UNIVERSITY OF CALIFORNIA, SAN DIEGO

Toward Engineering the Corticospinal Tract from Human Embryonic Stem Cells

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

in

Biomedical Sciences

by

Anders J. Smith

Committee in charge

Professor Binhai Zheng, Chair
Professor Sylvia Evans
Professor Joseph Gleeson
Professor Alysson Muotri
Professor Karl Willert

2014

The dissertation of Anders J. Smith is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2014

DEDICATION

This dissertation is dedicated to Cami, for all her love, support and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

It is hard to believe how close I am to reaching this major milestone, it seems that it was not long ago that I first set out on this journey. It has been a challenging and rewarding experience.

First and foremost, I would like to acknowledge Cami, my beautiful and supportive wife, and my parents who have always been so proud of their scientist.

Second, I would like to thank and acknowledge Binhai for his mentorship and direction over the past six years, as well as Sandy, my stem cell conspirator, and all the other members of the Zhang lab, past and present.

Finally, I would like to thank Drs. Hui Dong and Kim Barrett, who started me on the path to scientific discovery, my collaborator, James Weimann, the UCSD stem cell core, and all of my many friends and relatives whose support has made this accomplishment a possibility.

Chapter 2, in part is currently being prepared for submission for publication of the material. Smith, Anders & Hinckley, Sandy; Feltham, Casey; Rodriguez, Karen; Zapata, Arturo; Zhang, Sherry; Weimann, James; Zheng, Binhai. “Toward Engineering the Corticospinal Tract from Human Embryonic Stem Cells”. The dissertation author was the co-principal researcher/author of this material.

Chapter 3, in part is currently being prepared for submission for publication of the material. Smith, Anders; Zapata, Arturo; Zhang, Sherry; Feltham, Casey; Zheng, Binhai. “Efficient Labeling of Human Embryonic Stem Cells via BAC Mediated Transgenesis”. The dissertation author was the principal researcher/author of this material.

VITA

2003, B.S. in Biology (Neuroscience and Physiology), University of California, San Diego

2003-2006, Research Assistant, Acting Lab Manager, University of California, San Diego Medical Center, Division of Gastroenterology

2007, Research/Testing Assistant, Althea Dx, San Diego, CA

2014, Ph.D., Biomedical Sciences, University of California, San Diego, School of Medicine

PUBLICATIONS

Smith A, Contreras C, Ko KH, Chow J, Dong X, Tuo B, Zhang HH, Chen DB, Dong H. Gender-specific protection of estrogen against gastric acid-induced duodenal injury: stimulation of duodenal mucosal bicarbonate secretion. Endocrinology. 2008 Sep;149(9):4554-66. Epub 2008 May 22.

Sellers ZM, Mann E, **Smith A**, Ko KH, Giannella R, Cohen MB, Barrett KE, Dong H. Heat-stable enterotoxin of Escherichia coli (STa) can stimulate duodenal HCO3(-) secretion via a novel GC-C- and CFTR-independent pathway. FASEB J. 2008 May;22(5):1306-16. Epub 2007 Dec 20.

Smith AJ, Chappell AE, Buret AG, Barrett KE, Dong H. 5-Hydroxytryptamine contributes significantly to a reflex pathway by which

the duodenal mucosa protects itself from gastric acid injury. FASEB J. 2006 Dec;20(14):2486-95.

Dong H, **Smith A**, Hovaida M, Chow JY. Role of Ca2+-activated K+ channels in duodenal mucosal ion transport and bicarbonate secretion. Am J Physiol Gastrointest Liver Physiol. 2006 Dec;291(6):G1120-8. Epub 2006 Jun 8.

Tuo BG, Sellers ZM, **Smith AJ**, Barrett KE, Isenberg JI, Dong H. A role for CagA/VacA in Helicobacter pylori inhibition of murine duodenal mucosal bicarbonate secretion. Dig Dis Sci. 2004 Nov-Dec;49(11-12):1845-52.

Sellers ZM, Childs D, Chow JY, **Smith AJ**, Hogan DL, Isenberg JI, Dong H, Barrett KE, Pratha VS. Heat-stable enterotoxin of Escherichia coli stimulates a non-CFTR-mediated duodenal bicarbonate secretory pathway. Am J Physiol Gastrointest Liver Physiol. 2005 Apr;288(4):G654-63. Epub 2004 Oct 28.

Dong H, Sellers ZM, **Smith A**, Chow JY, Barrett KE. Na(+)/Ca(2+) exchange regulates Ca(2+)-dependent duodenal mucosal ion transport and HCO(3)(-) secretion in mice. Am J Physiol Gastrointest Liver Physiol. 2005 Mar;288(3):G457-65. Epub 2004 Oct 21.

ABSTRACT OF THE DISSERTATION

Toward Engineering the Corticospinal Tract from

Human Embryonic Stem Cells

by

Anders J. Smith

Doctor of Philosophy in Biomedical Sciences

University of California, San Diego, 2014

Professor Binhai Zheng, Chair

Human embryonic stem cells offer unprecedented potential for the study of human development and carry a significant promise for advancing the future of medical practice. HESCs are particularly useful for the study of tissues and organ systems that exhibit poor endogenous regenerative capabilities, like the adult CNS. The corticospinal tract is part of CNS, and is required for skilled voluntary movement of the hands, fingers, and toes. Fezf2 has been identified as an important marker and regulator of the

corticospinal motor neuron (CSMN) fate in mouse models, and is evolutionarily preserved from drosophila to humans.

Nevertheless, the role of Fezf2 in human cortical and CST development requires further elucidation. Here, using a Fezf2:EYFP knock-in hESC reporter line, we sought to clarify the role of Fezf2 in human cortical development, and attempted to generate human CSMNs.

Through directed differentiation, we were able to guide our cells through the process of cortical development. We found that Fezf2 is broadly expressed in early human cortical stem/ progenitor cells, and verified that these Fezf2 positive cells are mitotic, demonstrating that human Fezf2 expression is not limited to deep-layer post mitotic neurons. Further, we were able to verify that as differentiation progresses, so does the molecular identity of the progeny generated by these Fezf2 positive neural precursors. Gene expression analysis indicates that Fezf2 positive cells give rise to deep and superficial layer projection neurons, as well as glia, while Fezf2 negative cells give rise to other neural subtypes and cells expressing eye-associated markers. Through timing and FACS purification we were able to generate a population of post-mitotic neurons that were 94% Fezf2 positive, many of which co-expressed additional markers of deep-layer projection neurons.

Finally, Fezf2 positive cells transplanted into neonatal mice were able to maintain Fezf2 expression, adopted pyramidal morphologies, and

were able to correctly project down appropriate subcerebral motor pathways, even reaching the pyramidal decussation and the spinal cord.

Additionally presented here is the development and implementation of two BAC constructs for use in permanently labeling populations of hESCs and their progeny.

Chapter 1. Introduction of the Dissertation

1.1 Human Embryonic Stem Cells

Human embryonic stem cells (hESCs) are pluripotent cells, usually derived from the inner cell mass of a human blastocyst (Thomson, Itskovitz-Eldor et al. 1998). As in the case of the Hues9 cell line utilized in this work, these blastocysts most often donated from in vitro fertilization clinics, with signed, informed consent of embryo's donors (Cowan, Klimanskaya et al. 2004). The unparalleled value of these cells for biomedical research is derived from their pluripotent nature. Embryonic stem cells have the unique ability among human cell types to give rise to virtually any specific cell type in the adult human body (Thomson, Itskovitz-Eldor et al. 1998, Cowan, Klimanskaya et al. 2004). Certain diseases, such as type-one diabetes and Parkinson's disease are caused by the loss of just one or two specific cell types. Additionally, many systems in the adult human body, like the central nervous system (CNS) exhibit poor regenerative capacity following injury or degeneration.

The great hope of hESC based research is to develop cell replacement therapies, so that missing or damaged cell types can be regenerated and replaced, restoring healthy function. This is a worthy though lofty goal, and will likely take decades of dedicated research to attain. However, the usefulness of hESCs is not limited to cell replacement therapy.

Stem cell research allows us an unprecedented opportunity to study human development. Further, through directed differentiation of hESCs into specific cell types, we will have access to human cellular model systems that more accurately represent true human physiology in health and disease than any previous methods. These hESC derived model systems will lead to better understanding of specific cell types and their function, and will likely allow for better therapies and more streamlined drug discovery procedures, which will yield new treatment options for patients and physicians.

As mentioned briefly above, the human CNS exhibits a very poor regenerative capacity. Following injury or assault from degenerative disease, CNS deficits are largely permanent. The ability to generate neuronal subtypes of the human CNS, including the neurons of the cortex and spinal cord would be a great asset to studying human brain development as wells as methods to improve recovery following injury to these systems. Here we investigate the development of the human cortex, with special regard to the subcerebral projection neurons of the corticospinal tract, using a modified hESC line (Ruby and Zheng 2009).

1.2 The Mammalian Cortex

The mammalian neocortex, commonly referred to simply as the cortex, is an extremely complex, intricately organized six-layer structure, interwoven from hundreds of different neuronal subtypes and various glia.

It is the structure responsible for cognitive function, sensory perception and consciousness (Molyneaux, Arlotta et al. 2007). The neurons of the neocortex can be placed into two general categories, interneurons, which form close connections between different neurons, and projection neurons, which send out projections to relatively distant targets, including intracortical, subcortical and subcerebral regions. Projection neurons are glutamatergic and typically have a pyramidal morphology(Molyneaux, Arlotta et al. 2007, Molyneaux, Arlotta et al. 2007).

The six layers of the neocortex arise from radial glia (RG) and basal progenitors located in the ventricular and subventricular zones. The cortex forms in an inside out fashion, such that the deepest layers are born first, followed by the upper layers of the cortex (Figure 1.2). In the early stages of neocortical development, cells migrating from the ventricular zone give rise to the preplate, which eventually splits to form both the superficial marginal zone, and the deep subplate. The cortical plate, which will eventually give rise to the multiple layers of the neocortex, forms between these two layers. As new neurons are born, they migrate past earlier born neurons in the cortical plate, to their appropriate locations in the developing neocortex. Layer six, the deepest layer of the cortex forms first, followed sequentially by layer five, then four and then the more superficial layers (Molyneaux, Arlotta et al. 2007). (Figure 1.2) While there has been some literature suggesting the existence of distinct, layer restricted

progenitors (ie. neuronal progenitors that will give rise only to upper layer neurons), the predominant hypothesis of the scientific community, supported by decades of work, is that projection neurons of the neocortex arise from a common RG, whose potential becomes increasingly restricted as development progresses (Guo, Eckler et al. 2013). The fate of early born progenitors is potentially less restricted than that of late born cortical projection neurons. Indeed, it has been demonstrated that deep layer progenitors taken from one animal and transplanted into the superficial layers of another animal that is further along in cortical development, will give rise to the appropriate upper layer neurons in the recipient cortex, despite their deep-layer origins. It is, however, more difficult to persuade a late born progenitor to adopt a deep layer identity. This evidence further supports the concept of a common cortical NSC pool which gives rise to all layers of the cortex, with the eventual fate of the offspring of this progenitor pool being sequentially restricted and ultimately determined by birthdate and surrounding environmental cues.

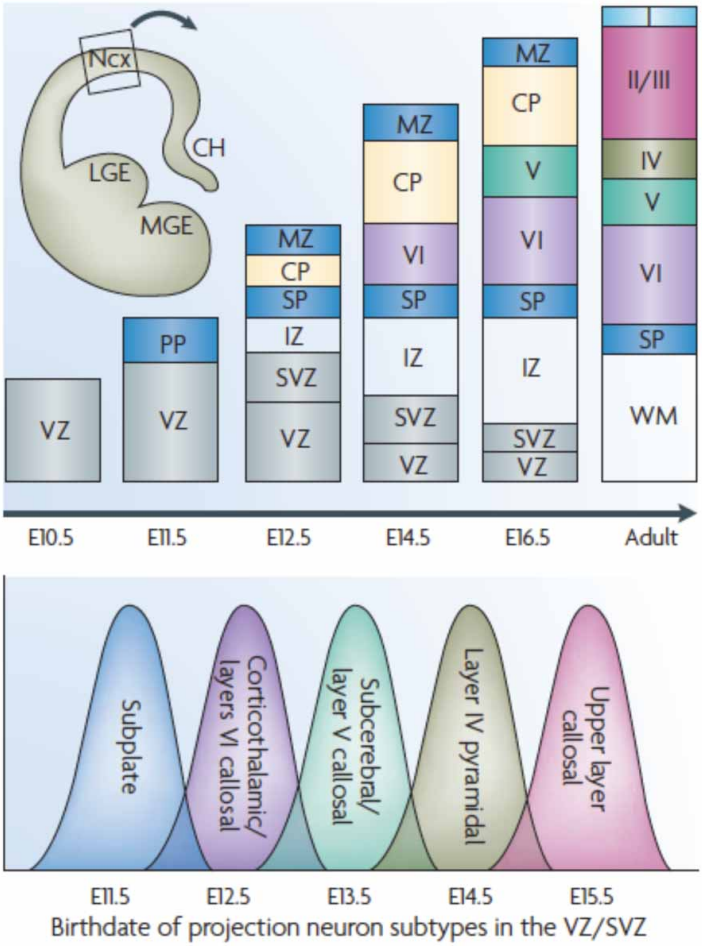

Figure 1.2. Schematic depicting how progenitors residing in the VZ and SVZ in mice produce projection neurons in an 'inside-out' fashion.

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 neurogenesis17. 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, Arlotta et al. 2007)

1.3 Corticospinal Motor Neurons

Corticospinal motor neurons (CSMNs) are subcerebral projection neurons that reside in layer five of the cortex. The axonal projections of these neurons extend down from the cortex, through the internal capsule, cerebral peduncle, pons, and medulla and out to the spinal cord (Figure 1.3). The nerve tract formed by these axons is known as the corticospinal tract (CST). The majority of these axons (~90%) cross the midline at the pyramidal decussation (those that do not cross the midline form the ventral CST). A subset of CSMNs synapse directly onto the alpha motor neurons which control the function of the forearms and hands (Dale Purves GJA, C Katz et al. 2001). Thus, CSMNs are particularly noteworthy, as they are vital for skilled, voluntary movement and fine motor control, particularly in the fingers and hands.

CSMNs are damaged or undergo degeneration in spinal cord injury and neurodegenerative diseases, respectively, and are therefore a clinically relevant neuronal subtype.

Because of the poor regenerative capabilities of the central nervous system, damage to the CST is largely permanent, leading to lifelong difficulties in movement, paralysis and lowered quality of life. The permanent and devastating nature of these injuries emphasizes the need for a better understanding of CSMNs, in order to eventually develop better

treatments for those suffering from damage or illness affecting the corticospinal tract.

While major strides have been made in understanding the process of cortical development (and therefore CSMN development) in mouse models, there are still many questions that need to be addressed. Likewise, though there are many conserved developmental pathways between mice and men, the human cortex differs significantly from the murine cortex. Additionally, the connectivity patterns of the CST also differ between mice and humans Thus, as one would expect, the human cortical development also differs from murine cortical development.

Additionally, due to obvious ethical reasons, many experiments performed in mouse models cannot be performed in living humans. Thus, in some cases, there are more points of uncertainty in our understanding of the human brain than in our understanding of the mouse brain. This is largely true of our understanding of human CSMN development. Because it is ethically impossible to directly study human cortical development in living subjects, a different model for human cortical and CST development is required in order to further progress our understanding. Here human embryonic stem cells (hESCs) offer us hope where other options fall short. As described above, hESCs are non-transformed human cells with the innate potential to self-renew, and to develop into any cell type of the human body, including all the cell types of the cortex, and therefore the

CST. Through careful manipulation of the biological factors influencing hESCs, it is possible to direct their differentiation from pluripotent stem cells to deep layer cortical neural stem cells, and from neural stem cells to specific neuronal subtypes, such as CSMNs.

While there are many signaling molecules, growth factors and transcription factors involved in cortical development one transcription factor that appears to be uniquely important for correct formation of the deep cortical layers, and more specifically for CSMNs is forebrain embryonic zinc finger-like protein 2 (Fezf2).

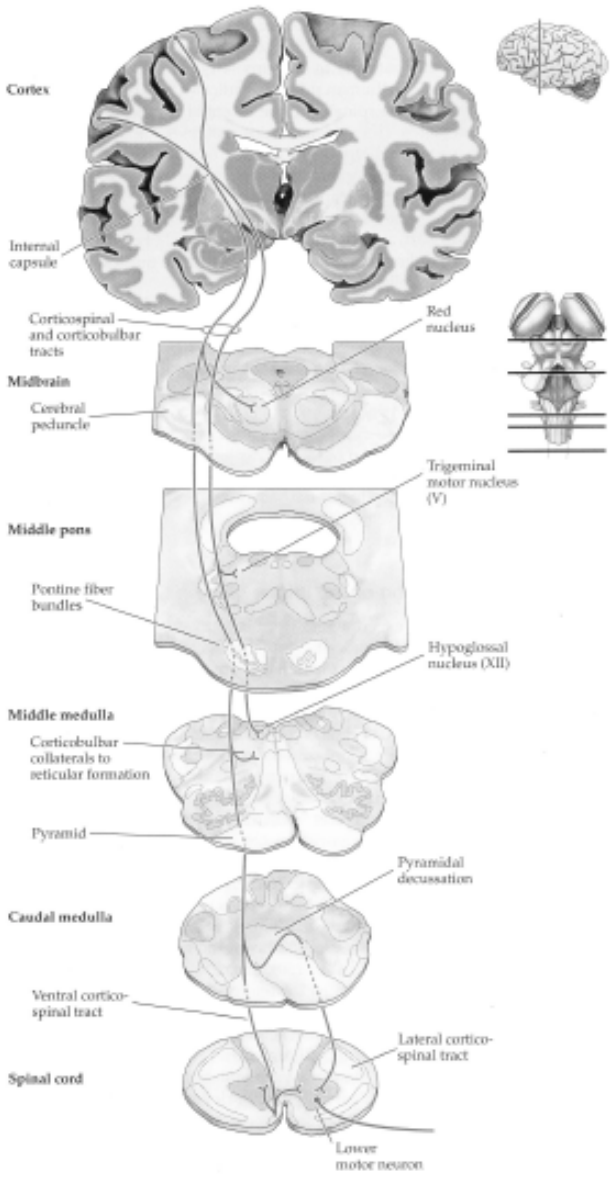

Figure 1.3. Illustration of the human CST.

Cell bodies residing inlayer V of the cortex extend axons subcerebrally through the cerebral peduncle, pons, medulla, pyramidal decussation and out into the spinal cord (Dale Purves GJA, C Katz et al. 2001).

1.4 The Transcription Factor Fezf2

The transcription factor Fezf2 has been identified as a key regulator in governing the development of CSMNs in murine models. Fezf2 is a transcription factor that is highly conserved in organisms ranging from Drosophila to humans (Molyneaux, Arlotta et al. 2005). In the mouse, Fezf2 is expressed at a constant level in subcerebral projection neurons, including CSMNs from E18.5 to P14 (Molyneaux, Arlotta et al. 2005). In the murine cortex, Fezf2 is specific to deep layer subcerebral projection neurons, such as CSMNs and corticotectal projection neurons.

1.4.1 Fezf2 Positive Progenitors

While relatively little attention has been addressed to the earliest manifestations of Fezf2 expression, Fezf2 is expression has been observed in the telenciphalic wall of the developing mouse brain at time points as early as E8.5 (Matsuo-Takasaki, Lim et al. 2000, Hirata, Suda et al. 2004). Fezf2 is expressed in the ventricular zone between E12.5 and E14.5, when subcerebral projection neurons are being born, and in the developing cortical plate where newly birthed subcerebral projection neurons would reside, strongly suggesting that Fezf2 positive progenitors are indeed giving rise to subcerebral projection neurons (Molyneaux, Arlotta et al. 2005) (Figure1.4.1).

Interestingly, however, the presence of Fezf2 positive RGs in the ventricular zone both significantly predates and persists long after the

birthdate of deep layer projection neurons(Guo, Eckler et al. 2013), indicating that while subcerebral projection neurons, including CSMNs, may indeed arise from Fezf2 positive progenitors, Fezf2 positive progenitors are likely not restricted to this fate. In fact, through elegant fate mapping experiments, it has been demonstrated, that RGs and basal progenitors that at some time in their lifespan expressed Fezf2, contribute to every layer of the murine cortex, and give rise to both projection neurons and glia (Guo, Eckler et al. 2013). It appears that a subset of RGs pass through a Fezf2 positive phase, before going on to generate what will become deep layer projection neurons, upper layer neurons, or glia, and the full significance of Fezf2 expression in this population remains unclear. Thus, much remains to be learned about the role of Fezf2 in the development of both the cortex in general, and subcerebral projection neurons specifically.

What's more, it remains to be seen if results observed in rodent models hold true in human cortical development. In order to address these questions, we have employed a previously generated Fezf2:EYFP hESC knock-in line (Ruby and Zheng 2009) to investigate the role of Fezf2 in human cortical development, with special regard to the generation of deep layer projection neurons, especially CSMNs. While hESCs have previously been used to model human cortical development in vivo, this is

the first time it, has been done in the context of following a molecularly identified neuronal subtype.

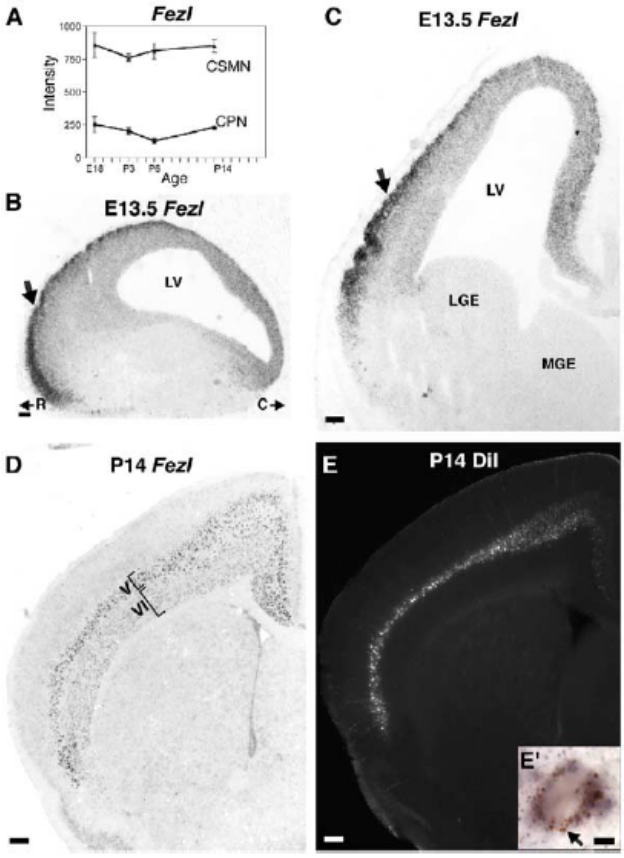

Figure 1.4.1. Fezl Is Expressed in Subcerebral Projection Neurons.

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, Arlotta et al. 2005)

1.4.2 Effects of Fezf2 Knockout

Perhaps the most powerful data supporting the key role of Fezf2 in CSMN development was the discovery that in genetically modified Fezf2 null (-/-) animals, the corticospinal tract completely fails to form. In a wild type mouse, layer V of the cortex is populated with large pyramidal, subcerebral projection neurons, including CSMNs and corticotectal projection neurons. In Fezf2 null animals, these neurons are absent, an observation confirmed by absence of Ctip2 positive neurons and both retrograde and anterograde labeling (Figure 1.4.2). The absence of these subcerebral projection neurons in not caused by cell death or abnormal migration, but rather appears to be due to the absence of the necessary molecular signal Fezf2. Additionally, the overall thickness of cortical layer V in Fezf2 null animals is reduced by 44%, while cortical layer VI is abnormally expanded, and layer VI neurons fail too mature properly (Molyneaux, Arlotta et al. 2005). This indicates that Fezf2 is likely involved in determining the laminar identity of layer V and VI neurons, and in establishing the boundary between these layers.

