strategy with Southern blot and PCR schemes. Unmarked blue boxes, exons of h*Fezf2*; Purple box, splice acceptor; YFP, enhanced Yellow Fluorescent Protein gene; Neo, Neomycin resistance gene; TK, Herpes Simplex Virus Thymidine Kinase gene; red arrow, loxP site; RV, EcoRV; P1 – P9 (arrowheads), PCR primers used to distinguish wild type, targeted and Cre excised alleles. (**B, C**) Representative Southern blots on targeted clones with either a 5' internal probe (**B**) or a 3' external probe (**C**). WT, wild type; G6 and B4, two targeted clones; kb, kilobase. (**D**) Representative PCR analysis on targeted clones. 3-primer PCR (top panel) was used in the initial screen; 2-primer PCRs (lower panels) followed by sequencing analysis were used to confirm correct targeting at both the left (5') and right (3') sides of homology. (**E**) PCR analysis to confirm Cre-mediated excision of the *Neo* cassette from a targeted clone. +, wild type allele; -, targeted allele before (Neo^+) or after (Neo^-) Cre excision. (**F**) PCR analysis to verify the absence of Cre plasmid integration in excised clones. Cre+ and Cre- g-DNA, Cre positive and negative mouse genomic DNA respectively (as controls); B4-7 and B4-17, two clones following Cre excision of the *Neo* cassette from the h*Fezf2* targeted clone B4.

from HUES-9 genomic DNA (see Materials and Methods); 3) exon 2 (carrying the ATG start codon) deleted to create a null allele for loss of function analysis; 4) a knockin reporter, the enhanced yellow fluorescent protein gene (*EYFP*), to track h*Fezf2* expressing cells; 5) a *Neo* positive selection marker flanked by two *loxP* sites so *Neo* can be excised by Cre expression in the event that it interferes with EYFP expression or a second round of gene targeting with the same vector is desired; 6) an *HSV-TK* gene as the negative selection marker.

A summary of all the electroporations performed to target h*Fezf2* with or without FIAU negative selection is shown in **Table 3.1**. In total, we obtained four targeted clones based on initial screens with Southern blot (1 clone) or PCR (3 clones) (data not shown). The integrity of homologous recombination at both the left (5') and right (3') arms of homology was verified by Southern blot with a 5' internal probe and a 3' external probe respectively (**Figure 3.2B, C**). This was further confirmed by a series of PCR analyses across both of the homology arms (**Figure 3.2D**) and sequencing analysis of the PCR products (data not shown). In addition, no other (unwanted random) integration events were detected with the 5' internal probe on the Southern blot in the targeted clones (**Figure 3.2C** and data not shown). Among the four targeted clones, FIAU negative selection was applied for two of them, which significantly enhanced the *relative* gene targeting frequency (expressed as the percentage of targeted clones among the stable transfectants successfully screened), a key measure in gene targeting, from ~0.3% (2 out of 675 clones

Table 3.1. Summary of electroporation experiments to target hFezf2 in HUES-9 cells

Exp No.	*n* of cells electroporated (millions)	μg DNA	*n* of drug-resistant colonies	*n* of clones screened	*n* of clones screened successfully	Electroporation efficiency (stable transfection)	*n* of targeted clones	Relative gene targeting frequency
Without FIAU selection								
1	12.5	10	183	58	149	1.5E-05	0	
2	15	10	234	68		1.6E-05	0	
3	15	20	246	67		1.6E-05	1	
4	15	10	449	192	150	3.0E-05	1	
5	15	10	499	192	182	3.3E-05	0	
6	15	10	265	162	90	1.8E-05	0	
7	15	10	368	162	104	2.5E-05	0	
Total	102.5	–	2,244	901	675	2.2E-05	2	0.3%
With FIAU selection								
8	10.5	15	192	107	58	1.8E-05	2	
9	10.5	10	187	183	72	1.8E-05	0	
Total	21	–	379	290	130	1.8E-05	2	1.5%

In calculating the relative gene targeting frequency, only clones successfully screened by polymerase chain reaction or Southern blot analysis (as indicated by the presence of at least the wild-type band) are counted.
Abbreviation: FIAU, 1-(2′-deoxy-2′-fluoro-§-D-arabinofuranosyl)-5-iodouracil.

successfully screened without FIAU selection) to ~1.5% (2 out of 130 clones successfully screened following FIAU selection). This level of enrichment for targeted clones (~5×) is consistent with our experience with mouse ES cells (usually 3-5×) [74] and is slightly higher than – but still comparable to – that predicted from our FIAU tests on HUES-9 cells (~2.5×, see above).

To determine whether targeted clones still express features of pluripotent stem cells, we examined the expression of several pluripotency markers by immunocytochemistry. Importantly, the expression of Oct4, Nanog, SSEA-4, TRA-1-60 and TRA-1-81 was readily detected in the targeted clones (**Figure 3.3**), indicating that these cells remained undifferentiated. Consistent with the expectation that h*Fezf2* is not expressed in ES cells, the expression of the EYFP knockin reporter was not detected in targeted clones by direct fluorescence (data not shown). This is not due to the failure of *EYFP* to serve as a reporter gene since upon neural differentiation, EYFP is expressed in a subset of the cells (see below).

To determine whether our gene targeting manipulations led to gross chromosomal abnormalities, we carried out karyotype analysis on three targeted clones (**Figure 3.4**). Two clones (G6, E11) exhibited the same normal karyotype as the parental HUES-9 cell line (with a naturally occurring pericentric inversion on Chromosome 9 that is observed in ~2% of the population with no known associated clinical abnormality) [55]. Most of the cells from the third clone (B4) also exhibited parental karyotype. However, a small percentage of the cells (15%; 3 out of 20 cells analyzed) from this clone carry

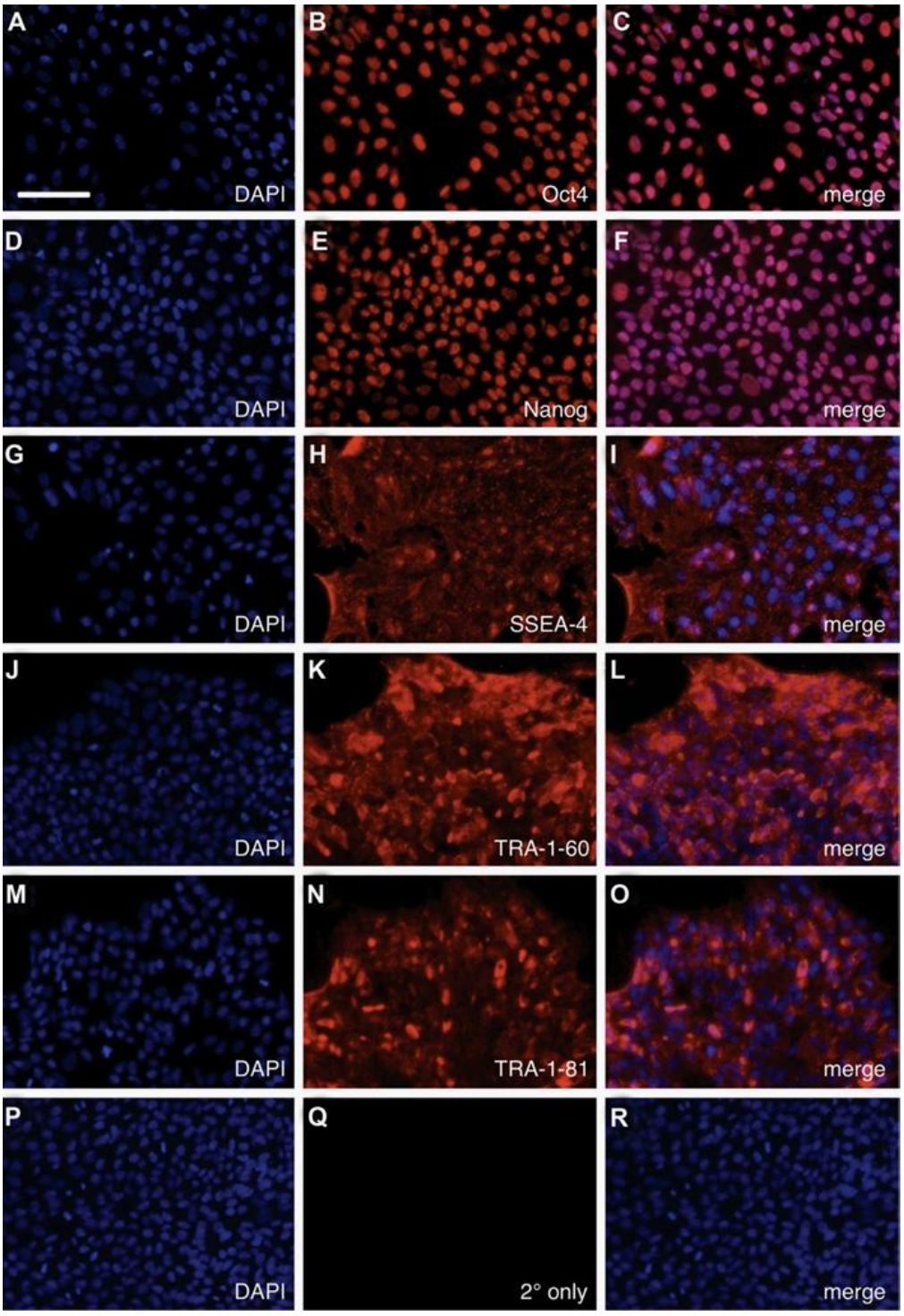


Figure 3.3. Expression of pluripotency markers in h*Fezf2* gene targeted HUES-9 cells by immunocytochemistry. 2° only, secondary antibody only (no primary antibody). Scale bar = 100 μm.

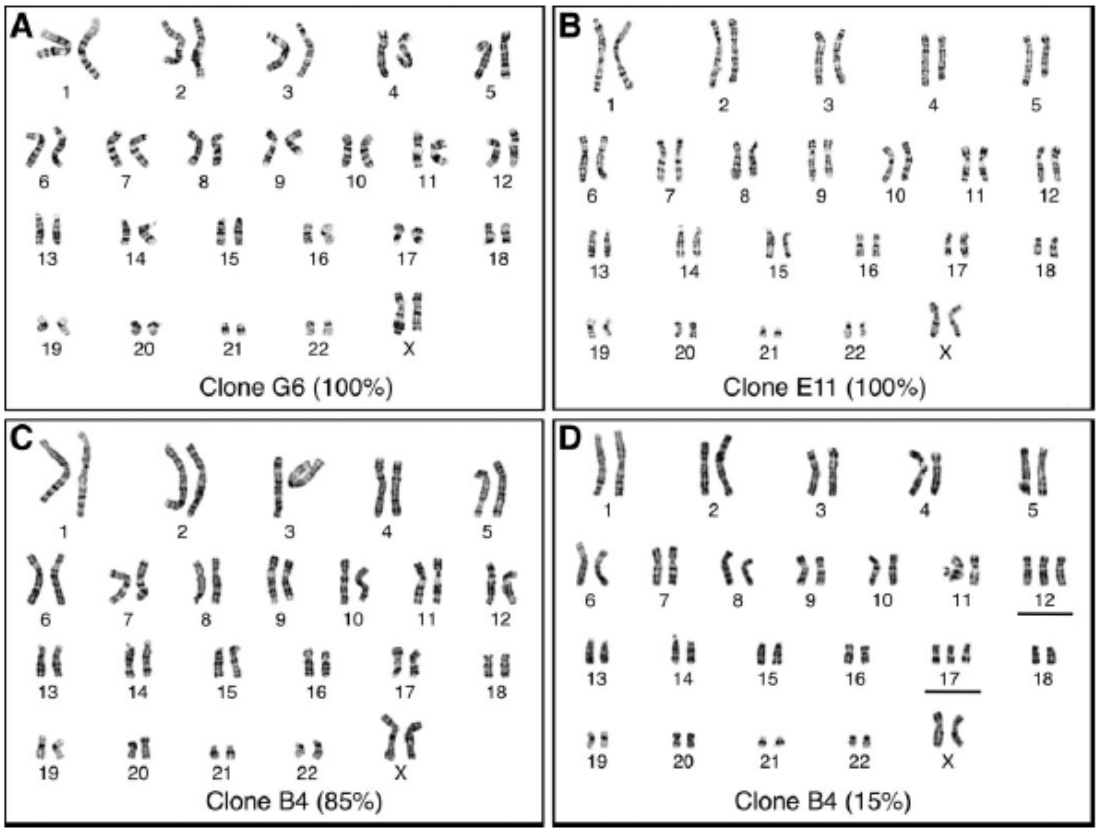


Figure 3.4. Karyotype analysis of h*Fezf2* gene targeted HUES-9 cells. 100% of the cells from clones G6 (**A**) and E11 (**B**) show normal karyotype. 85% of the cells from clone B4 show normal karyotype (**C**); 15% carry trisomies 12 and 17 (underlined in **D**). All cells carry a pericentric inversion on Chromosome 9 [inv(9)(p11q13)] as reported in the parental HUES-9 cells [55]. This inversion is observed in ~2% of the population with no known associated clinical abnormality.

trisomies 12 and 17 (**Figure 3.4D**). It has been reported that trisomies 12 and 17 occur frequently in human ES cell culture [55, 75]. Thus, these trisomies may have arisen spontaneously under standard culture conditions and were not necessarily a direct consequence of the unique procedures of gene targeting (e.g. electroporation).

