THE DIRECT REPROGRAMMING OF SOMATIC CELLS: ESTABLISHMENT OF A NOVEL SYSTEM FOR PHOTORECEPTOR DERIVATION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Melissa Mary Steward

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

December 2012

Purdue University

Indianapolis, Indiana

This thesis is dedicated to Anne McSherry Steward and the loving memory of Eleanor Mary Vahey. The grandmothers of the author provided unflinching support and inspiring examples of strength while articulating the value of independence, family and education. Working mothers were my first teachers of the critical concept ‘necessary and sufficient’. They have my gratitude, respect and love for everything they shared.

ACKNOWLEDGMENTS

I would like to thank many individuals for their support of this work. Firstly, I must thank Dr. Jason S. Meyer, my thesis advisor, and friend of over a decade, for always respecting and soliciting my contributions to science and education. His support and expertise are deeply appreciated for their influence on me as a developing scientist and person. I thank the members of my committee, Dr. Stephen Randall and Dr. Guoli Dai as well as my department chair, Dr. Simon Atkinson, for their support and expertise. I owe a debt of gratitude for significant support, both technical and personal, to my friend and collaborator Akshayalakshmi Sridhar. She is wise beyond her years. I thank Dr. Kathy Marrs, Dr. Mariah Judd and the NSF-funded GK-12 program, for providing professional and financial support, as well as the opportunity to teach in one of my favorite settings. I thank Meyer lab member Manav Gupta, the Biology Department and support staff of IUPUI, namely Sue Merrell, Shari Dowell and Kurt Kulhavy, for their technical contributions and assistance. I must thank my sister, Jennifer Steward and friend, Matthew Butcher, who have been indefatigable sources of love, support and motivation. Dr. Mark Kirk deserves special thanks as the first scientist to provide me a project and kindly, emphatically and ceaselessly encouraging and supporting my career and graduate studies. Thanks also to my many inspiring educators and supportive friends over the years, too numerous to name here and all those who went before me, shining the light as far as they could, allowing me to press even farther still. It appears to take a village to complete a thesis: I extend my deep and sincere gratitude to the numerous individuals who contributed to my success in this endeavor.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ALS	amyotrophic lateral sclerosis
BAM	A combination of three proneural transcription factors *Brn2, Ascl1* and *Myt1l*
bHLH	basic helix-loop-helix protein
BSC	Biological Safety Cabinet
CMV	cytomegalovirus
cDNA	complementary DNA
DAPI	4′,6-diamidino-2-phenylindole
DMEM	Dulbecco's modified eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
E16	embryonic day16
EDTA	ethylenediaminetetraacetic acid
ESCs	embryonic stem cells
FACS	fluorescence-activated cell sorting
FAD	familial Alzheimer′s disease
FBS	fetal bovine serum
GFP	green fluorescent protein
HBSS	Hank′s balanced salt solution
HEK293	human embryonic kidney cell line 293
iDA	induced dopaminergic
iMN	induced motor neuron
iN	induced neuronal cells
iPSCs	induced pluripotent stem cells

LCA	Leber′s congential amaurosis
MEF	mouse embryonic fibroblast
miRNA	micro ribonucleic acid
mRNA	messenger ribonucleic acid
MCSs	multiple cloning sites
MOIs	multiplicities of infection
P2	post-natal day 2
PBS	phosphate buffered saline
PGK	phosphoglycerate kinase
qRT-PCR	quantitative reverse transcriptase polymerase chain reaction
RP	retinitis pigmentosa
SCNT	somatic cell nuclear transfer
SMA	spinal muscular atrophy

ABSTRACT

Steward, Melissa Mary. M.S., Purdue University, December 2012. The Direct Reprogramming of Somatic Cells: Establishment of a Novel System for Photoreceptor Derivation. Major Professor: Jason S. Meyer.

Photoreceptors are a class of sensory neuronal cells that are deleteriously affected in many disorders and injuries of the visual system. Significant injury or loss of these cells often results in a partial or complete loss of vision. While previous studies have determined many necessary components of the gene regulatory network governing the establishment, development, and maintenance of these cells, the necessary and sufficient profile and timecourse of gene expression and/or silencing has yet to be elucidated. Arduous protocols do exist to derive photoreceptors *in vitro* utilizing pluripotent stem cells, but only recently have been able to yield cells that are disease- and/or patient-specific. The discovery that mammalian somatic cells can be *directly* reprogrammed to another terminally-differentiated cell phenotype has inspired an explosion of research demonstrating the successful genetic direct reprogramming of one cell type to another, a process which is typically both more timely and efficient than those used to derive the same cells from pluripotent stem cell sources. Therefore, the emphasis of this study was to establish a novel system to be used to determine a minimal transcriptional network capable of directly reprogramming mouse embryonic fibroblasts (MEFs) to rod photoreceptors. The tools, assays and experimental design chosen and established herein were designed and characterized to facilitate this determination and preliminary data demonstrated the utility of this approach for accomplishing this aim.

1 INTRODUCTION

The fields of developmental and regenerative biology have long sought to identify novel approaches for the repair of damaged and/or diseased tissue, including that of the nervous system. The mammalian central nervous system has been well documented as one with limited regenerative capabilities, due at least in part to an inhospitable environment for regeneration [1, 2]. In cases of injury and neurodegeneration, glial scarring, the lack of proliferating oligodendrocytes, and the presence of inhibitory factors can physically block or impair the regrowth of damaged neuronal axons and pathfinding of growth cones [3, 4]. In both injury-induced and neurodegenerative disorders, a toxic extracellular environment including widespread cell death and a general absence of growth-promoting signals has been described [4, 5]. The mulitude of factors contributing to the lack of regeneration in the mammalian central nervous system has been a significant limitation for the fields of mammalian developmental biology and regenerative medicine. A further limitation is a reduced ability to study the molecular mechanisms and sequelae of disease at the cellular level, in both developing and adult tissue. A lack of animal models for many disorders, as well as uncharacterized species differences in the pathways involved in injury, neurodegeneration and regeneration have hampered efforts to describe the underlying mechanisms controlling and contributing to these processes.

1.1 Pluripotent stem cells as models and therapeutic agents

When mouse embryonic stem cells (ESCs) were first derived in 1981 [6], followed by the derivation of human ESCs in 1998 [7], they provided a new model system for researchers to study developmental and disease processes at a cellular level. At the

same time, they represented a new potential therapeutic cellular agent for clinicians as a source for replacement cells in cases of neurodegeneration and injury.

ESCs are derived from the inner cell mass of a fertilized oocyte, and have two defining characteristics. They are pluripotent, which means that they can give rise to all the cell types of an adult organism, including all of the specific cell types of the central nervous system. They are also capable of self-renewal, which allows them to be cultured and expanded *in vitro* indefinitely, providing an unlimited source of cells for applications of research or therapeutics. However, one of the two major limiting attributes of ESCs as applied to the field of therapeutics is the fact that they are not patient-specific. Thus, these cells have an increased risk of rejection when transplanted into another individual. A second inherent risk involved with the transplantation of cells derived from a pluripotent cell source is the potential for delivering pluripotent, or undifferentiated and dividing, cells to the body

The derivation of induced pluripotent stem cells (iPSCs) in 2006 [10] represented a critical advance for regenerative medicine as the first opportunity to derive cells from a pluripotent source while circumventing the risk of immune rejection due to the ability to derive patient-specific lines. These iPSCs provided an opportunity to derive adult cell types via an indirect cellular reprogramming strategy, which could serve as the basis for pharmacological screening, disease-modeling and therapeutics such as cellular replacement or cell rescue enabled by transplantation. However, the second limiting attribute of ESCs as applied to therapeutics was not overcome with the advent of this new pluripotent cell source. The delivery of mitotically active, undifferentiated cells to a niche introduces a risk of tumorogenicity, i.e. tumor formation. Unregulated cell division and invasion of undifferentiated or inappropriately differentiated cells is a hallmark of certain forms of cancers. The advent of iPSCs did however, open wide the door for further innovative studies in cellular reprogramming. Directed *in vitro* differentiation of iPS cells prior to transplantation constitutes

one mechanism with which to minimize teratogenicity, but it does not exclude the possibility of even an exceedingly small number of cells avoiding differentiation *in vivo* application. An alternate strategy that would eliminate the tetratogenicity of iPS cell cultures would involve a direct reprogramming strategy. The demonstrated and replicated ability to genetically reprogram mammalian, adult, somatic cells to a pluripotent, mitotically-active cellular phenotype stood contrary to the long-standing tenet of biology that once cells become terminally differentiated, they cannot change their fate. If adult somatic cells could be genetically reprogrammed to a pluripotent state and further redifferentiated to specific adult cell phenotypes, the next question became: could these same adult cells be directly genetically reprogrammed to another cell fate?

1.2 Seminal studies in cellular reprogramming

Cellular reprogramming experiments conducted over the last 6 decades have laid a substantial foundation upon which the hypothesis and experimental design of this study are based. The work of Dr. John Gurdon and Dr. Shinya Yamanaka received the Noble Prize in Physiology or Medicine in 2012 for their significant and high impact discoveries in cellular reprogramming. Dr. Gurdon conducted the first experiment that successfully cloned an organism from a somatic cell source [8]. In this study, he used the process of somatic cell nuclear transfer (SCNT) established by Briggs and King [9]. This process involves the transplantation of the nucleus of a somatic cell to an enucleated, unfertilized oocyte. Cytoplasmic factors in the oocyte were found to be sufficient to reprogram the somatic nuclei to an effective earlier stage of development, allowing for the reinitiation of transcription of embryonic genes that were silenced in the adult cell and initiating cellular division of the oocyte based upon the genomic DNA of the somatic nucleus. Gurdon exploited this process to clone a new frog, *Xenopus laevis*, through the use of a nucleus from a gastrointestinal cell, removed from an adult frog. Rather than relying on undefined cytoplasmic

factors within an oocyte, Yamanaka′s work first demonstrated a genetic approach to reprogram mouse somatic cells to a pluripotent state via lentiviral delivery of a cocktail of four genes that govern pluripotency [10]. He dubbed the cells derived via this process induced pluripotent stem cells (iPSCs). Similar to embryonic stem cells, they were demonstrated to have the capacity to proliferate indefinitely in culture and differentiate both *in vivo* and *in vitro* to cell types derived from all three germ layers - ectoderm, mesoderm and endoderm. In between the time of these exciting discoveries, other groups demonstrated the direct reprogramming of fibroblasts to myoblasts via delivery of a single master transcriptional regulator *MyoD* [12], as well as the *in vivo* direct reprogramming of exocrine cells from the pancreas to insulin-secreting beta cells [13]. The implication of studies demonstrating these dramatic cell fate changes was that direct cellular reprogramming of somatic cells was possible utilizing a genetic approach.

1.3 Advantages of direct reprogramming over indirect reprogramming

There are several advantages afforded by direct reprogramming strategies when compared to those utilizing a pluripotent stem cell intermediary. While either strategy could be used to yield patient-specific cell populations, those derived via a direct reprogramming strategy can remain a mitotically inactive cell population. Indirect reprogramming strategies utilize pluripotent stem cells, which by definition are proliferative and can give rise to more undifferentiated cells, as well cells that are more differentiated. While *in vitro* protocols exist to differentiate these stem cells in substantial numbers and at high efficiencies and cell sorting using surface markers could purify these cells for many cell types, there remains an increased risk of transplanting undifferentiated cells, that could lead to tumor formation. Upon transplantation, directly reprogrammed cells would have a much lower risk of tumorigenicity, as the likelihood of introducing pluripotent stem cells to a new niche is significantly lower. Another advantage of using direct genetic reprogramming strategies is that they may

uncover novel genes involved in the gene regulatory network of the desired cell type. Many indirect reprogramming strategies utilizing *in vitro* differentiation of pluripotent stem cells to the final cell type involve adding soluble mitogens and growth factors to the cell culture media to differentiate stem cells, potentially activating or inactivating often innumerable and overlapping pathways in the cell. Direct genetic reprogramming strategies allow for the definition of elusive gene regulatory networks that are 'necessary and sufficient' for defined cellular phenotypes that are currently undescribed. Finally, direct reprogramming strategies are faster, more efficient and less arduous than those involving a pluripotent intermediary. For example, Marius Wernig′s group saw 20% conversion rates of fibroblasts to neuronal cells in 2 weeks time utilizing direct genetic reprogramming [18]. This efficiency, similar to that seen by many others, is orders of magnitude higher than that seen when establishing pluripotent stem cell lines, and on the order of weeks instead of months. For photoreceptors specifically, after the pluripotent cell lines are established, it takes up to another three months to derive photoreceptors from them [19]. None of these advantages conferred by direct genetic reprogramming affect their applicability when compared to cells derived via indirect reprogramming strategies. They can still be used for studies of development such as cell fate specification and for disease-modeling, as well as therapeutics such as cell replacement and rescue conferred by transplantation and also used for drug screening. Not only are none of these applications lost, some - such as transplantation applications - stand to be enhanced when cell populations are derived via direct reprogramming.

1.4 Differentiation and direct cellular reprogramming to neural phenotypes

Diseases of and injuries to the central and peripheral nervous system devastate the sensory experience and motor control of a significant portion of the population each year. Because of the prevalence and ramifications of these injuries and diseases, many efforts have focused on the replacement or rescue of neural cell populations once they

are damaged or lost. *In vitro* protocols already exist to derive specific neural and neuronal cell types from pluri- or multi-potent sources such as embryonic stem cells (ESCs), induced pluripotent stem cells (iPSCs) or neural stem cells [14–17, 19–21]. These protocols are often based upon culturing the stem cells in culture medium with fetal bovine serum and known proneural soluble growth factors. These factors are known to be involved in pathways governing neural specification *in vivo* and induce expression of neural-specific genes and positive feedback loops leading to the ultimate differentiation of pluripotent cells to neuronal phenotypes. Until very recently, cells derived via these protocols provided the best potential source for potential cellular replacement and rescue strategies, as well as pharmacological screening and disease-modeling.

The first successful *direct, genetic* reprogramming of mammalian somatic cells to a neuronal phenotype was published in 2010 [18] and since that time, many groups have utilized a similar experimental strategy to derive more specific neuronal cell types from mammalian somatic sources [22–27]. Vierbuchen et al. first used a strategy similar to the one employed by Yamanaka to derive induced pluripotent stem cells from fibroblasts [10]. In the landmark studies by Yamanaka group, they sought to reprogram terminally differentiated, somatic cells to a mitotically active pluripotent state. Thus, he tested the effects of viral delivery of combinations of transcription factors known to be active in embryonic stem cells and silenced in quiescent cell populations. These genes were therefore implicated to be involved in positively regulating pluripotency. Vierbuchen et al. hypothesized that a similar strategy could be used to derive neuronal cells directly from fibroblasts [18]. They defined a set of candidate transcription factors to test that were known or implicated to be involved in the processes governing pluripotency or were specific to neural cell populations. They started with a pool of 19 genes that were virally delivered combinatorially to mouse embryonic fibroblast (MEF) cells, and screened for neuronal conversion. They ultimately defined a combination of three factors, *Brn2, Ascl1* and *Myt1l* (BAM) that

could quickly and efficiently convert fibroblasts to neuronal cells. These neuronal cells were named induced neuronal (iN) cells and importantly were found to express multiple neural specific proteins, generate action potentials and form functional synapses when cultured with cortical neuronal or glial cells. These iN cell cultures contained inhibitory GABA-ergic neuronal cells, excitatory glutaminergic neuronal cells, as well as some iN cells expressing markers of cortical interneurons and other neuronal subtypes. Another important discovery from this study was the marked increase in efficiency and rapidity of neuronal conversion seen using this direct reprogramming strategy. They reported an approximate 20% conversion efficiency of infected cells within 2 weeks, wheras traditional methods for iPSC reprogramming typically report efficiencies of less than 0.1% and require several weeks for effective reprogramming. This exciting discovery spurred an explosion of studies in the neurosciences employing to use a similar approach to derive human iN cells, as well as specific neuronal cell types utilizing the same strategy. By delivering cell-specific transcription factors in combination with pro-neural genes such as those in the BAM cocktail to somatic cells, attempts were made to derive dopaminergic neurons or motor neurons, for example.