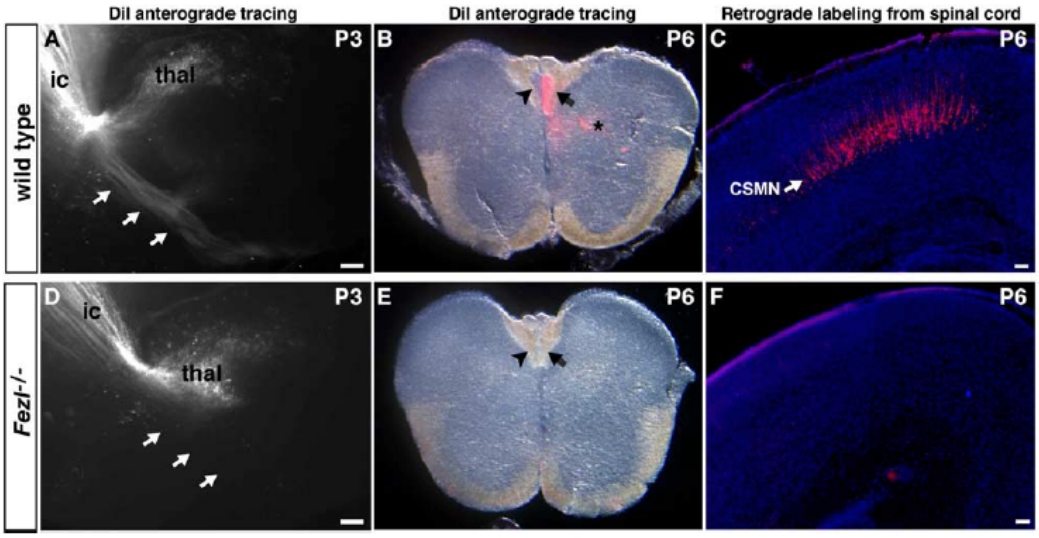

Figure 1.4.1. Figure 7. Absence of the Corticospinal Tract and Other Subcerebral Projections in Fezl−/− Mice.

(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).

1.4.3 Fezf2 Overexpression and Reprogramming

As described above, Fezf2 loss-of-function experiments have demonstrated the importance of Fezf2 in the development of subcerebral projection neurons like CSMNs. Fezf2 gain-of-function experiments have likewise highlighted the important role of this transcription factor. A Fezf2 overexpression construct was electroporated into the ventricle of developing mouse brain at E13.5. Cells that received this construct and

overexpressed Fezf2 were arrested in their migration and adopted a subcerebral like identity, demonstrated by Ctip2 expression, weak or absent Tbr1 expression, and the extension of axons through the internal capsule toward subcerebral targets (Molyneaux, Arlotta et al. 2005).

Additional studies have demonstrated that Fezf2 overexpression can reprogram cell identity toward a more subcerebral projection neuron fate. In 2010, Caroline Rouaux and Paola Arlotta demonstrated that by ectopically expressing Fezf2 in progenitors normally fated to become striatal medium spiny neurons, the fate of these progenitors could be redirected. Rather than becoming GABAergic medium spiny neurons, these progenitors gave rise to glutamatergic corticofugal neurons (Rouaux and Arlotta 2010). And in 2013, the same team published a paper where in they were able to reprogram layer II/III callosal projection neurons into layer V/VI corticofugal projection neurons through ectopic Fezf2 expression in embryonic and early postnatal mice (Rouaux and Arlotta 2013). Together, these experiments lend additional emphasis to the importance of the transcription factor Fezf2 in establishing subcerebral projection neuron identity in the murine cortex.

1.4.4 Fezf2 Interactions with Ctip2

The transcription factor Ctip2 is expressed in various regions in the brain, however, it has been identified as an important down stream effector of Fezf2 (Chen, Wang et al. 2008, Srinivasan, Leone et al. 2012).

As one would expect, Ctip2 is strongly expressed in developing layer V, corticofugal neurons, including corticotectal and corticospinal projection neurons. Ctip2 null animals exhibit a phenotype similar to Fezf2 null animals, in that, to a large extent, the CST fails to form. As discussed above, in Fezf2 null mice, progenitors in the subventricular zone migrate to the appropriate location in the cortex for the formation of corticofugal projection neurons, however, they then fail to take on the proper identity for subcerebral projection neurons. Not surprisingly, these would-be corticofugal Fezf2-negative neurons also fail to express Ctip2. Remarkably, ectopic expression of either Fezf2 or Ctip2 through electroporation at the appropriate time point in development is able to rescue the phenotype, permitting the generation of corticfugal neurons in layer V (Figure 1.4.4). Surprisingly, though ectopic Fezf2 expression is sufficient to rescue the corticospinal projection neuron phenotype, it does so without activating Ctip2 expression. Thus, though Ctip2 is an important down stream effector of Fezf2, and is even capable of replacing Fezf2, Fezf2 is capable of correctly directing corticospinal development through other down stream genes other than Ctip2 (Chen, Wang et al. 2008).

Interestingly, during development, layer V corticofugal projection neurons are born and develop at the same time and in the same region as layer V callosal projection neurons. Cortcofugal projection neurons are positive for both Fezf2 and Ctip2, where as layer V callosal projection

neurons are Fezf2 and Ctip2 negative, and are instead positive for Satb2. Through a system of mutual repression/ regulation these transcription factors govern the fate switch by which layer V projection neurons adopt one of these two identities.

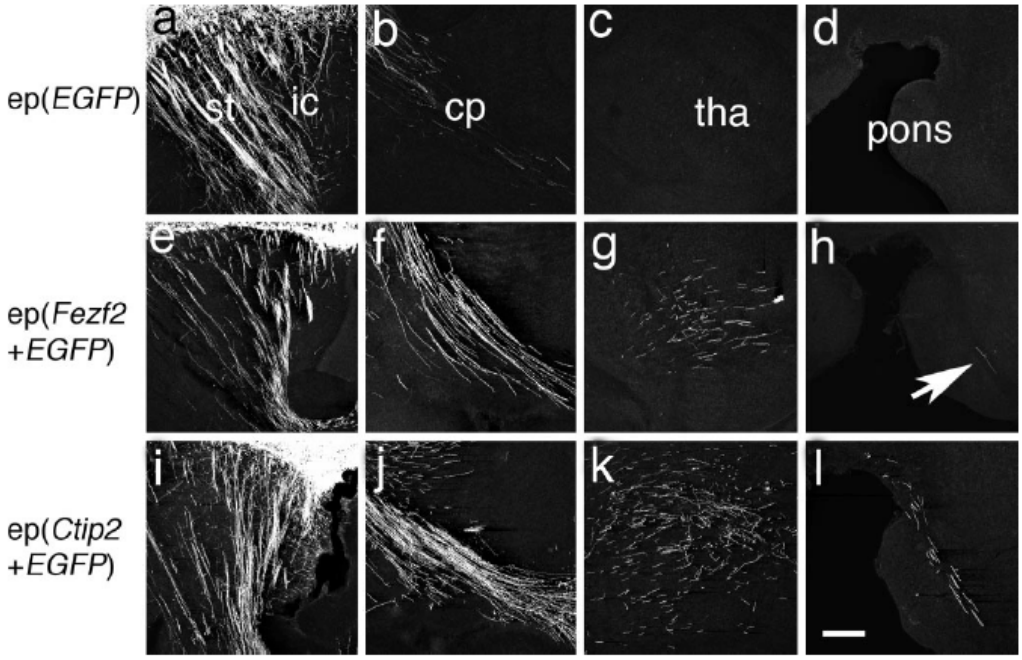

Figure1.4.4. Ectopic expression of Fezf2 or Ctip2 in wild-type mice is sufficient to alter the axon trajectories of upper-layer neurons, which normally form corticocortical connections, at P5.

(a–d) Axonal projections of layer 2/3 neurons electroporated withpCA-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). (i–l) Upper-layer neurons coelectroporated with pCA-Ctip2 and pCA-EGFP showed a pattern similar to that after electroporation with Fezf2, but an even greater number of labeled axons were present in the thalamus (k) and CST (l). ep, electroporation (expression construct); cp, cerebral peduncle; ic, internal capsule; st, striatum; tha, thalamus. (Scale bar, 200 μm.) (Chen, Wang et al. 2008)

1.4.5 Fezf2 Interactions with Satb2

Satb2 is another important transcription factor in cortical development. As noted briefly in the previous section, Satb2 governs the fate of callosal projection neurons, while Fezf2 and Ctip2 direct the fate of

subcerebral projection neurons like CSMNs. In addition to being expressed in callosally projecting neurons of layer V of the cortex, Satb2 is expressed in the superficial layers of the cortex.

In the previous section it was noted that Ctip2 is down regulated in the absence of Fezf2. Interestingly, in the absence of Fezf2, Satb2 expression is upregulated in the developing cortex, and would-be subcerebral projection neurons change their fate and instead become callosal projection neurons (Chen, Wang et al. 2008, Srinivasan, Leone et al. 2012). This indicates that Fezf2 represses Satb2 expression in subcerebral projection neurons, promoting the corticofugal fate.

Additionally, Satb2 binds to and represses the Ctip2 locus, and thus, not surprisingly, in the absence of Satb2, Ctip2 (but not Fezf2) is upregulated, causing what would otherwise become callosal projection neurons to adopt the corticofugal fate (Alcamo, Chirivella et al. 2008, Britanova, de Juan Romero et al. 2008, Srinivasan, Leone et al. 2012). Futher, in the absence of both Fezf2 and Satb2, Ctip2 expression is rescued in layer V of the cortex (Figure 1.4.5). Thus as Fezf2 serves to repress Satb2 in layer V of the cortex, promoting the corticofugal fate, and Satb2 acts to repress Ctip2 (down stream of Fezf2) to promote the callosal fate, the absence of both of these factors allows for normal expression of Ctip2. This experiment indicates that Fezf2 expression is not required for Ctip2 expression, and in fact, evidence supports the hypothesis that Fezf2

regulates Ctip2 in an indirect fashion by repressing Satb2 (Srinivasan, Leone et al. 2012).

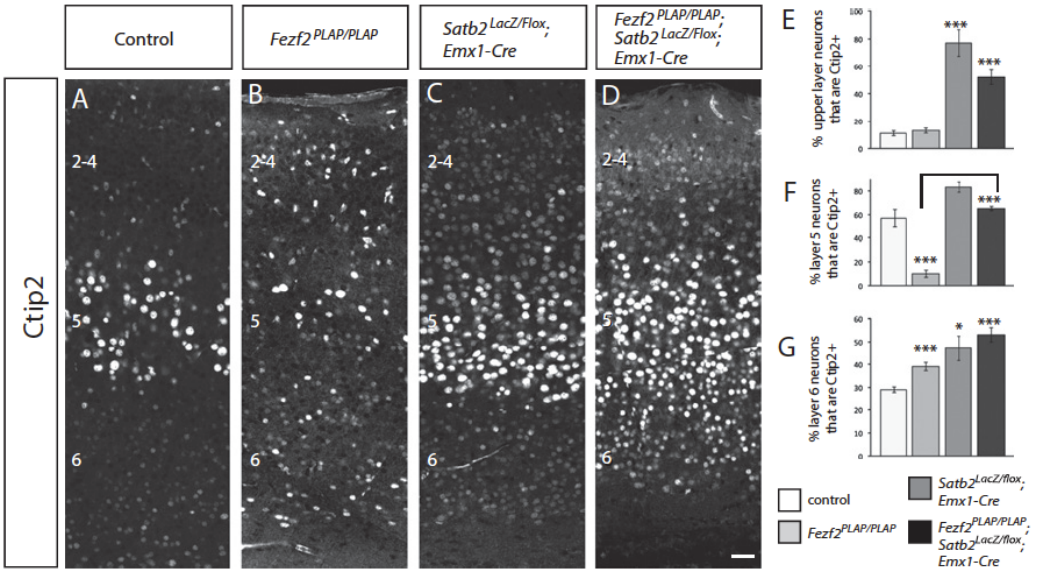

Figure 1.4.5. Ctip2 expression is restored in layer 5 neurons of Fezf2PLAP/PLAP;Satb2LacZ/flox;Emx1-Cre double mutants.

(A) In WT mice, the expression of Ctip2 protein is highest in layer 5 neurons, moderate in layer 6, and largely absent from the upper layers (2–4). (B) Fezf2 mutants show a severe reduction of Ctip2 expression in layer 5, whereas expression in layer 6 is increased. (C) Satb2 mutant mice show increased Ctip2 expression in layer 5, 6, and the upper layers. (D) Ctip2 expression is rescued in layer 5 of Fezf2PLAP/PLAP;Satb2LacZ/flox;Emx1-Cre double mutants. (A–D) Each panel is a composite of tiled confocal images to form the full figure. (Scale bar, 50 μm.) (E–G) Graphs depict the percentages of Ctip2+ neurons in the upper layers (E), layer 5 (F), and layer 6 (G) in various genotypes. Error bars are average ± SEM (n = 3 different animals). Confidence levels were calculated using a t test (*P < 0.5; **P < 0.05; ***P < 0.005, relative to controls) (Srinivasan, Leone et al. 2012).

1.4.6 Fezf2 Interactions with Tbr1

Tbr1 is another important transcription factor that plays a role in the determination of deep layer projection neuron identity as well as laminar fate during cortical development (Bedogni, Hodge et al. 2010, Han, Kwan

et al. 2011, McKenna, Betancourt et al. 2011, Srinivasan, Leone et al. 2012). In wild type animals, layer VI of the cortex is enriched for Tbr1 expression, and likewise, so are the corticothalamic projection neurons that reside in this layer (Molyneaux, Arlotta et al. 2005, McKenna, Betancourt et al. 2011). In this same region, Ctip2 is expressed at a much lower level. By contrast, in layer V of the cortex, where Fezf2 and Ctip2 are strongly expressed, Tbr1 expression is much lower. The interplay between these transcription factors is crucial for proper fate specification within these layers.

Fezf2 knockout mutants demonstrate an increase in Tbr1 expression in layer V, an expansion of layer VI and a decrease in the overall thickness of layer V. This is accompanied by an increase in thalamic innervation from cortical projection neurons. By contrast, in Tbr1 knockout mutants, there is an increase in Ctip2 expression in layer VI of the cortex, accompanied by a loss of corticothalamic neurons (Srinivasan, Leone et al. 2012). Interestingly, Ctip2 knockout mutants also demonstrate an increase of Tbr1 expression in layer V neurons, as well as an increase in cortical innervation of the thalamus. This indicates that Ctip2 (down stream of Fezf2) likely negatively regulates Tbr1. Conversely, Tbr1 has been shown to directly repress expression of Fezf2 (Han, Kwan et al. 2011).

Not surprisingly, given the antagonistic relationship between Satb2 and Ctip2, and between Ctip2 and Tbr1, there exists a positive relationship between Satb2 and Tbr1. It has been demonstrated that Satb2 binds directly to the Tbr1 genetic region, and promotes Tbr1 expression. This relationship between Satb2 and Tbr1 is further supported by the fact that over expression of Tbr1 in Satb2 knockouts can rescue the loss of callosal projection neurons usually observed with this genotype.

1.4.7 Transcription Factor Summary

In summary, Fezf2 is an important marker and regulator of the subcerebral projection neuron fate, and is crucial for corticospinal tract development. Fezf2 lies upstream of Ctip2, which is also strongly expressed in subcerebral projection neurons, including corticospinal motor neurons. Ectopic Ctip2 expression is sufficient to rescue the Fezf2 knockout phenotype. Ectopic Fezf2 expression can also rescue this phenotype, without inducing Ctip2 expression. Ctip2 is negatively regulated by Satb2 expression, which in layer V promotes the callosal projection neuron fate. Satb2 is negatively regulated by Fezf2, though perhaps not directly. Satb2 also binds to the Tbr1 genetic locus, increasing its expression. Tbr1 expression encourages the corticothalamic projection neuron fate usually observed in layer VI of the cortex. Ctip2 negatively regulates Tbr1 expression. Tbr1, in turn, negatively regulates Fezf2 expression (Bedogni, Hodge et al. 2010, Han, Kwan et al. 2011,

McKenna, Betancourt et al. 2011, Srinivasan, Leone et al. 2012) **(Figure 1.4.7)**.

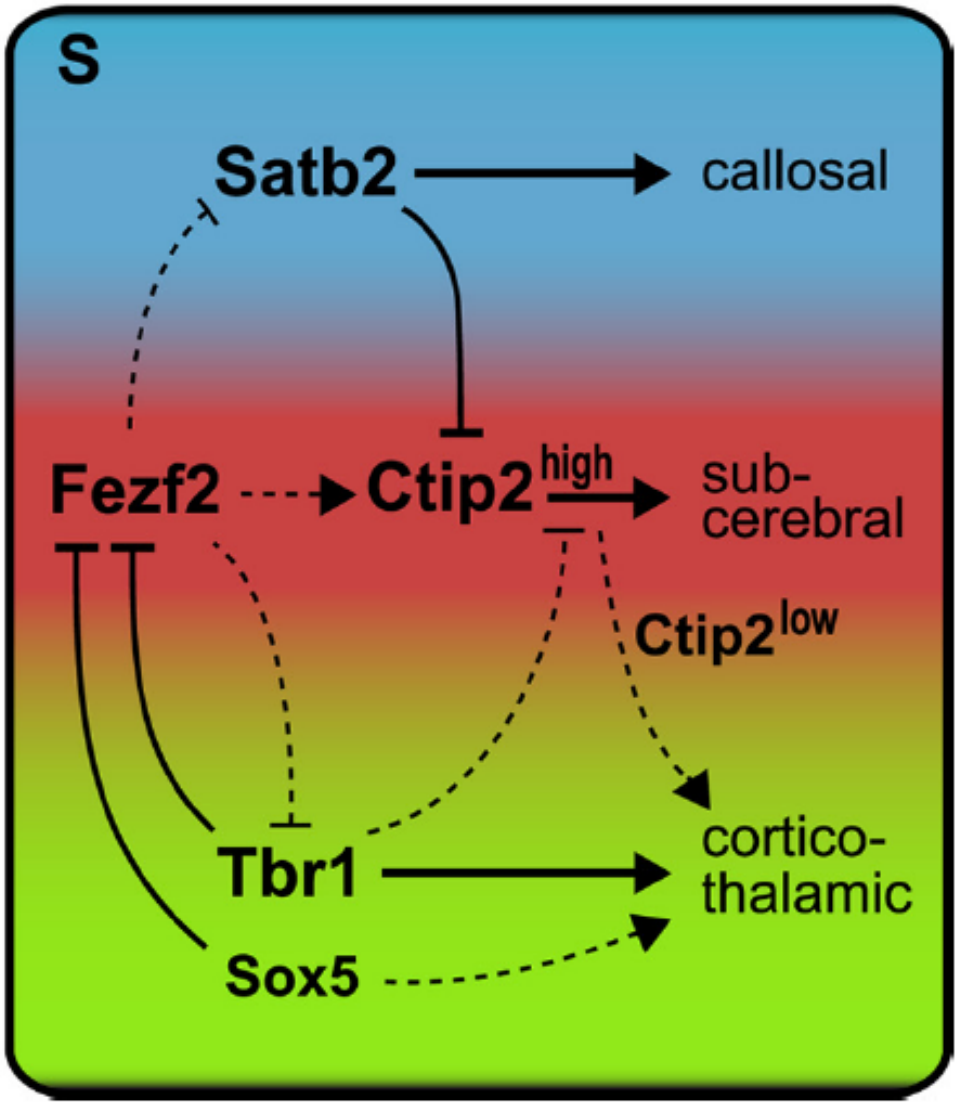

Figure 1.4.7. Proposed mechanisms for establishing cortical projection neuron identities.

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, Betancourt et al. 2011).

1.5 Cortical Development from Embryonic Stem Cells

The directed differentiation of hESCs into a specific desired cellular subtype is at best a challenging undertaking (Grivennikov 2008). The development of human CSMNs has been further complicated by the absence of a reliable Fezf2 antibody for IHC. Nevertheless, progress has been made in developing cortical neurons, and even some corticofugal neurons, initially in mouse models and later in models using human cells. In 2008, the Vanderhaeghen lab was able to recapitulate the major steps of corticogenesis in vitro using murine embryonic stem cells (Gaspard, Bouschet et al. 2008).

In 2009, Lorenz Studer published a protocol for highly efficient neural conversion of embryonic stem cells into neural precursors through Dual SMAD inhibition (Chambers, Fasano et al. 2009). This technique has had a powerful impact on the field of stem cell related neuroscience, and modifications of this technique have been widely and successfully employed. The differentiation technique described in this work is, in fact, a modification of the Dtuder dual SMAD inhibition protocol, adjusted to generate cortical projection neurons.

In 2010, James Weimann's lab was able to generate pyramidal projection neurons from murine embryonic stem cells, and demonstrated that they were able to integrate impressively well into the cortices of host

animals and send out appropriate projections (Ideguchi, Palmer et al. 2010) (Figure 1.5).

In 2013, the Vanderhaeghen group succeeded in producing a variety of cortical neurons, including pyramidal neuron subtypes, from human embryonic stem cells, and transplanted them into animal hosts. Some of these in vitro patterned projection neurons were able to project to visual, thalamic and even subcortical targets (Espuny-Camacho, Michelsen et al. 2013). Despite these successes much remains to be learned about human cortical development, and likewise the generation of corticospinal motor neurons from hESCs.

While hESCs have been used to generate pyramidal neurons, here we seek to elucidate the development of a specific molecularly defined neuronal subtype, the subcerebral projection neuron, and examine its generation from stem cell to progenitor, to adult neuron.

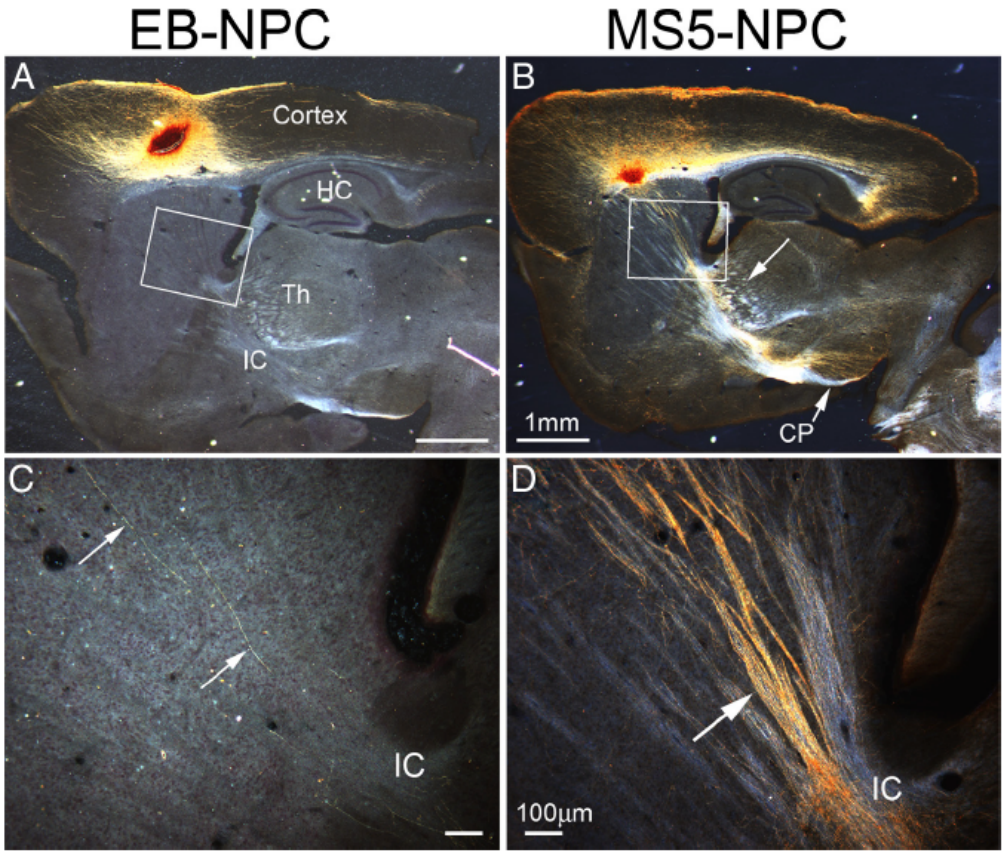

Figure 1.5. MS5 preconditioned ES-derived neurons project axons to subcortical areas. GFP$^+$ES cell-derived NPCs were transplanted into cortex of P2 mice.

