UCLA
UCLA Electronic Theses and Dissertations

Title
Who Am I: Blurring the Lines of Cellular Identity

Permalink
https://escholarship.org/uc/item/5k61n96j

Author
Ho, Chi Kent

Publication Date
2013

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
Los Angeles

Who Am I: Blurring the Lines of Cellular Identity

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

by

Chi Kent Ho

2013

© Copyright by

Chi Kent Ho

2013

ABSTRACT OF THE DISSERTATION

Who Am I: Blurring the Lines of Cellular Identity

by

Chi Kent Ho

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2013

Professor Michael A. Teitell, Co-chair

Professor Yi Sun, Co-chair

Cell identity used to be viewed as something static and each generation of organisms had their primordial germ cells separated early so that the somatic cells are fated to differentiate into the rest of the body. Pluripotency was a one-way street where things went from pluripotent to unipotent. However, today's understanding of cell identity has been massively shifted as the field of reprogramming has matured to a point where changing from one cell type to another is possible.

Attempts at dissecting the process of iPS reprogramming has helped to improve efficiency of reprogramming and lower tumorigenicity, but improving reprogramming speed proves to be a stronger barrier. And yet, many different cell types reprogram into iPS faster than with fibroblasts. By analyzing and comparing the gene expression profiles of two human cell

types that reprogram quickly (Amniotic-Fluid Derived Cells and Adipose Derived Stem Cells) and two human cell types that reprogram slowly (Human Foreskin Fibroblasts and keratinocytes), I found two sets of genes that could explain the differences in reprogramming speeds. PIAS3 and STAT3 seem to be the most promising candidates, but experiments with these two factors are inconclusive.

There has also been major progress in direct reprogramming or transdifferentiation between two lineage-specified cell types. One of the first cell types to be successfully transdifferentiated into in this new revitalization of reprogramming was the induced neurons. However, the neurobiology field is no longer dominated solely by the neuron as astrocytes have become more important in disease pathology as well. Therefore, I use the tried and true reprogramming approach that successfully reprogrammed fibroblasts into iPSCs and iNs to reprogram mouse and human fibroblasts into functional induced astrocytes. Mouse iAs require only the transcription factor, NFIA while human iA reprogramming requires NFIA in addition to TET2 and NOTCH. As a proof-of-principle, I apply the technology to reprogram Alexander's Disease patient-specific fibroblasts to show that the induced astrocytes can be used to create disease models *in vitro*. Finally, I use shRNA knockdown of the mutant protein that causes the mutant phenotype to show that the iA model is tractable to therapeutic testing.

The dissertation of Chi Kent Ho is approved.

Michael F. Carey

Samson Chow

Shuo Lin

Michael A. Teitell, Committee Co-chair

Yi E. Sun, Committee Co-chair

University of California, Los Angeles

2013

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my wife for not killing me for taking so long to get my PhD and for giving me a beautiful daughter that I will cherish every day in addition to reminding her to think carefully before attempting to get a PhD. I would also like to seriously thank my wife for bringing home the bacon because god knows I definitely don't win much of any kind of bread on a graduate student's salary. I do win at life though, since for all my flaws and really annoying habits and quirks, she still finds a way to love me and I love her for that.

I would also like to thank my mom for going through hell and swimming from Guangzhou to Hong Kong so that she can eventually immigrate to the US and give me a great life here. She worked endless hours and multiple jobs in the sweatshops of New York City to give me the best education in the world, so getting this PhD is kind of a culmination of her hopes and dreams for me. Her eternal Tiger Mother sort of disappointment in me has been slightly reduced as a result.

I'm incredibly thankful to my mentor, Dr. Yi Sun, who despite knowing nothing about me, decided to accept me into her lab as a transfer student from the east coast. She has also given me great inspiration with her feats of grant-writing and generally hardcore scientific pursuit. I never see her not excited about science, and that's something I'd love to inherit from her.

To Dr. Jun Kohyama, I am eternally grateful. I wouldn't be able to graduate if it wasn't for his help and unrelenting support in the iA work. His scientific knowledge and generally magical hands is something to be envied. He did trick me into eating some of the most horrible ramen in Japan once, so he's not all good, but definitely a giant among men. THANK YOU JUN!

I'd like to thank the CMB Training grant for providing training and support for me during my early career. They helped me out so much while I was starting out and their financial and scientific support was much appreciated. The invaluable lessons learned while TAing will always stay with me.

I'm also thankful to my many current and former lab mates who respect my tendency to keep to myself and not forcibly talk to me all day, except for Rosemarie who just can't help herself, so I won't blame her. The scientific discussions and journal clubs were a blast and I definitely learned a huge amount from them.

I'd like to thank the Academy. The National Academy of Sciences, that is. Hopefully, I'll be elected into it eventually, but actually, I just wanted to say that phrase because all the cool actors get to say it.

I'd like to thank anyone who actually reads this thesis for reading it, because I worked very hard on it so if no one reads it, well, that just sucks. So thank you. At least read these acknowledgements. I made it short. Thanks.

VITA

2001	B.S. Biomedical Engineering Columbia University New York, NY
2002-2005	Technician Rutgers University New Brunswick, NJ
2006	Transfer to UCLA MBI Program University of California, Los Angeles Los Angeles, CA
2007-2010	CMB Training Grant Teaching Assistant University of California, Los Angeles Los Angeles, CA

PUBLICATIONS

Tsoa RW, Coskun V, Ho CK, de Vellis J, Sun YE. Spatiotemporal differences in the origins of NG2 progenitors dictate their developmental potential in the mammalian forebrain. PNAS (Submitted).

Aramini JM, Sharma S, Huang YJ, Swapna GV, Ho CK, Shetty K, Cunningham K, Ma LC, Zhao L, Owens LA, Jiang M, Xiao R, Liu J, Baran MC, Acton TB, Rost B, Montelione GT. Solution NMR structure of the SOS response protein YnzC from Bacillus subtilis. Proteins. 2008 Jul;72(1):526-30.

Benach J, Wang L, Chen Y, Ho CK, Lee S, Seetharaman J, Xiao R, Acton TB, Montelione GT, Deng H, Sun R, Tong L. Structural and functional studies of the abundant tegument protein ORF52 from murine gammaherpesvirus 68. J Biol Chem. 2007 Oct 26;282(43):31534-41.

Rossi P., Theresa Ramelot S.S., Xiao R., Ho C.K., Ma L.C., Acton T.B., Kennedy M.A., Montelione G.T., 1H, 13C, and 15N resonance assignments for the protein coded by gene locus BB0938 of Bordetella bronchiseptica. J Biomol NMR. 2005 Nov;33(3):197.

Aramani, J.M., Huang, Y.J., Cort, J.R., Goldsmith-Fischman, S., Xiao, R., Shih, L.-Y., Ho, C.K., Liu, J., Rost, B., Honig, B., Kennedy, M.A., Acton, T.B., Montelione, G.T., Solution NMR structure of the 30S ribosomal protein S28E from Pyrococcus horikoshii. Protein Sci. 2003 Dec;12(12):2823-30.

CHAPTER 1

Introduction

1.1 Cell Identity and Reprogramming

Early understanding of cellular identity was very simple by today's standards. Different cells had different functions and their cellular identity was determined by their developmental pathway. Conrad H. Waddington even proposed a now-famous epigenetic landscape model for how cells go from pluripotent stem cells at the top of the landscape to a fully-differentiated somatic cell-type at the bottom of the landscape (Figure 1.1A, Waddington, C. H., 1957). But almost from the start, this model has been challenged time and again by great scientific experiments showing that cellular identity is far more complex and far more flexible than his model predicts. Only a year after Waddington published his model, Gurdon, Elsdale, and Fischberg published a paper in Nature in 1958 describing the successful technique later to be known as Somatic Cell Nuclear Transfer (SCNT) using *Xenopus laevis* oocytes. Their experiment showed that transferring a somatic cell nucleus into an oocyte allowed it to undergo normal development and become a sexually mature and fertile adult (Gurden et al., 1958). This showed that the assumed top to bottom progression of epigenetic development and cell identity presented in the epigenetic landscape was not absolute as a somatic cell-type that was at the "bottom" of the metaphorical hill could be coaxed to move back up to the "top" of the hill using this SCNT technique (Figure 1.1B). Following Gurdon's landmark finding, ten years later in 1968, Ernst Hadorn showed using imaginal discs in *Drosophila melanogaster* could undergo what he called transdetermination or a change in cell fate (Hadorn, 1968). This was yet another blow to Waddington's model showing that not only was vertical movement uphill possible, but lateral movement across the epigenetic landscape from one valley to another valley was also possible (Figure 1.1C).

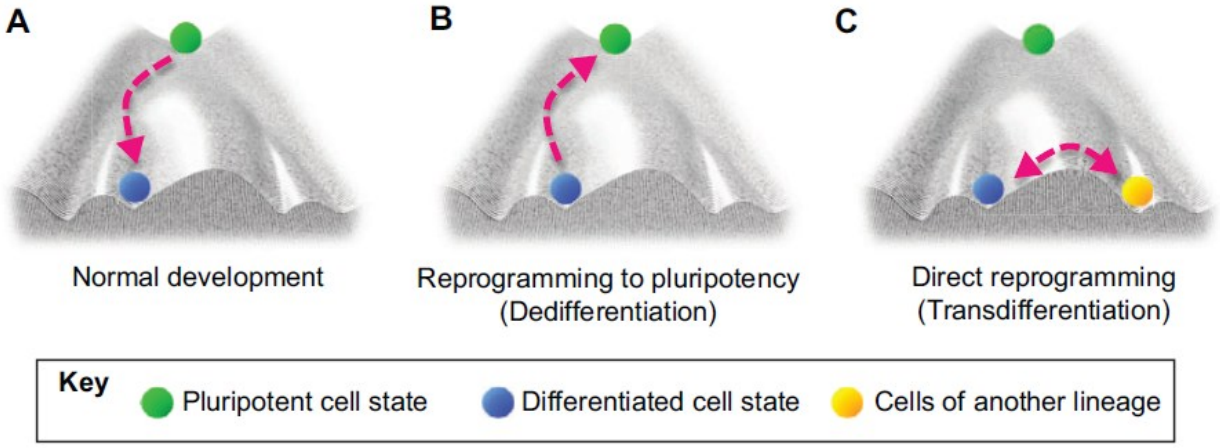

Figure 1.1: Waddington's epigenetic landscape. (A) The original description of the landscape was pretty straightforward. The ball at the top of the hill was considered the pluripotent state and the ball would roll down the valleys until it rested at its final cell state. (B) Gurdon's experiments with *Xenopus* oocytes showed that going back up the hill was possible. (C) Hadorn's experiments with *Drosophila* imaginal discs showed that changing cell fates was also possible. (Figure from Takahashi, 2012)

Over the many years since the key findings by Gurdon that cells could be reprogrammed by nuclear transfer and Hadorn that cell fate is not immutable, many other key experiments paved the way for the science of cellular reprogramming—where a cell of one identity is changed to another. It would take many experiments and many more years, however, with the maturation of the field of epigenetics before cellular reprogramming was achievable between vastly different cell types in a directed fashion. The field would take off once direct reprogramming of cells from a differentiated state to the pluripotent state was achieved by a fairly straightforward protocol. The realization that cellular identity is more tractable than previously imagined has opened up whole new fields of study that have been advancing tremendously ever since Waddington's strict model of cell identity was shown to be inadequate. Therefore, a more modern understanding of the epigenetic landscape consists of the possibility of moving up and down the hill as well as laterally between different valleys and in fact, some cells stay undifferentiated which could be represented as a local valley.

1.1.1 Direct Reprogramming of Cells

Although Gurdon's work with *Xenopus* and Hadorn's work with *Drosophila* showed that cell identity was mutable, their experiments were carried out by transferring either nuclei or cells into different environment where the reprogramming was done for them by a collection of unknown factors. As a result, reprogramming was very limited and the general scientific understanding of cell identity was still very much of the notion that cell types could be changed, but only in a limited number of special cases. One of the most remarkable experiments that showed the malleability of a cell's identity and the importance of epigenetics in determining cell identity was done in 1979 by Taylor and Jones. Using the nucleotide analog, 5-azacytidine (5-

azaC), Taylor and Jones were able to convert a line of mouse C3H10T1/2 embryonic fibroblasts into myogenic, adipogenic, and chondrogenic clonal lines (Taylor and Jones, 1979). DNA methylation changes were assumed to be involved as 5-azaC was the chemical that was able to affect these changes in cell identity. Building upon this finding, in 1987, Davis, Weintraub, and Lassar, in a groundbreaking experiment, showed that a single ectopically expressed transcription factor, Myoblast Determination 1 (MYOD1), was sufficient to convert the C3H10T1/2 fibroblasts into myoblasts (Davis et al., 1987). This finding was significant because it confirmed the ability to achieve directed reprogramming *in vitro* by expression of a single gene and showed that it was possible outside of a "natural" environment. In fact, the strategy and techniques used in this experiment to convert cells from one fate to another would prove to be applicable for other researchers trying to achieve directed cellular reprogramming into different cell types in future experiments. However, it would take the field a long time to catch up with the gravity and universality of their results.

The pioneering work of Davis et al. showed that MYOD1, a single master transcription factor was able to reprogram cells, but it would be over a decade later before more work of this type would be accomplished in blood cells. The field of reprogramming had not been born yet as researchers did not realize that reprogramming a cell from one type to another could be generally achievable given the correct factors. The next major advancement in reprogramming came in the year 2000 when Kondo et al. showed that a lymphoid progenitor could be forced to give rise to myeloid lineage blood cells by expressing interleukin-2 and granulocyte-macrophage colony-stimulating factor receptors (Kondo et al., 2000). This was relatively quickly followed by the reprogramming of astrocytes into neurons via Pax6 (Heins et al., 2002) and the reprogramming of differentiated B cells into macrophages using C/EBPα or C/EBPβ (Xie et al., 2004). At this

point, the burgeoning field of reprogramming was starting to take shape as reprogramming was become more and more commonplace. However, the metaphorical holy grail of reprogramming was still elusive—the reprogramming of somatic cells into pluripotent stem cells using defined factors. Thus far, pluripotent reprogrammed had only been achieved by Gurdon's SCNT and by hybrid cell fusion between somatic cells and either teratocarcinomas (Miller and Ruddle, 1976; or mouse Embryonic Stem (ES) cells (Tada et al., 2001). These feats were accomplished not by defined factors, but by environments similar to the oocyte cytoplasm in Gurdon's experiments.

1.2 Reprogramming into induced Pluripotent Stem Cells.

Though the advancement of reprogramming was happening at a rapid pace in the early 2000s, many researchers thought that reprogramming a somatic cell into a pluripotent stem cell would prove to be much more difficult. Reprogramming from a somatic cell type all the way up the metaphorical epigenetic landscape hill to nearly the most primitive state of pluripotency seemed like an overwhelming feat. Most reprogramming that had been accomplished up to that point was mostly done between two somatic cells derived from same germ layers. That is, except for one extremely important exception. One of the first experiments to show that reprogramming was even possible was done by Gurdon using *Xenopus* oocytes reprogramming the nucleus of a somatic cell. By the mid-2000s somatic cell nuclei from multiple species, including mammals like Dolly the cloned sheep (Wilmut et al., 2007), was shown to be reprogrammable into pluripotent stem cells by their respective species' oocytes. But the cytoplasm of an oocyte was so complex and uncharacterized and containing so many different factors that no one knew where to start. It seemed like direct reprogramming of a somatic cell to a pluripotent stem cell using defined factors was an extremely complicated pipe dream.

Despite these assumed challenges, the seemingly insurmountable was done in 2006 by Takahashi and Yamanaka by reprogramming mouse embryonic fibroblasts into what they dubbed mouse induced Pluripotent Stem cells or iPS cells (Takahashi and Yamanaka, 2006). The seemingly impossible was achieved using a relatively simple concept. The Yamanaka group created a reporter mouse fibroblast cell line that expressed βgeo (a fusion protein of β-galactosidase and the neomycin resistance gene) under the promoter of the F-Box Protein 15 (FBX15) gene, a marker of pluripotency (Figure 1.2A). Once a cell was reprogrammed into the pluripotent state, it would turn on the FBX15 gene and thus the βgeo cassette, which would confer G418 resistance. Using herculean effort, the group picked 24 transcription factors that were possibly important for pluripotency and thus potentially able to reprogram the fibroblasts and cloned each of them individually into separate retroviral vectors. With the cocktail of 24 transcription factors, they transduced them all retrovirally into fibroblasts and then cultured in mouse ES cell conditions with addition of G418. Astonishingly, the experiment worked and they saw the growth of what looked like ES cell colonies that were resistant to G418. Further experimentation with no less effort showed that only four factors were required for reprogramming into pluripotency (the so-called Yamanaka factors): Pou5f1 (Oct4), Sox2, Klf4, and c-myc (Figure 1.2B). Though this initial report of pluripotency had a few caveats (there was no germline transmission and tetraploid complementation was not achieved), it was the first time that somatic cells were reprogrammed to anything remotely resembling the pluripotent stem cell state using ectopically expressed transcription factors.

The work was quickly noticed by many researchers and a slew of papers quickly followed. These groups showed that by using better selection criteria, both germ-line transmission and tetraploid complementation using mouse iPS cells were possible using the four

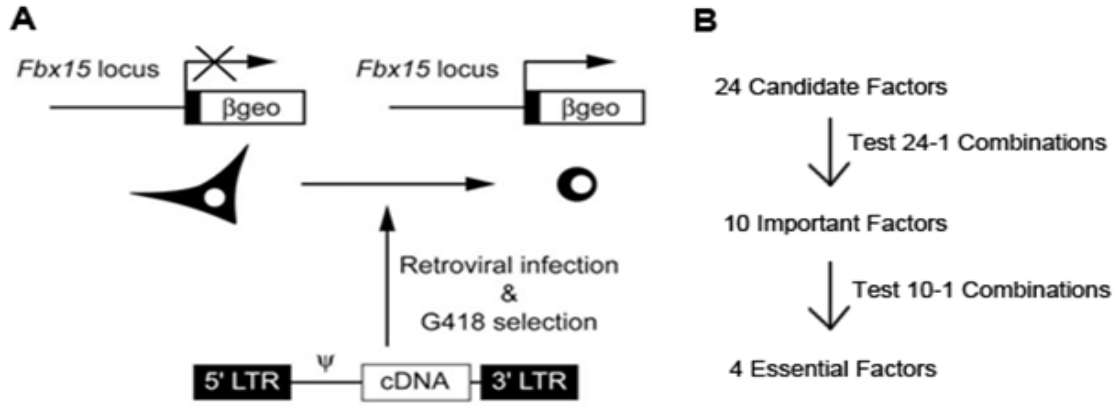

Figure 1.2 Yamanaka's strategy to create iPS. (A) They created a fibroblast line that had the βgeo gene driven by the Fbx15 promoter, such that upon successful reprogramming, the cells would become resistant to G418. These fibroblasts were then infected by a cocktail of retroviral particles that contained a combination of transcription factors and grown in mouse ES cell conditions with G418 selection. (B) They first tested 24 candidate factors and when that worked, they tested every combination of 23 factors to see if eliminating any 1 would abrogate reprogramming. They found 10 important ones and then tried each combination of 9 to see which factors were essential and found the 4 so-called Yamanaka Factors, of which 1 was later shown to be actually non-essential. (Figure adapted from Takahashi and Yamanaka, 2006)

Yamanaka factors (Wernig et al., 2007; Maherali et al., 2007; Meissner et al., 2007). Therefore, iPSCs were bona fide mouse ES cells, which are able to grow into a whole adult organism. The advent of iPSCs was the biggest breakthrough in reprogramming to the pluripotent state since Gurdon's SCNT experiments. And only a year later, the Yamanaka group published a paper showing that what was possible in mouse fibroblasts was also possible in human fibroblasts using the same 4 factors (Takahashi et al., 2007). It was also shown later that c-myc was actually very dispensable as an overexpressed factor in the reprogramming process (Nakagawa et al., 2008). These human iPS cells were also able to give rise to cells of all three germ layers in the teratoma assay, the only test of pluripotency available to human pluripotent stem cells. The idea that a somatic cell can be reprogrammed all the way back to the pluripotent state using only three factors and this was accomplishable even in human cells pushed the reprogramming field into a frenzy as researchers tried to reprogram other cell types into iPS cells.

1.2.1 iPS from Different Parental Cell Types

Although the original iPS generation was done using fibroblasts, that does not preclude other cell types from being able to be reprogrammed into iPS cells. Because iPS cells have been shown to have some epigenetic memory of the parental cell type (Chin et al., 2009), it may be preferred to choose a cell type more closely related to the eventual target cell type that the iPS cells will be differentiated into in order to minimize any possible side-effects that this property of iPS cells causes. However, different parental cell types could possibly require different factors to reprogram and investigators quickly attempted the generation of iPS from different parental cell types. Interestingly, the four Yamanaka factors have been shown to be quite robust in their ability to reprogram many other cell types besides the original fibroblasts that Yamanaka

originally used to generate iPS cells. The four factors or a subset of the four factors were able to reprogram cells from all germ layers including endoderm (Aoi et al., 2008; Stadtfeld et al., 2008), ectoderm (Aasen et al., 2008; Kim et al., 2009), and of course, mesoderm (Takahashi et al., 2006; Ye et al., 2009; Sun et al., 2009). Therefore, the four factors seem to be almost universal in their ability to reprogram into iPS.

Though the four Yamanaka factors were sufficient for many cell types, there was one cell type that required one additional factor before it was able to be reprogrammed into iPS. Terminally differentiated mouse B lymphocytes were unable to be reprogrammed with only the four Yamanaka factors, but rather required the addition of CCAAT/Enhancer Binding Protein, alpha (C/EBPα) or the knockdown of Paired Box 5 (Pax5) in order to reprogram the cells into macrophage-like cells before applying the four factors to push them all the way to the pluripotent state (Hanna et al., 2008). In other words, reprogramming of mature B lymphocytes required a step of dedifferentiation to a more immature state before they were able to be fully reprogrammed by the four Yamanaka factors. An added benefit of this experiment was that it showed unequivocally that the iPS reprogramming occurred in a terminally differentiated cell type rather than a rare population of stem cells since the resultant iPS showed the V-DJ recombination within their genome, an irreversible hallmark of terminally differentiated B lymphocytes.

One important aspect of reprogramming different parental cell lines was that different parental cell types reprogrammed at different rates and different efficiencies. Some cell types (like pancreatic β cells) reprogrammed similarly to fibroblasts at a slow rate in approximately 4 weeks and highly inefficient (less than 0.2% successfully reprogrammed cells) while others (like

Amniotic-Fluid Derived Cells) reprogrammed relatively quickly in about a week with relatively high efficiencies (~1%) (Li et al., 2009). There is no generally accepted molecular mechanism for the differences in reprogramming rates and efficiencies but this is not a well-studied subject area.

1.2.2 Additional Factors in iPS generation

Most of the different parental cells used for reprogramming showed that the four Yamanaka factors were generally useable in a variety of situations, but most of the different parent cells still suffered from one common problem: reprogramming was still relatively slow and extremely inefficient. On the other hand, reprogramming using cell fusion of human and mouse ES cells are quick (taking approximately 2 days) and efficient (with about 70% efficiency) (Tada et al., 2001). As a result, many researchers set out to find other factors or drugs that would be useful in enhancing reprogramming or increasing the speed of reprogramming to the pluripotent state. Much progress has been made in this aspect as many factors have been described in the literature on factors that can enhance or replace factors in the reprogramming process.

Of note, there were some factors that were part of Yamanaka's original screen, but were dropped by Yamanaka in lieu of the four factors he used in his final publication of mouse iPS. These include Nanog (Yu et al., 2007), Utf1 (Zhao et al., 2008), and Sall4 (Tsubooka et al., 2009; Plews et al., 2010). Another significant portion includes microRNAs, which were able to enhance speed and efficiency as well as replace factors (Subramanyam et al., 2011; Liao et al., 2011). Amazingly, expression of the mir302/367 cluster of microRNAs was able to reprogram without using any of the original four factors (Anokye-Danso et al., 2011)! However, because

microRNAs target so many different possible mRNAs, it is hard to tell the underlying mechanisms of how microRNAs affect reprogramming. Curiously, almost all enhancement factors mainly increased efficiency rather than speed, with the noted exception of the miRNAs. Speed seems to be a limiting factor in reprogramming except when using a select number of cell types other than fibroblasts.