When the BAM combination of transcription factors was initially delivered to human cell cultures, immature neuronal phenotypes were reported, along with significant cell death [23, 24]. It was quickly determined that the addition of another transcription factor, *NeuroD1* to the BAM cocktail resulted in the same neuronal attributes in human cells after 5-6 weeks as those seen in the mouse system in 2 weeks with the BAM combinatorial treatment alone. Neural-specific protein expression, action potentials and post-synaptic currents were observed [23]. The differential time-course of neuronal maturation seen when comparing the mouse and human system in direct reprogramming is similar to differences seen using mouse and human derived ESCs and iPSCs and may be reflective of a longer period of maturation during human gestation and *in vivo* development.

As dopaminergic neurons are affected in many neurodegenerative disorders, such as Parkinson′s disease and familial Alzheimer′s disease, replacement or rescue of these specific neurons holds great promise for strategies of regenerative medicine. Dopaminergic neurons were the first neuronal subtypes to be specified through genetic, direct reprogramming strategies [24–27]. Several independent studies reported different combinations of factors to derive these action potential-firing, tyrosine-hydroxylase positive, induced dopaminergic (iDA) cells from human and mouse fibroblast cells, with efficiencies reported approximating 10% of transduced cells, though only the delivery of *Ascl1, Nurr1* and *Lmx1a*, or the combination of these 3 genes with *Pitx3, Foxa2* and *En1* was capable of reprogramning cells that were characterized to release dopamine [25, 27]. Spinal motor neurons are another specific neuronal cell type that is known to be affected by disease-states including spinal muscular atrophy (SMA) and amyotrophic lateral sclerosis (ALS) or Lou Gehrig′s disease. Less than a month after reports about the direct reprogramming of human and mouse fibroblasts to induced dopaminergic (iDA) neuronal cells were published, the first study characterized the direct reprogramming of spinal motor neuronal cells as well [22]. Their highest efficiencies of conversion (around 5-10% in under 2 weeks) from mouse fibroblasts to induced motor neuron (iMN) cells were reported using the aforementioned BAM combination with the addition of four spinal motor neuron-specific factors, *Lhx3, Hb9, Isl1* and *Ngn2*. These iMN cells generated action potentials and responded to both excitatory and inhibitory neurotransmitters in culture, similar to ESC-derived and embryonic motor neurons. Addition of *NeuroD1* to the pool of these seven factors led to functional iMN cells reprogrammed from human ESC-derived fibroblasts as well, that were characterized as similar to their mouse counterparts in the study.

1.5 Specific neuronal subtypes as phenocopies and replacement cell sources

Once defined neuronal cell types could be specified using direct genetic reprogramming, the field was poised to ask if these directly reprogrammed iN cells could 1)

serve as reliable phenocopies for disease-states, 2) be demonstrated to integrate *in vivo* and 3) restore any function that had been lost associated with the particular disease pathology. Indeed, these questions have been addressed by several studies.

In the first study to derive iMN cells, using both mouse and human cells, iMN cell sensitivity to growth factor withdrawal was demonstrated similar to embryonic motor neurons [22]. The significant interest in the factors and pathways that confer neuronal survival in the context of injury and neurodegenerative disease states makes these cells a valuable *in vitro* tool for the study of motor neuron function, survival, disease, injury, and response to exogeneously added or removed defined factors. They further cocultured their iMN cells with glial cells derived from the SOD1 mutant mouse model of ALS, as it is known that motor neurons are selectively sensitive to toxic effects of mutant glia when compared to other neuronal cell types, such as spinal interneurons [28, 29]. They indeed demonstrated a reduction in iMN cells to an extent similar to that seen with ESC-derived motor neurons in this coculture system [28, 29]. They also found that iMNs derived from this mutant mouse model had impaired survival in culture when compared to wild-type derived iMNs. These findings in combination suggest that iMNs can serve as phenocopies for "both cell-autonomous and non-cell-autonomous contributors to motor neuron degeneration in ALS" [22]. Furthermore this group also used a rigorous test commonly used by the field of neuroscience to test the *in vivo* survival, migration ability, and response to *in vivo* axon guidance cues of these iMNs, testing their ability to contribute to the developing central nervous system. It was demonstrated that upon injection to the chick embryo neural tube, iMNs were able to survive *in vivo*, migrate to appropriate regions to integrate, and respond appropriately to *in vivo* axon guidance cues, as demonstrated by their axonal projections out of the spinal cord via the ventral horn towards the musculature.

Studies of induced dopaminergic (iDA) cells have taken the characterization of the utility of derived neuronal cells a step farther, demonstrating not only their abil-

ity to be derived from human patients with diseases such as familial and sporadic Alzheimer′s [24] or Parkinson′s [25, 27] and exhibit disease-specific phenotypes *in vitro* [24] as well as survive and integrate upon transplantation [25, 27]), but also that upon transplantation iDA cells were able to alleviate symptoms in a mouse model of Parkinsons disease [27]. Elevated dopamine levels were detected in the transplanted striatum of 6OHDA lesioned mice compared to controls and eight weeks after transplantation the animals with implanted cells showed significant reduction in amphetamine-induced rotation scores when compared to sham-injected or intact control-lesioned animals. While further studies need conducted aimed to increase the efficacy of such treatments, this important proof-of-principle establishes the utility of transplanted iDA cells to restore function in at least one animal model of human disease or injury.

All of these studies utilized a genetic approach to induce neuronal cells from fibroblasts. While there has been significant overlap in the particular genes or transcription factors specifically that were delivered, several groups have demonstrated similar cell phenotypes using various combinations. Interestingly, the group that reprogrammed fibroblasts from familial and sporadic Alzheimer′s disease patients used a 5-factor combination of genes to derive their iN cells that included *Brn2, Ascl1, Zic1, Olig2* and *Myt1l* further demonstrating that there are multiple pathways to a neural - even specific neuronal subtype - identity [24].

1.6 A model that accounts for direct cellular reprogramming

The paradigm of cellular biology during development once stated that cells undergo an irreversible process of increasing lineage commitment as they undergo differentiation, i.e. as cells develop and begin to differentiate, they become increasingly committed to a particular phenotype and once terminally differentiated, they cannot reinitiate cellular division or change cellular fate. However, an ever-rapidly growing number

of peer-reviewed studies has indicated and even characterized, events and outcomes completely contrary to this long-standing tenet of biology. If this relatively new, intriguing, expansive body of data cannot be reconciled with the previous biological model of development and cell fate commitment, then what model *does* exists to account for these phenomena that are observed and reproduced in such astounding numbers?

The gene expression network should be conceptualized as, and has indeed been demonstrated to be, a highly dynamic, multi-dimensional space. As an accepted rule, microarray data of global gene expression profiles demonstrates the highly dynamic nature of gene expression over time, as well as the variability within defined cell populations. For purposes of modeling, one should imagine each individual gene′s expression level as represented by an axis [31, 32]. The model depicted and defined by Huang [32] and Zhou and Huang [31] also describe particular positions within this multi-dimensional space that are states of gene expression that are low-energy for the cell to maintain. They name these states "attractor states". (Figure 1.1) shows a simplified gene network in which genes X1 and X2 cross-inhibit one another (a) and in (b) also positively feedback upon themselves. The third panel in each of these schematics graphically depicts the low energy 'attractor' states on the Z-axis of Quasi-potential [energy] (U) occupied by a cell governed by these feedback networks. Note that high expression of gene X1 along the y-axis coupled with low expression of gene X2 on the x-axis is depicted as an attractor state, S_1. A similarly stable but opposite gene expression profile exists at S_2, noting a cell′s state when it has a pattern of gene expression corresponding to low levels of gene X1 and high levels of gene X2. As noted in the figure legend, the "higher U is, the less stable that state is [31]". The cell reaches a low energy state by occupying a gene expression profile of what the authors named an 'attractor state'. Other intermediary gene expression profiles are less energy efficient, as indicated by their higher position on the Z-axis. The cell is therefore attracted to these basins of stability that are reflected by cellular pheno-

types, governed in part by gene expression feedback systems. Direct reprogramming strategies can therefore be considered two-fold in their approach. They seek to push the gene expression of a cell far enough out of it′s current attractor state and also nearest to the attractor state of the cellular phenotype desired.

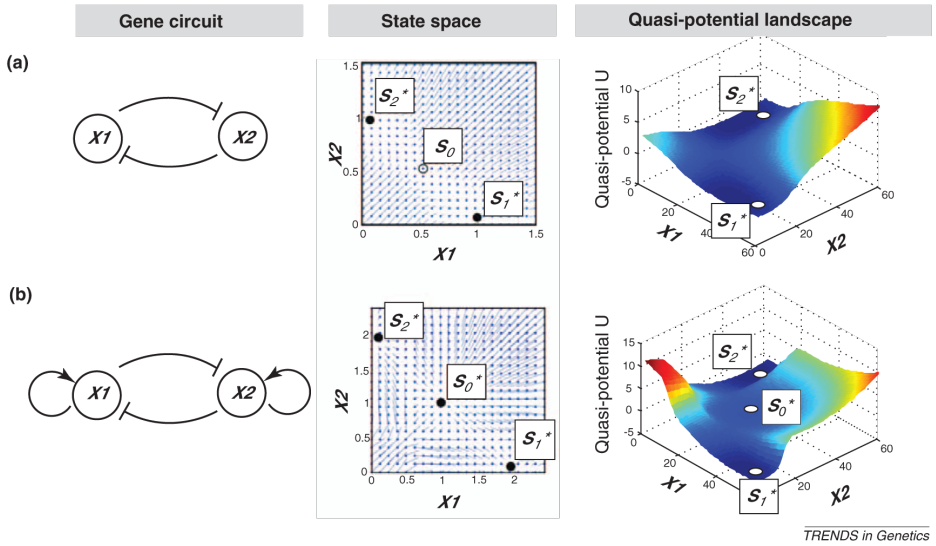

Figure 1.1. Simplified schematic of gene circuits and attractor states. Reprinted from Trends in Genetics, 27, Zhou JX, Huang S, Understanding gene circuits at cell-fate branch points for rational cell programming, pages 55-62, Copyright (2011), with permission from Elsevier. The self-stimulation (positive feedback) of genes X1 and X2 creates attractor state, S_0, representing a bipotent progenitor that is less stable that (attractor states) S_1 or S_2. "The quasi-potential landscape (right panel) offers a view on the global dynamics by assigning to each point S in the state space a 'quasi-potential' U(S) that is inversely related to the approximate relative stability of S, hence enabling the comparison of the relative 'depth' of attractors or any other point S. In this two-gene system, the state space is represented by the XY plane, whereas the Z-axis denotes U(S). The higher U is, the less stable that state is. Thus, the system is attracted to the lowest points = stable states = attractor states" [31].

Several useful predictions can be made using this model and many have been demonstrated to be true by the growing body of evidence put forth by studies of indirect and direct cellular reprogramming. First, since the state or phenotype of a cell at

any point in time is governed to a large extent by it′s gene expression profile, it is not permanent, even though relatively stable. If acted upon enough by outside factors that influence gene expression, a cell′s fate could be changed. This change would be the result of a significant enough change in gene expression, or enough energy added to the system, to overcome the stability gained by occupying its current phenotype. This model also predicts that the processes of cellular reprogramming do not need to be externally regulated throughout the entire process. Rather, it predicts that enough of a perturbation in the system can remove the cell from it′s current attractor state and that upon that perturbation, it will seek the nearest attractor state. This has been demonstrated by several groups that have used forced gene expression of lineage- and cell-specific genes to induce the cellular phenotypes they sought to induce from terminally differentiated cell types that typically have little to no expression of the specific genes delivered. This model also predicts that cells with more similar gene expression profiles can more easily be transitioned between. Another prediction of the model would be that the forced expression of specified genes may not be necessary. Rather, published by many independent groups, various combinations of genes involved in transcriptional regulation of cell-specific genes could provide enough change, likely due to positive feedback mechanisms and feed-forward systems that push gene expression towards a particular, desired attractor state.

1.7 Photoreceptors: A unique opportunity for direct reprogramming

Initial studies establishing direct reprogramming as a viable induction method to derive neuronal cell types were enabled by 1) a need for these specific cell types, as dictated by particularly problematic human disease pathologies and 2) a well-established body of literature identifying and delineating important gene regulatory networks of

the final, desired cell types. Photoreceptor cells of the retina constitute an additional cell type in which both of the requirements also exist, yet direct reprogramming of somatic cells to a photoreceptor fate has yet to be achieved.

The loss of sight, and the ensuing problems it brings are certainly among our most basic human fears. Almost 30% of the sensory input to the brain traces back to the retina, which is commonly referred to as the "window to the brain" [34–36]. The visual experience begins with photoreceptors, a unique class of neuronal sensory cells that are responsible for receiving light information that falls on the retina and converting that input to signals that the nervous system can process. The output of photoreceptors is integrated and processed first by interneurons of the retina before the information is transmitted to visual centers and others in the brain [37]. It should be no surprise then, that diseases deleteriously affecting photoreceptors are the primary cause of visual impairment or blindness in most retinal diseases, including macular degeneration, Lebers congential amaurosis (LCA), and retinal pigmentosa (RP), to name a few of the more common [36]. Therefore, cellular replacement strategies often have been aimed at protecting these important sensory cells as well as replacing them through transplantation, or by stimulating in vivo rescue or replacement by existing cell populations. Furthermore, studies and models of retinal degeneration could also provide valuable information about more general features of progressive neurodegeneration [38].

Photoreceptors are broadly classified into two main types: cones or rods. Cones respond to bright light and relay high resolution, color information. Rods on the other hand, function in low light and are a hundred-fold more light-sensitive than cones [36, 37, 39]. In mice and humans, 70-80 % of all cells in the neural retina are photoreceptors, with rods outnumbering cones 30:1 in mice, and 18-20:1 in humans [36, 41, 42], indicating that rod photoreceptors are the most abundant cell type in the retina of both mice and humans. While subtypes of cones exist expressing

different and singular visual pigments, the mammalian retina has only one rod opsin, rhodopsin, with a peak sensitivity around 500 nm [36, 37]. Lastly, and importantly, transplantation studies have demonstrated that rod precursor cells readily incorporate in the adult retina, differentiate, and form synaptic connections [43]. This study contrasted these rod progenitors with other progenitor or stem cells from various alternate stages of development that failed to integrate to the same extent as rod progenitors [43–47]. For these reasons- abundance, sensitivity, simplicity, and demonstrated integration- an abundance of research has focused on the gene regulatory networks of rod photoreceptors. Furthermore, the aforementioned reasons also make rod photoreceptor cells ideal targets for studies of direct cellular reprogramming, as well as excellent candidates for the first applications of directly reprogrammed cells to regenerative medicine, including transplantation experiments aimed at recovering vision in genetic or injury models where vision has been lost or impaired due to loss of photoreceptors.

1.8 Transcriptional dominance model of photoreceptor cell fate determination

Decades of research support the transcriptional dominance model (Figure 1.2) of photoreceptor cell fate determination put forth by Dr. Anand Swaroop [36]. While he states that "the molecular mechanisms that generate photoreceptor precursors from retinal progenitor cell remain uncharacterized", several players, including but not limited to, *CrX, Otx2, NrL, Nr2e3* and *RORβ* have been implicated as necessary in rod photoreceptor development [36]. Loss of any one of these genes leads to a complete, or almost complete loss of rod photoreceptors, or lack of expression of many important rod-specific phototransduction genes [36, 48–52]. It should also be noted that not one of these single genes has been sufficient to induce the differentiation of rod photoreceptors. However, the demonstrated overlapping targets of these genes, as well as the step-wise nature of photoreceptor differentiation from retinal progenitors and the increasingly likely multifactorial and transient nature of the terminal differ-

entiation process, makes identification of the genes which are necessary and sufficient a difficult task. So while some hierarchies of gene regulation and feedback loops are well-established, the 'necessary and sufficient' master transcription regulatory network continues to eludes researchers.