At P14, animals were killed, and brains were processed; GFP$^+$ axons and somata are golden brown because of the DAB reaction product. A, Transplantation of EB-NPCs into motor cortex.Manymature neurons are formed but only a few extend axons into the internal capsule (IC) (N=22). HC, Hippocampus. B, Transplantation of MS5-NPCs into motor cortex. At P14, many mature neurons are formed and extend axons into the internal capsule (IC), thalamus (Th) (arrow), and cerebral peduncle(CP)(arrow)(N=48). C, High-power image of boxed area in A, showing only two axons projecting into IC (arrows) in this section. A total of 11.0 +/- 11.6 axons pe ra transplanted hemisphere was observed (7– 8 sections). D, High-power image of the boxed area in B; hundreds of axons are apparent in fasciculated bundles entering the IC (arrow). Scale bars: A, B, 1 mm; C,D, 100µm (Ideguchi, Palmer et al. 2010).

1.6 Bacterial Artificial Chromosome (BAC) based transgenesis in hESCs

Additionally, presented here is a method for efficiently generating new transgenic hESC lines through BAC based transgenesis. Though

great strides have been made in genetically manipulating murine embryonic stem cells, the manipulation of hESCs has lagged behind (Giudice and Trounson 2008). A common means for labeling a desired cell type is the use of viral vectors to insert a fluorescent marker such as GFP into a target cell population. While effective in some studies, this method has its limitations. Namely, the size of the vector to be inserted is subject to the limitations of the viral vector selected, the efficiency of virally introduced genetic material is often negatively affected through epigenetic silencing (Xia, Zhang et al. 2007), and virally introduced fluorescent markers sometimes prove temporary, disappearing with repeated passages or multiple cell divisions (Macarthur, Xue et al. 2012).

Here we explore an alternative method for generating labeled hESC lines through the electroporation of modified bacterial artificial chromosomes (BACs). BACs are highly stable and relatively easy to work with. They have been employed successfully in the generation of brightly labeled transgenic mouse lines and were crucial to the success of the human genome project (Osoegawa, Mammoser et al. 2001, Wendl, Marra et al. 2001, Zeng, Kouprina et al. 2001, Zhang and Wu 2001, Waterston, Lander et al. 2002) and the Allen Brain Atlas (Boguski and Jones 2004, Boizuski 2005). BACs have also been used to modify hESCs (Song, Chung et al. 2010). In this work the generation of two modified BACs is described, as well as the use of these BACs to generate stably transfected,

constitutively active fluorescently labeled transgenic hESC lines. In the first, tdTomato was placed under the control of the Rosa26 locus, in the second, fluorescence was increased through the use of a CAG promoter and WPRE stabilizer. A variation of this construct was used successfully to generate a brightly labeled mouse line (Madisen, Zwingman et al. 2010). This BAC could be incorporated into virtually any embryonic stem cell line with relative ease, and would aid considerably in tracking transfected cells an their differentiated progeny either in vitro or in vivo.

1.7 Conclusion of the Introduction

HESCs offer unrivaled potential for the study of human development as well as the promise of aiding in the development of new biomedical discoveries and therapies, particularly for cell types that exhibit poor endogenous regenerative capabilities. The CST, a part of the central nervous system necessary for voluntary, fine motor control of the fingers hands and toes, is composed of one such cell type, CSMNs. CSMNs undergo injury or degeneration in spinal cord injuries and neurodegenerative diseases such as ALS. The transcription factor Fezf2 is an important marker and molecular determinant of the CSMN fate and is also expressed in cortical progenitors. Fezf2 is part of a complex network of transcription factors that regulate deep-layer laminar development as well as fate choice of deep layer projection neurons. Other key transcription factors include, Ctip2, which lies down steam of Fezf2, Satb2,

which promotes the callosal projection neuron fate, and Tbr1, which is strongly expressed in layer VI corticothalamic projection neurons. Our lab previously developed a fluorescent Fezf2:YFP knock-in reporter line. Here we implement this reporter line in order to study the role of Fezf2 in human cortical development and to help direct the differentiation of hESCs into putative CSMNs. Additionally we present a BAC-based means of hESC labeling with the fluorescent protein tdTomato with the goal of facilitating cell tracking.

1.8 References

Alcamo, E. A., L. Chirivella, M. Dautzenberg, G. Dobreva, I. Farinas, R. Grosschedl and S. K. McConnell (2008). "Satb2 regulates callosal projection neuron identity in the developing cerebral cortex." Neuron 57(3): 364-377.

Arlotta, P., B. J. Molyneaux, J. Chen, J. Inoue, R. Kominami and J. D. Macklis (2005). "Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo." Neuron 45(2): 207-221.

Bedogni, F., R. D. Hodge, G. E. Elsen, B. R. Nelson, R. A. Daza, R. P. Beyer, T. K. Bammler, J. L. Rubenstein and R. F. Hevner (2010). "Tbr1 regulates regional and laminar identity of postmitotic neurons in developing neocortex." Proc Natl Acad Sci U S A 107(29): 13129-13134.

Boguski, M. S. and A. R. Jones (2004). "Neurogenomics: at the intersection of neurobiology and genome sciences." Nature Neuroscience 7(5): 429-433.

Boizuski, M. S. (2005). "Neurogenomics: at the intersection of neurobiology and genome sciences." Faseb Journal 19(5): A1711-A1711.

Britanova, O., C. de Juan Romero, A. Cheung, K. Y. Kwan, M. Schwark, A. Gyorgy, T. Vogel, S. Akopov, M. Mitkovski, D. Agoston, N. Sestan, Z. Molnar and V. Tarabykin (2008). "Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex." Neuron 57(3): 378-392.

Chambers, S. M., C. A. Fasano, E. P. Papapetrou, M. Tomishima, M. Sadelain and L. Studer (2009). "Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling." Nat Biotechnol 27(3): 275-280.

Chen, B., L. R. Schaevitz and S. K. McConnell (2005). "Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex." Proc Natl Acad Sci U S A 102(47): 17184-17189.

Chen, B., S. S. Wang, A. M. Hattox, H. Rayburn, S. B. Nelson and S. K. McConnell (2008). "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 105(32): 11382-11387.

Cowan, C. A., I. Klimanskaya, J. McMahon, J. Atienza, J. Witmyer, J. P. Zucker, S. Wang, C. C. Morton, A. P. McMahon, D. Powers and D. A. Melton (2004). "Derivation of embryonic stem-cell lines from human blastocysts." N Engl J Med 350(13): 1353-1356.

Dale Purves GJA, D. F., Lawrence, A.-S. L. C Katz, James O McNamara, and S Mark and J. L. Williams (2001). Neuroscience, Sinauer Associates, Inc.

Espuny-Camacho, I., K. A. Michelsen, D. Gall, D. Linaro, A. Hasche, J. Bonnefont, C. Bali, D. Orduz, A. Bilheu, A. Herpoel, N. Lambert, N. Gaspard, S. Peron, S. N. Schiffmann, M. Giugliano, A. Gaillard and P. Vanderhaeghen (2013). "Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo." Neuron 77(3): 440-456.

Gaspard, N., T. Bouschet, R. Hourez, J. Dimidschstein, G. Naeije, J. van den Ameele, I. Espuny-Camacho, A. Herpoel, L. Passante, S. N. Schiffmann, A. Gaillard and P. Vanderhaeghen (2008). "An intrinsic mechanism of corticogenesis from embryonic stem cells." Nature 455(7211): 351-357.

Giudice, A. and A. Trounson (2008). "Genetic modification of human embryonic stem cells for derivation of target cells." Cell Stem Cell 2(5): 422-433.

Grivennikov, I. A. (2008). "Embryonic stem cells and the problem of directed differentiation." Biochemistry (Mosc) 73(13): 1438-1452.

Guo, C., M. J. Eckler, W. L. McKenna, G. L. McKinsey, J. L. Rubenstein and B. Chen (2013). "Fezf2 expression identifies a multipotent progenitor for neocortical projection neurons, astrocytes, and oligodendrocytes." Neuron 80(5): 1167-1174.

Han, W., K. Y. Kwan, S. Shim, M. M. Lam, Y. Shin, X. Xu, Y. Zhu, M. Li and N. Sestan (2011). "TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract." Proc Natl Acad Sci U S A 108(7): 3041-3046.

Hirata, T., Y. Suda, K. Nakao, M. Narimatsu, T. Hirano and M. Hibi (2004). "Zinc finger gene fez-like functions in the formation of subplate neurons and thalamocortical axons." Dev Dyn 230(3): 546-556.

Ideguchi, M., T. D. Palmer, L. D. Recht and J. M. Weimann (2010). "Murine embryonic stem cell-derived pyramidal neurons integrate into the cerebral cortex and appropriately project axons to subcortical targets." J Neurosci 30(3): 894-904.

Ip, B. K., N. Bayatti, N. J. Howard, S. Lindsay and G. J. Clowry (2011). "The corticofugal neuron-associated genes ROBO1, SRGAP1, and CTIP2 exhibit an anterior to posterior gradient of expression in early fetal human neocortex development." Cereb Cortex 21(6): 1395-1407.

Lee, E. C., D. Yu, J. Martinez de Velasco, L. Tessarollo, D. A. Swing, D. L. Court, N. A. Jenkins and N. G. Copeland (2001). "A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA." Genomics 73(1): 56-65.

Leone, D. P., K. Srinivasan, B. Chen, E. Alcamo and S. K. McConnell (2008). "The determination of projection neuron identity in the developing cerebral cortex." Curr Opin Neurobiol 18(1): 28-35.

Li, J., T. Jiang, B. Bejjani, E. Rajcan-Separovic and W. W. Cai (2003). "High-resolution human genome scanning using whole-genome BAC Arrays." Cold Spring Harbor Symposia on Quantitative Biology 68: 323-329.

Macarthur, C. C., H. Xue, D. Van Hoof, P. T. Lieu, M. Dudas, A. Fontes, A. Swistowski, T. Touboul, R. Seerke, L. C. Laurent, J. F. Loring, M. S. German, X. Zeng, M. S. Rao, U. Lakshmipathy, J. D. Chesnut and Y. Liu (2012). "Chromatin insulator elements block transgene silencing in engineered human embryonic stem cell lines at a defined chromosome 13 locus." Stem Cells Dev 21(2): 191-205.

Madisen, L., T. A. Zwingman, S. M. Sunkin, S. W. Oh, H. A. Zariwala, H. Gu, L. L. Ng, R. D. Palmiter, M. J. Hawrylycz, A. R. Jones, E. S. Lein and H. Zeng (2010). "A robust and high-throughput Cre reporting and characterization system for the whole mouse brain." Nat Neurosci 13(1): 133-140.

Matsuo-Takasaki, M., J. H. Lim, M. J. Beanan, S. M. Sato and T. D. Sargent (2000). "Cloning and expression of a novel zinc finger gene, Fez, transcribed in the forebrain of Xenopus and mouse embryos." Mech Dev 93(1-2): 201-204.

McKenna, W. L., J. Betancourt, K. A. Larkin, B. Abrams, C. Guo, J. L. Rubenstein and B. Chen (2011). "Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development." J Neurosci 31(2): 549-564.

Molyneaux, B. J., P. Arlotta, T. Hirata, M. Hibi and J. D. Macklis (2005). "Fezl is required for the birth and specification of corticospinal motor neurons." Neuron 47(6): 817-831.

Molyneaux, B. J., P. Arlotta and J. D. Macklis (2007). "Molecular development of corticospinal motor neuron circuitry." Novartis Found Symp 288: 3-15; discussion 15-20, 96-18.

Molyneaux, B. J., P. Arlotta, J. R. Menezes and J. D. Macklis (2007). "Neuronal subtype specification in the cerebral cortex." Nat Rev Neurosci 8(6): 427-437.

Osoegawa, K., A. G. Mammoser, C. Y. Wu, E. Frengen, C. J. Zeng, J. J. Catanese and P. J. de Jong (2001). "A bacterial artificial chromosome library for sequencing the complete human genome." Genome Research 11(3): 483-496.

Rouaux, C. and P. Arlotta (2010). "Fezf2 directs the differentiation of corticofugal neurons from striatal progenitors in vivo." Nat Neurosci 13(11): 1345-1347.

Rouaux, C. and P. Arlotta (2013). "Direct lineage reprogramming of post-mitotic callosal neurons into corticofugal neurons in vivo." Nat Cell Biol 15(2): 214-221.

Ruby, K. M. and B. Zheng (2009). "Gene targeting in a HUES line of human embryonic stem cells via electroporation." Stem Cells 27(7): 1496-1506.

Song, H., S. K. Chung and Y. Xu (2010). "Modeling disease in human ESCs using an efficient BAC-based homologous recombination system." Cell Stem Cell 6(1): 80-89.

Srinivasan, K., D. P. Leone, R. K. Bateson, G. Dobreva, Y. Kohwi, T. Kohwi-Shigematsu, R. Grosschedl and S. K. McConnell (2012). "A network of genetic repression and derepression specifies projection fates in the developing neocortex." Proc Natl Acad Sci U S A 109(47): 19071-19078.

Swaminathan, S., H. M. Ellis, L. S. Waters, D. Yu, E. C. Lee, D. L. Court and S. K. Sharan (2001). "Rapid engineering of bacterial artificial chromosomes using oligonucleotides." Genesis 29(1): 14-21.

Thomson, J. A., J. Itskovitz-Eldor, S. S. Shapiro, M. A. Waknitz, J. J. Swiergiel, V. S. Marshall and J. M. Jones (1998). "Embryonic stem cell lines derived from human blastocysts." Science 282(5391): 1145-1147.

Warming, S., N. Costantino, D. L. Court, N. A. Jenkins and N. G. Copeland (2005). "Simple and highly efficient BAC recombineering using galK selection." Nucleic Acids Res 33(4): e36.

Waterston, R. H., E. S. Lander and J. E. Sulston (2002). "On the sequencing of the human genome." Proceedings of the National Academy of Sciences of the United States of America 99(6): 3712-3716.

Wendl, M. C., M. A. Marra, L. W. Hillier, A. T. Chinwalla, R. K. Wilson and R. H. Waterston (2001). "Theories and applications for sequencing randomly selected clones." Genome Research 11(2): 274-280.

Xia, X., Y. Zhang, C. R. Zieth and S. C. Zhang (2007). "Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner." Stem Cells Dev 16(1): 167-176.

Xu, Y., Y. Shi and S. Ding (2008). "A chemical approach to stem-cell biology and regenerative medicine." Nature 453(7193): 338-344.

Zeng, C. J., N. Kouprina, B. Zhu, A. Cairo, M. Hoek, G. Cross, K. Osoegawa, V. Larionov and P. de Jong (2001). "Large-insert BAC/YAC libraries for selective re-isolation of genomic regions by homologous recombination in yeast." Genomics 77(1-2): 27-34.

Zhang, H. B. and C. C. Wu (2001). "BAC as tools for genome sequencing." Plant Physiology and Biochemistry 39(3-4): 195-209.

Zhao, S. Y. (2001). "A comprehensive BAC resource." Nucleic Acids Research 29(1): 141-143.

Chapter 2. Toward Engineering the Corticospinal Tract from Human Embryonic Stem Cells

2.1 Introduction

The mammalian cortex is comprised of several distinct neuronal subtypes organized during development into six layers (Molyneaux, Arlotta et al. 2007). These layers are born from radial glial cells (RGs) and basal progenitors in a temporally controlled, inside-out manner, such that deep layer neurons arise early in development, followed by upper-layer neurons (Chen, Schaevitz et al. 2005, Molyneaux, Arlotta et al. 2005, Guo, Eckler et al. 2013).

Corticospinal motor neurons (CSMNs) reside in layer five of the cortex and their axonal projections form the corticospinal tract (CST), the neural substrate for skilled voluntary movement and fine motor control (Dale Purves GJA, C Katz et al. 2001, Molyneaux, Arlotta et al. 2007, Leone, Srinivasan et al. 2008). CSMNs are damaged or undergo degeneration in spinal cord injury and neurodegenerative diseases, respectively, and are therefore a clinically relevant neuronal subtype. Because of the poor regenerative capabilities of the central nervous system, damage to the CST is permanent, leading to lifelong deficits in movement, paralysis and lowered quality of living. The chronic and devastating nature of these injuries emphasizes the need for a better understanding of CSMNs.

The transcription factor Fezf2 has been identified as a key regulator in governing the specification of CSMNs in vertebrate models (Arlotta,

Molyneaux et al. 2005, Chen, Schaevitz et al. 2005, Molyneaux, Arlotta et al. 2005). In the mature murine cortex, Fezf2 is restricted to deep layer subcerebral projection neurons, such as CSMNs and corticotectal projection neurons. In addition, Fezf2 is necessary for formation of the CST, as the corticospinal tract completely fails to form in Fezf2 null (-/-) animals(Molyneaux, Arlotta et al. 2005). Furthermore, Fezf2 can induce CSMN-like identity by overexpression in alternate cortical layers (Rouaux and Arlotta 2010, Rouaux and Arlotta 2013). Overall, fezf2 acts as a powerful transcription factor in regulating the CSMN lineage.

Fezf2 expression has also been identified in early neurogenesis of the developing mouse telencephalon(Molyneaux, Arlotta et al. 2005, McKenna, Betancourt et al. 2011, Guo, Eckler et al. 2013). Interestingly, the presence of these Fezf2 positive NPCs both significantly predates and persists long after the birthdate of CSMN, calling into question any correlation between this early Fezf2 expression and the Fezf2 expression observed in more mature deep layer subcerebral projection neurons. In fact, through elegant fate mapping experiments, it has been demonstrated that Fezf2 positive RGCs and basal progenitors contribute to every layer of the murine cortex, and give rise to both neurons and glia (Guo, Eckler et al. 2013). Despite these advances, the precise role of Fezf2 in the development of the human cortex, with its expanded progenitor domains and complex structures, remains unclear.

In order to address these questions, we have employed a previously generated Fezf2:EYFP hESC knock-in line (Ruby and Zheng 2009) to investigate the role of Fezf2 in human cortical development, in particular, the generation of deep layer projection neurons such as CSMNs. While hESCs have previously been used to model general human cortical development in vitro (Gaspard, Bouschet et al. 2008, Espuny-Camacho, Michelsen et al. 2013), it is of clinical importance to isolate and characterize specific neuronal subtypes. Our results are the first developmental mapping of a molecularly identified neuronal subtype of the human cortex. We describe a temporally controlled expression profile of Fezf2 from radial glia to subcerebral projection neurons.

2.2 Broad Expression of Fezf2 in Early Neural Development

Characterization of the Fezf2^{YFP} hESC line revealed broad, early expression of Fezf2 during neural induction. Using the Fezf2YFP reporter line, we were able to determine that the expression pattern of Fezf2 recapitulates human cortical development with NPCs arising first followed by a population of progenitors and neurons.

2.2.1 Highly efficient neural induction

Fezf2:YFP hESCs underwent a modified dual SMAD inhibition protocol to specifically generate cortical neurons (Figure 2.2.1 A). Immunostain for neural markers (Nestin, Sox2, and Tuj1) verified

successful neural induction (Figure 2.2.1 B). Fezf2 expression was clearly identified in live cells by YFP at Day 12, a period of late neural induction. Cortical precursors were further purified by manual passage to free floating spheres followed by whole sphere replating. After replating, the spheres extended radial glia-like processes and neuroblasts and young neurons started to migrate out (Figure 2.2.1 C).

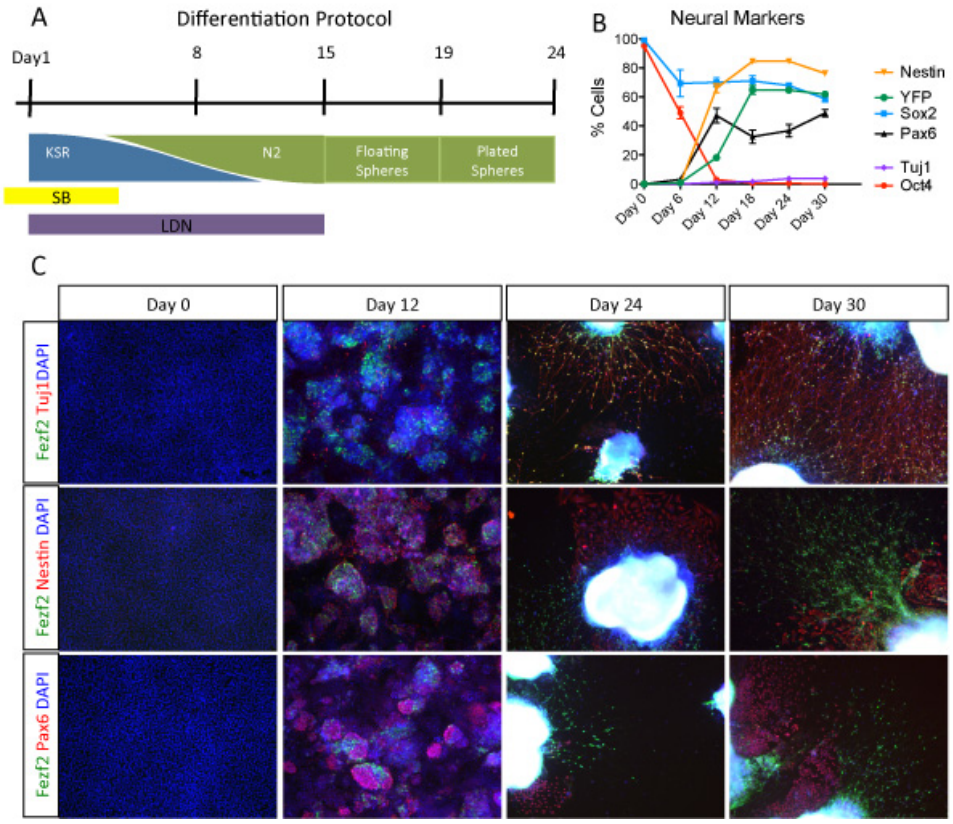

Figure 2.2.1 Highly efficient neural induction.

(A). Directed differentiation protocol, neural induction was initiated in confluent cultures via dual SMAD inhibition. On day 15 of differentiation, cells were manually passaged into floating spheres, four days later these spheres were plated down. (B) Quantification of neural marker expression throughout the course of early neural induction. (C) Representative images from the first 30 days of neural induction, demonstrating Tuj1, Nestin and Pax6 positive cells.

2.2.2 At differentiation day 24, most Fezf2^{YFP} positive cells are enriched for neural markers.

By day 24, Fezf2^{YFP} cells are enriched for neural markers with, nearly all Fezf2 positive cells co-expressing Nestin, and most expressing Sox2. YFP positive cells are likewise enriched for Pax6, compared to YFP negative Cells. (Figure 2.2.2 A, B). Detailed molecular characterization of Fezf2 during early developmental timepoints was performed to determine the pattern of expression (Figure 2.2.2 A, B). Cells were fixed at six-day intervals, starting with day 0 and proceeding through to day 30. Cells were co-stained for YFP, pluripotency and early neural markers.

As anticipated, pluripotency marker expression decreased steadily as neural induction initiated, with Oct4 expression falling rapidly from near ubiquitous expression to less than two percent expression by day 12. By day 18 Nestin expression increased from 0 to 84.5 +/- 1.8%. Fezf2^{YFP} expression was easily detectable by day 12 (expressed in 18.2+/- 1.4% of cells) and continued to increase, with 61.9+/- 1.3 % of cells expressing Fezf2 by day 30. Furthermore, Fezf2 positive cells were enriched for Pax6 (an early neural/forebrain marker) and Sox2 (a neural stem/progenitor marker) at early time points, and maintained high levels of expression throughout differentiation.

