Pairing Your SOX:
The Role of HMG-Domain Transcription Factors in Peripheral Nerve Myelination

by

Chetna Gopinath

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cellular and Molecular Biology)
in the University of Michigan
2018

Doctoral Committee:

Associate Professor Anthony Antonellis, Chair
Associate Professor Scott Barolo
Professor Roman Giger
Assistant Professor Shigeki Iwase
Professor Patricia Wittkopp

chetnag@umich.edu
2018
ORCID ID: 0000-0003-1117-5685

Dedication

To my pillars of strength Mom, Dad, and Anuj for your never ending love and support

Acknowledgements

This thesis would not have been possible without the support of a lot of people. Firstly, I would like to thank my mentor Tony Antonellis for all his support and motivation through the years. I joined Tony's laboratory as a Research Associate after my Master's degree. During this time I thought to myself as to how lucky the graduate students in his lab were and I could only wish that I could pursue my Ph.D. thesis in his laboratory. Tony is a fantastic advisor and I want to thank him for providing me with plenty of opportunities to grow and mature as a scientist. He has taught me to think about the big picture research questions, how to design experiments with proper controls, and how to carefully interpret the data. My only hope is to be at least half the amazing scientist he is. My writing and presentation skills have tremendously improved under his guidance. Tony has created a fun lab environment filled with smart and hardworking people and has always been very respectful of every lab member and puts people in his lab above everything. I want to thank him and his family for inviting Anuj and me to celebrate the holidays. Over the years, he has played a big part in me being the scientist I am today and I can't thank him enough for all his support.

I would like to thank my thesis committee- Drs. Barolo, Giger, Iwase and Wittkopp. Thank you so much for your feedback and useful suggestions, which have helped me through the course of

my PhD. I would like to thank Dr. Choya Yoon, a postdoctoral researcher in the Giger lab for her help with isolating primary Schwann cells. I would also like to thank Drs. Meisler and Kitzman and their laboratories for critical feedback and suggestions during our weekly lab meetings. I can't thank Dr. Isaac Jia enough for all his help with the bioinformatics analysis. I want to thank Dr. Amanda Pendleton for all her advice and help during my post-doc interviews.

I want to thank laboratories of Drs. Moran, Kalantry, and Kidd for always willing to share research supplies and equipment. I would like to thank all the core facilities at the University of Michigan. I want to thank our collaborators Drs. Svaren and Crawford for their help and willingness collaborate and share data.

I want to thank past and present members of the Antonellis lab, they are all fantastic scientists and very kind-hearted people. I want to thank Liz for always being helpful and for being instrumental in my thesis work. I want to thank Rebecca for numerous random "pop-culture" discussions. Molly, I will miss our early morning conversations and I can only hope your organizational skills have rubbed off on me. I would like to thank Natasha for her willingness to help with any task. I will miss our morning coffee walk! I want to thank past lab members- Stephanie, Bill, and Laurie for their support.

A special thank you to Dr. Fuller and the Cellular and Molecular Biology (CMB) program current and previous staff- Cathy Mitchell, Margarita Biekares, Patricia Ocelnik and Jessica Kijek for working so hard and taking good care of all the students. CMB is a great program and I have been fortunate to be a part of it. I would like to thank the University of Michigan and the

International Student Center for supporting international students and always making us feel welcome. I want to thank Dr. Laura Monschau of Counseling and Psychological Services (CAPS) for helping me tackle my imposter syndrome.

Grad school can be challenging at times and I would like to thank all my friends and family. Christina, thank you for always being there for me and supporting me over the years. I really hope we continue to stay in touch. I want to thank Bhushan, Uday, Clement, Bhavya, Harish, Akshat, Pooja, Prathyusha, Janani, Nandini, Sarika, Ganesh, Jayashree, and Kate. Thank you so much for being our support system in the United States and you guys are the reason we call Michigan our home. From numerous game nights to road trips we have taken, I will cherish every moment and I hope we continue to make new memories. I want to thank all my friends who are elsewhere in the US and in other parts of the world.

I am extremely fortunate to have a family that has always supported me and respected my career aspirations. Mom and dad thank you for teaching us the importance of education and sacrificing everything just to ensure that Chitra and I got the best education. You always put our needs and wants before your own and always encouraged us to pursue opportunities you never had. I have been able to achieve all my dreams because of you. I want to thank my sister Chitra and my brother-in-law Vivek for helping me settle down in the US and for always looking out for me. I want to thank my grandma, a truly exceptional human being who taught me to love and respect every being and I know she would have been so proud of me. I am thankful to my father-in-law, mother-in-law, Pati, Karthik, and Sonali for welcoming me into their family and for their support and motivation.

Before starting my Ph.D., I married my best friend, Anuj. Thank you for being my 24/7 tech support and for sitting through all my presentation rehearsals. I can't thank you enough for being patient during all my meltdowns and for keeping me happy everyday. Thank you for being supportive of my career aspirations and for agreeing to move to San Francisco. I couldn't have done this without you!

Table of Contents

Chapter 2
Stringent Comparative Sequence Analysis Reveals SOX10 as a Putative Inhibitor of Glial Cell Differentiation

Chapter 3
Exploring the Role of SOX6 in Peripheral Nerve Myelination

List of Figures

List of Tables

Abstract

Compact myelin is an innovative acquisition in jawed vertebrates that is formed by the wrapping of glial cell membranes around neuronal axons in the central and peripheral nervous systems. The major functions of compact myelin include providing trophic support to axons and increasing the propagation of nerve impulses, which has facilitated the escape and predatory behaviors of vertebrates. Schwann cells are neural crest derived glia that form compact myelin in the peripheral nervous system (PNS). The SRY-related HMG box 10 (SOX10) transcription factor is essential for all stages of Schwann cell development. Specifically, SOX10 activates the expression of key myelin genes in the PNS and has therefore been reported as a pro-myelination transcription factor. Previously identified SOX10 target genes have been shown to be critical for Schwann cell function. Thus, the identification of additional genes regulated by SOX10 will improve our understanding of myelination in the PNS. We developed a stringent, computational method for genome-wide identification of SOX10 response elements. Experimental validation of a set of predicted SOX10 response elements revealed *SOX5*, *SOX6*, and *NFIB* as novel SOX10 target genes. To further explore the utility of our computational data we compared our predictions to published SOX10 ChIP-seq data from rat sciatic nerve and our own DNase-seq data generated from cultured Schwann cells. This analysis—along with subsequent functional studies—revealed SOX10 response elements that map to *HES1, MYCN, ID2,* and *ID4*. Remarkably, *SOX5*, *SOX6*, *HES1*, *MYCN*, *ID2*, and *ID4* all encode proteins that inhibit

myelination. Thus, our computationally anchored strategy revealed a putative novel function for SOX10 in Schwann cells, which suggests a model where SOX10 activates the expression of genes that inhibit myelination during non-myelinating stages of Schwann cell development. We then deeply characterized the SOX10 response element at *SOX6*, which revealed that this element resides at a previously unreported alternative promoter that directs the expression of a specific mRNA isoform. SOX6 was previously reported to inhibit glial cell differentiation in the central nervous system; however, the role of SOX6 in Schwann cells has not yet been characterized. To explore the role of SOX6 in Schwann cells, we set out to identify SOX6 target loci via overexpression of SOX6 in culture Schwann cells followed by RNA-seq analysis. Gene ontology analysis of up-regulated genes revealed a putative role for SOX6 in Schwann cell proliferation and regeneration following nerve injury, which should be explored using further computational and functional studies. Importantly, the computational and functional datasets we present here will be valuable for the study of transcriptional regulation, SOX protein function, and glial cell biology.

Chapter 1

An Overview of Schwann Cell Biology

1.1 Introduction

Compact myelin is a lipid-rich layer formed by Schwann cells and oligodendrocytes along the axons of both the peripheral and central nervous systems respectively. The myelin sheath in jawed vertebrate organisms allows rapid conduction of action potentials along the nerves, which has contributed to fast responses to stimuli and offered the advantages of both predation and escape behaviors. The SRY-related HMG box 10 (SOX10) transcription factor is essential for all stages of peripheral nerve myelination. Indeed, mutations in *SOX10* and in SOX10 target genes cause several forms of demyelinating peripheral neuropathy. The main focus of my dissertation is to elucidate the transcriptional regulatory pathways important for peripheral nerve myelination by identifying novel SOX10 target loci. In this chapter, I provide an overview of how Schwann cells form myelin in the peripheral nervous system. Next, I present the evolution of myelin and discuss key differences between vertebrate myelin and invertebrate glial cell membranes. Finally, I discuss the transcriptional mechanisms important for vertebrate myelination and the role of SOX10 in this process.

1.2 An Introduction to the nervous systems

A typical vertebrate nerve consists of two cell types: neurons and glia. A neuron consists of a cell body, several branched projections called dendrites, and an axon (Fig. 1.1). Glial cells: (i) form the myelin sheath; (ii) provide trophic support to axons; and (iii) play a crucial role in nerve repair following injury. Schwann cells and oligodendrocytes are the glial cells that form compact myelin in the peripheral nervous system (PNS) and central nervous system (CNS), respectively (del Rio, 1921; Geren, 1954; Pérez-Cerdá et al., 2015). These cells form myelin in axon segments, thus generating periodic gaps called the nodes of Ranvier. The nodes of Ranvier contain high concentrations of sodium ion channels, which facilitate propagation of action potentials along the nerve; a process called saltatory (derived from the Latin word *saltare*, which means 'to jump') conduction (D'Este et al., 2017; Duflocq et al., 2008).

Compact myelin is specific to jawed vertebrates and is formed by the wrapping of glial cell membranes around the axons. Myelin was first observed in the eighteenth century by Antonie van Leeuwenhoek and his reports were based on observations using a peripheral nerve isolated from a cow (Rosenbluth, 1999; Tomassy et al., 2016). Several years later, in 1838, as part of his thesis, Robert Remak observed rabbit peripheral nerves during different stages of development. He noted that some nerves were surrounded by a sheath but that other nerves lacked this structure (Boullerne, 2016; Sakuta, 2010). He called the latter nerve bundles 'organic fibers', which are now referred to as Remak bundles. Remak bundles contain multiple axons surrounded by Schwann cells that are not myelinated. In the same year, Matthias Schleiden and Theodor Schwann were working on the 'cell theory', which proposed that all living things were composed of cells.

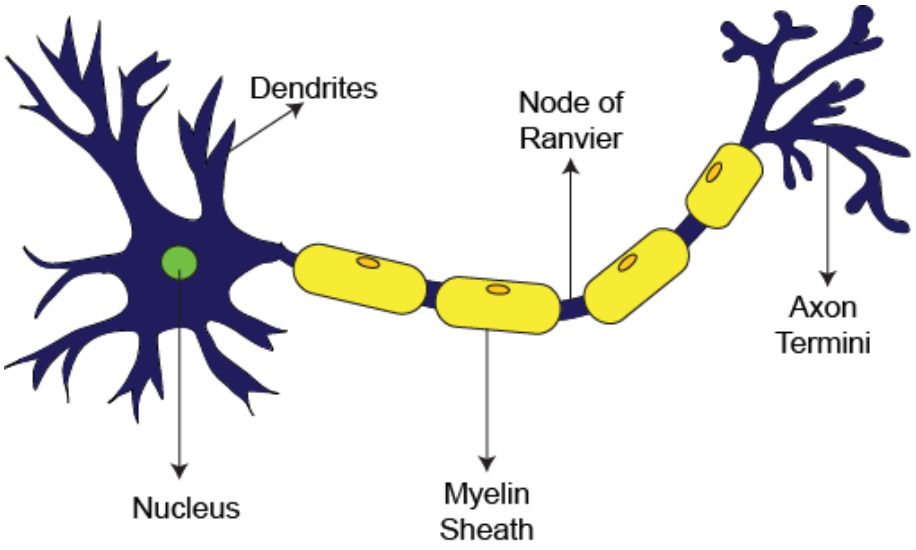


Figure 1.1 Structure of a neuron. The cell body and dendrites are depicted in blue and Schwann cells that form the myelin sheath are shown in yellow. The nuclei of the neuron and Schwann cells are shown in green and orange, respectively.

Schwann studied various animal tissues including peripheral nerves and noticed that the nerve bundles were composed of two types of fibers: white fibers and organic (or Remak) fibers. Schwann noted that the membrane of the white fibers was granulated and that the nuclei remained on the nerve fiber during various developmental stages. He proposed that the cells formed a continuous layer that came together to form the nerve fiber such that the nuclei always remained on the nerve fiber; these cells were later named Schwann cells (Boullerne, 2016; del Rio, 1921; Geren, 1954; Pérez-Cerdá et al., 2015). In 1854, the German pathologist Rudolf Virchow coined the term 'myelin', which is derived from the Greek word *myelos* (marrow) analogous to bone marrow; in the following years, the fatty substance surrounding the axon was referred to as myelin (Boullerne, 2016; Duflocq et al., 2008; Virchow, 1854).

1.3 Peripheral nerve myelination

Schwann cells are derived from neural crest cells (NCCs) (D'Este et al., 2017; Douarin et al., 1991), which are a transient multipotent population of cells that arise from the border as the neural plate folds to form the neural tube. NCCs separate themselves from the neural tube by a process called delamination (LaBonne and Bronner-Fraser, 1999; Rosenbluth, 1999; Tomassy et al., 2016) and undergo an epithelial to mesenchymal transition (EMT), a process by which the cells lose their epithelial properties to become mesenchymal cells. Following EMT, NCCs migrate to different locations within the embryo and—depending on the local environment—become restricted to generate distinct cell types. NCCs give rise to myriad cell types including melanocytes, Schwann cells, enteric nervous system neurons, and craniofacial cartilage (Anderson, 1993; Boullerne, 2016; Sakuta, 2010; Sauka-Spengler and Bronner-Fraser, 2008). Peripheral nerve myelination involves two developmental stages: (i) the transition of NCCs into

Schwann cell precursors (SCP); and (ii) the differentiation of SCPs into myelinating or non-myelinating Schwann cells (Jessen et al., 1994). SCP migrate along axonal tracts in the developing peripheral nerve (Zorick and Lemke, 1996). Communication between axons and Schwann cells is essential for myelination, axonal maintenance, and repair. Schwann cell development and myelination depend on mitogenic signals from the axons of the peripheral nerves. Conversion of SCP to immature Schwann cells entails gene expression changes, cell survival, and response to mitogens. Neuregulin-1 (NRG1) is a signaling molecule produced by axons that is essential for SCP survival in embryonic nerves; loss of *Nrg1* in mice is embryonic lethal and the embryos die by E11.5. Analysis of these mutant embryos revealed significantly reduced number of Schwann cell precursors suggesting that NRG1 is important for early stages of Schwann cell development (Meyer and Birchmeier, 1995). *NRG1* is one of the longest genes ~1.4MB long and undergoes alternative splicing to produce at least 15 different isoforms (Falls, 2003). Based on biological properties they are grouped into different types I-VI (Steinthorsdottir et al., 2004). Type III, in particular, is critical for the survival and proliferation of the SCP (Taveggia et al., 2005; Wolpowitz et al., 2000). NRG1 type III binds to receptor tyrosine kinases (Erbb2/3) that are expressed in Schwann cells (Levi et al., 1995) and this interaction elicits a cascade of regulatory pathways important for Schwann cell development (Britsch et al., 1998). ERBB2/3 form heterodimers and ERBB2 contains the catalytically active kinase domain and ERBB3 contains the NRG1 binding domain (Limpert and Carter, 2010). Early stages of Schwann cell development rely on NRG1/ErbB signaling and loss of *Nrg1, ErbB2,* or *ErbB3* in mice leads to absence of SCP (Meyer and Birchmeier, 1995; Riethmacher and Sonnenberg-Riethmacher, 1997; Woldeyesus et al., 1999). The NRG1/ErbB signaling pathway is critical for SCP survival and later for Schwann cell differentiation (Britsch et al., 1998; Leimeroth et al.,

2002a). Immature Schwann cells that associate with single large axons (>1 μm diameter) form the myelin sheath and those that associate with several small caliber axons (<1μm diameter) form the non-myelinating Remak bundles (Sherman and Brophy, 2005). Schwann cells possess remarkable plasticity and can dedifferentiate to an immature state resulting in demyelination. Demyelination of peripheral axons can occur in two ways: first, myelinating Schwann cells can lose axonal contact and dedifferentiate into the immature state following nerve injury to facilitate the recovery process; second, in patients with inherited demyelinating peripheral neuropathies the dedifferentiation pathway can be activated inappropriately due to the altered function of mutated genes. Once immature Schwann cells have associated with single axons, they exit the cell cycle to become pro-myelinating cells and initiate the myelination program (Fig. 1.2) (Jessen and Mirsky, 2005)

1.4 The developmental timeline of vertebrate myelination

Myelination is a dynamic process and occurs in a conserved temporal manner where axons in the PNS are myelinated before the CNS axons (Cravioto, 1965; Laurence, 2017). In 1965, Humberto Cravioto observed myelination in humans by isolating the sciatic nerves from 12-,14-,16-, and 22-weeks old fetuses. He made the following observations: (i) migrating Schwann cells were observed at 12-weeks; (ii) Schwann cells were mostly in the immature state at 14-weeks with the presence of occasional myelinated axons; (iii) myelinated axons were more frequent by 16-weeks; and (iv) a large number of axons were myelinated by 22-weeks (Cravioto, 1965; Jessen and Mirsky, 2005). In contrast, onset of PNS myelination is a postnatal event in rodents. Schwann cell migration occurs around embryonic (E) day 12-13 mice nerves (E14-15 in rat