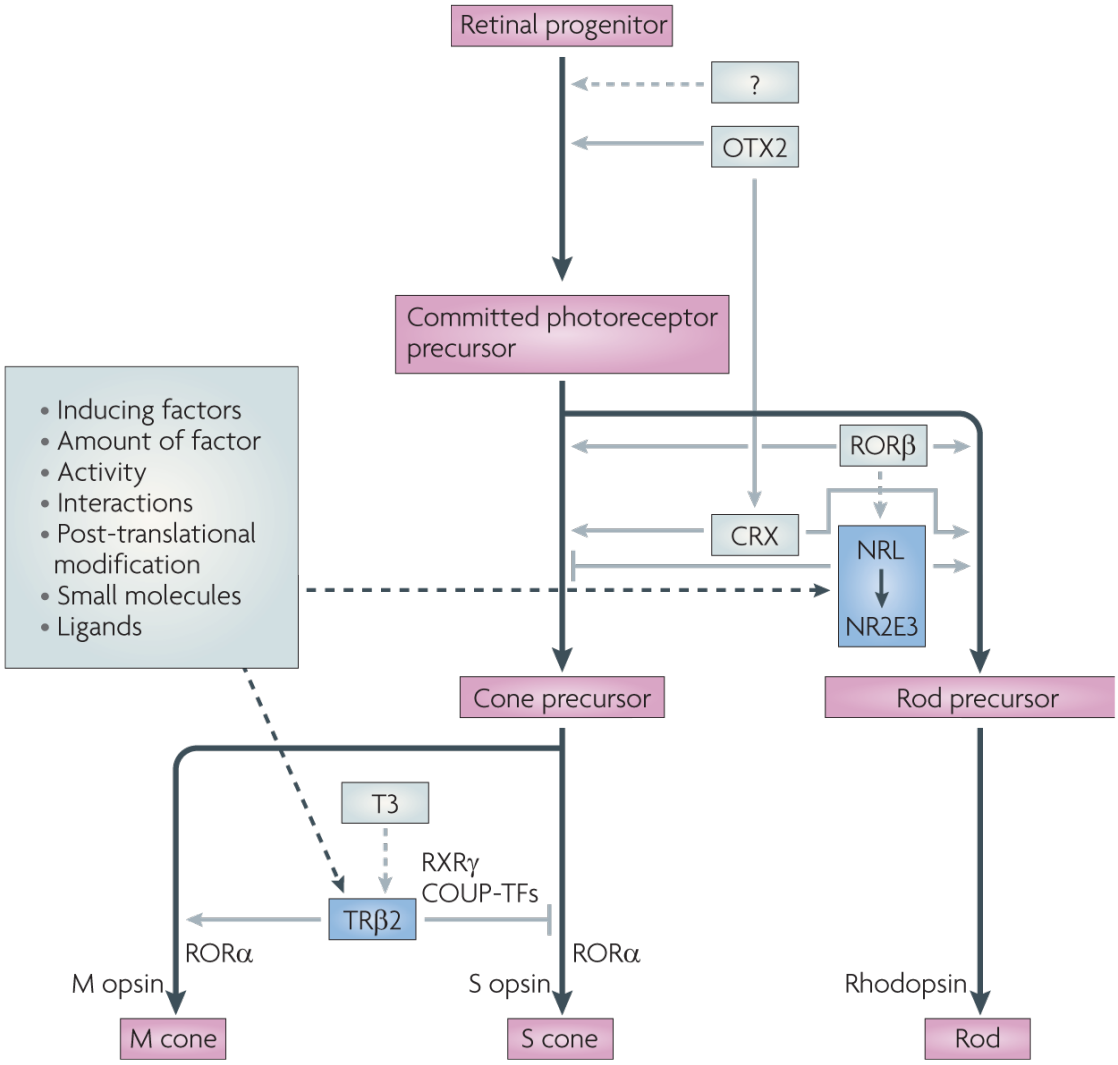

Figure 1.2. "Transcriptional dominance model of photoreceptor cell fate determination". Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Neuroscience (citation), copyright (2010). "A generic photoreceptor is formed under the control of homeobox protein OTX2 and other undetermined signals. This precursor is programmed to possess a 'default' S cone state, unless diverted by additional signals" [36].

Lacking this complete information, several groups have been successful in using ESCs and iPSCs to establish and apply *in vitro* protocols to derive cells that exhibit definitive properties of photoreceptors both when tested in culture and in transplantation

studies [14,15,19,20,53,54]. Meyer et al. [19] established a protocol for the derivation of photoreceptor cells from pluripotent stem cells that capitalized on known stages of development, and beautifully demonstrated a strong correlation between both the timecourse and gene and protein expression profiles of characteristic markers between their *in vitro* derived cells and normal *in vivo* development. This indirect reprogramming utilizing iPSCs demonstrated important proof-of-principle that photoreceptor cells can be derived *in vitro* and has since served as a standard in the field for the derivation of these cells [55, 56].

Direct reprogramming protocols however, have been much less arduous than those used to first establish iPSCs and then further differentiate them to the desired somatic cell type. Because a wealth of information exists about the underlying gene regulatory network governing photoreceptor development, photoreceptors are highly suited for direct reprogramming. Thus, efforts described within this thesis sought to capitalize on established systems and unique models in the fields of photoreceptor development and direct cellular reprogramming, with aims to establish approaches leading to the direct differentiation of rod photoreceptors from somatic cells.

2 ESTABLISHMENT OF A NOVEL SYSTEM FOR DERIVATION OF PHOTORECEPTORS VIA DIRECT REPROGRAMMING

The work herein described is aimed to design, characterize, establish and provide preliminary results on a system aimed at testing the hypothesis that somatic cells can be directly reprogrammed to a rod photoreceptor fate *in vitro*. The overall experimental aims include: the determination of candidate genes for reprogramming, cloning of these candidate genes into appropriate vectors, adaptation of a lentivirus system for gene delivery, generation of cells to use as a high-throughput screening-system for analysis of virally-infected or transfected somatic cells, providing proof-of-principle that these constructs lead to induction of gene and protein expression, collection of preliminary data demonstrating neuralization of somatic cells induced by combinations of known pro-neural genes, and finally, the induction of photoreceptor-like phenotypes in somatic cells.

2.1 Selection of candidate genes

While many of the genes and proteins involved in photoreceptor development and maintenance have been identified and are well-characterized, a relatively blind approach was undertaken to identify known and potentially novel transcription factors that govern these processes. Briefly, a list of candidate transcription factors was determined using published microarray datasets specific to early- and late-born rod photoreceptors.

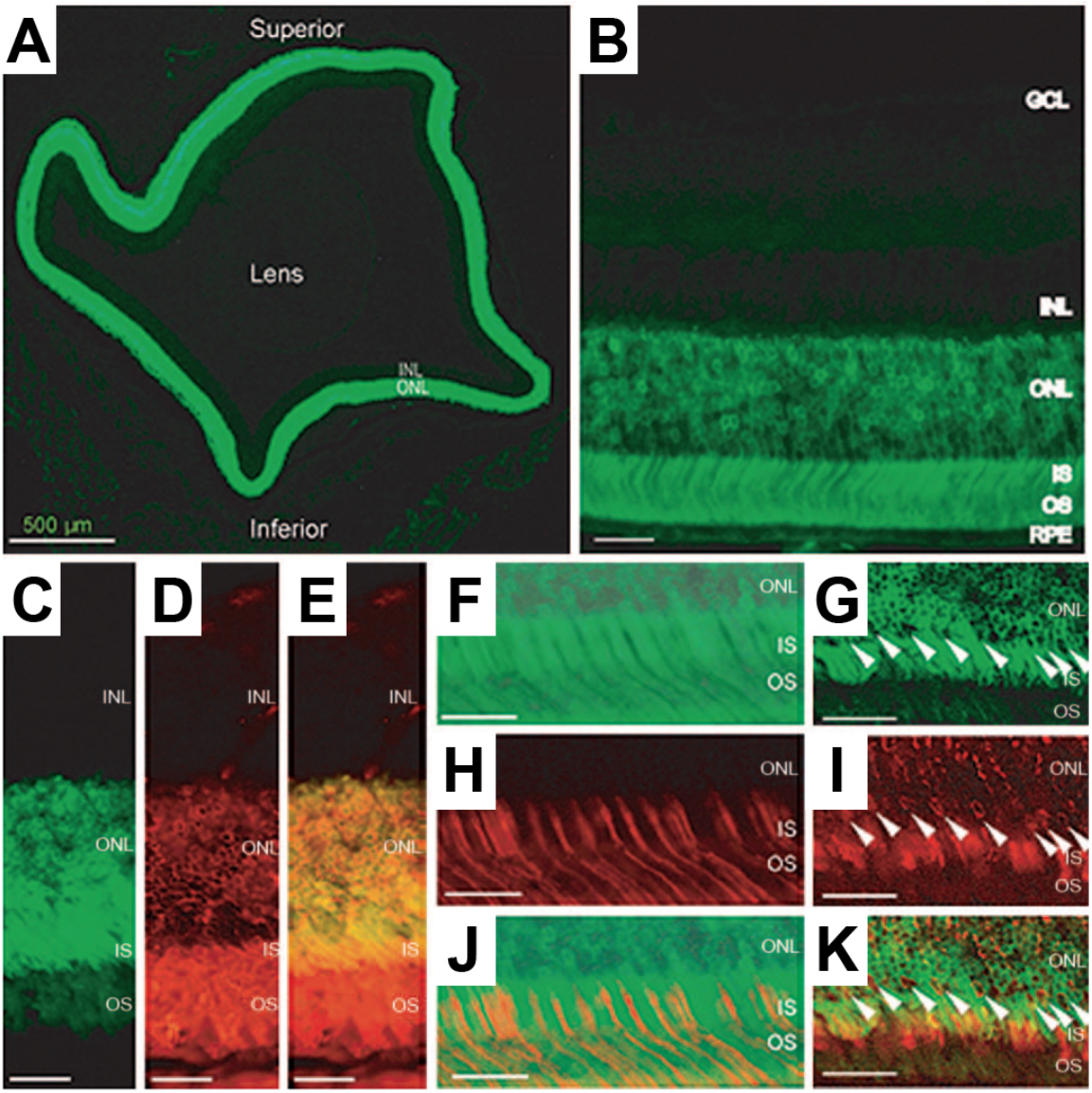

Figure 2.1. *NRL* promoter drives GFP-expression in rod, but not cone, photoreceptor cells in the retina. Copyright (2006) National Academy of Sciences, U.S.A. [57]. This figure is reprinted with permission from PNAS. (A) shows the specificity of GFP-NRL to the outer nuclear layer (ONL) of the entire adult retina in the GFP-NRL knock-in mouse. (B) demonstrates that not all cells in the ONL express GFP. (C, D, E) show immunolabelling with the rhodopsin antibody (red-D) completely overlaps (E) with GFP expression (C). (F, H, J) GFP expressing cells (F) show no overlap (J) with cells expressing the cone-specific marker, peanut agglutinin (H). (G, I, K) Photoreceptor cells are indicated by arrowheads, and cells expressing the cone-specific marker arrestin (red) shows no overlap with GFP-expressing (green) cells (K). This figure in sum demonstrates the specificity of NrL-promoter driven GFP-expression to rod photoreceptors in the retina [57].

Nrl is a basic motif-leucine zipper transcription factor that is specifically expressed in rod photoreceptors (Figure 2.1) [57, 66, 67]. Deletion of *Nrl* in mice results in a cone-only phenotype in the retina [51,57,68]. Interactions of *Nrl* with *Crx* and *Nr2e3*, along with other proteins, coordinate the expression of rod-specific genes [57–65] and down-regulate cone-specific gene expression [63, 64, 69, 70].

Microarray datasets utilizing FACS-sorted rod photoreceptors, enabled by a GFP-reporter construct under control of the *Nrl*-promoter, provided gene expression profiles at varying developmental time points [57]. Embryonic day 16 (E16) and postnatal day 2 (P2) were the earliest datasets collected, broadly reflecting genes expressed in early- and late-born rods, respectively [36, 42, 57, 71]. These datasets (GSE4051: GSM92633-36, and GSM92641-44, n=4) were then mined for transcription factors as identified by the RIKEN mouse library, yielding the following preliminary list of candidate transcription factors (Table 2.1).

Table 2.1
Number of transcription factors determined from data-mining listed by criteria

E16	P2	50% or more probes	All probes
Present	Present	494	222
Absent	Present	34	15
Present	Absent	148	71
Total transcription factors present		676	308

Data-mining of the E16 and P2 microarray data sets [57] yielded the above lists of transcription factor probes that were present or absent at each timepoint in 50% or more, or in all probes surveyed. For example, the first line is explained as 50% or more probes identified 494 transcription factors 'present' in the E16 as well as the P2

datasets. Of these 494 transcription factors, all probes identified 222 transcription factors 'present' in the E16 as well as the P2 dataset. In contrast, the second line is read as 34 transcription factors were identified as present by 50% of probes in the P2 dataset, but *absent* in the E16 dataset. Of these 34 transcription factors, 15 of them were found present in the P2 dataset but absent in the E16 date set using all probes. The total of 676 genes identified by 50% or more probes were further investigated. Mouse transcription factors were downloaded from RIKEN Mouse Transcription Factor Database.

A total of 676 transcription factors were identified as being expressed at either or both timepoints in 50% or more probes. This list was further narrowed by eliminating redundancy and cross-referencing published literature on each transcription factor identified to determine its potential role in governing photoreceptor development or direct reprogramming. The following criteria were used to identify candidates from the narrowed list of transcription factors: genes that are known to be 1) specifically expressed in neural tissue, 2) important for neural development, 3) implicated in epigenetic remodeling, 4) specifically expressed in the retina, and/or 5) important in retinal development. Using these specific criteria, 23 candidate genes were identified for their potential role in photoreceptor reprogramming (Table 2.2). Importantly, this search revealed many expected transcription factors as established by previous studies on rod photoreceptor differentiation and maintenance (e.g. *Crx, Nrl*), but also some novel candidates as well (e.g. *Blimp1, Sp4*).

Table 2.2
The list of 23 candidate genes and accession numbers, with their corresponding length in nucleotides

Candidate genes	Accession Number	Length in nucleotides
ASCL1	CCDS24101.1	696
BLIMP1	NM007548.3	2651
BMI1	BC056384.1	974
BRN2	CCDS18005.1	1338
CRX	NM007770.4	972
CTCF	BC049131.1	2211
DACH1	BC141130	2100
HDAC1	NM008228	1449
LHX2	BC055741.1	1221
MYBBP1A	BC052889.1	4035
MYT1	BC063252.1	3384
NEUROD1	NM010894.2	1074
NEUROD4	BC054391.1	993
NEUROGENIN2	BC055743.1	792
NR2E3	BC017521.1	1188
NRL	BC031440.1	714
OLIG2	BC051967.1	972
OTX2	BC017609.1	870
RAX	CCDS29311.1	1029
RORβ	BC024842.1	1344
SIX6	NM011384.4	741
SP4	NM011384.4	2349
ZIC1	BC060247.1	1344
Total gene basepairs cloned		34255

2.2 Establishment of a screening system for candidate genes

As a combinatorial approach utilizing a pool of 23 factors would yield a nearly insurmountable set of data to analyze given the number of combinations of a set number of transcription factors that can be designed, it was important to establish a screening system to narrow the pool of potential candidates with a high-throughput, efficient approach. As discussed previously, the initial studies demonstrating direct reprogramming used mouse embryonic fibroblasts (MEFs), in part due to their demonstrated high efficiencies in iPSC reprogramming [10]. For the purposes of reprogramming to rod photoreceptors, MEFs were derived from mice that have a Rhodopsin-GFP fusion knock-in as a replacement for native rhodopsin [33]. Since rhodopsin is specifically expressed in rod photoreceptor cells (Figure 2.2A) [33], MEFs derived from these animals will express the GFP-fusion protein only when they express the rod-specific gene, *Rhodopsin*. These MEF cells were used for experiments testing the ability of the 23 candidate transcription factors to induce expression of Rhodopsin-GFP (Figure 2.2B). Because *Opsin* gene expression is one of the final stages in photoreceptor differentiation [36], a GFP signal from these cells in culture can serve as a reliable screening tool for the conversion of fibroblasts to a rod photoreceptor identity.