3.3.5 Cre-mediated excision of the loxP-flanked Neo marker

The *loxP*-flanked *Neo* marker can be excised by transient Cre expression. Excision of the *Neo* marker serves two purposes: 1) to render the heterozygous targeted ES cells Geneticin sensitive so that a second round of gene targeting at the locus can be performed with the same gene targeting vector; 2) to avoid any potential undesired effect on the expression of the *EYFP* reporter gene as a precautionary measure because of documented evidence that a selection cassette, especially one with the bi-directional PGK promoter, may affect the expression of nearby genes [76-78].

Targeted cells were transiently transfected with a Cre plasmid. Transfected cells were plated at low densities to allow for clonal expansion in the absence of drug selection. This experiment also allowed us to estimate the efficiency of clonal expansion following electroporation for HUES-9 cells to be at ~3% when plated at a low density. Individual clones were picked and tested for Geneticin resistance in 96-wells. Successful Cre excision would lead to a loss of the *Neo* marker (**Figure 3.2A**). Out of 89 clones analyzed, 28 regained sensitivity to Geneticin, indicating a Cre excision efficiency of ~30%.

Correct Cre excision was confirmed by PCR analysis on Geneticin sensitive clones (**Figure 3.2E**). In addition, no undesired Cre plasmid integration was detected by PCR analysis using Cre specific primers (**Figure 3.2F**). We performed karyotype analysis on excised clones, focusing initially on clones excised from the parental clone B4 (with trisomies 12 and 17 in ~15% of the cells). Out of three excised clones analyzed, all exhibited normal karyotype in 100% of the cells analyzed (~20 cells analyzed per clone, data not shown). This result indicates that homogenous, karyotypically normal cells can be readily obtained after one round of clonal expansion from a HUES-9 population containing a subset of karyotypically abnormal cells.

Following Cre-mediated excision of *Neo*, we made an initial attempt to target the second h*Fezf2* allele with the same gene targeting vector. Out of 149 colonies effectively screened, none were doubly targeted. In addition to the relatively low number of colonies screened, this may be due to a lower expected targeting frequency for two reasons: 1) the second round of gene targeting may lead to re-targeting of the first targeted allele, effectively reducing the expected targeting frequency for the second allele by half; 2) the first round of gene targeting may have occurred preferentially at the allele that has less amount of sequence discrepancy with the targeting vector, and consequently targeting the second allele would entail homologous recombination with an increased amount of sequence discrepancy (see Discussion).

3.3.6 EYFP knockin reporter is expressed upon neural differentiation

To determine whether the knockin *EYFP* gene is functional as a fluorescent reporter, we took advantage of a published protocol of neural differentiation from human ES cells [59], as a first step towards the development of an efficient differentiation paradigm to generate EYFP positive neurons or neuronal progenitors from h*Fezf2* targeted HUES-9 (HUES-9-h*Fezf2)* cells. Briefly, following embryoid body formation, HUES-9 cells were grown in neural induction media to allow for the development of neural rosette-like structures. Rosettes were then replated sequentially in neural precursor media and neural differentiation media. Upon neural differentiation, native EYFP signals were readily observed in a subset of cells that partially overlap with cells positive for OTX2 (**Figure 3.5**), a forebrain/midbrain marker [79, 80]. Thus, the *EYFP* knockin reporter gene can be induced following neural differentiation of h*Fezf2* targeted HUES-9 cells.

3.4 Discussion

Here we demonstrated successful gene targeting at the h*Fezf2* locus in HUES-9 human ES cells. An *EYFP* reporter gene was targeted into the h*Fezf2* locus, making it possible to track the differentiation of human ES cells along the lineage towards subcortical projection neurons including corticospinal neurons [31, 32]. This would be useful both for the development of an efficient *in vitro* differentiation protocol and for assessing *in vivo* differentiation and integration of HUES-9 cells or its derivatives towards subcortical neurons in an

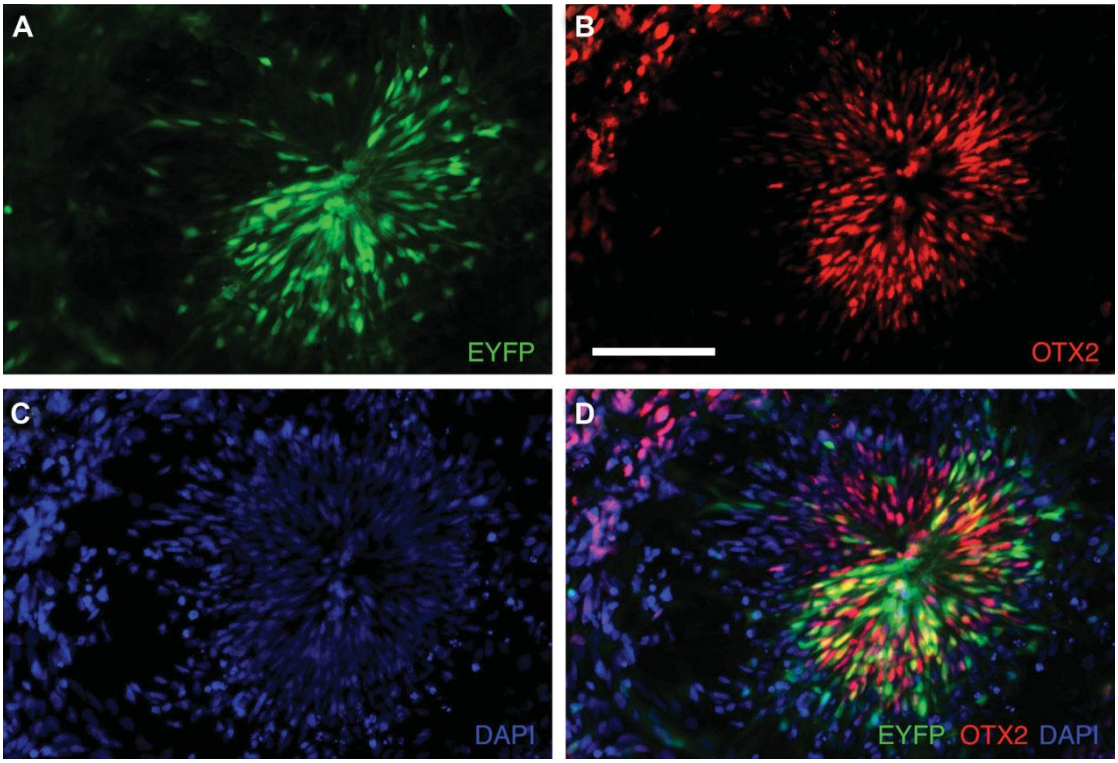


Figure 3.5. *EYFP* reporter gene expression upon neural differentiation of h*Fezf2* gene targeted and Cre excised HUES-9 cells. EYFP (in green) was detected with direct fluorescence (native signals); OTX2 (in red) was detected with immunofluorescence. Yellow signals indicate areas of overlap between EYFP and OTX2. Note that OTX2 is concentrated in the nucleus while EYFP is more evenly distributed between the cytoplasm and the nucleus. Scale bar = 100 μm.

animal transplantation model [81].

Our data indicate that, similar to mouse ES cells, HUES cells are highly amenable to targeted modifications without any significant loss of pluripotency marker expression or karyotypic stability. Trisomies 12 and 17 arise in culture occasionally but even when this occurs homogenous karyotypically normal cell populations can be readily obtained after one round of clonal expansion. In this regard, the robust capability of HUES-9 cells for clonal expansion (at a ~24% efficiency) offers a great advantage. In comparison, the H1.1 cell line has been reported to exhibit a clonal expansion efficiency of ~1% when plated at a low density [68]. A high clonal expansion efficiency is also important for genetic modifications that require low density plating after electroporation, including Cre-mediated excision of a *loxP*-flanked selection marker. In this regard, we observed ~30% efficiency of Cre mediated excision, which is comparable to that observed with mouse ES cells [82].

Two methods have been shown to enhance the efficiency of clonal expansion of human ES cells. Treatment with the ROCK inhibitor Y-27632 was shown to increase the clonal expansion efficiency of KhES cells from ~1% to ~27% [72]. Treatment with a neurotrophin cocktail was shown to effect a similar increase in the clonal expansion efficiency of H1 and H9 cells [71]. However, the general applicability of these methods to other human ES cell lines remains to be seen, and the long-term effect of such treatments on the pluripotency, full differentiation potential *in vitro* and *in vivo*, or tumorigenicity of human ES cells awaits further validation. In contrast, the intrinsically high clonal expansion

efficiency of HUES-9 cells presents a unique advantage in genetic manipulation.

Among the few published reports on electroporation-based gene targeting in human ES cells, three groups reported gene targeting at the *Hprt* locus in male (XY) ES cells [65, 67, 68]. Loss of *Hprt* (residing on the X-chromosome so there is only one copy of *Hprt* in XY cells) leads to drug resistance to 6-Thioguinine, which can be selected for in culture. This strategy is not generally applicable to most other loci in the genome since there is no negative selection scheme available for most human genes. Estimated gene targeting frequencies at *Hprt* (in the absence of negative selection) range from 0.2% to 2% under normal culture conditions [65, 67, 68]. Two groups reported gene targeting at non-selectable genes using a promoterless gene trap targeting vector [66, 68]. This strategy relies on the gene of interest being expressed in human ES cells and consequently cannot be extended to genes not expressed in human ES cells such as h*Fezf2*. The reported targeting frequencies vary, from 2.3% at *ROSA26* to 27 - 40% at *POU5F1* (encoding Oct4). To date, only one group published a detailed report on gene targeting at a locus not expressed in human ES cells [64]. However, the targeting frequency at this locus (*MIXL1*), a key measure for the robustness of gene targeting, was not reported.

Here we provide the first report on the gene targeting frequency at a gene not expressed in human ES cells using an electroporation-based gene targeting strategy that is applicable to all genes. With negative selection but

without promoterless gene trap enrichment, we observed a targeting frequency of ~1.5% at h*Fezf2*. Our results have several important implications. First, this targeting frequency is sufficient for routine application since up to several hundred clones can be screened in one experiment. Second, negative selection significantly improves the targeting frequency (~5×) and may prove crucial when targeting genes not expressed in human ES cells. Without negative selection the targeting frequency at h*Fezf2* is only ~0.3%, significantly lower than that with the promoterless gene trap targeting strategy (2.3% at *ROSA26* and up to 40% at *POU5F1*) [66, 68]. It is possible that in addition to the enrichment provided by promoterless gene trap, the genes expressed in human ES cells are inherently more accessible to gene targeting due to a potentially more open chromatin structure. Third, with the caveat that gene targeting frequencies vary from locus to locus so a single locus may not be representative, our targeting frequency at h*Fezf2* is at the lower end of what is usually observed in mouse ES cells with the same targeting strategy and a comparable total length of homology. This might be partly due to the hybrid nature of human ES cells, which makes it impossible, for most genes, to construct a targeting vector that is isogenic to both alleles of the same gene. In this regard, our sequencing analysis identified four single nucleotide polymorphisms between the two HUES-9 alleles within the total region of homology. In addition, PCR errors introduced during long range genomic PCR, which are frequently observed even with the use of high fidelity amplification systems, will also presumably contribute to a lower targeting frequency.

Although it has been shown that gene targeting in human cells (including human ES cells) does not require the use of isogenic DNA [65, 67, 83], the impact of the amount of sequence polymorphism between the targeting vector and the genomic allele to be targeted on gene targeting frequencies in human ES cells has not been evaluated systematically.

Two other distinct strategies have been used to achieve gene targeting in human ES cells. A Zinc-finger nuclease-based strategy coupled with integrase-defective lentiviral vector delivery has been used to target the *CCR5* locus in human ES cells [84]. However, the level of expertise required for the design and validation of the zinc finger nuclease presents a challenge. In addition, whether this strategy only introduces the intended genetic change without any off-target effect remains unknown. More recently, an adenovirus-based vector delivery strategy was used to target the *Hprt* locus [85]. It remains to be seen whether this result can be extended to non-selectable genes, including those that are not expressed in human ES cells. In general, the viral vector-based gene targeting strategies may have a distinct advantage for human ES cell lines that have not yet been adapted for enzymatic passage.

Our data indicate that HUES-9 cells are amenable to targeted modifications similar to mouse ES cells with an electroporation-based method. HUES-9 cells exhibit a high clonal expansion efficiency and survive electroporation well. Thus, the perception that human ES cells are less amenable to genetic modification has been primarily based on observations made on cells that were developed under mechanical passage. Cells

developed under enzymatic passage can better tolerate clonal expansion. Human ES cells developed or adapted for enzymatic passage and efficient clonal expansion are robust substrates for targeted modifications via electroporation. Because all HUES cell lines were developed under the same conditions, our procedures on HUES-9 are presumably transferable to other HUES lines. Along with their other positive features including a more "pristine" state of X-chromosome inactivation and a wide range of developmental potential as a group [86, 87], genetic amenability of HUES cells will make them valuable tools in basic research and therapeutic development.

3.5 Conclusion

We have successfully targeted the h*Fezf2* gene in HUES-9 human embryonic stem cells, which will facilitate the study of differentiation of HUES-9 cells towards subcortical projection neurons including corticospinal neurons in vitro and in animal transplantation models. Similar to mouse ES cells, HUES-9 cells exhibit a high efficiency of clonal expansion and can be readily modified genetically with an electroporation-based method without any significant loss of pluripotency marker expression or karyotypic stability.

Chapter 3, in full, is a reprint of the material as it appears in Ruby and Zheng, Stem Cells 2009. The dissertation author is the primary author of this paper. All statements were accurate as of the time of publishing.