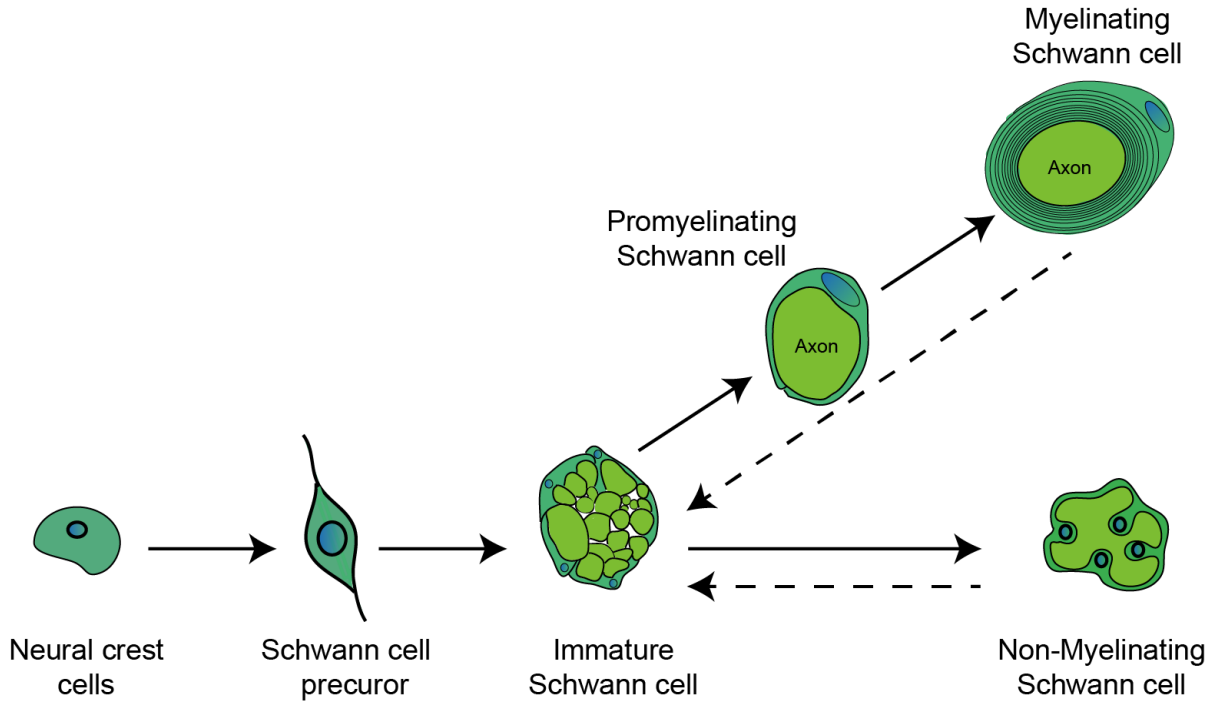


Figure 1.2 Stages of Schwann cell development. Neural crest cells give rise to Schwann cell precursors, which migrate along axons and differentiate to form immature Schwann cells. Immature Schwann cells have two fates: (i) those that are associated with single axonal segments become promyelinating Schwann cells and eventually form the myelin sheath; and (ii) those that are associated with multiple axons form non-myelinating Schwann cells. Dashed arrows indicate that these cells types can revert back to the immature stage. Adapted from Jessen and Mirsky, 2005.

nerves), immature Schwann cells are observed in E13-E15 mice nerves (E15-17 in rat nerves), and myelination begins after birth. Gestation time is about 19-21 days, 21-23 days for mice and rats, respectively (Jessen and Mirsky, 2005; Zalc et al., 2008).

1.5 The evolution of myelination

The myelin sheath has three major functions: (i) increase the conduction velocities along the axons; (ii) maintain axon integrity; and (iii) provide trophic factors to axons. The myelin sheath increases nerve conduction velocities by several magnitudes compared to an unmyelinated nerve and this enables a faster response to stimulus. The conduction velocity of an unmyelinated axon is about 0.5-10 meter per second (m/s) and that for a myelinated axon can be as high as 150 m/s. From an evolutionary standpoint, the two ways to improve conduction velocities are to either increase the axonal diameter or insulate the axons with myelin. The propagation speed of an action potential is directly proportional to the axon diameter in unmyelinated axons and this mechanism works well for invertebrates considering their smaller body sizes (Schweigreiter et al., 2006; Zalc et al., 2008). However, most vertebrates have larger bodies and myelin offers a more efficient way of improving the conduction velocity compared to increasing axonal diameter (Schweigreiter et al., 2006; Zalc et al., 2008). Thus, ecological conditions, the need for rapid escape, and large body sizes may have selected for the emergence of myelin in jawed vertebrates (Schweigreiter et al., 2006; Zalc et al., 2008).

Myelin was first observed in placoderms about 425 million years ago (Zalc et al., 2008). By comparing the length and the diameter of oculomotor nerves from fossilized placoderms skulls and their similarly-sized immediate ancestors (osteostraci), it was noted that the nerve in

placoderms was ten times longer compared to osteostraci, with the diameter in both species remaining constant. It was thus proposed that the only plausible way for placoderms to increase the length of the nerve while maintaining a constant diameter was a layer of myelin along the placoderm axons. Placoderms were active predators with hinged jaws. Since the structures that form the jaw are derived from the neural crest cells, it was proposed that the appearance of Schwann cells—also a neural crest derived cell type—may be coupled to the evolution of PNS myelin to facilitate predatory and escape behaviors (Zalc, 2016; Zalc et al., 2008). Myelination has not been observed in invertebrates, however, annelids and anthropods possess myelin-like layers surrounding their axons (Bullock et al., 1984; Davis et al., 1999; Pereyra and Roots, 1988; Roots et al., 1991) (Fig. 1.3).

All vertebrate species except agnathostomes (the jawless hagfish and lamprey) have myelinated axons (Bullock et al., 1984; Smith et al., 2013). Although lampreys lack myelin sheaths they do have glial cells and, interestingly, sequencing the lamprey genome revealed the presence of myelin-associated proteins (Smith et al., 2013). Peripheral myelin protein (PMP22), myelin protein zero (MPZ), and myelin proteolipid protein (PLP) are components important for the formation of myelin in the peripheral and central nervous systems and these genes were identified in the lamprey genome. Combined with the high conservation of the protein-coding sequences of these loci between lamprey and jawed vertebrates, these data suggest that the evolution of regulatory pathways (*e.g.*, transcription factors and cis-acting regulatory sequences) in gnathostomes led to their recruitment for the formation of myelin (Bullock et al., 1984; Smith et al., 2013).

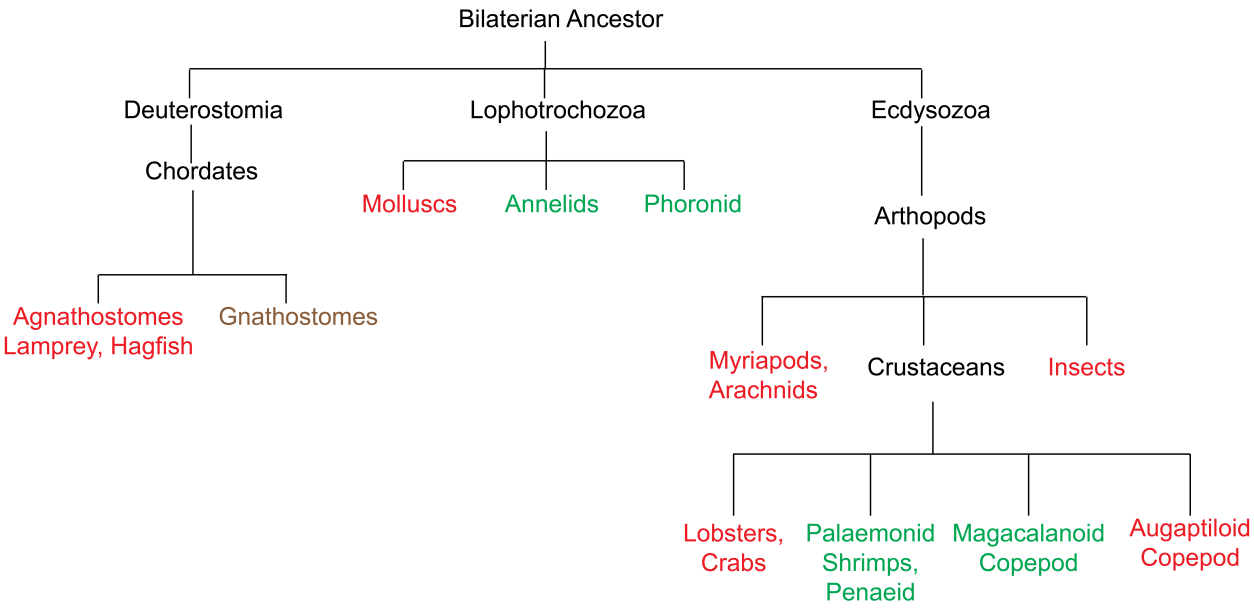


Figure 1.3 Myelin Evolution. A simplified phylogenetic tree showing the presence of myelin across the animal kingdom. Species labeled in green contain myelin-like layers around axons, those shown in red do not contain myelin and Gnathostomes have compact myelin. Figure adapted from Knowles 2017

1.6 A Comparison of invertebrate and vertebrate myelin

The axons in all chordates with the exception of jawless fish are surrounded by a compact, multi-layered myelin sheath (Bullock et al., 1984; Davis et al., 1999; Pereyra and Roots, 1988; Roots et al., 1991). In contrast, invertebrate axons associate with a loosely wound myelin-like ensheathment (Davis et al., 1999; Friede and Bischhausen, 1982; Michailov et al., 2004; Pereyra and Roots, 1988; Roots et al., 1991). The key differences of invertebrate and vertebrate myelin with respect to morphology, conduction velocity, and lipid and protein content are discussed below.

1.6.1.Morphology

Glial cells in vertebrates form compact myelin and the morphology of myelin is conserved across species. The cell membrane of Schwann cells and oligodendrocytes are wrapped around axons to form compact myelin in the peripheral and central nervous systems respectively. The thickness of the myelin sheath depends on the axonal diameter (Friede and Bischhausen, 1982; Michailov et al., 2004) and this process is controlled by Neuregulin-1 signaling in the peripheral nerve (Michailov et al., 2004; Schweigreiter et al., 2006). The myelin-like layer in invertebrates has variable morphology (Roots and Lane, 1983; Schreiner et al., 2007; Schweigreiter et al., 2006). In the earthworm *Lumbricus terrestris,* the glial cells wrap around giant axons to form a compact structure similar to what is observed in vertebrates (Cardone and Roots, 1991; Pusch et al., 1998; Roots and Lane, 1983; Schreiner et al., 2007); where as in the crayfish *Procambrus clarkii* it is loosely wound and number of layers can vary (Cardone and Roots, 1991; Heuser and Doggenweiler, 1966; Pusch et al., 1998).

Another striking difference between vertebrate and invertebrate myelin is the position of the glial cell nuclei. The location of the nuclei can vary in invertebrates and can be found in any location within the myelin-like layer (Bunge et al., 1989; Heuser and Doggenweiler, 1966; Schweigreiter et al., 2006). In contrast, vertebrate glial cell nuclei consistently reside within the outermost layer of compact myelin (Bunge et al., 1989; Günther, 1976; Schweigreiter et al., 2006).

1.6.2 Conduction velocity

The conduction velocity of action potentials along the axons is higher in vertebrates compared to invertebrates. For example, the conduction velocity of the myelinated giant axon (diameter ~90μm) of the earthworm *Lumbricus terrestris* is about 30 m/s (Günther, 1976; Lee et al., 1999; Schweigreiter et al., 2006) and that of a rat axon (diameter ~4.5μm) is about 59 m/s (Garbay et al., 2000; Lee et al., 1999; Schweigreiter et al., 2006). Faster conduction velocities enable rapid response to changes in the environment and offer a selective advantage.

1.6.3 Lipid and protein content

A key feature that distinguishes myelin from other cell membranes is a high lipid-to-protein ratio (Bürgisser et al., 1986; Garbay et al., 2000; Waehneldt et al., 1986). Myelin in vertebrates is mainly composed of phopholipids, cholesterol, galactocerebroside, sulphatide, and small amounts of protein (Bürgisser et al., 1986; Jeserich et al., 1990; Roots, 2008; Waehneldt et al., 1986). Myelin-like layers in annelids and crustaceans are also composed of phospholipids and cholesterol; however, this layer in these species lack galactocerebroside and sulphatide (Roots, 2008; Ryu et al., 2007). Crustaceans have a higher lipid to protein ratio (15:1) compared to vertebrate and annelids (2:1) (Laurence, 2017; Roots, 2008; Ryu et al., 2007).

Both compact myelin in vertebrates and myelin-like structures in invertebrates contain low molecular weight proteins. The protein components of the invertebrate myelin-like sheath is unclear and have not been fully catalogued (Cardone and Roots, 1996; Laurence, 2017; Monuki et al., 1993; Ryu et al., 2007). Immunoblot analysis of the ventral nerve cord of the earthworm *Lumbricus terrestris* suggested a simple protein expression pattern with 80, 42, and 30 kDa sized proteins (Campagnoni, 1988; Cardone and Roots, 1996; Saavedra et al., 1989).

Protein composition within the myelin sheath differs across different vertebrate species and between the CNS and the PNS. Myelin basic protein (MBP) is one of the main components of both CNS and PNS myelin in vertebrate species. MBP is a cell adhesion molecule which facilitates myelin compaction (Campagnoni, 1988; Givogri et al., 2001; Nawaz et al., 2013; Saavedra et al., 1989; Zorick et al., 1999). The *GOLLI* gene contains 10 exons in humans and undergoes alternative splicing to produce different *MBP* isoforms (Givogri et al., 2001; Roach et al., 1985). *Mbp* null mutant mice, called shiverer, arose spontaneously and display severe CNS hypomyelination. Interestingly, the PNS was only moderately affected (Greenfield et al., 1973; Roach et al., 1985; Topilko et al., 1994) likely due to the fact that PNS myelin contains about half the amount of MBP detected in CNS myelin. The most abundant protein in the PNS is Myelin Protein Zero (MPZ or P_0), a transmembrane glycoprotein (Greenfield et al., 1973; Lemke and Axel, 1985; Topilko et al., 1994). The cytoplasmic domains of MBP and MPZ share sequence homologies (Bai et al., 2011; Jones et al., 2011; Lemke and Axel, 1985; Saavedra et al., 1989; Tai and Smith, 1984; Waehneldt and Jeserich, 1984) so, the absence of *Mbp* may be compensated by *Mpz* in the peripheral nerves of the shiverer mutant mice.

MPZ encodes a highly conserved cell adhesion molecule that is the major component of peripheral myelin in all vertebrates; interestingly, MPZ is also expressed in the CNS myelin sheath of aquatic vertebrates (Bai et al., 2011; Jones et al., 2011; Lanwert and Jeserich, 2001; Saavedra et al., 1989; Tai and Smith, 1984; Waehneldt and Jeserich, 1984). Mammals express a single isoform of MPZ but several alternative isoforms have been detected in teleost and cartilaginous fish (Giese et al., 1992; Lanwert and Jeserich, 2001; Tai and Smith, 1984; Waehneldt and Jeserich, 1984). Heterozygous null *Mpz* mice display normal myelinated axons, which could be due to the presence of *Mbp*. However, homozygous null mice exhibit tremors, motor coordination abnormalities, and severe hypomyelination (Giese et al., 1992; Martini et al., 1995). Mice deficient of both *Mbp* and *Mpz* display severe PNS demyelination and die at about 5-6 weeks of age (FOLCH and LEES, 1951; Martini et al., 1995). In terrestrial vertebrates, proteolipid protein (PLP) is the major component of the CNS myelin (FOLCH and LEES, 1951; Puckett et al., 1987) and in mammals, PLP is expressed in Schwann cells but it is excluded from the myelin sheath (Puckett et al., 1987; Snipes et al., 1992). In addition to MBP, MPZ, and PLP, other proteins such as peripheral myelin protein (PMP22) (Mikol and Stefansson, 1988; Snipes et al., 1992), myelin-associated glycoprotein (MAG) (Mikol and Stefansson, 1988) are associated with the process of myelination in vertebrates.

Protein composition of myelin-like layer in invertebrates remains obscure and the presence of cell adhesion proteins such as MBP, MPZ, PLP, and PMP22 may have led to the formation of compact myelin in vertebrates. Sequencing the lamprey genome, a jawless vertebrate, has provided interesting evidence of the presence of myelin proteins such as PLP, MPZ, and PMP22 in spite of not having myelinated axons (Smith et al., 2013). *SOX* genes have been identified

throughout the animal kingdom and are transcriptional regulators that are important for many developmental processes. In particular SOXE proteins (SOX8, SOX9, and SOX10) are critical for neural crest specification and differentiation and SOX10 is critical for myelination (discussed in detail later) (Guth and Wegner, 2008). Interestingly, SoxE (*SoxE1*, *SoxE2*, and *SoxE3*) genes have been identified in lampreys and *SoxE3* displays limited sequence homology to *SOX9* (Lakiza et al., 2011). Whereas, *SoxE1* and *SoxE2* do not share any sequence similarity with either SOX8 or SOX10 and are not orthologous (Guth and Wegner, 2008; Lakiza et al., 2011). Owing to the importance of SOX10 in myelination, vertebrate-specific neofunctionalization of SOXE proteins and altered transcriptional regulation of genes important for myelination may have contributed to formation of compact myelin in gnathostomes. Additionally, evolution of cis-regulatory sequences required for the expression of these myelin genes may have evolved in vertebrates, which could explain the differences we observe between compact myelin and myelin-like layer in invertebrates. Thus, a thorough understanding of the regulatory pathways important for myelination will provide unique insights into the evolution of myelin in jawed vertebrates. The aim of my Ph.D. thesis research is to investigate the transcriptional regulation of genes important for vertebrate myelination.

1.7 Transcriptional regulation of peripheral nerve myelination

1.7.1 An overview of transcription

According to the central dogma of molecular biology: (i) DNA is transcribed into RNA, and this is a reversible process where RNA can be reverse transcribed to DNA; and (ii) RNA is translated into proteins. Transcription is a process catalyzed by the enzyme RNA polymerase in which information from DNA molecules is transcribed to a single-stranded messenger RNA (mRNA)

molecule. Bacteria contain a single RNA polymerase enzyme, in contrast to eukaryotes, which contain three RNA polymerases: RNA polymerase I, II, and III. RNA polymerase II (Pol II) transcribes all protein-coding genes and RNA polymerase I and III transcribe genes that code for transfer RNA, ribosomal RNA, and other small RNAs (Chambon, 1975).