1.3 Direct Reprogramming Rejuvenated

Although the creation of iPS cells created a whole new avenue of studying diseased cells, reprogramming technology is not limited to pluripotent stem cells only. One of the original experiments to show the possibility of direct reprogramming was Davis et al's discovery that fibroblasts could be directly reprogrammed into myoblasts using MYOD1. Yamanaka's discovery of iPS cells did not detract from the field of direct reprogramming, but instead, actually intensified research into reprogramming. Before the arrival of iPS to the reprogramming field, researchers were trying to find single master transcription factors similar to MYOD1 for myoblasts and little progress was made using that approach. Yamanaka's important contribution to the field was that he showed a relatively simple method of discovering the necessary factors in reprogramming one cell type to another and that a combination of relatively few factors rather than a single master transcription factor was what was needed to reprogram. His method could be condensed into the following protocol:

1. Infect fibroblasts (or another cell type) with a pool of retroviruses that code for a cocktail of different transcription factors that are highly expressed within or required for the maintenance of the target cell type's identity. Most researchers start with over 10 retroviruses.

2. Culture the transduced cells in media that allows for growth or maintenance of the target cell type. Selection would enhance success rate further.

3. Wait for a reporter to express or stain the mixture of cells after intervals of time to monitor for successful reprogramming events.

4. Systematically eliminate extraneous or counterproductive transcription factors one by one.

5. Repeat step 1 with all combinations of the cocktail with 1 less transcription factor than before and proceed through steps 2-4 as normal. Repeat protocol until the only minimum number of transcription factors necessary for reprogramming is left.

This protocol has proven to be highly effective for many target cell types and many successful applications of the protocol have since been reported. In essence, Yamanaka's work built on the foundation established by his predecessors, who showed such a process was even possible and that transcription factors were the key. Yamanaka simply established a protocol to streamline the process such that new discoveries can be made with limited epigenetic knowledge of the cell type in question.

Using Yamanaka's approach, several groups have shown that direct reprogramming was achievable using fibroblasts as the initial cell type and attempting reprogramming with a cocktail of many transcription factors before narrowing the critical transcription factors down to a handful. This approach was used to produce induced neurons (Vierbuchen et al., 2010), induced cardiac myocytes (Ieda et al., 2010), and induced hematopoietic progenitor cells (Szabo et al.,

2010). While in 2011, this strategy was applied to produce hepatic (Huang et al., 2011; Sekiya and Suzuki, 2011) and chondrogenic (Hiramatsu et al., 2011) lineages from fibroblasts and even neural subtypes like dopaminergic (Kim et al., 2011) and spinal motor neurons (Son et al., 2011). However, there are many more disease-relevant cell types that have yet to be reprogrammed into, and there is still much to learn about the process of direct reprogramming.

1.4 Reprogramming as a Method of Disease Modeling

The discovery of the ability to make pluripotent stem cells without the use of oocytes or pre-implantation embryos was a major boon to the field of stem cell biology. The ethical quandaries and complexities that embryonic stem cell technology posed to researchers were avoided when dealing with iPS cells, which were directly reprogrammed from somatic cells using transcription factors. Since the somatic cells and the reprogrammed iPS cells had the same genome and it was a matter of a change in epigenetics, an obvious application of iPS technology was to model complex disease so that researchers had a tool to test therapeutic strategies or discover mechanisms of the disease in a dish. This technology was especially useful for diseases that involve cell types that were hard to attain in large quantities like neurons and cardiomyocytes. By taking patient-specific somatic cells which were already available in the clinical setting or cell banks of diseased patients, researchers could reprogram them to a pluripotent stem cell state that has the potential to give rise to all cell types of the human body including the cell type that presents the pathological phenotype of the disease. The iPS could also self-renew, meaning even if there was a limited supply of the donor disease cell, reprogramming into iPS gave scientists a near endless supply of the cells containing the mutated genome.

Work on this field exploded after Yamanaka's seminal discovery and iPS cells were created and differentiated into the target cell types for multiple diseases including neurodegenerative disorders (like Parkinson's, Huntington's, and Alzheimer's), cardiac diseases (like CPVT, Dilated cardiomyopathy), metabolic diseases (like Type-1 Diabetes and Hurler syndrome), and developmental diseases (like Hutchinson-Gilford progeria and Rett's Syndrome) as well as many other types (Reviewed by Trounson et al., 2012). Many of these disease models, especially the ones resulting from a congenital mutation, resulted in a reported phenotype and some even showed new or verified existing therapeutic approaches. On the other hand, complex multifactorial diseases like immune system diseases were less successful in the approach. Therefore, this approach to studying a disease has its limitations and is not a magic bullet that can be used for every type of disease.

1.4.1 Reprogramming into iPS versus Direct Reprogramming in Disease Modeling

With the ability to change from one cell type to another as well as from a differentiated state to a pluripotent state and then back down to a third cell type, an obvious use of reprogramming technology was to model diseases in specific cell types that are affected by the disease, which is now an area of intense interest. Disease modeling with afflicted cell types would allow researchers to test hypotheses of disease mechanism and possible treatments. The field is still developing at a rapid pace and there are still many aspects of reprogramming that are not well understood. Since disease modeling is so powerful and achievable through reprogramming to the pluripotent state, researchers scrambled to make patient-specific iPS of many diseases. However, direct reprogramming offers a different avenue with its own advantages and disadvantages when modeling diseases compared to using the iPS route. Neither

method is perfect, but to understand a disease more fully, both methods can be complementary and invaluable.

Reprogramming into iPS offers a few very critical advantages when compared to direct reprogramming. Primary amongst these is a virtually unlimited supply of self-renewing pluripotent stem cells that can differentiate to all cell types of the human body. Another major advantage for iPS is that developmental disorders could be more adequately studied when differentiation of cells from a progenitor population could be part of the pathology of the disease. Finally, there is a wealth of knowledge in terms of differentiation protocols and disease models already established from ESCs from rodent and primate sources. Even though these advantages are significant, these advantages are also tempered with some problems. For example, the cultured iPS over many passages tend to accumulate genetic abnormalities and differentiation protocols tend to produce heterogeneous populations since they are coming from a pluripotent source and purification of the cell type of interest away from other progenies is often required. The iPS lines generated from reprogramming are also somewhat heterogeneous and no two lines are the same so reproducibility can be an issue.

On the other hand, direct reprogramming also offers its own advantages when compared to reprogramming into iPS. The most important advantage is speed. The main problem with reprogramming first to the iPS cell stage and then differentiate to the cell type of interest is that it is a two-step process. Furthermore, both steps are very slow and usually inefficient. Generally, generating, purifying clonal populations, and verifying iPS lines takes months of time and then depending on the target cell type, differentiation can take a few months' time as well. In contrast, many of the direct reprogramming protocols take a week. Another important advantage is that

direct reprogramming generally does not require oncogenic transcription factors like c-Myc or Klf4 and furthermore, since directly reprogrammed cells do not reach the pluripotent stage, there is little risk of teratoma formation or other forms of tumorigenesis. Finally, reprogramming into iPS requires the parental cells be able to proliferate as cellular replication is a requirement of iPS reprogramming. In contrast, direct reprogramming does not require cell division and can be done on terminally differentiated cell types. Direct reprogramming is not without its flaws either. The field is still in its infancy stages as direct reprogramming stagnated for a long period of time until revitalization from Yamanaka's work. As a result, direct reprogramming is less well-studied compared to ES cells and differentiation, which is well established since the discovery of mouse ES cells. Since iPS cells are similar to ES cells, a lot of the work done in the latter can be applied to the former. Also, a rather laborious task necessary to achieve direct reprogramming is to discover and clone a large set of possibly relevant transcription factors that is required to cause the transdifferentiation. The advantage of iPS is that the four Yamanaka factors are nearly universal and can be used on multiple cell types to reprogram to the iPS cell state. The starting cell type may be a factor to consider when attempting direct reprogramming. Fortunately, fibroblasts have proven to be quite flexible in this regard as they seem to be a good starting material for all types of reprogramming.

Introduction to the Thesis

Reprogramming is still an area of intense study and the field is moving quickly in making new discoveries. Since Yamanaka's lab showed that reprogramming with a cocktail of different transcription factors could change cellular identity even all the way up the epigenetic landscape to the pluripotent state, researchers have been trying to understand this process of iPS reprogramming and a lot of progress has been made in understanding the mechanics and the kinetics of reprogramming iPS. Despite these new insights, the process is still a great mystery and inefficient and slow, especially in fibroblasts. Understanding this process better and finding out how to affect this process will give further insights into cellular identity as well as how to manipulate and improve reprogramming as a whole, not just for iPS reprogramming.

The interesting aspect of iPS reprogramming is that so far, the discovered additional factors that help reprogram has been only able to increase efficiency of reprogramming such that more cells eventually become reprogrammed or the successful clones are of higher quality, but the time it takes has not been improved. The only time faster iPS reprogramming occurs is when using a different cell type to start. In this case, there are many cell types that reprogram faster and more efficiently than fibroblasts, which was the cell type originally used by Yamanaka and Lasser in their work. However, why do different cell types reprogram at different speeds. Was it simply that they started at a higher place in the epigenetic landscape such that they were closer to the pluripotent state? Data from reprogramming neural progenitor cells and hematopoietic stem cells seem to run counter to that notion as these stem cells reprogram at a similar speed to fibroblasts. On the other hand, more differentiated cells like hepatocytes have been reported to be more quickly reprogrammable. Therefore, a more likely scenario is that some cell types

express some critical factors that might speed up or help the reprogramming process such that not only efficiency, but the time needed to reprogram is affected. If this scenario is true, then it would be prudent to compare the transcriptomes of different cell types that reprogram at different speeds to see if the faster reprogramming cell types share any transcription factors or chromatin remodeling proteins that may give them the edge over their slower counterparts.

In this work, I compare 4 different parental cell types and how they reprogram compared to what was reported in the literature. Then I compare the transcriptomes of the four parental cell types to see if any expressed proteins would be able to help account for the speed differences. Using advanced expression analysis known as Weighted Gene Co-expression Network Analysis (WGCNA), I find that PIAS3 expression seems to correspond with faster reprogramming while the target of PIAS3, STAT3 is implicated in cells slower reprogramming. However, tests with these two factors seem to show that they alone are not the determinants of the speed of reprogramming.

Another area of concentrated research is in direct reprogramming of cells to different cell types. Direct transdifferentiation is extremely useful in studying and modeling diseases, but most attempts reported so far have been concentrated into reprogramming into a select few cell types, notably neurons and different subtypes of neurons. This is a result of the fact that brain tissue is difficult to obtain and also because neurological diseases are still not well understood and a major public health problem. Astrocytes are neural cells that are emerging as a major player in the pathology of some neurological diseases and are no longer considered to be simply support cells for neurons. Despite this fact, no literature on reprogramming into astrocytes has yet been published. Moreover, since there is a lack of this technology, no disease models have

been created by reprogramming patient-specific cells into induced astrocytes. The ability to directly reprogram into astrocytes would allow scientists and clinicians to better understand the role astrocytes play in diseases and how better to treat those diseases.

In this work, I use Yamanaka's approach to reprogram fibroblasts into induced astrocytes (iAs) first in mouse. I then show that they express astrocyte markers and can functionally support neurons *in vitro*. I also show that human fibroblasts can also be reprogrammed into induced astrocytes and use the technology to reprogram patient-specific fibroblasts from the human astrocyte disease—Alexander's disease. These human diseased iAs show phenotypes similar to those reported in mouse models of the disease and shRNA knockdown is able to ameliorate the disease phenotype to some extent.

REFERENCES

Aasen T, Raya A, Barrero MJ, Garreta E, Consiglio A, Gonzalez F, Vassena R, Bilić J, Pekarik V, Tiscornia G, Edel M, Boué S, Izpisúa Belmonte JC (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol.* **26**: 1276-1284.

Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, Zhang Y, Yang W, Gruber PJ, Epstein JA, Morrisey EE (2011). Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* **8**: 376-388.

Aoi T, Yae K, Nakagawa M, Ichisaka T, Okita K, Takahashi K, Chiba T, Yamanaka S (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* **321**: 699-702.

Chin MH, Mason MJ, Xie W, Volinia S, Singer M, Peterson C, Ambartsumyan G, Aimiuwu O, Richter L, Zhang J, Khvorostov I, Ott V, Grunstein M, Lavon N, Benvenisty N, Croce CM, Clark AT, Baxter T, Pyle AD, Teitell MA, Pelegrini M, Plath K, Lowry WE (2009). Induced pluripotent stem cells andembryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* **5**: 111–123

Davis RL, Weintraub H, Lassar AB (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**: 987-1000.

Gurdon JB, Elsdale TR, Fischberg M (1958). Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* **182**: 64-65.

Hadorn E (1968). Transdetermination in cells. *Sci. Am.* **219**: 110-120.

Hanna J, Markoulaki S, Schorderet P, Carey BW, Beard C, Wernig M, Creyghton MP, Steine EJ, Cassady JP, Foreman R, Lengner CJ, Dausman JA, Jaenisch R (2008). Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* **133**: 250-264.

Heins N, Malatesta P, Cecconi F, Nakafuku M, Tucker KL, Hack MA, Chapouton P, Barde YA, Gotz M (2002). Glial cells generate neurons: the role of the transcription factor Pax6. *Nat. Neurosci.* **5**: 308-315.

Hiramatsu K, Sasagawa S, Outani H, Nakagawa K, Yoshikawa H, Tsumaki N (2011). Generation of hyaline cartilaginous tissue from mouse adult dermal fibroblast culture by defined factors. *J. Clin. Invest.* **121**: 640-657.

Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, Hu Y, Wang X, Hui L (2011). Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* **475**: 386-389.

Ieda M, Fu JD, Delgado-Olguin P, Vedantham V, Hayashi Y, Bruneau BG, Srivastava D (2010). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**: 375-386.

Li C, Zhou J, Shi G, Ma Y, Yang Y, Gu J, Yu H, Jin S, Wei Z, Chen F, Jin Y (2009). Pluripotency can be rapidly and efficiently induced in human amniotic fluid-derived cells. *Hum. Mol. Genet.* **18**: 4340-4349.

Liu C, Yu C, Yuan K, Lou J, Ding M, Deng H (2008). Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell* **3**: 475-479.

Liao B, Bao X, Liu L, Feng S, Zovoilis A, Liu W, Xue Y, Cai J, Guo X, Qin B, Zhang R, Wu J, Lai L, Teng M, Niu L, Zhang B, Esteban MA, Pei D (2011). MicroRNA cluster 302-367 enhances somatic cell reprogramming by accelerating a mesenchymal-to-epithelial transition. *J Biol Chem* **286**:17359-173564.

Kim JB, Sebastiano V, Wu G, Araúzo-Bravo MJ, Sasse P, Gentile L, Ko K, Ruau D, Ehrich M, van den Boom D, Meyer J, Hübner K, Bernemann C, Ortmeier C, Zenke M, Fleischmann BK, Zaehres H, Schöler HR (2009). Oct4-induced pluripotency in adult neural stem cells. *Cell* **136**:411-419.

Kim J, Su SC, Wang H, Cheng AW, Cassady JP, Lodato MA, Lengner CJ, Chung CY, Dawlaty MM, Tsai LH, Jaenisch R (2011). Functional integration of dopaminergic neurons directly converted from mouse fibroblasts. *Cell Stem Cell* **9**: 413-419.

Kondo M, Scherer DC, Miyamoto T, King AG, Akashi K, Sugamura K, Weissman, I L (2000). Cell-fate conversion of lymphoid-committed progenitors by instructive actions of cytokines. *Nature* **407**: 383-386.

Maherali N, Sridharan R, Xie W, Utikal J, Eminli S, Arnold K, Stadtfeld M, Yachechko R, Tchieu J, Jaenisch R, Plath K, Hochedlinger K (2007) Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell* **1**: 55-70

Meissner A, Wernig M, Jaenisch R (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nat Biotechnol* **25**: 1177-1181.

Miller RA, Ruddle FH (1976). Pluripotent teratocarcinoma-thymus somatic cell hybrids. *Cell* **9**:45-55.

Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, Yamanaka S (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* **26**: 101-106.

Plews JR, Li J, Jones M, Moore HD, Mason C, Andrews PW, Na J (2010). Activation of pluripotency genes in human fibroblast cells by a novel mRNA based approach. *PLoS One*. **5**: e14397.

Sekiya S, Suzuki A (2011). Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* **475**: 390-393.

Son EY, Ichida JK, Wainger BJ, Toma JS, Rafuse VF, Woolf CJ, Eggan K (2011). Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell* **9**: 205-218.

Stadtfeld M, Brennand K, Hochedlinger K (2008). Reprogramming of pancreatic beta cells into induced pluripotent stem cells. *Curr Biol.* **18**:890-894.

Subramanyam D, Lamouille S, Judson RL, Liu JY, Bucay N, Derynck R, Blelloch R (2011). Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat Biotechnol* **29**: 443-8.

Sun N, Panetta NJ, Gupta DM, Wilson KD, Lee A, Jia F, Hu S, Cherry AM, Robbins RC, Longaker MT, Wu JC (2009). Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **106**: 15720-15725.

Szabo E, Rampalli S, Risueño RM, Schnerch A, Mitchell R, Fiebig-Comyn A, Levadoux-Martin M, Bhatia M (2010). Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* **468**: 521-526.

Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr. Biol.* **11**: 1553-1558.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**: 861-872.

Takahashi K (2012). Cellular reprogramming – lowering gravity on Waddington's epigenetic landscape. *J. of Cell Sci.* **125**: 2553-2560.

Takahashi K, Yamanaka S (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663-676.

Taylor SM, Jones PA (1979). Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. *Cell* **17**: 771-779.

Trounson A, Shepard KA, DeWitt ND (2012). Human disease modeling with induced pluripotent stem cells. *Curr. Opin. Genet. Dev.* **22**: 509-516.

Tsubooka N, Ichisaka T, Okita K, Takahashi K, Nakagawa M, Yamanaka S (2009). Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. *Genes Cells* **14**: 683-694.

Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Südhof TC, Wernig M (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**: 1035-1041.

Waddington CH (1957). The Strategy of the Genes. London: George Allen and Unwin.

Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, Bernstein BE, Jaenisch R (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**: 318-324.

Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH (1997). Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**: 810-813.

Xie H, Ye M, Feng R, Graf T (2004). Stepwise reprogramming of B cells into macrophages. *Cell* **117**: 663-676.

Ye Z, Zhan H, Mali P, Dowey S, Williams DM, Jang YY, Dang CV, Spivak JL, Moliterno AR, Cheng L (2009). Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. *Blood* **114**:5473-5480.

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**: 1917-1920.

Zhao Y, Yin X, Qin H, Zhu F, Liu H, Yang W, Zhang Q, Xiang C, Hou P, Song Z, Liu Y, Yong J, Zhang P, Cai J, Liu M, Li H, Li Y, Qu X, Cui K, Zhang W, Xiang T, Wu Y, Zhao Y, Liu C, Yu C, Yuan K, Lou J, Ding M, Deng H (2008). Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell* **3**: 475-479.

CHAPTER 2

Materials and Methods

2.1 DNA constructs and DNA transfections

The four mouse Yamanaka factor retroviral constructs were a kind gift from the Katherine Plath lab (). The four human Yamanaka factor retroviral constructs were bought from Addgene (Takahashi et al., 2007). NFIA was subcloned from pCHNFI-A4 available from Addgene into the pMXs vector and then subcloned into pMYs-IRES-Puro for NFIA-IP construct. NICD was subcloned from the pCAGGS-NICD vector available via Addgene. TET2CD was cloned from cDNA by taking the C-terminal part of the protein starting from K1023. RESTVP16 was created by subcloning REST (NRSF) from pHR'-NRSF-CITE-GFP available from Addgene and inserted into the pMYs-VP16-IRES-GFP vector. $ALK3^{QD}$ was subcloned from pCMV5-Alk3QDHA, a kind gift from Dr. Attisano. THRAVP16 was created by subcloning the THRA (Erb-alpha) from pAdTrack-CMV FLAG Rev-Erb alpha available from Addgene into the pMYs-VP16-IRES-GFP vector. STAT3 and PIAS3 were subcloned from vectors gifted from Dr. Ke Shuai's lab. PGKneolox2DTA.2 vector backbone was obtained from Addgene and the homology arms were cloned sequentially from genomic DNA isolated from HFF wild-type cells.

DNA transfections were done using Bio-T (Bioland Scientific), Lipofectamine-2000 (Invitrogen), or Nucleofector kits (Lonza) following manufacturer's suggested protocols.

2.2 Cell culture

All commercially available cells were maintained according to manufacturer's recommended protocol. HFF and human keratinocytes were obtained from American Type Culture Collection (ATCC). Human Keratinocytes-SFM and ADSC was obtained from Invitrogen. AFDC was a kind gift from Dr. Ying Jin's Lab and was grown under conditions

previously described (Li et al., 2009). Plat-E and Plat-A cells were a kind gift from Dr. Kitamura's lab and cultured as described previously (Morita et al., 2000). U251 astrocytoma cells were purchased from Sigma-Aldrich. Alexander's Disease patient-specific fibroblasts were obtained from the Coriell Institute.

MEFs were dissected from E17.5 CF1 mice following a previously described protocol (Takahashi and Yamanaka, 2006). Human ES and iPS cells were maintained in media as previously described (Wu et al., 2007). Mouse ES and iPS cells were maintained on gelatin-coated dishes with mitotically inactivated mouse embryonic fibroblasts and in media as previously described (Ito et al., 2010). Human NPC differentiation from iPS and ESC both followed ESC differentiation protocol previously described (Wu et al., 2007). Mouse NPC differentiation from iPS and ESC both followed ESC differentiation protocol previously described (Ge et al., 2002). E14.5 primary mouse cortical neurons were cultured in serum containing medium as previously described (Martinowich et al., 2003). Primary mouse astrocytes were cultured from P1 mice as previously described (Allaman et al., 2004). Neuronal co-culture was done following previously described protocol (Jones et al., 2011).

2.3 Reprogramming

Mouse iPS reprogramming followed previously published protocol (Takahashi and Yamanaka, 2006). Human iPS reprogramming was done following the respective protocols for the cell types (Takahashi et al., 2007; Aasen et al., 2008; Li et al., 2009; Sun et al., 2009). However, for comparative purposes, they were also reprogrammed using a unified protocol. Briefly, 100,000 cells were plated in a single well of a 6-well plate and transduced with the four Yamanaka factors as previously described (Takahashi et al., 2007). The cells were then grown in

their respective medias for two days before being passaged by 0.25% trypsin/EDTA dissociation onto irradiated MEFs and then were grown as if they were iPS for the needed times before they were picked and propagated.

Reprogramming of mouse and human astrocytes followed similar protocols with the main difference being the 4 days of culture for mouse astrocytes and 7 days for human astrocytes and the number of factors involved. Briefly, Plat-A (for human) or Plat-E (for mouse) were passaged and counted so that 4-5million cells were plated on a 10cm plate. The next day, Bio-T transfection reagent (Bioland Scientific) was used to transfect the viral vectors. 6 hours later, media was changed to fresh media. 24 hours later, HFF or MEF cells were passaged and plated in 6-well plates coated with PO/FN using 100,000cells. Another 24 hours after, (48 hours after viral vector transfections), the media of the fibroblasts were changed for the supernatant containing the viral particles with 4μg/mL polybrene (Sigma). Another 24 hours after that (72 hours after viral vector transfections), the media was changed to DMEM/F12 with 10% FBS and 1X B27. Finally, another 24 hours after that, the media was changed once again with puromycin selection.

2.4 Gene Expression Microarrays and analysis

One-color gene expression microarrays were done on the Agilent 4x44k human gene expression microarray platform (G4845A) following manufacturer's protocols and data was extracted using Feature Extraction 9.5 software (Agilent). The data was then transformed into log space and subject to quantile normalization using the Limma (Smyth and Speed, 2003). Afterwards, the ComBat algorithm was run on the data (Johnson et al., 2007). Data was then filtered for the top 50% of highest covariance in the data across all the microarrays. Finally, the

normalized, corrected data was analyzed using WGCNA (Oldham et al., 2006) with the following parameters: signed weighted correlation network; softPower=28; minClusterSize=50; deepsplit=0; cutHeight=0.3. Hub gene analysis was done using methodology described in Horvath and Dong, 2008.

2.5 Immunocytochemistry and Alkaline Phosphatase staining

Cells on coverslip were fixed (4% paraformaldehyde (PFA) at room temperature), processed and stained (overnight at 4°C) for immunofluorescence analysis using following antibodies as previously described (Get et al., 2002): mouse anti-MAP2 (Sigma); rabbit anti-synapsin (Millipore); goat anti-SOX2 (Santa Cruz); mouse anti-OCT4 (Santa Cruz); rabbit anti-GFAP (Sigma); mouse anti-GFP (Roche); mouse anti-SSEA1 (Abcam); mouse anti-SSEA4 (Abcam); mouse anti-Tra-1-81 (Millipore); mouse anti-Nanog (Santa Cruz); rabbit anti-c-myc (Sigma); mouse anti-Nestin (Millipore); rabbit anti-SMA (Abcam); mouse anti-CRYAB (Abcam); guinea pig anti-GLT-1 (Millipore); Hoechst 33342 staining was used to label nuclei.