3.6 References

1. Mak TW. Gene targeting in embryonic stem cells scores a knockout in Stockholm. Cell. 2007;131:1027-1031.

2. Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. Science. 1998;282:1145-1147.

3. Giudice A, Trounson A. Genetic modification of human embryonic stem cells for derivation of target cells. Cell Stem Cell. 2008;2:422-433.

4. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet. 2005;6:507-512.

5. Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007;131:861-872.

6. Yu J, Vodyanik MA, Smuga-Otto K, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007;318:1917-1920.

7. Park IH, Arora N, Huo H, et al. Disease-Specific Induced Pluripotent Stem Cells. Cell. 2008.

8. Service RF. Nobel Prize in chemistry. Three scientists bask in prize's fluorescent glow. Science. 2008;322:361.

9. Davis RP, Ng ES, Costa M, et al. Targeting a GFP reporter gene to the MIXL1 locus of human embryonic stem cells identifies human primitive streak-like cells and enables isolation of primitive hematopoietic precursors. Blood. 2008;111:1876-1884.

10. Di Domenico AI, Christodoulou I, Pells SC, et al. Sequential Genetic Modification of the hprt Locus in Human ESCs Combining Gene Targeting and Recombinase-Mediated Cassette Exchange. Cloning Stem Cells. 2008;10:217-230.

11. Irion S, Luche H, Gadue P, et al. Identification and targeting of the ROSA26 locus in human embryonic stem cells. Nat Biotechnol. 2007;25:1477-1482.

12. Urbach A, Schuldiner M, Benvenisty N. Modeling for Lesch-Nyhan disease by gene targeting in human embryonic stem cells. Stem Cells. 2004;22:635-641.

13. Zwaka TP, Thomson JA. Homologous recombination in human embryonic stem cells. Nat Biotechnol. 2003;21:319-321.

14. Cowan CA, Klimanskaya I, McMahon J, et al. Derivation of embryonic stem-cell lines from human blastocysts. N Engl J Med. 2004;350:1353-1356.

15. Guhr A, Kurtz A, Friedgen K, et al. Current state of human embryonic stem cell research: an overview of cell lines and their use in experimental work. Stem Cells. 2006;24:2187-2191.

16. Molyneaux BJ, Arlotta P, Hirata T, et al. Fezl is required for the birth and specification of corticospinal motor neurons. Neuron. 2005;47:817-831.

17. Chen B, Schaevitz LR, McConnell SK. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. Proc Natl Acad Sci U S A. 2005;102:17184-17189.

18. Ramirez-Solis R, Davis AC, Bradley A. Gene targeting in embryonic stem cells. Methods Enzymol. 1993;225:855-878.

19. Pyle AD, Lock LF, Donovan PJ. Neurotrophins mediate human embryonic stem cell survival. Nat Biotechnol. 2006;24:344-350.

20. Watanabe K, Ueno M, Kamiya D, et al. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat Biotechnol. 2007;25:681-686.

21. Mansour SL, Thomas KR, Capecchi MR. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature. 1988;336:348-352.

22. Bradley A, Zheng B, Liu P. Thirteen years of manipulating the mouse genome: a personal history. Int J Dev Biol. 1998;42:943-950.

23. Draper JS, Smith K, Gokhale P, et al. Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat Biotechnol. 2004;22:53-54.

24. Manley NR, Barrow JR, Zhang T, et al. Hoxb2 and hoxb4 act together to specify ventral body wall formation. Dev Biol. 2001;237:130-144.

25. Olson EN, Arnold HH, Rigby PW, et al. Know your neighbors: three phenotypes in null mutants of the myogenic bHLH gene MRF4. Cell. 1996;85:1-4.

26. Scacheri PC, Crabtree JS, Novotny EA, et al. Bidirectional transcriptional activity of PGK-neomycin and unexpected embryonic lethality in heterozygote chimeric knockout mice. Genesis. 2001;30:259-263.

27. Zhang SC, Wernig M, Duncan ID, et al. In vitro differentiation of transplantable neural precursors from human embryonic stem cells. Nat Biotechnol. 2001;19:1129-1133.

28. Simeone A, Acampora D, Gulisano M, et al. Nested expression domains of four homeobox genes in developing rostral brain. Nature. 1992;358:687-690.

29. Simeone A, Puelles E, Acampora D. The Otx family. Curr Opin Genet Dev. 2002;12:409-415.

30. Muotri AR, Nakashima K, Toni N, et al. Development of functional human embryonic stem cell-derived neurons in mouse brain. Proc Natl Acad Sci U S A. 2005;102:18644-18648.

31. Araki K, Imaizumi T, Okuyama K, et al. Efficiency of recombination by Cre transient expression in embryonic stem cells: comparison of various promoters. J Biochem. 1997;122:977-982.

32. Sedivy JM, Vogelstein B, Liber HL, et al. Gene targeting in human cells without isogenic DNA. Science. 1999;283:9a.

33. Lombardo A, Genovese P, Beausejour CM, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol. 2007;25:1298-1306.

34. Suzuki K, Mitsui K, Aizawa E, et al. Highly efficient transient gene expression and gene targeting in primate embryonic stem cells with helper-dependent adenoviral vectors. Proc Natl Acad Sci U S A. 2008;105:13781-13786.

35. Osafune K, Caron L, Borowiak M, et al. Marked differences in differentiation propensity among human embryonic stem cell lines. Nat Biotechnol. 2008;26:313-315.

36. Silva SS, Rowntree RK, Mekhoubad S, et al. X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc Natl Acad Sci U S A. 2008;105:4820-4825.

Chapter 4

Directed differentiation of HUES-9-h*Fezf2* hESCs towards a corticospinal fate

4.1 Abstract

Generating a specific cell type of interest at high enough efficiency and purity to address both basic biology questions and to use as a potential source for cellular replacement therapies is a major challenge for the field of regenerative medicine. To increase the efficiency of current differentiation protocols, it is essential to have a method to identify the cell type of interest among a mixed population of cells. Using our HUES-9-h*Fezf2* fluorescent reporter cell line, we have developed a differentiation protocol to generate Fezf2-expressing cells. Importantly, because Fezf2-expressing cells express EYFP, live cells can be purified via FACS for further analysis. Our protocol is able to generate ~3% Fezf2-expressing cells as determined by FACS analysis. The ability to purify live cells allows for continued culturing of purified cells and transplantation of purified cells into an animal host to assess survival, differentiation, and cell type identity (i.e. type of neuron as determined by size, morphology, axon projections, etc). These tools provide the opportunity to study human cells in ways that were previously inaccessible, and offer enormous potential for advancing basic biology research and progressing towards development of cellular replacement therapies for a variety of neurological conditions.

4.2 Introduction

Current differentiation protocols generate a wide variety of cell types, and are generally very inefficient at generating a particular cell type at high

purity. In addition, some protocols require the use of a co-culture system with another cell type, such as co-culture with a mouse endoderm-like cell line (END2) for cardiomyocytes[1] or the PA6 or MS5 stromal feeder cell line for neurons[2]. To perform in depth studies on a particular cell type of interest, it is critical to be able to generate a large quantity of cells at high purity. Before any cellular therapies can be possible, it is essential to develop differentiation protocols that are efficient and do not require the use of animal cells or animal products.

Directing an ES cell towards a specific cell fate is a difficult task for any cell type. The complexity of the human brain, specifically the cortex, adds an additional dimension of complexity to the task of generating a specific type of neuron. There are multiple decisions a cell must make along its path to successfully acquire a specific cortical neuron identity.

Along the pathway from ES cell to cortical neuron, the first major decision that must be made is for a cell to adopt a neural identity. At this stage, a cell is then committed to the neural lineage, but is still able to make all types of neural cells. In the central nervous system, three major neural cell types are neurons, astrocytes, and oligodendrocytes. Neurons are the basic elements for wiring of the brain. They are electrically excitable cells that transmit signals through electrical and chemical signaling. Neurons have three major parts: the cell body, dendrites, and an axon. The cell body is the powerhouse of the cell, containing the cell nucleus. From the cell body, numerous dendrites extend and connect to other axons to receive signals

and relay them back to the cell body. While neurons may have numerous dendrites, they extend only one axon. In general, the role of this axon is to conduct electrical signals away from the cell body, to communicate with other cells. Astrocytes function as support cells, playing a wide range of roles, such as providing nutrients to the surrounding tissue, helping to maintain the blood brain barrier, and participating in the repair process after injury. Oligodendrocytes also serve as support cells, but have a more specialized role. Their main function is the insulation of axons, which is accomplished by wrapping the axons with myelin sheaths that are basically membranes made by oligodendrocytes (in the CNS) or Schwann cells (in the PNS). Myelin is critical for the function of many types of axons, as it allows axons to send rapid and efficient signals.

Once an ES cell has committed to the neural lineage, the next major step towards becoming a CSMN is to adopt a forebrain fate. There are a number of steps involved in the process of a general neural progenitor committing to a dorsal forebrain fate. The first step is to commit to become a neural cell of the central nervous system, as opposed to the peripheral nervous system. During neurulation, neural plate folds into a cylinder-like structure called neural tube that would become the future CNS, leaving out the neural crest cells some of which will become the future PNS. The central nervous system consists of the brain and spinal cord. Once a cell has committed to become a brain cell (not a spinal cord cell), the next step is to acquire an area-specific identity within the brain. The brain is divided into

three areas: the forebrain, midbrain, and hindbrain. The cerebral cortex, where CSMNs reside, is located in the forebrain. However, once a cell acquires a forebrain fate, there is still a wide variety of cell types it could become.

After acquiring a forebrain fate, the final steps involve acquiring a very specialized cellular identity within the forebrain to become a CSMN. The cortex is divided into six layers, with each layer containing a variety of cell types. Layer V contains CSMNs, as well as a variety of other related subcerebral projection neurons. Therefore, it is necessary not only to acquire a layer V cortical identity, but also to ensure a CSMN identity is chosen. Due to the complexity of the human brain, directing differentiation towards a specific cell type has proven to be difficult. To increase the efficiency of generating a particular cell type, it is important to provide cells with cues to instruct them which cell type to become. Based on our understanding of mammalian nervous system development from model systems, such as the mouse, we can apply these principles to direct differentiation towards a desired fate. We can use known markers of different cell types to monitor the progression of differentiation and determine if some or all of the mouse developmental pathways are conserved in human cells.

4.3 Results

To achieve efficient neural induction, it is important to uniformly direct cells towards a neural fate. To improve the efficiency of neural induction, it is

important to identify pathways that regulate development of the three different lineages: mesoderm, endoderm, and ectoderm. Inhibition of SMAD signaling has been shown to be an effective method to achieve efficient neural induction. SMADs are a class of intracellular proteins that transduce extracellular transforming growth factor beta (TGF-β) signals to the nucleus to activate downstream TGF-β signaling[3-8]. Studies in frog have demonstrated BMP inhibitors, such as noggin, play a critical role in neural induction[9]. Noggin has been shown to play a similar role in human neural development[10], and has been used in differentiation protocols to increase the efficiency of neural induction[11, 12]. In addition, a small molecule drug, SB431542, has been shown to improve the efficiency of neural induction[13]. SB431542 inhibits the Activin/Lefty/TGF-β pathways by preventing phosphorylation of the ALK family of receptors. Recently, it has been shown that dual SMAD inhibition, using noggin and SB431542 in combination, achieves greater than 80% neural induction[14].

In addition to general neural induction, cells must be directed to acquire a dorsal identity in order to become CSMNs. One of the major patterning factors responsible for dorsal/ventral patterning is sonic hedgehog (Shh). Shh is secreted from the notochord and ventral floor plate of the developing neural tube to create a gradient that directs dorsal/ventral patterning[15]. The highest concentrations of Shh are found in the most ventral portions of the neural tube, with lower levels in the more dorsal areas. CSMNs are found in the dorsal forebrain, therefore inhibition of Shh signaling during

times when cells are responding to environmental cues to determine their cell fate is a potential mechanism to direct cells towards a dorsal fate.