A transcriptional unit is the region of DNA that contains appropriate signals (such as initiation and termination) which results in the generation of a nascent transcript. The main steps of transcription are: (i) formation of the preinitiation complex; (ii) promoter clearance; (iii) elongation; and (iv) termination. RNA polymerase initiates transcription by binding to a DNA segment called the promoter (Nevins, 1983). Pol II along with the general transcription factors forms a preinitiation complex (PIC) that binds to a core promoter sequence to initiate transcription. Promoters often contain an AT-rich sequence called the TATA box or Goldberg Hogness box, which is located 25-30 bases upstream from the transcription start site (TSS) (Corden et al., 1980).

A general transcription factor, TFIID, binds to the TATA-box and recruits other transcription factors and Pol II resulting in the formation of the PIC (Roeder, 1996). Binding of the PIC to the promoter is sufficient to drive low levels of transcription, a process referred to as basal transcription. Promoter sequences are usually asymmetric so the RNA polymerase can bind only in one direction and the template strand is chosen based on the location and orientation of the promoter (Roeder, 1996). Following the assembly of the PIC, TFIIH unwinds the DNA strands that reside roughly 12-14 base pairs around the TSS to create a transcription bubble (Roeder, 1996). Pol II facilitates the polymerization of the first nucleotide and RNA synthesis occurs in

the 5' to 3' direction. Within this transcription bubble, Pol II releases short transcripts called abortive transcripts (Munson and Reznikoff, 1981). Pol II contains a highly conserved domain called the C-terminal repeat domain (CTD), which is essential for Pol II function. It consists of 25-52 tandem repeats of the heptad sequence $Y_1S_2P_3T_4S_5P_6S_7$, with serine 2 and serine 5 capable of undergoing phosphorylation (Corden, 1990). The phosphorylation patterns change as transcription proceeds and each phosphorylation pattern recruits different proteins which serve different functions (Phatnani and Greenleaf, 2006). Phosphorylation of the CTD disrupts Pol II interactions with the basal transcription machinery and Pol II enters the elongation phase of transcription (Jiang et al., 1996). As the nascent RNA is being produced, the eukaryotic pre-mRNA undergoes processing to form the mature mRNA and the phosphorylation status of the CTD recruits the factors needed for producing the mature mRNA. The pre-mRNA undergoes 5'-capping, splicing, and 3'-end processing and these events occur co-transcriptionally to ultimately form the mature RNA (Proudfoot et al., 2002).

1.7.2 The function of cis-regulatory elements in transcriptional regulation

Cis-acting regulatory elements (CREs) are DNA sequences (typically, but not exclusively, non-coding sequences) associated with transcriptional regulation that play an important role in the spatial and temporal expression of a given locus; CREs have been shown to be critical for specifying gene expression patterns required for establishing diverse cell types. CREs usually contain consensus motifs (specific DNA sequences) for transcription factors and binding of TFs to these motifs elicits a transcriptional response. Promoters, enhancers, repressors, and insulators are well-characterized classes of CREs (Noonan and McCallion, 2010) (Fig. 1.4). CREs can drive basal levels of transcription (promoters), increase transcription (enhancers), inhibit

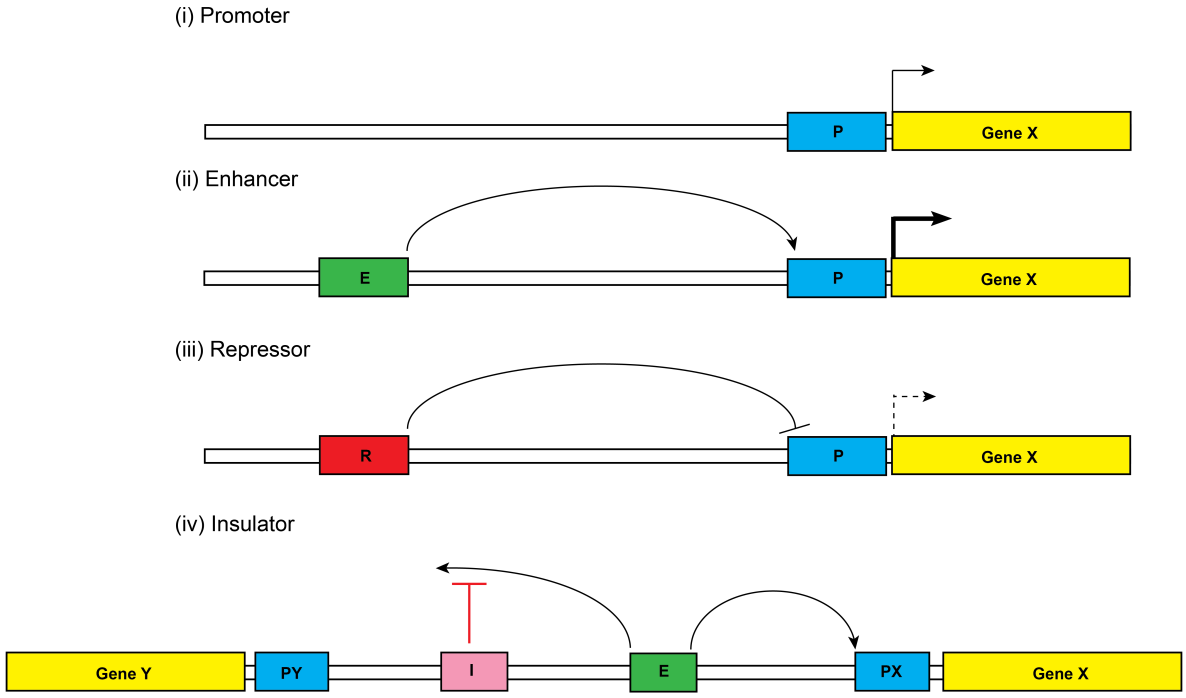


Figure 1.4 Schematic of the function of cis-acting regulatory elements. (i) The transcriptional machinery binds to the promoter (P) and leads to basal levels of gene expression. (ii) Enhancer (E) acts on the promoter in a position- and orientation-independent manner to activate gene expression. (iii) Repressor (R) acts on the promoter to decrease gene expression. (iv) Enhancer (E) activates the expression of gene X and the insulator (I) element prevents ectopic expression of gene Y. PX and PY are the promoters of Gene X and Y respectively. Adapted from Noonan and McCallion, 2010.

transcription (repressors), or prevent ectopic gene transcription (insulators) (Fig. 1.4). The core promoter element consists of a TATA box, an initiator element, and the transcription start site (TSS), which defines the direction of transcription. Only about 24% of human genes contain a TATA-box and most human promoters are GC-rich (Yang et al., 2007). CpG islands are regions of DNA that contain multiple cytosine-guanine repeats and ~70% of human promoters reside within these sequence elements (Saxonov et al., 2006). CpG islands play a critical role in regulating transcription by influencing the local chromatin environment. The cytosine in the CpG dinucleotide can undergo methylation to form 5-methylcytosine and about 80% of CpG islands in the human genome are methylated (Bird, 1999). Importantly, methylation of the CpG islands is inversely related to transcriptional activity and, as such, this process is an important mechanism for regulating gene expression (Razin and Riggs, 1980; Vardimon et al., 1982).

Most eukaryotic genes have a single promoter; however, certain genes in the human genome have multiple promoters with about three putative alternative promoters residing at each locus with more than one such element (Kimura et al., 2006). Use of alternative promoters—and alternative first exons—significantly contributes to the diversity of protein products expressed from a single gene. For example, a series of sequential, internal promoters can lead to a series of N-terminally truncated proteins that may have different cellular functions (Fogarty et al., 2016). In contrast, alternative promoters can produce transcripts with identical open reading frames (ORFs); 60-80% of alternative promoters produce protein with identical ORFs (Landry et al., 2003). Here, differential 5' untranslated regions may alter gene function, for example by altering the cellular location of a transcript or the rate of protein translation (Phelps et al., 1998). In addition to promoters, distal cis-acting regulatory elements such as enhancers, repressors, and

insulators can affect the expression levels of a given locus. Enhancers are DNA sequences that stimulate or enhance gene expression by exerting an effect on gene promoters. Enhancers are enriched with short DNA motifs that serve as binding sites for transcription factors. Binding of transcription factors to enhancers results in the recruitment of co-activators and chromatin remodeling proteins and that results in an open chromatin configuration that allows increased transcription. The first enhancer was identified in the animal virus SV40 in 1981 (Banerji et al., 1981). The authors transiently transfected HeLa cells with a plasmid containing the rabbit hemoglobin β1 gene and did not detect transcription. However, when they co-transfected a plasmid that contains a fragment from the Simian Virus 40 (SV40), it resulted in a 200-fold increase in hemoglobin β1 transcript levels (Banerji et al., 1981). To narrow down the region that caused enhanced hemoglobin β1 levels, the authors created several deletion mutants of the SV40 DNA fragment and tested their ability to increase transcription. This revealed a 72-bp fragment of the SV40 DNA that caused increased expression of the hemoglobin β1 transcript (Banerji et al., 1981). The DNA fragment within the mouse immunoglobulin heavy chain locus was the first eukaryotic enhancer discovered (Banerji et al., 1983). The authors cloned this enhancer fragment in a plasmid that expresses β-globin gene reporter. They cloned the enhancer at different locations (upstream and downstream) relative to the reporter gene and performed transient transfection as described above to measure the activity of the enhancer. This fragment was able to stimulate the transcription of β-globin gene in a distance- and orientation-independent manner (Banerji et al., 1983).

DNA is wrapped around nucleosomes, with each core nucleosome being an octamer with two copies each of the histone proteins H2A, H2B, H3, and H4. The amino termini of these histone

proteins can undergo post-translational modifications such as methylation and acetylation, and these modifications alter chromatin structure and affect gene transcription. DNase I is an enzyme that digests nucleosome depleted regions of DNA (open chromatin) and regulatory regions such as promoters, enhancers are located in such regions (Boyle et al., 2008a).

Enhancers can be classified as inactive, poised, primed, or active and can be distinguished by the chromatin accessibility and the type of histone modifications they are associated with (Heinz et al., 2015; Shlyueva et al., 2014). Chromatin associated with inactive enhancers is highly compacted and lacks chromatin modifications (Heinz et al., 2015). Enhancers are considered to be poised if they contain histone H3 lysine 27 trimethylation (H3K27me3), a Polycomb protein-associated repressive mark (Heinz et al., 2015; Zentner et al., 2011). Poised enhancers may exhibit some basal activity and signal dependent binding of transcription factors can cause them to become active (Heinz et al., 2015; Samstein et al., 2012). Primed enhancers, which represent an intermediate class of enhancers, are associated with H3K4 monomethylation (Zentner et al., 2011). Lineage specific factors recognize these chromatin marks and bind primed enhancers and regulates genes in a cell-specific manner (Heinz et al., 2015). H3K27 acetylation distinguishes active from poised or primed enhancers (Creyghton et al., 2010). Some of the prominent features of enhancers are: (i) they can act on gene promoters in a distance- and orientation-independent manner; (ii) they are sensitive to DNase I treatment; (iii) they contain sequence motifs for transcription factors; and (iv) the histones within the vicinity of enhancers can be modified as follows: poised (H3K27me3+, H3K27ac-); primed (H3K4me1+, H3K27ac-); and active (H3K4me1+, H3K27ac+).

With recent advances in sequencing technologies several genome-wide studies have identified enhancers based on the above properties. DNase-seq experiment was developed to identify the regions of open chromatin within the entire genome (Boyle et al., 2008a; Crawford et al., 2006). In addition to DNase-seq other techniques such as MNase-seq (Schones et al., 2008), FAIRE-seq (Giresi et al., 2007) and ATAC-seq (Buenrostro et al., 2013) have been developed to identify genome-wide distributions of open chromatin regions. Chromatin immunoprecipitation combined with massive parallel sequencing (ChIP-seq) involves crosslinking proteins to DNA, using an antibody against the protein of interest to isolate the protein-DNA complex, recovering the DNA fragments, and performing high-throughput DNA sequencing to identify DNA sequences bound to specific proteins (e.g., transcription factors or histones). Various studies—including the encyclopedia of DNA elements (ENCODE) consortium—have used the above techniques to identify thousands of enhancer elements in various species. The aim of the ENCODE project is to identify and catalogue all functional elements within the human genome (ENCODE Project Consortium et al., 2007). The ENCODE project has generated several ChIP-seq data sets across multiple cell types using antibodies against various histone marks, co-activator p300, and specific transcription factors. These techniques—and the associated, freely available resources—have proven invaluable for characterizing transcriptional hierarchies important for specific cell populations, and in this thesis I employ them for studies on Schwann cell biology.

1.8 The transcription factor SOX10 is essential for peripheral nerve myelination

There are two opposing transcriptional programs that control myelination in the peripheral nervous system: transcription factors that inhibit myelination and transcription factors that

promote myelination. Negative regulators of myelination are important during early Schwann cell development to inhibit ectopic myelination as precursors divide and migrate along peripheral nerve axons. As Schwann cell development progresses toward radial sorting and membrane extension, positive regulators increase the expression of genes important for the process of myelination. The transcription factors c-Jun and SOX2 (Jessen and Mirsky, 2008; Parkinson et al., 2008; Roberts et al., 2017a) are negative regulators of myelination and are expressed early during Schwann cell development. In contrast, SOX10, OCT6, and EGR2 are well known for their role in promoting myelination (Jessen and Mirsky, 2005). Interestingly, *OCT6* and *EGR2* expression begin at the point of myelin formation while *SOX10* is expressed throughout the Schwann cell (and neural crest) lineage. This observation raises the question of how myelination is inhibited in the context of a pro-myelination factor such as SOX10.

1.8.1 The SRY-related HMG-box 10 (SOX10) transcription factor

The SRY-related HMG-box (SOX) family of transcription factors plays an important role in regulating genes during embryonic development. The *sex-determining region Y* (*SRY*) gene encodes a protein, which contains a 75 amino acid protein domain called the high mobility group (HMG) box. The HMG box recognizes specific DNA sequences and facilitates binding of SRY to DNA (Harley et al., 1992). There are two types of HMG-box proteins: (i) those containing multiple HMG boxes, which bind DNA independent of the sequence; and (ii) those containing a single HMG box, which bind DNA in a sequence-dependent manner (Prior and Walter, 1996). Proteins in the SOX family of transcription factors have been a single HMG domain and all bind DNA in a sequence-dependent manner (Gubbay et al., 1990; Travis et al., 1991; Waterman and Jones, 1990). SOX proteins recognize the sequence 5'-(A/T)(A/T)CAA(A/T)G-3' within the

minor groove of DNA (Harley et al., 1994; Laudet et al., 1993). Based on the presence of an HMG domain and amino acid sequence similarities, 20 SOX proteins have been classified into eight groups: SOXA-H (Bowles et al., 2000). In spite of recognizing the same DNA sequence, SOX proteins from different groups have acquired distinct biological functions and spatio-temporal expression patterns (Kamachi and Kondoh, 2013; Kondoh and Kamachi, 2010).

In vertebrate species, the SOXE group includes SOX8, SOX9, and SOX10 (Guth and Wegner, 2008). The transcripts encoding SOX8, SOX9, and SOX10 were identified using RT-PCR on total RNA isolated from 11.5 days post coitum (dpc) mouse embryos (Wright et al., 1993). SOXE genes are expressed in the early neural crest cells (NCCs) and are involved in multiple stages of neural crest development (Cheung and Briscoe, 2003; Reiprich et al., 2008). SOX9 expression precedes SOX8 and SOX10 and there is evidence that SOX9 induces the expression of SOX10 (Cheung and Briscoe, 2003). SOX10 expression begins in late, pre-migratory NCCs and plays a crucial role in the development and specification of PNS neurons and glial cells (Kuhlbrodt et al., 1998; Stolt and Wegner, 2010). In the PNS, most neurons and glial cells (Schwann cells) are derived from the NCCs and expression of SOX8 and SOX9 are down regulated before the cells commit to a neuronal or glial fate (Weider and Wegner, 2017). As the neurons mature SOX10 expression is repressed, whereas Schwann cells continue to express SOX10 (Kuhlbrodt et al., 1998). Melanocytes are the other neural crest derived cell type that continue to express SOX10 (Mollaaghababa and Pavan, 2003). In addition to NCC derived cell lineages, SOX10 is important for oligodendrocyte development and myelination in the central nervous system (Li et al., 2007)

Although these three SOX proteins share >90% amino acid identity, they are not entirely functionally redundant. Replacing the *Sox10* locus with *Sox8* in mice was not sufficient to fully rescue the phenotype of *Sox10* null mice. *Sox8* was able to fully rescue the loss of *Sox10* in the PNS, however *Sox8* was unable to rescue the defects observed in melanocytes, oligodendrocytes, and the enteric nervous system (Kellerer et al., 2006). Recent data reveal that *Sox8* was able to rescue the loss of *Sox10* during later stages of oligodendrocyte differentiation. Specifically, deleting *Sox8* along with *Sox10* in differentiating oligodendrocytes in mice caused severe myelination defects suggesting that both *Sox8* and *Sox10* are required for maintaining myelination in the central nervous system (Turnescu et al., 2017). SOX10 is essential for all stages of Schwann cell development. Mice homozygous for a conditional null allele of *Sox10* lack Schwann cells leading to impaired development of the PNS (Britsch et al., 2001). Furthermore, ablation of *Sox10* expression at various stages of Schwann cell development (through the use of a conditional null allele and stage-specific CRE recombinase transgenes) results in defective PNS myelination. Specifically, deleting *Sox10* at the immature Schwann cell stage in mice causes absence of myelinating and non-myelinating Schwann cells (Finzsch et al., 2010). The peripheral nerves of these mice are translucent (due to the lack of myelin) compared to the opaque nerves from wild-type littermates and the mutant mice die by five-seven weeks after birth (Finzsch et al., 2010). Additionally, conditional deletion of *Sox10* in myelinating Schwann cells in mice leads to severe hypomyelination and impairs Schwann cell homeostasis (Fröb et al., 2012). Thus, SOX10 is indispensable for Schwann cell development, maintenance, and myelination.