In the timeline of our differentiation protocol, Day 24 represents a peak of Fezf2 NPC population and was used as a starting point for later

neuronal differentiation studies. Early in neural induction a majority of NPC express Fezf2, a population that is maintained throughout early neurogenesis. The widespread expression of Fezf2 during these early time points indicates that specificity of this molecular marker is developmentally determined.

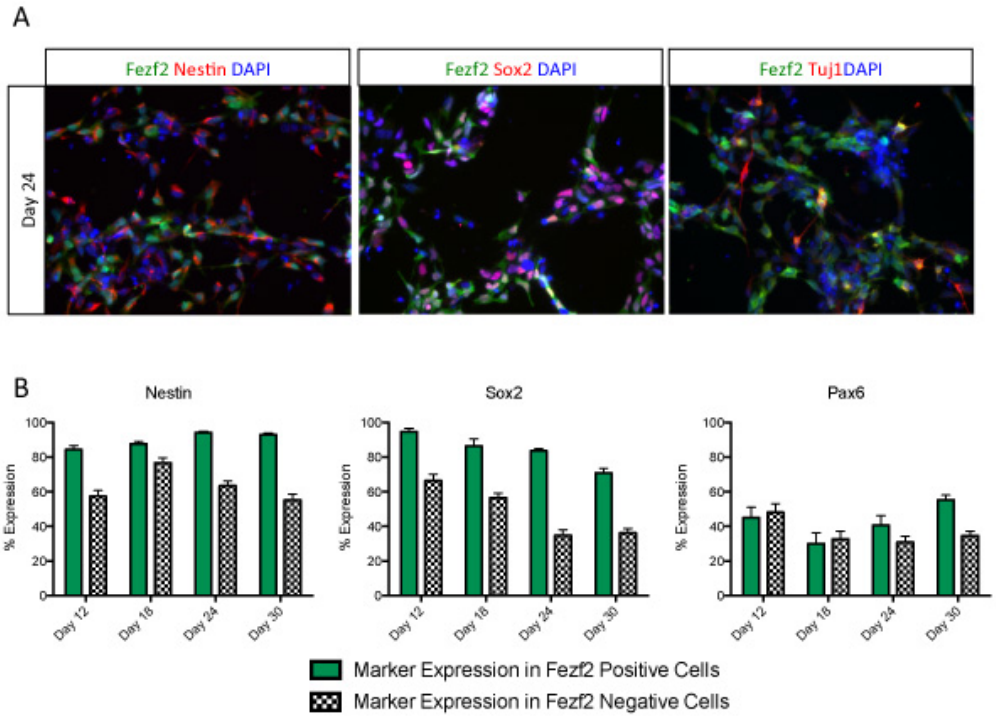

Figure 2.2.2. By differentiation day 24, most Fezf2 positive cells co-express neural stem/ progenitor markers.

(A) Representative images from acutely re-plated day 24 cells expressing YFP and (from left to right) Nestin, Sox2, and Pax6. (B) Quantification of Neural marker expression in YFP positive and negative cells. By differentiation day 24 and after, YFP positive cells are enriched for neural markers compared to YFP negative cells.

2.3 Fezf2 is Expressed in Mitotic Neural Precursors

Previously, Fezf2 mRNA has been identified in the ventricular zone in vivo, but with no cell type delineation. Cortical neuroepithelial cells in vitro

develop in a rosette structure and concomitantly arise with Fezf2 expression (Figure 2.3 A). Further evidence for Fezf2 in human neural precursor cells include co-labeling with the dividing radial glia marker pVimentin and neural precursor marker Nestin (Figure 2.3 B, C). Neural precursor cells are mitotic and time-lapse imaging of endogenous YFP signal showed cell division in Fezf2 NPC (Figure 2.3 D). 626 Fezf2^{YFP} cells over a 30-hour period were analyzed. During this period we did not observe any change in Fezf2 regulation either from YFP- to YFP+ or vice versa. Furthermore, Fezf2 was always passed to both daughter cells after division, although the half-life of YFP complicates this analysis in the 30hr time-lapse period. These results indicate Fezf2 neurons originate from Fezf2 NPC. Quantification of neural stem cell mitosis was divided into symmetric and asymmetric according to the plane of division. No difference was seen in the percentage of dividing Fezf2 cells between early and late points of differentiation. Early in neurogenesis the Fezf2 cells were evenly split between symmetric and asymmetric, while a late neurogenesis time point exhibited 3-fold more symmetric divisions. These results suggest Fezf2 is expressed in both stem and progenitors relating to the ventricular and subventricular zones.

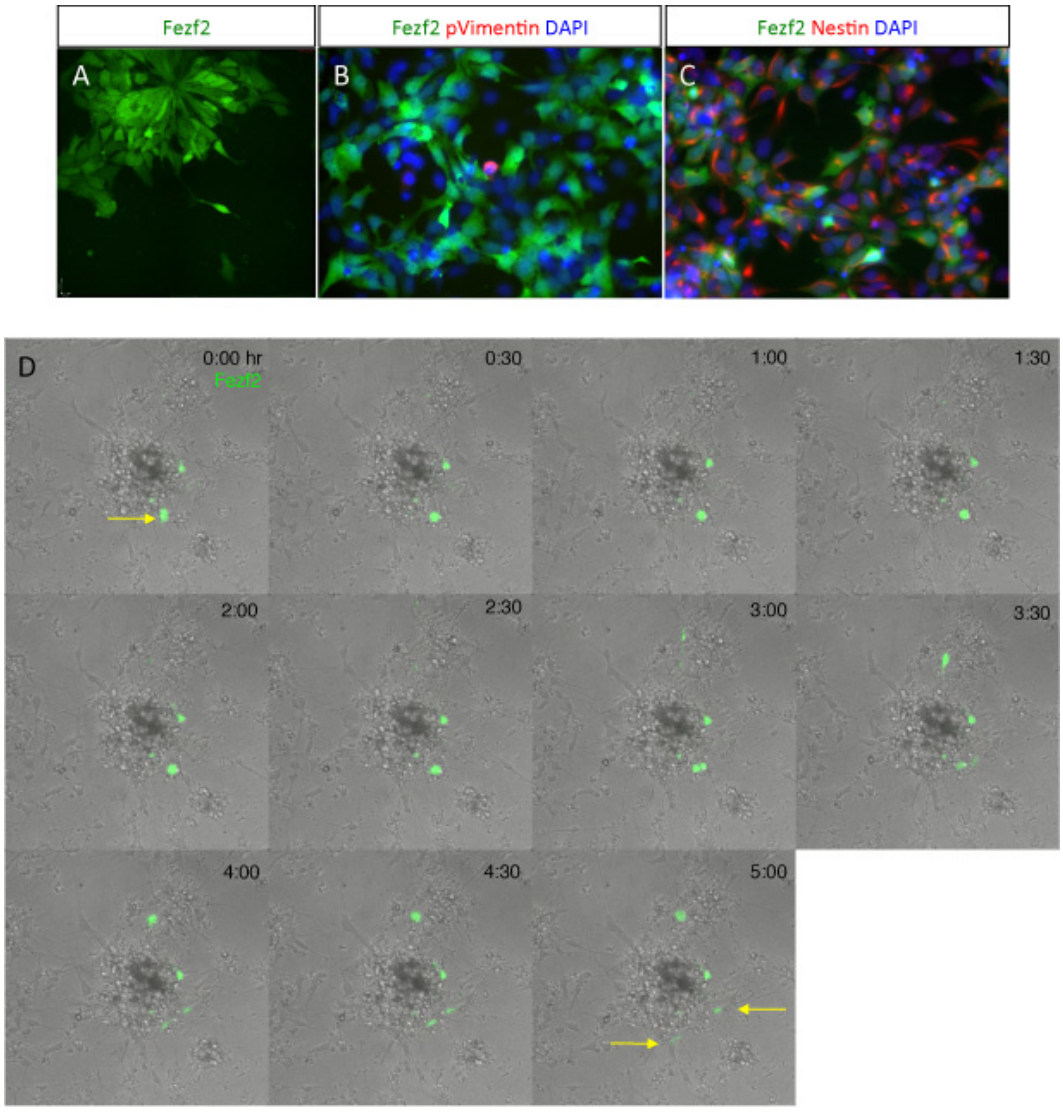

Figure 2.3. Fezf2 is expressed in mitotic neural precursors.

(A) Representative image of a Fezf2^{YFP} positive neural rosette. (B) Dividing Fezf2^{YFP} positive cell co-expressing pVimentin, a marker for dividing radial glia. (C) The neural marker Nestin is highly co-expressed in Fezf2^{YFP} positive cells. (D) Time lapse images of a dividing Fezf2^{YFP} positive cell (yellow arrow).

2.4 Fezf2 Positive NPCs Generate Deep Layer Cortical Projection Neurons

In order to determine the neural subtypes generated from Fezf2 positive progenitors, NSC rich spheres from day 24, day 60, and day 90 were dissociated and sorted via FACS. FACS analysis revealed a spectrum of $Fezf2^{YFP}$ intensities, and samples were collected from $Fezf2^{YFP}$ negative, low, and high expressing populations (Fig.2.4 A). RNA was extracted from each of these populations, and the remaining cells were re-plated at 100,000 cells/cm2 to encourage neuron generation. Twenty-one days later, RNA was extracted from some of these cells, and the remainder were fixed and evaluated by immunostain (Figure 2.4 D). This process is summarized in Figure 2.4 B. Interestingly, the percent of cells composing each population of progenitors shifted as time in culture progressed. At day 24, the Fezf2YFP negative population accounted for 6.4% of gated cells, the low population for 31.0%, and the high population for 62.5%. By day 60 this had shifted to negative 22.2%, low 45.7%, high 32.1%, and at day 90, 42.9%, 39.8%, and 17.3% respectively (Fig.2.4 C). This may represent a change in the overall composition of the NPC spheres across the different time points.

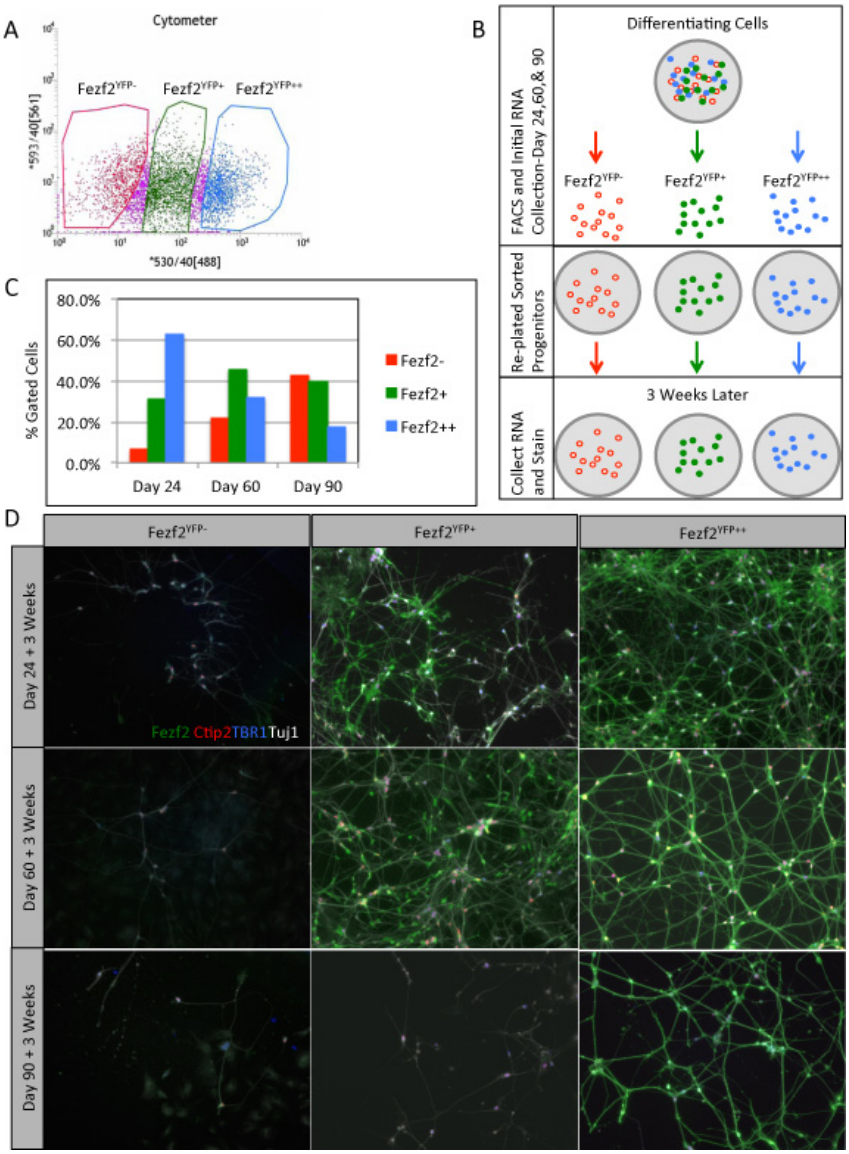

Figure 2.4. Fezf2 positive cells generate deep layer cortical projection neurons.

(A) Representative FACS plot illustrating the three populations collected: Fezf2^{YFP} negative (-), Fezf2^{YFP} low (+) and Fezf2^{YFP} high (++) expressing cells. (B) Cells from each population were collected at differentiation day 24, 60 and 90. RNA was collected immediately, after sort and remaining cells were re-plated at low density. Three weeks later, RNA was collected from these re-plated cells. Remaining cells were fixed and stained. (C) Percent of gated cells making up each population. (D) Progeny of FACS purified cells twenty-one days post sort, stained for YFP, early neuronal marker Tuj1,and deep layer markers Ctip2 and Tbr1.

2.4.1 Evaluation of Fezf2 positive populations via immunostain

As described above, three weeks after re-plating, cells from each time point were stained for Fezf2, Tuj1, an early neuronal marker, Ctip2, a marker highly expressed in layer V, and Tbr1, highly expressed in layer VI (Figure 2.4 D). As one might anticipate, very few Fezf2 positive cells were found in the progeny of the Fezf2^{YFP-} populations, while the progeny of the Fezf2 positive populations were rich in Fezf2 expressing cells. The one exception to this rule is the Fezf2 low expressing population of the day 90 time point. The progeny of this population were very low in Fezf2 positive cells, and more closely resembled the progeny of the Fezf2 negative populations. This may be due to the relatively advanced age of these low expressing progenitors, they may represent a population of cortical progenitors that have advanced past their Fezf2 positive stage.

2.4.1.1 Both time point and Fezf2 expression influence the identity of purified cells and their Fezf2 positive progeny

Twenty-one days post FACS, the progeny of Fezf2^{YFP-}, Fezf2^{YFP+}, and Fezf2^{YFP++} cells were fixed and stained for YFP, indicating Fezf2 expression, Tuj1, indicating a post-mitotic neuronal identity, and DAPI to mark individual cell nuclei. Notably, the identity of the Fezf2 positive cells within these populations was greatly influenced by time point. In all three populations derived from the day 24 time point, the majority of Fezf2 positive cells were not yet post mitotic neurons. Of Fezf2 positive cells

derived from the day 24 YFP- population, only 33.6 +/- 15.3% expressed Tuj1. In Fezf2 positive cells from the YFP+ population, only 41.3 +/- 4.1% expressed Tuj1, and Fezf2 positive cells from the YFP++ population, 36.4 +/- 3.3% were Tuj1 positive (Figure 2.4.1.1 A). In all three populations derived from day 60 time point, the opposite was true. By day 60+21 of differentiation, the majority of Fezf2 positive cells were neurons, with 75.0 +/- 18.7% expressing Tuj1 in the YFP- population, 77.0 +/- 8.4 expressing Tuj1 in the YFP+ population, and 88.6 +/- 2.5% expressing Tuj1 in the YFP++ population (Figure 2.4.1.1 B). Interestingly, the progeny of the day 90 populations were more diverse. The Fezf2 positive cells of the day 90+21 YFP- and YFP+ populations each showed nearly equal distribution between Tuj1 positive and Tuj1 Negative cells. In the YFP- population, 55.8 +/- 30.0% of Fezf2 positive cells expressed Tuj1, and in the YFP+ population, 51.4 +/- 9.6% were Tuj1 positive. In the YFP++ population however, Fezf2 positive cells were enriched for Tuj1, expression, with 77.5 +/- 3.7coexpressing Tuj1 (Figure 2.4.1.1 C). This may indicate that at differentiation day 24+21, the majority of Fezf2 positive cells are still neural progenitors and NSCs, while by day 60+21, Fezf2 positive cells are clearly enriched for neuronal identity, representing potential deep layer projection neurons. At day 90+21, only the progeny of the brightest YFP population are still enriched for the neuronal identity. Fezf2 positive cells in the other two populations may represent Fezf2 positive progenitors, or

possibly even other cell types, such as glia or eye cells, undergoing a Fezf2 positive phase.

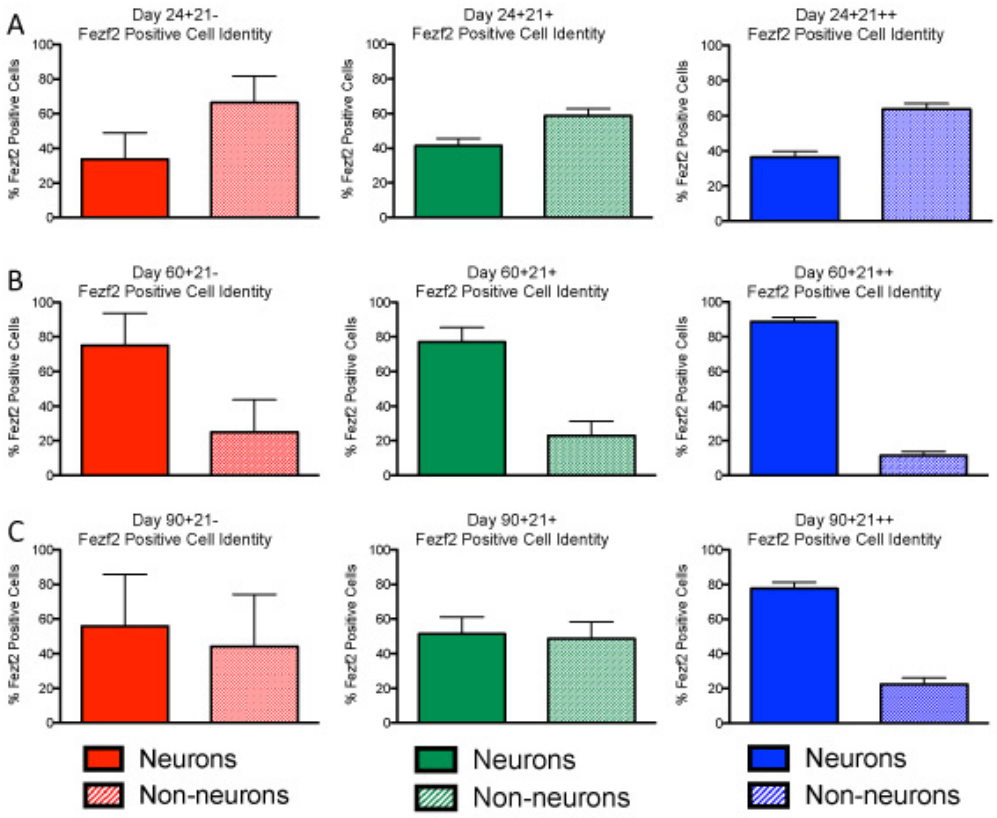

Figure 2.4.1.1. Both time point and Fezf2 expression influence the identity of purified cells and their Fezf2 positive progeny.

(A) The majority of day 24+21 Fezf2 positive cells are Tuj1 negative. (B) Most day 60+21 Fezf2 positive cells are Tuj1 positive neurons. (C) Day 90+21 Fezf2 positive cells from both the YFP- and YFP+ populations are about equal parts neuronal and non-neuronal, while the majority of Fezf2 positive cells from the YFP++ population are Tuj1 positive neurons.

2.4.1.2 Peak production of Fezf2 positive neurons occurs at day 60+21 in the YFP++ Population

The percentage of cells expressing Fezf2 in each population and across all time points correlated with the intensity of YFP expression in its parent population. That is, the Fezf2^{YFP++} populations gave rise to more Fezf2 positive cells than did the Fezf2^{YFP+} populations, and the Fezf2^{YFP+} populations gave rise to more Fezf2 positive cells than did the YFP- populations. The highest percentage of Fezf2 positive cells was observed in the day 60+21 YFP++ population, in which 90 +/- 1.3% of cells were Fezf2 positive (Figure 2.4.1.2 A). Likewise, the highest percentage of Tuj1 expression was also observed at this time point in the same population, with 84.6 +/- 2.8% of cells expressing Tuj1 (Figure 2.4.1.2 B). Not surprisingly, the highest percentage of cells co-expressing Fezf2 and Tuj1 also occurs in this same population at this same time point, with 79.6 +/- 2.0% of cells positive for both Tuj1 and Fezf2 (Figure 2.4.1.2 C). In fact, at day 60+21, in the YFP++ population, 94 +/- 1.2% of neurons are Fezf2 positive, indicating a highly purified population of presumptive deep layer neurons (Figure 2.4.1.2 D).

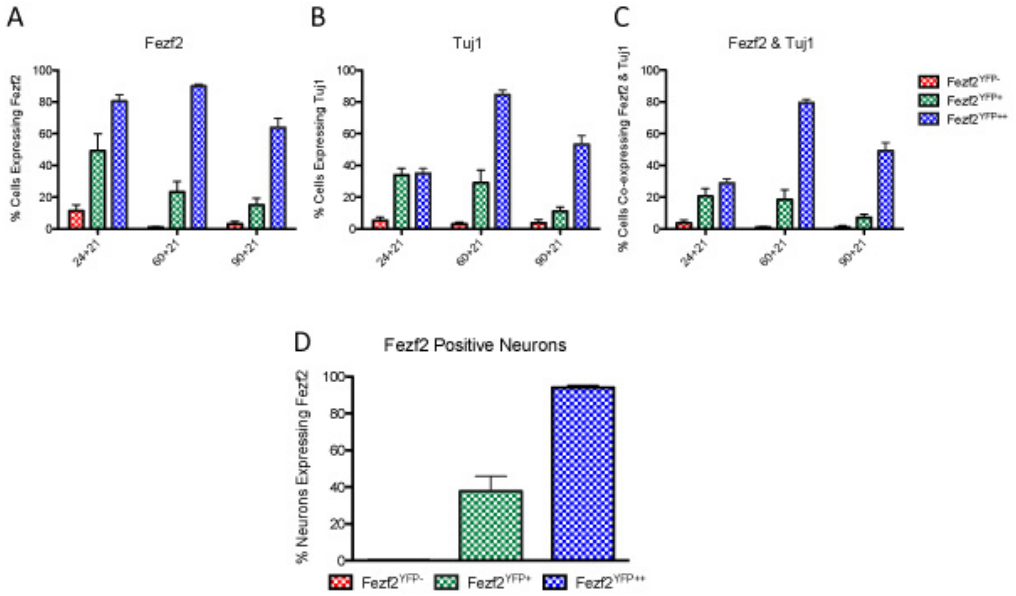

Figure 2.4.1.2. Peak production of Fezf2 positive neurons occurs at day 60+21 in the Fezf2^{YFP++} population.

(A) Across all three time points, the progeny of the Fezf2^{YFP++} population exhibited the highest percentage of Fezf2 positive cells, While progeny of the Fezf2^{YFP-} population showed the lowest percentage of Fezf2 positive cells. (B) The highest percentage of Tuj1 positive cells was observed at day 60+21 in the progeny of the Fezf2^{YFP++} population. (C) Likewise the highest percentage of Fezf2/ Tuj1 co-expression was also observed at day 60+21 in the progeny of the Fezf2^{YFP++} population. (D) In fact, 94 +/- 1.3% of neurons in the day 60+21, Fezf2^{YFP++} population were Fezf2 positive, strongly indicating a deep layer neuronal identity.