4.3.1 Efficient neural induction via dual SMAD inhibition

We tested the efficiency of neural induction via this dual SMAD inhibition method using two different media: the published KSR/N2 media published and a chemically defined media called N2B27. A side-by-side comparison was performed to look at expression of key neural markers in each media condition to determine if the media used has a detectable impact on differentiation. On day 5 of differentiation, expression of Oct4, Pax6, and Nestin were assessed to analyze the transition from the ES cell stage towards neural induction **(Figure 4.1)**. Oct4 is a pluripotency marker, indicating cells maintaining an undifferentiated stem cell-like state; nestin is a general neural stem cell marker; Pax6 is also a neural stem cell marker and has been shown to play an important role in genesis of cortical pyramidal neurons[16]. In the N2B27 condition, Oct4 expression was detected in very few cells; the vast majority of cells had downregulated Oct4 expression by day 5. In contrast, a large proportion of cells in the KSR/N2 condition continued to maintain Oct4 expression throughout the first 5 days of differentiation. Both conditions produced Nestin-positive cells, but there was a marked difference in Pax6 expression. Cells cultured I n N2B27 showed uniform Pax6 expression throughout virtually all cells; KSR/N2 produced patches of Pax6 expression, but much fewer cells expressed Pax6 as compared to the N2B27 condition. Taken

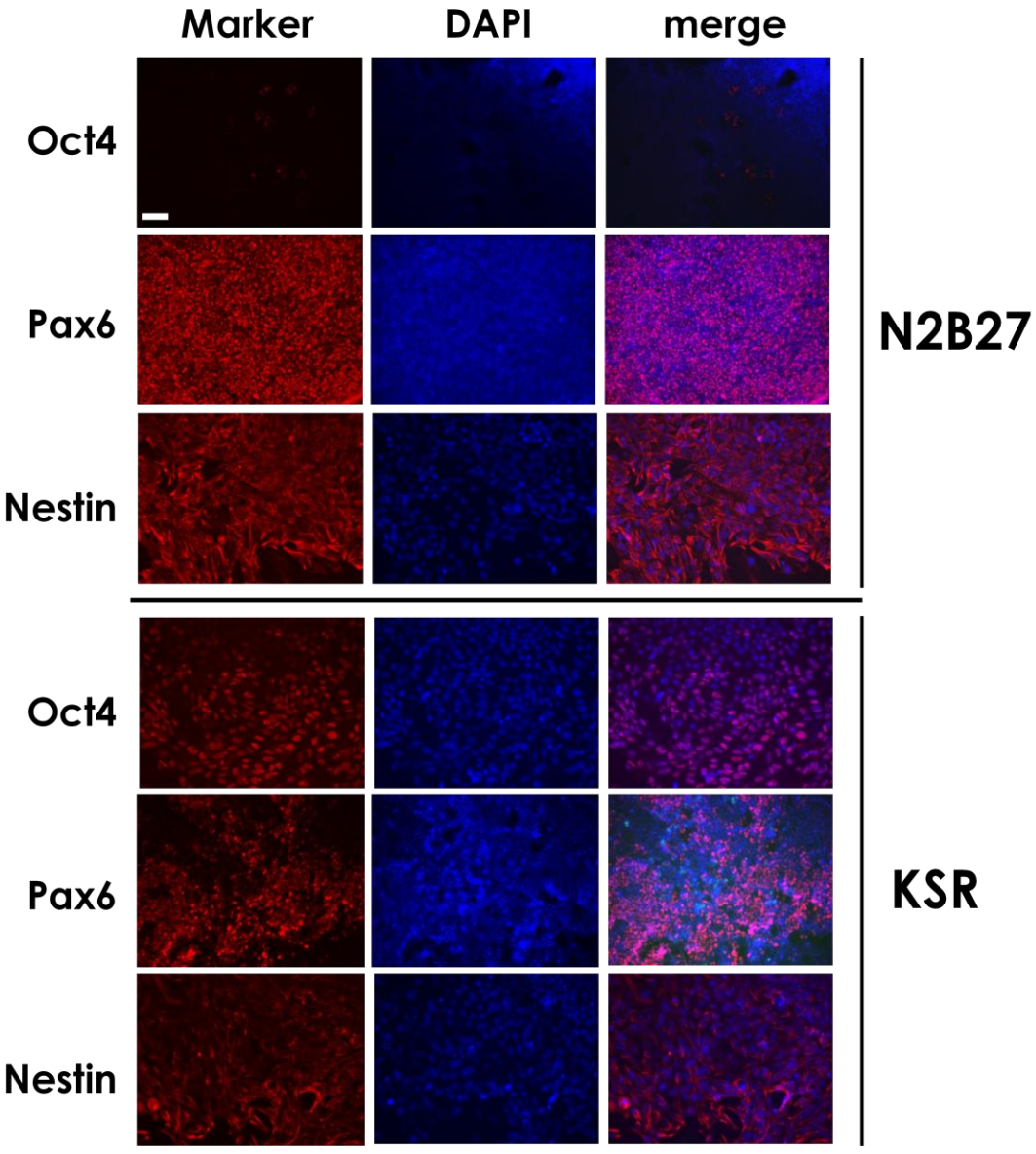


Figure 4.1: Comparison of pluripotency and neural marker expression at day 5 of differentiation using two different differentiation media. scale bar = 100 µm. Abbreviations: DAPI, 4',6-diamidino-2-phenylindole; KSR, knockout serum replacement.

together, this information suggests the rate of differentiation varies between these two conditions, with cells cultured in N2B27 demonstrating faster downregulation of the pluripotency marker Oct4 and upregulation of the neural marker Pax6.

To further analyze differences between the two conditions, marker analysis was also performed at day 10 of differentiation. Otx2 expression was analyzed in addition to the continued analysis of Oct4, Nestin, and Pax6 **(Figure 4.2)**. Otx2 is a neural marker specific for anterior neuroectoderm. In both conditions, no Oct4-positive cells were detected, indicating all cells had exited their initial embryonic stem cell state. Similar to day 5, Nestin expression was detected throughout both cultures. Pax6 expression continued to increase in the KSR/N2 condition, showing similar expression to the N2B27 culture at day 5. N2B27 cells also showed a high proportion of Pax6-positive cells, but the culture was not as homogeneously Pax6-positive as was seen at day 5, suggesting some cells that had expressed Pax6 at day 5 may already be downregulating expression. The most striking differences were seen in Otx2 expression. While both cultures produced some Otx2-positive cells, there was a pronounced difference in expression pattern and frequency. The majority of cells cultured in N2B27 stained positive for Otx2; however, KSR/N2 cells showed only a small subset of cells positive for Otx2, and these cells almost always occurred in clusters. This data further supports the conclusion that these two different conditions result in different rates of differentiation, with the N2B27 cells progressing at a faster rate than KSR/N2 cells. In addition, N2B27

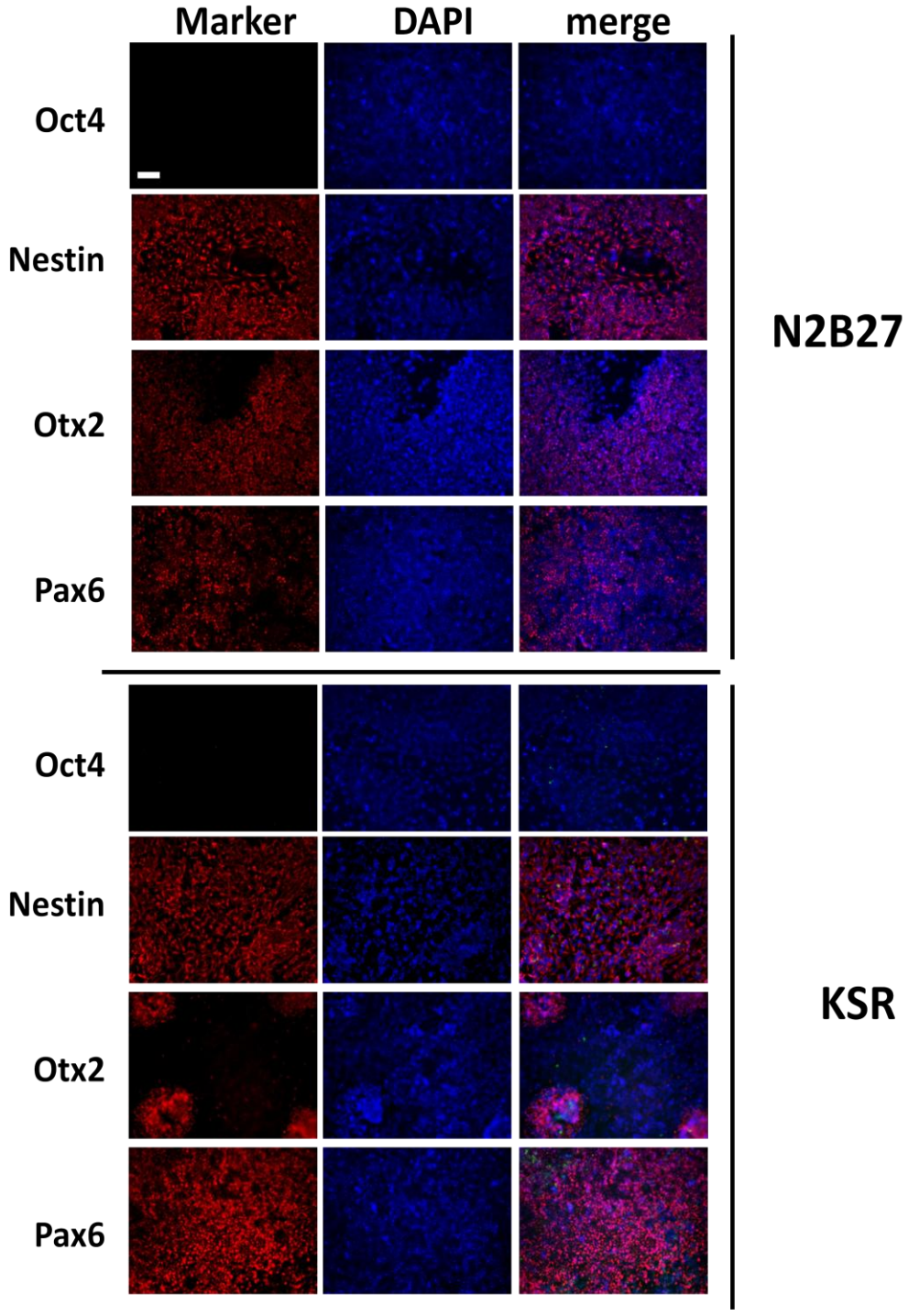


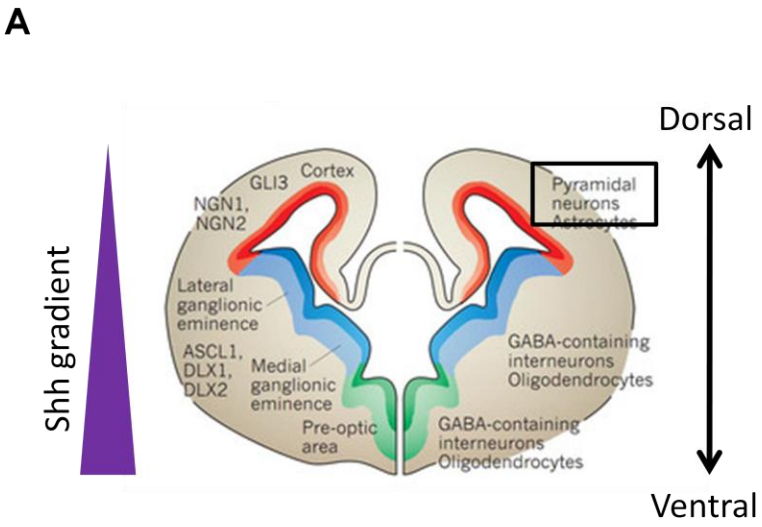
Figure 4.2: Comparison of pluripotency and neural marker expression at day 10 of differentiation using two different differentiation media. scale bar = 100 µm. Abbreviations: DAPI, 4',6-diamidino-2-phenylindole; KSR, knockout serum replacement

produces a more homogenous culture of cells that are more uniformly Otx2-positive, indicating this culture system may be more efficient at generating forebrain progenitor cells. For these reasons, all future experiments were performed using the N2B27 method for differentiation.

4.3.2 Patterning towards dorsal forebrain via inhibition of sonic hedgehog

Once a cell has committed to the neural lineage, the next step towards acquiring a CSMN fate is to adopt a dorsal identity. Patterning of the neural tube along the dorsal-ventral axis defines compartments of progenitor cells, which give rise to different classes of neurons[17]. The secreted protein, sonic hedgehog (Shh) is a key morphogen involved with induction of ventral neural identity[18]. Therefore, inhibition of Shh signaling should result in cells with a dorsal neural identity. To test this, we treated cells with cyclopamine, an inhibitor of Shh signaling, during differentiation to prevent induction of a ventral identity.

To inhibit Shh signaling, cells were treated with cyclopamine during days 2 through 10 of differentiation. To analyze the effect of cyclopamine treatment on efficiency of generating Fezf2-expressing cells, ES cells were differentiated for 19 days, and then subjected to semi-quantitative RT-PCR analysis to measure Fezf2 expression **(Figure 4.3)**. Cells differentiated without any patterning factors showed a low level of Fezf2 expression. Dual SMAD inhibition treatment increased Fezf2 expression greater than 3-fold as compared to no treatment. Combined treatment with dual SMAD inhibitors



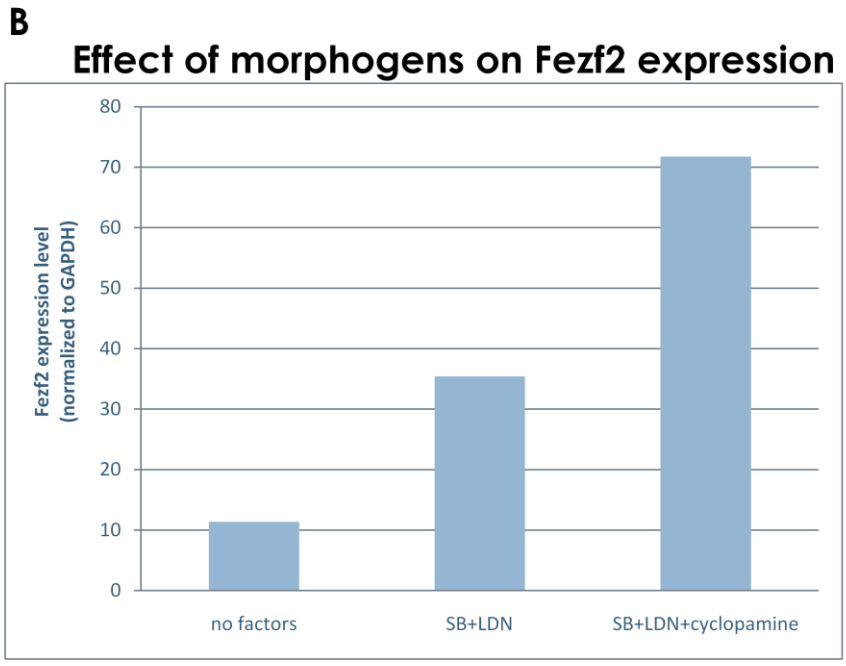


Figure 4.3: Effect of inhibiting SMAD and Shh signaling on Fezf2 expression level. (A) Schematic showing Shh gradient. Shh expression is highest in ventral areas, and decreases in more dorsal zones, where pyramidal neurons are located. (B) semi-q-RT-PCR analysis at day 19 of differentiation. Fezf2 levels were normalized to GAPDH. SB+LDN are dual SMAD inhibitors SB431542 and LDN-193189.

and cyclopamine revealed an added 2-fold increase in Fezf2 expression as compared to treatment with dual SMAD inhibitors alone; a 7-fold increase was observed when compared to cells differentiated without any patterning factors.