1.8.2 Structure of the SOX10 protein

SOX10 encodes a 466 amino-acid protein that contains four major domains: a dimerization domain, an HMG domain, a K2 domain, and a transactivation domain (Fig. 1.5). The dimerization (DIM) domain is unique to the SOXE group and facilitates homodimerization and heterodimerization with other SOX proteins (Huang et al., 2015; Peirano and Wegner, 2000). SOX10 can function as a monomer or a dimer through binding a single (*e.g.*, 5'-ACAAA-3') or a head-to-head (*e.g.*, 5'-ACAAAnnnnnnTTTGT-3') consensus sequence, respectively (Peirano and Wegner, 2000). The length of the intervening spacer sequence between the monomeric sites can vary; however, a six-base pair spacer is most frequently observed (Schlierf et al., 2002). The HMG domain enables binding of SOX10 to DNA in a sequence-specific manner. Binding of SOX10 to the minor groove of DNA induces a bend in the DNA helix facilitating recruitment of several co-factors and thus, SOX10 acts as an architectural protein. The bending angle is 75-80° for a monomeric site and 103.5-122° for a dimeric site (Peirano and Wegner, 2000; Schlierf et al., 2002). The function of the K2 domain is poorly defined but there is evidence that this domain possess transactivation potential (Schepers et al., 2000). The K2 domain has also been called a "cell-specific transactivation domain" based on evidence from transgenic mice where the endogenous *Sox10* locus was replaced with a SOX10 construct lacking the K2 domain (Schreiner et al., 2007). Development of neural crest cells and oligodendrocytes were unaffected in mice lacking the K2 domain. However, loss of the K2 domain affected Schwann cell myelination, melanocytes, and the enteric nervous system. Schwann cells from the above mice were restricted in the promyelinating stage and did not express myelin genes such as *Mbp* and *Mpz* (Schreiner et al., 2007) suggesting that the K2 domain is important for Schwann cell myelination. The

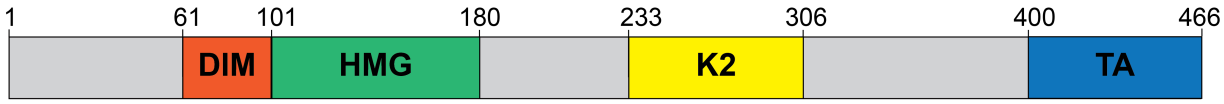


Figure 1.5 Structure of the SOX10 protein. The major domains of SOX10 are displayed from N- to C-terminus: dimerization (DIM; orange) domain, high mobility group (HMG; green) domain, K2 domain (yellow), and transactivation (TA; blue) domain. Numbers indicate amino acid positions along the protein.

transactivation (TA) domain is within the C-terminus of the SOX10 protein which is important for its regulatory function (Pusch et al., 1998). In addition, the TA domain physically interacts with positive transcription elongation factor b (P-TEFb) suggesting that SOX10 plays a role in transcription elongation (Arter and Wegner, 2015).

1.8.3 The function of SOX10 in Schwann cells

Each stage of Schwann cell development is tightly regulated by transcription factors and, not surprisingly, SOX10 plays a central role in this process. Our current understanding is that SOX10 directs a promyelinating program in Schwann cells by activating the expression of several key genes. NRG1/ErbB signaling is critical for early stages of Schwann cell development. While SOX10 is not known to regulate *NRG1*, SOX10 is known to regulate its receptor *ErbB3* (Prasad et al., 2011) (Fig. 1.6A). Binding of axonally derived ligand NRG1 to ERBB3 elicits a signaling cascade, which promotes Schwann cell differentiation(Leimeroth et al., 2002b). At the immature Schwann cell stage, SOX10 activates the expression of *OCT6*, a POU domain transcription factor. SOX10 binds to the Schwann cell-specific enhancer (SCE) located 10kb downstream of *OCT6* to activate gene expression (Jagalur et al., 2011; Mandemakers et al., 2000). Although *OCT6* is expressed in other cell types in the central nervous system and the skin, deleting the 4.3kb long SCE in mice resulted in Schwann cell specific ablation of *Oct6* expression (Ghazvini et al., 2002). SOX10 physically interacts with and recruits the chromatin remodeling protein called BRG1 to SCE (Weider et al., 2012) (Fig. 1.6B). Conditional deletion of *Brg1* in mouse Schwann cells leads to stalled Schwann cell differentiation and reduced *Oct6* expression, and these mice suffer from severe peripheral neuropathy (Ryu et al., 2007; Weider et al., 2012). The activation of *OCT6* also depends on

axonal signals that increase intracellular cyclic AMP (cAMP) levels. cAMP activates Protein Kinase A, an enzyme that phosphorylates CREB, and phosphorylated CREB activates the expression of target genes (Lee et al., 1999). Additionally, the SCE contains a consensus motif for the CREB transcription factor (Jagalur et al., 2011), suggesting that SOX10 and CREB regulate *OCT6*. After the induction of *OCT6*, the immature Schwann cells transition into the promyelinating stage. In this stage, SOX10, OCT6, and BRN2 activate expression of the transcription factor locus *Early Growth Response 2 (EGR2/KROX20)*. The three proteins bind to the myelinating Schwann cell element (MSE) located 35kb downstream of *EGR2* (Ghislain and Charnay, 2006). Activation of *EGR2* also depends on the Mediator complex. The Med12 subunit interacts with SOX10 and binds to the MSE (Vogl et al., 2013) (Fig. 1.6C). Importantly, the myelination program is arrested in the promyelination stage in *Egr2* null mice (Topilko et al., 1994). Once EGR2 is expressed, SOX10 and EGR2 synergistically activate the expression of genes important for myelination; indeed, SOX10 and EGR2 are critical for the formation and maintenance of the myelin sheath (Bremer et al., 2011; Decker et al., 2006; Fröb et al., 2012; Topilko et al., 1994). Genome-wide studies have been performed that reveal co-occupancy of SOX10 and EGR2 at several loci (Srinivasan et al., 2012). These and other studies have revealed that SOX10 and ERG2 synergistically regulate the expression of Peripheral Myelin Protein 22 (*PMP22)* (Jones et al., 2011)·(Jones et al., 2012), Myelin Basic Protein *(MBP)* (Denarier et al., 2005), Myelin Protein Zero *(MPZ)* (LeBlanc et al., 2006) and Connexin32 *(GJB1)* (Bondurand et al., 2001; Jang et al., 2010) (Fig. 1.6D). Taken together, these have added to our current understanding that SOX10 regulates a promyelination program in Schwann cells by activating key genes at different developmental stages of Schwann cell development.

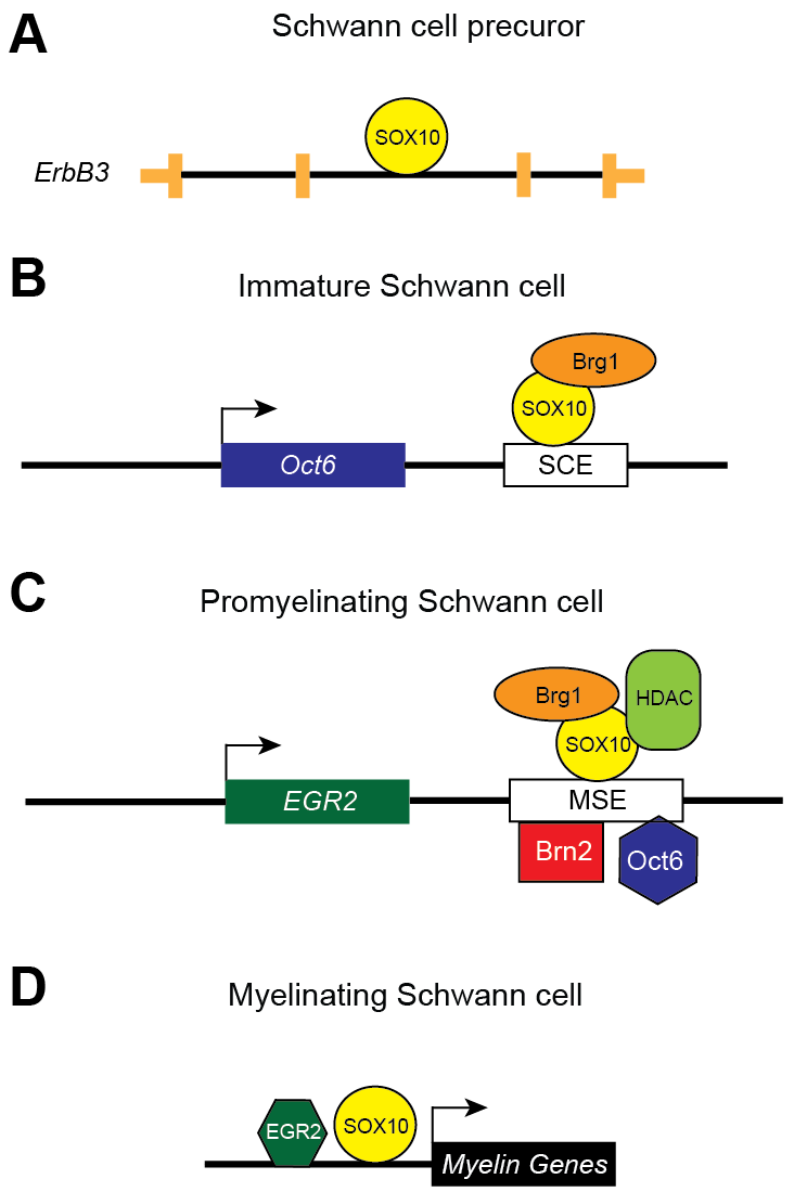


Figure 1.6 A transcriptional hierarchy for Schwann cell development. (A) SOX10 activates the expression of ERBB3 in Schwann cell precursors. (B) SOX10, along with the chromatin remodeler BRG1, activates the expression of OCT6 via a Schwann cell specific enhancer (SCE) in immature Schwann cells.(C) In the promyelinating stage, SOX10 activates the expression of EGR2 via the myelinating Schwann cell enhancer (MSE). Other factors such as OCT6, BRN2, BRG1, and histone deacetylase (HDAC) bind to the MSE. (D) SOX10 and EGR2 act synergistically to activate expression of myelin genes. Loci are not drawn to scale.

1.8.4 The function of SOX10 in other cell types

SOX10 expression begins as the neural crest cells start to migrate and neural crest cells give rise to many distinct cell types including melanocytes, enteric nervous system, sensory neurons, and Schwann cells (Cheung and Briscoe, 2003; Erickson and Reedy, 1998) consistent with SOX10 being important for the development of other neural crest derivatives. Furthermore, SOX10 is expressed in developing and differentiated oligodendrocytes—cells that do not arise from the neural crest—indicating that SOX10 is critical for myelination in general. In melanocytes, SOX10 regulates the expression of microphthalmia-associated transcription factor (*MITF*) (Lee et al., 2000). During the later stages of melanocyte development, SOX10 and MITF synergistically activate the expression of dopachrome tautomerase (*DCT*), which functions in melanin synthesis (Jiao et al., 2004). In the enteric nervous system, SOX10 and PAX3 activate the expression of the tyrosine kinase receptor locus *c-RET,* which is important for the development of these cells (Lang et al., 2000). Finally, SOX10 is expressed in the oligodendrocyte progenitors in the central nervous system (CNS) and plays a crucial role in oligodendrocyte myelination (Stolt et al., 2002). In oligodendrocyte precursor cells, OLIG2 regulates the expression of *SOX10* by binding to a distal enhancer U2 or MCS4 (Antonellis et al., 2008; Küspert et al., 2011). Once induced, SOX10 and OLIG2 regulate the myelination program in the CNS by activating the expression of myelin genes such as proteolipid protein (*PLP*), myelin-associated glycoprotein (*MAG*), and *MBP* (Stolt et al., 2002). SOX10 is well known as an activator of gene expression; however, SOX10 has been shown to repress gene expression in Bergmann glial cells in chick brain. Specifically, SOX10 represses the expression of chick kainate binding protein, which is a member of the ionotropic glutamate receptor family (Cruz-Solis et al., 2009).

One of the earliest studies on the function of SOX10 were performed in a spontaneous mutation that arouse in mice termed Dominant megacolon (*Dom*). These mice contain a frame shift mutation in the *Sox10* coding region that abolishes the transactivation domain (Herbarth et al., 1998; Lane and Liu, 1984; Southard-Smith et al., 1998). $Dom^{-/-}$ mice are embryonic lethal and heterozygous animals display defects in neural crest derived cell lineages. Loss of *Sox10* in heterozygous *Dom* mice causes pigmentation defects (due to loss of melanocytes), loss of enteric ganglia, and loss of glial cells and neurons in the PNS (Herbarth et al., 1998; Southard-Smith et al., 1998). Similar defects were observed in the zebrafish mutant *colorless* (*cls*) (Dutton et al., 2001; Kelsh and Eisen, 2000). Another spontaneous mouse mutant of Sox10 called $Sox10^{Hry}$ is characterized by pigmentation defects and reduced number of enteric neurons (Antonellis et al., 2006). This phenotype is associated with reduced *Sox10* expression caused by a ~16kb deletion of non-coding sequences located ~40kb upstream of *Sox10* TSS. It was then shown that the deletion contains non-coding sequences important for *Sox10* expression (Antonellis et al., 2006). Detailed characterization of the ~16kb deletion revealed three *cis*-acting regulatory elements that contribute to the phenotype observed in $Sox10^{Hry}$ mice (Antonellis et al., 2008).

1.8.5 Mutations in SOX10 and in SOX10 target genes cause demyelinating peripheral neuropathy

Mutations in the human *SOX10* gene cause several neurocristopathies and, interestingly, different *SOX10* mutations correlate strongly with disease severity and expressivity. SOX10 mutations can cause Waardenburg-Hirschsprung or Waardenburg-Shah disease (WS4) (Pingault et al., 1998) or a more severe syndrome called peripheral demyelinating neuropathy, central dysmyelinating leukodystophy, Waardenburg, and Hirschsprung disease (PCWH) (Inoue et al.,

2002). WS4 is a combination of Waardenburg syndrome and Hirschsprung disease; individuals with WS4 display pigmentation defects, hearing loss, heterochromia, and aganglionic megacolon (Omenn and McKusick, 1979). PCWH is characterized by defects in Schwann cells, oligodendrocytes, melanocytes, and enteric neurons (Inoue et al., 2002). The distinct phenotypes observed depends on the type of *SOX10* mutation. Disease-associated *SOX10* mutations that result in a premature stop codon early in the transcript undergo nonsense-mediated decay (NMD) resulting in haploinsufficiency and causing WS4. Mutations in *SOX10* that result in premature stop codons later in the transcript escape NMD and result in PCWH. In the latter situation, the mutant protein acts in a dominant-negative manner on the remaining wild-type SOX10 protein and resulting in a greater than 50% reduction in SOX10 function, thus giving rise to a more severe phenotype (Inoue et al., 2004). The peripheral neuropathy phenotype observed in individuals with PCWH is similar to Charcot-Marie-Tooth (CMT) disease. CMT, also referred to as hereditary motor and sensory neuropathy, has a prevalence of one in 2,500 individuals worldwide. It is divided into two major subclasses based on electrophysiological studies: CMT type 1 and CMT type 2. (Warner et al., 1999). CMT1 primarily affects the Schwann cells and CMT2 affects the axons of the peripheral nerves. Demyelinating CMT can be divided into two major forms based on the inheritance pattern: autosomal dominant (CMT1) and autosomal recessive (CMT4) (Brennan et al., 2015).

CMT1 is an early onset disorder and patients typically present with symptoms during the first two decades of life. Symptoms include foot deformity, distal muscle weakness, sensory loss, and muscle atrophy (Saporta et al., 2011). As noted above, the myelin sheath acts as an electrical insulator and allows rapid propagation of electrical impulses. The speed at which the electric

impulses are conducted along the nerve is called the motor nerve conduction velocity (MNCV). The MNCV of patients suffering from CMT1 is <40 meters per second (m/s); normal ranges are 50-60m/s (Saporta et al., 2011). The most common cause of CMT1 is a 1.4MB duplication on chromosome 17p11.2, which accounts for 60-70% of CMT cases (Lupski et al., 1991; Saporta et al., 2011). Interestingly, many forms of demyelinating CMT are caused by mutations in SOX10 target genes that are critical for myelination such as *PMP22* (Lupski et al., 1991), *MPZ* (Hayasaka et al., 1993), *EGR2/KROX20* (Warner et al., 1998) and *GJB1/CX32* (Bergoffen et al., 1993). Indeed, mutations within regulatory regions at many of these loci have been identified in patients with CMT. For example, duplications of distal regulatory elements at *PMP22*, not including the transcriptional unit, are observed in patients with a mild form of CMT (Weterman et al., 2010). Similarly, pathogenic variants have been identified in the promoter region of *GJB1*; and functional studies have revealed that the mutations reside within SOX10 binding sites and that they reduce SOX10 binding (Kabzińska et al., 2011). Recently, deletion of this promoter has been identified in patients with CMT and the deletion segregated with disease (Kulshrestha et al., 2017). It has long been noted that patients with CMT disease display severe variability in age of onset and disease severity, even among patients with an identical pathogenic variant. This phenotypic variability suggests that there are other genetic factors that could modify disease severity. Thus, identifying additional SOX10-responsive regulatory sequences will improve our understanding of myelination and novel sequences to screen for mutations and modifiers of peripheral neuropathy.

1.9 Negative regulators of myelination

Positive and negative signaling pathways in Schwann cells ensure that myelination occurs at the correct developmental stage and similar pathways regulate remyelination following peripheral nerve injury. One of the notable features of myelinating Schwann cells is their plasticity, that is, the ability to dedifferentiate to form immature Schwann cells following injury (Fig 1.2). Damage to the peripheral nerve can be due to physical injury or to genetic mutations as observed in patients with CMT1. Following nerve injury, peripheral axons degenerate and myelinating Schwann cells lose axonal contact leading to demyelination (Jessen and Mirsky, 2008). Subsequently, macrophages are activated and recruited to the site of injury and phagocytize myelin and axonal debris (Martini et al., 2008). In the event that an axon begins to regenerate, Schwann cells reinitiate the myelination program.