2.4.1.3 Ctip2 and Tbr1 are expressed in Fezf2 positive neurons

While the highest percentage of neurons occurs at day 60+21 in the YFP++ population, the neurons that do exist in the YFP++ populations of all three time are highly Fezf2 positive (Figure 2.4.1.3 A). In order to better characterize these neurons, they were co-stained for YFP, Ctip2 and Tbr1, as mentioned above. Interestingly, among the Fezf2 positive neurons of the YFP++ populations, the highest percentage of Ctip2 co-expression occurs at day 24+21 (80.7 +/- 2.7%). By day 60+21 this value as dropped

to 63.5 +/- 3.1% and by day 90+21 it is 35.1 +/- 8.8%. Nevertheless, this means that a significant number of Fezf2 positive neurons are co-express Ctip2 at all three time points, indicating a likely deep-layer identity, and a possible CSMN identity (Figure 2.4.1.3 B). Unlike Ctip2 expression, Tbr1 co-expression in Fezf2 positive neurons remains relatively constant across all three time points. At day 24+21 Tbr1 is co-expressed by 71.1 +/- 2.9% of Fezf2 positive neurons, at day 60+21, 67.9 +/- 2.4% and at day 90, 66.3 +/- 4.4%. Once again this significant percentage of Fezf2/Tbr1 co-expression indicates a deep layer identity. Though Tbr1 is most strongly expressed in layer VI of the cortex, it is also weekly expressed in parts of layer V, and is expressed in the more superficial layers of the cortex. Thus these Tbr1 positive, Fezf2 positive neurons could represent layer V or layer VI neurons, or possibly neurons still in the process of determining their fate. Indeed, a significant percentage of Fezf2 positive neurons in all three time points co-express Fezf2, Ctip2, and Tbr1. At day 24+21, 62.9 +/- 3.5% of Fezf2 positive neurons co-express both Ctip2 and Tbr1, at day 60+21, 61.6 +/- 2.6, and at day 90+21, 34.1 +/- 8.6. This large degree of co-expression was somewhat unanticipated, as Tbr1 is generally thought to promote the Corticothalamic projection neuron fate (Bedogni, Hodge et al. 2010, Han, Kwan et al. 2011, McKenna, Betancourt et al. 2011), while Ctip2 and Fezf2 encourage the CSMN fate (Arlotta, Molyneaux et al. 2005, Alcamo, Chirivella et al. 2008, Ip, Bayatti et al. 2011, Srinivasan, Leone et

al. 2012). Likely, many of these neurons though maintaining a deep layer identity, have yet to determine their fate. In order to further ascertain the identity of these neurons and their sister cells, mid-throughput gene expression analysis was carried out on all initial sorted populations and their progeny.

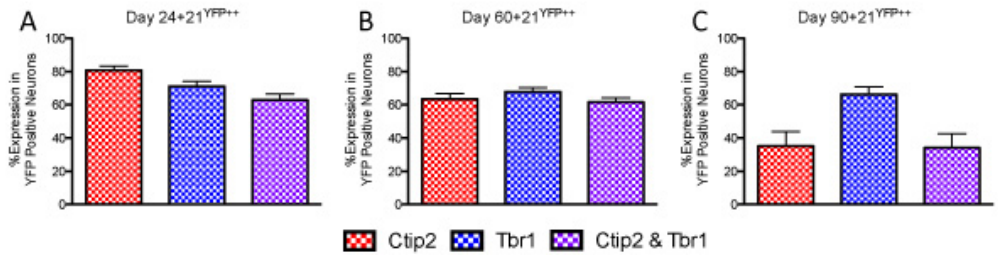

Figure 2.4.1.3. Ctip2 and Tbr1 are expressed in Fezf2 positive neurons.

(A) Fezf2 positive neurons in the Fezf2^{YFP++} population at day 24+21 show the highest percentage of colocalization with Ctip2, compared to the same population at other time points. (B) By day 60+21, Ctip2/ Fezf2 colocalization has decreased slightly, while Tbr1/Fezf2 coexpression is roughly the same. (C) At day 90+21, Fezf2/Ctip2 expression is at its lowest, but still represents nearly 40% of Fezf2 positive neurons. At all three time points, a significant percentage of Fezf2 positive neurons coexpress both Ctip2 and Tbr1.

2.4.2 Gene expression analysis of Fezf2 positive cells identifies populations of deep-layer projection neurons, neural progenitors, and glia

In order to further explore the identity of the cells making up the $Fezf2^{YFP-}$, $Fezf2^{YFP+}$, and $Fezf2^{YFP++}$ populations and their progeny, RNA was collected at day 0, 12, 24, 24+21, 60, 60+21, 90, and 90+21, converted to cDNA and analyzed via Fluidigm Biomark Gene Expression platform against a battery of carefully selected candidate genes. Gene expression was then normalized to Day 0 and control genes Gapdh and Rpl19.

2.4.2.1 Hierarchical clustering reveals populations with shared expression signatures

Hierarchical clustering of these data highlighted several unique clusters including populations rich in deep layer projection neurons, cortical progenitors, and stem cells (Figure 2.4.2.1). Day 60+21++, Day 60+21+ and Day 90+21++ show similar expression patterns, and are enriched for markers common in deep layer projection neurons such as Sox5, Tbr1, Ctip2, Er81, Crym, HCN1, and Vglut1.

Day 90++, 24+21++, 24+21+, 60+, and 60++ highly express progenitor markers such as pax6, sox1, sox2, Foxg1, and Tbr2 , as well as weekly expressing deeplayer neuronal markers, potentially indicating a

mixed population of many progenitors and some deep layer projection neurons. Day 24++ and 24+ also highly express progenitor markers.

The Fezf2 negative populations also cluster together and are enriched for markers common to the human eye (Pax6, Chx10, and Cryab), as well as markers for mesoderm, mesencephalon, diencephalon, hindbrain, neural crest and the cajal retzius zone. Interestingly, Day 90+ and Day 90+21+ also cluster with the Fezf2 negative populations. This was perhaps to be expected as immunostain of the day 90+21+ cells reveled few neurons and little Fezf2 expression (Fig. 3D).

The two stem cell populations evaluated also clustered together, highly expressing Oct4.

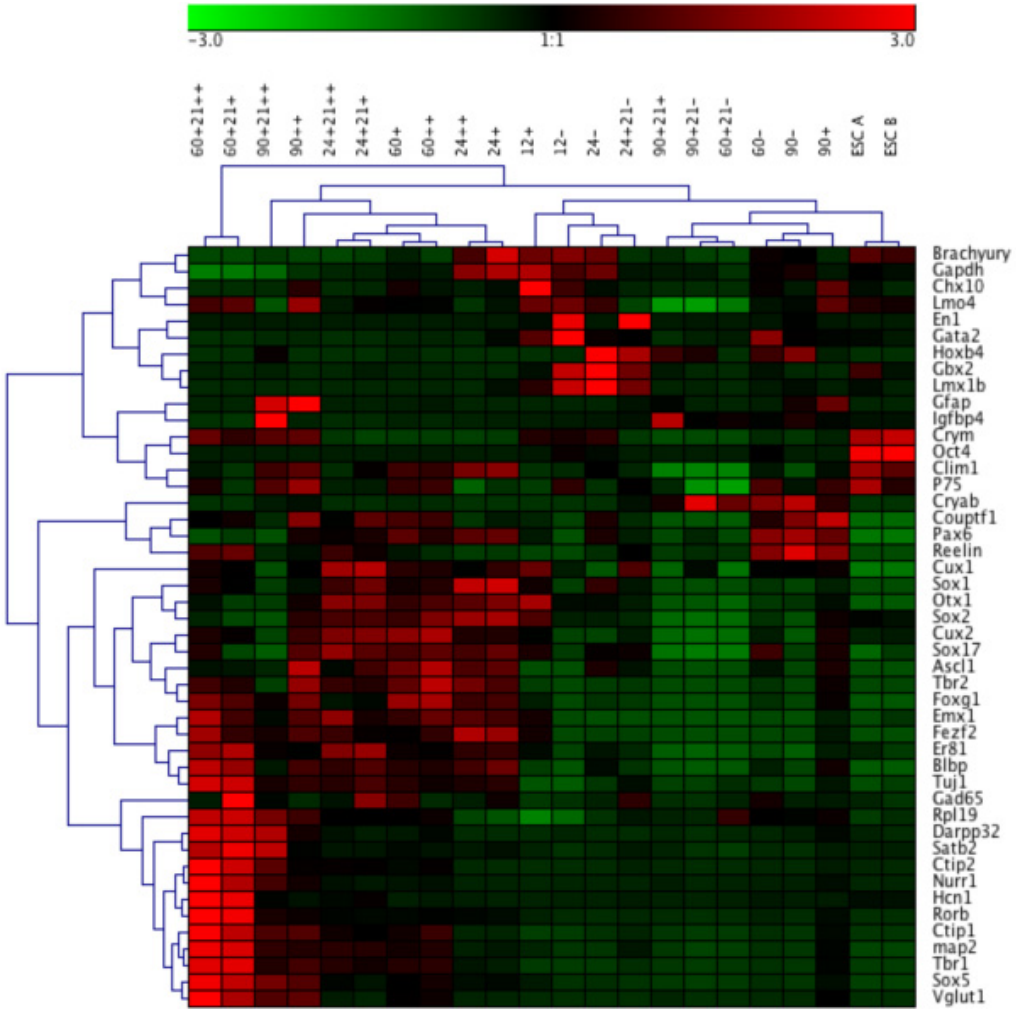

Figure 2.4.2.1 Hierarchical clustering reveals populations with shared expression signatures.

The first three columns on the left show a similar expression pattern, and are enriched for deep layer projection neuron markers. The next six columns also display a similar expression pattern, and are enriched for NSC/NPC markers, as well as some projection neuron markers. The next cluster is composed primarily of Fezf2 negative cells, and is enriched for a variety of non-cortical markers, Brachyury, Chx10, Hoxb4, and Cryab. Finally, the two hESC populations cluster together on the far right side, and are enriched in stem cell markers like Oct4.

2.4.2.2 Gene expression analysis indicates that Fezf2 positive cells give rise to multiple neural subtypes, including CSMNs, callosal projection neurons, and glia.

During Cortical development, Reelin is expressed in the most superficial layer of the cortex, as well as the deep Cajal Retzius zone. Interestingly, Reelin was found to be expressed in both Fezf2 positive and negative populations and their progeny (Figure 2.4.2.2 A). Interestingly, at the time of FACS Reelin is most highly expressed in the YFP- populations, however, twenty-one days later, Reelin is usually most highly expressed in the progeny of Fezf2 positive cells. This indicates that Reelin positive cells may arise from Fezf2 positive cells, but likely mature into Fezf2 negative cells.

Satb2, which is expressed in layer V callosal projection neurons, as well as superficial layers is also highly expressed in the progeny of Fezf2 positive progenitors at day 60+21, and is also highly expressed in the progeny of Fezf2YFP++ progenitors at day 90+21, suggesting that Satb2 positive neurons also arise from Fezf2 positive NPCs (Figure 2.4.2.2 B).

GFAP, a glial/ astrocytic marker was highly expressed only at the latest time points, primarily in Fezf2 positive populations, indicating that Fezf2 positive progenitors likely give rise to populations of astrocytes as well (Figure 2.4.2.2 C).

Examination of classical layer specific markers demonstrates that layer V and VI projection neuron production, indicated by Ctip2 and Tbr1 expression, respectively, reaches a peak among Fezf2 positive cells at day 60+21. Satb2, which is expressed in layer V callosal projection neurons, as well as superficial layers is also highly expressed in the progeny of Fezf2 positive progenitors at day 60+21, and is also highly expressed in the progeny of $Fezf2^{YFP++}$. progenitors at day 90+21, suggesting that Satb2 positive neurons also arise from Fezf2 positive NPCs (Figure 2.4.2.2 D-F).

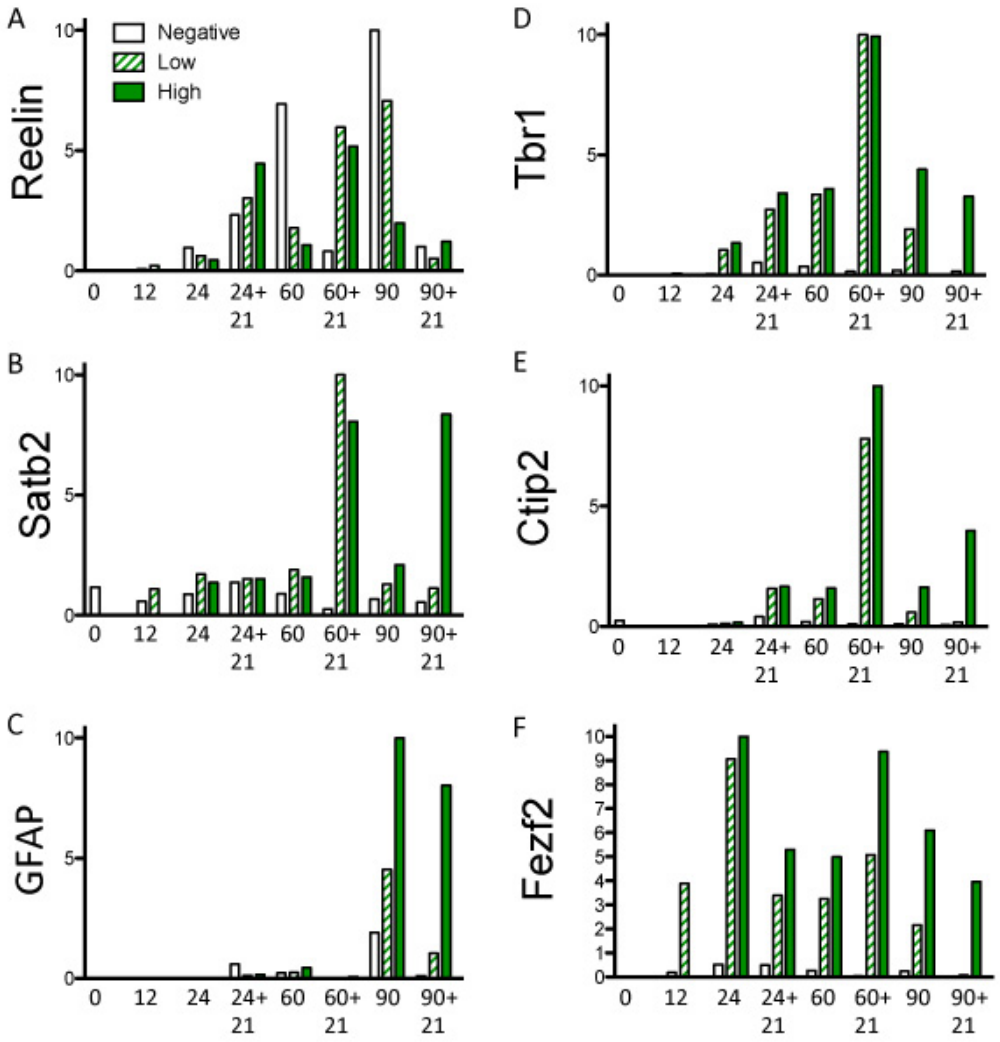

Figure 2.4.2.2. Gene expression analysis indicates that Fezf2 positive cells give rise to multiple neural subtypes, including CSMNs, callosal projection neurons, and glia.

(A) Reelin is expressed in the deep Cajal Retzius zone as well as in the most superficial layer of the cortex. At the time of FACS, Reelin is most highly expressed in the YFP negative populations, however, 21 days post-sort, Reelin is more highly expressed in the progeny of the Fezf2 positive populations, suggesting that Reelin positive cells are derived from Fezf2 positive progenitors, but are primarily Fezf2 negative. (B) In the murine cortex, Satb2 is highly expressed in layer V callosal projection neurons and in layer II – IV neurons. Gene expression analysis reveals that Satb2 is highly expressed in the progeny of Fezf2 positive day 60 and day 90 cells, indicating that Fezf2 positive NPCs/ NSCs give rise to these neural subtypes as well. (C) GFAP is expressed in astrocytes and other glia, and here, is expressed in Fezf2 positive and negative cells at day 90, as well as in the progeny of Fezf2 positive day 90 cells, indicating that glia may arise from Fezf2 positive precursors. (D, E) Deep layer markers Ctip2 and Tbr1 are expressed in Fezf2 positive cells and their progeny, peaking at day 60+21. (F) Fezf2 expression corresponds with YFP intensity.

2.5 Evaluation of Fezf2 Positive Cells Through In Vivo Transplantation

Generating an appropriate molecular signature for deep-layer projection neurons in vitro as very encouraging, however, the primary definition of a CSMN is that it projects along descending motor pathways from the cortex to the spinal cord. In order to further evaluate the identity of our Fezf2 positive NPCs, cells from day 24 of differentiation were sorted via FACS and Fezf2 positive cells were transplanted into the cortices of neonatal mice. Several weeks later, these cells were examined to assess their molecular identity and to determine if they had achieved an appropriate morphology. Finally, the axonal projections of these transplanted cells were evaluated to see if they extended along appropriate pathways and whether they could appropriate subcerebral targets.

2.5.1 Fezf2 positive cells achieve a pyramidal morphology and maintain an appropriate molecular identity in vivo.

Differentiation day 24 FACS purified Fezf2 positive cells were transplanted into the brains of neonatal (p0 to p6) mice. Four weeks post transplant, Fezf2 positive cells exhibited a pyramidal morphology, typical of corticospinal motor neurons (Figure 2.5.1 A, B). In layer V of the developing cortex, CSMNs and callosal projection neurons are born at the

same time. Due to the antagonistic relationship between the transcription factors governing this fate choice, Fezf2 and Ctip2 supporting the CSMN fate and Satb2 supporting the callosal fate, mature projection neurons in this region can express either Fezf2 or Satb2, but not both (Figure 2.5.1 C). Transplanted cells follow this rule, appropriately expressing either Fezf2 or Satb2. Remarkably, as these cells were originally FACS purified Fezf2 positive, this Satb2 positive Fezf2 negative population arose from Fezf2 positive progenitors. This strongly indicates that Fezf2 positive NPCs are not restricted to generating only subcerebral projection neurons. Additionally, transplanted Fezf2 positive neurons were able to maintain a layer V molecular identity, evidenced by Fezf2 /Ctip2 co-expression. This was true even when the transplant was in an environment other than layer V of the cortex, indicating that transplanted cells were capable of maintaining their molecular identity independent of immediate environment (Figure 2.5.1 D).

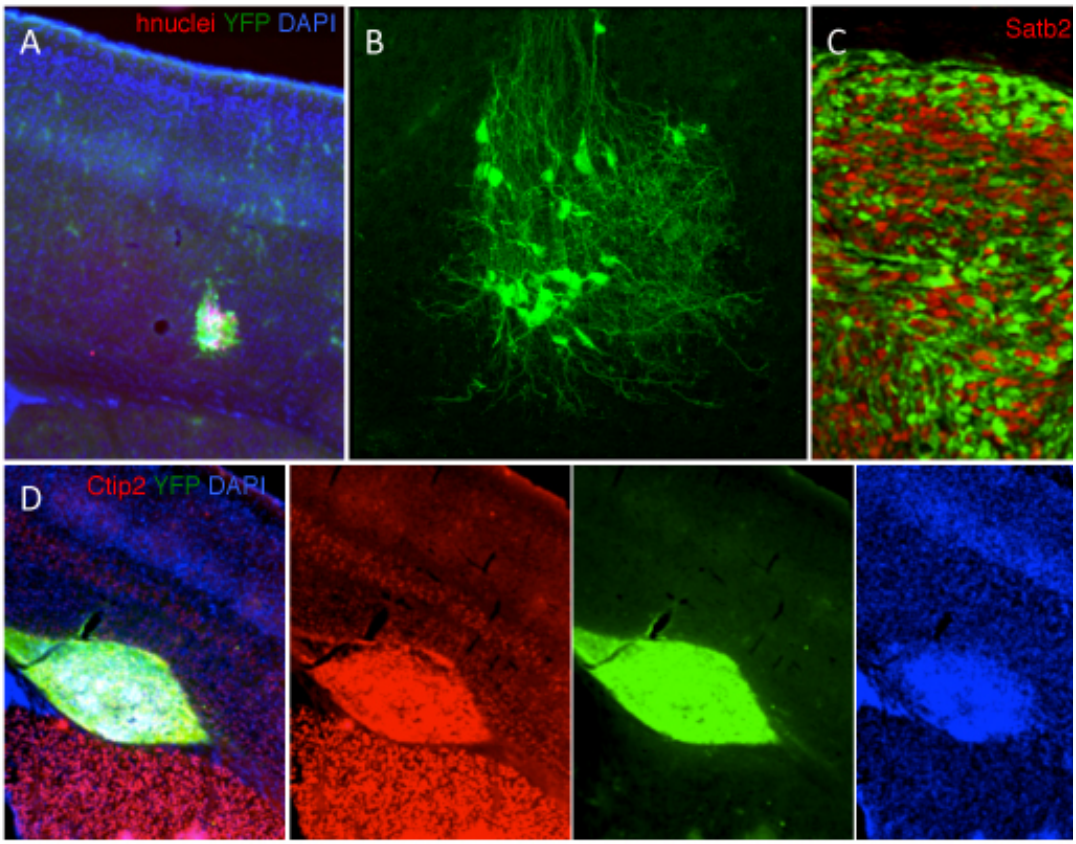

Figure 2.5.1. Fezf2 positive cells achieve a pyramidal morphology and maintain an appropriate molecular identity in vivo.

(A, B) Transplanted Fezf2^{YFP} positive cells exhibited a mature pyramidal morphology. (C) FACS purified, Fezf2 positive cells gave rise to both Fezf2 positive neurons and Fezf2 negative, Satb2 positive cells in vivo. (D) Transplanted Fezf2 positive cells were able to maintain a layer V morphology, evidenced by Fezf2/ Ctip2 coexpression, even in ectopic locations.

2.5.2 Transplanted Fezf2 positive neurons projected to appropriate subcerebral targets in vivo, even reaching the spinal cord.

Some neuronal subtypes, such as serotonergic neurons or dopaminergic neurons can be classified largely by their ability to produce

a unique neurotransmitter (*eg*. serotonin or dopamine). This is not the case with CSMNs, as they, like all excitatory projection neurons in the cortex, produce glutamine as their primary neurotransmitter. Perhaps the single most distinctive feature that defines CSMNs is the trajectory of their axonal projections. As neither the laminar structure of the cortex nor the long distance targets of CSMNs are present in vitro, the axonal trajectory of our Fezf2 positive neurons could only be assessed through in vivo transplantation. This presents several difficulties. First, the mouse cortex differs significantly from the human cortex, and so may the axon guidance systems that direct murine CSMN axons to their targets. Secondly, by the time of transplantation, the host CST has already formed, and the axon guidance systems that directed these axons to their targets are no longer in place. Nevertheless, transplanted Fezf2 positive cells were able to project axons through the internal capsule, the cerebral peduncle, the pons, the pyramidal tract, the pyramidal decussation, and out to the spinal cord (Figure 2.5.2).