4.3.3 A directed differentiation method to generate cortical neurons

The combined effects observed using an adherent culture system, dual SMAD inhibition, and inhibition of Shh signaling allowed us to develop a streamlined protocol to generate cortical neurons **(Figure 4.4A)**. Starting at the embryonic stem cell stage, cells are cultured at least one passage feeder-free prior to differentiation until confluent. To prepare for differentiation, cells are passaged onto growth factor reduced Matrigel and cultured in N2B27 media with fgf until cells reach confluence. To begin differentiation, fgf is removed and dual SMAD inhibitors are added to the culture for the first ten days of differentiation. Starting on day 2, cyclopamine is also added to inhibit Shh signaling. On day 11, all patterning factors are removed. At this point, the cells are neural progenitor cells. Cells can be passaged and maintained in fgf if a renewable neural progenitor population is desired. To continue along the differentiation path, fgf is added at a high concentration to induce the final mitotic division and allow cells to become mature neurons. The media is changed to neural differentiation media on day 12 with no exogenous factors added. Cells have been maintained up to 30 days of differentiation with half media changes daily starting at day 12.

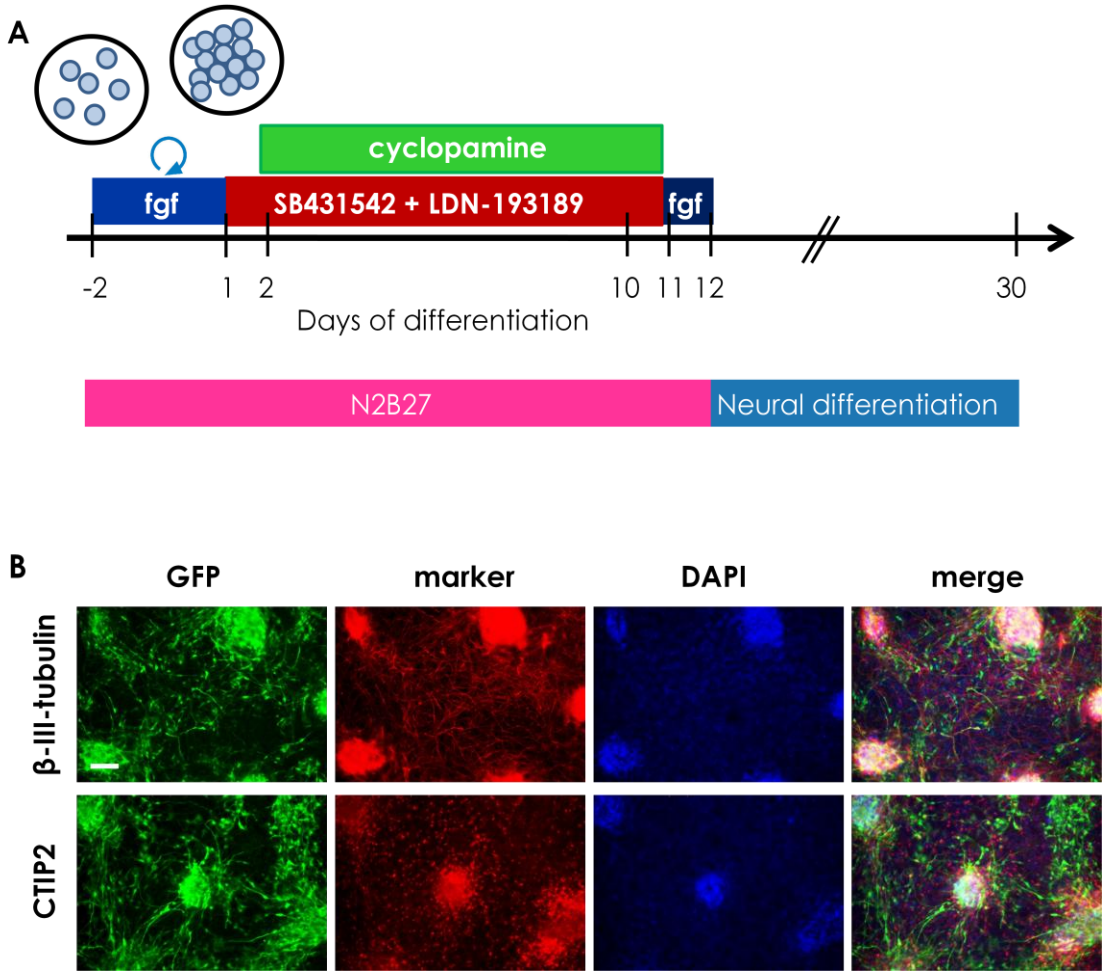


Figure 4.4: Directed differentiation of HUES-9-h*Fezf2* cells. (A) Schematic of the differentiation scheme to generate Fezf2+ cells from human embryonic stem cells. (B) Immunostaining of cells at day 30 of differentiation. Endogenous EYFP expression was amplified by staining with a GFP antibody. scale bar = 100 μm. Abbreviations: DAPI, 4',6-diamidino-2-phenylindole; GFP, green fluorescent protein.

4.3.4 Fezf2-expressing cells also express β-III-tubulin and CTIP2 late in differentiation

At early time points, our protocol gives a high percentage of neural cells; however, it is also important to look at later time points to determine if this protocol is capable of generating Fezf2+ cells, as assessed by the Fezf2-EYFP reporter. To analyze this, cells were differentiated for 30 days, then analyzed for expression of key markers. Immunostaining revealed a number of cells co-expressing Fezf2 (as shown by EYFP reporter expression) and β-III-tubulin, an early neural marker (Tuj1, **Figure 4.4B**). In mouse, it has been shown that Fezf2 is expressed in neurons, but not astrocytes or oligodendrocytes. Co-expression of Fezf2 and β-III-tubulin demonstrates in human cells, Fezf2 is also expressed in neurons.

One of the best characterized genes associated with Fezf2 function is CTIP2. CTIP2 is a transcription factor positively regulated by Fezf2 that plays a key role in differentiation of subcerebral projection neurons, including extension of axonal projections[19-22]. Analysis at day 30 showed a subset of Fezf2-expressing cells also express CTIP2 **(Figure 4.4B)**. Additionally, the order and timing of CTIP2 expression appears to be conserved in human cells, as CTIP2 is not detected via immunostaining until approximately day 20 of differentiation, whereas Fezf2 is detected starting around day 10 (data not shown). This data provides evidence for conservation of the role of Fezf2 between mouse and human, and the potential for Fezf2 to be a powerful tool to study human corticospinal motor neurons.

4.3.5 Time course analysis of developmental markers

To look more in depth at the progression of our culture system, RT-PCR analysis of some key developmental markers was performed at a series of time points throughout differentiation **(Figure 4.5)**. Pluripotency markers Oct4 and Nanog were quickly downregulated within the first few days of differentiation. This data is consistent with immunostaining analysis that showed very few cells still maintaining Oct4 expression at day 5 **(see Figure 4.1)**. Sox2 expression was relatively consistent throughout differentiation, which was to be expected since Sox2 is involved in maintaining pluripotency in ES cells, but also plays an important role in neural stem cells.

To look at the timeline of neural induction, we analyzed a panel of neural markers. Pax6 was detected starting at the day 6 time point. This data is consistent with immunostaining, which showed Pax6 expression in the majority of cells analyzed at day 5. Additional neural markers, including Sox1, Otx1, and Mash1, all show a similar pattern of expression: absence of detectable levels in ES cells with upregulated expression during the first week of differentiation as cells are committing to a neural fate. Importantly, Sox1 is detected at an earlier time point than Otx1 and Mash1, consistent with the fact that in mice Sox1 is expressed earlier than the other markers.

To look at differentiation towards a projection neuron phenotype, Fezf2, COUP-TF1, and CTIP2 expression was analyzed. COUP-TF1 is another

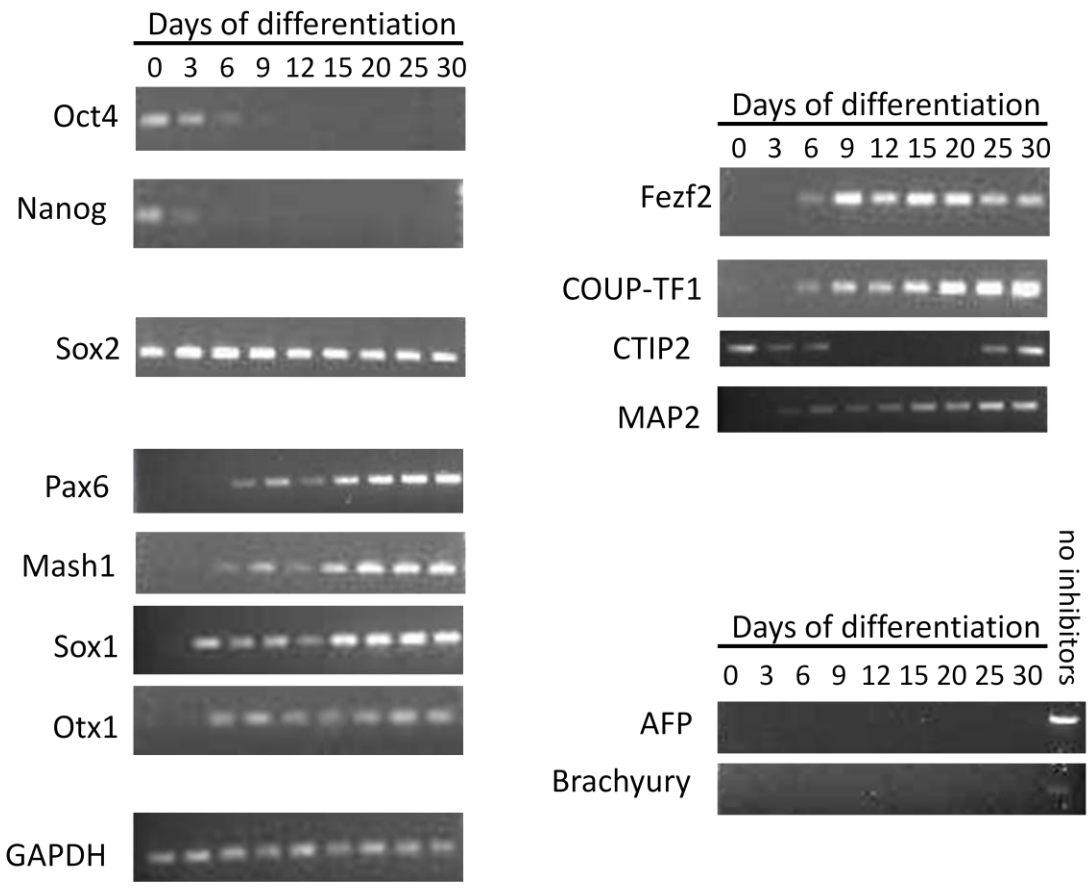


Figure 4.5: RT-PCR analysis of key markers throughout differentiation. RNA samples were harvested from cells at multiple time points during differentiation to assess expression of pluripotency, neural, and cortical markers as cells differentiated. GAPDH was run as a loading control.

transcription factor that has been shown to play an important role in projection neuron specification[23]. Fezf2 and COUP-TF1 showed similar patterns of expression, first being detected at day 6 of differentiation. This is consistent with mouse development, where Fezf2 is first expressed in neural progenitor cells. Therefore, it would be expected that if these pathways are conserved in human cells, Fezf2 induction would show a similar pattern of induction as other neural markers. To look downstream of Fezf2, CTIP2 expression was also analyzed. While there is some CTIP2 RNA detected at early time points, it is quickly downregulated. Expression is then detected again at day 25 of differentiation, increasing at day 30. This is consistent with immunostaining, which, following the initial downregulation, is unable to detect CTIP2 protein until after approximately 20 days of differentiation. Taken together, the expression timeline of a subset of key developmental markers indicate there is some conservation of developmental pathways between mouse and human cells.

Analysis of our modified dual SMAD inhibition differentiation method clearly demonstrates efficient neural induction; however, it is also important to look at other germ layer markers to determine if substantial amounts of either endoderm or mesoderm are also being produced in this culture system. To address this, we looked at alphafetoprotein (AFP, endoderm) and Brachyury (mesoderm) expression in our culture system, as compared to culturing without any inhibitors. Neither AFP nor Brachyury were detected at any time points in the dual SMAD inhibition culture. In contrast, both markers were

detected in cultures not treated with inhibitors. This demonstrates that (1) our cells are capable of differentiating into cells of all three germ layers and (2) dual SMAD inhibition is efficient at blocking differentiation towards endoderm and mesoderm lineages, resulting in efficient neural induction.