Alkaline phosphatase staining was done using Chemicon's kit and following manufacturer's instructions.

2.6 Dendrite length and synapsin1 punctate quantification

Neuronal co-culture measurements were done by immunostaining with mouse anti-MAP2 antibody (Sigma) and rabbit anti-synapsin (Millipore). The coverslips were then imaged using a fluorescent microscope and the image was analyzed and quantified using ImageJ software (Abramoff et al., 2004).

2.7 TALENs assembly

TALENs was designed and cloned following protocol published by the Zhang group and using the kit ordered from Addgene (Sanjana et al., 2012). The following sequences were targeted: Mutant-specific left: tgg cca agc cag acc tca cc; Mutant-specific right: ttg cct cat act gcg tgc ag; WT-specific right: ttg cct cat act gcg tgc gg; Non-specific left: tca ccg cag ccc tga aag ag; Non-specific right: ttg ctg gac gcc att gcc tc.

2.8 Teratoma assay

The cells were harvested by dispase and collagenase IV treatment and centrifuged. They were then washed and resuspended in Hank's Balanced Salt Solution (Gibco). $1/4^{th}$ of the cells from a confluent 10cm dish was injected subcutaneously to each dorsal flank of a SCID mouse. 2-6 months afterwards, tumors were dissected, measured, and fixed 4% paraformaldehyde in PBS. The tissue was then embedded in paraffin, sliced, and subjected to hematoxylin and eosin staining.

2.9 shRNA knockdown and quantitative real-time PCR

shRNA knockdown was done using Clontech's Lenti-X shRNA2 system following manufacturer's protocol for oligo design, annealing, and cloning. Allele specificity design was done following previously mentioned methods (Schwarz et al., 2006; Geng and Ding, 2008) using the following oligos: P16 Forward: GAT CCg ccc tga aag aga tct gca TTC AAG AGA tgc aga tct ctt tca ggg cTT TTT TG; P16 Reverse: AAT TCA AAA AAg ccc tga aag aga tct gca TCT CTT GAA tgc aga tct ctt tca ggg cG; P13 Forward: GAT CCg aga tct gca cgc agt atg TTC AAG AGA cat act gcg tgc aga tct cTT TTT TG; P13 Reverse: AAT TCA AAA AAg aga tct gca cgc

agt atg TCT CTT GAA cat act gcg tgc aga tct cG; P1314 Forward: GAT CCg aga tat gca cgc agt atg TTC AAG AGA cat act gcg tgc ata tct cTT TTT TG; P1314 Reverse: AAT TCA AAA AAg aga tat gca cgc agt atg TCT CTT GAA cat act gcg tgc ata tct cG.

For qPCR, total RNA was extracted using Trizol reagent (Invitrogen) following manufacturer's protocol. The cDNA was made using Superscript III SuperMix (Invitrogen) following manufacturer's protocol. For real-time PCR, Fast SYBR Master Mix (Life Technologies) was used following manufacturer's protocol.

REFERENCES

Aasen T, Raya A, Barrero MJ, Garreta E, Consiglio A, Gonzalez F, Vassena R, Bilić J, Pekarik V, Tiscornia G, Edel M, Boué S, Izpisúa Belmonte JC (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol.* **26**: 1276-1284.

Abramoff MD, Magalhaes PJ, Ram SJ (2004). Image Processing with ImageJ. *Biophotonics International* **11**: 36-42.

Allaman I, Pellerin L, Magistretti PJ (2004). Glucocorticoids modulate neurotransmitter-induced glycogen metabolism in cultured cortical astrocytes. *J. Neurochem.* **88**: 900-908.

Ge W, Martinowich K, Wu X, He F, Miyamoto A, Fan G, Weinmaster G, Sun YE (2002). Notch signaling promotes astrogliogenesis via direct CSL-mediated glial gene activation. J. Neurosci. Res. **69**: 848-860.

Geng CM, Ding HL (2008). Double-mismatched siRNAs enhance selective gene silencing of a mutant ALS-causing allele. *Acta Pharmacol. Sin.* **29**: 211-216.

Horvath S, Dong J (2008). G eometric interpretation of gene coexpression network analysis. *PLoS Comput. Biol.* **4**: e1000117.

Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**: 1129-1133.

Johnson, WE, Rabinovic, A, and Li, C (2007). Adjusting batch effects in microarray expression data using Empirical Bayes methods. *Biostatistics* **8**: 118-127.

Jones EV, Bernardinelli Y, Tse YC, Chierzi S, Wong TP, Murai KK (2011). Astrocytes control glutamate receptor levels at developing synapses through SPARC-beta-integrin interactions. *J. Neurosci.***31**: 4154-4165.

Li C, Zhou J, Shi G, Ma Y, Yang Y, Gu J, Yu H, Jin S, Wei Z, Chen F, Jin Y (2009). Pluripotency can be rapidly and efficiently induced in human amniotic fluid-derived cells. *Hum. Mol. Genet.* **18**: 4340-4349.

Oldham MC, Horvath S, Geschwind DH (2006). Conservation and evolution of gene coexpression networks in human and chimpanzee brains. Proc. Natl. Acad. Sci. U.S.A. **103**: 17973-17978.

Martinowich K, Hattori D, Wu H, Fouse S, He F, Hu Y, Fan G, Sun YE (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* **302**: 890-893.

Morita S, Kojima T, Kitamura T (2000). Plat-E: an efficient and stable system for transient packaging of retroviruses. *Gene Ther.* **7**: 1063-1066.

Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F (2012). A transcription activator-like effector toolbox for genome engineering. *Nat. Protoc.* **7**: 171-192.

Schwarz DS, Ding H, Kennington L, Moore JT, Schelter J, Burchard J, Linsley PS, Aronin N, Xu Z, Zamore PD (2006). Designing siRNA that distinguish between genes that differ by a single nucleotide. *PLoS Genet.* **2**: e140.

Smyth GK, Speed, TP (2003). Normalization of cDNA microarray data. *Methods* **31**: 265-273.

Sun N, Panetta NJ, Gupta DM, Wilson KD, Lee A, Jia F, Hu S, Cherry AM, Robbins RC, Longaker MT, Wu JC (2009). Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **106**: 15720-15725.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**: 861-872.

Takahashi K, Yamanaka S (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663-676.

Wu H, Xu J, Pang ZP, Ge W, Kim KJ, Blanchi B, Chen C, Südhof TC, Sun YE (2007). Integrative genomic and functional analyses reveal neuronal subtype differentiation bias in human embryonic stem cell lines. *Proc. Natl. Acad. Sci. U.S.A.* **104**: 13821-13826.

CHAPTER 3

Assessing the gene expression differences of four parental cell types that contribute to iPS reprogramming speed

3.1 Introduction

Though Gurdon, with his SCNT experiments, and Davis, with his MYOD1 experiments, showed that reprogramming cell identity was possible, it wasn't until Yamanaka combined the work of his two predecessors did the field of reprogramming really take off and become a major field of study. Since then, the process of iPS reprogramming has been studied using many methods. The underlying mechanisms that cause reprogramming are still quite mysterious, though many important aspects of the process have been elucidated.

One of the first questions to be studied was the question of why iPS reprogramming was so inefficient and took such a long time for the process to complete as earlier studies reprogramming with cell fusion was very quick (Tada et al., 2001). To this end, many groups studied whether replacing or adding additional factors could change the dynamics, efficiency, or speed up the reprogramming process. Early work with different chemical additives such as 5-azaC (Mikkelson et al., 2008), valproic acid (Huangfu et al., 2008), and ascorbic acid (Wang et al., 2011) had promising results. These additives seemed to increase efficiency as well as the transition from partial reprogramming to the fully reprogrammed state, but none of them improved the rate-limiting step, which was the dedifferentiation from the fibroblast state to the partially reprogrammed state. Other researchers tried to add additional protein factors to try to increase efficiency as well. Utf1 (Zhao et al., 2008), Esrrb (Feng et al., 2009), Sall4 (Tsubooka et al., 2009; Plews et al., 2010), L-myc (Nakagawa et al., 2010), and Glis1 (Maekawa et al., 2011) were all found to increase efficiency but again, not speed. Finally, miRNAs have been shown to be able to reprogram iPS in conjunction with the Yamanaka factors or even in lieu of the factors (Subramanyam et al., 2011; Liao et al., 2011). Of especial note, one set of miRNAs, the

mir302/367 cluster of microRNAs, in combination with Valproic Acid (VPA) reprogrammed fibroblasts without any of the Yamanaka factors. Surprisingly, this combination of mir302/367 with VPA was able to increase efficiency ***and*** speed when used for reprogramming (Anokye-Danso et al., 2011). This shows that there are important epigenetic changes that can be induced that actually increase reprogramming speed. Since miRNAs can potentially target a pool of different mRNAs, it is hard to discern what critical genes are being regulated that can help reprogramming speed. But since miRNAs alter expression of proteins, it could be gene expression patterns that alter reprogramming speed.

Another question that was quickly investigated was whether or not fibroblasts were uniquely able to be reprogrammed into iPS or would it be possible to reprogram other cell types into iPS. Yamanaka's original report of iPS reprogramming used fibroblasts as the parental cell type to reprogram into iPS. Later, it was shown that other cell types such as liver and stomach cells (Aoi et al., 2008), pancreatic b cells (Stadtfeld et al., 2008), keratinocytes (Aasen et al., 2008), mature B lymphocytes (Hanna et al., 2008), peripheral blood cells (Ye et al., 2009), adipose stem cells (Sun et al., 2009), neural stem cells (Kim et al., 2009), amniotic fluid-derived cells (Li C et al., 2009), and many others. Interestingly, some of these cell types reprogrammed faster and more efficiently than the others while some were as slow as fibroblasts. It seemed that different cell types intrinsically had different reprogramming efficiencies and speeds. Since reprogramming efficiency is easily modulated by addition of other factors as mentioned earlier, the fact that different cell types reprogrammed with variable efficiencies was not too surprising. On the other hand, the time it took to reprogram a cell was more difficult to change using additional factors and seemed to be dependent on cell type.

An open question that still remains is what determines the reprogramming speed of a cell. Since nuclear reprogramming by cell fusion or SCNT occurs rapidly, that means iPS reprogramming speed still has huge potential in speed gains. Results from other labs show that different cell types reprogram at different speeds. And as shown by the fact that miRNA are able to increase reprogramming speed, it is very likely that protein and gene expression is a possible modulator of speed. Therefore, I wanted to investigate what differences in gene expression patterns of different cell types can possibly explain the different reprogramming speeds.

3.2 Results

3.2.1 Reprogramming Mouse iPS Using the 4 Yamanaka Factors

In order to establish the basic protocol for reprogramming cells using defined, ectopically expressed transcription factors, I started by replicating the results presented by Shinya Yamanaka's group in 2006 (Takahashi and Yamanaka, 2006) by reprogramming Mouse Embryonic Fibroblasts (MEFs) into mouse induced Pluripotent Stem Cells (iPSCs). Although I was interested in human reprogramming, as a proof-of-principle, mouse cells were used first because mouse (miPSCs) are easier to maintain and much faster and more efficient to reprogram as compared to human pluripotent stem cells. Successfully reprogramming mouse cells into miPSCs was a great stepping stone to the more difficult method of reprogramming human cells into human iPSCs.

The first step was to clone the four factors into retrovirus expression vectors as was done by the Yamanaka lab. Retrovirus expression vectors allows for robust and sustained expression of the gene products in mammalian MEFs as well as high infection rates (Figure 3.1A). In particular, an added benefit of the pMXs expression vector system is that the integrated expression cassettes for the exogenous genes have been shown to be silenced upon successful iPS generation (Yao et al., 2004). Fortunately, the mouse versions of the four Yamanaka factors (OKSM) were already cloned into pMXs vectors by the Katherine Plath group. They were kind enough to share the constructs with me and these constructs were used for all the following mouse iPS reprogramming experiments. The cells used were the Plat-E cells that were kindly given to us by the Kitamura group. These were the same cells used by the Yamanaka group in their original reprogramming experiments. Plat-E cells were designed specifically to be used

with the pMXs viral vector system for simple transfection and viral packaging at extremely high titers and therefore, high infection rates (Morita et al., 2000).

Following the original published protocol by Takahashi and Yamanaka, I attempted to establish the technology in the lab using MEFs as the parental cell type. Initial attempts were difficult because we had no reporters of successful reprogramming and the only method of positive colony selection was by colony morphology. This was different from Yamanaka's original paper, which used MEFs from transgenic mice that had an Fbx15 driven GFP reporter, making their selection of miPS much easier. After many attempts, though, I was able to generate a few lines of miPS clones. The miPS generated stained positively for alkaline phosphatase activity (Figure 3.1B), which is often used as an early indicator of successful reprogramming. However, alkaline phosphatase activity is generally considered to be less stringent as it also stains partially reprogrammed colonies as well as fully reprogrammed colonies. Therefore, I also used immunofluorescence to stain for pluripotent markers (Figure 3.1C). OCT4 was used as a positive control since this was one of the proteins ectopically expressed by the retrovirus. However, NANOG and SSEA-1 are both endogenous as neither was transduced into the cells via retrovirus. The miPS colonies also had typical mESC colony morphology and maintained pluripotency over multiple passages.

3.2.2 Reprogramming Huntington's Disease Mouse Model iPS

With the success of the initial experiments, I went on to attempt to reprogram Huntington's Disease (HD) mouse model MEFs that were being studied by a colleague in the lab. The HD model cell line was created by a BAC transgene for huntingtin. Using the same protocol as the successful reprogramming experiment, I was able to reprogram these MEFs into miPS.

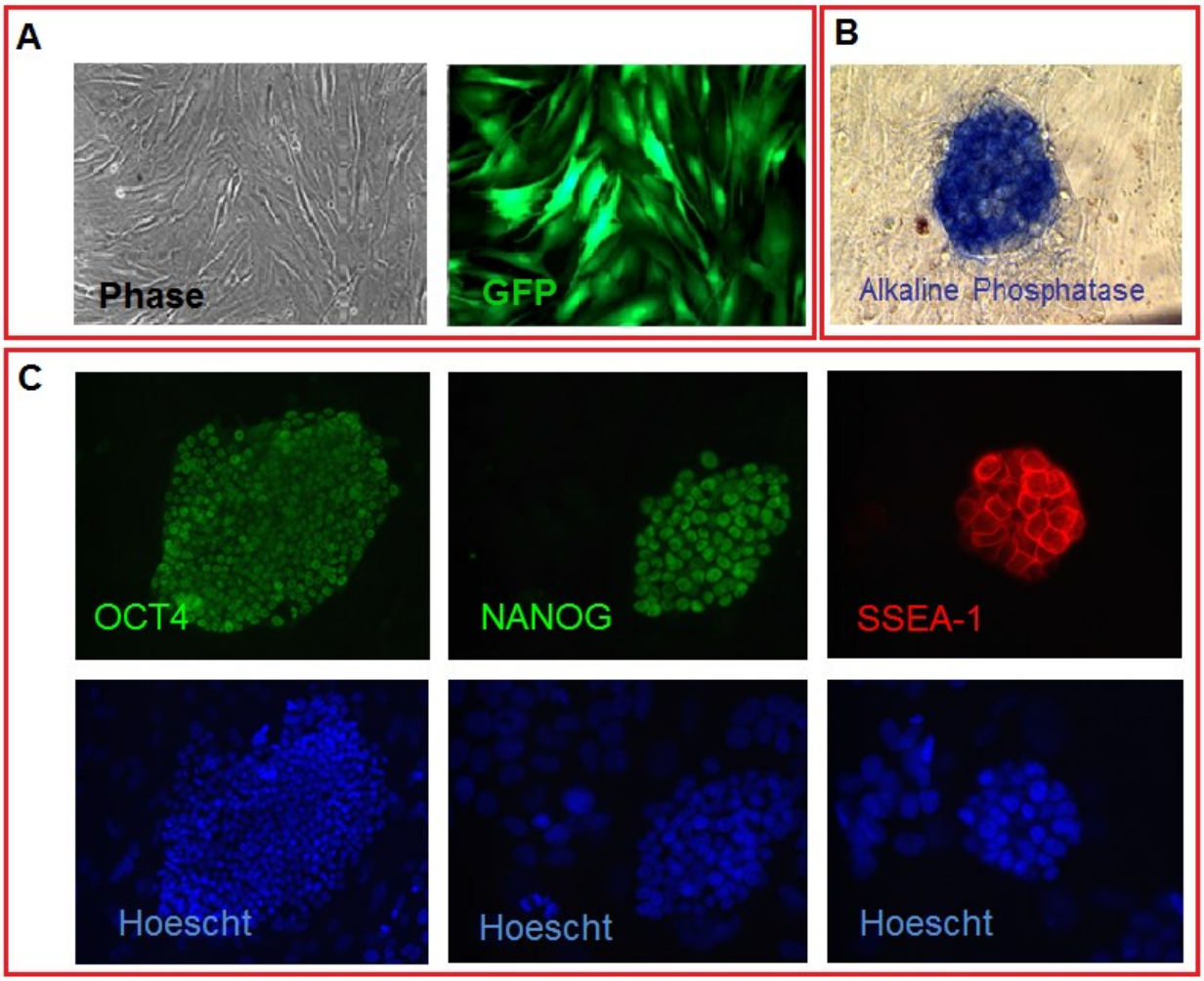

Figure 3.1 Reprogramming MEFs into mouse iPS. (A) pMXs retroviral system allows for very high transfection rates and high expression of exogenous proteins like GFP. (B) Alkaline Phosphatase is an early marker of iPS reprogramming. (C) OCT4, NANOG, and SSEA-1 are all markers of pluripotency. NANOG and SSEA-1 are endogenous markers. Also, the formation of compact colonies is indicative of iPS cells.

Because I had experience reprogramming the wild-type MEFs, this reprogramming went much more smoothly. The HD miPS appeared normal compared to wild-type miPS and mESCs in almost all regards including pluripotency markers and colony morphology. However, one minor difference seemed to be passage time with the HD iPS proliferating at a slower rate and thus takes a longer time between passages as compared to wild-type iPS.

Since HD is a neurodegenerative disease, I differentiated these HD model miPS into NPCs using an established protocol within the lab (Ge et al., 2002). Cells were collected at the miPS and NPC stage and assessed via qPCR for the expression of the endogenous pluripotency markers before and after differentiation. Figure 3.2A shows that the endogenous expression of Nanog, Oct4, and Sox2 are all very robust and comparable to the wild-type ESCs control during the miPS stage. For the NPC stage, Oct4 and Nanog should be silenced while Sox2 should be still expressed, although at a lower level. The results (Figure 3.2B) show that Nanog is drastically reduced during differentiation as expected, but Oct4 levels were not reduced as dramatically even in wild-type controls. This could be due to the fact that the differentiation is a mixed culture and there could still be strong levels of Oct4 in the undifferentiated cells mixed into the culture. Sox2 expression was also expressed at a lower level, as expected. The results were somewhat promising though the experiments would need to be repeated. However, since the HD model miPS was not my focus, I let my colleague continue those experiments while I moved on to human iPS.

3.3.1 Reprogramming Human iPS Using the 4 Yamanaka Factors

Having learned a lot from reprogramming the MEFs into miPS, I then started reprogramming Human Foreskin Fibroblasts (HFF) into hiPS. The protocol was similar to MEFs,

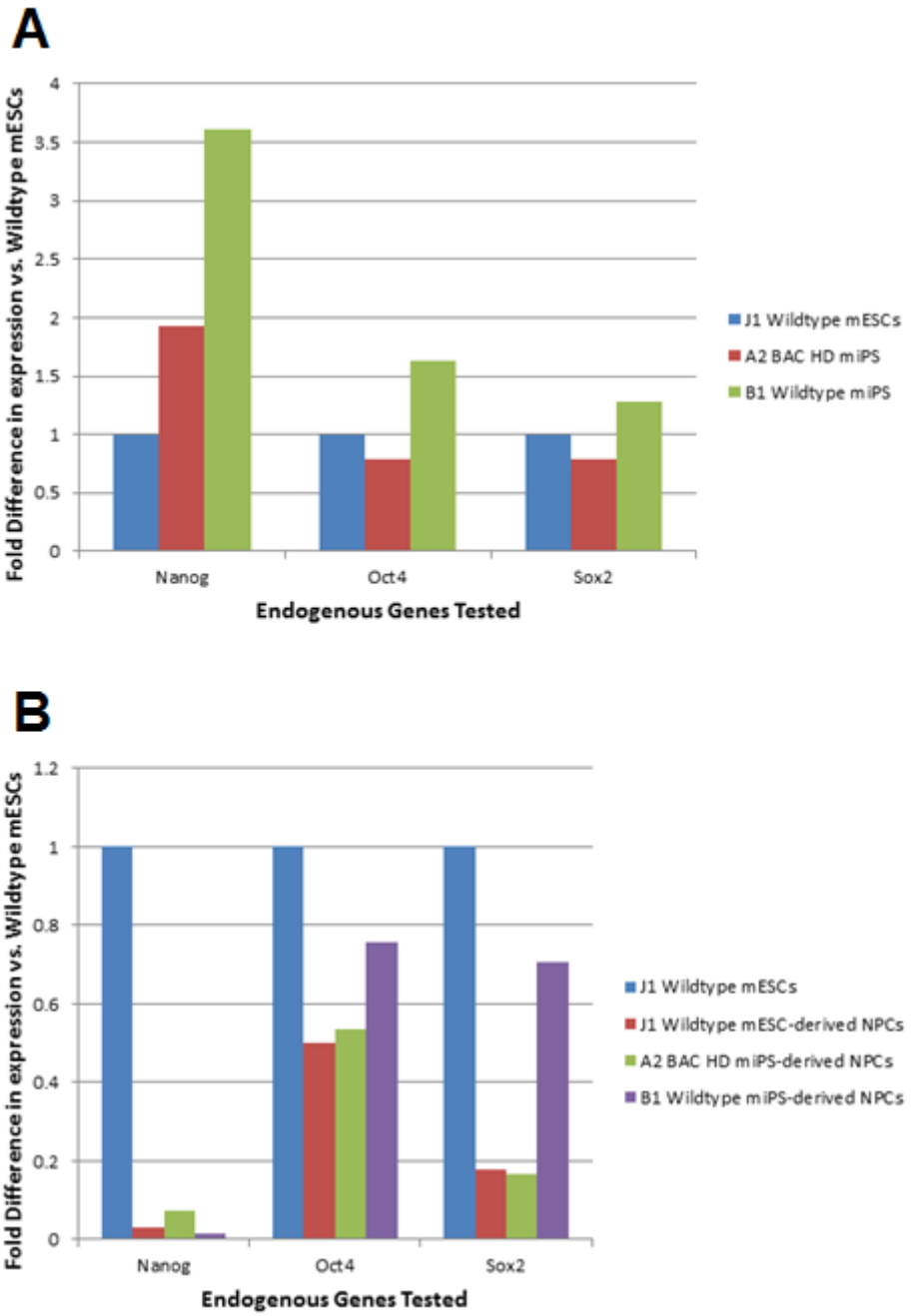

Figure 3.2 Endogenous pluripotency markers of iPSCs and ESCs and NPCs. (A) J1 wild-type mESCs were used as control to compare the endogenous expression levels of Nanog, Oct4, and Sox2 by qPCR in Huntington's Disease miPS and wild-type miPS. (B) J1 wild-type mESCs were used as control to compare the endogenous expression levels of Nanog, Oct4, and Sox2 by qPCR in Huntington's Disease miPS-derived-NPCs and wild-type miPS-derived-NPCs.

but took a much longer time. Reprogramming into hiPS was much more challenging than the initial attempts to reprogram miPS because not only was there no marker for selection of successfully reprogrammed colonies, but passaging of the hiPS and hESCs was much more difficult and much more involved. Also, hiPS and hESCs are much more difficult to maintain in an undifferentiated state compared to miPS, which do not require bFGF feeding regularly. Fortunately, the four human factors in the pMXs vector were commercially available from Addgene, and that is what I used for the reprogramming experiments.

Since hiPS was a much more challenging protocol, many of my early attempts at reprogramming HFF into hiPS failed. However, I was eventually able to reprogram the HFF into hiPS after many failed experiments. These hiPS had the flat colony morphology with bright edges that typical of human ESCs and stained positive for alkaline phosphatase activity (Figure 3.3A). Furthermore, they expressed the correct pluripotency markers including endogenous nuclear markers like Nanog and endogenous surface markers like Tra-1-81 and SSEA4 (Figure 3.3B). Since these hiPS looked like ESCs and also expressed the correct markers, I tried to differentiate these cells into NPCs by embryoid body. The iPS cells were able to form embryoid bodies with similar morphology to control ESCs and differentiate into Nestin-positive NPCs, and endogenous marker of NPCs (Figure 3.4). Finally, the most stringent test for pluripotency in human cells is the teratoma assay. Upon injection of hiPS into the leg of a Severely Combined Immuno-Deficient (SCID) mouse, the hiPS line was able to form a teratoma (Figure 3.5A). Furthermore, sectioning and staining with hematoxylin and eosin showed cells representing all three germ layers (Figure 3.5B). Therefore, the cells could be considered fully reprogrammed hiPS cells since partially reprogrammed cells do not give rise to teratomas that have cells of all three germ layers.