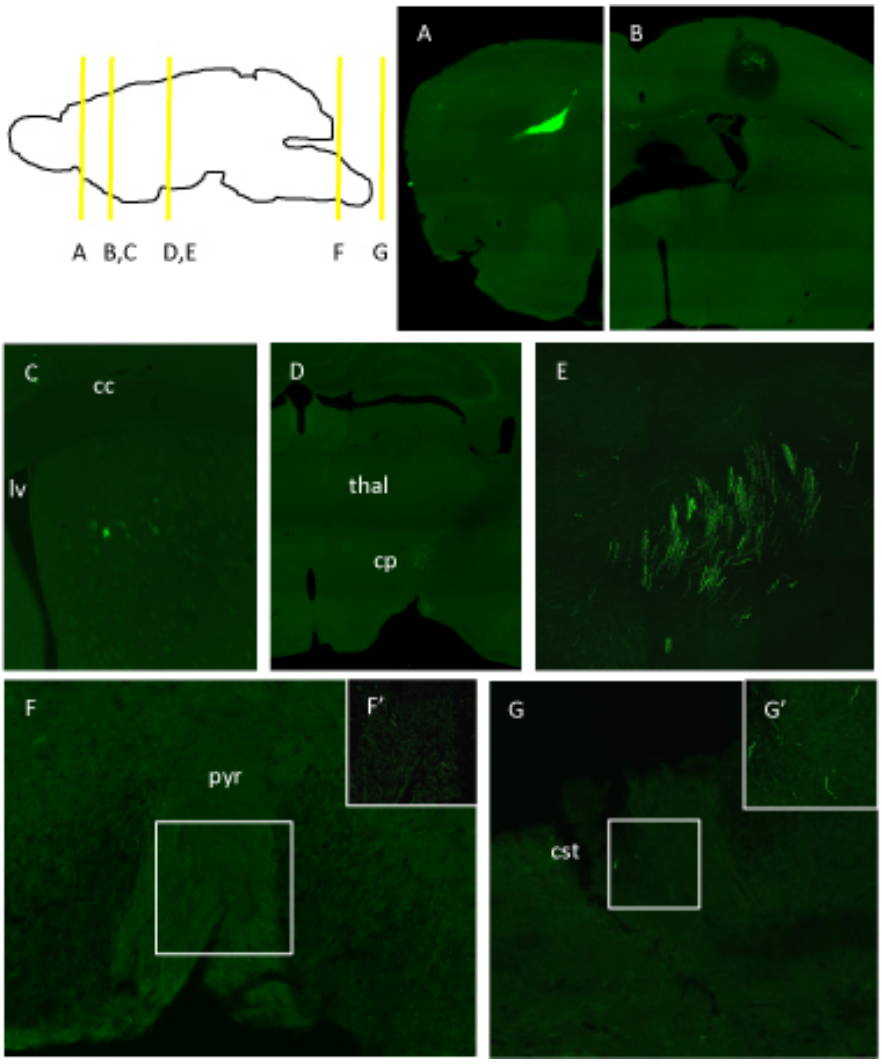

Figure 2.5.2. Transplanted Fezf2 positive neurons projected to appropriate subcerebral targets in vivo, even reaching the spinal cord.

(A, B) Representative images of transplanted Fezf2 positive cells. (B, C) Fezf2 axons at the level of the internal capsule, (D, E) the cerebral peduncle, (F, F') the pyramidal decussation, and (G, G') the spinal cord.

Chapter 2, in part is currently being prepared for submission for publication of the material. Smith, Anders & Hinckley, Sandy; Feltham, Casey; Rodriguez, Karen; Zapata, Arturo; Zhang, Sherry; Weimann, James; Zheng, Binhai. "Toward Engineering the Corticospinal Tract from Human

Embryonic Stem Cells". The dissertation author was the co-principal researcher/author of this material.

2.6 Materials and Methods

2.6.1 Stem cell culture

Hues9 hESCs were cultured on mouse embryonic fibroblasts in HUES media as described previously (Ruby and Zheng 2009), except that cells were supplemented with 20ng/ml bfgf. For one to two passages immediately prior to FACS, cells were grown under feeder free conditions on matrigel (BD Biosciences), and fed daily with mouse embryonic fibroblast (MEF) conditioned media (MCM) supplemented with 20ng/ml bfgf.

2.6.2 Fluorescence assisted sell Sorting

Cells were grown under feeder free conditions for two passages prior to FACS, on matrigel, and fed daily with MEF conditioned media supplemented with bfgf. Cells were dissociated with accutase, resuspended and passed through a 40μm cell strainer. Cells were then centrifuged at 1000 RPM for 4 minutes. Cells were resuspended in sort buffer, consisting of Ca++/Mg++ free HBSS (Life Technologies) with 5mM EDTA and 1%FBS, at a concentration of about 6 million cells/ ml, and kept on ice until FACS. Cells were sorted into collection tubes containing

collection media comprised of Ca++/Mg++ free HBSS with 20% KSR. FACS was carried out at the UCSD Stem Cell Core using either the FACS Aria II (BD Biosciences) or the Influx (BD Biosciences), depending on availability.

2.6.3 Stem cell differentiation

Stem Cells were differentiated into neurons using dual SMAD inhibition. Cells were grown in feeder-free conditions on matrigel for one passage prior to induction of differentiation. Cells were then trypsinized, and re-plated on plates coated with growth factor-reduced matrigel (BD Biosciences). Cells were fed with MEF-conditioned media supplemented with 20ng/ml bfgf until 90 to 100% confluent. Cells were then fed with 50% MCM and 50% hESDiff Media (Knockout DMEM with 15% KSR, 1X NEAA and 1x Glutamax, all from Life Technologies) supplemented with 20 ng/ml bfgf and 5μM SB431542 (SB, Stemgent), for one day. Followed by 100% hESDiff media supplemented with 5μM SB431542 and 100nM LDN-193189 (LDN, Stemgent) for 4 days, 75% hESDiff/ 25% N2 Media (DMEM/F12 and 1X N2 supplement, both from Life Technologies-Gibco) supplemented with 100nM LDN for 2 days, 50% hESDiff/ 50% N2 Media supplemented with 100nM LDN for 2 days, 25% hESDiff/ 75% N2 with 100nM LDN for 2 days, and then 100% N2 supplemented with 100nM LDN for 3 days. On day 14 of differentiation, cells were manually passaged into spheres. Spheres were then allowed to float for 4 days in

N2 Media. On day 19 of differentiation, cells were plated down on plates coated with growth factor-reduced matrigel, in N2 Media. Cells were fed with half media changes every other day until day 24. On day 24 of differentiation media was switched to Neural Differentiation Media (NDM) consisting of Neurobasal Media (Life Technologies) with 1X Glutamax (Life Technologies), 0.5X N2 Supplement (Life Technologies), and B27 Supplement without Vitamin A (Life Technologies), after which half media changes were made as needed. Cells were then either fixed directly with 4%PFA or dissociated and re-plated at a dilution of 1:3 and fixed after settling overnight. In the case of FACS purified, re-plated cells, cells from each population were plated at a concentration of 100,000 cells /cm2 and permitted to grow for 21 days, after which they were fixed with 4% PFA.

2.6.4 Immunofluorescence and quantification

Cells and brain slices were fixed with 4%PFA. After fixation, the following Primary antibodies were used: Ctip2: rat anti Ctip2 (abcam, ab18465), Human Nuclei: mouse anti Human Nuclei (Millipore, MAB1281), Nestin: mouse anti Nestin (Millipore, MAB5326), Oct4: mouse anti Oct4 (Santa Cruz Biotechnology, SC-5279), Pax6: rabbit anti Pax6 (Covance, PRB-278P-100), Satb2: mouse anti Satb2 (abcam, ab51502) and rabbit anti Satb2 (abcam, ab34735), Sox2: rabbit anti Sox2 (abcam, ab97959), Tbr1: rabbit anti Tbr1 (abcam, ab31940), Tuj1: mouse anti Tuj1 (Covance, MMS-435P-250), YFP: chicken anti GFP (Life Technologies, A10262).

Secondary antibodies used include: 488 goat anti chicken (Life Technologies, A11039), 546 donkey anti rabbit (Life Technologies A10040), 546 donkey anti mouse (Life Technologies, A10036), 546 goat anti rat (Life Technologies, A11081), 680 goat anti mouse (Life Technologies, A21058), 680 donkey anti goat (Life Technologies, A21084), 680 donkey anti rabbit (Life Technologies, A10043). In the case of the four color stain for YFP, Ctip2, Tbr1 and Tuj1, a tertiary stain with Streptavardin was employed to enhance Tbr1 fluorescence in the blue channel. In all other conditions, DAPI was used to mark cell nuclei in the blue channel. For each condition, at least three images were taken in each of at least three wells (a minimum of 9 images). Cells from each image were then counted in imageJ. The values shown here are the mean and SEM, with each image having equal weight.

2.6.5 Gene Expression Analysis

RNA was extracted form FACS purified cells using RNEasy extraction kit (Qiagen). cDNA was generated using Superscript Vilo Mastermix (Life Technologies). Assays were purchased from Deltagene. Fluidigm Biomark analysis was carried out according to manufacturer's instructions.

2.6.6 Cell transplantation

Cells were differentiated as described above for 24 days. At day 24 of differentiation, cells were dissociated with accutase for 10 to 20 minutes, accompanied by trituration. Dissociated cells were then suspended in NDM, passed through a 40μm cell strainer, and centrifuges for 4 minutes at 100 rpm. Excess media was aspirated and the cell pellet was re-suspended in sort buffer. Cells were then purified for YFP positive cells via FACS. Purified Fezf2:YFP positive cells were diluted to a concentration of 30,000 cells / μl, prior to transplantation. Neonatal pups, P0 to P6, were anesthetized using an ice water bath to induce hypothermia. Anesthetized mice were then placed on a ice on a platform within the stereotactic injection device. The skin was partially pulled back from the skull, and the skull was perforated using a metal needle tip. Approximately 15,000 cells were injected into the pup cortices through these perforations via a pulled glass needle, under the control of a stereotactic device. The skin flaps were then pushed back to their original position and sealed using veterinarian glue. Mouse pups regained consciousness upon warming in a water bath, and were then returned to the care of their mothers. Two to ten weeks later, mice were sacrificed and fixed via perfusion with 4%PFA. Brains were then dissected out and submerged in 4%PFA, followed by sucrose. Fixed brains were sectioned into 50μM sections via cryostat, stained and imaged.

2.7 Discussion

The two primary goals of this work were: first, to evaluate the role of Fezf2 in cortical development, and second, to generate CSMNs from hESCs. Fezf has long been recognized as an important marker and molecular determinant for murine CSMNs (Chen, Schaevitz et al. 2005, Molyneaux, Arlotta et al. 2005, Chen, Wang et al. 2008, Rouaux and Arlotta 2010). It has also been observed that Fezf2 is expressed both in developing CSMNs in layer V of the murine cortex, and in the neural stem and progenitor cell populations of the ventricular and subventricular zones. Interestingly, relatively little attention has been paid to these stem and progenitor cell populations. It has however been observed that these Fezf2 positive radial glia both predate and persist long after the generation of the deep layer projection neurons of the cortex where Fezf2 is know to be expressed. Thus, while Fezf2 is expressed in many of the progenitors that give rise to these deep-layer projection neurons, it appears that Fezf2 positive progenitors do not represent a lineage restricted deep-layer progenitor population. Indeed, recently, through elegant fate mapping experiments in genetically modified mice, it has been determined that stem and progenitor cells that at one point expressed Fezf2, give rise to neurons that populate every layer of the adult cortex, and even glia (Guo, Eckler et al. 2013).

In this work we have found that Fezf2 is likewise expressed in human cortical progenitor cells, and further that these cells are mitotic. This means that in human cells, Fezf2 is not only expressed in post-mitotic young neurons, but also in actively dividing neural stem cells, including radial glia.

Additionally, we have demonstrated that these Fezf2 positive neural precursors do indeed give rise to deep layer projection neurons. This is evidenced by Fezf2-Ctip2 and Fexf2-Tbr1 co-expressing neurons arising from FACS purified Fezf2 positive cells. Additional markers, including those characteristic of CSMNs, were also found to be expressed in these populations. In addition to the anticipated deep-layer, subcerebral projection neurons, other neural cell types also arose from Fezf2 positive progenitors, including Satb2 positive neurons. Satb2 is known to be expressed in the Fezf2 negative callosal projection neurons of cortical layer V, as well as in superficial layer projection neurons (Alcamo, Chirivella et al. 2008, Britanova, de Juan Romero et al. 2008). The observation that neural cell types expressing Satb2 arose from Fezf2 positive progenitors indicates that in humans, as in mice, Fezf2 positive stem and progenitor cells do give rise to Fezf2 negative cortical neurons, exhibiting markers consistent with both deep and superficial layers. Indeed, purified Fezf2 positive cells transplanted into mice gave rise to both Fezf2 positive neuronal populations and Satb2 positive, Fezf2 negative neurons.

Additionally, gene expression analysis of late stage Fezf2 positive progenitors and their progeny revealed an upregulation of GFAP expression in both Fezf2 negative and Fezf2 positive populations, suggesting that Fezf2 positive progenitors give rise to glia as well as projection neurons. Thus Fezf2 positive stem/ progenitor cells are not restricted to generating only subcerebral projection neurons, but instead represent an important precursor to cortical neurons of both deep and superficial layers as well as glia.

Interestingly, many of the Fezf2 positive neurons generated in vitro, co-expressed Fezf2/Ctip2 and Tbr1. While co-expression of these transcription factors does exist in a subset of deep-layer neurons in the murine cortex, the frequency with which this co-expression was observed was surprising, particularly because the Fezf2/Ctip2 pathway is generally thought to develop in opposition to the Tbr1 mediated pathway (Han, Kwan et al. 2011, McKenna, Betancourt et al. 2011), and vice versa. In the case of these neurons, it may be this reciprocal relationship that leads to the high level of co-expression. Morphologically, the Fezf2 positive neurons generated in vitro are still relatively immature. While they do exhibit the customary triangular shaped cell body of pyramidal neurons, the apical dendrites are difficult to distinguish, and the degree of arborization is minimal, especially when compared to those generated in vivo from transplanted cells (Figure 2.5.1 A, B). Thus our in vitro neurons,

in the absence of cortical environmental cues, are likely immature and it is possible the fate choice between the subcerebral and corticothalamic fates is not yet complete. It may be possible to rectify this through adaptations to the late stage differentiation protocol. Nevertheless, through directed differentiation, timing and FACS purification, we were able to generate a population of neurons that were 95% positive for Fezf2, in vitro.

As demonstrated above, many transplanted Fezf2 positive neurons were able to extend axons selectively down appropriate subcerebral motor pathways, even reaching the pyramidal decussation and the spinal cord. This is remarkable considering that the endogenous axonal guidance systems would no longer be in place at the time of transplant. However, not all transplanted cells exhibited this degree of selectivity. Very large transplants in particular projected to both appropriately and ectopically, possibly because they were large enough to create a unique microenvironment within the larger environment of the host cortex. Some smaller transplants also projected to targets inappropriate for subcerebral projection neurons, such as projecting across the corpus callosum, or along other white matter tracts. This suggests that molecular identity alone is not sufficient to assure appropriate axonal trajectory. There is likely a dependence on graft size as well as precise graft location in determining the successfulness of transplanted neurons in projecting to their intended

targets. Even so, many transplanted neurons were able to maintain their molecular identity, achieve a mature pyramidal morphology, and project to correct subcerebral targets.

2.8 References

Alcamo, E. A., L. Chirivella, M. Dautzenberg, G. Dobreva, I. Farinas, R. Grosschedl and S. K. McConnell (2008). "Satb2 regulates callosal projection neuron identity in the developing cerebral cortex." Neuron **57**(3): 364-377.

Arlotta, P., B. J. Molyneaux, J. Chen, J. Inoue, R. Kominami and J. D. Macklis (2005). "Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo." Neuron **45**(2): 207-221.

Bedogni, F., R. D. Hodge, G. E. Elsen, B. R. Nelson, R. A. Daza, R. P. Beyer, T. K. Bammler, J. L. Rubenstein and R. F. Hevner (2010). "Tbr1 regulates regional and laminar identity of postmitotic neurons in developing neocortex." Proc Natl Acad Sci U S A **107**(29): 13129-13134.

Britanova, O., C. de Juan Romero, A. Cheung, K. Y. Kwan, M. Schwark, A. Gyorgy, T. Vogel, S. Akopov, M. Mitkovski, D. Agoston, N. Sestan, Z. Molnar and V. Tarabykin (2008). "Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex." Neuron **57**(3): 378-392.

Chen, B., L. R. Schaevitz and S. K. McConnell (2005). "Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex." Proc Natl Acad Sci U S A **102**(47): 17184-17189.

Chen, B., S. S. Wang, A. M. Hattox, H. Rayburn, S. B. Nelson and S. K. McConnell (2008). "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 **105**(32): 11382-11387.

Dale Purves GJA, D. F., Lawrence, A.-S. L. C Katz, James O McNamara, and S Mark and J. L. Williams (2001). Neuroscience, Sinauer Associates, Inc.

Espuny-Camacho, I., K. A. Michelsen, D. Gall, D. Linaro, A. Hasche, J. Bonnefont, C. Bali, D. Orduz, A. Bilheu, A. Herpoel, N. Lambert, N. Gaspard, S. Peron, S. N. Schiffmann, M. Giugliano, A. Gaillard and P.

Vanderhaeghen (2013). "Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo." Neuron **77**(3): 440-456.

Gaspard, N., T. Bouschet, R. Hourez, J. Dimidschstein, G. Naeije, J. van den Ameele, I. Espuny-Camacho, A. Herpoel, L. Passante, S. N. Schiffmann, A. Gaillard and P. Vanderhaeghen (2008). "An intrinsic mechanism of corticogenesis from embryonic stem cells." Nature **455**(7211): 351-357.

Guo, C., M. J. Eckler, W. L. McKenna, G. L. McKinsey, J. L. Rubenstein and B. Chen (2013). "Fezf2 expression identifies a multipotent progenitor for neocortical projection neurons, astrocytes, and oligodendrocytes." Neuron **80**(5): 1167-1174.

Han, W., K. Y. Kwan, S. Shim, M. M. Lam, Y. Shin, X. Xu, Y. Zhu, M. Li and N. Sestan (2011). "TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract." Proc Natl Acad Sci U S A **108**(7): 3041-3046.

Ip, B. K., N. Bayatti, N. J. Howard, S. Lindsay and G. J. Clowry (2011). "The corticofugal neuron-associated genes ROBO1, SRGAP1, and CTIP2 exhibit an anterior to posterior gradient of expression in early fetal human neocortex development." Cereb Cortex **21**(6): 1395-1407.

Leone, D. P., K. Srinivasan, B. Chen, E. Alcamo and S. K. McConnell (2008). "The determination of projection neuron identity in the developing cerebral cortex." Curr Opin Neurobiol **18**(1): 28-35.

McKenna, W. L., J. Betancourt, K. A. Larkin, B. Abrams, C. Guo, J. L. Rubenstein and B. Chen (2011). "Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development." J Neurosci **31**(2): 549-564.

Molyneaux, B. J., P. Arlotta, T. Hirata, M. Hibi and J. D. Macklis (2005). "Fezl is required for the birth and specification of corticospinal motor neurons." Neuron **47**(6): 817-831.

Molyneaux, B. J., P. Arlotta, J. R. Menezes and J. D. Macklis (2007). "Neuronal subtype specification in the cerebral cortex." Nat Rev Neurosci **8**(6): 427-437.

Rouaux, C. and P. Arlotta (2010). "Fezf2 directs the differentiation of corticofugal neurons from striatal progenitors in vivo." Nat Neurosci **13**(11): 1345-1347.

Rouaux, C. and P. Arlotta (2013). "Direct lineage reprogramming of post-mitotic callosal neurons into corticofugal neurons in vivo." Nat Cell Biol **15**(2): 214-221.

Ruby, K. M. and B. Zheng (2009). "Gene targeting in a HUES line of human embryonic stem cells via electroporation." Stem Cells **27**(7): 1496-1506.

Srinivasan, K., D. P. Leone, R. K. Bateson, G. Dobreva, Y. Kohwi, T. Kohwi-Shigematsu, R. Grosschedl and S. K. McConnell (2012). "A network of genetic repression and derepression specifies projection fates in the developing neocortex." Proc Natl Acad Sci U S A **109**(47): 19071-19078.

Chapter 3. Effective Fluorescent Labeling of Human Embryonic Stem Cells via BAC Mediated Transgenesis

3.1 Introduction

Human embryonic stem cells (hESCs) offer unprecedented potential for science and medicine (Thomson, Itskovitz-Eldor et al. 1998, Cowan, Klimanskaya et al. 2004). The pluripotent nature of these cells makes them ideal tools for developmental research. Because they are non-transformed human cells, they may also be especially valuable for high throughput screens and the development of new pharmaceutical agents (Xu, Shi et al. 2008). Likewise, hESCs appear to be optimal candidates for eventual cell transplant and cell replacement therapies, particularly for tissue types that exhibit poor endogenous regenerative capabilities such as the human central nervous system. Despite their great promise, there remain some obstacles that must be overcome before hESCs can be utilized to their full potential. One obstacle that is frequently encountered is the need for an effective means to label populations of hESCs in order track them and their progeny (Xia, Zhang et al. 2007).

While there exist some hESC lines that have been modified to express persistent fluorescence, restricting research to only these lines would greatly limit the number available options. Another alternative is to use lentivirus vectors to label a desired population of cells. However, genes introduced into hESCs via lentivirus often undergo silencing (Xia, Zhang et al. 2007). Even successfully labeled cells have been known to diminish in or eliminate their marker expression over time, with multiple

passages, or during the course of differentiation (Macarthur, Xue et al. 2012). Additionally, some laboratories may be poorly equipped or reluctant to handle viral vectors.

Bacterial artificial chromosomes (BACs) offer another means to produce clonally stable transgenic cells (Song, Chung et al. 2010). A BAC consists of a large span of DNA from a species of interest, attached to a bacterial backbone, and can be easily replicated in bacteria. BACs are considerably larger than basic plasmids, reaching sizes up to ~200kb or more. This large size means that a BAC may contain a gene of interest as well as all of its endogenous upstream and down stream regulators. Because BACs can be replicated in bacteria, they can also be targeted for modification through highly efficient recombineering in any of the recombineering bacterial strains (Lee, Yu et al. 2001, Swaminathan, Ellis et al. 2001, Warming, Costantino et al. 2005). When transfected into mammalian cell lines, BACs are highly stable. BACs have been used extensively for molecular biology and gene expression research, for example, implementation of BACs was crucial for the success of the human genome project (Osoegawa, Mammoser et al. 2001, Zhang and Wu 2001, Zhao 2001, Waterston, Lander et al. 2002, Li, Jiang et al. 2003). Additionally, the Allen Brain Atlas has made extensive use of genetically labeled BACs in generating their expansive gene expression data (Boguski and Jones 2004, Boizuski 2005).

Here we strive to develop a non-viral, genetic tool, which can be incorporated with relative ease into virtually any stem cell or iPC line resulting in bright, persistent, clonally stable fluorescence. Through the use of recombineering techniques and BAC transgenesis, we have generated two BAC constructs, which utilize the ROSA26 locus to drive fluorescent protein expression in hESCs. Our initial construct utilized no additional promoter or enhancer regions aside from the ROSA26 locus. Shortly after electroporation into the HUES9 hESC line, this BAC effectively labeled hESCs with ubiquitous, clonally stable fluorescence. What's more, fluorescence persisted throughout neuronal differentiation, demonstrating the BAC's effectiveness as a marker in both stem cells and their differentiated progeny. Seeking to further enhance the intensity of fluorescence, we generated an additional BAC, which incorporates a CAG promoter in conjunction with a WPRE enhancer into the ROSA 26 BAC construct. These modifications were successful in significantly enhancing the fluorescence of transfected hESCs. As with the initial construct, fluorescence is passed on effectively from parent to daughter cells. Taken together, these data indicate that the BACs we have generated are indeed a broadly applicable tool for the effective labeling and tracking of embryonic stem cells and their progeny.

3.2 BAC Modification and Electroporation Yields Effective hESC Transgenesis

We desired to create a non-viral genetic tool that could be incorporated into virtually any hESC line, and result in ubiquitous, persistent fluorescence in it's recipients and their progeny. Towards this end, we selected the ROSA26 locus containing murine BAC, RP23-401D9, and constructed a targeting vector to incorporate tdTomato cDNA between the first and second exons of the ROSA26 locus. Incorporation of the targeting vector into the BAC was accomplished via homologous recombination in the recombineering bacterial strain SW106. Homologous recombination was verified by both acquired drug resistance, and by PCR across upstream and downstream homology arms. In this Modified BAC, tdTomato was incorporated with a splice acceptor, leaving its expression to be driven solely by the ROSA 26 locus (Figure 3.2 A, B). As we wished to demonstrate that this BAC could be successfully incorporated into virtually any ESC line, even one that had been previously modified, we chose to transfect our labeled BAC into an already existing Fezf2:EYFP hESC knock-in line (Ruby and Zheng 2009). HESCs were transfected via electroporation. 48 hours after electroporation, transfected cells were selected for via G418/Neomycin selection. After 14 days of selection, up to 46% of surviving colonies were visibly fluorescent. Interestingly, virtually

all colonies exhibited a degree of mosaic expression (Figure 3.2 C-F). Fluorescent colonies were manually isolated and expanded.