4.3.6 Purification of Fezf2-expressing cells via FACS

To study Fezf2-expressing cells in depth, it is critical to be able to purify Fezf2-expressing cells from non-expressers. To achieve this, we performed FACS sorting on live cells at day 19 of differentiation. This time point was chosen for a number of reasons. Fluorescence (resulting from Fezf2 reporter expression) is not detected until approximately day 10 of differentiation, therefore any time point earlier than 10 days is not feasible due to lack of Fezf2 expression. In addition, it has been shown in mice Fezf2 is expressed more broadly in neocortical progenitors at early time points, and expression is refined to layer V cells later in development. Therefore, we chose a later time point to increase the probability of Fezf2-expressing cells representing layer V neuronal lineage, as opposed to a progenitor that may or may not later become a layer V neuron. On the other hand, because neurons become more fragile and sensitive to dissociation as they mature, we did not want to sort very late stage neurons due to the risk of neuron death while preparing cells for sorting. For all these reasons, we chose day 19 as a time point for our initial purification of Fezf2-expressing cells.

Cells prepared for sorting at day 19 of differentiation showed a good survival rate, with approximately 70-80% of cells alive throughout the sort, as determined by propidium iodide. Sorting for EYFP, it was determined that approximately 3% of surviving cells were Fezf2+ **(Figure 4.6)**. To confirm the purity of our sorting conditions, sorted samples were run on the BD FACS Canto. The EYFP- population was 99.7% pure and the EYFP+ population was 97% pure, indicating our sorting conditions result in pure populations with very few contaminating cells.

4.4 Discussion

Here we report a directed differentiation method to increase the efficiency of generating Fezf2-expressing neurons. In mice, Fezf2 has been shown to be critical for the birth and specification of CSMNs[22]. To assess if these same pathways are conserved in human cells, it is highly desirable to be able to visualize cells expressing Fezf2 in order to characterize these cells in greater depth. Using our targeted fluorescent reporter line, we have successfully visualized Fezf2-expressing cells both in live culture as well as through immunostaining for the fluorescent marker (EYFP).

To develop a directed differentiation protocol, we applied known developmental mechanisms from other species to direct cells towards our desired cell fate. Dual SMAD inhibition achieves highly efficient neural induction by inhibiting differentiation towards the other germ layers,

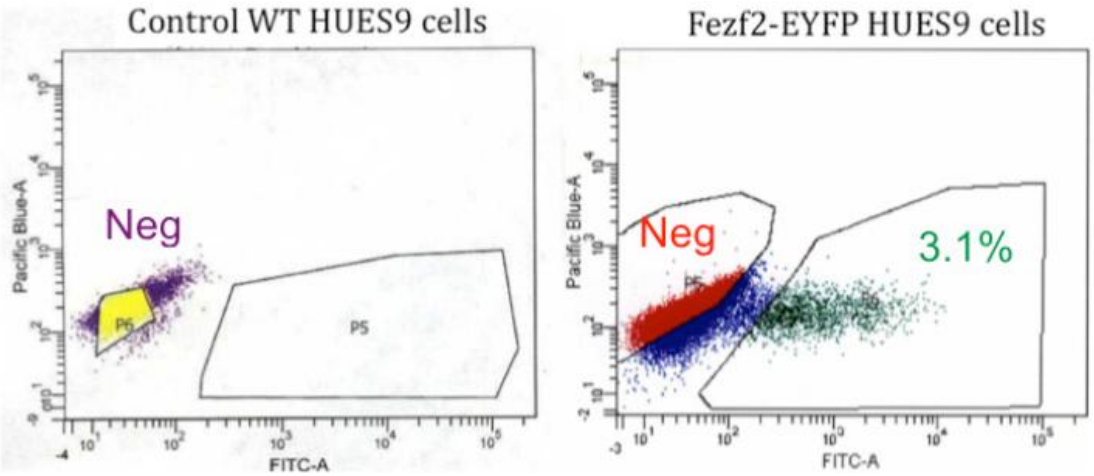


Figure 4.6: Purification of Fezf2-expressing cells via FACS. Cells were sorted for EYFP on day 19 of differentiation. Control wild type HUES-9 cells were used as a negative control to set gates. Approximately 3% of Fezf2-EYFP cells expressed Fezf2 as determined by EYFP expression.

mesoderm and endoderm. To analyze the progression of cells during the differentiation process, immunostaining was performed to monitor loss of pluripotency markers and simultaneous induction of neural markers. We initially compared two different neural media recipes to determine if the composition of media affects the differentiation process. Analysis revealed there was a difference in the two systems, with the timeline of differentiation appearing to be slightly different. In addition, cells cultured in N2B27 media were more homogenous across the plate and showed more efficient induction of Otx2, a marker expressed in anterior neuroectoderm. For these reasons, we chose to use the N2B27 system for our future cultures. The pluripotency marker, Oct4, showed rapid downregulation, with only few cells maintaining expression by day 5. As expected, cells uniformly expressed Pax6 and Nestin at day 5, indicating efficient neural induction. By day 10, Otx2 expression was also detected in a majority of the cells, indicating an anterior neuroectoderm identity.

To pattern the cells towards a dorsal fate, inhibition of sonic hedgehog signaling via treatment with cyclopamine was applied. Cyclopamine treatment showed approximately a 2-fold increase in Fezf2 expression as compared to dual SMAD inhibition alone, and a striking 7-fold increase over cultures not treated with any patterning factors. Analysis of cultures at 30 days of differentiation revealed Fezf2-expressing cells also expressed the neuronal marker β-III-tubulin, indicating Fezf2-expressing cells are neurons. Immunostaining for CTIP2, a marker downstream of Fezf2 in mice, showed a

subset of CTIP2-expressing cells also expressed Fezf2. This data is an important piece of evidence suggesting some of the mechanisms of CSMN development and specification are conserved between mouse and humans.

Further evidence suggesting conservation of developmental pathways between mouse and human was shown through a time course study looking at some key markers throughout differentiation. Pluripotency markers *Oct4* and *Nanog* were rapidly downregulated, as expected. Neural markers *Sox1*, *Mash1*, and *Otx1* were upregulated early in differentiation, at times consistent with neural induction. Markers of alternate germ layers, *AFP* and *Brachyury*, were not detected in our directed differentiation culture method at any time point, indicating efficient neural induction with undetectable levels of mesoderm and endoderm cells. Looking more specifically at makers of subcerebral projection neurons, *Fezf2* and *COUP-TF1* were detected approximately one week after differentiation, in a time course similar to when protein expression is first detected via our fluorescent reporter. The downstream marker, *CTIP2*, was detected late in expression, much later than Fezf2 expression, and also at a time consistent with when immunostaining is first able to detect protein expression.

The presence of cells co-expressing Fezf2 and CTIP2 is enticing evidence suggesting Fezf2 plays a similarly important role in human CSMN development as has been shown in mice. Determining if Fezf2 plays the same master regulatory role for CSMN development in human cells is critical. Elegant studies have been done in mice demonstrating the importance of

Fezf2 and its role in CSMN fate determination. However, if these mechanisms are not conserved in humans, these studies will not be translatable to human cells. The evidence we present here is the first analysis of human Fezf2-expressing cells. Our data suggests there is some conservation between mouse and human CSMN development, indicating much of the mouse work might be transferred to human studies. This has vast implications for both basic biology research and for the potential of developing future therapies.

4.5 Conclusions

We present here the first analysis of human Fezf2-expressing cells, using our targeted fluorescent reporter line. Results from this analysis suggest at least some key developmental pathways are likely to be conserved between mouse and human cells. This would allow for translation of the mouse studies to human cells, and also provide a platform to perform meaningful in vivo assays via cell transplantation in animal models that cannot be done in humans. Taken together, building on existing knowledge in the literature in combination with our development and utilization of a knockin reporter line, we have developed a directed differentiation method that is highly efficient at neural conversion and is capable of generating Fezf2-expressing neurons. These neurons (or their Fezf2-expressing progenitors) can then be purified through FACS sorting via the fluorescent reporter for further analysis to decipher their molecular identity and functional properties in relation to the CSMN cell fate via a range of in vitro and in vivo assays.

4.6 References

1. Mummery CL, Ward D, Passier R. Differentiation of human embryonic stem cells to cardiomyocytes by coculture with endoderm in serum-free medium. Curr Protoc Stem Cell Biol. 2007;Chapter 1:Unit 1F 2.

2. Kawasaki H, Mizuseki K, Nishikawa S, et al. Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity. Neuron. 2000;28:31-40.

3. Attisano L, Wrana JL. Mads and Smads in TGF beta signalling. Curr Opin Cell Biol. 1998;10:188-194.

4. Attisano L, Wrana JL. Signal transduction by the TGF-beta superfamily. Science. 2002;296:1646-1647.

5. Heldin CH, Miyazono K, ten Dijke P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature. 1997;390:465-471.

6. Massague J. TGF-beta signal transduction. Annu Rev Biochem. 1998;67:753-791.

7. Whitman M. Smads and early developmental signaling by the TGFbeta superfamily. Genes Dev. 1998;12:2445-2462.

8. Wrana JL. Crossing Smads. Sci STKE. 2000;2000:re1.

9. Smith WC, Harland RM. Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in Xenopus embryos. Cell. 1992;70:829-840.

10. Valenzuela DM, Economides AN, Rojas E, et al. Identification of mammalian noggin and its expression in the adult nervous system. J Neurosci. 1995;15:6077-6084.

11. Elkabetz Y, Panagiotakos G, Al Shamy G, et al. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes Dev. 2008;22:152-165.

12. Lee H, Shamy GA, Elkabetz Y, et al. Directed differentiation and transplantation of human embryonic stem cell-derived motoneurons. Stem Cells. 2007;25:1931-1939.

13. Smith JR, Vallier L, Lupo G, et al. Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. Dev Biol. 2008;313:107-117.

14. Chambers SM, Fasano CA, Papapetrou EP, et al. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol. 2009;27:275-280.

15. Ribes V, Briscoe J. Establishing and interpreting graded Sonic Hedgehog signaling during vertebrate neural tube patterning: the role of negative feedback. Cold Spring Harb Perspect Biol. 2009;1:a002014.

16. Mo Z, Zecevic N. Is Pax6 critical for neurogenesis in the human fetal brain? Cereb Cortex. 2008;18:1455-1465.

17. Jessell TM. Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat Rev Genet. 2000;1:20-29.

18. Ulloa F, Marti E. Wnt won the war: antagonistic role of Wnt over Shh controls dorso-ventral patterning of the vertebrate neural tube. Dev Dyn.239:69-76.

19. Chen B, Schaevitz LR, McConnell SK. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. Proc Natl Acad Sci U S A. 2005;102:17184-17189.

20. Chen B, Wang SS, Hattox AM, et al. The Fezf2-Ctip2 genetic pathway regulates the fate choice of subcortical projection neurons in the developing cerebral cortex. Proc Natl Acad Sci U S A. 2008;105:11382-11387.

21. Chen JG, Rasin MR, Kwan KY, et al. Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. Proc Natl Acad Sci U S A. 2005;102:17792-17797.

22. Molyneaux BJ, Arlotta P, Hirata T, et al. Fezl is required for the birth and specification of corticospinal motor neurons. Neuron. 2005;47:817-831.

23. Tomassy GS, De Leonibus E, Jabaudon D, et al. Area-specific temporal control of corticospinal motor neuron differentiation by COUP-TFI. Proc Natl Acad Sci U S A.107:3576-3581.

Chapter 5

Whole genome gene expression analysis of Fezf2-EYFP+ cells

5.1 Abstract

Direct studies into human CSMN development are not practical due to inaccessibility. Fezf2 has been shown to be a critical gene involved in mouse CSMN specification. To begin to study human CSMNs, we utilized our *Fezf2* fluorescent reporter cell line to purify and analyze Fezf2+ cells. Using whole genome gene expression analysis, 450 significant genes were identified that were differentially expressed between Fezf2+ and Fezf2- populations. This data offers a first look at presumptive human CSMNs and offers insight into mechanisms involved in human CSMN specification and guidance for future studies.

5.2 Introduction

Advances in molecular biology techniques have equipped researchers with a variety of tools to study a particular cell type of interest in great detail. These tools have become very powerful, requiring less sample input and providing large amounts of high quality data. One such powerful tool that offers great insight into the expression level of genes as a result of treatment, disease, or during different developmental stages is the DNA microarray.

One of the numerous advancements in DNA microarray technology is the increase in the number of probes that can be spotted onto one array. This has allowed for whole genome gene expression analysis on a single array, providing powerful insight into changes in gene expression levels between experimental samples. The DNA microarray is often a first step in analyzing a

new cell type to take a comprehensive and unbiased look at which genes and pathways may be most important in the particular system being studied.

This approach was applied to the study of mouse corticospinal motor neurons (CSMNs) in an attempt to identify key genes that may play a critical role in CSMN development and specification[1]. A number of genes were identified that were specific to CSMN, and not expressed in closely related cell types, including Fezf2. Subsequent studies have validated the critical role Fezf2 plays in CSMN in mice[1-7], demonstrating the power of the microarray, and the potential to gain guidance for future studies.

The role of Fezf2 has not yet been studied in human cells, nor is known to what extent CSMN development is conserved between the mouse and human systems. Our preliminary studies looking at key markers, including CTIP2, indicate there is some conservation between the two systems. However, looking only at genes that have been identified in the mouse system limits the scope of analysis, and does not allow for identification of genes that may be of importance in human development that do not play a role in mouse development. To perform an unbiased, whole genome analysis of gene expression levels in the human system, we performed DNA microarray analysis on a purified population of Fezf2-expressing cells as assessed with the *Fezf2-EYFP* reporter. This data is the first genome-wide analysis of human cells of presumptive CSMN lineage and may guide future studies of this scientifically and clinically important cell population.