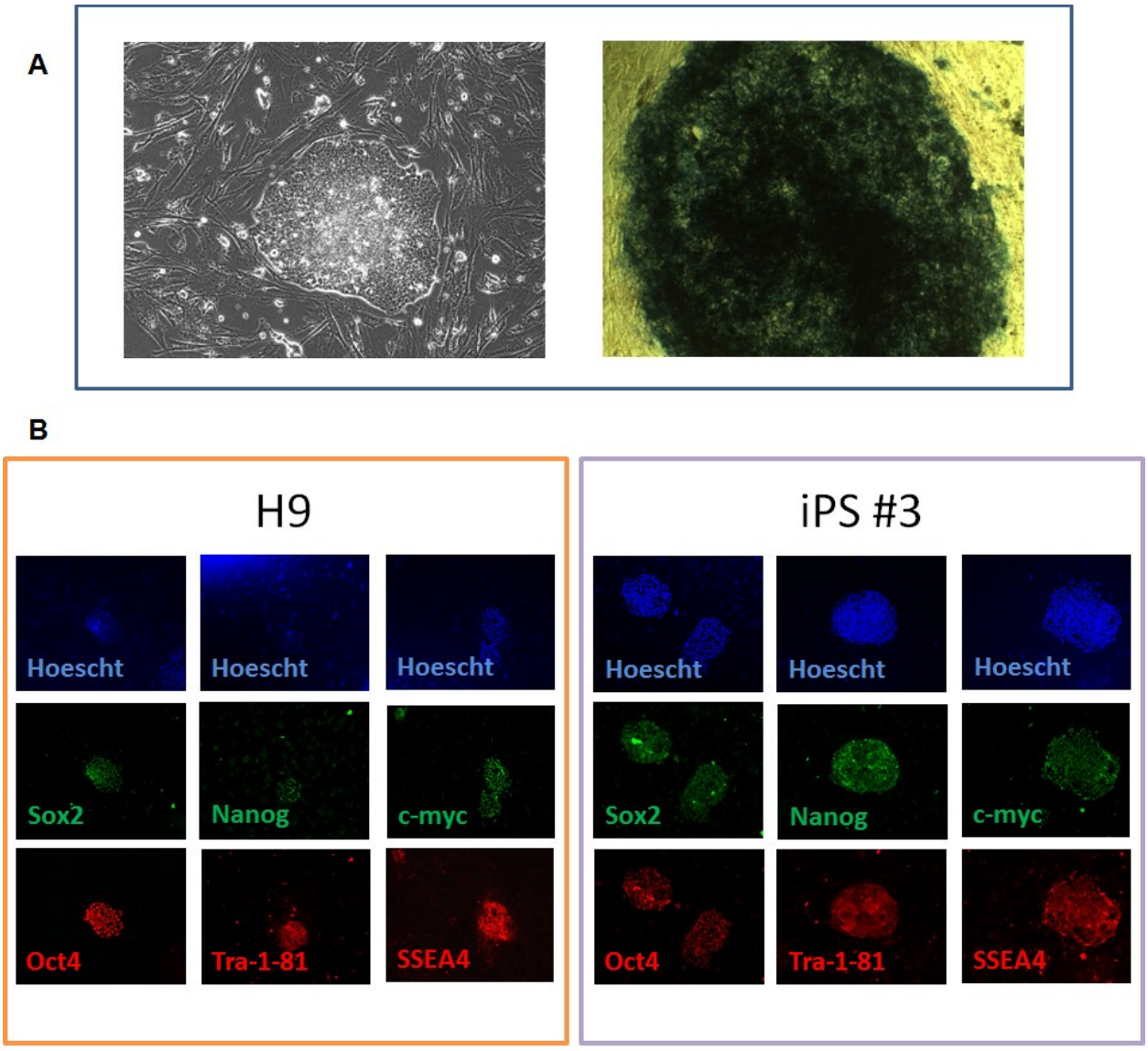

Figure 3.3 hiPS generated from reprogramming HFFs are similar to ESCs. (A) hiPS have typical ESC morphology and have alkaline phosphatase activity. (B) H9 and iPS derived from HFFs stain positive for many of the human pluripotency markers.

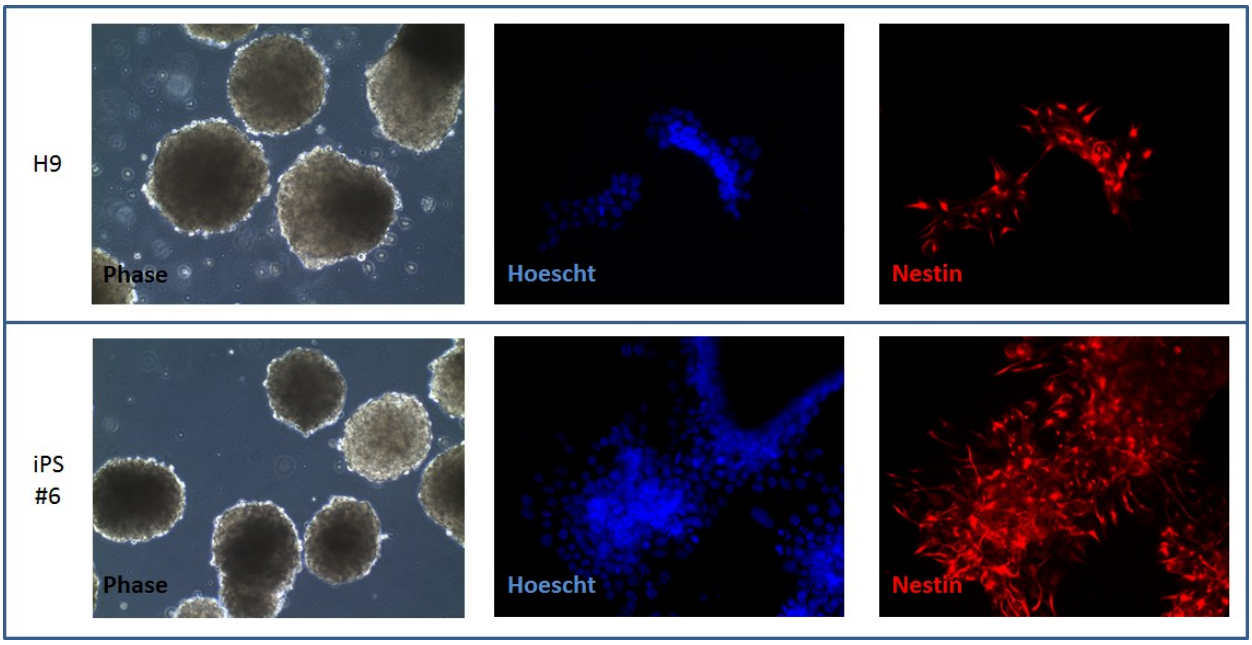

Figure 3.4 H9 and hiPS differentiation. H9 was used as a control and were compared to hiPS in their ability to differentiate into NPCs via embryoid body formation. The NPC marker Nestin was used as a marker for successful reprogramming.

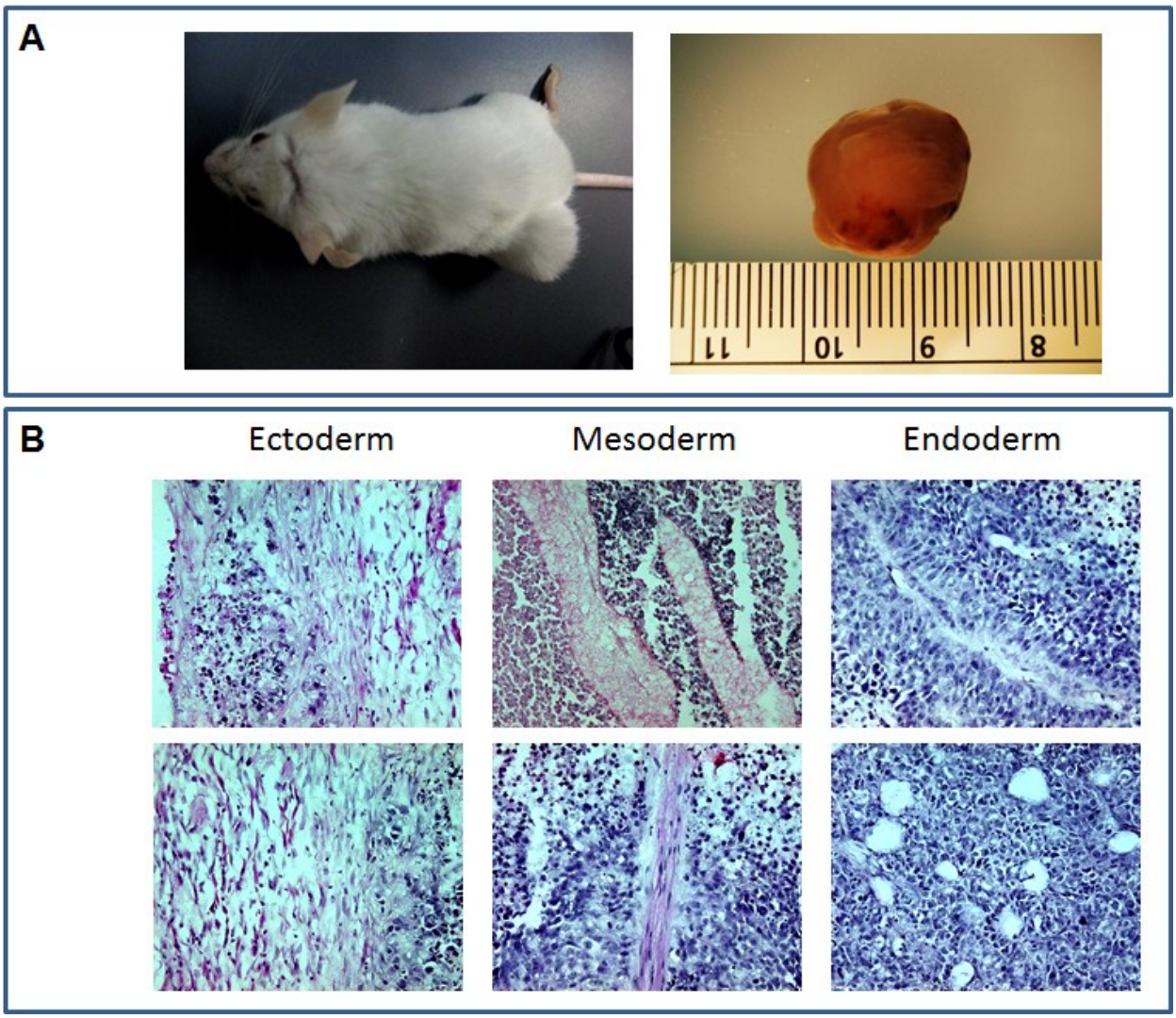

Figure 3.5 hiPS form teratomas. The hiPS injected into SCID mice form teratomas that have cells from all three germ layers. This is the most stringent test for human pluripotent stem cells.

3.3.2 Reprogramming Four Different Parental Cell Types

With the success of reprogramming HFF into hiPS, I went on to investigate the differences in reprogramming speeds of different parental cell types. In the literature, three cell types that I had access to reprogrammed at faster speeds than fibroblasts while still using only the four Yamanaka factors. These included human keratinocytes, Amniotic Fluid Derived Cells (AFDCs), and Adipose Derived Stem Cells (ADSCs). Therefore, I wanted to reprogram these 3 different cell types and compare them to Human Foreskin Fibroblasts (HFF). Keratinocytes were commercially available from two vendors and so I tried using both types of keratinocytes. All five of these cell types were reprogrammed by transducing all four of the Yamanaka factors so that there would be no variation in the reprogramming protocol.

Interestingly, some of my reprogramming results were different from those reported in the literature. The control HFFs reprogrammed as expected such that the colonies produced were ready to be picked within 4 weeks of viral infection. ADSCs also reprogrammed as expected and were ready to be picked within 2 weeks. On the other hand, AFDC reprogrammed slower than reported, though still faster than HFFs. AFDC colonies were ready to be picked within about 2 weeks as compared to the 1 week reported in the original paper (Li C et al., 2009). Finally, the two keratinocyte lines were the most different from that reported in the literature, which were reported to be ready to be picked within 3 weeks (Aasen et al., 2008). The Invitrogen keratinocytes had no successful reprogramming at all and the ATCC keratinocytes took a full 4 weeks before they were ready to be picked, similar to the control HFF. The differences in the keratinocytes reprogramming speeds could be a result of differences in keratinocyte derivation and maintenance compared to the ones used in the original report. The

keratinocytes used in the original paper by Aasen et al. were from juvenile foreskins and from single individuals while the ATCC keratinocytes were from pooled neonatal foreskin fibroblasts. The differences in age and most likely derivation protocols could account for the possible differences in reprogramming. The Invitrogen keratinocytes were specially adapted to their proprietary SFM medium and could only grow in that special medium. Thus, the Invitrogen keratinocytes were likely to be fundamentally different from normally derived keratinocytes such as the ATCC and the ones from Aasen et al. Overall, the reprogramming results were mixed. However, the differences were somewhat of a blessing in disguise because the four cell types had four different reported reprogramming speeds. AFDC, ADSC, keratinocytes, and HFF were reported to be ready to be picked after one, two, three, and four weeks, respectively. So rather than comparing 4 cell types with 4 different reprogramming speeds, instead I could now compare two cell types that reprogrammed within the same 2 weeks, namely the AFDCs and ADSCs, and two cell types that reprogrammed within the same 4 weeks, namely the HFFs and the keratinocytes from ATCC. AP staining at two weeks shows the differences in that AFDC and ADSC had fewer, larger colonies that were ready to be picked, HFF and ATCC keratinocytes had more numerous, but also smaller colonies that needed two more weeks to grow before they were ready to be picked (Figure 3.6). The reprogramming results are summarized in Table 3.1.

Parental Cell Line	Source	Source in Literature	Reported Speed	Speed Achieved	Citation
AFDC	Ying Jin Lab	Primary derived	1 week	2 weeks	Li et al., 2009
ADSC	Invitrogen	Primary derived	2 weeks	2 weeks	Sun et al., 2009
ATCC keratinocytes	ATCC	Primary derived	3 weeks	4 weeks	Aasen et al., 2008
Keratinocytes-SFM	Invitrogen	Primary derived	3 weeks	Did not fully reprogram	Aasen et al., 2008
HFF	ATCC	Cell Applications, Inc.	4 weeks	4 weeks	Takahashi et al., 2007

Table 3.1 Summary of reprogramming four cell types into iPS. ADSC and HFF reprogrammed as expected. AFDC and ATCC keratinocytes were slightly slower and keratinocytes-SFM from Invitrogen did not fully reprogram.

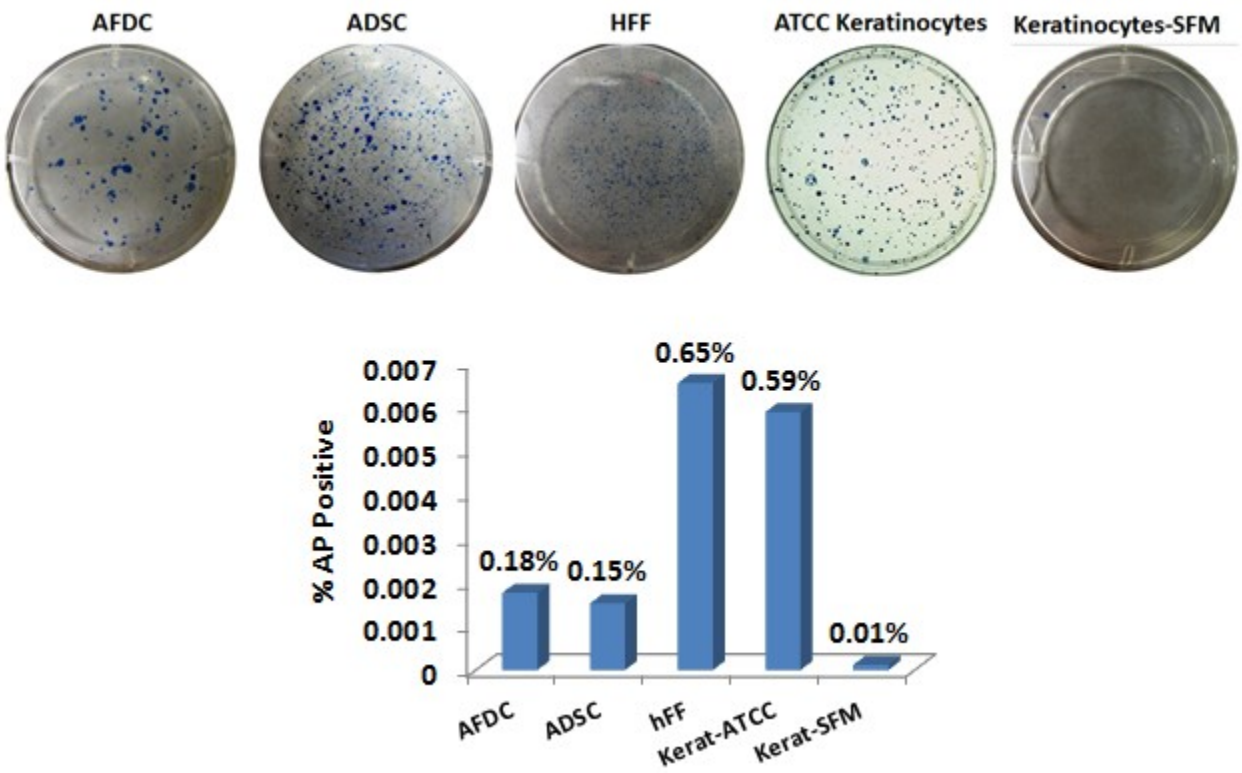

Figure 3.6 AP staining of iPS derived from four cell types. Alkaline phosphatase staining shows that the AFDC and ADSC were able to form larger, more developed colonies while the HFF and ATCC keratinocytes formed many smaller, less developed colonies. Keratinocytes-SFM did not successfully reprogram and had negligible AP staining.

3.3.3 Expression Analysis of Four Parental Cell Lines

In order to compare the four different parental cell lines expression profiles, I ran one-color Agilent 4x44k human gene expression microarrays in triplicate. The libraries were prepared according to the manufacturer's protocol and gene expression data was extracted from the image using Agilent's Feature Extraction software. Upon initial analysis of gene expression differences, a strong pattern was observed. As shown in Figure 3.7, the Principle Component Analysis (PCA) of all the data showed that the strongest component was the fact that the ectodermic keratinocytes were vastly different and separated from the rest of the other cell types, which morphologically appeared similar to fibroblasts and are cells of the mesoderm layer so this was not too surprising. The second strongest component, however, was easily separated by the experiment replicate number, meaning that there was a strong batch effect. Batch effects are very common in microarray experiments and the strong batch effect could mask any differences that are vital for the analysis. As a result of the strong similarities the three fibroblast-like cell types (AFDC, ADSC, and HFF) and the strong batch effect, I decided that the normal gene expression analysis pipeline would be insufficient for my specific experiment since that would likely show mainly the differences based on cell type and experimental replicate number rather than the underlying reprogramming dynamics.

Pioneering bioinformatics work has been done by the Geschwind and Horvath groups at UCLA. They have developed a more sophisticated method of gene expression analysis called Weighted Gene Co-expression Network Analysis or WGCNA (Oldham et al., 2006). The main idea of this method of gene expression analysis is that they create a network of interconnected genes that seem to co-express together and group these genes together as a module that might be

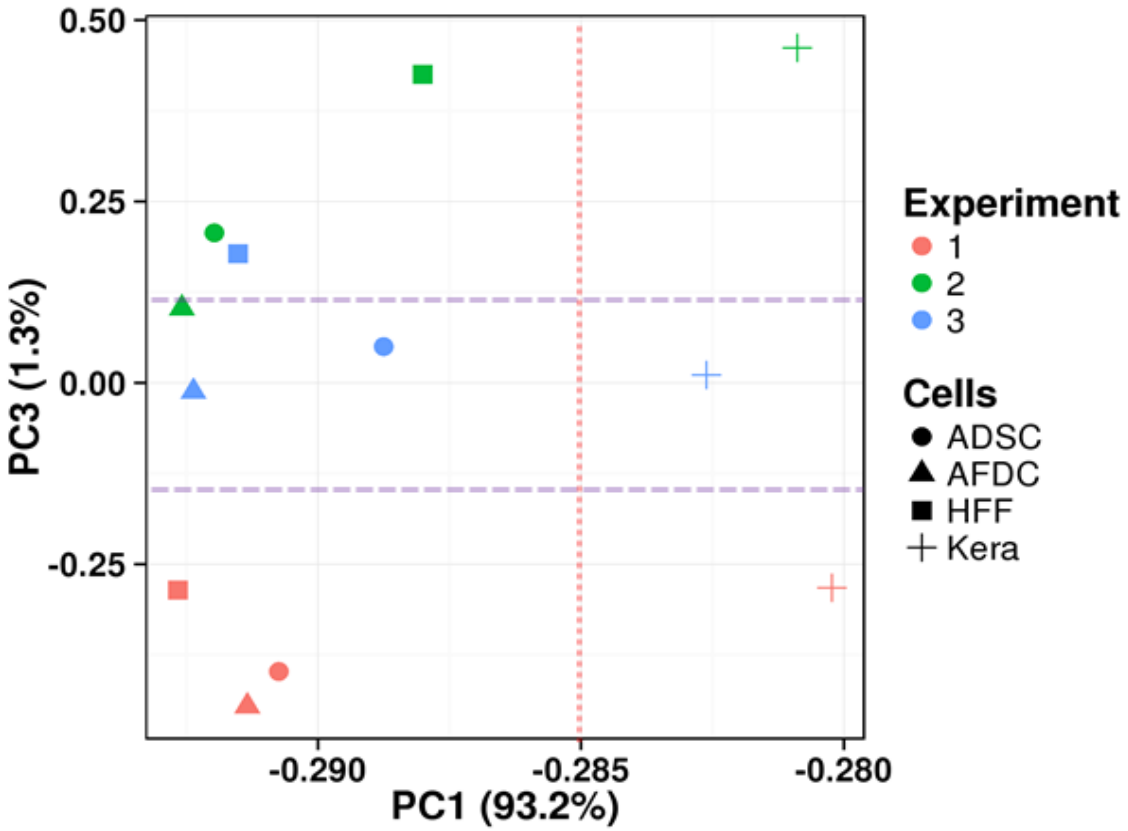

Figure 3.7 PCA analysis of the gene expression data from four cell types. Principle Component 1 (PC1) represents the divide (red dotted line) of the cell types with fibroblast-like clustering to the left of the graph while keratinocytes clustered to the right. PC2 represents the batch effect, with each color being separated along the vertical axis (dashed purple lines).

working together for a specific biological function. Therefore, I expected that using WGCNA would be able to help me find different modules corresponding to different characteristics of the data set like a rather large module for genes expressed mainly in fibroblast-like cells and another large module for genes expressed mainly in keratinocytes. More importantly, though, was the ability to find a module for genes co-expressed in fast reprogramming cells and genes co-expressed in slow reprogramming cells. The batch effect would be eliminated by a simple algorithm called ComBat, which has been used by other researchers to eliminate the batch effect (Johnson et al., 2007).