Negative regulators of myelination are often highly expressed during early stages of Schwann cell development, and then down regulated as myelination occurs. Furthermore, a subset of these regulators may be induced after nerve injury to facilitate the dedifferentiation process (Jessen and Mirsky, 2008). To date, a number of transcription factors have been characterized as myelin inhibitors including c- JUN, SOX2, and EGR1/3 (Jessen and Mirsky, 2008). c-JUN is a ubiquitously expressed transcription factor that is a component of the AP-1 complex. c- JUN is expressed during early stages of Schwann cell development and blocks myelination by repressing *EGR2* expression (Parkinson et al., 2008). Additionally, c- JUN expression is induced post nerve injury (and in the nerves of patients with CMT1) and mediates the Schwann cell dedifferentiation process (Hutton et al., 2011; Klein et al., 2014; Parkinson et al., 2008). Recently, genome-wide analyses of gene expression changes were assessed in mature rat

peripheral nerves pre- and post-nerve injury. The authors performed ChIP-seq for H3K27ac marks using sham and injured rat sciatic nerves, which revealed an enrichment of c-JUN motifs in injury-induced enhancers at genes known to be upregulated post injury (Hung et al., 2015). Other transcription factors such as NOTCH1, Inhibitor of DNA-binding 2 (ID2) are also known to repress myelination (Quintes and Brinkmann, 2017; Woodhoo et al., 2009). Thus, a balance between positive and negative regulators is essential for proper myelination during development and to facilitate recovery after injury. However, little is known about how these negative regulators are activated during development and how the balance between the two opposing myelination programs is achieved.

1.10 Summary

Myelin is a multi-layered sheath formed by glial cells that wrap a membrane around axons in the peripheral and central nervous system. This sheath is unique to vertebrates and seems to have co-evolved with jaws and certain sections of the brain. The myelinating glia of the peripheral and central nervous system are Schwann cells and oligodendrocytes, respectively. Myelination is a dynamic process orchestrated by transcriptional regulators. One gene that is essential for Schwann cells encodes the SRY-related HMG box 10 (SOX10) transcription factor. SOX10 is expressed during all stages of Schwann cell development and directs a transcriptional hierarchy that promotes PNS myelination. Importantly, mutations in *SOX10* and in SOX10 target genes result in demyelinating neuropathy. Thus, identifying and characterizing the full panel of SOX10 target loci will provide novel insights into SOX10 function, Schwann cell biology, and the pathogenesis of demyelinating neuropathy. My dissertation focuses on defining regulatory pathways important for Schwann cells by identifying SOX10 target genes.

In chapter 2, I will discuss the computational pipeline we developed to identify conserved SOX10 consensus sequences. I performed *in vitro* assays to validate a handful of putative SOX10 response elements and these data revealed a novel, and rather surprising, role for SOX10 in Schwann cells. Specifically, that SOX10 activates the expression of negative regulators of myelination including *SOX5*, *SOX6*, *NOTCH1, ID2, ID2, MYCN*, and *HES5*. In chapter 3, I provide evidence for SOX10 regulating the expression of *SOX6* via an alternative promoter. To further study the role of SOX6 in Schwann cells, I performed expression studies to identify direct and indirect SOX6 target genes. Finally, in chapter 4, I will discuss the impact of my thesis work and present outstanding questions that can be pursued for future investigation.

Chapter 2

Stringent Comparative Sequence Analysis Reveals SOX10 as a Putative Inhibitor of Glial Cell Differentiation

2.1 Introduction

Schwann cells produce the myelin sheath in the peripheral nervous system (PNS), which allows rapid saltatory conduction and long-range communication between the central nervous system and innervated muscles and sensory organs. Schwann cell development is directed by a transcriptional hierarchy that promotes the expression of proteins important for migration along peripheral nerves, radial sorting of axons, and the initiation of myelination (Jessen and Mirsky, 2005; Stolt and Wegner, 2016). Atop this hierarchy sits the transcription factor SOX10, which is critical for the development and long-term function of Schwann cells (Kuhlbrodt et al., 1998)and is expressed during all stages of Schwann cell development (Britsch et al., 2001; Kuhlbrodt et al., 1998).

Three major lines of evidence underscore the importance of SOX10 for the function of Schwann cells. First, ablation of *Sox10* expression in mouse models causes: (***i***) a lack of Schwann cells

when performed during early development (Britsch et al., 2001); (***ii***) lethality due to peripheral neuropathy when performed in immature Schwann cells (Finzsch et al., 2010); and (***iii***) demyelination of peripheral nerves when performed in terminally differentiated Schwann cells (Bremer et al., 2011). Second, dominant-negative *SOX10* mutations cause an autosomal dominant disease characterized by peripheral demyelinating neuropathy, central dysmyelinating leukodystrophy, Waardenburg-Shah syndrome, and Hirschsprung disease (Inoue et al., 1999; 2004); the non-PNS phenotypes reflect the role of SOX10 in other neural crest derivatives (*i.e.*, melanocytes and enteric neurons) and in oligodendrocytes. Finally, mutations in SOX10 target genes including those encoding peripheral myelin protein 22 (*PMP22*), myelin protein zero (*MPZ*), and gap junction beta 1 (*GJB1*) cause demyelinating peripheral neuropathy (Bondurand et al., 2001; Jones et al., 2011; 2007; LeBlanc et al., 2006; Peirano et al., 2000).

The identification of additional SOX10 response elements and target loci will provide important information on the process of myelination in the peripheral nerve as well as novel target sequences to scrutinize for mutations and modifiers of peripheral neuropathy. Indeed, genome-wide analyses have been essential for characterizing SOX10 biology in Schwann cells (Lopez-Anido et al., 2015; Srinivasan et al., 2012); however, these efforts have primarily focused on identifying positive regulators of myelination by examining tissues or cells in a myelinating state. Thus, less-biased approaches are needed to complement the above studies and to identify functions of SOX10 outside of the regulation of promyelinating loci.

In this chapter, I describe a stringent computational strategy to rapidly predict SOX10 response elements in the human genome. Combined with molecular functional studies, this strategy

revealed that SOX10 positively regulates the expression of *SOX5*, *SOX6*, *NOTCH1*, *HMGA2*, *HES1*, *MYCN*, *ID4*, and *ID2*. Interestingly, each of these genes has a known or predicted role in the negative regulation of glial cell differentiation. As such, we identified a putative novel role for SOX10 in Schwann cells and present a model where SOX10 activates the expression of negative regulators of myelination to temper the pro-myelinating program during non-myelinating stages of Schwann cell development.

The author did all of the work presented in this chapter except the computational pipeline, which was generated by Drs. Law, Antonellis, and Prasad. The Svaren laboratory at the University of Wisconsin performed the qRT-PCR analyses. The DNase hypersensitivity assay was performed by Dr. Crawford's laboratory at Duke University.

2.2 Methods

2.2.1 Computational identification and prioritization of SOX10 consensus sequences

To identify all SOX10 consensus sequences in the human genome we downloaded individual text files for each human chromosome (hg18) from the UCSC Human Genome Browser (Kent et al., 2002) and wrote a Perl script that examines each file for the consensus sequences (using a regular expression analysis) 'ACACA' or 'ACAAD'; where 'D' is a G, T, or A nucleotide. To identify two SOX10 consensus sequence monomers that are oriented in a head-to-head manner (and that may represent a dimeric SOX10 binding site) we wrote a second Perl script that examines the human chromosome text files and reports each 'ACACA' or 'ACAAD' consensus sequence that is five to 10 base pairs 5' to the reverse complement of this consensus sequence: 'TGTGT' or 'HTTGT', where 'H' represents an A, C, or T nucleotide.

To identify genomic sequences that are identical between human, mouse, and chicken we downloaded the vertebrate (44 species) multiz alignment (maf) files from the UCSC Human Genome Browser (hg18) and extracted the alignments for human, mouse, and chicken. Next, we utilized the program ExactPlus (Antonellis et al., 2006) to identify all human sequences that are at least 5 base pairs long and identical between all three species. All subsequent computational analyses that assess for overlap between these and other datasets were performed using the UCSC Table Browser (Karolchik et al., 2004) and the 'intersection' tool. For these analyses we employed UCSC Genome custom tracks containing each: (**1**) human RefSeq (hg18) protein-coding sequence to exclude coding sequences; (**2**) human RefSeq entry (hg18) plus 2.5 kb upstream and 2.5 kb downstream of the transcriptional unit to identify regions that map to known genes; (**3**) SOX10 ChIP-seq peak in the rat genome (rn5) using HOMER analysis (Heinz et al., 2010) of previously described P15 sciatic nerve data sets; and (**4**) DNase-seq peak in the rat genome (rn5) that has an F-Seq (Boyle et al., 2008b) score of at least 0.08.

To identify the 57 loci with a known or predicted role in peripheral nerve myelination we performed the following PubMed searches in September 2014: (**1**) each gene name plus 'Schwann'; and (**2**) each gene name plus 'Myelin'. We also searched for each gene name plus 'Schwannoma' in the GEO Profiles database at NCBI to determine if gene expression is depleted upon treatment with SOX10 siRNA (Lee et al., 2008).

2.2.2 Luciferase reporter gene expression constructs

Primers containing attB1 and attB2 Gateway cloning (Thermo Fisher Scientific) sequences were designed for each region identified by our computational pipeline and were synthesized by IDT.

The primers were resuspended using ultrapure water (Thermo Fisher Scientific cat # 10977023) to make 200μM stock solutions and diluted 1:10 prior to PCR. Regions were PCR-amplified using human or rat genomic DNA and subsequently cloned into pDONR221 using BP clonase (Thermo Fisher Scientific cat # 11789020). Protocol used for each BP reaction: 2μL PCR amplified product, 0.5μL pDONR221, 1.5μL TE buffer, and 1μL BP clonase. The reaction was incubated at room temperature for one hour. To stop the reaction, 1μL Proteinase K was added and the reaction was incubated at 37°C for ten minutes. 3μL of above reaction was mixed with 12.5μL Top10 One shot *E.coli* cells (Thermo Fisher Scientific cat # C404003) and incubated on ice for 15 minutes. The bacteria were heat shocked at 42°C for 45 seconds and were recovered by adding 62.5μL SOC growth media (Thermo Fisher Scientific cat # C404003) and incubating at 37°C with shaking. The transformation reaction was plated on 50mg/mL kanamycin selective plates. The plates were incubated overnight at 37°C. Individual colonies were picked and grown in 7mL of kanamycin selective media. Plasmid DNA was isolated using the Qiagen miniprep kit (Qiagen cat # 27104) and was subjected to BsrG1 (NEB cat # R0575S) digestion to ensure presence of the insert. The following protocol was used for BsrG1 reactions: 6.75μL ultrapure water, 1μL (150ng) plasmid DNA, 1μL bovine serum albumin (BSA), 1μL NEB buffer 2, and 0.25μL BsrG1. The reaction was incubated at 37°C for one hour. The reaction products were run on a 1% agarose tris-EDTA-borate (TBE) gel to identify pDONR221 plasmids that contained our insert. These pDONR221 constructs were sequence verified using Sanger sequencing and cloned upstream of an E1B minimal promoter (Antonellis et al., 2006) directing luciferase expression in the forward (pE1B Forward) or reverse (pE1B Reverse) orientation using LR clonase (Thermo Fisher Scientific cat # 11791100). Mutagenic primers to delete SOX10 consensus sequences were designed using an online tool called QuikChange primer design

(http://www.genomics.agilent.com/primerDesignProgram.jsp) and were synthesized by IDT. The primers were re-constituted to 1250ng/μL and subsequently diluted 1:10 prior to the reaction. Site directed mutagenesis performed using the QuikChange II Mutagenesis Kit (QuikChange II cat # 200521) using the following protocol: 5μL 10X buffer, 1μL pDONR construct (50ng/μL), 1μL forward primer, 1μL reverse primer, 1μL dNTPs, 3μL Quik solution, 38μL ultrapure water and 1μL Pfu tag polymerase. After amplification, 1μL Dpn1 enzyme was added to the reaction and incubated at 37°C for two hours. Following Dpn1 digestion, ethanol precipitation was performed using the following protocol: Quikchange reaction, 150μL 100% ethanol, and 5μL 3M sodium acetate. The reaction was incubated at -80°C for two hours and then centrifuged at 13200 rpm for 30 minutes at 4°C. After centrifugation, the supernatant was discarded and the pellet was dried using vacuum centrifuge for five minutes. The dried pellet was re-suspended in 15μL ultrapure water and transformed into Top10 one shot *E.coli* cells as described above. The resulting mutant constructs were sequence verified to ensure that the SOX10 consensus site was deleted and that no other mutations were generated. Subsequently, the constructs were recombined into the pE1B luciferase vector as described above.

2.2.3 Cell culture and luciferase assays

Cultured rat Schwann cells (S16)(Goda et al., 1991; Toda et al., 1994) and mouse motor neuron derived cells (MN1) (Salazar-Grueso et al., 1991) were maintained in Dulbecco's modified Eagle's medium (Thermo Fisher Scientific cat # ILT12430054) supplemented with 10% fetal bovine serum (Thermo Fisher Scientific cat # 26140-079), 2mM L-glutamine (Thermo Fisher Scientific cat # ILT25030081), and 1X Penicillin-Streptomycin (Thermo Fisher Scientific cat # ILT15070063). For luciferase assays, $1x10^4$ cells were plated in each well of a 96-well culture

plate (Corning Life Sciences cat # 07-200-565) and incubated overnight at 37°C in 5% CO_2. Cells were transfected using Lipofectamine 2000 (Thermo Fisher Scientific 11668-019). Transfections were performed using the following protocol for each well: 25µL Opti-MEM (Thermo Fisher Scientific cat # 31985062) was mixed with 0. 25µL Lipofectamine 2000 and this mixture was incubated for 10 minutes at room temperature (Mix 1). 200ng of the pE1B construct (Antonellis et al., 2006) was mixed with 2ng of pCMV-Renilla in 25µL of Opti-MEM (Mix 2). Prior to transfection, pE1B plasmid for each region was diluted to 200ng/µL and pCMV-Renilla was diluted to 2ng/µL. After the 10-minute incubation 25µL of Mix 1 was added to Mix 2, mixed well, and incubated at room temperature for 20 minutes. The cells were washed with 60µL 1XPBS and 50µL of transfection mixture was added to each well. Cells were incubated at 37°C for 4 hours and then grown in 75µL standard growth media for 48 hours. For overexpression studies, 100ng of a construct to express wild-type or E189X SOX10 (Inoue et al., 2004), wild-type SOX8, SOX9, or EGR2 was included in the transfection reaction. 48 hours after transfection the cells were washed with 60µL 1XPBS and subsequently lysed in 20µL of 1X passive lysis buffer [4µL of 5X Passive Lysis Buffer (Promega cat # E1980) and 16µL ultrapure water]. The plate was shaken for one hour at room temperature. 10µL of the lysate was transferred to a 96-well white assay plate (Corning Life Sciences cat # CLS3789) and a Dual Luciferase Assay (Promega cat # E1980) was performed using a luminometer to determine the activities of luciferase and renilla. Luciferase activity was normalized to renilla activity and the fold increase in luciferase activity was calculated relative to the empty control vector (set to 1), which does not contain a genomic insert. For each genomic segment, three independently isolated reporter gene constructs were studied in eight technical replicates for a total of 24 reactions per segment.

2.2.4 Standard and quantitative RT-PCR

Total RNA was isolated from S16 and MN1 cells. $1x10^5$ cells were plated in a 6-well plate (USA Scientific cat # CC7682-7506) and incubated overnight at 37°C in 5% CO_2. Transfections were performed using the following protocol for each well: 1mL Opti-MEM (Thermo Fisher Scientific cat # 31985062) was mixed with 10µL Lipofectamine 2000 and this mixture was incubated for 10 minutes at room temperature (Mix 1). 4µg of wild-type or E189X SOX10 (Inoue et al., 2004) 1mL of Opti-MEM (Mix 2). After the 10-minute incubation 1mL of Mix 1 was added to Mix 2, mixed well, and incubated at room temperature for 20 minutes. The cells were washed with 2mL 1XPBS and 2mL of transfection mixture was added to each well. Mock transfections were performed in the absence of DNA. Cells were incubated at 37°C for 4 hours and then grown in 3mL standard growth media. After 72 hours, total RNA was isolated from the transfected cells using the RNeasy kit (Qiagen cat # 74104). Subsequently, cDNA was synthesized using 1µg of total RNA and the High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific 4368814). The following protocol was used for cDNA synthesis: 3.2µL ultrapure water, 2µL 10X random primers, 2µL 10X RT buffer, 0.8µL 25X dNTPs, 1µL RNase inhibitor, 1µL reverse transcriptase, and 1µg RNA in 10µL ultrapure water. The cycler conditions are as follows: 25°C for 10 minutes, 37°C for two hours, and 85°C for five minutes. For standard RT-PCR was performed on isolated cDNA using gene specific primers. A PCR for *β-actin* served as a positive control. All PCR products were subjected to DNA sequencing to confirm specificity. For quantitative RT-PCR, RNA was purified from three independent rat sciatic nerves at the P1, P15, and adult time points using the RNeasy Lipid kit (Qiagen cat # 74804), and quantitative RT-PCR was performed as described (Gokey et al., 2012).

2.2.5 siRNA-mediated depletion of SOX10

Control siRNA (siControl 1, Ambion catalog number AM4611) or Sox10 siRNA (siSox10 1, Life Technologies catalog number s131239) were transfected into S16 cells as described using the Amaxa Nucleofection system following the manufacturer's instructions. At 48 hours post-transfection, RNA was isolated using Tri-Reagent (Ambion) and analyzed by quantitative RT-PCR as described (Gokey et al., 2012). The Svaren laboratory performed this experiment.