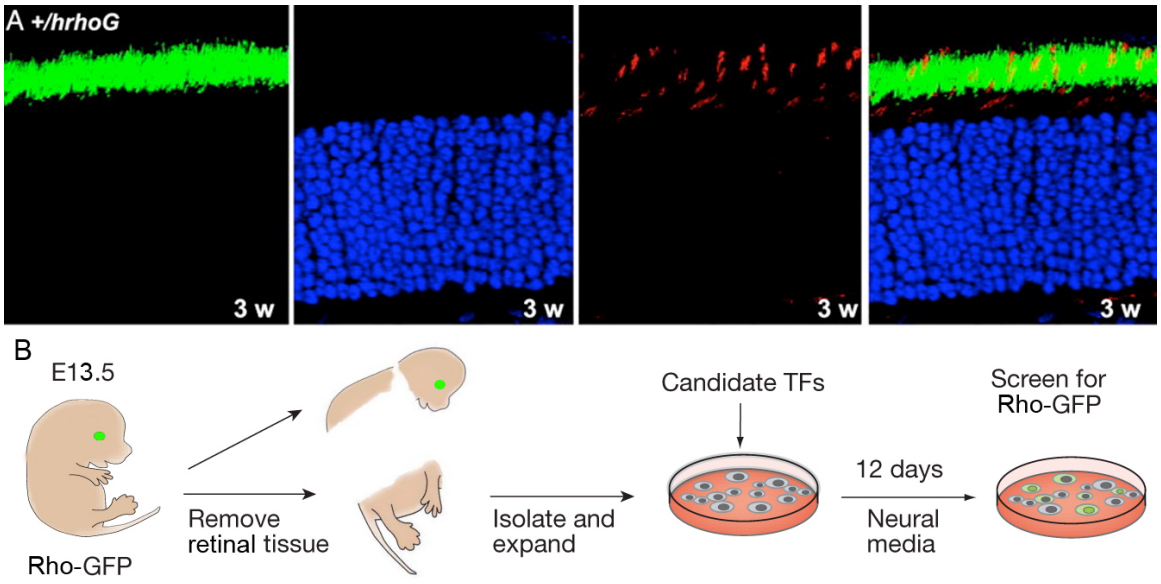

Figure 2.2. Specificity of Rhodopsin-GFP fusion to rod photoreceptor cells and schematic of experimental design utilizing MEFs derived from rhodopsin-GFP mice. (A) Retinal sections from 3-week-old mice expressing the rhodopsin-GFP fusion protein demonstrate left to right the expression of the GFP in rod photoreceptor outer segments in the outer nuclear layer (ONL), DAPI-labeled nuclei of the ONL (blue), rhodamine peanut aggluttin staining of cone sheaths (red), and a merge image demonstrating no overlap of GFP-expression with rhodamine-labeled cells [33]. Copyright (2004) National Academy of Sciences, U.S.A. [33]. Figure A is reprinted with permission from PNAS. (B) Experimental schematic adapted from [18] delineating the usage of MEF cells derived from Rhodopsin-GFP fusion knock-in mice to screen candidate transcription factors for their ability to directly reprogram fibroblasts to rod photoreceptor cells.Figure B was adapted by permission from Macmillan Publishers Ltd: Nature [18], copyright (2010).

2.3 Lentiviral expression construct modifications

A third-generation lentivirus system, utilizing a modified pCSCIGW vector as the transfer vector, was used for cloning of candidate genes and subsequent virus production.These plasmids were generously provided by Dr. Scott Witting, from the

Indiana University School of Medicine. The expression construct originally contained a cytomegalovirus (CMV) promoter and an IRES-GFP reporter sequence following the gene of interest (Figure 2.3).

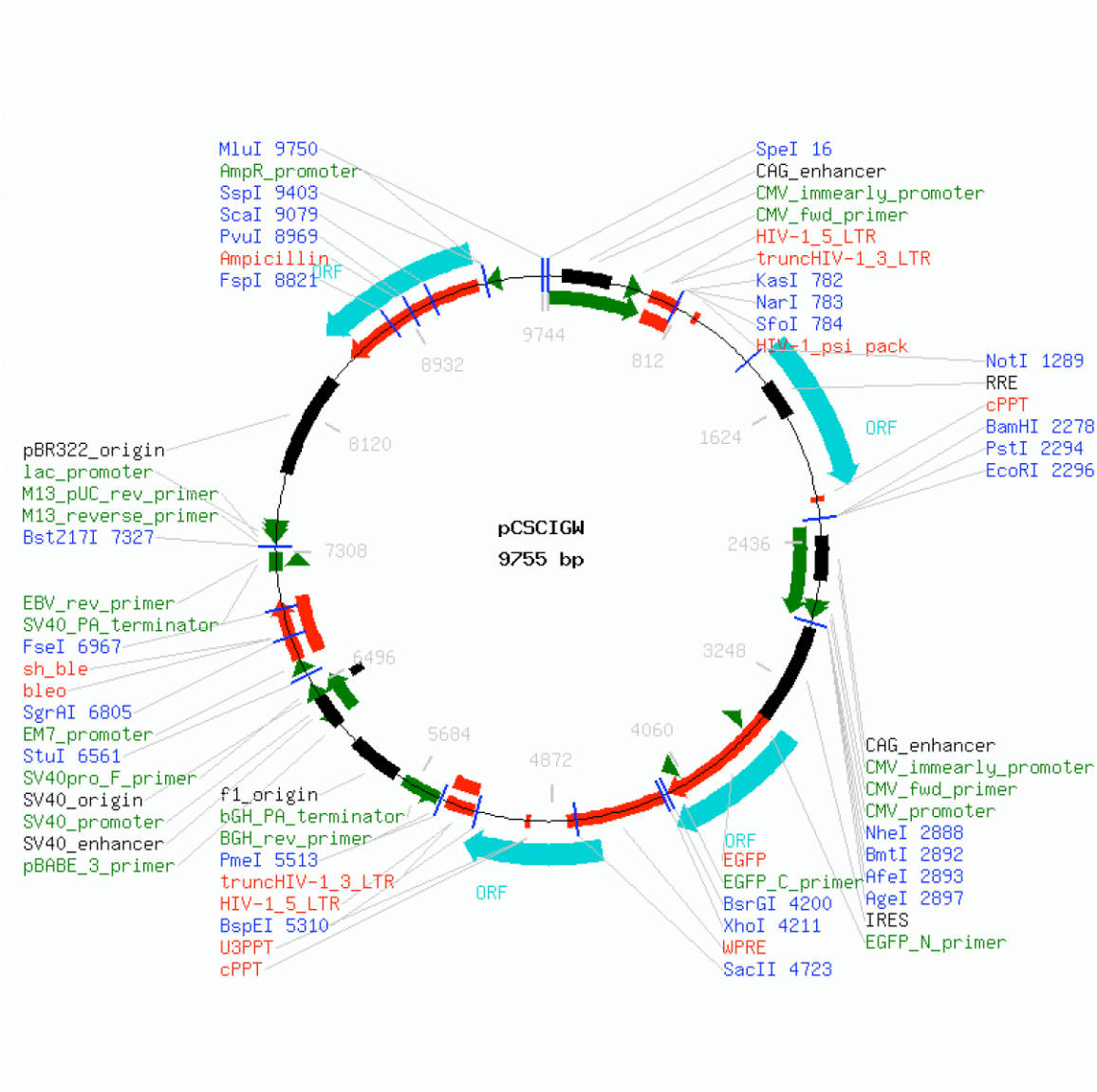

Figure 2.3. The modification and confirmation of the viral expression construct. The original viral expression plasmid map, pCSCIGW, provided by Dr. Scott Witting of Indiana University School of Medicine

Published data suggested that a more appropriate promoter for these applications would be that of phosphoglycerate kinase (PGK), because of demonstrated higher expression in undifferentiated and neural cell types, whereas the CMV promoter has been known to be silenced in certain cell types [72, 73]. Therefore, the CMV promoter was cut out, and replaced with the PGK promoter that was PCR-amplified from the pLenti-PGK-GFP plasmid, graciously provided again by Dr. Scott Witting. This PCR product was run out on a 1% low melting point agarose gel (Figure 2.4), was purified using the Wizard Gel Purification Kit (Promega), and ligated to the pGEM bacterial expression vector (Promega). JM109 bacterial cells (Promega) were transformed with this construct per the manufacturer's instruction and cells were plated overnight on selective plates. The following day, single colonies were picked, and grown for 16 hours at 37 °C. DNA was isolated from these bacterial cultures using the QIAprep Spin Miniprep kit (Qiagen), and the PGK promoter was excised using EcoRI and NheI cut sites at the 5′ and 3′ ends respectively. This product was was separated from the pGEM backbone on a 1% low melting point agarose gel, was further purified using the Wizard Gel Purification Kit (Promega) and an overnight ligation reaction was used to insert it into the pCSCIGW backbone that had also been gel-purified after being linearized using EcoRI and NheI (Figure 2.5). JM109 bacterial cells (Promega) were transformed with this ligation product per manufacturer′s instruction, and plated overnight on selective plates. The following day, ten single colonies were picked, and grown for 16 hours at 37 °C. Plasmid DNA was isolated from these bacterial cultures using the QIAprep Spin Miniprep kit (Qiagen). AscI and MluI were initially used to confirm the PGK promoter insertion in these clones, as there was one MluI recognition sequence on the pCSCIGW backbone, and the PGK promoter itself had one novel AscI recognition site within its sequence. This test identified 2 clones that yielded the two predicted fragments of 2716 bps and 7215 bps (Figure 2.6). The current sequence was confirmed by the Indiana University DNA Sequencing Core Facility using custom sequencing primers (Figure 2.7). Due

to the 5′ and 3′ long-terminal repeats (LTRs) on the lentiviral expression plasmid, recombination-incompetent (mcrB mrr genotype) Stbl3 cells (Invitrogen) were used for all subsequent transfections containing the lentivirus construct.

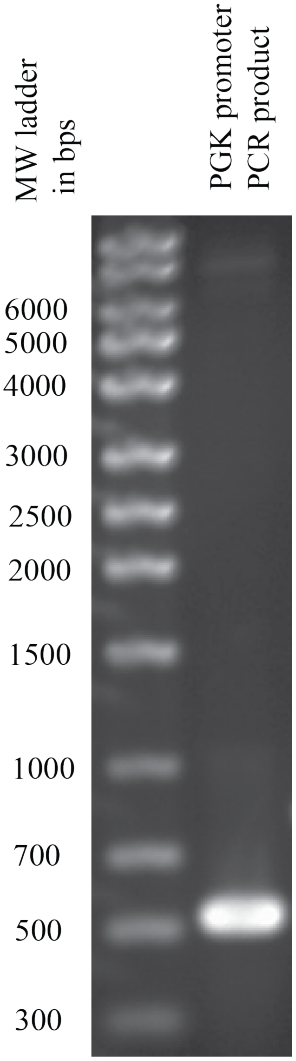

Figure 2.4. Agarose gel showing the PCR-amplified PGK promoter used for ligation to pCSCIGW backbone

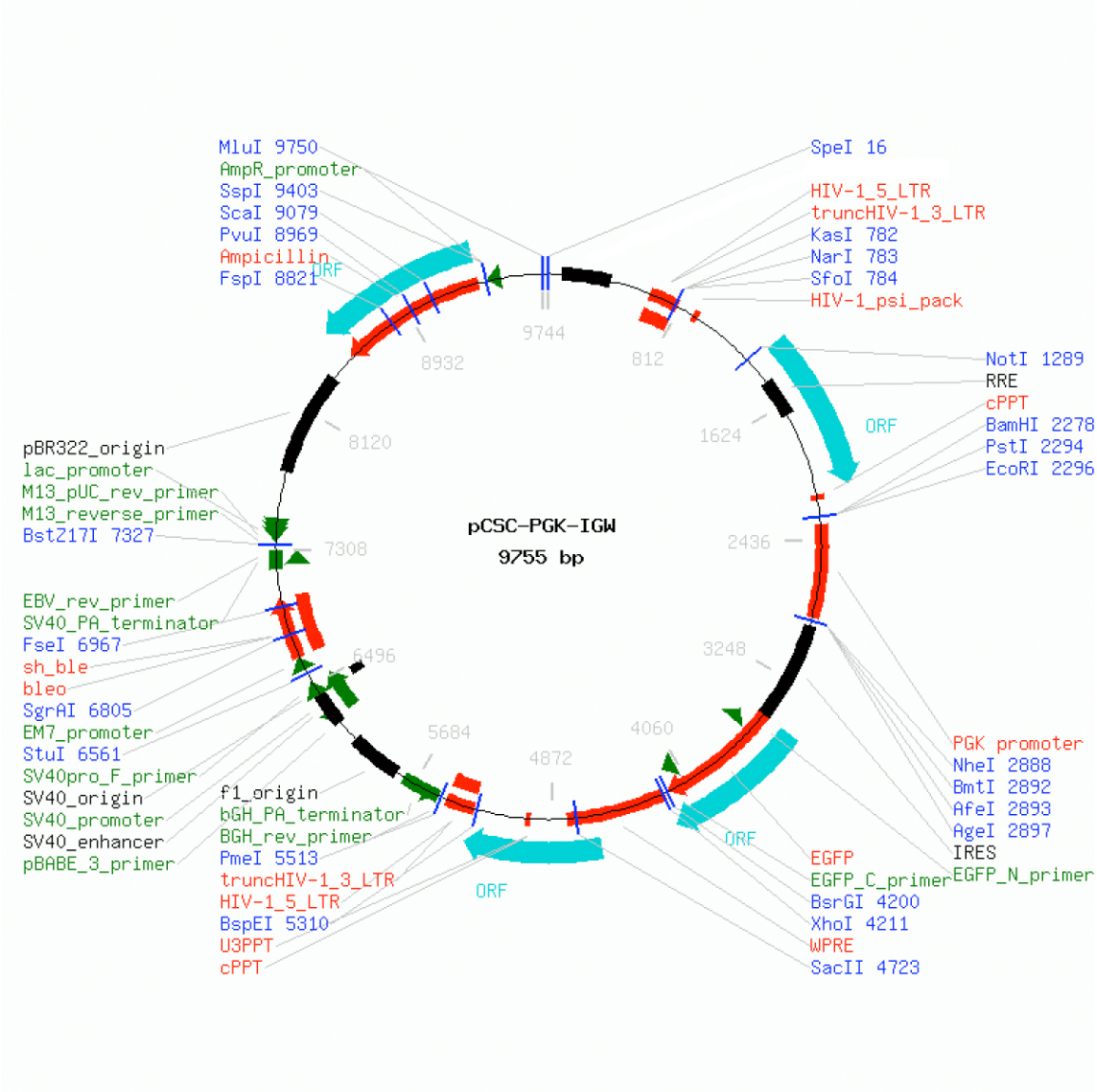

Figure 2.5. The modified pCSC-PGK-IGW viral expression construct, after the PGK promoter was inserted to replace the CMV promoter. This construct was used for all gene insertion and expression experiments.