In order to verify that WGCNA could be useful for my analysis, I did a test using publicly available gene expression data from the Caltech ENCODE database. The database included RNA-seq data from multiple cell types including many cell lines, cells from different tissues, and human ESCs. I used WGCNA to find an ESC-specific modules and important factors in ESC identity and reprogramming. Since there is a large number of cell types that are not ESCs, finding an ESC-specific module should be relatively easy. After downloading and reformatting the huge dataset, I ran WGCNA on the data and as a result, was able to extract an ES-specific module (Figure 3.8) consisting of 892 genes. I then used the distance matrix of the co-expression network created by WGCNA to calculate the most important genes within that ES-specific module and ranked them in a list. Looking into the module, I found important genes such as Oct4, Nanog, Fbx15, Sox2, Utf1, Glis1, amongst many others that were shown to be important for reprogramming. Although most of these important genes (Nanog, Sox2, Utf1, Fbx15) were ranked within the top 10 of the list, many of these important genes (Glis1 and crucially Oct4) were ranked very low on the module's weighted distance. Klf4 and c-myc, which

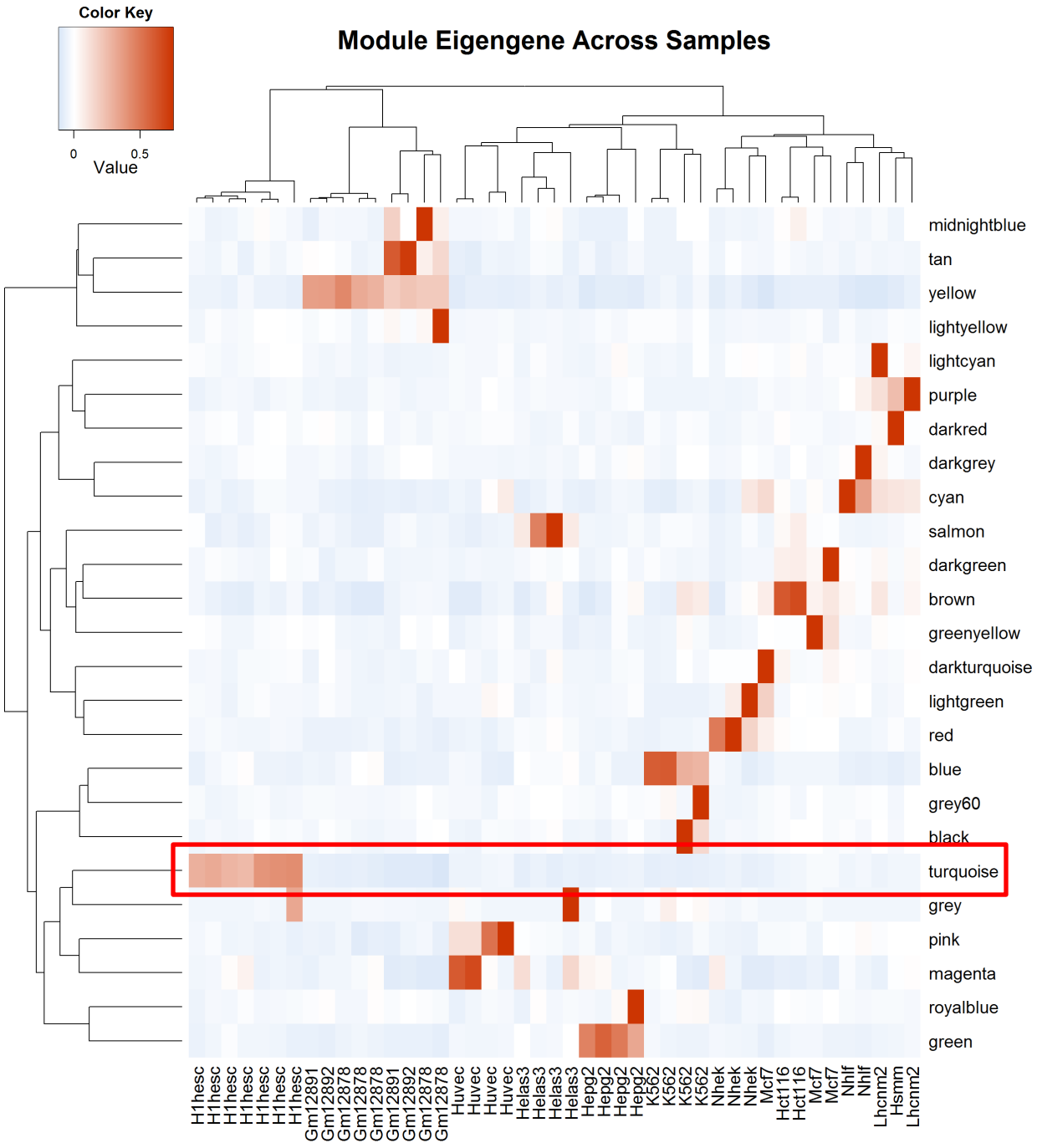

Figure 3.8 WGCNA analysis of ENCODE data. The ESCs clustered separately from the rest of the cell types and the module labeled as turquoise (red box) represents a module that is specific to ES cells.

are not ES-specific were not on that module list. However, the results were encouraging as it showed that WGCNA could easily distinguish the ES cells from the rest of the cells and could isolate a module that was important for pluripotency and reprogramming. However, the caveat was that the ranking created by the distance matrix was not necessarily indicative of the factors that would prove to be most important.

With the promising though mixed results of the test using the ENCODE data, I processed the data of the four parental cell lines using WGCNA as well (Figure 3.9). As expected, there was a very large module that was keratinocyte-specific and another large module that was fibroblast-like specific. There was also a module for each specific cell type, which also makes sense. Importantly, there was a module that had genes upregulated in the slow reprogramming cells and downregulated in the fast reprogramming cells (the "Slow Module"). Conversely, there was also a module that had genes upregulated in the fast reprogramming and downregulated in the slow reprogramming cells (the "Fast Module"). Looking at the network distance weights, Pias3 was the most important gene in the Fast Module according to WGCNA. Another interesting gene that ranked highly in the Fast Module was Gadd45b, which has been implicated along with Gadd45a in active demethylation (Barreto et al., 2007) and demethylation is an important part of reprogramming (Maherali et al., 2007). Also, Sall4 was found in the fast module and had already been shown to increase reprogramming efficiency (Plews et al., 2010; Tsubooka et al., 2009). Finally, Kdm4c was also in the fast module. Kdm4c has been shown to be highly expressed in ES cells (Katoh and Katoh, 2007) and also help iPS reprogramming efficiency (Chen et al., 2013). For the Slow Module, Eif1ay was the most important gene according to WGCNA. A potentially important gene is the Npm1 gene, which regulates p53 and Ink4a/Arf (Cheng et al., 2007), both of which have been shown to be a key component regulating

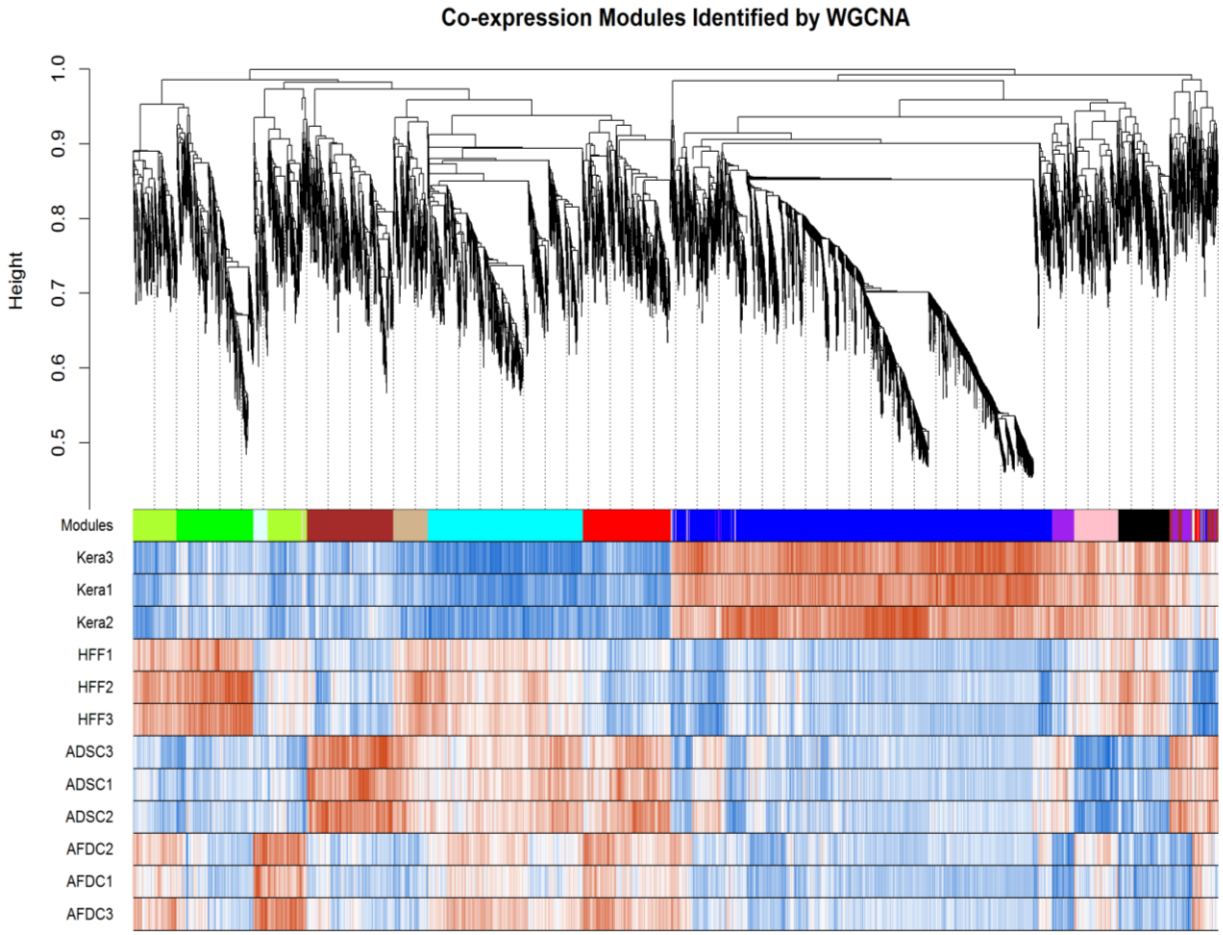

Figure 3.9 WGCNA analysis of the four parental cell lines. The largest module, the blue module represents genes upregulated in keratinocytes specifically. The second largest module, the turquoise module represents genes upregulated in fibroblast-like cells specifically. The red module represents genes upregulated in cells that reprogram within 2 weeks (the fast module) and the black module represents genes upregulated in cells that reprogram within 4 weeks (the slow module.

iPS reprogramming (Marion et al., 2009; Li H et al., 2009). Of especial interest was the fact that Stat3, the target of the inhibition of Pias3 (Protein Inhibitor of Activated Stat 3), was part of the Slow Module. It would make sense if Stat3 limits reprogramming speed but Pias3 would be able to abrogate the effect of Stat3 and result in increasing reprogramming speed. Furthermore, both of these genes are heavily involved in cell signaling as part of the Jak/Stat pathway. The Jak/Stat pathway has been implicated in differentiation (reviewed in Harrison, 2012) and therefore, Pias3 made sense as a possible important factor in speeding up iPS reprogramming, which is a form of dedifferentiation. As a result, these two genes seemed like very good and logical candidates to attempt to modulate the reprogramming speeds of the four cell types. Table 3.2 and Table 3.3 list some of the top weighted genes in the fast module and slow module, respectively.

Name	Description	Wght
PIAS3	SUMO E3-ligase. Regulates STAT3 signaling.	138.7523
APBB1IP	Stands for Amyloid-beta precursor Protein-Binding, family B, member 1 Interacting protein. Involved in Ras activation and actin cytoskeletal remodeling.	138.0052
AGPAT6	1-AcylGlycerol-3-Phosphate O-AcylTransferase 6. Involved in signal transduction and lipid biosynthesis.	130.447
MAP3K3	Mitogen-Activated Protein Kinase Kinase Kinase 3 (MAPKKK3). Mediates activation of the NF-kappa-B, AP1 and DDIT3 transcriptional regulators.	130.1985
GADD45B	Mediates activation of stress-responsive MTK1/MEKK4 and MAPKKK. Involved in active demethylation?	109.0730
SALL4	Transcription factor that can enhance reprogramming to iPS.	64.85556
KDM4C	aka JMJD2C. Demethylates H3K9me3 to me2 and H3K36me3 to me2.	50.84108

Table 3.2 List of genes in the red module. Pias3, Gadd45b, Sall4, and Kdm4c could all be potentially important genes that affect iPS reprogramming.

Name	Description	Wght
EIF1AY	Required for Ribosomal translational initiation and elongation.	79.95503
RAB11FIP1	A Rab11 GTPase effector protein involved in the endosomal recycling process.	74.46802
NPM1	Ribosomal biogenesis, histone assembly, cell proliferation, and regulation p53 and ARF.	73.18331
RPLP0	Component of the ribosomal 60S subunit.	73.07044
TBC1D8	May act as a GTPase-activating protein for Rab family protein.	72.67004
CD9	Differentiation, adhesion, and signal transduction, and tumor suppression.	68.51048
STAT3	Activated in response to cytokines including IFNs, EGF, IL5, IL6, HGF, LIF and BMP2.	16.59215

Table 3.2 List of genes in the black module. Npm1 and Stat3 are potentially important genes in regulation of iPS reprogramming.

3.4.1 Reprogramming with STAT3 and PIAS3

Since Pias3 was listed as the top hit in the Fast Module and Stat3 was listed within the Slow Module, they were natural targets to attempt to alter the speed of reprogramming. Therefore, I subcloned the two genes from two vectors kindly given by the Shuai lab into the pMXs vector so that they can be retrovirally introduced with the other four Yamanaka factors. For this experiment, I transduced the two fast reprogramming cell types, AFDC and ADSC, with the four Yamanaka factors along with Stat3, which might slow down their reprogramming speed to something more similar to the 4 weeks needed by the "slow" reprogramming cells. On the other hand, the two slow reprogramming cell types, HFF and ATCC keratinocytes, were transduced with the four Yamanaka factors along with Pias3, which might speed up their reprogramming to the 2 weeks needed by the "fast" reprogramming cells. The cells were then assessed to see if they would reprogram faster or slower compared to their controls. Unfortunately, there was no obvious effect of the two factors on their respective cell types as

they all reprogrammed at the normal, expected speeds. It is possible that Pias3 alone or Stat3 alone is not sufficient to alter the reprogramming speeds of the different cell types since it is possible that a combination of changes is needed for reprogramming speed changes as implicated by the success of miRNAs in combination with VPA to speed up the reprogramming speed. Another possibility is that Pias3 and Stat3 are actually the wrong candidate genes. As revealed by my test using the ENCODE data, WGCNA's rankings are do not necessarily correspond to the biologically most important gene and is a limitation of the approach. In other words, the best candidate could be another, possibly more obscure, gene that is listed less prominently within the module and there would be more study required to elucidate that possibility.

REFERENCES

Aasen T, Raya A, Barrero MJ, Garreta E, Consiglio A, Gonzalez F, Vassena R, Bilić J, Pekarik V, Tiscornia G, Edel M, Boué S, Izpisúa Belmonte JC (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol.* **26**: 1276-1284.

Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, Zhang Y, Yang W, Gruber PJ, Epstein JA, Morrisey EE (2011). Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* **8**: 376-388.

Aoi T, Yae K, Nakagawa M, Ichisaka T, Okita K, Takahashi K, Chiba T, Yamanaka S (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* **321**: 699-702.

Barreto G, Schäfer A, Marhold J, Stach D, Swaminathan SK, Handa V, Döderlein G, Maltry N, Wu W, Lyko F, Niehrs C (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445**: 671-675.

Chen J, Liu H, Liu J, Qi J, Wei B, Yang J, Liang H, Chen Y, Chen J, Wu Y, Guo L, Zhu J, Zhao X, Peng T, Zhang Y, Chen S, Li X, Li D, Wang T, Pei D (2013). H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat. Genet.* **45**: 34-42.

Cheng K, Grisendi S, Clohessy JG, Majid S, Bernardi R, Sportoletti P, Pandolfi PP (2007). The leukemia-associated cytoplasmic nucleophosmin mutant is an oncogene with paradoxical functions: Arf inactivation and induction of cellular senescence. *Oncogene* **26**: 7391-7400.

Feng B, Jiang J, Kraus P, Ng JH, Heng JC, Chan YS, Yaw LP, Zhang W, Loh YH, Han J, Vega VB, Cacheux-Rataboul V, Lim B, Lufkin T, Ng HH (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat. Cell Biol.* **11**: 197-203.

Ge W, Martinowich K, Wu X, He F, Miyamoto A, Fan G, Weinmaster G, Sun YE (2002). Notch signaling promotes astrogliogenesis via direct CSL-mediated glial gene activation. J. Neurosci. Res. **69**: 848-860.

Hanna J, Markoulaki S, Schorderet P, Carey BW, Beard C, Wernig M, Creyghton MP, Steine EJ, Cassady JP, Foreman R, Lengner CJ, Dausman JA, Jaenisch R (2008). Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* **133**: 250-264.

Harrison DA (2012). The Jak/STAT pathway. *Cold Spring Harb. Perspect Biol.* **4**.

Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, Muhlestein W, Melton DA (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat. Biotechnol.* **26**: 1269-1275.

Johnson, WE, Rabinovic, A, and Li, C (2007). Adjusting batch effects in microarray expression data using Empirical Bayes methods. *Biostatistics* **8**: 118-127.

Katoh Y, Katoh M (2007). Comparative integromics on JMJD2A, JMJD2B and JMJD2C: preferential expression of JMJD2C in undifferentiated ES cells. Int. J. Mol. Med. **20**: 269-273.

Li C, Zhou J, Shi G, Ma Y, Yang Y, Gu J, Yu H, Jin S, Wei Z, Chen F, Jin Y (2009). Pluripotency can be rapidly and efficiently induced in human amniotic fluid-derived cells. *Hum. Mol. Genet.* **18**: 4340-4349.

Li H, Collado M, Villasante A, Strati K, Ortega S, Cañamero M, Blasco MA, Serrano M (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* **460**: 1136-1139.

Liu C, Yu C, Yuan K, Lou J, Ding M, Deng H (2008). Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell* **3**: 475-479.

Liao B, Bao X, Liu L, Feng S, Zovoilis A, Liu W, Xue Y, Cai J, Guo X, Qin B, Zhang R, Wu J, Lai L, Teng M, Niu L, Zhang B, Esteban MA, Pei D (2011). MicroRNA cluster 302-367 enhances somatic cell reprogramming by accelerating a mesenchymal-to-epithelial transition. *J Biol Chem* **286**:17359-173564.

Maekawa M, Yamaguchi K, Nakamura T, Shibukawa R, Kodanaka I, Ichisaka T, Kawamura Y, Mochizuki H, Goshima N, Yamanaka S (2011). Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* **474**: 225-229.

Maherali N, Sridharan R, Xie W, Utikal J, Eminli S, Arnold K, Stadtfeld M, Yachechko R, Tchieu J, Jaenisch R, Plath K, Hochedlinger K (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell* **1**: 55-70.

Marión RM, Strati K, Li H, Murga M, Blanco R, Ortega S, Fernandez-Capetillo O, Serrano M, Blasco MA (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* **460**: 1149-1153.

Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, Bernstein BE, Jaenisch R, Lander ES, Meissner A (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**: 49-55.

Morita S, Kojima T, Kitamura T (2000). Plat-E: an efficient and stable system for transient packaging of retroviruses. *Gene Ther.* **7**: 1063-1066.

Nakagawa M, Takizawa N, Narita M, Ichisaka T, Yamanaka S (2010). Promotion of direct reprogramming by transformation-deficient Myc. *Proc. Natl. Acad. Sci. USA* **107**: 14152-14157.

Okita K, Ichisaka T, Yamanaka S (2007). Generation of germline-competent induced pluripotent stem cells. *Nature* **448**: 313-317.

Oldham MC, Horvath S, Geschwind DH (2006). Conservation and evolution of gene coexpression networks in human and chimpanzee brains. Proc. Natl. Acad. Sci. U.S.A. **103**: 17973-17978.

Plews JR, Li J, Jones M, Moore HD, Mason C, Andrews PW, Na J (2010). Activation of pluripotency genes in human fibroblast cells by a novel mRNA based approach. *PLoS One*. **5**: e14397.

Subramanyam D, Lamouille S, Judson RL, Liu JY, Bucay N, Derynck R, Blelloch R (2011). Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat Biotechnol* **29**: 443-8.

Sun N, Panetta NJ, Gupta DM, Wilson KD, Lee A, Jia F, Hu S, Cherry AM, Robbins RC, Longaker MT, Wu JC (2009). Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **106**: 15720-15725.

Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr. Biol.* **11**: 1553-1558.

Takahashi K, Yamanaka S (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663-676.

Tsubooka N, Ichisaka T, Okita K, Takahashi K, Nakagawa M, Yamanaka S (2009). Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. *Genes Cells* **14**: 683-694.

Wang T, Chen K, Zeng X, Yang J, Wu Y, Shi X, Qin B, Zeng L, Esteban MA, Pan G, Pei D (2011). The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. *Cell Stem Cell* **9**: 575-587.

Yao S, Sukonnik T, Kean T, Bharadwaj RR, Pasceri P, Ellis J (2004). Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications. *Mol. Ther.* **10**: 27–36.

Ye Z, Zhan H, Mali P, Dowey S, Williams DM, Jang YY, Dang CV, Spivak JL, Moliterno AR, Cheng L (2009). Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. *Blood* **114**:5473-5480.

Zhao Y, Yin X, Qin H, Zhu F, Liu H, Yang W, Zhang Q, Xiang C, Hou P, Song Z, Liu Y, Yong J, Zhang P, Cai J, Liu M, Li H, Li Y, Qu X, Cui K, Zhang W, Xiang T, Wu Y, Zhao Y,

CHAPTER 4

Direct reprogramming of human Alexander's Disease patient-specific fibroblasts into induced astrocytes that can serve as a human *in vitro* disease model

4.1 Introduction

Yamanaka's work in iPS reprogramming took the stem cell field by storm, but a secondary effect of his contribution to the scientific community was a boost to direct reprogramming. Since his reported success of iPS reprogramming in mouse, other scientists have applied his straightforward protocol to reprogram fibroblasts into other cell types. This application of the technology is especially useful for disease modeling in cells that are normally difficult to obtain and study such as the diseases of the CNS. It is not surprising, therefore, that one of the first successes reported using the same approach was in mouse induced neurons (Vierbuchen et al., 2010) as well as dopaminergic neurons (Kim et al., 2011) and motor neurons (Son et al., 2011). All of these neurons are relevant in complex human neurological diseases and can be used for disease modeling. By simply taking fibroblasts of a patient, one could make disease-relevant cell types in a dish and can be a powerful approach to studying pathological mechanisms as well as therapeutic approaches.

Though much work has been done in reprogramming human cells into neurons of all kinds, the generally neglected glial subtype of neural cells like oligodendrocytes or astrocytes have yet to be reprogrammed into. Astrocytes have historically been relegated to simply support cells for neurons but more recently these support cells have been shown to be more important than original thought (reviewed in Molofsky et al., 2013). They have been shown to play a role in many neurodevelopmental and neurodegenerative diseases like Rett Syndrome, Fragile X, Alexander's Disease, and Amyotrophic Lateral Sclerosis (ALS) as well as others. Thus, it is important to develop induced astrocytes so that these diseases could be better studied. In fact, because a lot of neurological diseases have cell-autonomous as well as cell non-autonomous

effects, the ability to reprogram mutant fibroblasts to induced neurons or induced astrocytes or even both and then use them in a co-culture system is a great motivation for making induced astrocytes to study the dynamics and interplay between the two intimately linked cell types.

Alexander's Disease (AxD) is a unique neurological disease in that its pathology can be attributed to a genetic mutation in the GFAP gene specifically misregulated in astrocytes (Yoshida and Nakagawa, 2011). As such, any induced astrocytes could be used to make a human model of Alexander's Disease, which thus far has been mainly studied in mouse models. Since it is a fatal neurodegenerative disease, any models created to study it would be a boon to the biomedical research field and mouse models are not always sufficient to discover the mechanisms of the human version of the disease. Furthermore, any hopes of screening drug or protein therapies would be much more easily done in a cell model of the disease, especially when it comes to mutation specific effects and penetration problems.

To this end, I reprogrammed fibroblasts into induced astrocytes and showed that they are functional. Furthermore, I reprogrammed human wild-type and AxD patient-specific fibroblasts into iAs to make a human model of AxD astrocytes that responds to changes in GFAP levels as has been shown in the mouse models of AxD (Hagemann et al., 2009). The induced astrocyte reprogramming technology can also be applied to the study of other neurological diseases in the future.

4.2 Results

4.2.1 Candidate factors in reprogramming MEFs into induced-astrocytes

With the successful reprogramming of iPS by Yamanaka, iNs by Wernig, and other cell types from fibroblasts using the relatively straightforward approach pioneered by Yamanaka. In short, these groups infected fibroblasts with a cocktail of different viruses that ectopically express transcription factors that are important for the cell identity of the target cell type of interest. This is the approach I used to make induced astrocytes. I chose to start with mouse cells because mouse cells seem to be easier to handle, easier to reprogram, and faster to reprogram as shown by the success of iPS and iNs first in mouse. It has been shown that mouse fibroblasts reprogram into iPS (Takahashi and Yamanaka, 2006) much faster than human fibroblasts reprogram into iPS (Takahashi et al., 2007) when using the same factors. Furthermore, reprogramming human fibroblasts into iNs (Pang et al., 2011) require one more factor, NeuroD1, than required to reprogram mouse fibroblasts into iNs (Vierbuchen et al., 2010).