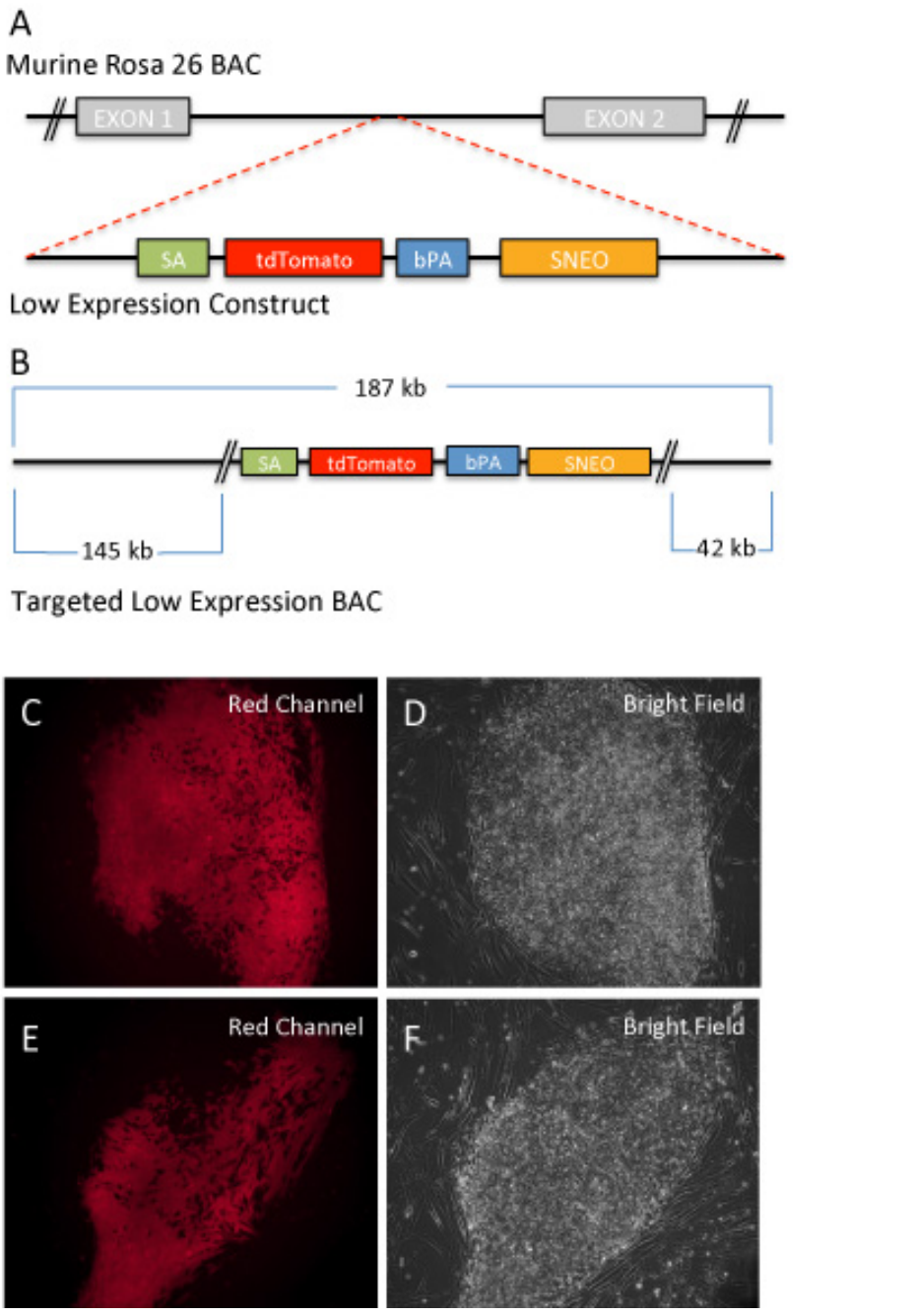

Figure 3.2. BAC modification and electroporation yields effective hESC transgenesis.

(A) Vector targeting construct for BAC recombineering. (B) Successfully modified low expression BAC construct for hESC electroporation. (C-F) TdTomato expressing colonies generated through BAC transgenesis and subsequent neomycin based selection showing both red channel (C & E) and bright field (D & F) images of two colonies.

3.3 Observed Mosaic Expression was Overcome via FACS and Fluorescent Expression was Maintained Across Multiple Passages and Throughout Differentiation.

As stated above, virtually all fluorescent colonies exhibited mosaic tdTomato expression, even after fourteen days of G418 selection. In order to obtain a purified population of fluorescent cells, expanded colonies were sorted via FACS, and fluorescent cells were then re-plated (Figure 3.3 A, B). These purified cells were then passaged every two to four days for five or more passages and expanded to three separate plates, after which they were once again, evaluated via FACS to assess the long term stability of fluorescent expression. FACS analysis demonstrated that even after multiple passaging events (five or more), 93.2 +/- 0.4% of cells were still fluorescent, indicating that once purified, tdTomato expression remained persistent (Figure 3.3 C).

We next sought to determine if the tdTomato labeled stem cells would maintain fluorescence throughout the course of differentiation. Fluorescent cells were differentiated towards a CNS fate through dual SMAD inhibition. After 14 days of differentiation, the cells were manually passaged into spheres. These spheres were permitted to float unattached for four days, before being plated down onto growth factor-reduced matrigel coated plates. These spheres were then permitted to continue differentiating for an additional 5 days before being re-plated at low

density. Three weeks later, cultures were fixed and analyzed via ICC. In order to verify that neurons were indeed present, these cells were stained for tdTomato, Fezf2:EYFP, and Tuj1, cell nuclei were stained with DAPI (Figure 3.3 D-H). Imaging revealed that these differentiated cells were positive for tdTomato, incuding those that were also positive for Tuj1, indicating a postmitotic neuronal identity. Presence of Fezf2:EYFP demonstrates that this cell line's utility as a fluorescent reporter knock-in line remained unhindered by the inclusion of the tdTomato BAC.

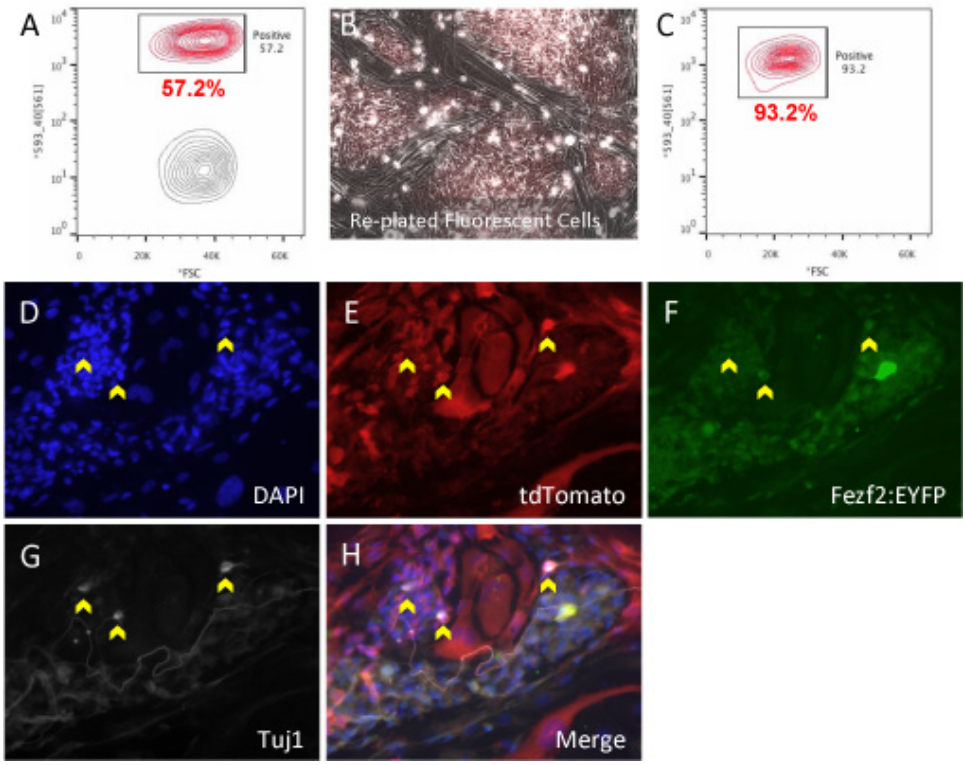

Figure 3.3. Observed mosaic expression was overcome via FACS and fluorescent expression was maintained across multiple passages and throughout differentiation.

(A) Mosaic expressing transgenic colonies were expanded and purified by FACS. (B) Purified tdTomato positive cells were re-plated and grown in culture for a minimum of five passages before being re-evaluated by FACS. (C) Even after multiple passaging events, 93.2 +/- 0.4% of cells remained positive. (D-H) Purified cells then underwent CNS differentiation. Fluorescence was maintained as cells differentiated and was even expressed in post-mitotic neurons (yellow arrows).

3.4 Construct Modification Significantly Strengthened Fluorescent Expression

Following the success of our initial construct, we attempted to dramatically increase the fluorescent intensity of our cells by incorporating a CAG promoter and WPRE stabilizer into our BAC construct. Madisen et. al previously published a CAG-loxP-Stop-loxP- tdTomato-WPRE-

PGK/Neo construct named Ai9, which they had successfully used to generate a Cre inducible tdTomato mouse line (Madisen, Zwingman et al. 2010). We obtained the Ai9 plasmid from AddGene, and transformed it into the SW106 recombineering strain. We then induced cre-recombination to eliminate the stop codon, and replaced the PGK/Neo with SNeo, to allow for effective selection in both prokaryotic and eukaryotic cells (Figure 3.4 A). This new construct was then used to once again target murine BAC RP23-401D9. The targeting vector was incorporated into the BAC via homologous recombination in the recombineering bacterial strain SW106. As before, recombination was verified via acquired selection resistance and PCR across both homology arms.

This new high expressing BAC was then electroporated into the Hues9 Fezf2:EYFP hESC line. Electroporated cells underwent G418 selection for two weeks, after which fluorescent colonies were manually selected and expanded. Once again, Mosaic expression of tdTomato was observed among fluorescent colonies, even after an additional expansion at low density (Figure 3.4 B-D). A comparison between the two BACs demonstrates a significant increase in both visual intensity and gene expression. In cells containing the high tdTomato expressing BAC, a 95ms exposure time was sufficient to achieve a suitable intensity for imaging, while a 1200ms exposure time was required for cells containing the

tdTomato low expressing BAC (Figure 3.4 E-I). QRT-PCR evaluation of two high and two low expressing FACS purified cell lines demonstrated an increase in tdTomato expression of up to eleven fold in the high expressing lines (Figure 3.4 J).

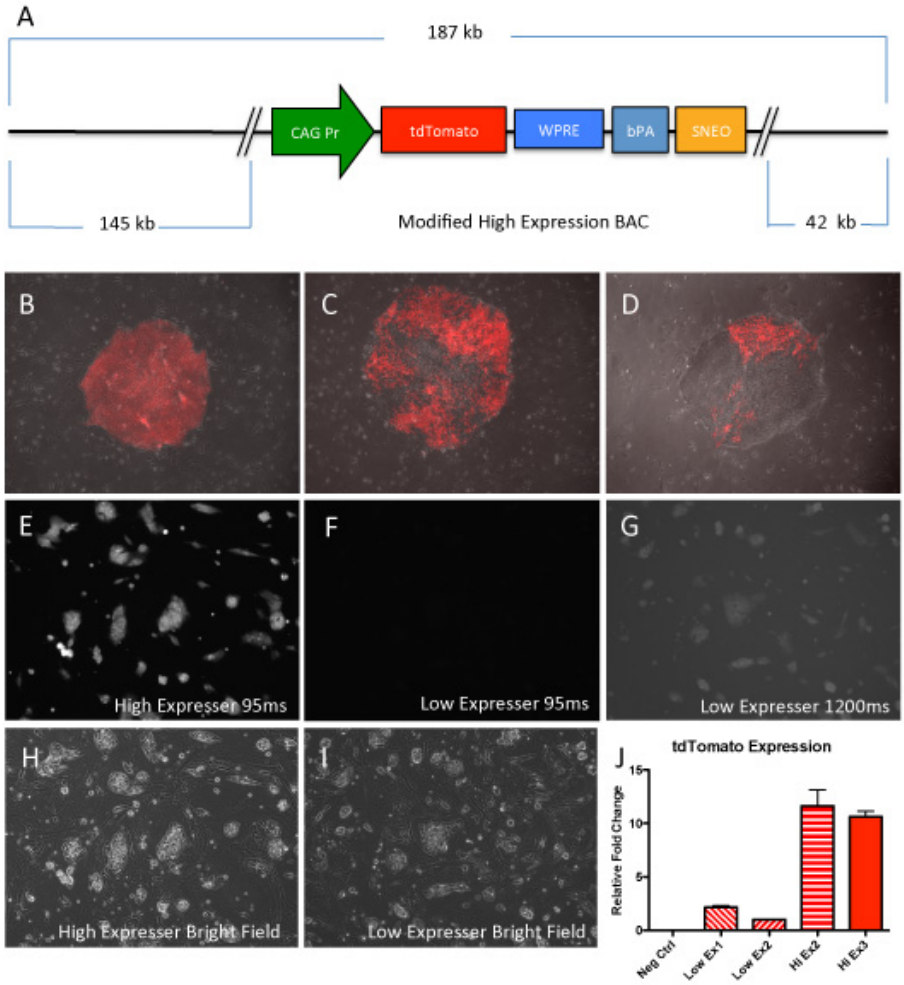

Figure 3.4. Construct modification significantly strengthened fluorescent expression.

(A) A new high expression BAC construct was generated, incorporating a CAG promoter and WPRE stabilizer. (B-D) Representative images of tdTomato expressing colonies generated with the new high expression BAC. (E-G) Transgenic tdTomato high expressing cells are much brighter than the previously generated low expressing cells. (E) High expressing cells are clearly visible with just a 95 ms exposure time, while low expressing transgenic cells are invisible with this brief exposure time (F). An exposure time more than 12 times longer is required to visualize low expressing cells (G). (H &I) Bright-field images of high expressing and low expressing cells respectively. (J) QRT-PCR confirms that tdTomato RNA expression is much higher in cells generated with high expression BAC compared to low expressing cells.

3.5 Mosaic Expression was Overcome via FACS, and Persisted Through Multiple Passages and Throughout Differentiation

As before, virtually all high-expressing colonies exhibited some degree of mosaic expression. Pure fluorescent populations were obtained via FACS. Once again, a FACS purified population of fluorescent cells was expanded, passaged multiple times and was then re-evaluated by FACS. After at least 5 passages, 95.9 +/- 0.1% of cells were still fluorescent, demonstrating stable expression of the fluorescent gene (Figure 3.5 A, B).

Following purification via FACs, a high expressing clone was expanded and submitted to our central nervous system (CNS) differentiation protocol. As anticipated, these high-expressing cells were able to differentiate into various cell types, including post-mitotic neurons. As with the low expressing cells, fluorescence was maintained throughout the differentiation procedure, and was clearly observable in neurons and their projections. Once again the integrity of the Fezf2:EYFP reporter was maintained (Figure 3.5 C-G).

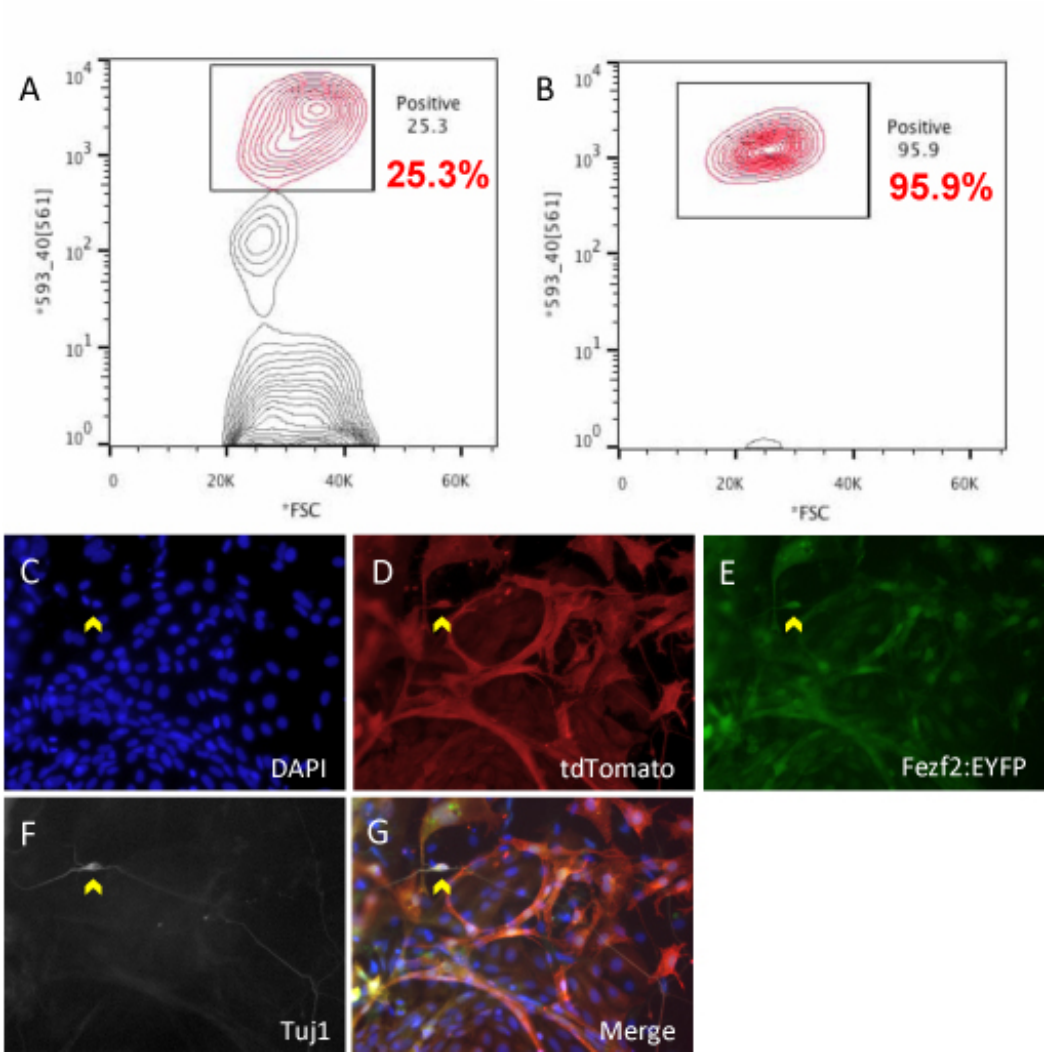

Figure 3.5. Mosaic expression was overcome via FACS, and persisted through multiple passages and throughout differentiation.

(A) Mosaic expressing colonies were purified via FACS. (B) After a minimum of five passages, purified cells were again analyzed by FACS, and were found to be 95.9 +/- 0.1% pure. (C-G) Purified high expressing tdTomato positive cells underwent CNS differentiation. Fluorescence was maintained as cells differentiated and was even expressed in post-mitotic neurons (yellow arrow).

3.6 Causes of Mosiac Expression

We hypothesized that the mosaic expression observed in newly formed colonies was due either to cross-feeding (a phenomenon in which the presence of resistant cells enables the survival of neighboring non-resistant cells), or epigenetic silencing of the fluorescent gene. In order to determine which was occurring, six mosaic colonies (three containing the low-expressing BAC and three containing the high expressing BAC) were expanded and then sorted into fluorescence positive and negative populations via FACS (Figure 3.6.1 A, B). DNA was then extracted from these different populations immediately following the sort. The remaining sorted cells were re-plated into separate individual wells for further evaluation. Despite our precautions (Figure 3.6.1 C) and the relatively high efficiency of FACS, there were still some unavoidable errors in sorting (a small percentage of fluorescent cells was sorted in with the negative population). These miss-sorted cells were visible among re-plated populations (Figure 3.6.2 A,B).

3.6.1 Quantitative PCR indicates that epigenetic silencing does not play a role in mosaic expression.

The DNA collected from each of the purified fluorescence positive and negative populations, along with a non-transfected negative control, was evaluated by qPCR for the presence or absence of tdTomato DNA.

Each of the positive populations was clearly positive for tdTomato. The six negative populations demonstrated a near absence of tdTomato DNA, showing a profile virtually identical to negative control (except for the small amount of tdTomato DNA attributed to the miss-sorted cells) (Figure 3.6.1 D). These results indicate a significant role for cross-feeding, rather than epigenetic silencing, in the survival of the fluorescence negative cells during selection.

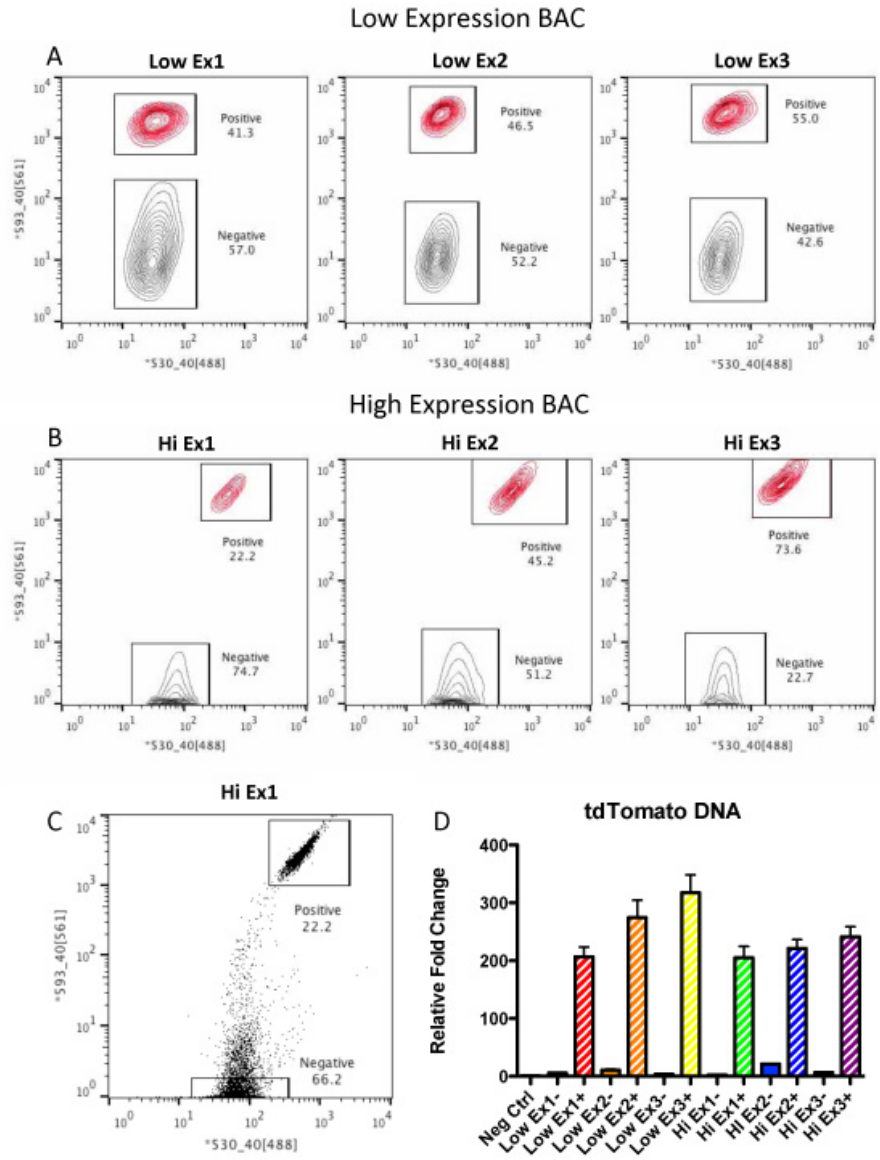

Figure 3.6.1 Quantitative PCR indicates that epigenetic silencing does not play a role in mosaic expression.