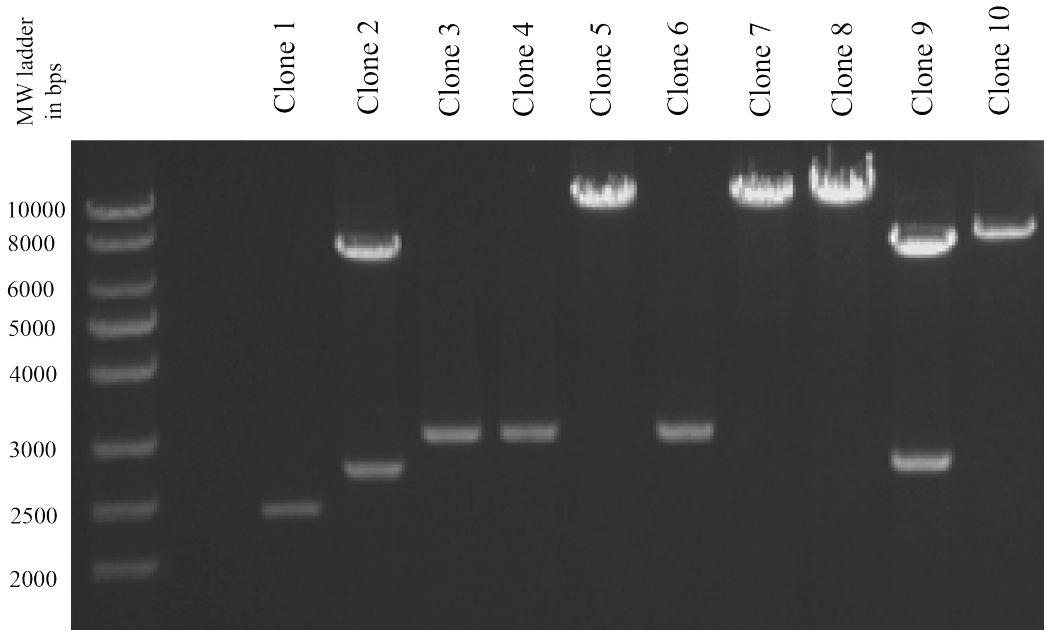

Figure 2.6. Agarose gel showing clones 2 and 9 of the pCSC-PGK-IGW ligation cut with MluI and AscI yielding the 2 expected products. Both of these clones had the PGK-promoter correctly sequenced in full by the Indiana University Core Facility using custom sequencing primers.

FORWARD 5'- GGT GGA GAG AGA GAC AGA GAC AG -3'

REVERSE 5'- CCT TGC ATT CCT TTG GCG AGA G -3'

Figure 2.7. Custom sequencing primers used to sequence confirm the PGK promoter in the pCSC-PGK-IGW backbone.

As the IRES-GFP sequence on the original pCSCIGW plasmid was not only unnecessary, but would interfere with the proposed screening system utilizing rhodopsin-GFP reporter MEF cells, multiple cloning sites on the construct were used to introduce genes of interest while simultaneously removing the IRES-EGFP segment of the pCSC-PGK-IGW construct, including NheI, BmtI, AfeI, or AgeI 5′ to the insertion

site, as well as BsrGI or XhoI 3′ to the IRES-EGFP sequence(Figure 2.5). This modified expression construct and multiple cloning sites were used for all gene expression experiments.

2.4 Cloning strategies for the 23 gene candidate constructs

2.4.1 PCR amplification techniques

Initially, commercially-available gene constructs on bacterial plasmids were ordered for all available genes to be used for PCR amplification, gel purification, and subsequent ligation to the lentiviral expression construct (Open Biosystems, Source BioScience). Multiple cloning sites (MCSs) on both the 5′ and 3′ ends of the gene insertion site enabled the construction of primer pairs designed with restriction enzyme recognition sequences that were not present in the sequence of the gene of interest. Primers for PCR amplification of genes were designed with restriction enzyme recognition sequences that were immediately followed by insertion of the first six bases of Kozak sequence (GCCRCCAUGG) on the 5′ end, followed by complementary bases of the 5′ and 3′ gene sequence, respectively. The restriction enzyme recognition sequences were added to enable the gene′s insertion into the lentiviral expression construct, while the Kozak sequence (the six bases added in addition to the start codon of the gene) was included to enhance the efficiency of translation of the gene product of interest after delivery to somatic cells (Tables 2.3, 2.4).

PCR protocol modifications to obtain the correct gene sequences were necessary for some genes included the addition of PCR enhancers, such as Betaine or DMSO, as well as varying concentrations of template, polymerase, and magnesium chloride. Additionally, PCR reactions were performed using a gradient of temperatures to identify optimal conditions for each gene′s amplification. For example, PCR amplification using primers designed for *Olig2* initially yielded no product of the right size (986 bps)

Table 2.3
Primers used for PCR-amplification of genes that add restriction enzyme recognition sequences to the 5′ and 3′ ends. All primers are listed 5′ to 3′.

CTCF	
FORWARD	ATA TAT GCT AGC GCC ACC ATG GAA GGT GAG GCG GTT GA
REVERSE	ATA TAT CTC GAG GCC ACC TCA CCG GTC CAT CAT GCT GAG G
LHX2	
FORWARD	ATA TAT GCT AGC GCC ACC ATG CTG TTC CAC AGT CTG
REVERSE	ATA TAT CTC GAG GCC ACC TTA GAA AAG GTT GGT AAG AGT CGT TTG TGA AG
MYBBPIA	
FORWARD	ATA TAT ACC GGT GCC ACC ATG GCG GAG ATG AAG AGC C
REVERSE	ATA TAT TGT ACA GCC ACC TCA AGG TGT CTG CAC TCT CCT GCT GG
MYT1	
FORWARD	ATA TAT GCT AGC GCC ACC ATG AGC TCA GAA AGT GAT GAC AAG C
REVERSE	ATA TAT CTC GAG GCC ACC CTA GAC CTG AAT GCC CC
NEUROD1	
FORWARD	ATA TAT ACC GGT GCC ACC ATG ACC AAA TCA TAC AGC GAG A
REVERSE	ATA TAT CTC GAG GCC ACC CTA ATC GTG AAA GAT GGC ATT AAG CT

Table 2.4

Primer Sequences(N-O)

NEUROGENIN2	
FORWARD	ATA TAT ACC GGT GCC ACC ATG TTC GTC AAA TCT GAG ACT CTG GAG
REVERSE	ATA TAT CTC GAG GCC ACC CTA GAT ACA GTC CCT GGC G
NRL	
FORWARD	ATA TAT GCT AGC GCC ACC ATG GCT TTC CCT CCC AGT C
REVERSE	ATA TAT CTC GAG GCC ACC TCA GAG GAA GAG GTG TGT GTG GTC
NR2E3	
FORWARD	ATA TAT GCT AGC GCC ACC ATG AGC TCT ACA GTG GC
REVERSE	ATA TAT CTC GAG GCC ACC CTA GTT TTT GAA CAT GTC ACA CAG GAG
OLIG2	
FORWARD	ATA TAT GCT AGC GCC ACC ATG GAC TCG GAC GCC AG
REVERSE	ATA TAT CTC GAG GCC ACC TCA CTT GGC GTC GGA GGT GAG
OTX2	
FORWARD	ATA TAT GCT AGC GCC ACC ATG ATG TCT TAT CTA AAG CAA CCG CCT
REVERSE	ATA TAT CTC GAG GCC ACC TCA CAA AAC CTG GAA TTT CCA TGA GGA

visible on a 1% agarose gel. However, addition of 5% DMSO lead to a light band of this size visible, and addition of 10% DMSO at 62.5 °C yielded the most specific product (Figure 2.8). The original protocol and these modifications are covered in detail in the methods chapter (CHAPTER 3).

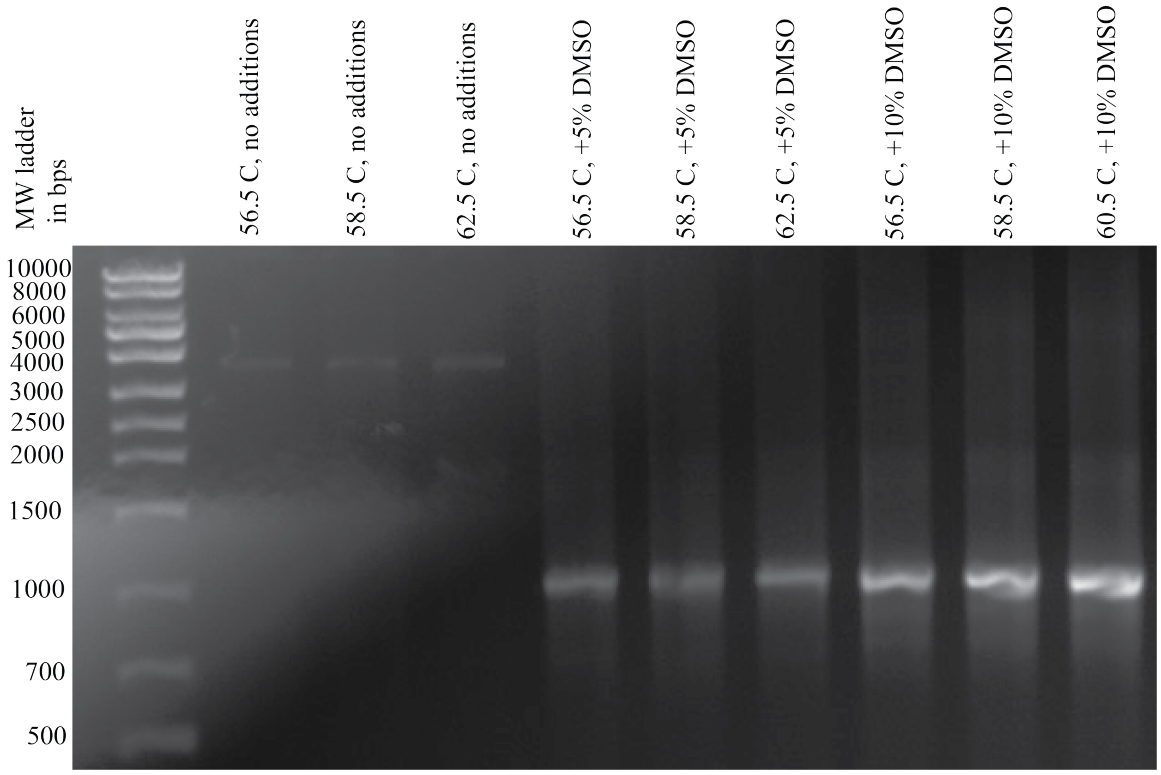

Figure 2.8. PCR-amplification and optimization experiments for the *Olig2* gene on an 1% agarose gel

Successfully PCR-amplified products were separated on a 1% low melting point agarose gel alongside the pCSC-PGK-IGW vector that had been digested with corresponding restriction enzymes. These were gel-purified and ligated together. Stbl3 bacterial cells were transformed with this ligation product per the manufacturer′s instruction, and plated overnight onto selective plates. The following day, single colonies were picked, and grown for 16 hours at 37 °C. Plasmid DNA was isolated from these bacterial cultures using the QIAprep Spin Miniprep kit (Qiagen). This

DNA was test cut using the restriction enzymes with recognition sequences at the 5′ and 3′ ends, and run out on a 1% agarose gel. DNA isolated from clones that had excision products corresponding to the expected size of the inserted genes were sequenced by the Indiana University DNA Sequencing Core Facility.

While PCR protocol modifications lead to some additional genes being successfully amplified and ultimately sequence-confirmed, many genes of interest were simply not amenable to PCR amplification under any conditions tested, with or without protocol modifications. All failures at PCR amplification were likely attributable to GC-rich regions or templates of these particular genes.

Therefore, for several remaining gene sequences, a second strategy was utilized: genes were cut out of the commercially provided bacterial expression construct with appropriate restriction enzymes, and ligated into a bacterial transfer vector, either the pGEM (Promega), pLITMUS (New England Biosystems), or pCR2.1-TOPO system (Invitrogen). This provided an additional method to add restriction enzyme recognition sequences to the 5′ and 3′ ends of these genes that were compatible with the modified pCSC-PGK-IGW lentiviral construct. Gene sequence compatibility with available restriction sequences, as well as restriction sites on the original plasmids, determined the system used for each remaining gene independently. Similar to PCR amplification, this strategy was successful for several genes, but still, not all remaining gene sequences were compatible with these systems.

2.4.2 Direct commercial custom gene synthesis

The final subset of genes of interest were either not at all commercially or academically available, or were unable to be synthesized/modified by any of the previous strategies. Thus, these gene sequences were obtained through custom gene synthesis approaches (Integrated DNA Technologies) that included insertion on a bacterial ex-

pression vector. Such gene synthesis included sequence verification of both strands, as well as plasmid mapping and generation of a FASTA sequence file. For the remaining genes of interest, restriction enzyme sequences not contained within the gene but within the multiple cloning sites of the lentiviral expression construct were added to the 5′ and 3′ end of the gene sequence. These plasmids were expressed in chemically-competent JM109 (Promega) cells, the DNA purified, and the gene of interest excised by restriction digest. Successfully excised gene products were separated on a 1% low melting point agarose gel alongside the modified pCSC-PGK-IGW vector that had been digested with corresponding restriction enzymes. These were gel-purified and ligated together. Stbl3 bacterial cells were transformed with this ligation product per the manufacturers instructions, and plated overnight onto selective plates. The following day, single colonies were picked, and grown for 16 hours at 37 °C. Plasmid DNA was isolated from these bacterial cultures using the QIAprep Spin Miniprep kit (Qiagen). This DNA was test cut using the restriction enzymes with recognition sequences at the 5′ and 3′ ends, and separated on a 1% agarose gel. DNA isolated from clones that had excision products corresponding to the predicted size of the inserted genes were sequenced by the Indiana University DNA Sequencing Core Facility.

2.5 Sequence-confirmation of lentiviral expression constructs

Prior to the use of lentiviral expression vectors containing candidate genes, it was necessary to sequence verify the gene sequences to ensure proper transcription of these transcription factors. Sequencing primers were thus designed approximately 100 basepairs up- and down-stream of the gene insertion site of the lentiviral construct (Table 2.5). Each of the 23 gene constructs was sequenced-confirmed by sequencing services provided by the Indiana University DNA Sequencing Core Facility. Most sequence analyses confirmed approximately 500-800 base pairs in either direction, so for several of the larger gene constructs, specific internal primers were designed to confirm the entire internal sequence (Tables 2.6, 2.7). For each of 23 gene constructs,

complete sequence confirmation was obtained for every nucleotide. The lone exception to this was the Mybbp1a gene (4035 bps), which had been PCR-amplified and had a single nucleotide base change at base 2804, an A to G substitution, resulting in a change in amino acid at residue 1112, from lysine to arginine, maintaining the positively charged residue at this amino acid.

Table 2.5
Custom sequencing primers for all gene insertions into the pCSC-PGK-IGW expression plasmid

FORWARD	5′ AATAGCGGCTGCTCAGCA 3′
REVERSE	5′ GGCATTAAAGCAGCGTATCCAC 3′

2.6 Restriction enzyme-excision confirmation of large-scale plasmid DNA

Once each clone was sequence-confirmed, 100 mL cultures of bacteria expressing the construct were grown overnight at 37 °C, and the plasmid was subsequently purified. Each plasmid DNA sample was again confirmed by restriction enzyme excision of a product of the genes corresponding size, from a backbone of approximately 9000 kbs, to confirm that no contamination or mislabeling had taken place after plating from glycerol stocks (Figure 2.9).