2.2.6 DNase hypersensitivity site identification

DNase-seq was performed with three biological replicates of rat Schwann (S16) cells at passage numbers five, eight, and 14. Each replicate contained ~20 million cells frozen into 1 mL of freezing media. Cells were thawed and DNase-seq libraries generated as previously described (Song and Crawford, 2010) with the exception of adding a 5' phosphate to linker 1 to increase the ligation efficiency. DNase-seq libraries from three replicates were pooled into one lane of an Illumina Hi-Seq 2000. Raw reads were aligned to the rat (rn5) genome using Bowtie (Langmead, 2010) and unique mapping with up to two mismatches allowed in an alignment. For the three samples, 69.2% (36,295,401), 70.8% (43,564,606), and 67.9% (39,579,719) of the reads mapped to rn5. Peaks were called using F-Seq and the default settings(Boyle et al., 2008b). For the three samples: 502,787 (sample 1), 438,254 (sample 2), and 412,267 (sample 3) peaks were identified. 149,342 were shared among all three samples. We next used sample 2 as a representative experiment and extracted all DNase-seq peaks with an F-Seq score [52] of at least 0.08. This revealed a set of 31,845 peaks (7.3%) that were used to prioritize SOX10 response elements (see section 'Computational identification and prioritization of SOX10 consensus sequences' above).

Data from the three samples have been submitted to GEO (GSM2166058, GSM2166059 and GSM2166060). The Crawford laboratory performed this experiment.

2.3 Results

2.3.1 Genome-wide prediction of SOX10-responsive transcriptional regulatory elements

SOX10 binds to a well-defined consensus sequence ('ACACA' or 'ACAAD'; where 'D' is a G, T, or A nucleotide) as a monomer or as a dimer when two consensus sequences are oriented in a head-to-head fashion (Peirano and Wegner, 2000; Srinivasan et al., 2012). To identify putative SOX10 binding sites in the human genome, we wrote a Perl script to scan each human chromosome and report all occurrences of the above SOX10 consensus sequence. This revealed over 33 million monomeric consensus sequences and ~549,000 dimeric consensus sequences with an intervening sequence of five to 10 base pairs.

Multiple-species conservation analysis is an effective approach for predicting non-coding DNA sequences with a role in transcriptional regulation (Antonellis and Green, 2008). Importantly, functionally validated SOX10 binding sites have been identified in non-coding genomic sequences that are conserved between human and chicken (Antonellis et al., 2008; Gokey et al., 2012; Hodonsky et al., 2012). To prioritize the large dataset of SOX10 consensus sequences, we aligned the human, mouse, and chicken genomes and identified all genomic sequences that are five base pairs or longer (the length of the monomeric SOX10 consensus sequence) and that are identical between these three species.

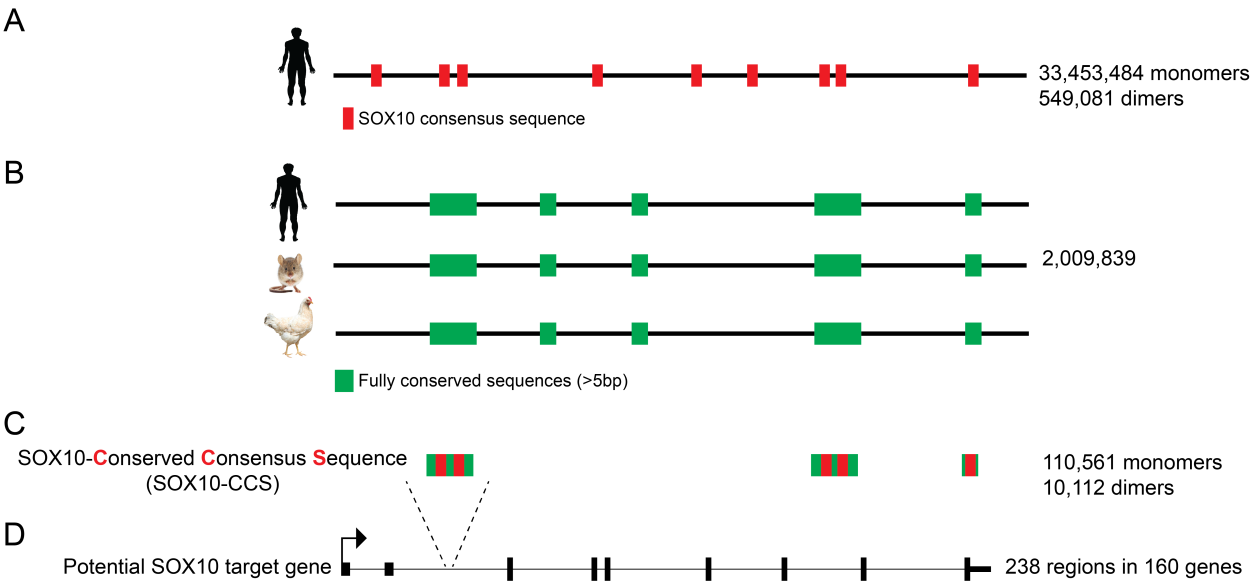


Figure 2.1. Computational pipeline to predict SOX10-responsive genomic regions. (A) Each of the 33,453,484 monomeric (depicted as a red box) and 549,081 dimeric SOX10 consensus site identified in the human genome (HG18). (B) Multi-species conservation analysis identified 2,009,829 conserved genomic segments (longer than five base pairs, shown in green) among human, mouse, and chicken. (C) Compared data sets generated in (A) and (B) and identified 110,561 monomeric and 10,112 dimeric SOX10 conserved consensus sequences (SOX10-CCS) that are conserved in human, mouse, and chicken. (D) Prioritized dimeric conserved SOX10 sites within introns and proximal promoter elements of RefSeq genes. The arrowhead indicates annotated transcription start site.

This revealed over two million conserved coding and non-coding genomic segments. To develop a panel of prioritized SOX10 consensus sequences for functional studies, we used the rationale that: (**1**) focusing on conserved dimeric SOX10 consensus sequences will enrich for *bona fide* SOX10 binding sites; (**2**) focusing on non-coding sequences will deprioritize sequences that are conserved due to the function of the gene product; and (**3**) focusing on proximal promoter and intronic sequences will provide a candidate target gene for further studies. Thus, we compared the above datasets to identify dimeric SOX10 consensus sequences that are conserved between human, mouse, and chicken (including the intervening sequence), reside in non-coding sequences, and map to an intron or 2.5 kb upstream or downstream of a known (RefSeq) human gene. This revealed 238 genomic sequences at 160 loci for further study. To determine the efficacy of our approach, we further prioritized the above 238 genomic segments by identifying the subset that map to loci with a known or predicted role in myelination. This revealed 57 genomic sequences at 32 loci with a conserved, dimeric SOX10 consensus sequence that resides within an intron or directly upstream of a myelin-related transcriptional unit; we named these elements SOX10 Conserved Consensus Sequences (Fig. 2.1).

2.3.2 Seven conserved SOX10 consensus sequences display regulatory activity in Schwann cells

Using our computational pipeline, we identified 57 regions that harbor conserved head-to-head SOX10 consensus sequences at loci with a known or predicted role in myelination. To test if these sequences are active in Schwann cells *in vitro*, a region surrounding each consensus sequence was amplified from human genomic DNA and cloned upstream of a minimal promoter directing the expression of a luciferase reporter gene. The regulatory activity of each genomic

segment was tested in cultured rat Schwann (S16) cells (Goda et al., 1991; Toda et al., 1994), which express endogenous SOX10 (Hodonsky et al., 2012). The luciferase expression directed by each genomic segment was determined in luciferase activity assays compared to a control vector with no genomic insert ('Empty'). Seven of the 57 genomic segments demonstrated a greater than 2.5-fold increase in luciferase activity compared to the empty vector in S16 cells (Fig. 2.2): SOX10-CCS-01 (3.7-fold increase; maps to *PAX7*), SOX10-CCS-13 (54-fold increase; maps to *SOX6*), SOX10-CCS-18 (82-fold increase; maps to *SOX5*), SOX10-CCS-19 (49-fold increase; maps to *SOX5*), SOX10-CCS-39 (5.9-fold increase; maps to *TCF7L2*), SOX10-CCS-43 (25-fold increase; maps to *BCAS3*), and SOX10-CCS-51 (2.6-fold increase; maps to *NFIB*). These data suggest that these seven genomic sequences (Table 1) are potential SOX10 response elements.

2.3.3 The SOX10 consensus sequence is required for the orientation-independent activity of three regulatory elements at *SOX5*, *SOX6*, and *NFIB*

To determine if the regulatory activity of the seven genomic segments is dependent on the orientation of the DNA sequence, we retested the activity of each segment in both the 'forward' and 'reverse' orientation relative to a construct with no genomic insert ('Empty') within our reporter gene construct in S16 cells. This revealed three genomic segments that enact a greater than 2.5-fold increase in luciferase activity in both orientations (Fig. 2.3): SOX10-CCS-13 (72-fold forward and 9-fold reverse), SOX10-CCS-19 (70-fold forward and 33-fold reverse), and SOX10-CCS-51 (4-fold forward and 9-fold reverse). To assess the specificity of these results to Schwann cells, we tested each of the seven genomic segments in both orientations in cultured

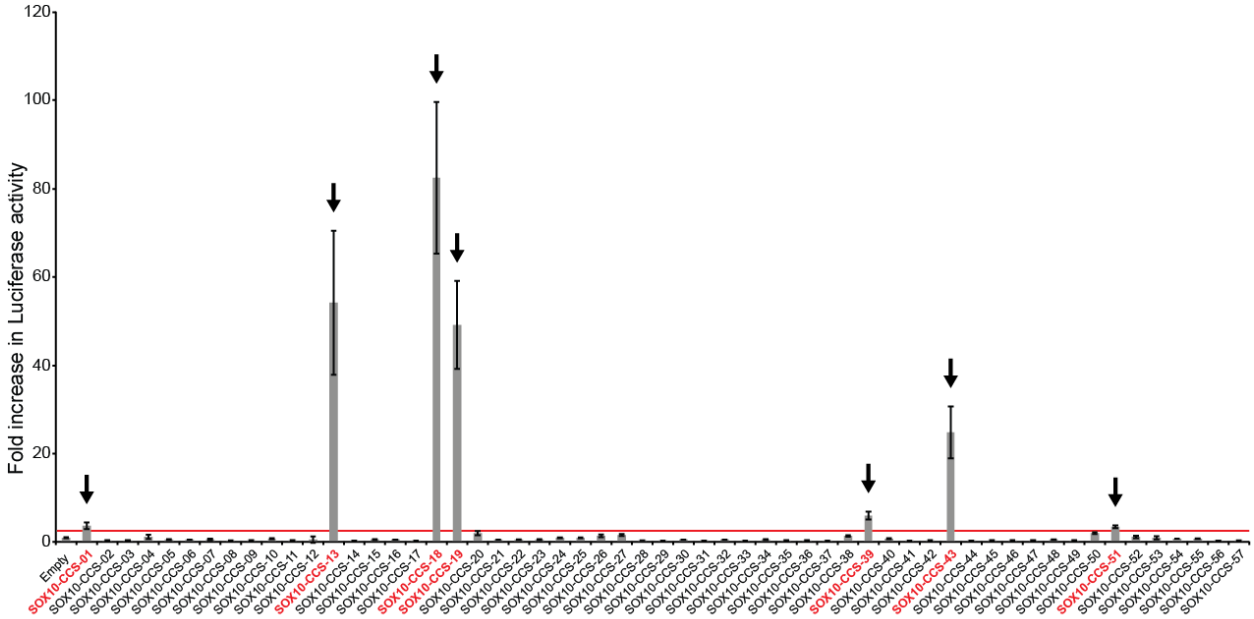


Figure 2.2. Seven regions demonstrate regulatory activity in Schwann cells. Each of the 57 genomic segments containing prioritized SOX10 consensus sequences was cloned upstream of a luciferase reporter gene and tested for enhancer activity in cultured Schwann (S16) cells. Luciferase data are expressed relative to a control vector that does not harbor a genomic insert ('Empty'). Regions that display a greater than 2.5-fold increase (red line) in luciferase activity are indicated in red text and by an arrow. Error bars indicate standard deviations.

Element ID	*Locus*	*UCSC Coordinates*[1]	*SOX10 Consensus Sequence*[2]
SOX10-CCS-01	*PAX7*	chr1:18,854,300-18,855,055	**ACAAA**CTCATTAAA**CTTGT**
SOX10-CCS-13	*SOX6*	chr11:16,334,301-16,335,278	**ACAAT**CAAGC**ATTGT**
SOX10-CCS-18	*SOX5*	chr12:24,058,988-24,059,872	**ACAAA**AATGT**ATTGT**
SOX10-CCS-19	*SOX5*	chr12:24,059,397-24,060,164	**ACACA**GAACATT**ATTGT**
SOX10-CCS-39	*TCF7L2*	chr10:114,894,980-114,895,808	**ACAAT**CCCCAAGATT**TTTGT**
SOX10-CCS-43	*BCAS3*	chr17:56,683,905-56,684,657	**ACACA**TTAATAACGT**TTTGT**
SOX10-CCS-51	*NFIB*	chr9:14,299,332-14,299,796	**ACAAT**CTGTTCTT**TGTGT**

Table 2.1. Seven genomic segments with regulatory activity in Schwann cells

[1]Coordinates refer to the March 2006 UCSC Genome Browser Human assembly (hg18).

[2]SOX10 consensus sequences are indicated in bold, underlined text.

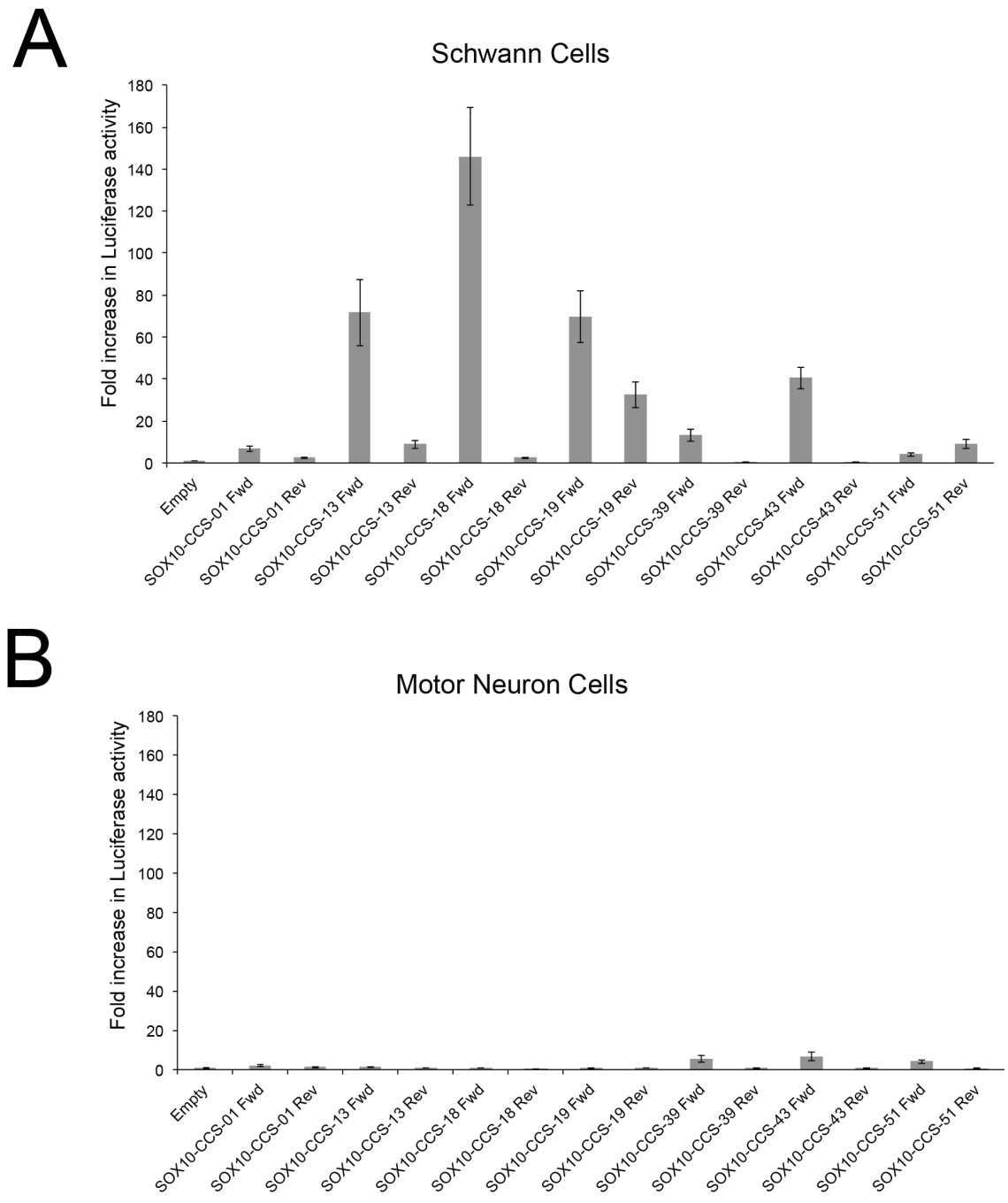


Figure 2.3. Four genomic segments are active in Schwann cells. (A) The seven active regions from Figure 2.2 were tested in forward and reverse orientation in rat Schwann (S16) cells (A) and mouse motor neuron cells (B). Luciferase data are expressed relative to a control vector without a genomic segment ('Empty'). Error bars indicate standard deviations and arrows and lines indicate genomic segments that are active in both orientations.

mouse motor neurons (MN1 cells)(Salazar-Grueso et al., 1991), which do not express endogenous SOX10 (Hodonsky et al., 2012). None of the genomic segments enact a greater than 2.5-fold increase in luciferase activity in both orientations in MN1 cells suggesting that our data in S16 cells is Schwann-cell specific; however, three had low levels of activity in only the forward orientation in MN1 cells: SOX10-CCS-39 (5.5-fold), SOX10-CCS-43 (6.7-fold), and SOX10-CCS-51 (4-fold) (Fig. 2.3).

To test the necessity of the conserved SOX10 consensus sequence for the observed activity associated with the seven genomic segments described above, we deleted the dimeric SOX10 site along with the intervening sequence in each construct (ΔSOX10) and compared the activity to the wild-type genomic segment using the more active orientation. This revealed three genomic segments that display at least a 50% reduction in activity upon deleting the SOX10 consensus sequence (Fig. 2.4): SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51. Combined, our data are consistent with these three genomic segments—at the *SOX6*, *SOX5*, and *NFIB* loci, respectively—representing Schwann cell enhancers that harbor functional SOX10 binding sites.