The experimental approach I used is quite simple and similar to that used by Yamanaka's group and Wernig's group (Figure 4.1A):

1. Dissect MEFs from E17.5 mouse embryos and expand them for a source of fibroblasts for reprogramming.
2. Infect the MEFs with a pool of retroviruses that express candidate transcription factors that are important for astrocyte identity and differentiation.
3. Culture the transduced cells in astrocyte media on polyorthinine and fibronectin coated plates to allow induced astrocytes to attach and survive.

4. Since MEFs only took 3 days to reprogram into Tuj1-positive iNs (Vierbuchen et al., 2010), I decided four days should be sufficient for the reprogramming to occur. After the 4 days of reprogramming, stain for the astrocyte marker Glial Fibrillary Acidic Protein (GFAP) to see if any MEFs reprogram into astrocytes. If no reprogrammed cells could be detected, then culture for longer and stain again at a later time.
5. Upon successful reprogramming, repeat protocol but with the same cocktail except less one transcription factor to test if any transcription factor is required or counterproductive. Keep doing this until the minimum number of transcription factors that are required are left.

The most critical step in the protocol is actually the decision of which factors to use to attempt to reprogram the cells. By choosing the right factors, reprogramming should be achievable. Also, since cell signaling is so important in neural development and astro-gliogenesis, activation of these signaling pathways or using factors that are the final transcriptional effector of these pathways would be extremely useful. Most successful attempts at reprogramming have used transcription factors that are important for the normal differentiation or physiological development of the target cell type. For example, Achaete-Scute Complex Homolog 1 (Ascl1) was shown to be critical in iN reprogramming and is a known factor in neuronal commitment and differentiation (Vierbuchen et al., 2010; Pang et al., 2011).

Successful reprogramming requires a good understanding of the target cell type and the developmental and differentiation process that leads to the target cell fate. Also,

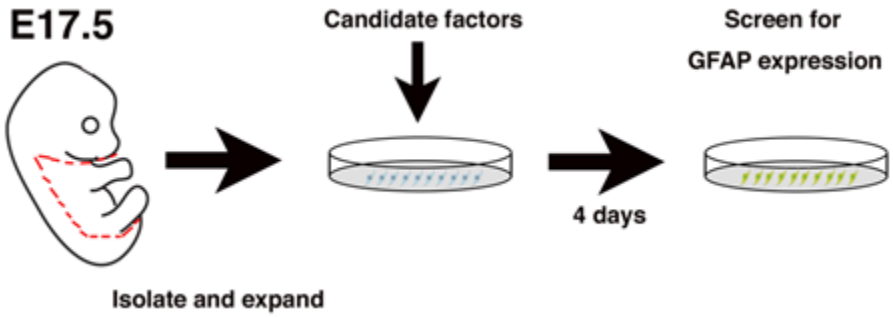

Figure 4.1 Reprogramming MEFs into induced Astrocytes. The protocol used is similar to that used by Yamanaka in reprogramming fibroblasts into iPS and Wernig in reprogramming fibroblasts into iNs.

transcription factors that are either specific for or is highly expressed in the target cell type make great potential candidate factors for reprogramming. Furthermore, for To that end, I chose seven such factors: Notch1, Nuclear Factor I/A (Nfia), Ten Eleven Translocation 2 (Tet2), RE1-Silencing Transcription Factor (Rest), the BMP signaling pathway, the thyroid hormone signaling pathway, and the Jak/Stat signaling pathway.

Notch1 has been heavily implicated as an important protein in neural development and differentiation, especially important is its role in suppressing neuronal and oligodendrocyte differentiation and promotion of the astrocyte lineage (Ge et al., 2002; Grandbarbe et al., 2003). This pro-astrocyte function of Notch1 is mediated by direct binding of the Notch intracellular domain (NICD) to the Gfap promoter and activating the gene. Since the full length Notch1 is a membrane protein that becomes cleaved and transported to the nucleus upon Notch signal activation, I only cloned the Notch1 intracellular domain (NICD), which is the domain that mediates transcriptional regulation.

The nuclear transcription factor, Nfia is actually downstream of Notch signaling activation and is a critical gene in gliogenesis as well as being necessary to suppress the neuronal cell fate (Deneen et al., 2006). Therefore, Nfia, especially in combination with NICD, would be a good candidate for reprogramming.

Tet2 has been recently implicated in the active DNA demethylation pathway (Guo et al., 2011) and with our lab's previous work showing that astrocyte differentiation is suppressed by DNA methylation (Fan et al., 2005), Tet2 seems like a likely candidate. Also, a recent unpublished finding in the lab is an upregulation of Tet2 during astrocyte

differentiation. Unfortunately, the Tet2 protein is actually quite large and the full length would not be able to be packaged in the viral particles. However, Tet2 is unique in the Tet family as the only protein in the family without a canonical CxxC DNA-binding domain and instead, it interacts via its catalytic domain with IDAX for genome targeting (Ko et al., 2013). As a result, the Tet2 catalytic domain (TET2CD) should be sufficient for reprogramming, and so I cloned that fragment into the retroviral backbone.

REST is a chromatin remodeling protein that has been associated with silencing neuronal genes in non-neuronal cells. However, its role has recently been expanded to include glial differentiation and cellular identity as well as gene activation in addition to its traditional role of silencing, especially when it comes to the astrocyte cell fate (Abrajano et al., 2009). Since REST has been traditionally associated with transcriptional repression in most contexts, I added the viral VP16 transactivation domain in order to have REST bind and activate transcription as that seems to be more pro-astrocyte.

$ALK3^{QD}$ is a constitutively active form of the Bone Morphogenetic Protein (BMP) type I receptor (also known as BMPR1A) that activates the BMP pathway (Shepherd et al., 2010). The BMP pathway via Smad1 is critically involved in switching neural fate from a neuronal one to an astrocytic fate (Nakashima and Taga, 2002) and therefore, constitutively active BMP signaling might be crucial to reprogramming.

Along with BMP, the Jak/Stat pathway has been shown to promote astro-gliogenesis and therefore, activating the Jak/Stat pathway via providing Leukemia Inhibitory Factor (LIF) to the cells should activate the Jak/Stat pathway and promote astro-gliogenesis. In fact, Jak/Stat has

been shown to act with Notch (Miller and Gauthier, 2007) and DNA methylation of the Gfap locus to regulate the onset of astrogliogenesis (Fan et al., 2005).

Finally, thyroid hormone processing is an important role of astrocytes and the balance of the levels of thyroid hormone receptor alpha 1 (THRA) and thyroid hormone receptor beta 1 (THRB) has been shown to modulate astro-gliogenesis (Morte et al., 2004). By expressing THRA fused with VP16 in fibroblasts and then providing the ligand, tri-iodothyronine (T_3), I hope to hyper-activate the transcriptional activation of the astro-gliogenesis genes that THRA normally turn on during development of astrocytes.

4.2.2 Reprogramming MEFs into GFAP-positive cells

The first experiment was done using the E17.5 MEFs with all seven candidate factors. The cells were allowed to reprogram for four days and then they were fixed and stained for GFAP. Amazingly, some cells stained positive for GFAP while the negative control (MEFs infected with pMXs-GFP) had no GFAP positive staining (Figure 4.2A). The efficiency was extremely low though, but in most previous reprogramming reports in the literature, the full cocktail of many factors give rise to a lower number of positive reprogramming events as some factors tend to be counterproductive. Since seven factors were able to reprogram, I next tried with the four transcription factors only (NICD, TET2CD, NFIA, and REST-VP16) since transcription factors seem to be the most important from previous reprogramming reports. With the four transcription factors, they also were able to reprogram successfully, showing GFAP staining (Figure 4.2A). With the success of the 4 factors, the next step was to try with all the combinations of 4 – 1 factors to see which factors were essential and which ones were counterproductive. By taking away NICD, TET2CD, or REST-VP16, the three combinations of

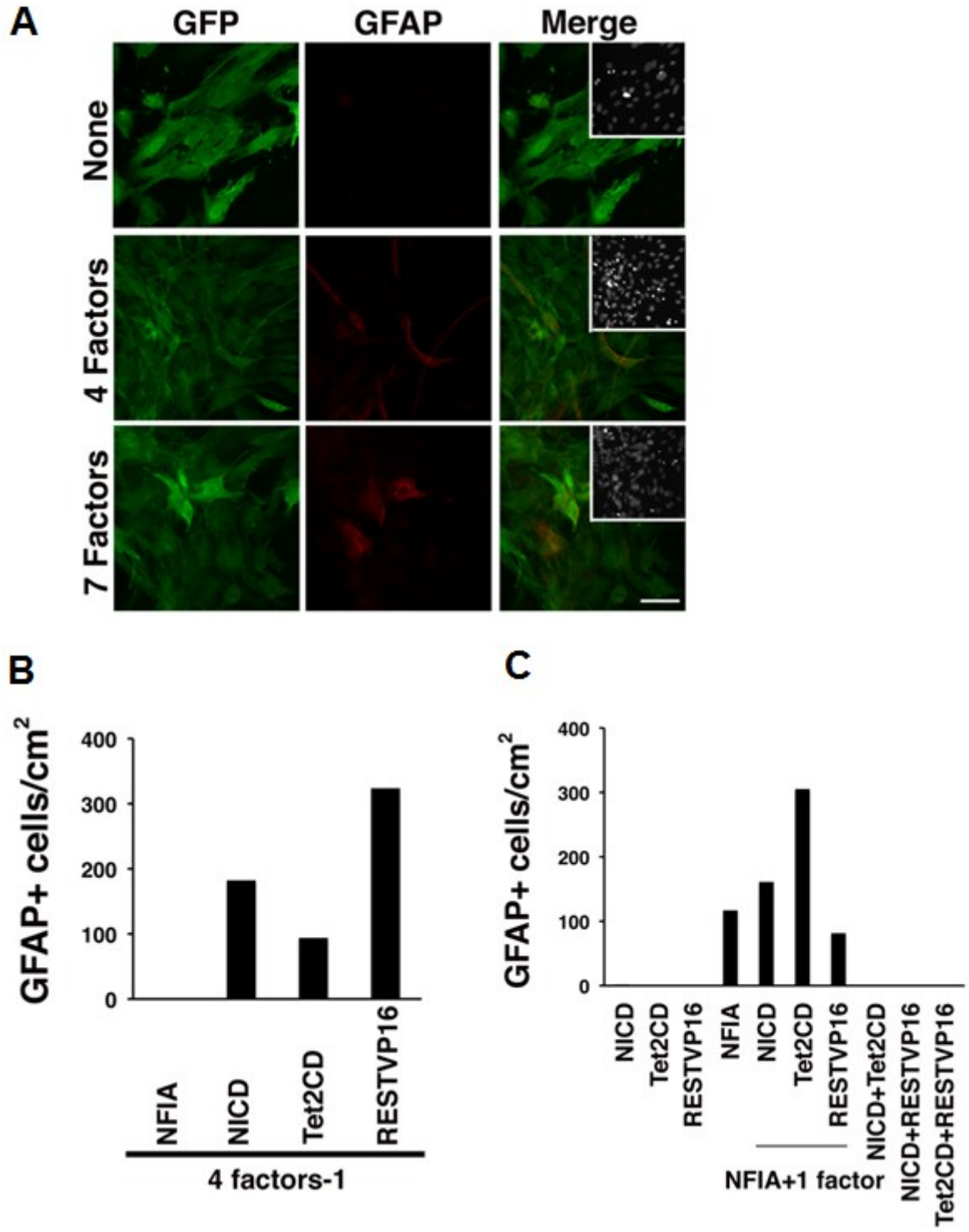

Figure 4.2 Narrowing down the required factors for iA reprogramming. (A) Both the combination of 7 factors and 4 factors gave rise to GFAP-positive cells. (B) Reprogramming with each combination of 4 factors minus 1. (C) Reprogramming with each of the four factors by themselves and all the combinations of 2 factors.

factors showed that none of those factors were completely necessary for the reprogramming to occur. Any combination that involved NFIA seemed to successfully reprogram. However, NFIA seemed to be necessary for reprogramming to occur as taking that away and only using only NICD, TET2CD, and REST-VP16 together was not enough to reprogram any cells to become GFAP positive (Fig 4.2B).

Since NFIA seemed to be the most important factor, I tested to see if NFIA was both necessary **and** sufficient to reprogram MEFs into GFAP positive cells. In addition, to show that NFIA was unique in this possibility, I also tested each of the factors individually and in combinations of 2 without NFIA. Surprisingly, NFIA by itself could successfully reprogram MEFs, but the other factors either in combination or by themselves could not reprogram the MEFs as long as NFIA was not present (Figure 4.2C). Though NFIA was necessary and sufficient to reprogram the MEFs, reprogramming could possibly be enhanced by the other transcription factors. The next step was to test if reprogramming with NFIA in combination with each of the other transcription factors. As shown in Figure 4.2C, by combining NFIA with NICD or TET2CD, reprogramming efficiency was able to be increased. Tet2CD seemed to have a stronger effect than NICD. Interestingly, REST-VP16 actually decreased the reprogramming efficiency of NFIA when used in combination. Since NFIA was sufficient to reprogram the MEFs into GFAP-positive cells, I decided to proceed with further experiments using only NFIA for reprogramming.

4.2.3 GFAP-positive cells express genes typical of astrocytes

The first step in characterizing these GFAP-positive cells was to make sure that they were no longer fibroblasts and that they had other astrocyte markers expressed. MEFs express alpha

smooth muscle actin (SMA, also known as ACTA) and it is considered a general fibroblast marker expressed in a variety of different types of fibroblasts. Figure 4.3A shows that the control fibroblasts do not express GFAP but express SMA. In contrast, the cells transduced with NFIA do not express SMA, but express GFAP. In fact, the two markers are generally mutually exclusive in the cultures. Furthermore, cells that were transduced with NFIA were stained for GLAST or SLC1A3, which is a glutamate transporter that is highly expressed in astrocytes. The reprogrammed cells had ~25% GLAST-positive staining while control cells that were only transduced with GFP had 0% GLAST-positive staining (Figure 4.3B). Although GFAP is the most commonly used general astrocyte marker, it has been shown to be expressed in NPCs (Garcia et al., 2004) and to be heterogeneously expressed across different types of astrocytes with fibrous astrocytes having high expression and protoplasmic astrocytes having low expression (Cahoy et al., 2008). ALDH1L has been suggested as a superior marker that is more specific to astrocytes of all types. Co-staining with both GFAP and ALDH1L show very good overlap of co-expression and quantification shows that there is no cells in the control fibroblasts that express ALDH1L while in the NFIA transduced cells, ~10% of the cells express ALDH1L (Figure 4.3C). Since immunofluorescence shows that the GFAP cells are also positive for ALDH1L, they are most likely astrocytes. To further confirm their identity, I ran qPCR on the cells along with a positive control of dissected astrocytes from P1 mice. As Figure 4.4 shows, the MEFs transduced with NFIA express GFAP, GLAST, GLT1, and SPARC, which are also all expressed by astrocytes but not the fibroblasts transduced with pMYs-IRES-GFP. Interestingly, GFAP expression levels are a small fraction (~2000 fold less) of control astrocytes, but as shown by Cahoy et al., this is not out of the ordinary as different astrocytic types have different levels of GFAP expression and these astrocytes could very well be low GFAP-expressing in relation to P1

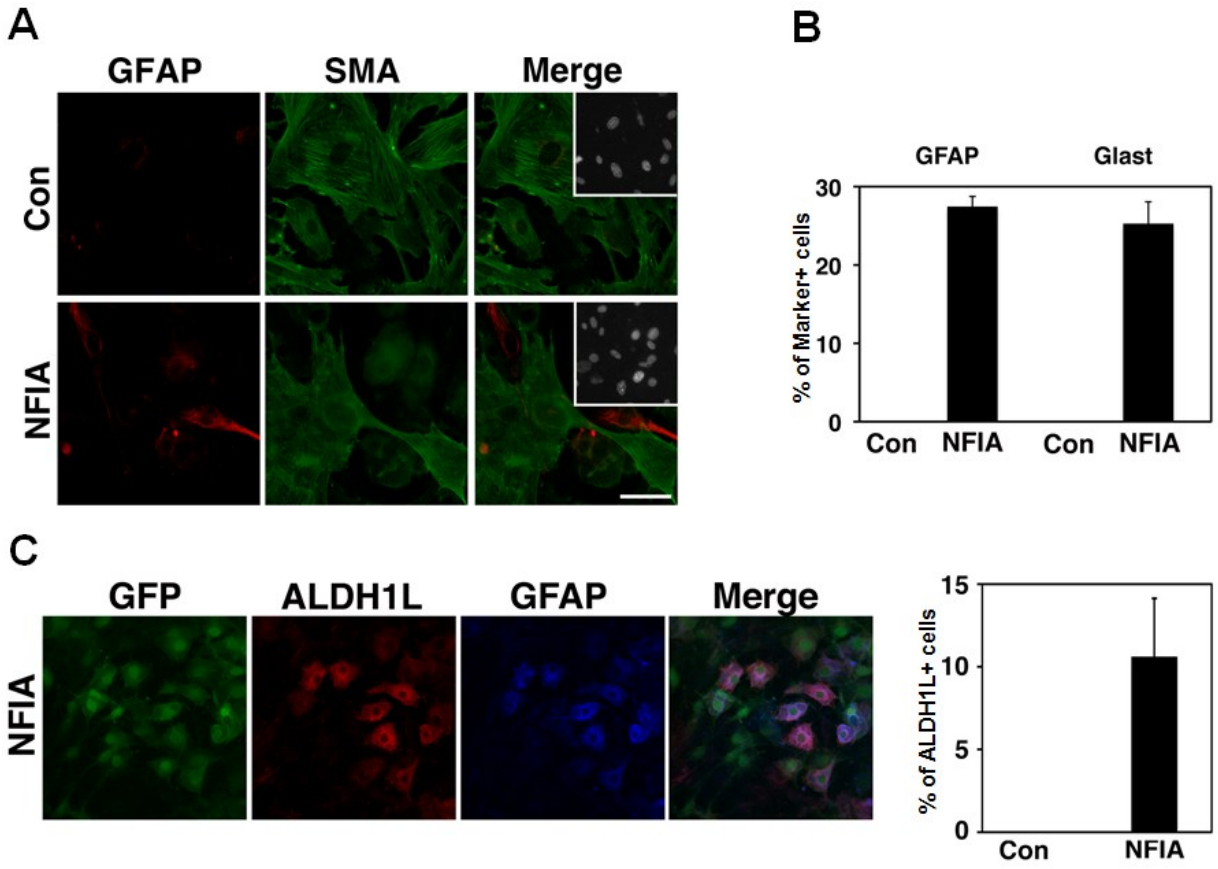

Figure 4.3 Fibroblast and astrocyte marker immunofluorescence. (A) Cells were stained with GFAP and SMA. Control is uninfected while NFIA are transduced with the NFIA virus. (B) Cells transduced with NFIA-IRES-GFP and control cells that was transduced with only IRES-GFP was stained for GFAP or GLAST and quantified. (C) Cells transduced with NFIA-IRES-GFP virus were assessed for GFP, ALDH1L, and GFAP expression by immunofluorescence. Quantification was done on control cells that were transduced with only GFP and cells transduced with NFIA-IRES-GFP virus.

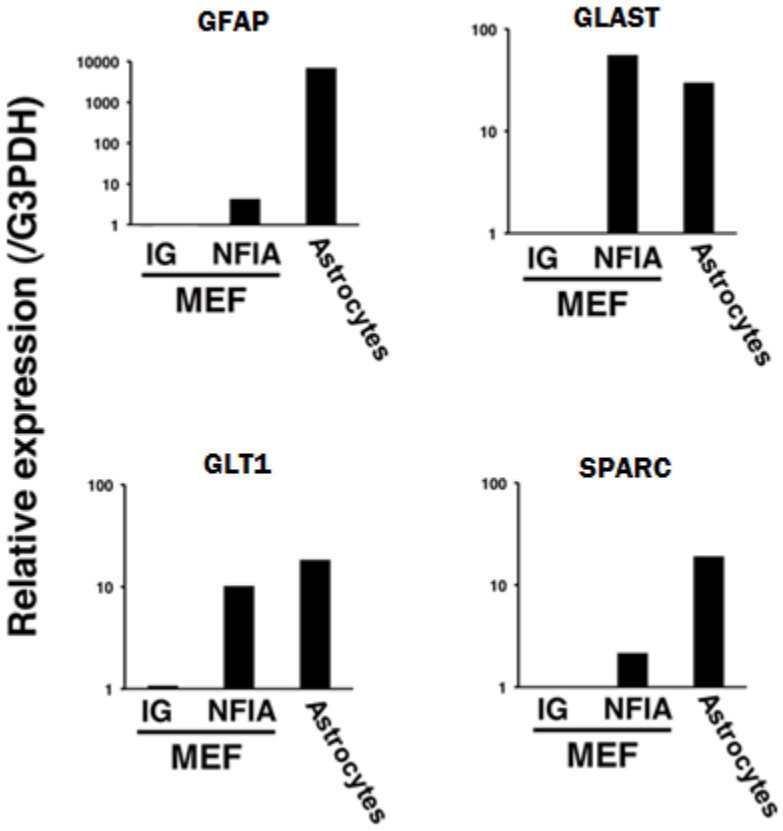

Figure 4.4 qPCR of Astrocyte markers. Relative expression of different astrocyte markers compared to the housekeeping gene, Glyceraldehyde 3-phosphate dehydrogenase (G3PDH). IG is the negative control cells which were transduced with IRES-GFP. NFIA were the cells transduced with NFIA-IRES-GFP. Astrocytes are mouse astrocytes dissected from P0 mice.

mouse astrocytes. Since the control MEFs have no expression of GFAP whatsoever, it is obvious that the cells transduced with NFIA are very different from MEFs.

Although there was successful reprogramming, the GFAP-positive cells were only about 10% or less of the total number of cells. Furthermore, the MEFs are able to grow in the 10% FBS astrocyte medium while the GFAP-positive cells generally did not grow. As a result, the culture would eventually be taken over by the MEFs. In order to alleviate this problem and to achieve a higher purity of GFAP-positive cells, I cloned an Internal Ribosomal Entry Site (IRES) sequence followed by the Puromycin resistance gene and added it to the 3' end of the NFIA gene in the pMXs vector to create pMXs-NFIA-IRES-Puro (NFIA-IP). Using this as the vector to reprogram, I was able to select for the cells that were transduced with NFIA and more likely to reprogram by adding puromycin into the growth media. With this improvement, the culture of the GFAP-positive cells was stable with no overgrowth of MEFs.

4.2.3 GFAP-positive cells support dendritic growth and synaptic formation of neurons

One of the key functions of astrocytes is their ability to support neuronal growth and maturation. The classic test of astrocyte function is to test if co-culture of these GFAP-positive cells can promote dendritic outgrowth. Cortical neurons were dissected from E14 mice and co-cultured with the GFAP-positive cells or cultured without any support cells. They were then stained with the neuronal marker MAP2 to visualize the dendrites of the neurons. The average dendrite length per neuron was then measured and compared between control and co-culture with GFAP-positive cells. As seen in Figure 4.5A, GFAP-positive cells do indeed support dendritic growth as those cultured with the GFAP-positive cells had, on average, almost seven times longer dendrites than the neurons grown without support cells.