5.3 Results

Initial analysis of Fezf2+ cells (Chaper 4) has implicated conservation of some of the mechanisms described in the mouse corticospinal system. However, these methods are limited in scope since only genes identified as playing a key role in mouse corticospinal development have been examined. To take a comprehensive, unbiased look at all genes that may play a role in human cells, we performed gene expression analysis. Whole genome gene expression analysis comparing Fezf2+ to Fezf2- cells resulted in identification of 450 highly significant genes, based on a false discovery rate (FDR) <0.01. (**Figure 5.1)**.

5.3.1 Conservation of gene expression between mouse CSMN and human Fezf2+ cells

Genes that play an important role in CSMN development and specification in mice have been identified[1]. These genes were identified by gene expression analysis comparing purified populations of CSMN to related neuron types, including callosal projection neurons and corticotectal projection neurons. However, because direct studies of human CSMNs are not feasible due to inaccessibility, other methods are necessary to study CSMN function in human cells.

To address this question, we performed whole genome gene expression analysis on a purified population of Fezf2+ neurons, and compared expression

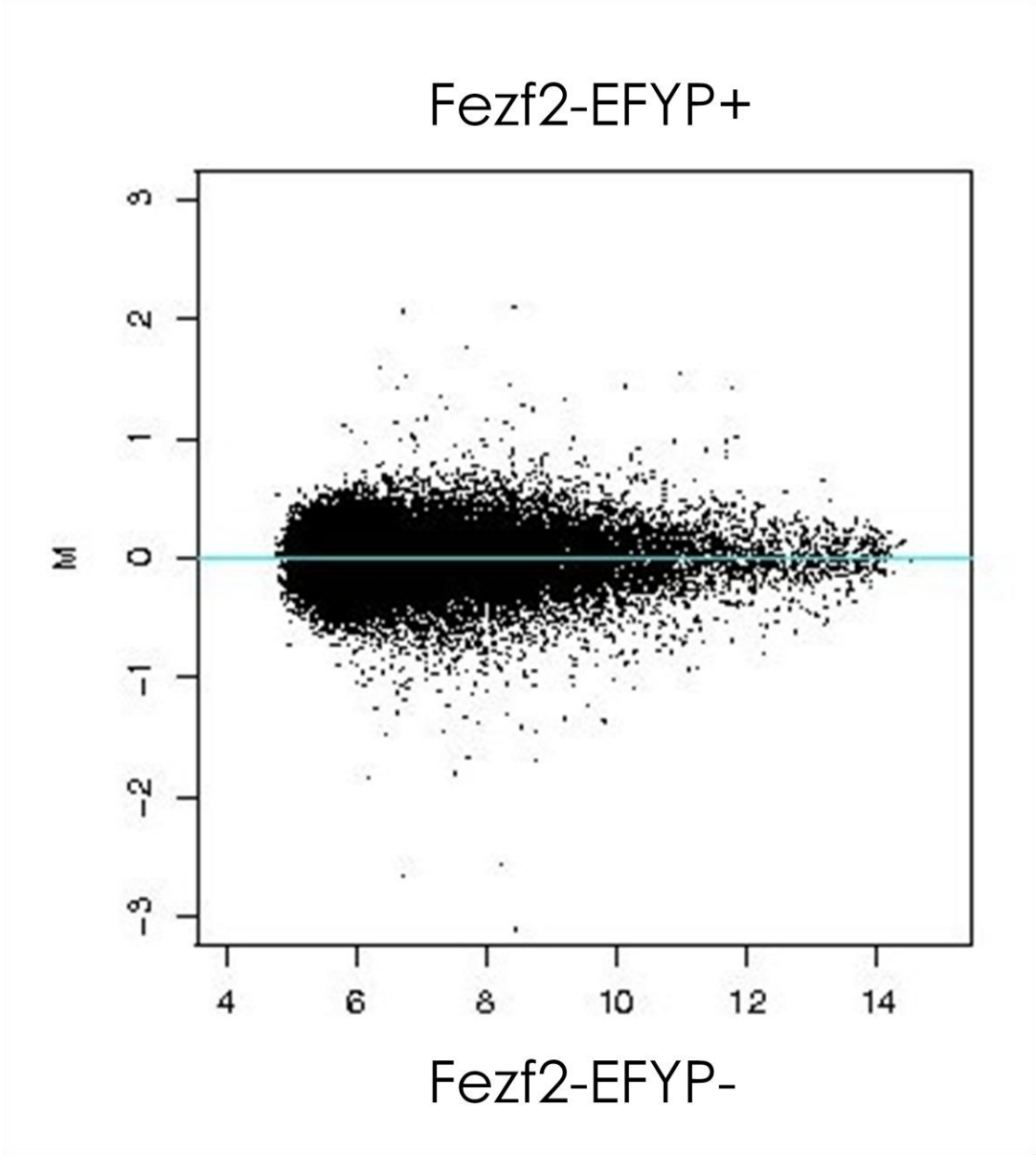


Figure 5.1: Scatter plot of normalized microarray data. Gene expression analysis comparing Fezf2+ versus Fezf2- populations of cells reveals numerous genes differentially regulated between the two populations. The x-axis represents the absolute expression level of each gene expressed as a log2 ratio (e.g. 6.5 on the x-axis represents an expression value of ~90). The y-axis represents the fold change in expression between populations, also expressed as a log2 ratio (e.g. 2 on the y-axis represents a 4-fold change in expression).

to the Fezf2- population. As expected, Otx1, an anterior neural marker, was upregulated in Fezf2+ as compared to the Fezf2- population. Additionally, markers for early neocortical progenitors such as FoxG1 and EMX2 were also upregulated in the Fezf2+ population. Expression of fgf8 was also significantly lower in Fezf2+ cells, which is to be expected as fgf8 is expressed along the midbrain/hindbrain border[8, 9], and *Fezf2* labels forebrain cells.

To look more in depth into area-specific gene expression within the cortex, we examined gene expression changes of known genes identified to play a key role in mouse CSMN development and related neuron types. A number of genes shown to be involved in mouse CSMN were also upregulated in human Fezf2+ cells, including NR2F1 (COUP-TF1, birth and timing of CSMNs), lumican (intermediate CSMN development), and encephalopsin (late CSMN development) **(Table 5.1)**. In addition, genes that have been identified as being expressed in related projection neurons, but not expressed in CSMNs were downregulated in the Fezf2+ population. This includes dkk3 and LMO4, which are expressed in callosal projection neurons, but not in CSMNs.

Expression of Tbr1, a gene expressed highly in layer VI neurons, was also found to be downregulated in Fezf2+ cells. This is an important finding as Fezf2 has been shown to be expressed in progenitor cells of both layers V and VI, but is later downregulated in layer VI and maintained in layer V. Lower Tbr1

Table 5.1: Summary of a subset of differentially regulated genes. Genes upregulated in Fezf2+ cells as compared to Fezf2- cells are indicated with a (+); Genes downregulated in Fezf2+ cells are indicated with a (-).

Gene	Change	Function
OTX1	+	anterior
FOXG1	+	early neocortical progenitors
EMX2	+	early neocortical progenitors
Encephalopsin (OPN3)	+	CSMN late development
lumican	+	CSMN intermediate development
NR2F1 (COUP-TF1)	+	timing and birth of CSMN
fgf8	-	midbrain/hindbrain border
TBR1	-	high expression in layer VI
dkk3	-	negative for CSMN
LMO4	-	negative for CSMN
CD47	+	cell surface antigen
CD24	-	cell surface antigen
CD83	+	cell surface antigen
CD63	+	cell surface antigen
CD46	-	cell surface antigen
CD9	-	cell surface antigen
CD164	-	cell surface antigen
ROBO1	-	axon guidance
ROBO2	+	axon guidance
ROBO3	+	axon guidance
SLIT2	+	axon guidance
SEMA5A	-	axon guidance
SEMA3A	+	axon guidance
SEMA3C	-	axon guidance
SEMA4D	-	axon guidance
SEMA4C	+	axon guidance
SEMA3E	+	axon guidance

expression in the Fezf2+ population suggests, at least at this stage, the majority of cells expressing Fezf2 are layer V cells. This data, in combination with markers of CSMN identity, suggests Fezf2 is a marker for CSMNs in human cells, as it is for mouse CSMNs.

5.3.2 Newly identified genes

To identify genes that may play an important role in human cells, we looked at genes that showed the highest fold change between populations **(Table 5.2)**. Analysis revealed a number of these genes are zinc finger genes, similar to Fezf2. The gene showing the largest differential regulation between the two populations is zinc finger protein 22 (ZNF22/KOX15) which is expressed almost 10-fold less in Fezf2+ cells as compared to Fezf2- cells. There is currently not much known about ZNF22, other than it likely plays a role in transcriptional regulation. It has also been identified as potentially playing a role in tooth development[10]. Also largely downregulated in Fezf2+ cells is transcription factor CP2 (TFCP2), with a 7-fold difference in expression. TFCP2 has been shown to bind a variety of promoters, including fibrinogen, alpha-globin, SV40, and HIV-1 promoters[11-13]. Polymorphisms in TFCP2 have been linked in many studies to being involved in the pathogenesis of Alzheimer's disease[14-17].

The gene showing the largest upregulation in Fezf2+ cells is zinc finger protein 69 (ZNF69). Very little is known about this gene other than it likely plays a role in transcriptional regulation via DNA binding and it is located on

Table 5.2: Summary of genes with the highest fold change in expression. Genes upregulated in Fezf2+ cells as compared to Fezf2- cells have a positive fold change; genes downregulated in Fezf2+ cells have a negative fold change.

Gene	Fold change
Zinc finger protein 22 (KOX 15) (ZNF22)	9.787
Transcription factor CP2 (TFCP2)	7.007
Zinc finger protein 69 (ZNF69)	-4.306
TSPY-like 5 (TSPYL5)	3.655
Ribosomal protein L39-like (RPL39L)	-3.716
Catalase (CAT)	-2.849
Protocadherin beta 5 (PCDHB5)	2.489
Chromosome 16 open reading frame 33	2.351
Heat shock 70kDa protein 2 (HSPA2)	2.216
Collagen, type III, alpha 1 (COL3A1)	2.161
Distal-less homeobox 5 (DLX5)	2.042
Solute carrier family 44, member 2 (SLC44A2)	2.006
Solute carrier family 43, member 2 (SLC43A2)	1.996
Tripartite motif-containing 4 (TRIM4), transcript	-1.936
TEK tyrosine kinase, endothelial	-1.930
Ephrin-B1 (EFNB1)	1.948
Mesoderm specific transcript homolog(MEST)	1.902
Solute carrier family 16, member 14 (SLC16A14)	1.902

chromosome 22 along with numerous other zinc finger genes[18, 19]. Other genes showing much higher expression in Fezf2+ cells include ribosomal protein L39-like (RPL39L) and catalase (CAT). Further studies will be necessary to determine what role these genes might play in human CSMN development.

5.3.3 Potential cell surface antigen signature

One common method to purify a cell type of interest is to identify a cell surface antigen signature that is unique to the desired cell type, and use this signature to purify cells via FACS. Gene expression analysis revealed a series of CD genes that are differentially expressed between the two populations, offering a potential starting point to identify a unique signature for Fezf2+ cells. This would allow for purification of Fezf2+ cells from all stem cell lines, regardless of whether or not they contain a fluorescent reporter.

In the Fezf2+ population, CD47, CD 63, and CD 83 were upregulated; in contrast, CD9, CD 24, CD46, and CD164 were downregulated. Therefore, combination of these markers may allow for purification of a CD9-/CD24-/CD46-/CD47+/CD63+/CD83+/CD164- population of cells that are Fezf2+. This can be tested by sorting cells expressing the *Fezf2* fluorescent reporter and determining if the CD cell surface signature is successful at distinguishing between YFP+ and YFP- cells.

5.3.4 Expression of axon guidance molecules

The process of finding the correct path for an axon to follow in order to reach its intended destination is a complex process that requires multiple guidance cues along the way. Gene expression analysis revealed a number of guidance molecules that show differential expression between Fezf2+ and Fezf2- cells. These molecules include members of the Eph/ephrin, ROBO/Slit, and semaphorin/plexin families.

The Eph-ephrin family of genes play an important role in axon guidance, among other things, and typically exert a repulsive effect to cells. Fezf2+ cells showed significantly lower expression of receptors EphA4, EphA10, and EphB3. Expression of ligands ephrinA1 and ephrinB1 were also significantly lower in Fezf2+ cells. In contrast, expression of receptor EphB3 was significantly increased in the Fezf2+ population.

Similar to the Eph/ephrin family, ROBO/Slit work together as a receptor/ligand team. Members of the ROBO family are highly conserved transmembrane glycoproteins that serve as receptors for the Slit family of ligands. Similar to ephrins, Slit molecules participate in axon guidance and generally exert a repulsive effect on cells. Three of the four ROBO receptors were identified as differentially expressed, with lower ROBO1 expression and higher ROBO2 and ROBO3 expression in the Fezf2+ population. Of the three Slit ligands, only Slit2 showed significantly different expression, with increased expression in Fezf2+ cells.

Semaphorins are a family of secreted and membrane proteins that act as guidance cues for axonal growth cones. Plexins, along with neuropilins, act as the receptors for semaphorins. Plexin BI is the only member of the plexin family identified in the gene expression analysis, and was found to be expressed at significantly higher levels in Fezf2+ cells. While only one receptor was identified as being differentially regulated, a variety of semaphorins were identified. Fezf2+ cells showed higher expression of SEMA3A, SEMA3E, and SEMA4C and decreased expression of SEMA3C, SEMA4D, and SEMA5A when compared to the Fezf2- population. Together, this data may provide important insight into molecules that are important for path finding in CSMNs.