Table 2.6
Custom sequencing primers for BRN2, BLIMP1 and CTCF

Internal Primer Name	Sequence(5′ to 3′)
BRN2542	ATC AAG CCC TCG GTG GTG GTA
BRN2899	ATG AGC CAC ACC ATG CAG A
BRN21197	GAA CAT GTG CAA GCT GAA GC
BLIMP1192	CTC TTG GAA AAA CGT GTG GG
BLIMP1688	AAA ACC TGG CTG CCT GTC AGA A
BLIMP11161	CTT CCG TCT TCC ACA ACT CCA A
BLIMP11652	CCA CAG TGC CTT CTC CCT TA
BLIMP12161	TGT GCC CTG CCA AGT TTA CGC AAT
CTCF501	GAA AGT GGG GGC CAA TGG AGA AGT GGA GAC ACT A
CTCF899	GCA GAG CAT TCA GAA CAG TGA CCC TCC TGA GGA ATC
CTCF1311	TCA TTG TCC CCA TTG TGA CAC TGT CAT AGC CCG A

Table 2.7
Custom sequencing primers for MYBPP1A and MYT1

Internal Primer Name	Sequence(5′ to 3′)
MYBBP1A7	GAG ATG AAG AGC CCC ACG AAA GCT
MYBBP1A500	ACC CCA ACC ACT TAC AGG GAC A
MYBBP1A994	GGG GAG AAC ATG GTT ATT TCT AAG CCC C
MYBBP1A1481	AGA CGA AGC AGC ACT TCT CCT T
MYBBP1A1997	AGG TGG TCC GGA GTG TAT TTG GTC
MYBBP1A2517	TGA GCA CCC CCT GAT CCT GGA ACT ACT TGA
MYBBP1A2999	ACC CAG TGA TCT GTA AGA ACC TGC TTC C
MYBBP1A3507	GCG GAA GAA AAA GGG ATT CTT GCC AGA GAC C
MYT6	CTC AGA AAG TGA TGA CAA GCG AGC
MYT501	CCT CCT AAA TCT GGG CCA AA
MYT995	GCC CAG AGC TAT CTA GTC CTA AAC CT
MYT1501	AGC AAC CGC AAT ACT CAC AG
MYT2205	GGA ACT CTG GAC TTG AGC ATG CAT
MYT2487	TGA CAA GAG CCT CAG AAA CCT C
MYT3000	GGA GAA TGA TGA GGA GAT CAA GCA GC

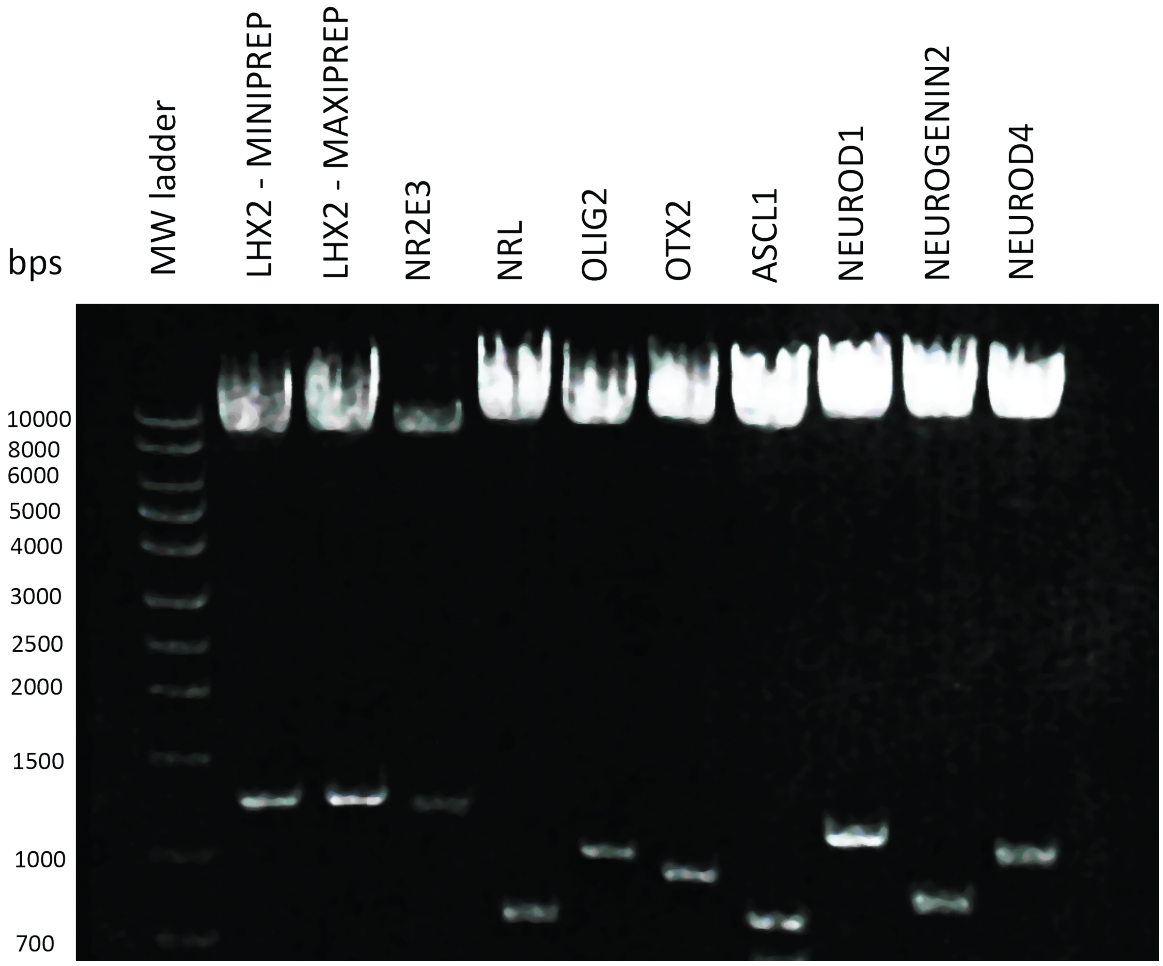

Figure 2.9. Proper gene excision from the pCSC-PGK-IGW backbone. Plasmid DNA was purified from 100 mL cultures after gene sequence confirmation, cut with restriction enzymes, and imaged on a 1% agarose gel.

2.7 Lentivirus production: protocol optimization

Lentivirus was chosen as the method of gene delivery due to the ability to yield a high percentage of stable transformations. Expression vectors containing candidate genes were transfected into HEK293 cells along with the third-generation viral packaging plasmids VSVG, GAG-POL and REV and viral supernatant was collected 36 hours later. Initial attempts at virus production yielded low-titer virus, as evidenced by low-

infection coupled with minimal cell death in treated cultures of cells. Optimization of the protocol included multiple changes to the initial method, including shortening the incubation time, precise cell counts and quantifications for the amount of viral plasmids added to HEK293 cells, and lead to infection rates over 90%, as demonstrated by EGFP-expression using a viral construct with the IRES-EGFP reporter intact for all protocol optimization experiments (Figure 2.10). This optimized protocol is detailed in the Methods Chapter.

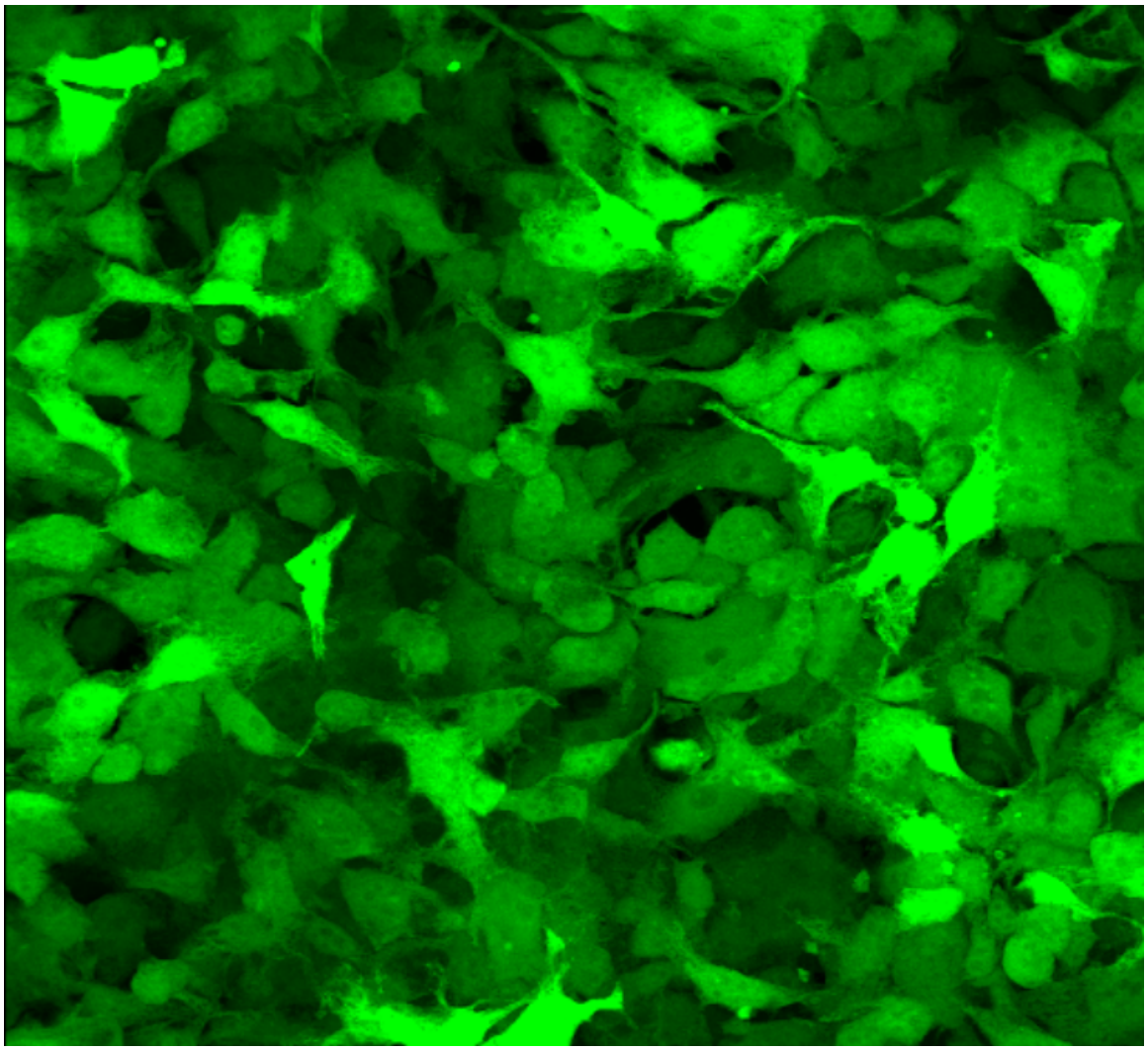

Figure 2.10. Optimization of the lentivirus production and delivery protocols yielded infection in over 90% of cells, as demonstrated by EGFP-expression in HEK293 cells 60 hours after infection with the pCSC-PGK-IGW construct with the IRES-EGFP reporter construct intact.

2.8 Demonstration of experimental feasibility and utility of constructs

To test the ability of candidate genes to be expressed in cells, and detected upon analysis, expression vectors of these candidate genes were initially transfected into HEK293 cells via calcium phosphate transfection. This protocol is detailed in Chap-

ter 3, methods. Approximately 60 hours following transfection, samples were collected for subsequent analysis. Commercially available antibodies exist for several of the proteins of interest, and preliminary immunocytochemistry experiments demonstrated upregulated protein expression when these constructs were delivered using a calcium phosphate transfection system. Immunostaining for ASCL1, LHX2 and OTX2 demonstrated upregulation of expression of each of these transcription factors when delivered individually as compared to controls (Figure 2.11). The proper nuclear localization of these candidate transcription factors was observed in a majority of cells after transfection. Induction of expression of these proteins thus demonstrated the viability of these constructs for subsequent direct reprogramming applications.

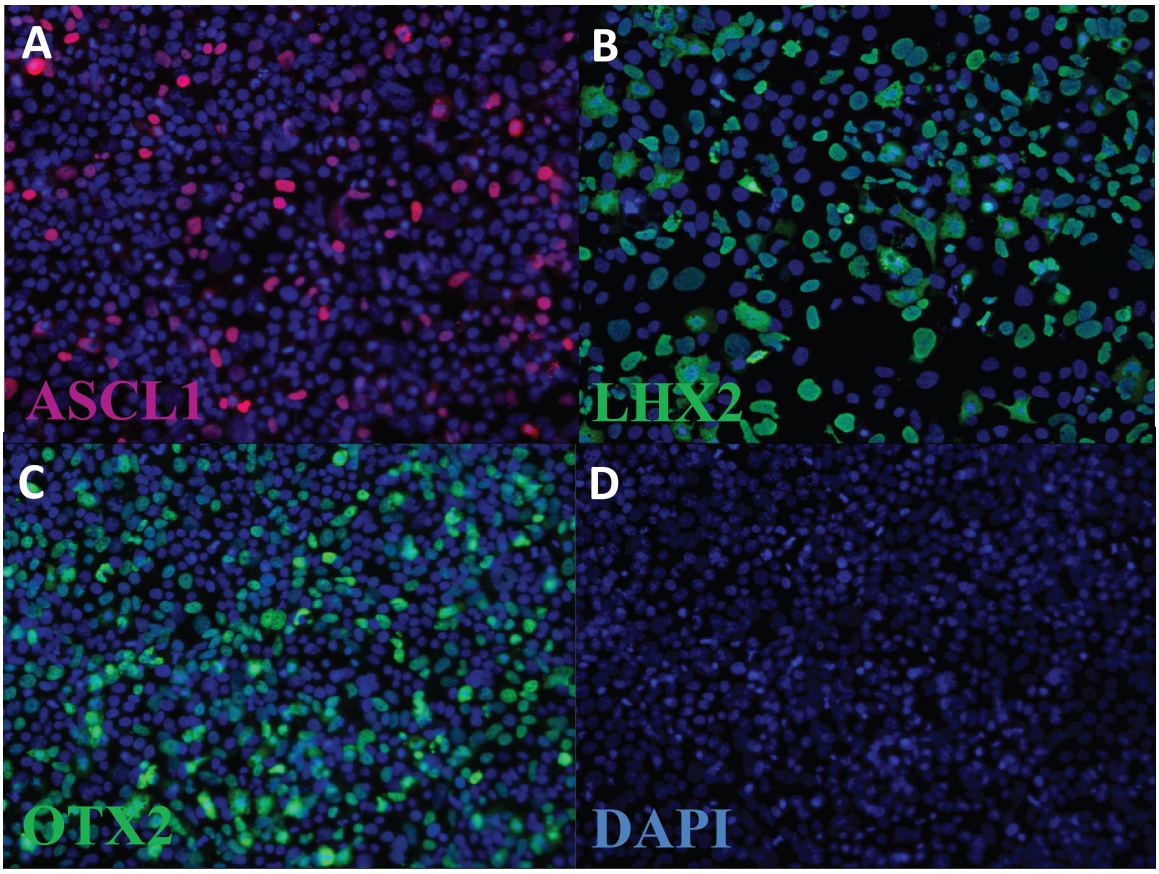

Figure 2.11. Upregulation of protein expression induced in HEK293 cells via calcium phosphate delivery of pCSC-PGK-IGW gene constucts. (A) ASCL1 (B) LHX2 and (C) OTX2 expression induced by construct delivery via calcium phosphate transfection to HEK293 cells, 60 hours post-transfection.(D) CONTROL image. (D) Negative control representation of no expression of ASCL1, LHX2 or OTX2 in untransfected cells via immunostaining. All nuclei are DAPI-labeled (blue).

2.9 Reprogramming of somatic cells through delivery of transcription factors

Previous studies have demonstrated the feasibility of cellular reprogramming to a neural lineage through the delivery of key transcription factors [18,22–26]. To demonstrate the ability of the synthesized constructs containing transcription factors to lead to neural reprogramming of MEF cells, the plasmids containing the transcription fac-

tors *Ascl1* and *Brn2* were delivered in combination to cultures of wild-type MEF cells. The next day, the culture medium was changed and cells were maintained for a total of 2 weeks with media changes every 2-3 days. After a total of 2 weeks post-infection, cells were fixed with 4% paraformaldehyde and analyzed by immunocytochemistry for the presence of neuronal phenotypes as a result of transcription factor-based reprogramming. At this timepoint, numerous cells had adopted elaborate neuronal morphologies, as well as the expression of the neuronal-specific protein βIII-tubulin (Figure 2.12A). Uninfected control experiments lacked the expression of βIII-tubulin.