Cells from three low expressing (A) and three high expressing (B) mosaic colonies were sorted via FACS, and negative and positive populations were collected, and DNA was collected form each population. (C) Representative FACS dot plot of high expressing cells. In order to obtain as pure a representative population as possible, only clearly positive and clearly negative cells were collected. Intermediate expressing cells were ignored. Even so a small percentage of miss-sorting error persisted. (D) DNA extracted from each of these six positive and six negative populations was evaluated via qPCR. Results demonstrated that positive populations contained tdTomato DNA while negative populations did not.

3.6.2 Additional drug selection supports a role for cross-feeding in mosaic tdTomato expression.

In order to further explore these results, re-plated fluorescence positive and negative cells were expanded, and then submitted to a second round of G418 selection. Re-plated positive and negative cell populations were permitted to expand for four to five days prior to initiation of G418 selection. As stated above, FACS purification is not 100% efficient, and thus, we were not surprised to see a small number of fluorescence positive cells in our purified negative populations (Figure 3.6.2 A, B). G418 selection persisted for ten days, during which time, each population was imaged daily. As anticipated given the above results, the fluorescence-negative cells were largely eliminated by selection, except in the areas immediately surrounding fluorescent cells. Fluorescence-positive cells, however, were unaffected by selection and continued to proliferate, indicating once again, that the mosaic expression observed in initial colonies was largely due to cross-feeding.

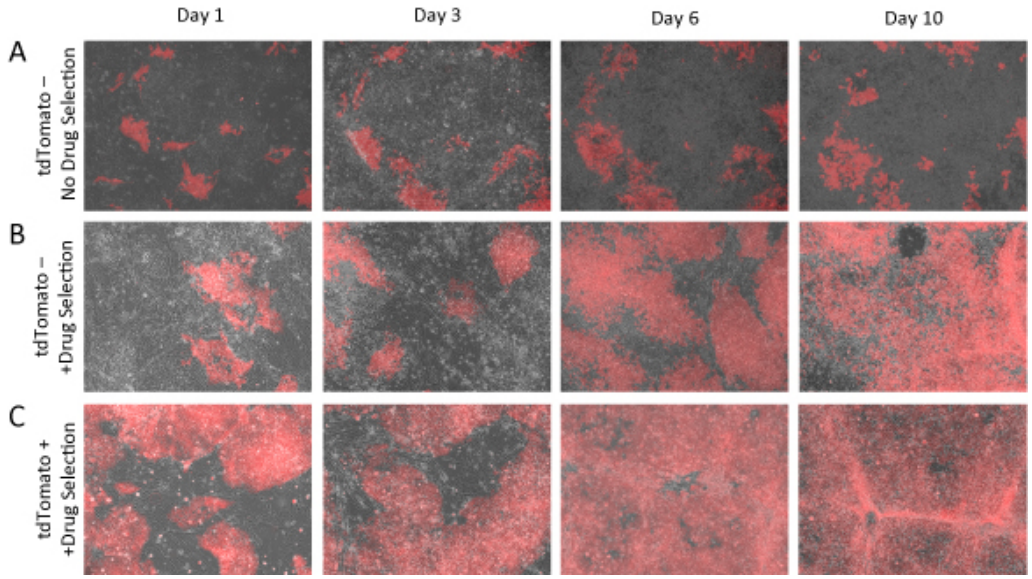

Figure 3.6.2 Additional drug selection supports a role for cross-feeding in mosaic tdTomato expression.

FACS purified tdTomato negative (A & B) and positive (C) cells were re-plated and expanded. When 70 to 90% confluent these cells were monitored for an additional 10 days in either the absence (A) or presence (B & C) of selection drug G418. Both negative populations include a small percentage of miss-sorted tdTomato positive cells. Without further selective pressure, these cells remain a minority (A). In the presence of Selective pressure, tdTomato negative cells quickly die off, except in close proximity to tdTomato positive cells (B). TdTomato positive populations are unaffected by selection (C).

3.7 Materials and Methods

3.7.1 Targeting Vectors and BAC Modification

The low expressing targeting vector is illustrated in Figure 3.2 A. Briefly, it consists of a 500 base pair up stream homology arm, followed by a splice acceptor, tdTomato, bPA, a special neomycin resistance gene (SNeo), and the downstream 500 base pair homology arm. Homology arms were cloned from Genomic DNA using Expand High Fidelity PCR kit (Roche). Sneo, cloned from plasmid PL452, is a Neomycin resistance factor under the control of both a eukaryotic PGK promoter and a prokaryotic EM7 promoter, which allows Neomycin to be used for selection in both mammalian and bacterial cultures.

The high expressing targeting vector illustrated in Figure 3.3 A was generated through modification of the Ai9 plasmid, originally created by Madison et al (Madisen, Zwingman et al. 2010). This plasmid was electroporated into the SW106 bacterial strain using Gene Pulser II (Biorad) at 1.75 kV, 25 μF. The floxed stop was then eliminated through the induction of CRE expression through the introduction of 0.1% L(+)Arabinose into the culture media. SNeo was cloned from plasmid PL452 using an Expand High Fidelity PCR kit (Roche). The PGK/Neo was then excised from Ai9 using restriction enzymes Asc1 and Fse1 (NEB) and SNeo was ligated into its place using Rapid DNA Ligation kit (Roche). This Modified version of Ai9 (MAI9) was then used as a targeting vector.

Murine BAC RP23-401D9 was obtained from Children's Hospital Oakland Research Institute (CHORI) as an LB stab. BAC DNA was isolated using Macherey-Nagel's Nucleobond Xtra Maxi Plus kit. Purified BAC DNA was then electroporated into recombineering strain SW106 using Biorad Gene Pulser II at 1.75 kV, 25 μF. BAC containing bacteria were selected for via chloramphenicol selection (12.5 μg/ml). Incorporation of the complete BAC was verified by PCR. Targeting of the ROSA26 locus within the BAC was carried out within the SW106 bacteria strain as per the manufacturers instructions. Successful homologous recombination was verified by PCR across both homology arms. Targeted BACs were isolated using the Nucleobond Xtra Maxi Plus kit (Macherey-Nagel). Prior to electroporation into stem cells, 100 ug of purified BAC was linearized using enzyme PI-SceI (NEB). Linearized BAC DNA was precipitated, spooled, and dissolved in 250 μl 5%TE overnight at room temperature.

3.7.2 Stem Cell Culture

Hues9 hESCs were cultured on mouse embryonic fibroblasts in HUES media as described previously (Ruby and Zheng 2009), except that cells were supplemented with 20ng/ml bfgf. For at least two passages immediately prior to electroporation, cells were grown under feeder free conditions on matrigel (BD Biosciences), and fed daily with mouse

embryonic fibroblast (MEF) conditioned media (MCM) supplemented with 20ng/ml bfgf.

3.7.3 Electroporation and Selection

When cells reached 70 to 90% confluence, they were dissociated enzymatically into single cells using 0.5% trypsin. Trypsin was quenched with HUES Media and cells were filtered through a 40µm cell strainer (BD Falcon). Cells were then centrifuged at 1000 RPM for 4 minutes and resuspended in 650 µl of HUES Media. Cells were then combined with 100µg of BAC DNA suspended in 250µl 5% TE, mixed by pipetting, and incubated on ice for 5 minutes. Cell/ DNA mixture was then transferred to a cold Gene Pulser IIcuvette (0.4 cm gap) and electroporated with Gene Pulser II (Bio-Rad) at 320V, 200µF. The cuvette was then placed in the cell incubator at 37ºC for 7 minutes to allow cells to recover after which cells were plated onto Neomycin-resistant MEFs. Geneticin G418 Sulfate (G418, Life Technologies) selection commenced 48 hours after electroporation at 50µg/ ml, and continued for 13 days. 14 days after electroporation, individual drug-resistant fluorescent colonies were manually collected and expanded.

3.7.4 Fluorescence Assisted Cell Sorting

Cells were grown under feeder free conditions for at least one passage prior to FACS, on matrigel, and fed daily with MEF conditioned

media supplemented with bfgf. Cells were trypsinized, resuspended and passed through a 40µm cell strainer. Cells were then centrifuged at 1000 RPM for 4 minutes. Cells were resuspended in sort buffer, consisting of Ca++/Mg++ free HBSS (Life Technologies) with 5mM EDTA and 1%FBS, at a concentration of about 6 million cells/ ml, and kept on ice until FACS. Cells were sorted into collection tubes containing collection media comprised of Ca++/Mg++ free HBSS with 20% KSR. FACS was carried out at the UCSD Stem Cell Core using either the FACS Aria II (BD Biosciences) or the Influx (BD Biosciences), depending on availability.

3.7.5 Stem Cell differentiation

Stem Cells were differentiated into neurons using dual SMAD inhibition. Cells were grown in feeder-free conditions on matrigel for one passage prior to induction of differentiation. Cells were then trypsinized, and re-plated on plates coated with growth factor-reduced matrigel (BD Biosciences). Cells were fed with MEF-conditioned media supplemented with 20ng/ml bfgf until 90 to 100% confluent. Cells were then fed with 50% MCM and 50% hESDiff Media (Knockout DMEM with 15% KSR, 1X NEAA and 1x Glutamax, all from Life Technologies) supplemented with 20 ng/ml bfgf and 5µM SB431542 (SB, Stemgent), for one day. Followed by 100% hESDiff media supplemented with 5µM SB431542 and 100nM LDN-193189 (LDN, Stemgent) for 4 days, 75% hESDiff/ 25% N2 Media (DMEM/F12 and 1X N2 supplement, both from Life Technologies-Gibco)

supplemented with 100nM LDN for 2 days, 50% hESDiff/ 50% N2 Media supplemented with 100nM LDN for 2 days, 25% hESDiff/ 75% N2 with 100nM LDN for 2 days, and then 100% N2 supplemented with 100nM LDN for 3 days. On day 14 of differentiation, cells were manually passaged into spheres. Spheres were then allowed to float for 4 days in N2 Media. On day 19 of differentiation, cells were plated down on plates coated with growth factor-reduced matrigel, in N2 Media. Cells were fed with half media changes every other day until day 24. On day 24 of differentiation media was switched to Neural Differentiation Media (NDM) consisting of Neurobasal Media (Life Technologies) with 1X Glutamax (Life Technologies), 0.5X N2 Supplement (Life Technologies), and B27 Supplement without Vitamin A (Life Technologies). After which half media changes were made as needed. Cells were then either fixed directly with 4%PFA or dissociated and re-plated at 100,000 cells /cm2 and then fixed with 4% PFA.

3.7.6 Immunofluorescence

Cells were fixed with 4% PFA. After fixation, the following Primary antibodies were used: tdTomato: rabbit anti dsRed (Clonetech, 632496), Tuj1: mouse anti Tuj1 (Covance, MMS-435P-250), and YFP: chicken anti GFP (Life Technologies, A10262). The following secondary antibodies were employed: 488 goat anti chicken (Life Technologies, A11039), 546

donkey anti rabbit (Life Technologies A10040), and 680 goat anti mouse (Life Technologies, A21058). Cell bodies were stained with DAPI.

3.8 Discussion

HESCs have virtually unlimited potential due to their ability to become virtually any cell type (Thomson, Itskovitz-Eldor et al. 1998, Cowan, Klimanskaya et al. 2004). This makes them a powerful tool for developmental studies and modeling, pharmaceutical discoveries, and potentially, cell replacement therapies. While this branch of research is still in its infancy, already great advances are being made. Initially, the genetic manipulation of hESCs lagged behind the advances being made with murine embryonic stem cells. However with the advent and implementation of relatively new tools such as talens and CRISPR, there will likely be a major increase in the number of genetically modified hESC lines, beyond what is already available. Here we have developed and introduced a non-viral means for incorporating persistent, bright fluorescence into virtually any cell line. This fluorescence persists across multiple passaging events and throughout differentiation, making it an ideal means for labeling and tracking cells and their progeny. Even the weaker of our two fluorescent constructs was bright enough to effectively label both cell bodies and neurites for live imaging in vitro.

While not the primary objective of this research project, we were also able to investigate the phenomenon of cross-feeding during neomycin

selection. We were initially surprised to observe such a large degree of mosaic expression among fluorescent colonies. Essentially every colony, regardless of which construct had been used to generate it, exhibited some degree of mosaic expression. As explained above, it appears likely that cross-feeding led to the survival and expansion of non-transfected cells in the near vicinity of successfully transfected cells. While in this case, the majority of non-fluorescent cells were easily eliminated by FACS, achieving a purified population of transfected cells may prove more difficult with constructs lacking immediate fluorescent expression. This will be important to bear in mind for future transgenesis experiments.

Another topic of potential interest that was unexpectedly touched upon by this project was the possible issue of tdTomato toxicity or interference at high levels of expression in differentiating cells. While the majority of the low-expressing clones we tested differentiated easily into neurons, as anticipated, only two of the four high expressing clones tested was able to differentiate effectively. Whether the failure of the other three clones was due to tdTomato toxicity, disruptive integration of the construct, or some other interference with differentiation pathways remains to be determined. However, as these difficulties were observed more frequently in high expressing cell lines, it appears likely that high expression of tdTomato plays a role in the observed difficulty in differentiating these cells.

Despite these challenges, we were successful in achieving stable, bright ubiquitous fluorescence in a previously modified hESC reporter cell line. Here, it has been demonstrated these cell lines effectively label cells and their differentiated progeny in vitro. These cell lines will likely prove equally easy to track when differentiated and transplanted into in vivo model systems. Additionally, because the BAC construct contains a red fluorescent protein, it can be incorporated into any previously existing YFP or GFP labeled cell line without disrupting imaging of the original marker. What's more, this study demonstrates that BAC based transgenesis is a feasible and stable means of generating new transgenic hESC lines.

Chapter 3, in part is currently being prepared for submission for publication of the material. Smith, Anders; Zapata, Arturo; Zhang, Sherry; Feltham, Casey; Zheng, Binhai. "Efficient Labeling of Human Embryonic Stem Cells via BAC Mediated Transgenesis". The dissertation author was the principal researcher/author of this material.

3.9 References

Boguski, M. S. and A. R. Jones (2004). "Neurogenomics: at the intersection of neurobiology and genome sciences." Nature Neuroscience 7(5): 429-433.

Boizuski, M. S. (2005). "Neurogenomics: at the intersection of neurobiology and genome sciences." Faseb Journal 19(5): A1711-A1711.

Cowan, C. A., I. Klimanskaya, J. McMahon, J. Atienza, J. Witmyer, J. P. Zucker, S. Wang, C. C. Morton, A. P. McMahon, D. Powers and D. A. Melton (2004). "Derivation of embryonic stem-cell lines from human blastocysts." N Engl J Med 350(13): 1353-1356.

Lee, E. C., D. Yu, J. Martinez de Velasco, L. Tessarollo, D. A. Swing, D. L. Court, N. A. Jenkins and N. G. Copeland (2001). "A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA." Genomics 73(1): 56-65.

Li, J., T. Jiang, B. Bejjani, E. Rajcan-Separovic and W. W. Cai (2003). "High-resolution human genome scanning using whole-genome BAC Arrays." Cold Spring Harbor Symposia on Quantitative Biology 68: 323-329.

Macarthur, C. C., H. Xue, D. Van Hoof, P. T. Lieu, M. Dudas, A. Fontes, A. Swistowski, T. Touboul, R. Seerke, L. C. Laurent, J. F. Loring, M. S. German, X. Zeng, M. S. Rao, U. Lakshmipathy, J. D. Chesnut and Y. Liu (2012). "Chromatin insulator elements block transgene silencing in engineered human embryonic stem cell lines at a defined chromosome 13 locus." Stem Cells Dev 21(2): 191-205.

Madisen, L., T. A. Zwingman, S. M. Sunkin, S. W. Oh, H. A. Zariwala, H. Gu, L. L. Ng, R. D. Palmiter, M. J. Hawrylycz, A. R. Jones, E. S. Lein and H. Zeng (2010). "A robust and high-throughput Cre reporting and characterization system for the whole mouse brain." Nat Neurosci 13(1): 133-140.

Osoegawa, K., A. G. Mammoser, C. Y. Wu, E. Frengen, C. J. Zeng, J. J. Catanese and P. J. de Jong (2001). "A bacterial artificial chromosome library for sequencing the complete human genome." Genome Research 11(3): 483-496.

Ruby, K. M. and B. Zheng (2009). "Gene targeting in a HUES line of human embryonic stem cells via electroporation." Stem Cells 27(7): 1496-1506.

Song, H., S. K. Chung and Y. Xu (2010). "Modeling disease in human ESCs using an efficient BAC-based homologous recombination system." Cell Stem Cell 6(1): 80-89.

Swaminathan, S., H. M. Ellis, L. S. Waters, D. Yu, E. C. Lee, D. L. Court and S. K. Sharan (2001). "Rapid engineering of bacterial artificial chromosomes using oligonucleotides." Genesis 29(1): 14-21.

Thomson, J. A., J. Itskovitz-Eldor, S. S. Shapiro, M. A. Waknitz, J. J. Swiergiel, V. S. Marshall and J. M. Jones (1998). "Embryonic stem cell lines derived from human blastocysts." Science 282(5391): 1145-1147.

Warming, S., N. Costantino, D. L. Court, N. A. Jenkins and N. G. Copeland (2005). "Simple and highly efficient BAC recombineering using galK selection." Nucleic Acids Res 33(4): e36.

Waterston, R. H., E. S. Lander and J. E. Sulston (2002). "On the sequencing of the human genome." Proceedings of the National Academy of Sciences of the United States of America 99(6): 3712-3716.

Xia, X., Y. Zhang, C. R. Zieth and S. C. Zhang (2007). "Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner." Stem Cells Dev 16(1): 167-176.

Xu, Y., Y. Shi and S. Ding (2008). "A chemical approach to stem-cell biology and regenerative medicine." Nature 453(7193): 338-344.

Zhang, H. B. and C. C. Wu (2001). "BAC as tools for genome sequencing." Plant Physiology and Biochemistry 39(3-4): 195-209.

Zhao, S. Y. (2001). "A comprehensive BAC resource." Nucleic Acids Research 29(1): 141-143.

Chapter 4. Conclusion and Further Directions

4.1 Conclusion

4.1.1 Fezf2 positive stem/ progenitor cells play a significant role in human cortical development, and the development of CSMNs

Human embryonic stem cells offer enormous potential for the study of human development and regenerative medicine. Due to their ability to self-renew and differentiate into any cell type of the human body, they bring hope for the development new research models and new therapies, particularly for systems that display poor endogenous regenerative capabilities, like the human CNS. The CST, part of the CNS, is one such system. The CST is vital for skilled voluntary movement of the hands, fingers and toes, and is damaged or undergoes degeneration in spinal cord injury, stroke and neurodegenerative diseases, like ALS. The CST is composed of neurons called CSMNs. Fezf2 is a molecular marker and regulator essential for the development of CSMNs during normal corticogenesis in mice, and is expressed in subcerebral projection neurons in humans. In addition to being expressed in post-mitotic CSMNs, Fezf2 is known to be expressed in cortical stem/ progenitor cells in the ventricular and sub-ventricular zones during cortical development in mice. Recently, it has been demonstrated that these Fezf2 positive populations give rise to cortical projection neurons, not just in layer V and VI, but in

every layer of the cortex, and even give rise to glia. However the role of Fezf2 in human corticogenesis remains understood. Here, using an hESC Fezf2:EYFP knock-in reporter line, we investigated the role of Fezf2 in human cortical development, and sought to generate human CSMNs from hESCs. In addition, in this work we described the generation and utilization of two BACs, each employed to permanently label hESCs and their progeny with persistent fluorescent expression.

We have demonstrated, through directed differentiation, that Fezf2 is broadly expressed in human cortical stem/ progenitor cells, evidenced by coexpression of Fezf2 with multiple early neural markers. Further we have demonstrated via immunostain and time-lapse imaging, that these Fezf2 positive cells are mitotic.

Through in vitro lineage tracing, starting at multiple time points, we have demonstrated that Fezf2 positive cells give rise to both Fezf2 positive and Fezf2 negative cell types, including neurons that show characteristics of deep-layer projection neurons like CSMNs.

Mid-throughput gene expression data from Fezf2 positive neural precursors and their progeny have indicated that Fezf2 positive precursors give rise to multiple neuronal subtypes, including deep-layer sun cerebral and thalamic projection neurons, callosal projection neurons, superficial layer neurons and glia. These data indicate that in humans, as in mice,

Fezf2 positive NSCs and NPCs give rise to a broad diversity of cortical progeny.

With regard to selectively gerating CSMNs, careful timing and FACS purification of Fezf2 positive cells during differentiation allowed us to generate a population of post-mitotic neurons that were 94% positive for Fezf2 and exhibited morphological and molecular characteristics of immature CSMNs.

Additionally, many Fezf2 positive cells, when transplanted into neonatal mice, maintained Fezf2 expression, co-expressed the CSMN marker Ctip2, and displayed a pyramidal morphology. Further, transplanted Fezf2 positive neurons were able to send axonal projections through appropriate subcerebral targets, even reaching the pyramidal decussation and the spinal cord- despite the absence of endogenous guidance molecules. Together these data suggest a CSMN identity.

4.1.2 BAC mediated transgenesis effectively labels hESCs and their progeny.

In this work, I also described the generation of two separate BAC constructs, developed by incorporating tdTomato into the murine ROSA26 BAC RP23-401D9. The first of these BACs incorporated tdTomato through the use of a splice acceptor alone, placing transgene expression entirely under the control of the ROSA26 locus. While this BAC construct was effective in labeling hESCs and their progeny, the intensity was relatively

low. In order increase fluorescent intensity, a new BAC construct was generated in which the tdTomato transgene was flanked by an upstream CAG promoter and a downstream WPRE stabilizer. These modifications resulted in an up to eleven-fold increase in tdTomato expression relative to the previous construct, and cells transfected with this new construct were visibly far brighter. As efore, fluorescence in transfected hESCs persisted through multiple passages and throughout differentiation.

4.2 Future Directions

In this work we demonstrated that Fezf2 is broadly expressed in early cortical stem/ progenitor cells, and that these cells give rise to a variety of neuronal subtypes and glia, including likely CSMNs. However we did not demonstrate whether or not Fezf2 is necessary for the generation of all of these neuronal subtypes. In the Fezf2 knockout mouse, there is an absence of layer V subcerebral projection neurons, an overall expansion of layer VI, and callosal abnormality- the upper layers, however, are unperturbed. An interesting follow up would be to generate a Fezf2 knockout line by targeting the remaining Fezf2 allele in the existing Fezf2:YFP knock-in line. This would allow the lab to repeat the assays presented here in the absence of Fezf2. Analysis of these cells and their progeny would yield a better understanding of which neural sub-types are dependent on Fezf2 for their development.

Additionally, putative CSMNs could be used for basic axonal injury studies in vivo. We have already plated Fezf2 positive stem/ progenitor cells in microfluidics chambers. These chambers consist of a cell body chamber and an axonal compartment, connected by narrow grooves passing under a wall. Cells plated in the cell body chamber extend their axons through the grooves into the axonal compartment. Axonal injury can then be incurred by slicing the axons in the axonal chamber. Molecular responses in each individual chamber than can be assessed, and the axons can be monitored to determine their ability to regenerate. Small molecules could likewise be assayed to determine their effect on neural protection or regeneration. Because these neurons are non-transformed human cells, exhibiting characteristics of CSMNs, they may be more directly representative of true CSMNs than mouse or cancer cells lines

One difficulty we encountered with our in vivo transplantation studies was finding an effective way to locate transplanted human axons that had ceased to express Fezf2. Staining for human neurofilament yielded some success, but failed to stain all axons. This became clear when some clearly visible Fezf2 positive axons failed to stain. A clear next step for the tdTomato BAC transfected line would be to differentiate and transplant it in vivo. If successful, it should make identifying both Fezf2 positive and Fezf2 negative axons a simple matter, as even the weak expressor is bright enough to effectively label neurites in vitro.

If these steps can be taken successfully, it will serve to further increase our understanding of human cortical development and the role of Fezf2 in this process. What's more, it will add to the knowledge that will eventually lead to better more effective treatments for sufferers of spinal cord injury, stroke and ALS.