2.3.4 SOX10 is required for the activity of the three regulatory elements at *SOX5*, *SOX6*, and *NFIB*

To test if SOX10 induces the activity of SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51, we co-transfected each reporter gene construct with or without a construct to express wild-type SOX10 in MN1 cells, which do not express endogenous SOX10 (Hodonsky et al., 2012; Inoue et al., 2004). Subsequently, we compared the activity of each construct in the presence or absence of SOX10 expression.

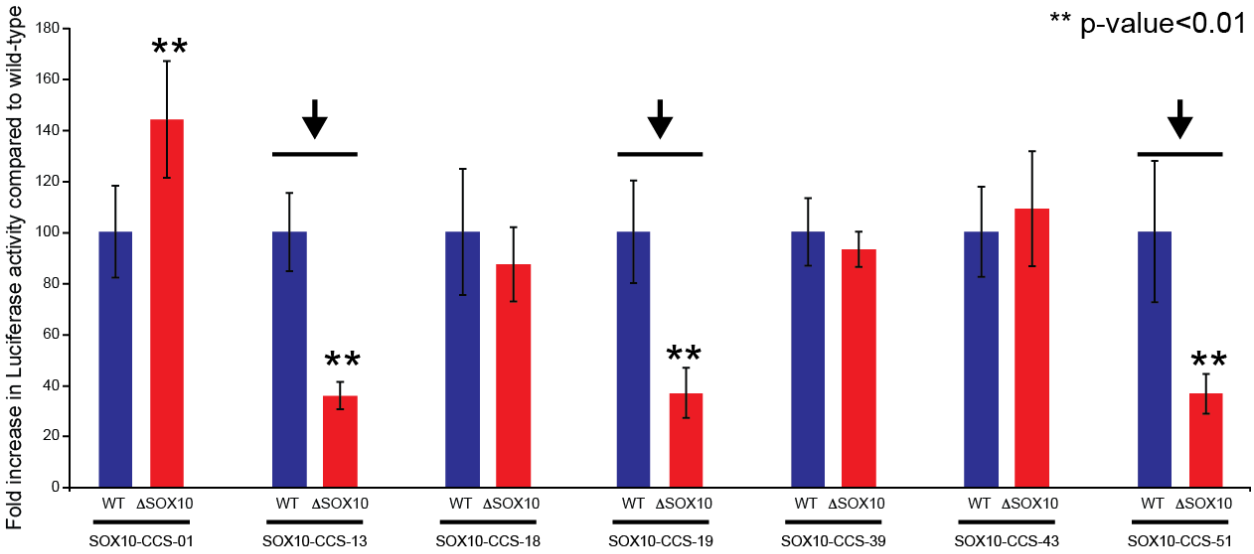


Figure 2.4. Three genomic regions require the SOX10 consensus sequence for their luciferase activity. Luciferase reporter gene constructs containing either the wild-type sequence (WT) or the sequence lacking the SOX10 consensus sequence(s) (ΔSOX10) were transfected into S16 cells and tested in luciferase assays. The luciferase activity associated with each ΔSOX10 construct is expressed relative to the respective wild-type construct. Error bars indicate standard deviations and arrows and lines indicate genomic segments with a required SOX10 consensus sequence. P-value calculated using two-tailed student t-test.

There was a ~1,000-fold increase in the activity of SOX10-CCS-13 and a ~200-fold increase in the activities of SOX10-CCS-19 and SOX10-CCS-51 in the presence of SOX10 (Fig. 2.5). SOX8, SOX9, and SOX10 belong to the SOXE family of transcription factors, which bind to nearly identical sequence motifs (Stolt and Wegner, 2010).To test if SOX10 specifically regulates SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51, we co-transfected each reporter construct with or without a construct to express SOX8 or SOX9 in MN1 cells (Fogarty et al., 2016; Hodonsky et al., 2012)and compared the effect on regulatory activity with that induced by SOX10 . In the presence of SOX8 we observed a ~140-fold, ~75-fold, and ~50 fold increase in the activity of SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51, respectively. In the presence of SOX9 we observed a ~350-fold, ~150-fold, and ~80-fold increase in the activity of SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51, respectively. Importantly, SOX8 and SOX9 did not increase the luciferase activity of these regions to the same level as SOX10, suggesting that SOX10 has a higher affinity for the sequences within SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51 (Fig 2.6).

SOX10 is known to synergistically interact with other transcription factors to enact gene expression [OCT6, BRN2, and EGR2 in Schwann cells, OLIG2 and MYRF in oligodendrocytes (Emery, 2013), and PAX3 and MITF in melanocytes (Harris et al., 2010)]. Thus, we wanted to determine if the well-characterized co-factor in Schwann cells (EGR2) works synergistically to activate these elements. EGR2, a master regulator of Schwann cell myelination, is regulated by SOX10, OCT6, and BRN2 (Ghislain and Charnay, 2006). SOX10 and EGR2 synergistically regulate key myelin genes such as *PMP22* (Jones et al., 2011), *MPZ* (Jones et al., 2007), and *GJB1* (Bondurand et al., 2001).

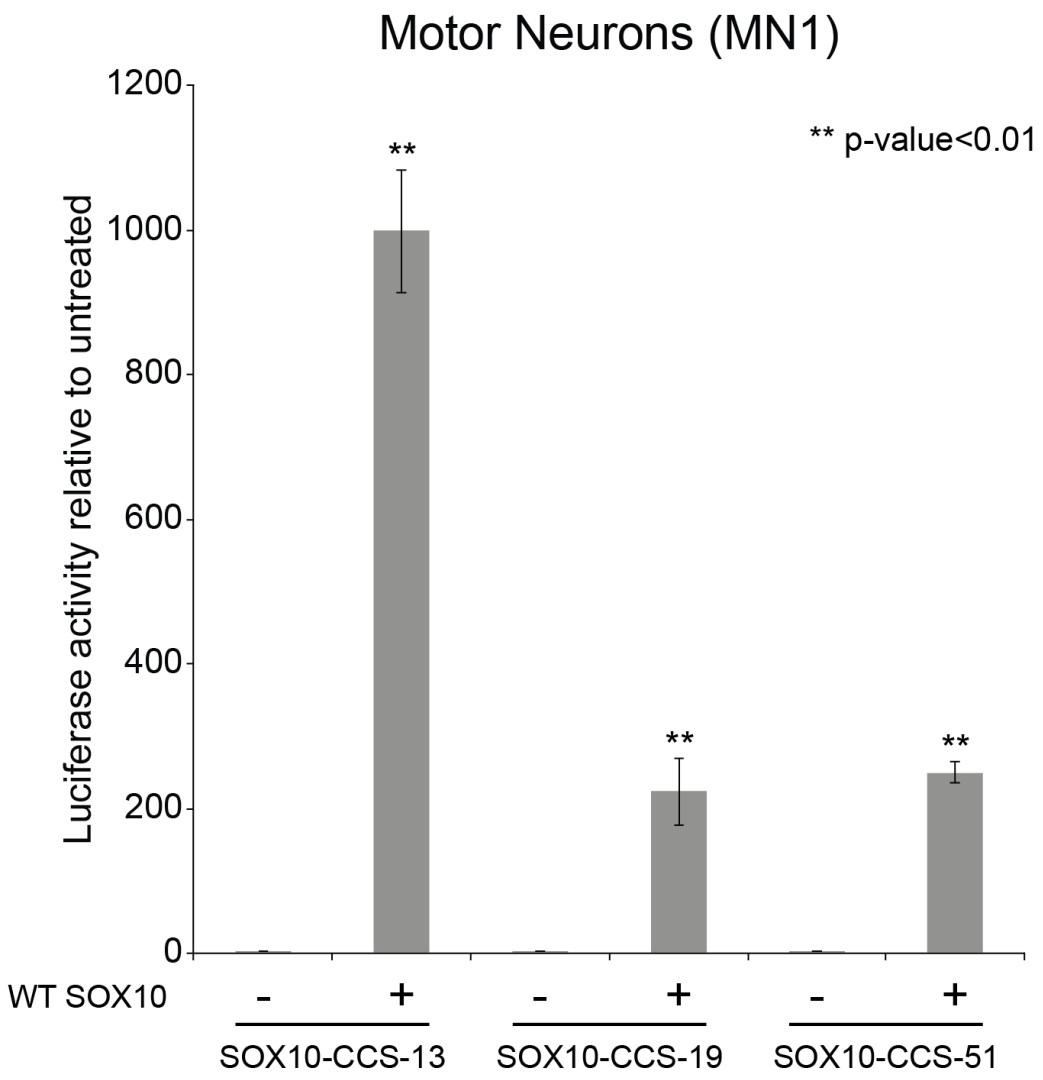


Figure 2.5. SOX10 is sufficient for the regulatory activities of SOX10-CCS-13, SOX10-CCS-19 and SOX10-CCS-51. Luciferase reporter gene constructs harboring SOX10-CCS-13, SOX10-CCS-19 or SOX10-CCS-51 were transfected into mouse motor neurons (MN1) with or without a construct to express wild-type SOX10. The luciferase activity associated with each construct in the presence of SOX10 is expressed relative to that of the construct in the absence of SOX10. P-value calculated using two-tailed student t-test.

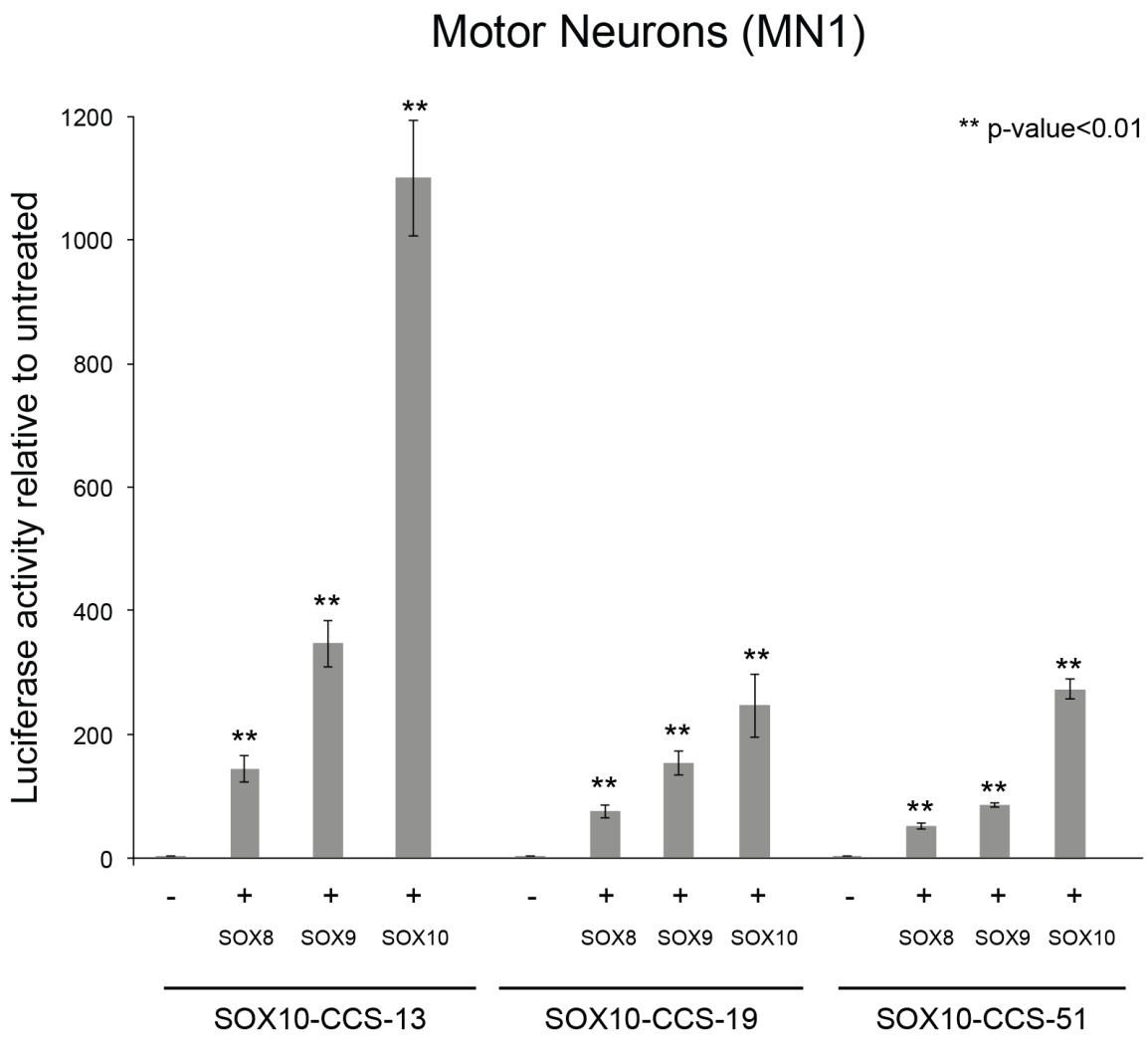


Figure 2.6. SOX8 and SOX9 also increase the regulatory activity of SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51. Luciferase reporter gene constructs harboring SOX10-CCS-13, SOX10-CCS-19, or SOX10-CCS-51 were transfected into mouse motor neurons (MN1) with constructs to express SOX8 or SOX9. The luciferase activity associated with each construct in the presence of SOX8 or SOX9 is expressed relative to that of the same construct in the absence of these transcription factors. Error bars indicate standard deviations. Please note that the SOX10 data are identical to those in Figure 2.5 and are included to facilitate a comparison. P-value calculated using two-tailed student t-test.

We co-transfected SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51 reporter constructs with a construct to express EGR2 and SOX10 in MN1 cells and compared the effect on regulatory activity with that induced by SOX10 alone. In the presence of EGR2 we observed a moderate increase in luciferase activity of SOX10-CCS-13 (~2.2-fold), SOX10-CCS-19 (~12-fold) and SOX10-CCS-51 (~10-fold) (Fig 2.7). However, in the presence of both EGR2 and SOX10 we did not see an increase in activity above that induced by SOX10 alone (even though an equivalent amount of SOX10 expression vector was transfected in each experiment). These data suggest that the three regions are primarily regulated by SOX10 and that EGR2 and SOX10 do not act synergistically upon them.

To determine if SOX10 is necessary for the activity of SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51 in Schwann cells, S16 cells were transfected with each SOX10-CCS luciferase reporter gene construct along with a construct to express a dominant-negative mutant form of SOX10 (E189X), which interferes with the function of endogenous SOX10 (Inoue et al., 2004). Importantly, E189X SOX10 has been shown to specifically reduce the activity of genomic elements. While the strategies presented here also are unable to fully capture all SOX10 binding sites, the combination of multiple datasets and methodologies generally yields a stronger segments harboring SOX10 binding sites in luciferase assays (Brewer et al., 2014). We observed a greater than 85% reduction in the activity of all three genomic segments upon co-transfection with E189X SOX10 (Fig. 2.8). Combined, our data indicate that SOX10 is required for the *in vitro* enhancer activity of SOX10-CCS-13, SOX10-CCS-19, and SOX10-CCS-51.

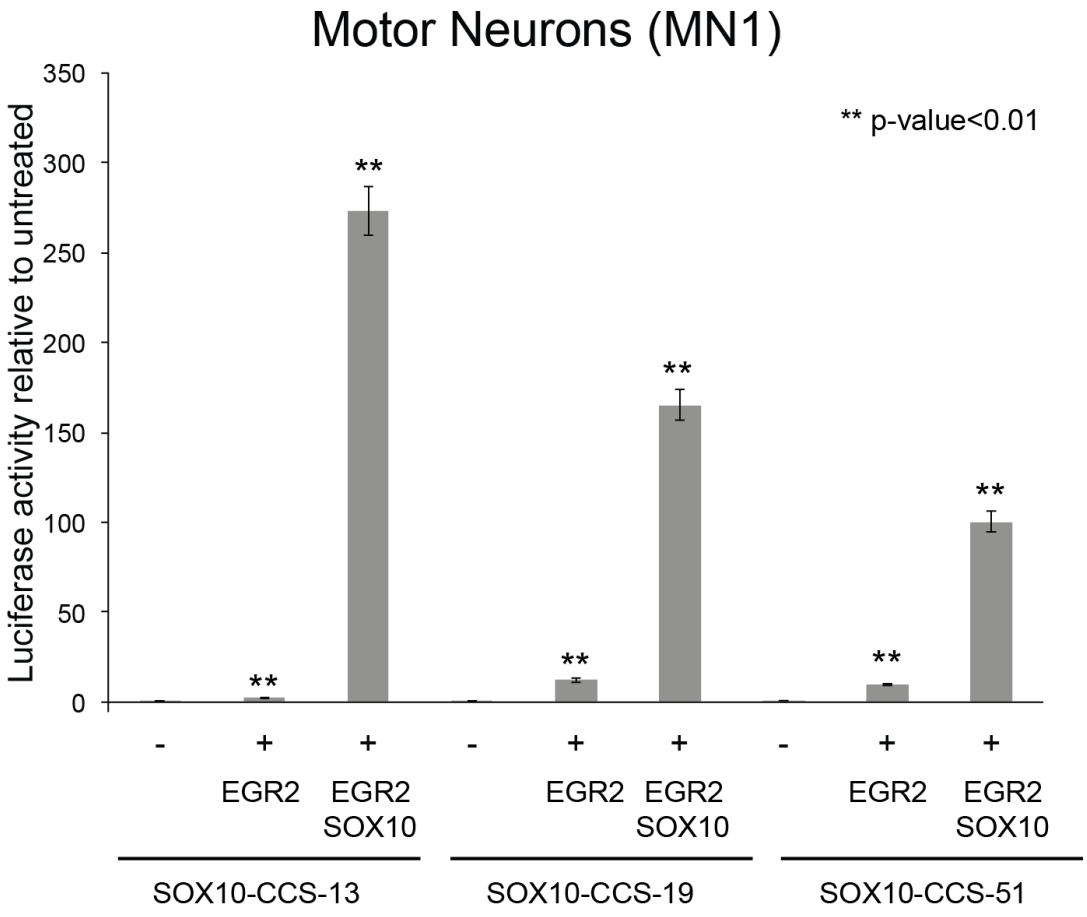


Figure 2.7. EGR2 does not act synergistically with SOX10 to activate SOX10-CCS-13, SOX10-CCS-19, or SOX10-CCS-51 in vitro. Luciferase reporter gene constructs harboring SOX10-CCS-13, SOX10-CCS-19, or SOX10-CCS-51 were transfected into mouse motor neurons (MN1) with constructs to express EGR2 and/or SOX10. The luciferase activity associated with each construct in the presence of the transcription factor(s) is expressed relative to that of the untreated reporter construct. Error bars indicate standard deviations. Please note that the SOX10 data are identical to those in Figure 2.5 and are included to facilitate a comparison. P-value calculated using two-tailed student t-test

2.3.5 SOX10 regulates the expression of genes that inhibit myelination

Our stringent computational and functional analyses rapidly identified a previously unreported SOX10-responsive promoter at the *SOX6* locus. Importantly, this finding was facilitated by the knowledge of a well-defined SOX10 consensus sequence and reports that SOX10 binding sites can be conserved among vertebrate species including human and chicken (Antonellis et al., 2008; Gokey et al., 2012; Hodonsky et al., 2012). To determine if our conservation analysis combined with whole genome datasets can reveal a set of high-confidence SOX10 response elements for further study we: (**1**) utilized available SOX10 ChIP-seq data generated from rat Schwann cell nuclei *in vivo* (Srinivasan et al., 2012); (**2**) performed DNase-seq on cultured rat Schwann (S16) cell nuclei; and (**3**) identified 67,482 non-coding SOX10 monomeric consensus sequences conserved between human, mouse, and chicken (data not shown), and converted them to the rat genome [rn5; 61,133 (90.5%) were successfully converted]. Intersecting these three data sets revealed 214 rat genomic segments that harbor conserved SOX10 consensus sequences and that map to SOX10 ChIP-seq and DNase-seq peaks (these genomic segments were computationally extracted as SOX10 ChIP-seq peaks). To determine if this approach identified specific biological pathways, we extracted the name of the rat RefSeq gene closest to each region—the 214 genomic segments map to 191 known genes —and performed a gene ontology search using the overrepresentation test for biological processes (geneontology.org). This analysis revealed 183 biological processes with a p-value less than 0.05 and 37 biological processes that showed a greater than five-fold enrichment compared to the human genome. Ten of the identified biological processes directly relate to myelinating glia, which all resided in the top 14 enriched terms (Table 2.2).