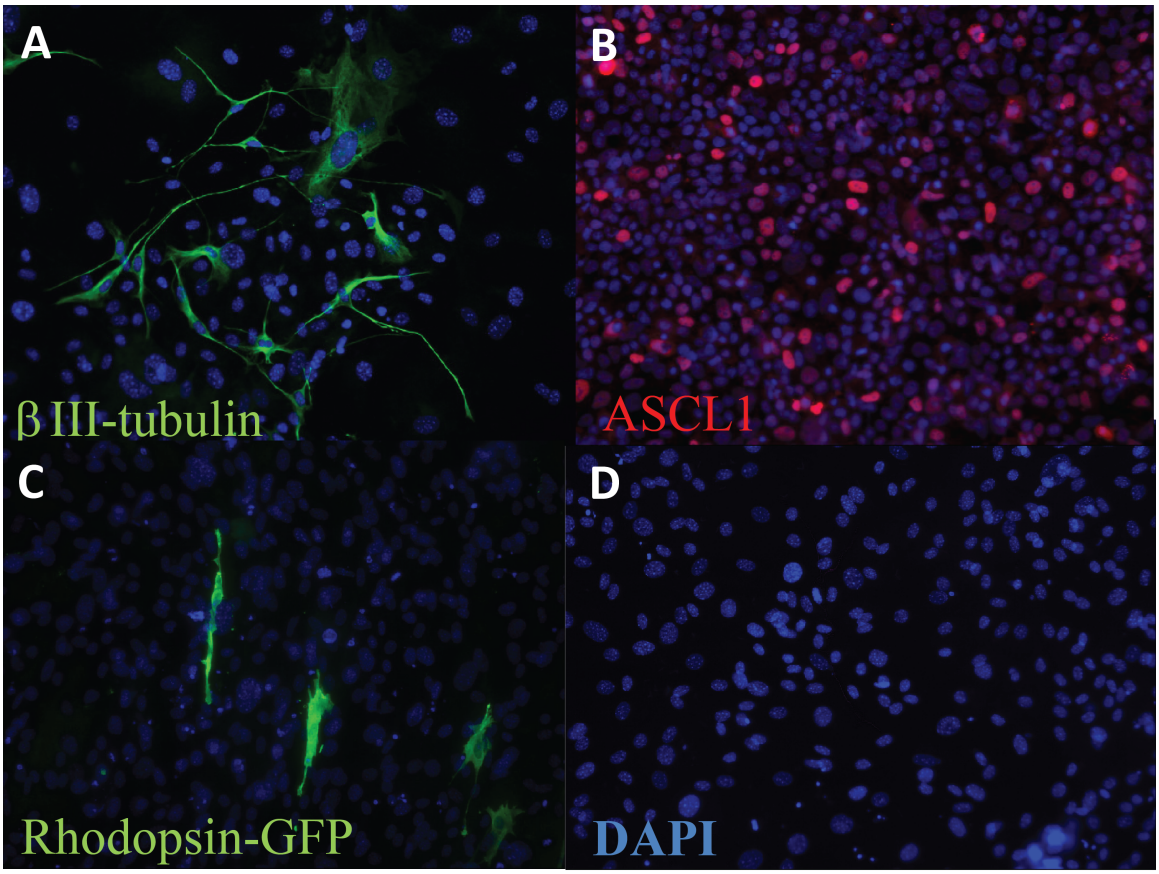

Figure 2.12. Downstream phenotypic and protein expression changes induced in MEF cells after construct delivery (A) βIII-tubulin expression in MEF cells, 2 weeks post-infection with a combination of Brn2 and Ascl1-containing virus. (B) *Ascl1* immunolabeling of cells in (A). (C) 60h post-infection with the Olig2-containing construct, MEF cells derived from the Rhodopsin-GFP fusion knock-in mouse were observed expressing GFP. (D) Negative control representative no expression of Ascl1, Brn2 or Olig2 in untransfected cells via immunostaining. All nuclei are DAPI-labeled (blue).

Given the ability to derive neuronal cell types through cellular reprogramming strategies, it was then asked whether or not other transcription factors could lead to the adoption of retinal photoreceptor-associated phenotypes. Using the previously identified candidate transcription factors described (Table 2.2), these genes were transfected into MEF cells derived from the Rhodopsin-GFP fusion knock-in mouse. Within 60

hours of following transfection, samples were collected for analysis via immunocytochemistry to test for the upregulation of transcription factors as well as potential morphological changes as a result of gene transfection. While most genes tested resulted in the upregulation of gene expression within these MEF cells (Figure 2.12B), it was observed that the introduction of the transcription factor *Olig2* lead not only to an increase in OLIG2 expression within these cells, but also to an apparent expression of green fluorescent protein, indicative of Rhodopsin-GFP expression throughout the cytosol, presumably suggesting that OLIG2 may play an important role in the activation of Rhodopsin expression (Figure 2.12C). This finding was exciting, as it importantly demonstrated both the utility of the experimental design and system utilized to identify candidate genes and demonstrated their effects on somatic cells, as well as the ability of designed constructs to induce rhodopsin-promoter driven expression of GFP in MEF cells. Only several months ago, elegant clonal lineage studies first identified Olig2 as a transcription factor expressed in a subset of retinal progenitor cells from specific stages of development, producing daughter cells biased towards amacrine cell interneurons and rod photoreceptors [30]. The current study described herein utilizing cellular reprogramming strategies was undertaken a year and a half before this finding was published, reinforcing the utility of choosing candidate genes based initially on the rod-specific microarray datasets and then narrowed by the criteria defined earlier in this thesis. As the remaining constructs are further tested and characterized using the Rhodopsin-GFP knock-in MEF cells, more exciting findings such as this stand to be discovered.

3 DETAILED METHODS

3.1 MEF derivation

The protocol used to derive MEF cells from the Rhodopsin-GFP fusion knock-in mice is from the website http://www.molgen.mpg.de [75]. Timed matings were set up, and fetuses were harvested at approximately embryonic day 13.5 of gestation. Pregnant female mice were euthanized with CO_2, and cervically dislocated to ensure death. A vertical superficial incision was made in the abdomen. The fetuses were dissected from the uterus individually and put in a 10 cm dish of 1XPBS. One fetus at a time was cut and washed in a new 10 cm plate of PBS. Each embryo was dissected out of the amnion and placenta and put in its own 60 mm-dish of PBS. The embryo head was dissected away to ensure that no retinal cells were included in MEF harverst. The heart and liver were removed and discarded. The remaining tissue was washed with 2 mL sterile 1X PBS and tissue was placed in a fresh 60mm dish. Tissue was thoroughly minced with scapels and 2 ml of 0.25% trypsin with EDTA was added. Tissue was incubated at 37 °C for 10 minutes and shook twice during this incubation. After incubation, embryo tissue was vigorously pipetted up and down until it was a single cell suspension (approximately 15 times). This 2 mL single cell suspension was transferred to a 15 mL conical tube with 8 mL DMEM media with 10% FBS and 1% Penicillin/Streptomycin and pipetted up and down 5 times. Samples were left for 5 minutes to ensure settling of non-dissociated tissue and then the supernatant was transferred to a T-75 flask, and labeled 'passage 1'. The next day the media was changed. Cells were split every 3-4 days with 0.25% trypsin with EDTA at a ratio of 1:4.

3.2 Cloning strategies

3.2.1 PCR amplification

Initial attempts at PCR cloning utilized primers with restriction enzyme recognition sites added to the 5′ and 3′ ends (Tables 2.3, 2.4) and the Easy-A high-fidelity PCR cloning kit (Agilent) with the manufacturers recommendations for reagent dilutions, cycle lengths, and temperatures. Modifications to the protocol included: addition of PCR enhancers including 1%, 5%, or 10% (DMSO) and/or 0.75 M Betaine (Sigma), supplementation of 1-2.5 mM MgCl for final concentrations ranging from 3-4.5 mM, addition of 5 amplification cycles, gradient PCR +/-4 °C in 1 °C steps, an increase in primer concentration from 100 ng/μL to 250 ng/μL, and allowing an additional 1 minute for extension.

3.2.2 Serial bacterial expression vector cloning

Genes of interest were restriction enzyme excised from commercial suppliers bacterial expression vectors, isolated using the Wizard gel-purification kit (Promega), and ligated into the pGEM (Promega), pLITMUS (New England Biosystems) or pCR 2.1-TOPO system (Invitrogen) using the Quick ligation kit (New England Biosystems). Upon transformation of 5μL of ligation product into the JM109 (Promega) bacterial strains per manufacturers instruction, cells were plated onto selective plates overnight. The following day, single colonies were picked, and cultures were grown for 16 hours at 37 °C. Plasmid DNA was subsequently purified using the QIAprep Spin Miniprep kit (Qiagen). Genes of interest were re-excised with restriction enzymes that were gene and viral vector compatible. These products, along with compatibly cut viral vector backbone, were gel-purified, and ligated together using the quick ligation kit (New England Biosystems) and Stbl3 cells (Invitrogen) were transformed. After plating overnight onto selective plates, single colonies were picked, cultures grown, DNA

purified, and restriction enzyme test cut, samples were sent for sequencing to the Indiana University DNA Sequencing Core Facility with custom sequencing primers.

3.2.3 Genes custom ordered from Integrated DNA Technologies

Lyophilized DNA was resuspended at 20 ng/μl in water, and 5μL was used to transform JM109 bacterial cells (Promega). Genes were excised from the provided plasmid using restriction enzyme sites that were engineered at the 5′ and 3′ ends. Compatibly cut gene products and lentiviral expression vector were separated on a 1% low melting point agarose gel, isolated using The Wizard gel purification kit (Promega), and ligated together using the quick ligation kit (New England Biosystems) and used to transform Stbl3 bacterial cells (Invitrogen). Transformed cells were grown overnight at 37 °C on LB/agar plates in the presence of carbenicillin. The next day, colonies were picked, and added to 2 mL of LB containing carbenicillin (50μg/ml). Cultures were grown overnight at 37 °C, and plasmid DNA was purified via the QIAprep Spin Miniprep kit (Qiagen). Plasmids were then restriction enzyme test cut, and samples were sent for sequencing to the Indiana University DNA Sequencing Core Facility with custom sequencing primers (Table 2.6, 2.7).

3.3 Cell culture

HEK293 or MEF cells were plated in T75 flasks or onto poly-ornithine/laminin-coated coverslips and allowed to grow overnight. Cell culture media consisted of DMEM supplemented with 10% fetal bovine serum, 1X sodium pyruvate, 1X MEM non-essential amino acids, and 1X penicillin/streptomycin. The following day, media was changed and cells were monitored until they reached confluency. At confluency, cells were passaged with 0.25% trypsin with EDTA and split at a ratio of 1:4.

3.4 Virus production

HEK293 cells were plated at $5x10^6$ cells per T75 flask. An excel spreadsheet (provided by Dr. Scott Witting of Indiana University School of Medicine) was used for molar calculations for all plasmids. At the end of the next day, transfer plasmid, packaging plasmids, $CaCl_2$ solution and water, were added to one 5 mL cryovial and labeled 'DNA Master mix'. 2X HBSS was added to another 5mL cryovial. 12 mL of fresh media was added to the HEK293 cells. A vortex in the BSC was set to 4.5, and the vial of HBSS was continually vortexed. Using a P1000, the DNA master mix was added dropwise, but fairly quickly, to the vortexing mix. A 1 minute timer was started immediately after adding all the DNA master mix and this solution was incubated in the hood for exactly 1 minute. As complex size increases rapidly over time, any deviation from this 1 minute incubation will decrease the efficiency of delivery to cells. The DNA/HBSS mix was added to the HEK293 cell flask. The flask was tilted and calcium phosphate/DNA master mix was added dropwise on top of the media. The solution was mixed well by tilting back and forth and turning the flask before moving onto the next set. Flasks were placed back in the incubator overnight. The following morning, media was removed and replaced with 12 mLs of HEK293 cell media. transfected cells were allowed to grow for another 35 hours and the supernatant was collected. Media was filtered through a .45μm filter while in the BSC and twelve 1mL aliquots were made. 1-500μL was used to infect MEF or HEK293 cells for initial titration experiments to empirically determine the concentrations necessary for each expression construct that yielded over 90% infection. For most expression constructs used in the experiments described, 50μL was sufficient, though MOIs remain undetermined at the time of writing this thesis.

3.5 Calcium phosphate transfection

For each 10 cm dish, DNA was dissolved with 450μL water, 50μL of 2.5M $CaCl_2$ and mixed. This solution was slowly added dropwise into 500μL of BES, mixed thoroughly

and incubated in the dark for 20 minutes. After, this solution was slowly dropped into the culture dish, and distributed evenly by gentle rocking of the plate. Cells were left to grow in the 37 °C/5% CO2 incubator for 12 hours. After 12 hours, media was replaced. Cells were grown for approximately another 48 hours before being fixed for the immunocytochemistry analysis described below.

3.6 Immunocytochemistry

Cells were fixed with 4% paraformaldehyde for ten minutes, washed three times with PBS, and permeabilized with 0.2% Triton X-100 in PBS for 10 minutes. Samples were blocked in PBS with 10% donkey serum for one hour to prevent non-specific antibody binding. Primary antibodies were subsequently applied in 5% donkey serum/0.1% Triton X-100 at recommended dilutions and incubated overnight at 4 °C. Following three washes with PBS, fluorophore-conjugated secondary antibodies and DAPI were applied at a 1:1000 dilution in 5% donkey serum/0.1% Triton X-100 for one hour at room temperature. After 3 washes in PBS, coverslips were mounted on slides and imaged on a Leica 5500 upright microscope with epifluorescence.

4 CONCLUSIONS, FUTURE EXPERIMENTS AND IMPLICATIONS

4.1 Conclusions

This thesis undertook the design, establishment and characterization of a system aimed to test the hypothesis that mouse embryonic fibroblasts, MEF cells, could be genetically, directly reprogrammed to rod photoreceptors. Importantly, provision of preliminary data in support of this hypothesis was also provided. The aims of this study were sevenfold: to determine candidate genes for direct reprogramming studies, cloning of candidate genes into appropriate vectors, adapting a lentiviral system for gene delivery, generating cells for use as a high-throughput screening system for analysis of virally-infected or transfected somatic cells, provision of proof-of-principle that the designed constructs led to induction of protein expression, collecting preliminary data demonstrating the neuralization of somatic cells induced by combinatorial delivery of known pro-neural genes and lastly, the induction of photoreceptor-like phenotypes in somatic cells.

Each of these aims was addressed successfully. First, 23 candidate genes were identified and all 23 were successfully cloned or synthesized and inserted into a modified lentiviral expression construct. The lentivirus production and delivery protocol was modified to enhance induction of protein expression in over 90% of somatic cells infected, as observed by using a GFP-reporter construct for optimization experiments. MEF cells were derived from a mouse model with a Rhodopsin-GFP fusion knock-in in place of native rhodopsin, enabling the high-throughput screening of the conversion of these MEF cells using GFP-expression. As *Rhodopsin* expression is specific to rod photoreceptors and one of the final stages of photoreceptor development, this

is a reliable indication of MEF conversion to rod photoreceptor phenotypes. Antibodies to the proteins produced by several constructs were commercially available, and allowed for the demonstration of induction of protein expression when these constructs were delivered singularly or in combination to MEF cells. Preliminary data demonstrated that the combinatorial delivery of constructs containing the pro-neural genes *Ascl1* and *Brn2* led to the recapitulation of published work [18] demonstrating expression of the neural-specific protein, βIII-tubulin. And lastly, utilizing the MEF cells derived from Rhodopsin-GFP-fusion knock-in mouse, delivery of *Olig2* alone was demonstrated to induce expression of GFP, indicating its ability to activate the rhodopsin promoter. While each of these aims was addressed, work remains to be done to completely characterize and test the systems established. Furthermore, while proof-of-principle was the aim and evidence was provided or exciting experiments aimed at optimizing the direct reprogramming of fibroblasts to rod photoreceptor have barely begun.