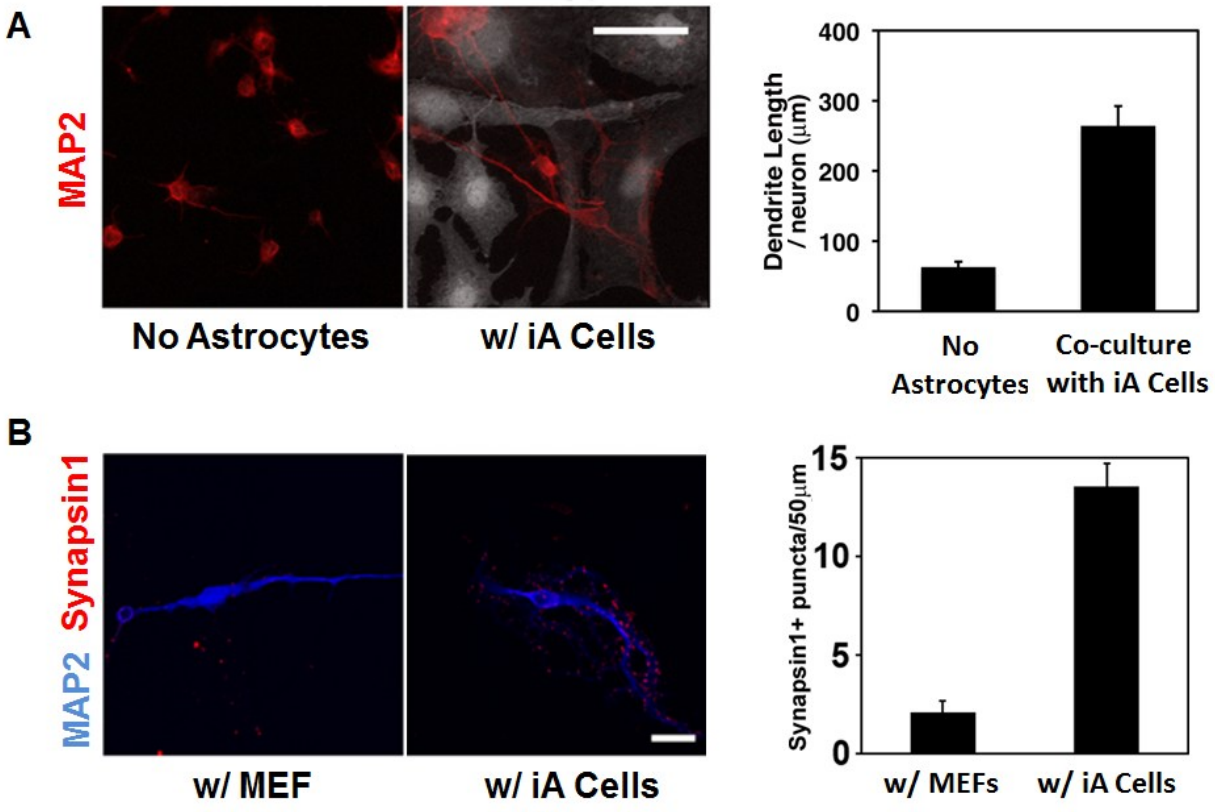

Figure 4.5 Neuronal co-culture. (A) Cortical neurons dissected from P0 mice were either co-cultured with no astrocytes or with iA cells. Average dendrite length per neuron was measured by staining processes with MAP2 antibody. (B) Cortical neurons dissected from P0 mice were either co-cultured with MEFs or with iA cells. The cultures were stained with MAP2 to visualize dendrites and Synapsin1 to visualize synaptic connections. The average Synapsin1 punctate per 50 μm was measured.

Besides supporting neural growth and maturation, astrocytes also function to promote synaptic formation between neurons. E14 cortical neurons were dissected and co-cultured with either MEFs or the GFAP-positive cells. They were then stained with MAP2 to visualize the neurons and their dendrites. In addition, the cultures were stained with Synapsin1 to visualize the synaptic vesicles in the neurons. Average Synapsin1 punctate density was counted and the GFAP-positive cells did indeed increase the number of Synapsin1 punctate of the stained neurons by about 6.5 fold (Figure 4.5B). Because these cells express astrocyte markers and can support neuronal function, they can be considered bona fide induced Astrocytes (iAs).

4.3.1 HFF can be reprogrammed into iAs using three factors

Since MEFs were shown to be able to be reprogrammed into mouse iAs, I wondered if human fibroblasts could be reprogrammed to iAs as well using a similar approach. Previous reprogramming studies have shown that transcription factor combinations that reprogram MEFs can also reprogram human fibroblasts, although in the case of iNs, human fibroblasts required an additional factor compared to MEFs—NeuroD1. Also, the reprogramming took longer in the human cells compared to the mouse cells, which is typical of other human reprogramming experiments such as for human iPS reprogramming (Takahashi and Yamanaka, 2006; Takahashi et al., 2007). For human fibroblasts, I used the HFFs from ATCC. GFAP staining was once again used as the test for successful reprogramming, but was tested after a week to account for possibly longer reprogramming times needed.

Since NFIA was both necessary and sufficient to reprogram MEFs into mouse iAs, the most obvious experiment was to test if NFIA was also necessary and sufficient to reprogram HFFs into human iAs. Unlike in mouse, NFIA was unable to reprogram HFFs into iAs by itself

(Figure 4.6). Since NFIA was insufficient to reprogram HFFs into iAs, I tested to see if the other transcription factors, Tet2CD and NICD, individually could reprogram HFFs, but not surprisingly, neither of the other two transcription factors individually could reprogram HFFs successfully. So it seemed that for HFFs, there would be a combination of factors required to reprogram them into iAs, unlike the MEFs that only required NFIA. Since mouse iA reprogramming efficiency increased when using NFIA in combination with either Tet2CD or NICD, I thought that the three of the factors together or a combination of NFIA with either additional factor might be sufficient to reprogram HFFs into iAs. As Figure 4.6A shows, in fact, all three factors must be used together in order to have reprogramming succeed. With the lack of any of the three, no GFAP-positive cells arise.

To ensure that the human iAs, like the mouse iAs, no longer express SMA, I stained the human iAs with SMA antibody and found that they no longer express SMA, unlike the control fibroblasts (Figure 4.6B). Just like the mouse iAs it was found that SMA was mutually exclusive with GFAP. Thus, it is possible to reprogram human fibroblasts into iAs, though requiring three factors (TET2CD and NICD in addition to NFIA) rather than just NFIA as in MEFs reprogramming. Furthermore, the reprogramming took a full week rather than four days.

4.3.2 Human iAs can be used as a model of Alexander's Disease *in vitro*

Reprogramming technology is especially useful for modeling diseases in the pathological cell type when that cell type is difficult to obtain or rare. In the case for astrocytes, they are difficult to obtain and one human disease that has been shown to manifest in astrocytes is Alexander's Disease (AxD). Fortunately, the Coriell Cell Repository has human fibroblasts from a patient with a mutation in the early onset form of Alexander's disease where a point mutation in

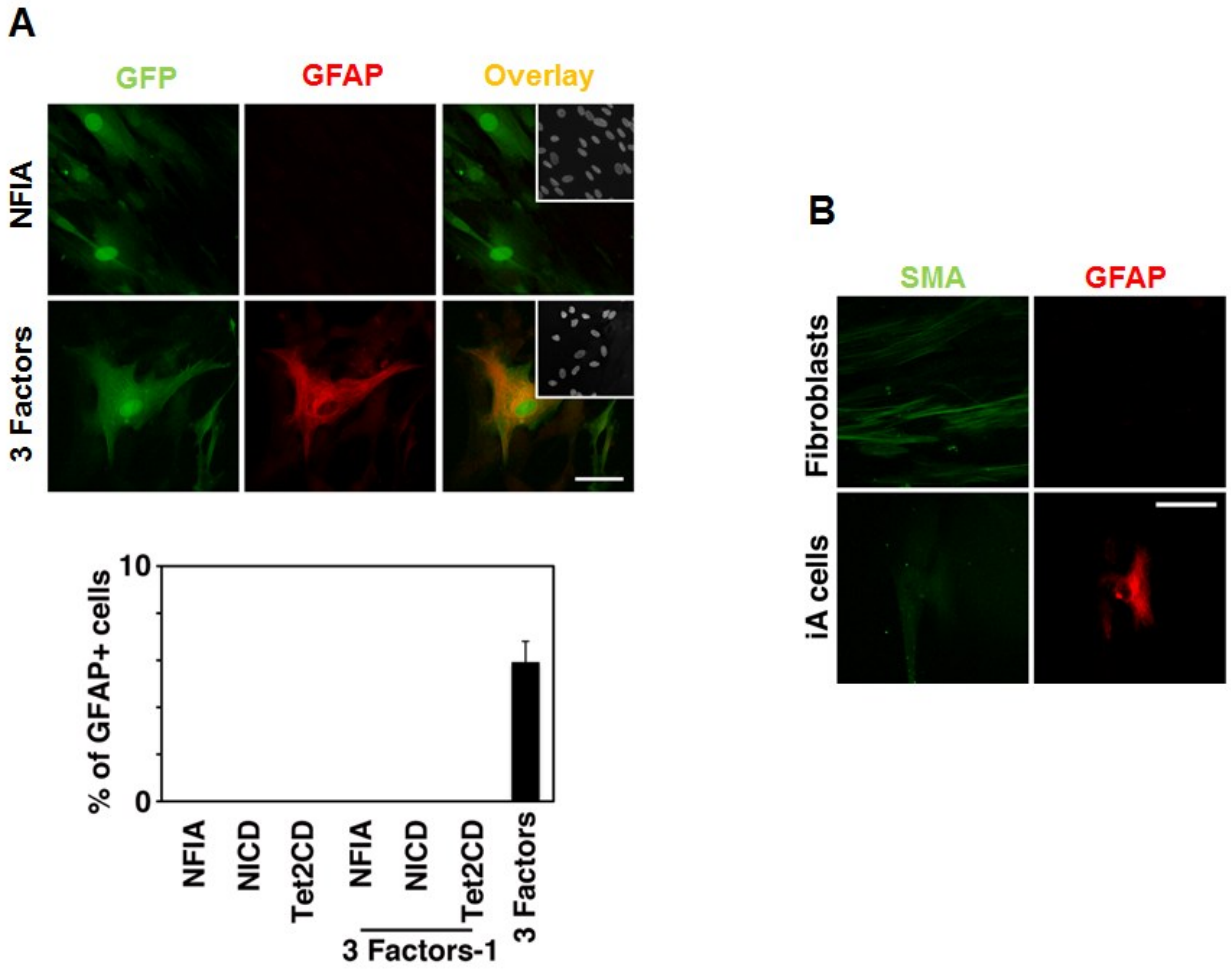

Figure 4.6 Human iA reprogramming. (A) NFIA was used alone or in combination with NICD and/or TET2CD for reprogramming into human iAs. GFAP staining was used as reporter. (B) HFF and human iA cells were stained for SMA and GFAP.

the GFAP gene causes arginine 369 to a cysteine (R239C mutation, Figure 4.7). This mutation has been shown to be a dominant negative form which causes the GFAP intermediate filament molecules to aggregate and form what's known as Rosenthal fibers—the hallmark of Alexander's Disease (Tang et al., 2006; Tang et al., 2010). Using these fibroblasts available from Coriell, I could show that these human iAs would be a good candidate for modeling the disease.

After obtaining the cells from Coriell, I first tested to make sure that the cells did, in fact contain the C715T mutation in the Gfap gene such that the protein will have a R239C mutation. Fortunately, the C715T mutation actually introduces a BglII site into the locus, so I designed primers that flanked that area in Exon4 and PCR amplified the fragment from genomic DNA obtained from control HFF cells as well as the AxD fibroblasts. Afterwards, I digested the PCR fragment with BglII to reveal that indeed, there was the C715T mutation and the AxD fibroblasts had the additional BglII site compared to the control (Figure 4.7).

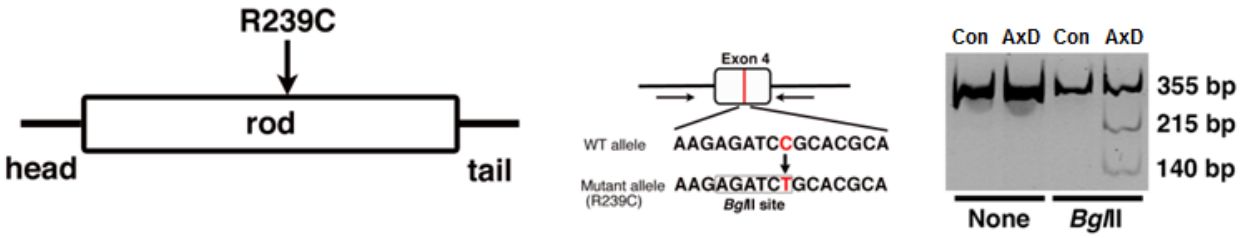

Figure 4.7 AxD patient-specific fibroblasts. The R239C mutation occurs in the rod domain of the intermediate filament protein GFAP. This point mutation is caused by a C715T mutation in exon 4 of the gene, which introduces a BglII site. Presence of the mutation can be tested by PCR amplifying the local fragment and digesting with BglII.

Having verified that the patient-specific fibroblasts had the mutation as expected, I proceeded to test if reprogramming these fibroblasts could be used as an *in vitro* disease model of Alexander's Disease. One of the most distinguishing phenotypes of an astrocyte afflicted with the R239C mutation is GFAP aggregates that take on the form of Rosenthal fibers so showing that iAs from these patient-specific fibroblasts had that phenotype would lend credence to iAs being able to be useful for disease models. It was reported that U251 cells, a GFAP-positive cell line derived from a malignant human astrocytoma tumor which are commercially available, when overexpressing the R239C mutation displayed such a phenotype (Tang et al., 2010). As such, I could use it as a great positive control for the desired phenotype. To this end, I cloned the R239C mutant form of GFAP into the pMYs-IRES-GFP vector to be used to package retrovirus to infect U251 cells. As negative controls, I would use the normal HFFs reprogrammed into iAs as well as the U251 cells transduced with IRES-GFP rather than the mutant GFAP. I cannot use U251 cells transduced with wild-type GFAP as a control as that would actually also cause aggregation and would serve as a poor negative control (Koyama and Goldman, 1999).

I reprogrammed the AxD fibroblasts and the HFFs at the same time using the 3 factors (NFIA-IP, TET2CD, and NICD) and allowed them to reprogram for a week as normal. I also transduced the U251 control cells with either GFP or the R239C-GFAP- IRES-GFP. Afterwards, I fixed the AxD iAs, the HFF iAs, and the U251 control cells and stained for GFAP to see the differences between the AxD iAs and the three control cell types. The contrast in the staining of GFAP in iAs was striking. In the AxD iAs, many different GFAP aggregate types could be found including dense Rosenthal fibers, dot aggregation, as well as what appeared to be large inclusion bodies. Most of the wild-type HFF iAs looked normal and fibrous, though a small percentage did have some inclusion bodies (Figure 4.8A). The number of cells with aggregates

or densed fibers was counted in the HFF iAs and the AxD iAs. The HFF iA controls only had about 10% of the cells with inclusion bodies while 38% of the AxD iAs displayed GFAP aggregates. Even more impressive was the complete lack of dense Rosenthal fibers in the HFF iAs while close to 15% of the AxD iAs showed the Rosenthal fiber phenotype (Figure 4.8B). The positive control U251 cells transduced with the mutant R239C GFAP showed a similar phenotype to that of the AxD cells, though with a smaller percentage of aggregates (28%) and a higher percentage of Rosenthal fibers (41%). This difference could be attributable to the fact that the U251 cells were overexpressing the R239C mutant while the AxD iAs were likely expressing the R239C mutant at a much lower level. The differences in the ratios of wild-type to mutant protein could cause a shift in the aggregation dynamics as hinted in the work of the Goldman lab (Koyama and Goldman, 1999; Tang et al., 2006; Tang et al., 2010). In comparison, the negative control U251 wild-type cells transduced only with GFP were similar to the normal HFF iAs (Figure 4.8B). The presence of Rosenthal fibers and the increase of aggregates in the Alexander's Disease patient-specific induced astrocytes form strong supporting evidence that the mutant AxD iAs can serve as a human cell model for Alexander's Disease.

Besides the formation of Rosenthal fibers, Alexander's diseased astrocytes have also been reported to have decreased level of GLT1 (Slc1a2), a glutamate transporter vital to astrocytes' role in the CNS (Tian et al., 2010). Moreover, the aggregation and misfolding of GFAP causes the upregulation and accumulation of alpha-B crystallin (CRYAB), a heat-shock protein that tries but ultimately fails to ameliorate the GFAP aggregates by recruitment to proteasomes in AxD astrocytes (Iwaki et al., 1989; Der Pang et al., 2006). Since alpha-B crystallin is not normally upregulated unless misfolded intermediate filaments are present, it would be relatively easy to see upregulation of the protein in AxD iAs. Therefore, I stained the AxD iAs and control

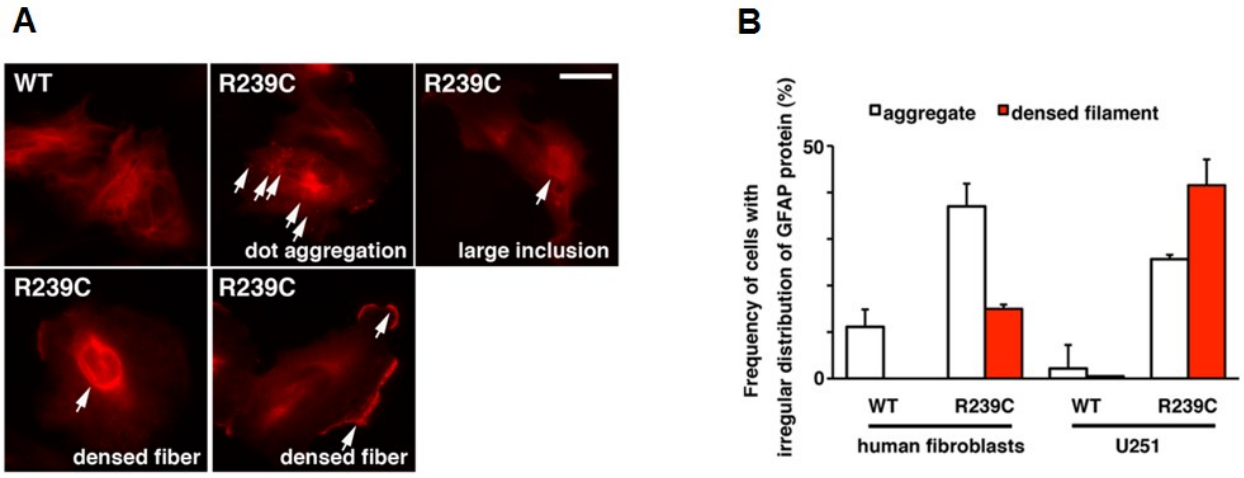

Figure 4.8 GFAP aggregates in AxD iAs. (A) GFAP staining of Wild-type (WT) HFF iAs displayed mostly loose fiber staining, as expected, while a fraction of the AxD iAs (R239C) stained for GFAP displayed dot aggregation, dense Rosenthal fibers, and large inclusion bodies. (B) Quantification of GFAP aggregate and densed fiber staining in HFF iAs (WT human fibroblasts), AxD iAs (R239C human fibroblasts), U251 astrocytoma cells transduced with IRES GFP (WT U251), and U251 astrocytoma cells transduced with R239C-GFAP-IRES-GFP (R239C U251).

HFF iAs for CRYAB and GFAP to look for upregulated and accumulated CRYAB protein in the GFAP-positive cells. Undeniably, the control iAs had no detectable levels of CRYAB and GFAP double positive cells while around 50% of the AxD iAs had positive CRYAB staining (Figure 4.9A). To see if the GLT1 expression level reduction phenotype is also recapitulated in the AxD iAs, I also stained the AxD iAs and the control HFF iAs for GLT1 and GFAP. I then counted the number of GLT1 and GFAP double positive cells relative to all GFAP-positive cells. As seen in Figure 4.9B, the AxD iAs had a greatly decreased number of double positive cell percentage (10%) compared to that of the HFF iAs (40%). To ensure that the lower number of double positive cells was not caused by the relative low sensitivity of the immunofluorescence technique, I also ran quantitative real-time PCR against GLT1 expression levels and normalized to the GFAP expression levels to ensure that the mixed population did not skew the data. And indeed, the GLT1 expression levels were decreased as measured by qPCR (~11 fold decrease).

Since the AxD iAs showed increased GFAP aggregation with the formation of Rosenthal fibers as well as increased expression of CRYAB and decreased expression of GLT1 while the control wild-type HFF iAs generally lacked any of these phenotypes, the AxD iAs are likely a good human cell model for Alexander's Disease. So far, only mouse models have been created for Alexander's disease, so this human cell model would be of great use in studying the human pathology of the disease.

4.3.3 Repairing the Alexander's Disease mutation in the iA model using gene editing

Because the Alexander's Disease model is a cell model, the model can be used to test if repairing the R239C mutation in the GFAP gene would ameliorate the disease phenotypes when the fibroblasts are converted into iAs. Successfully ameliorating the disease would be a proof of

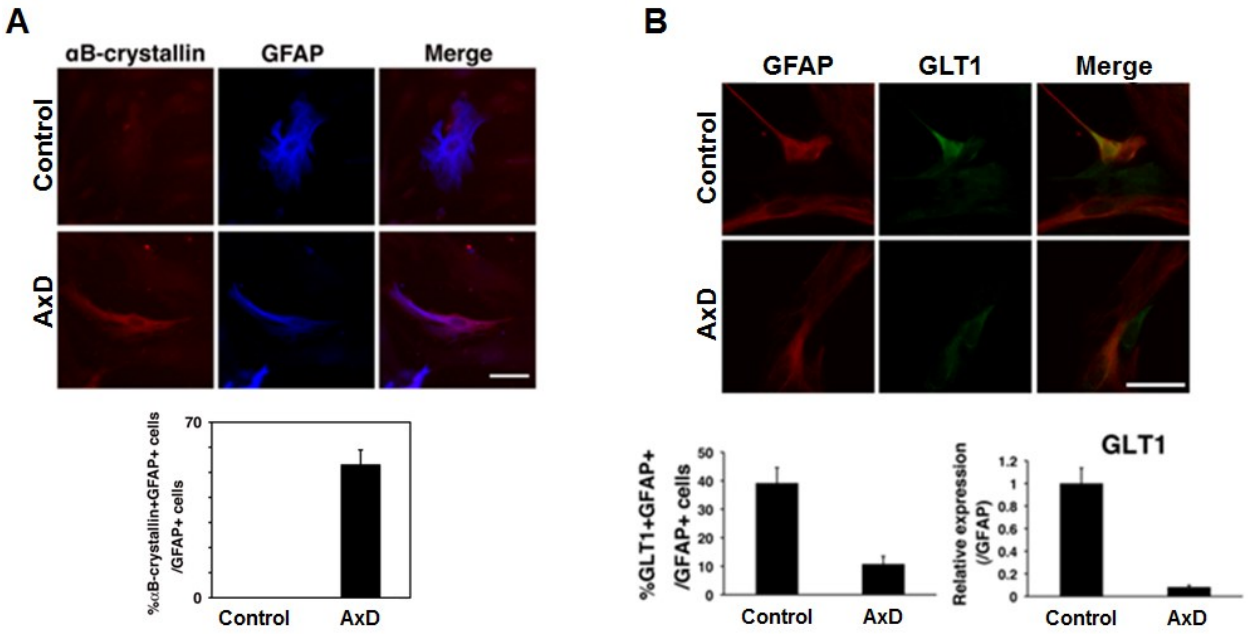

Figure 4.9 Increased CRYAB and decreased GLT1. (A) GFAP and αB-crystallin staining in HFF iAs (Control) and AxD iAs (AxD) show AxD iAs with strong αB-crystallin levels. GFAP and αB-crystallin double-positive cells were counted as a percentage of total GFAP-positive cells (B) GFAP and GLT1 staining in HFF iAs (Control) and AxD iAs (AxD) show AxD iAs with lowered GLT1 levels. GFAP and GLT1 double positive cells were counted as a percentage of total GFAP cells. qPCR was run on GLT1 relative to GFAP expression in HFF iA and AxD iA cultures.

principle that this model could be used to study therapeutic strategies in treating Alexander's Disease. The mutation in the AxD fibroblasts I used was single nucleotide change that was from a cytosine to a thiamine in the 715th nucleotide (C715T) which results in a single amino acid change (R239C). An elegant method to repair the mutation would be to use gene editing technology to replace the C715T mutation with the corrected cytosine so that the AxD fibroblasts would have a wild-type genotype again. Indeed, this would also create the ideal isogenic control for future experiments with the model.

A recent innovation in the field of gene editing was the so called TALEN (Transcription Activator-like Effector Nucleases) technology which allows for specific recognition of custom DNA sequences and subsequent cleavage in order to create a double-strand break (Reviewed in Mussolino and Cathomen, 2012). Normally, the cell will then undergo Non-Homologous End Joining (NHEJ) or Homology Directed Repair (HDR) to try to repair the double stranded break. NHEJ normally introduces indel mutations that are not desired for my experiment and therefore, I had to pursue HDR repair. HDR repair requires the introduction of a template so that the cell can copy the sequence and incorporate the desired new fixed sequence to repair the double strand break site introduced by the TALENs. Therefore, I had to design and clone the two flanking TALEN constructs that specifically recognizes the mutant site and not the wild-type sequence. Afterwards, I needed to transfect the constructs along with a linearized template to try to induce the homology directed repair machinery.