5.4 Discussion

Whole genome gene expression analysis offers great insight into mechanisms that may play an important role in the differentiation and guidance of specific neuronal lineages. While much has been learned about genes that are involved in the development and specification of mouse CSMNs, it is unknown how much of those pathways is conserved in human CSMNs. Because direct studies of human CSMNs is not practical due to inaccessibility, the best method to gain insight into human CSMNs is to direct ES cells towards a corticospinal fate, and then analyze this population. To do this, we used a *Fezf2* reporter line to label cells of presumptive CSMN lineage, purified cells expressing *Fezf2*, and performed whole genome gene expression analysis comparing the Fezf2+ population to the Fezf2- population.

As a result, 450 genes were identified as being significantly different in expression level between the two populations. While extensive follow-up studies will be necessary to determine which genes play an important role, initial analysis reveals a number of genes expressed highly in mouse CSMNs are also upregulated in the Fezf2+ population, lending support to indicate Fezf2+ cells may in fact be represent the human CSMN lineage. Additionally, a variety of axon guidance molecules were identified, which may lend insight to which molecules play a key role in guiding the axons of CSMNs to their final target. Identifying which repulsive molecules are involved in axon path finding for CSMNs may also provide valuable clues for spinal cord injury studies. As repulsive guidance cues may be inhibitory to axon regrowth, targeting such repulsive molecules may aid in relieving some of the inhibitory environment that exists after spinal cord injury and contributes to the inhibition of regeneration of damaged axons.

While the microarray data shows upregulation in the Fezf2+ population of genes shown to be important for mouse CSMN development, this is not sufficient to definitively state Fezf2+ cells are CSMNs. It has also been shown in mice that Fezf2 is expressed in other related cell types, and at least transiently, is expressed outside of layer V. Therefore, because it is not possible to determine the exact layer identity of a cell in culture, it is important to note that additional markers would be of great benefit to identifying human CSMN in vitro. Currently, *Fezf2* is the most specific known marker for the CSMN lineage, and is therefore the ideal starting place to begin studying human

CSMNs. Further studies will provide greater insight into the identity of human CSMN and may identify a marker or panel of markers that specifically identify human CSMNs and can separate them from related neuron lineages.

5.5 Conclusions

Described here is the first look at gene expression in human Fezf2+ cells. Gene expression analysis revealed 450 highly significant genes that are differentially expressed between Fezf2+ and Fezf2- cells. Of these genes, some are genes identified as being important for mouse CSMN development. A number of these genes were upregulated in Fezf2+ cells, supporting the hypothesis that *Fezf2* is a marker for CSMNs in human cells also. Numerous axon guidance molecules were also identified, lending insight to which receptors and ligands may play a key role in path finding for CSMNs. Additional analysis of this data will be necessary to achieve a comprehensive understanding and may even lead to identification of a marker or panel of markers that specifically label human CSMNs.

5.6 References

1. Arlotta P, Molyneaux BJ, Chen J, et al. Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. Neuron. 2005;45:207-221.

2. Chen B, Schaevitz LR, McConnell SK. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. Proc Natl Acad Sci U S A. 2005;102:17184-17189.

3. Chen B, Wang SS, Hattox AM, et al. The Fezf2-Ctip2 genetic pathway regulates the fate choice of subcortical projection neurons in the

developing cerebral cortex. Proc Natl Acad Sci U S A. 2008;105:11382-11387.

4. Chen JG, Rasin MR, Kwan KY, et al. Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. Proc Natl Acad Sci U S A. 2005;102:17792-17797.

5. McKenna WL, Betancourt J, Larkin KA, et al. Tbr1 and Fezf2 Regulate Alternate Corticofugal Neuronal Identities during Neocortical Development. J Neurosci.31:549-564.

6. Molyneaux BJ, Arlotta P, Hirata T, et al. Fezl is required for the birth and specification of corticospinal motor neurons. Neuron. 2005;47:817-831.

7. Rouaux C, Arlotta P. Fezf2 directs the differentiation of corticofugal neurons from striatal progenitors in vivo. Nat Neurosci.13:1345-1347.

8. Sunmonu NA, Li K, Guo Q, et al. Gbx2 and Fgf8 are sequentially required for formation of the midbrain-hindbrain compartment boundary. Development.138:725-734.

9. Ohuchi H, Yoshioka H, Tanaka A, et al. Involvement of androgen-induced growth factor (FGF-8) gene in mouse embryogenesis and morphogenesis. Biochem Biophys Res Commun. 1994;204:882-888.

10. Gao Y, Kobayashi H, Ganss B. The human KROX-26/ZNF22 gene is expressed at sites of tooth formation and maps to the locus for permanent tooth agenesis (He-Zhao deficiency). J Dent Res. 2003;82:1002-1007.

11. Bing Z, Reddy SA, Ren Y, et al. Purification and characterization of the serum amyloid A3 enhancer factor. J Biol Chem. 1999;274:24649-24656.

12. Romerio F, Gabriel MN, Margolis DM. Repression of human immunodeficiency virus type 1 through the novel cooperation of human factors YY1 and LSF. J Virol. 1997;71:9375-9382.

13. Huang HC, Sundseth R, Hansen U. Transcription factor LSF binds two variant bipartite sites within the SV40 late promoter. Genes Dev. 1990;4:287-298.

14. Banerjee D, Nandagopal K. Potential interaction between the GARS-AIRS-GART Gene and CP2/LBP-1c/LSF transcription factor in Down syndrome-related Alzheimer disease. Cell Mol Neurobiol. 2007;27:1117-1126.

15. Bertram L, Parkinson M, McQueen MB, et al. Further evidence for LBP-1c/CP2/LSF association in Alzheimer's disease families. J Med Genet. 2005;42:857-862.

16. Schahab S, Heun R, Schmitz S, et al. Association of polymorphism in the transcription factor LBP-1c/CP2/LSF gene with Alzheimer's disease and major depression. Dement Geriatr Cogn Disord. 2006;22:95-98.

17. Xu Y, Kim HS, Joo Y, et al. Intracellular domains of amyloid precursor-like protein 2 interact with CP2 transcription factor in the nucleus and induce glycogen synthase kinase-3beta expression. Cell Death Differ. 2007;14:79-91.

18. Aubry M, Marineau C, Zhang FR, et al. Cloning of six new genes with zinc finger motifs mapping to short and long arms of human acrocentric chromosome 22 (p and q11.2). Genomics. 1992;13:641-648.

19. Dunham I, Shimizu N, Roe BA, et al. The DNA sequence of human chromosome 22. Nature. 1999;402:489-495.

Chapter 6

Conclusion and Future Directions

6.1 Conclusion

Human embryonic stem cells offer great promise both as a tool for studying developmental processes and as a renewable source of cells for regenerative medicine. One obstacle that must be overcome before stem cells can be widely used for regenerative medicine is the ability to efficiently direct them from the embryonic state towards a particular cell type. As we gain a greater understanding of the mechanisms controlling development and cell type specification, we will improve our ability to develop effective therapeutic treatments for patients suffering from a variety of diseases.

I have successfully targeted the h*Fezf2* gene in HUES-9 human embryonic stem cells, which will facilitate the study of differentiation of HUES-9 cells towards subcortical projection neurons, including corticospinal neurons, in vitro and in animal transplantation models. Importantly, this strategy is applicable to all genes, and not just those genes expressed in ES cells. As expected, there is no fluorescence detected in undifferentiated HUES-9-h*Fezf2* ES cells, since *Fezf2* is not expressed in ES cells. However, upon differentiation, a subset of cells express EYFP, indicating the fluorescent reporter is functional.

Using the *Fezf2* reporter cell line, I performed the first studies looking at *Fezf2*-expressing human cells. Multiple lines of evidence, including immunostaining and RT-PCR analysis, suggest at least some key developmental pathways are conserved between mouse and human cells. This would allow for translation of the mouse studies to human cells, and also provide a platform to perform meaningful in vivo assays that cannot be done

in humans. Taken together, this data demonstrates that building upon existing literature, I have developed a directed differentiation method that is highly efficient at neural conversion and is capable of generating Fezf2-expressing neurons. These neurons can then be purified through FACS sorting via the fluorescent reporter for further analysis to determine if these cells are in fact CSMNs.

Whole genome gene expression analysis of purified Fezf2+ cells revealed 450 highly significant genes, as defined by a false discovery rate (FDR) less than 0.01. Pathway analysis showed numerous neural and brain development pathway genes are differentially expressed between Fezf2+ and Fezf2- cells. This is to be expected since these cells are the result of a neural differentiation protocol. In addition, a number of genes identified in the mouse system as being involved in corticospinal motor neuron development were up regulated in Fezf2+ cells. This is compelling evidence to support the hypothesis that there is some evolutionary conservation between mouse and human corticospinal development. More importantly, this provides added support that *Fezf2* is a good marker for the human CSMN lineage.

While there are numerous lines of evidence supporting the hypothesis that *Fezf2* is a marker for the human CSMN lineage, as has been shown in mice, this evidence is not sufficient to definitively say that Fezf2+ cells are indeed CSMNs. The ultimate experiment that is necessary to determine if Fezf2-expressing cells are CSMNs, or have the potential to become CSMNs, is

transplantation of cells into an animal model. If cells send axons from the cortex to the spinal cord, then it can be concluded that Fezf2+ cells have the potential to become human CSMNs.

6.2 Future Directions

The *Fezf2* fluorescent reporter cell line makes possible a number of important experiments that would otherwise not be possible. The fluorescent reporter provides an ideal system for testing potential drugs or molecules that may increase the percentage of Fezf2-expressing cells. The ability to FACS purify live *Fezf2*-expressing cells provides a source of cells to perform in depth analysis and gain greater insight into the molecular identity and developmental potential of these cells. Additionally, purified *Fezf2*-expressing cells can be transplanted into an animal model to determine where these cells send their projections.

Because human ES cells are capable of forming all cell types in the human body, it is not surprising the efficiency of generating any one cell type is inefficient. To increase the efficiency, it is important to have a method of identifying your cell type of interest to assess efficiency of various methods. The Fezf2 fluorescent reporter cell line provides a fluorescent read out in live cells of Fezf2 expression. Because Fezf2 expression can be detected in real time in living cells, it is possible to utilize this cell line to test potential drugs or molecules that have the ability to affect Fezf2 expression. The most efficient method to achieve this goal is to perform high throughput screening of

libraries of compounds. This type of analysis may not only reveal a compound that can increase the efficiency of differentiation, but identifying the target of the compound may also reveal some important information about pathways or other genes that regulate or are regulated by *Fezf2*.

One of the major questions that has not been addressed thus far is the time course of *Fezf2* expression. There may be multiple waves of *Fezf2* expression, and cells expressing *Fezf2* at different times may represent related or separate lineages of cells. It has been shown in the mouse system that *Fezf2* is expressed in the progenitors for both layer V and VI cells, but is down regulated in layer VI cells and upregulated in layer V once the cells are finished migrating. However, this has not yet been studied in human cells to determine if a similar pattern of expression is maintained between species.

Purification of live Fezf2+ and Fezf2- cells at different times during differentiation will allow for continued culturing of the purified populations. Cells can then be monitored over time to determine if all Fezf2+ cells retain expression, or if some turn off Fezf2 at later time points. The opposite can also be analyzed: culturing Fezf2- cells and looking at later time points to determine if any cells turn on Fezf2 expression will address the question of whether there are multiple waves of Fezf2 expression.

While *Fezf2* has been shown to be critical for CSMN development, it is not exclusive to the CSMN lineage. Therefore, *Fezf2* expression alone is not sufficient to definitively say a cell is a CSMN. The only way to demonstrate a cell is a CSMN is to look at where the cell projects its axon to. Since CSMN are

the only cells in the cortex to send their projections to the spinal cord, transplantation into an animal model is an important step to address the molecular identity of Fezf2+ cells. One of the key experiments that can be done with the *Fezf2* fluorescent reporter cell line is to transplant Fezf2-expressing cells into the mouse brain and allow the cells to integrate, differentiate, and extend axons. Because EYFP is detected in both the cell body and axon, location of the cell bodies as well as their projections can be detected via fluorescence. This will allow for analysis of survival, integration, migration, and projections.

The first transplantation experiment to be performed should be transplantation of Fezf2+ cells into the cortex. Depending on the age of the mouse at time of transplantation, cells should be targeted to layer V, if layers have been formed. At very early ages (i.e. embryonic and early post natal) layers will not yet have formed, and therefore layer-specific targeting is not possible. If these experiments reveal Fezf2+ cells sending projections to the spinal cord, it can then be concluded that Fezf2-expressing cells have the ability to become CSMNs.

To take it one step further, the next series of experiments would involve transplanting Fezf2-expressing cells to other areas of the brain. This is important because if cells are placed in other areas of the brain, they will not be under the influence of the usual guidance cues and environmental signals CSMNs are typically exposed to. Therefore, if Fezf2+ cells are able to extend axons to the spinal cord when they are ectopically located, then it can be

concluded these cells are fated to become CSMNs. It has been shown in mice that Fezf2 over expression is sufficient to direct cells towards a CSMN fate, and can also redirect cell fate from an unrelated cell type (i.e. medium spiny neuron) to a CSMN fate. If the role of Fezf2 is shown to be as critical in human cells as has been demonstrated in the mouse system, this will be a major advance for the field and will provide an excellent tool for advancing therapies for neurological disorders such as spinal cord injury and amyotrophic lateral sclerosis.