4.2 Future experiments continuing the project presented herein

The bulk of the work contained herein aimed to establish and confirm the tools necessary to test the hypothesis that mouse embryonic fibroblasts could be directly reprogrammed to rod photoreceptors. Previous studies described and the data presented establish the proof-of-principle that the direct reprogramming of photoreceptors is feasible using the constructs and systems established herein, although future experiments will be necessary to confirm and expand upon these results, as well as further optimize the system. To aid in these efforts, all 23 genes of interest have been sequence-confirmed in the modified pCSC-PGK-IGW lentiviral expression construct, and micro- to milligram quantities of each of these constructs have been restriction enzyme excision-confirmed and are aliquoted and logged in long-term storage. Furthermore, for many of these constructs, viral supernatants are also aliquoted and stored. These tools will serve as the basis for many subsequent experiments, as well

as used for further characterization studies. Preliminary data has demonstrated the utility of these constructs generally, but each one still needs careful, individual characterization. In order to achieve the highest efficiency of infection and presumably greatest degree of reprogramming, viral supernatant titers need to be defined, cellular toxicity determined, and multiplicities of infection (MOIs) established once a qPCR-based analysis system is optimized. While preliminary data demonstrates upregulation of protein expression for all constructs tested, either delivered via calcium phosphate transfection, or using the viral supernatants, this preliminary proof-of-principle will have to be demonstrated and characterized for all remaining constructs at both the mRNA and protein levels.

The most exciting data from this project is yet to come, as combinatorial delivery of the constructs is largely untested at this time. Combinatorial testing of candidate transcription factors allows for the establishment of the optimal set of genes for reprogramming, yielding the highest efficiency, while eliminating those that are unnecessary and perhaps even inhibitory. Several strategies have been outlined and developed to begin this process, utilizing the rod-specific, Rhodopsin-promoter-driven GFP-reporter MEF cells. The systems to be tested, roughly in this order, include: using a complete pool of all 23 factors, a system testing each factor in combination with the pro-neural factor *Ascl1*, a system delivering *Brn2, Ascl1, Myt1* (BAM)+ one other factor at a time, one testing the effects of *Ascl1*+ the top candidates identified from both +1 systems, and finally one testing the effects of removing each candidate one at a time from the combination of top candidates. Each of these strategies can be processed in a high-throughput screening scenario using the Rhodopsin-GFP MEF cells generated. This experimental design should result in defining a minimum pool for the direct reprogramming of MEF cells to a rod photoreceptor phenotype.

4.3 Implications of work resulting in directly reprogrammed rod photoreceptors

The definition of a minimum pool of transcription factors that can directly reprogram somatic cells to a rod photoreceptor fate has staggering implications. Firstly, it would be the first demonstration of direct reprogramming of somatic cells to a sensory neuronal cell type. It would also provide researchers a system in which to determine and characterize the earliest events in photoreceptor development, including necessary and sufficient gene expression levels, which despite great efforts, continue to evade characterization in the most well-designed systems and experiments to date. It may uncover genes currently not known to be involved in the process of photoreceptor development. Furthermore, it would provide a much faster strategy to derive and develop photoreceptors *in vitro*, which would be patient-specific, thereby rapidly increasing opportunities leading to the development of therapeutic agents. For example, photoreceptors reprogrammed from somatic fibroblasts could be utilized for cell replacement for degenerative disorders such as macular degeneration or retinitis pigmentosa. Additionally, when derived from patients with known genetic diseases of the retina, these cells could be used as a novel platform for disease modeling and pharmacological screening, particularly to study the pathology of a disease over time. Furthermore, this system will allow for the design of experiments to test cellular interactions of healthy and diseased cells, as well as interactions with cell types of neighboring and surrounding tissue in co-culture systems, as well as characterization of reprogrammed cells abilities to secrete survival factors and rescue diseased cells.

LIST OF REFERENCES

LIST OF REFERENCES

[1] Fawcett JW. Overcoming inhibition in the damaged spinal cord. J Neurotrauma. 2006;23:371-383.

[2] Yiu G, He Z. Glial inhibition of CNS axon regeneration. Nat Rev Neurosci. 2006;7:617-627.

[3] Bahr M, Bonhoeffer F. Perspectives on axonal regeneration in the mammalian CNS. Trends Neurosci. 1994;17:473-479.

[4] Horner PJ, Gage FH. Regeneration in the adult and aging brain. Arch Neurol. 2002;59:1717-1720.

[5] Horner PJ, Gage FH. Regenerating the damaged central nervous system. Nature. 2000;407:963-970.

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

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

[8] Gurdon JB, Elsdale TR, Fischberg M. Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. Nature. 1958;182:64-65.

[9] Briggs R, King TJ. Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. Proc Natl Acad Sci U S A. 1952;38:455-463.

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

[11] Noggle S, Fung HL, Gore A, et al. Human oocytes reprogram somatic cells to a pluripotent state. Nature. 2011;478:70-75.

[12] Davis RL, Weintraub H, Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell. 1987;51:987-1000.

[13] Zhou Q, Brown J, Kanarek A, et al. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. Nature. 2008;455:627-632.

[14] Hirami Y, Osakada F, Takahashi K, et al. Generation of retinal cells from mouse and human induced pluripotent stem cells. Neurosci Lett. 2009;458:126-131.

[15] Lamba DA, Karl MO, Ware CB, et al. Efficient generation of retinal progenitor cells from human embryonic stem cells. Proc Natl Acad Sci U S A. 2006;103:12769-12774.

[16] Hu BY, Zhang SC. Directed differentiation of neural-stem cells and subtype-specific neurons from hESCs. Methods Mol Biol. 2010;636:123-137.

[17] Li XJ, Hu BY, Jones SA, et al. Directed differentiation of ventral spinal progenitors and motor neurons from human embryonic stem cells by small molecules. Stem Cells. 2008;26:886-893.

[18] Vierbuchen T, Ostermeier A, Pang ZP, et al. Direct conversion of fibroblasts to functional neurons by defined factors. Nature. 2010;463:1035-1041.

[19] Meyer JS, Shearer RL, Capowski EE, et al. Modeling early retinal development with human embryonic and induced pluripotent stem cells. Proc Natl Acad Sci U S A. 2009;106:16698-16703.

[20] Osakada F, Ikeda H, Sasai Y, et al. Stepwise differentiation of pluripotent stem cells into retinal cells. Nat Protoc. 2009;4:811-824.

[21] Zhang XQ, Zhang SC. Differentiation of neural precursors and dopaminergic neurons from human embryonic stem cells. Methods Mol Biol. 2010;584:355-366.

[22] Son EY, Ichida JK, Wainger BJ, et al. Conversion of mouse and human fibroblasts into functional spinal motor neurons. Cell Stem Cell. 2011;9:205-218.

[23] Pang ZP, Yang N, Vierbuchen T, et al. Induction of human neuronal cells by defined transcription factors. Nature. 2011;476:220-223.

[24] Qiang L, Fujita R, Yamashita T, et al. Directed conversion of Alzheimer's disease patient skin fibroblasts into functional neurons. Cell. 2011;146:359-371.

[25] Caiazzo M, Dell'Anno MT, Dvoretskova E, et al. Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. Nature. 2011;476:224-227.

[26] Pfisterer U, Kirkeby A, Torper O, et al. Direct conversion of human fibroblasts to dopaminergic neurons. Proc Natl Acad Sci U S A. 2011;108:10343-10348.

[27] Kim J, Su SC, Wang H, et al. Functional integration of dopaminergic neurons directly converted from mouse fibroblasts. Cell Stem Cell. 2011;9:413-419.

[28] Di Giorgio FP, Carrasco MA, Siao MC, et al. Non-cell autonomous effect of glia on motor neurons in an embryonic stem cell-based ALS model. Nat Neurosci. 2007;10:608-614.

[29] Nagai M, Re DB, Nagata T, et al. Astrocytes expressing ALS-linked mutated SOD1 release factors selectively toxic to motor neurons. Nat Neurosci. 2007;10:615-622.

[30] Hafler BP, Surzenko N, Beier KT, et al. Transcription factor Olig2 defines subpopulations of retinal progenitor cells biased toward specific cell fates. Proc Natl Acad Sci U S A. 2012;109:7882-7887.

[31] Zhou JX, Huang S. Understanding gene circuits at cell-fate branch points for rational cell reprogramming. Trends Genet. 2011;27:55-62.

[32] Huang S. Reprogramming cell fates: reconciling rarity with robustness. Bioessays. 2009;31:546-560.

[33] Chan F, Bradley A, Wensel TG, et al. Knock-in human rhodopsin-GFP fusions as mouse models for human disease and targets for gene therapy. Proc Natl Acad Sci U S A. 2004;101:9109-9114

[34] Rodeick RW. The First Steps in Seeing. Sinauer Associates Publishers, Sunderland, Massachusetts; 1998.

[35] Dowling JE. The Retina: An Approachable Part of the Brain. Belknap Press, Harvard Univ. Press; 1987.

[36] Swaroop A, Kim D, Forrest D. Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. Nat Rev Neurosci. 2010;11:563-576.

[37] Larry Squire DB, Floyd Bloom, Sascha du Lac, Anirvan Ghosh, Nicholas C. Spitzer. Fundamental Neuroscience: Academic Press; 2008

[38] Clarke G, Collins RA, Leavitt BR, et al. A one-hit model of cell death in inherited neuronal degenerations. Nature. 2000;406:195-199.

[39] Luo DG, Xue T, Yau KW. How vision begins: an odyssey. Proc Natl Acad Sci U S A. 2008;105:9855-9862.

[40] Carter-Dawson LD, LaVail MM. Rods and cones in the mouse retina. I. Structural analysis using light and electron microscopy. J Comp Neurol. 1979;188:245-262.

[41] Roorda A, Williams DR. The arrangement of the three cone classes in the living human eye. Nature. 1999;397:520-522.

[42] Carter-Dawson LD, LaVail MM. Rods and cones in the mouse retina. II. Autoradiographic analysis of cell generation using tritiated thymidine. J Comp Neurol. 1979;188:263-272.

[43] MacLaren RE, Pearson RA, MacNeil A, et al. Retinal repair by transplantation of photoreceptor precursors. Nature. 2006;444:203-207.

[44] Chacko DM, Rogers JA, Turner JE, et al. Survival and differentiation of cultured retinal progenitors transplanted in the subretinal space of the rat. Biochem Biophys Res Commun. 2000;268:842-846.

[45] Sakaguchi DS, Van Hoffelen SJ, Theusch E, et al. Transplantation of neural progenitor cells into the developing retina of the Brazilian opossum: an in vivo system for studying stem/progenitor cell plasticity. Dev Neurosci. 2004;26:336-345.

[46] Van Hoffelen SJ, Young MJ, Shatos MA, et al. Incorporation of murine brain progenitor cells into the developing mammalian retina. Invest Ophthalmol Vis Sci. 2003;44:426-434.

[47] Young MJ, Ray J, Whiteley SJ, et al. Neuronal differentiation and morphological integration of hippocampal progenitor cells transplanted to the retina of immature and mature dystrophic rats. Mol Cell Neurosci. 2000;16:197-205.

[48] Furukawa T, Morrow EM, Li T, et al. Retinopathy and attenuated circadian entrainment in Crx-deficient mice. Nat Genet. 1999;23:466-470.

[49] Jia L, Oh EC, Ng L, et al. Retinoid-related orphan nuclear receptor RORbeta is an early-acting factor in rod photoreceptor development. Proc Natl Acad Sci U S A. 2009;106:17534-17539.

[50] Koike C, Nishida A, Ueno S, et al. Functional roles of Otx2 transcription factor in postnatal mouse retinal development. Mol Cell Biol. 2007;27:8318-8329.

[51] Mears AJ, Kondo M, Swain PK, et al. Nrl is required for rod photoreceptor development. Nat Genet. 2001;29:447-452.

[52] Nishida A, Furukawa A, Koike C, et al. Otx2 homeobox gene controls retinal photoreceptor cell fate and pineal gland development. Nat Neurosci. 2003;6:1255-1263.

[53] Lamba DA, McUsic A, Hirata RK, et al. Generation, purification and transplantation of photoreceptors derived from human induced pluripotent stem cells. PLoS One. 2010;5:e8763.

[54] Meyer JS, Howden SE, Wallace KA, et al. Optic vesicle-like structures derived from human pluripotent stem cells facilitate a customized approach to retinal disease treatment. Stem Cells. 2011;29:1206-1218.

[55] Carr AJ, Vugler AA, Yu L, et al. The expression of retinal cell markers in human retinal pigment epithelial cells and their augmentation by the synthetic retinoid fenretinide. Mol Vis. 2011;17:1701-1715.

[56] Mellough CB, Sernagor E, Moreno-Gimeno I, et al. Efficient Stage Specific Differentiation of Human Pluripotent Stem Cells Towards Retinal Photoreceptor Cells. Stem Cells. 2012.

[57] Akimoto M, Cheng H, Zhu D, et al. Targeting of GFP to newborn rods by Nrl promoter and temporal expression profiling of flow-sorted photoreceptors. Proc Natl Acad Sci U S A. 2006;103:3890-3895.

[58] Chen S, Wang QL, Nie Z, et al. Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. Neuron. 1997;19:1017-1030.

[59] Lerner LE, Gribanova YE, Ji M, et al. Nrl and Sp nuclear proteins mediate transcription of rod-specific cGMP-phosphodiesterase beta-subunit gene: involvement of multiple response elements. J Biol Chem. 2001;276:34999-35007.

[60] Mitton KP, Swain PK, Chen S, et al. The leucine zipper of NRL interacts with the CRX homeodomain. A possible mechanism of transcriptional synergy in rhodopsin regulation. J Biol Chem. 2000;275:29794-29799.

[61] Rehemtulla A, Warwar R, Kumar R, et al. The basic motif-leucine zipper transcription factor Nrl can positively regulate rhodopsin gene expression. Proc Natl Acad Sci U S A. 1996;93:191-195.

[62] Yoshida S, Mears AJ, Friedman JS, et al. Expression profiling of the developing and mature Nrl-/- mouse retina: identification of retinal disease candidates and transcriptional regulatory targets of Nrl. Hum Mol Genet. 2004;13:1487-1503.

[63] Chen J, Rattner A, Nathans J. The rod photoreceptor-specific nuclear receptor Nr2e3 represses transcription of multiple cone-specific genes. J Neurosci. 2005;25:118-129.

[64] Cheng H, Aleman TS, Cideciyan AV, et al. In vivo function of the orphan nuclear receptor NR2E3 in establishing photoreceptor identity during mammalian retinal development. Hum Mol Genet. 2006;15:2588-2602.

[65] Cheng H, Khanna H, Oh EC, et al. Photoreceptor-specific nuclear receptor NR2E3 functions as a transcriptional activator in rod photoreceptors. Hum Mol Genet. 2004;13:1563-1575.

[66] Swain PK, Hicks D, Mears AJ, et al. Multiple phosphorylated isoforms of NRL are expressed in rod photoreceptors. J Biol Chem. 2001;276:36824-36830.

[67] Coolen M, Sii-Felice K, Bronchain O, et al. Phylogenomic analysis and expression patterns of large Maf genes in Xenopus tropicalis provide new insights into the functional evolution of the gene family in osteichthyans. Dev Genes Evol. 2005;215:327-339.

[68] Daniele LL, Lillo C, Lyubarsky AL, et al. Cone-like morphological, molecular, and electrophysiological features of the photoreceptors of the Nrl knockout mouse. Invest Ophthalmol Vis Sci. 2005;46:2156-2167.

[69] Oh EC, Cheng H, Hao H, et al. Rod differentiation factor NRL activates the expression of nuclear receptor NR2E3 to suppress the development of cone photoreceptors. Brain Res. 2008;1236:16-29.

[70] Peng GH, Ahmad O, Ahmad F, et al. The photoreceptor-specific nuclear receptor Nr2e3 interacts with Crx and exerts opposing effects on the transcription of rod versus cone genes. Hum Mol Genet. 2005;14:747-764.

[71] Young RW. Cell differentiation in the retina of the mouse. Anat Rec. 1985;212:199-205.

[72] Teschendorf C, Warrington KH, Jr., Siemann DW, et al. Comparison of the EF-1 alpha and the CMV promoter for engineering stable tumor cell lines using recombinant adeno-associated virus. Anticancer Res. 2002;22:3325-3330.

[73] Xia X, Zhang Y, Zieth CR, et al. Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. Stem Cells Dev. 2007;16:167-176.

[74] Yang N, Ng YH, Pang ZP, et al. Induced neuronal cells: how to make and define a neuron. Cell Stem Cell. 2011;9:517-525.

[75] Greber B. Isolation and handling of primary mouse embryonic fibroblasts (MEFs), 2011.