Unfortunately, there are many technical hurdles that I must overcome in order for the gene editing approach to succeed. Primary amongst these is the fact that induced-astrocyte cells do not proliferate so I could not directly use the TALEN gene editing machinery in the iAs since

any edited iAs would be impossible to purify, expand, and detect the nucleotide change. However, human fibroblast cultures do contain a stem cell population called MUSE cells that do proliferate, but a typical fibroblast culture only contains around 1% MUSE cells (Kuroda et al., 2010). Therefore, any gene editing experiments would have to be done in fibroblasts so that it can be purified and expanded before reprogramming into iAs. Another major problem is that human fibroblasts seem to be difficult to transfect. I tested multiple cationic-lipid-based and electroporation-based protocols and reagents, but could only achieve approximately 2-5% transfection rates. Because the template for HDR would have to be introduced by transfection, this poses a great problem. Finally, the general recombination rate of successful HDR using TALEN technology is actually around 1-16%, depending on transfection efficiency of the template (Miller et al., 2011; Ding et al., 2012) Therefore, I would have an efficiency of successful HDR repair occurring in a proliferative MUSE cell of approximately 0.0002% to 0.008% or less than 1 cell to 24 cells in 300,000 fibroblasts (the typical number of cells in a single confluent well of a 6-well plate). In other words, there would be an extremely low possibility of success with the approach as designed.

Rather than using a simple single stranded oligonucleotide as reported by the Cowan group (Ding et al., 2012), I decided that using a selection based strategy would increase the chances of success. I made a template that included positive-negative selection such that when homology directed repair was successful, the neomycin resistance cassette would be incorporated into the intron upstream of exon 4 of the GFAP gene near the mutation site and I could select the cells using G418 to find successful repair events (Figure 4.10). Any mistakes in recombination would result in the integration of the diphtheria toxin A subunit (DTA) which would kill the cell. Furthermore, upon successful isolation of clones that had the repaired genome, two flanking loxP

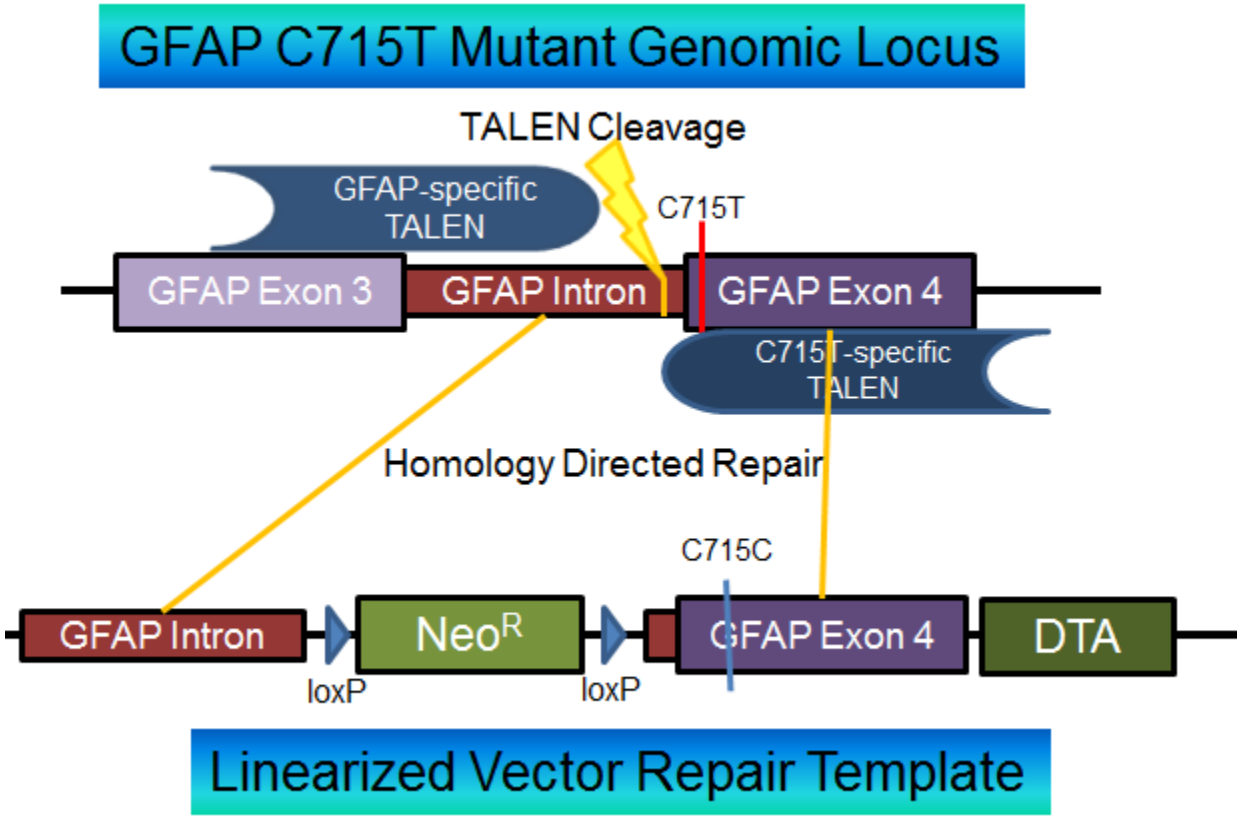

Figure 4.10 Homology Directed Repair of R239C GFAP mutation. Positive-negative selection based repair of the R239C GFAP mutant locus (C715T) in AxD patient-specific fibroblasts using TALENs to recognize the mutant sequence specifically.

sites would allow CRE-mediated recombination and removal of the neomycin cassette to ensure less genomic alterations.

Despite my best efforts, the combination of extremely low transfection rates with a tiny population of MUSE cells and finally relatively low homologous recombination rates, I was unable to isolate any successfully gene edited clones. Selection with G418 did have some cells that survived, indicating successful homologous recombination in a subset of the cells, but they were non-proliferative and most likely not MUSE cells. Therefore, I was unable to propagate the cells to test for their genotype and subsequent reprogramming into iAs. Even with the selection strategy, the success rate was too low to continue with the experiment.

4.3.4 Ameliorating the Alexander's Disease phenotypes in the iA model using shRNA

Though gene editing would have been the preferred method of repairing the AxD iAs, a simpler to accomplish method would be to simply knockdown the mutant form of GFAP in the AxD iAs using short hairpin RNAs (shRNAs). Since the R239C mutant form of GFAP is a dominant negative mutation, knocking down the expression of the mutant GFAP could ameliorate the disease phenotypes. However, knocking down the wild-type version would prevent identifying the iAs as successfully reprogrammed and would call into question their identity as astrocytes. Therefore, I needed to design my shRNA to be relatively specific to the mutant form. I designed 3 shRNA hairpins along with a scrambled shRNA. Two of the hairpins, named P16 and P13, target the mutant site by incorporating the C to T mutation into the 16^{th} or 13^{th} position of the shRNA sequence, which has been shown to be generally able to distinguish single nucleotide polymorphisms (Schwarz et al., 2006). The third design, named P1314, incorporates a second mutation so that the shRNA would have two mismatches with the wild-

type GFAP and only one mismatch with the mutant R239C GFAP. The two mismatched design is expected to have lower knockdown but more specificity compared to the two single mismatch designs (Geng and Ding, 2008).

To test these shRNA designs, I would need cells that express the mutant form of GFAP and the wild-type form of GFAP as well. Unfortunately, I did not have any human astrocytes expressing the R239C mutant form of GFAP. However, I did have the R239C mutant GFAP and wild-type GFAP genes cloned into the retroviral pMYs-IRES-GFP backbone. I used these constructs and transfected the plasmids along with the shRNAs into HEK293T cells to test the knockdown efficiency of each construct. After tweaking the levels of mutant and wild-type GFAP transfection levels a little, I was able to see that knockdown of GFAP and specificity varied across the designs (Figure 4.11A). P13 had the most efficient knockdown (around 40% knockdown), but was very non-specific. On the other hand, P1314 had lower overall knockdown, but was more specific to the mutant GFAP (25% knockdown in mutant only). Finally, P16 didn't really knockdown at all. Therefore, I decided to proceed with both P13 and P1314 to see if either of the two would be able to help ameliorate the AxD phenotypes in the iAs.

With the relatively poor shRNA knockdown levels (maximal of 40% knockdown), I decided to test out reprogramming AxD fibroblasts with the shRNA cassettes integrated via lentivirus using the Lenti-X shRNA2 construct from Clontech. After packaging and infecting the AxD fibroblasts with the shRNAs, I checked infection rates using the zsGreen fluorescent protein that is integrated as part of the pLenti-X system. One hundred percent of the cells were green and therefore, I proceeded to use these cells to transduce the 3 factors required to reprogram into iAs. After the normal one week of reprogramming, I stained the cells with GFAP

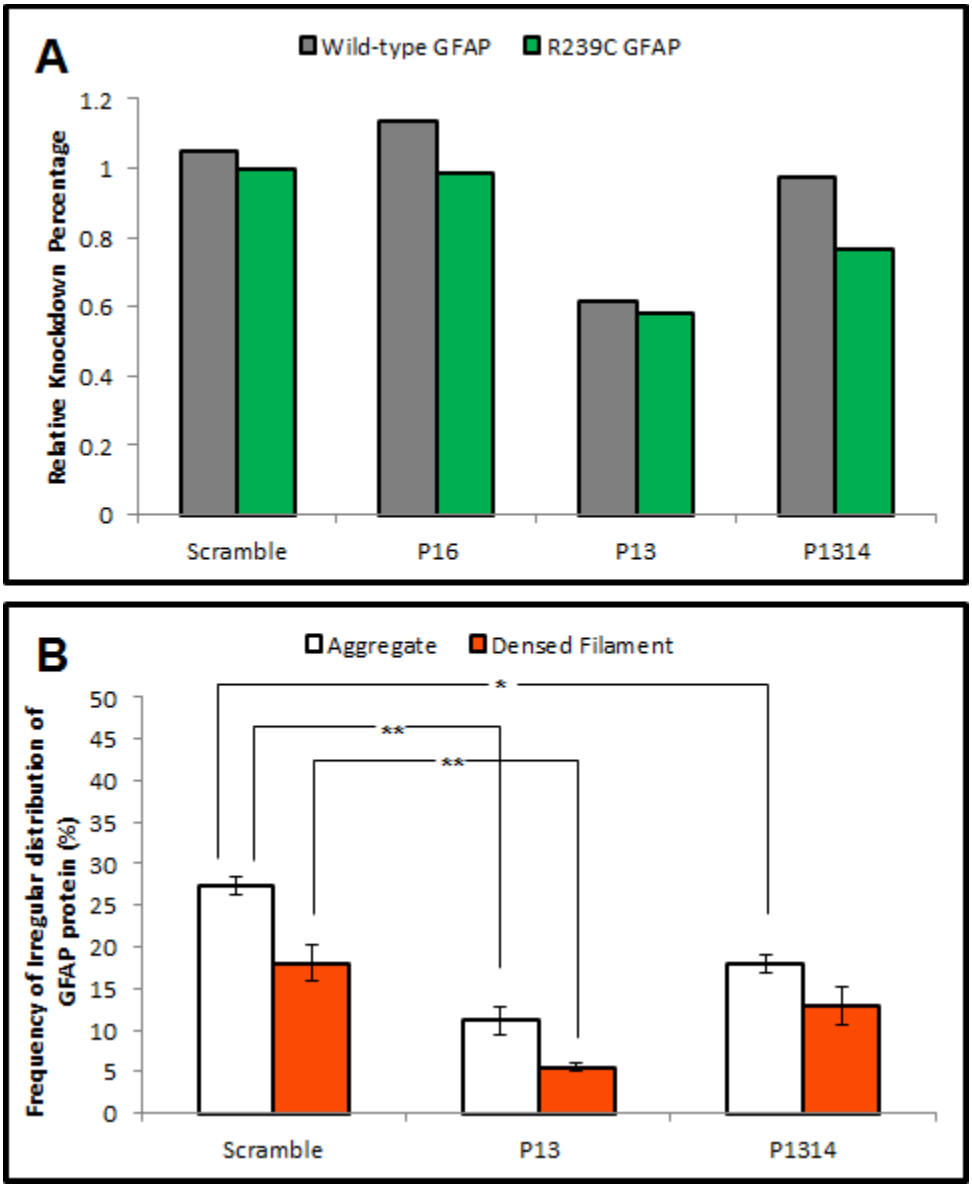

Figure 4.11 AxD iA phenotype after shRNA knock-down of GFAP. (A) Three shRNAs (P16, P13, and P1314) were designed and tested along with the negative control (Scramble) for knockdown of overexpressed Wild-type GFAP or R239C GFAP in 293T cells. Data normalized against a no shRNA control. (B) P13 and P1314 along with the negative control were separately transduced into AxD fibroblasts. They were then reprogrammed into iAs and stained with GFAP. Afterwards, cells with aggregates and densed fibers were counted as a percentage of total cells counted. * $p < 0.05$; ** $p < 0.005$

to see how the GFAP organization in the different samples differed. The difference was immediately noticeable in the P13 samples when compared to the scrambled shRNA control. There were a lot more normal cells in the P13 samples. The P1314 samples also had an increase in the proportion of normal cells, but it was only noticeable after counting (Figure 4.11B). However, the number of cells with aggregates were consistently higher than the number of cells with densed fibers throughout all the samples, which was also the case when the AxD iAs were first tested (Figure 4.8). This could be a feature of the AxD iA model system.

There are two aspects of the shRNA knockdown that are quite fascinating. First, even with relatively low expression knockdown, there was a reduction in the number of aggregates and densed fibers in the AxD iAs. And secondly, even though P13 was less specific and knocked down both mutant and wild-type GFAP, there was a stronger rescue of the phenotype using P13 compared to P1314. The most simple explanation of the perceived discrepancy could simply be a disconnect between the RNA levels measured by the qPCR in the shRNA tests and the actual protein levels of GFAP in the cells. This is especially relevant because the GFAP proteins were overexpressed in 293T cells for the qPCR, while the AxD iAs did not have the same level of overexpression of GFAP and therefore knockdown levels could be drastically different. However, the simple explanation can explain the first observation but not necessarily the second observation. Instead, the explanation that I think is more biologically relevant and can explain both interesting observations is the fact that the levels of both wild-type GFAP and mutant R239C GFAP affect the aggregation dynamics of GFAP as shown by the work of the Goldman group (Koyama and Goldman, 1999; Tang et al., 2006; Tang et al., 2010). The aggregation dynamics seems to be heavily influenced by the overall GFAP levels and not just the mutant GFAP levels. Even elevated levels of wild-type GFAP without the presence of any

mutant GFAP could cause aggregation in astrocytes. Therefore, knocking down both the wild-type GFAP and mutant GFAP simultaneously and more strongly could explain the better rescue of P13, which had around 40% knockdown of both proteins as compared to P1314 that had only 25% knockdown of only the mutant protein (effectively only lowering total GFAP levels by 12.5% assuming equal levels of mutant and wild-type protein).

REFERENCES

Abrajano JJ, Qureshi IA, Gokhan S, Zheng D, Bergman A, Mehler MF (2009). Differential deployment of REST and CoREST promotes glial subtype specification and oligodendrocyte lineage maturation. *PLoS One* **4**: e7665.

Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, Xing Y, Lubischer JL, Krieg PA, Krupenko SA, Thompson WJ, Barres BA (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J. Neurosci.* **28**: 264-278.

Deneen B, Ho R, Lukaszewicz A, Hochstim CJ, Gronostajski RM, Anderson DJ (2006). The transcription factor NFIA controls the onset of gliogenesis in the developing spinal cord. *Neuron* **52**: 953-968.

Der Perng M, Su M, Wen SF, Li R, Gibbon T, Prescott AR, Brenner M, Quinlan RA (2006). The Alexander disease-causing glial fibrillary acidic protein mutant, R416W, accumulates into Rosenthal fibers by a pathway that involves filament aggregation and the association of alpha B-crystallin and HSP27. *Am. J. Hum. Genet.* **79**: 197-213.

Ding Q, Lee YK, Schaefer EA, Peters DT, Veres A, Kim K, Kuperwasser N, Motola DL, Meissner TB, Hendriks WT, Trevisan M, Gupta RM, Moisan A, Banks E, Friesen M, Schinzel RT, Xia F, Tang A, Xia Y, Figueroa E, Wann A, Ahfeldt T, Daheron L, Zhang F, Rubin LL, Peng LF, Chung RT, Musunuru K, Cowan CA (2012). A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* **12**: 238-251.

Fan G, Martinowich K, Chin MH, He F, Fouse SD, Hutnick L, Hattori D, Ge W, Shen Y, Wu H, ten Hoeve J, Shuai K, Sun YE (2005). DNA methylation controls the timing of astrogliogenesis through regulation of JAK-STAT signaling. *Development* **132**: 3345-3356.

Garcia AD, Doan NB, Imura T, Bush TG, Sofroniew MV (2004). GFAP-expressing progenitors are the principal source of constitutive neurogenesis in adult mouse forebrain. *Nat. Neurosci.* **7**: 1233-1241.

Ge W, Martinowich K, Wu X, He F, Miyamoto A, Fan G, Weinmaster G, Sun YE (2002). Notch signaling promotes astrogliogenesis via direct CSL-mediated glial gene activation. J. Neurosci. Res. **69**: 848-860.

Geng CM, Ding HL (2008). Double-mismatched siRNAs enhance selective gene silencing of a mutant ALS-causing allele. *Acta Pharmacol. Sin.* **29**: 211-216.

Guo JU, Su Y, Zhong C, Ming GL, Song H (2011). Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle* **10**: 2662-2668.

Grandbarbe L, Bouissac J, Rand M, Hrabé de Angelis M, Artavanis-Tsakonas S, Mohier E (2003). Delta-Notch signaling controls the generation of neurons/glia from neural stem cells in a stepwise process. *Development* **130**: 1391-1402.

Iwaki T, Kume-Iwaki A, Liem RK, Goldman JE (1989). Alpha B-crystallin is expressed in non-lenticular tissues and accumulates in Alexander's disease brain. *Cell* **57**: 71–78.

Kim J, Su SC, Wang H, Cheng AW, Cassady JP, Lodato MA, Lengner CJ, Chung CY, Dawlaty MM, Tsai LH, Jaenisch R (2011). Functional integration of dopaminergic neurons directly converted from mouse fibroblasts. *Cell Stem Cell* **9**: 413-419.

Ko M, An J, Bandukwala HS, Chavez L, Aijö T, Pastor WA, Segal MF, Li H, Koh KP, Lähdesmäki H, Hogan PG, Aravind L, Rao A (2013). Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* **497**: 122-126.

Koyama Y, Goldman JE (1999). Formation of GFAP cytoplasmic inclusions in astrocytes and their disaggregation by alphaB-crystallin. *Am. J. Pathol.* **154**: 1563-1572.

Kuroda Y, Kitada M, Wakao S, Nishikawa K, Tanimura Y, Makinoshima H, Goda M, Akashi H, Inutsuka A, Niwa A, Shigemoto T, Nabeshima Y, Nakahata T, Nabeshima Y, Fujiyoshi Y, Dezawa M (2010). Unique multipotent cells in adult human mesenchymal cell populations. *Proc. Natl. Acad. Sci. U.S.A.* **107**: 8639-8643.

Miller FD, Gauthier AS (2007). Timing is everything: Making neurons versus glia in the developing cortex. *Neuron* **54**: 357–369.

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**: 143-148.

Molofsky AV, Krencik R, Ullian EM, Tsai HH, Deneen B, Richardson WD, Barres BA, Rowitch DH (2013). Astrocytes and disease: a neurodevelopmental perspective. *Genes Dev.* **26**: 891-907.

Morte B, Manzano J, Scanlan TS, Vennström B, Bernal J (2004). Aberrant maturation of astrocytes in thyroid hormone receptor alpha 1 knockout mice reveals an interplay between thyroid hormone receptor isoforms. *Endocrinology* **145**: 1386-1391.

Mussolino C, Cathomen T (2012). TALE nucleases: tailored genome engineering made easy. *Curr. Opin. Biotechnol.* **23**: 644-650.

Nakashima K, Taga T (2002). Mechanisms underlying cytokine-mediated cell-fate regulation in the nervous system. *Mol. Neurobiol.* **25**: 233-244.

Pang ZP, Yang N, Vierbuchen T, Ostermeier A, Fuentes DR, Yang TQ, Citri A, Sebastiano V, Marro S, Südhof TC, Wernig M (2011). Induction of human neuronal cells by defined transcription factors. *Nature* **476**: 220-223.

Schwarz DS, Ding H, Kennington L, Moore JT, Schelter J, Burchard J, Linsley PS, Aronin N, Xu Z, Zamore PD (2006). Designing siRNA that distinguish between genes that differ by a single nucleotide. *PLoS Genet.* **2**: e140.

Shepherd TG, Mujoomdar ML, Nachtigal MW (2010). Constitutive activation of BMP signalling abrogates experimental metastasis of OVCA429 cells via reduced cell adhesion. *J. Ovarian Res* **3**: 5.

Son EY, Ichida JK, Wainger BJ, Toma JS, Rafuse VF, Woolf CJ, Eggan K (2011). Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell* **9**: 205-218.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**: 861-872.

Takahashi K, Yamanaka S (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663-676.

Tang G, Perng MD, Wilk S, Quinlan R, Goldman JE (2010). Oligomers of mutant glial fibrillary acidic protein (GFAP) inhibit the proteasome system in alexander disease astrocytes, and the small heat shock protein alphaB-crystallin reverses the inhibition. *J. Biol. Chem.* **285**: 10527-10537.

Tang G, Xu Z, Goldman JE (2006). Synergistic effects of the SAPK/JNK and the proteasome pathway on glial fibrillary acidic protein (GFAP) accumulation in Alexander disease. *J. Biol. Chem.* **281**: 38634-38643.

Tian R, Wu X, Hagemann TL, Sosunov AA, Messing A, McKhann GM, Goldman JE (2010). Alexander disease mutant glial fibrillary acidic protein compromises glutamate transport in astrocytes. *J. Neuropathol. Exp. Neurol.* **69**: 335-345.

Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Südhof TC, Wernig M (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**: 1035-1041.

Yoshida T, Nakagawa M (2011). Clinical aspects and pathology of Alexander disease, and morphological and functional alteration of astrocytes induced by GFAP mutation. *Neuropathology* **32**: 440-446.

CHAPTER 5

Conclusions

Ever since Waddington proposed a simple model to explain cell identity, the understanding of the epigenetic mechanisms of cell identity maintenance and differentiation has been evolving. Through the seminal work of Gurdon, Davis, and Yamanaka, the field has gone through many renaissances and advanced drastically to today's understanding that a cell identity is highly plastic and can be altered by manipulating the transcription machinery and epigenetics of the cell. Yamanaka's streamlining of the reprogramming process when he first reprogrammed iPS cells can be said to have revolutionized and rejuvenated the field, leading to many important revelations since his first reports of mouse iPS.

However, Yamanaka's approach to reprogramming still used a lot of black-box epigenetic changes and mechanisms that are still being elucidated today. One of the most interesting questions in iPS reprogramming is why is it so slow and inefficient? The reprogramming process does not necessarily have to be a slow and inefficient as shown by reprogramming via cell fusion. In fact, different cell types reprogram at different speeds even though they all have a rather large epigenetic "hill" to overcome to return to pluripotency. Indeed, how high up the epigenetic hill a cell starts does not seem to determine the speed of reprogramming. By studying these different cell types' expression profiles, I have found two modules using WGCNA that could possibly contain genes that are important for the alteration of iPS reprogramming speed. PIAS3 and STAT3, though intriguing candidates, does not seem to be able to modify this fundamental characteristic of four Yamanaka factor iPS reprogramming. It could be that a combination of factors are necessary or that other genes within those two modules play a more primary role in the differences in reprogramming speeds of the four cell types.

Though iPS reprogramming is powerful and extremely useful, direct reprogramming from one cell lineage to another can be even more relevant for studying disease models. Luckily, the process that Yamanaka pioneered can be easily applied to direct reprogramming as well. Using this process, I was able to reprogram both mouse and human fibroblasts into induced astrocytes that can support neuronal function and express astrocytic markers. The induced astrocyte technology can be used for disease modeling as human astrocytes are difficult to obtain as are most cells of the brain. As a proof of principle, I applied the iA technology to reprogramming Alexander's Disease patient-specific fibroblasts. The AxD iAs display the typical phenotypes that are found in the mouse models and human diseased astrocytes, namely Rosenthal fibers as well as elevated levels of CRYAB and depressed levels of GLT1. By using shRNA knockdown, I have shown that this human iA model of AxD can be used to affect improvement of the disease phenotype. Thereby, the human AxD iA model could be used for future studies on therapeutic approaches to the human disease as well.

The reprogramming field is still rapidly developing and new breakthroughs are made every day. Eventually, the enigmas of iPS reprogramming will be fully understood and the ethical benefits of the technology will allow stem cell technology to advance to a whole new level. In the future, I imagine that making disease models of different cell types would be a relatively easy task whether by direct reprogramming or by iPS reprogramming followed by differentiation. By opening up the epigenetics of cell identity, the pioneering work of Gurdon, Lasser, and Yamanaka has truly created a whole new world of discovery.