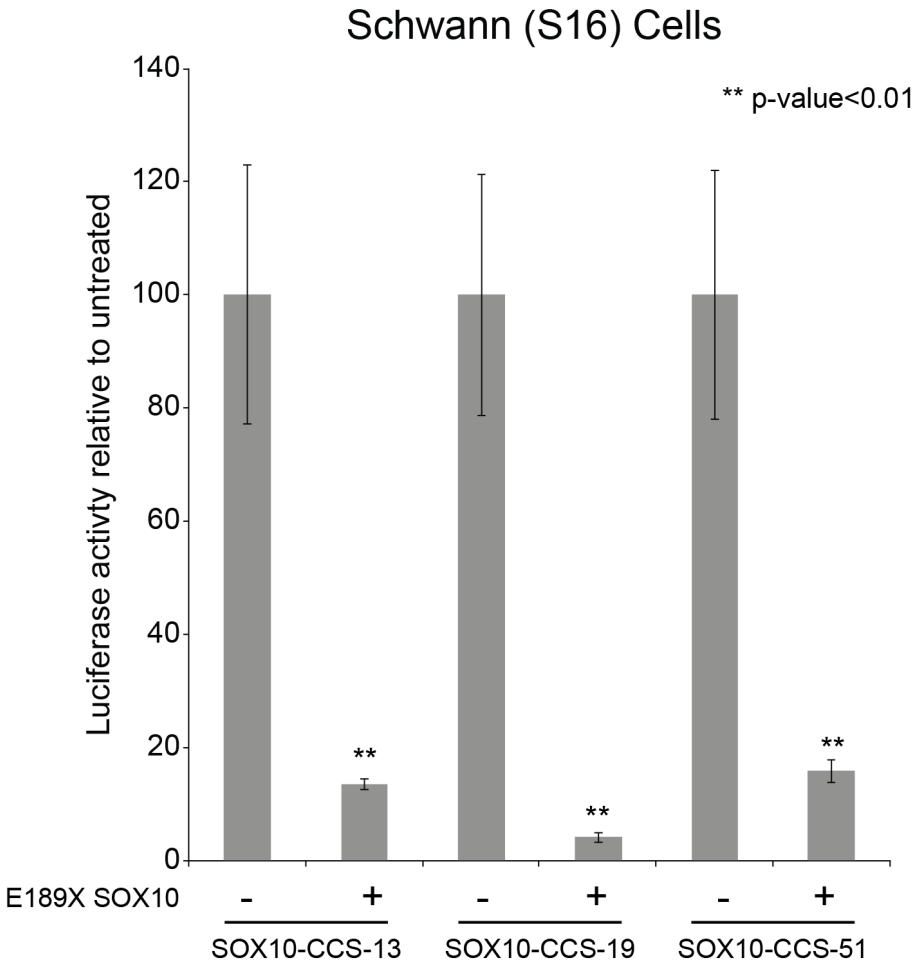


Figure 2.8. SOX10 is necessary for the luciferase activity of the three genomic regions. Luciferase reporter gene constructs harboring SOX10-CCS-13, SOX10-CCS-19 or SOX10-CCS-51 were transfected into rat Schwann (S16) cells with or without a construct to express dominant-negative (E189X) SOX10. The luciferase activity associated with each construct in the presence of E189X SOX10 is expressed relative to that of the construct in the absence of E189X SOX10. Error bars indicate standard deviations in both panels. P-value calculated using two-tailed student t-test.

Therefore, this combined strategy provided a highly confident set of 214 SOX10-response elements at 191 loci for future functional studies aimed at better understanding the biological process of myelination.

Interestingly, three of the 10 gene ontology biological processes that relate to myelination specifically relate to negative regulation of gliogenesis, which was due to the presence of six genes: *NOTCH1, HMGA2*, *HES1*, *MYCN*, *ID4*, and *ID2* (Table 2). Computational analyses revealed eight SOX10 consensus sequences within DNase-seq and SOX10 ChIP-seq peaks at these six loci (Table 2.3). To determine if *NOTCH1*, *HMGA2*, *HES1*, *MYCN*, *ID4*, and *ID2* harbor *bona fide* SOX10 response elements, we amplified genomic regions surrounding the SOX10 consensus sequences using rat genomic DNA and cloned each genomic segment (in both the 'forward' and 'reverse' orientation) upstream of a minimal promoter directing luciferase expression. The regulatory activity of each genomic segment was tested in S16 cells as described above. This revealed five genomic segments ('regions' or 'R') that directed reporter gene activity at least 2.5-fold higher than the empty control vector in both orientations: *Notch1-R1* (4.7-fold forward and 56-fold reverse), *Hmga2-R2* (93.7-fold forward and 87-fold reverse), *Hes1-R1* (22-fold forward and 7.6-fold reverse), *Mycn-R1* (28-fold forward and 16-fold reverse) and *Id2-R1* (8.9-fold forward and 4.1-fold reverse) (Fig. 2.9A). Regions *Notch1-R2* (7.6-fold) and *Id4-R1* (8.6-fold) directed reporter gene activity at least 2.5-fold higher than the empty control vector only in the forward orientation (Fig. 2.9A).

GO biological process	*Homo Sapiens*	*Our List*	*Expected*	*P-value*
Negative regulation of oligodendrocyte differentiation	12	4	0.1	2.74E-02
Negative regulation of glial cell differentiation	25	6	0.21	6.34E-04
Regulation of astrocyte differentiation	25	6	0.21	6.34E-04
Regulation of oligodendrocyte differentiation	28	5	0.23	3.29E-02
Negative regulation of gliogenesis	34	6	0.28	3.77E-03
Oligodendrocyte differentiation	60	10	0.5	9.43E-07
Regulation of gliogenesis	74	10	0.61	6.96E-06
Regulation of glial cell differentiation	54	7	0.45	3.23E-03
Glial cell differentiation	135	13	1.12	1.22E-06
Gliogenesis	168	13	1.39	1.66E-05

Table 2.2. Gene Ontology annotations for loci. Number in columns labeled Human, Our list, Expected refer to number of genes in the human genome associated with the GO term, number of genes from our list associated with the GO term, and number of genes expected by chance. The gene name in the column labeled ‘Loci ‘are those from our list

Hmga2-R1 was not active in either orientation and was excluded from further analysis. Thus, we identified seven genomic sequences at six loci (*NOTCH1*, *HMGA2*, *HES1*, *MYCN*, *ID4*, and *ID2*) that display regulatory activity in Schwann cells. To determine if the identified SOX10 consensus sequences (Table 3) are important for the regulatory activity of the seven active regions described above (Fig. 2.9A) we deleted the SOX10 consensus sequence from each construct (termed 'ΔSOX10' in Fig. 2.9B) and compared the activity to the wild-type construct using the more active orientation. *Notch1-R1*, *Notch1-R2*, *Hmga2-R2*, *Hes1-R1*, and *Id2-R1* contain dimeric SOX10 consensus sequences, which were deleted along with the intervening sequence. *Mycn-R1* contains a monomeric consensus sequence (ΔSOX10-1) and a dimeric consensus sequence (ΔSOX10-2), which were independently deleted. *Id4-R1* contains a dimeric consensus sequence with a 20 base-pair intervening sequence. Since this intervening sequence is longer than those previously observed for validated dimeric SOX10 binding sites (Antonellis et al., 2008; Brewer et al., 2014; Gokey et al., 2012; Hodonsky et al., 2012; Jones et al., 2012; Peirano and Wegner, 2000) we studied each monomer independently. Specifically, we deleted the dimeric consensus sequence along with intervening sequence (ΔSOX10-1), the first monomer only (ΔSOX10-2), and the second monomer only (ΔSOX10-3). Deleting the SOX10 consensus sequences in regions *Notch1-R1*, *Hmga2-R2*, *Mycn-R1* (ΔSOX10-2), *Id4-R1* (ΔSOX10-1 and ΔSOX10-3), and *Id2-R1* reduced luciferase activity in S16 cells by at least 50% (Fig. 2.9B), indicating that the SOX10 consensus sequences in these five regions are important for their regulatory activity. In contrast, deleting the SOX10 consensus sequences in *Notch1-R2* and *Hes1-R1* did not reduce the enhancer activity associated with these genomic segments (Fig. 2.9B). We have shown that SOX10 directly regulates *SOX6* in Schwann cells (see above).

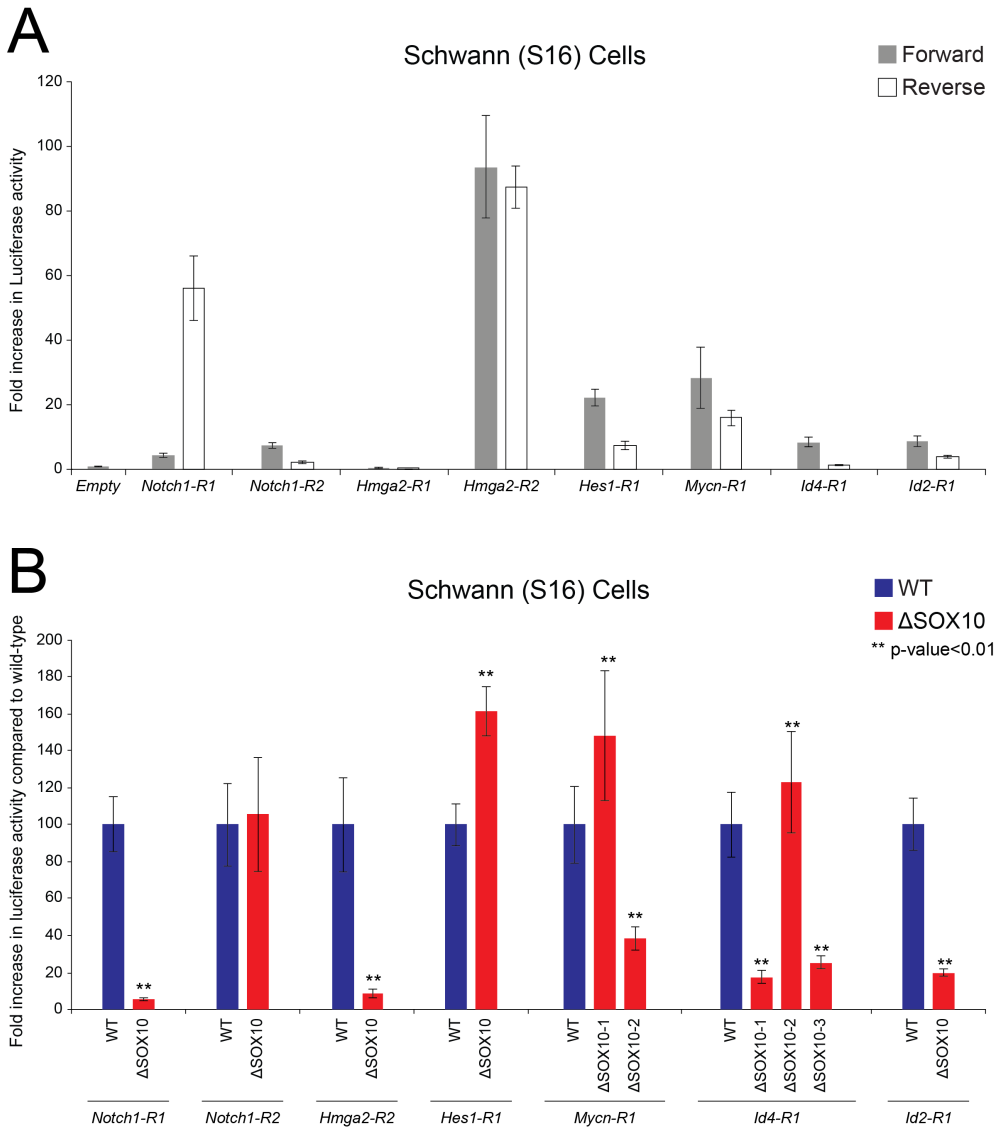


Figure 2.9 SOX10 activates the expression of inhibitors of glial cell differentiation.(A) Eight genomics segments at the rat *Notch1*, *Hmga2*, *Hes1*, *Mycn*, *Id4*, and *Id2* loci were cloned upstream of a luciferase reporter gene in both the forward and reverse orientations and tested for luciferase activity in rat Schwann (S16) cells. Luciferase data are expressed relative to a control vector with no genomic insert ('Empty'). Error bars represent standard deviations. (B) The conserved SOX10 consensus sequence(s) were deleted in each of the seven regions that were active in Fig. 6A (see text for details). Luciferase reporter gene constructs containing the wild-type sequence (WT) or the sequence lacking the SOX10 consensus sequence(s) (ΔSOX10) were transfected into S16 cells and luciferase assays performed. Luciferase activities are expressed relative to the wild-type expression constructs and error bars represent standard deviations.

Element ID	*UCSC Coordinates*[1]	*SOX10 Consensus Sequence*[2]
Notch1-R1	chr3:9,307,836-9,308,296	**ACAAT**GGGGCC**TCTGT**
Notch1-R2	chr3:9,308,175-9,309,096	**ACAAT**CGGC**TTTGT**
Hmga2-R1	chr7:65,390,088-65,391,287	CTTAG**ACACA**GCACTT
Hmga2-R2	chr7:65,427,912-65,428,606	**ACACA**GGCCCCTC**TTTGT**
Hes1-R1	chr11:77,415,315-77,415,779	**TGTGT**GAGCGCCA**TGTGT**
Mycn-R1	chr6:51,229,947-51,230,533	**ACAAT**GGCCTC**TTTCT**ACAG**ACAAT**
Id4-R1	chr17:18,701,460-18,702,118	**ACAAA**AACAGCAGTAAATGGAGGCC**TTTGT**
Id2-R1	chr6:53,090,794-53,091,254	**ACAAG**AAACAC**ATTGT**

Table 2.3. Eight genomic segments within loci that inhibit glial cell differentiation

[1]Coordinates refer to the March 2012 UCSC Genome Browser Rat assembly (rn5).

[2]SOX10 consensus sequences are indicated in bold, underlined text.

To determine if SOX10 positively regulates the expression of *Notch1*, *Hmga2*, *Hes1*, *Mycn*, *Id4*, *Id2*, and *Sox5* in cultured rat Schwann (S16) cells we again utilized the *Sox10* siRNA that has been shown to efficiently down-regulate *Sox10* expression(Gokey et al., 2012; Lopez-Anido et al., 2015). After isolation of mRNA at 24 hours post-transfection, qRT-PCR shows that *Sox10* depletion in S16 cells results in the reduced expression of all of the above genes except for *Hmga2* (Fig. 2.10A). To directly test if *NOTCH1*, *HMGA2*, *HES1*, *MYCN*, *ID4*, *ID2*, *SOX5*, and *SOX6* are developmentally regulated during myelination *in vivo* we examined mRNA levels at three timepoints in rat sciatic nerve (n=3 at each timepoint). P1 corresponds to the onset of myelination, P15 is a peak timepoint of myelination in the PNS, and adult sciatic nerve is a timepoint where active myelination has subsided. Interestingly, the expression of all seven genes tested (*Notch1*, *Hmga2*, *Hes1*, *Mycn*, *Id4*, *Id2*, *Sox5*, and *Sox6*) are highest at P1 and then repressed at P15 and adult, consistent with a role in repressing precocious myelination (Fig. 2.10B).

2.4 Discussion

Previous efforts have identified SOX10 binding sites and target genes in Schwann cells (Lopez-Anido et al., 2015; Srinivasan et al., 2012). These efforts, and others (Lee et al., 2008), have utilized a variety of experimental methodologies to identify putative SOX10 regulatory elements across diverse SOX10-positive tissues. While each approach uncovered novel SOX10 response elements, no single method has been successful in the identification of all SOX10 response elements.

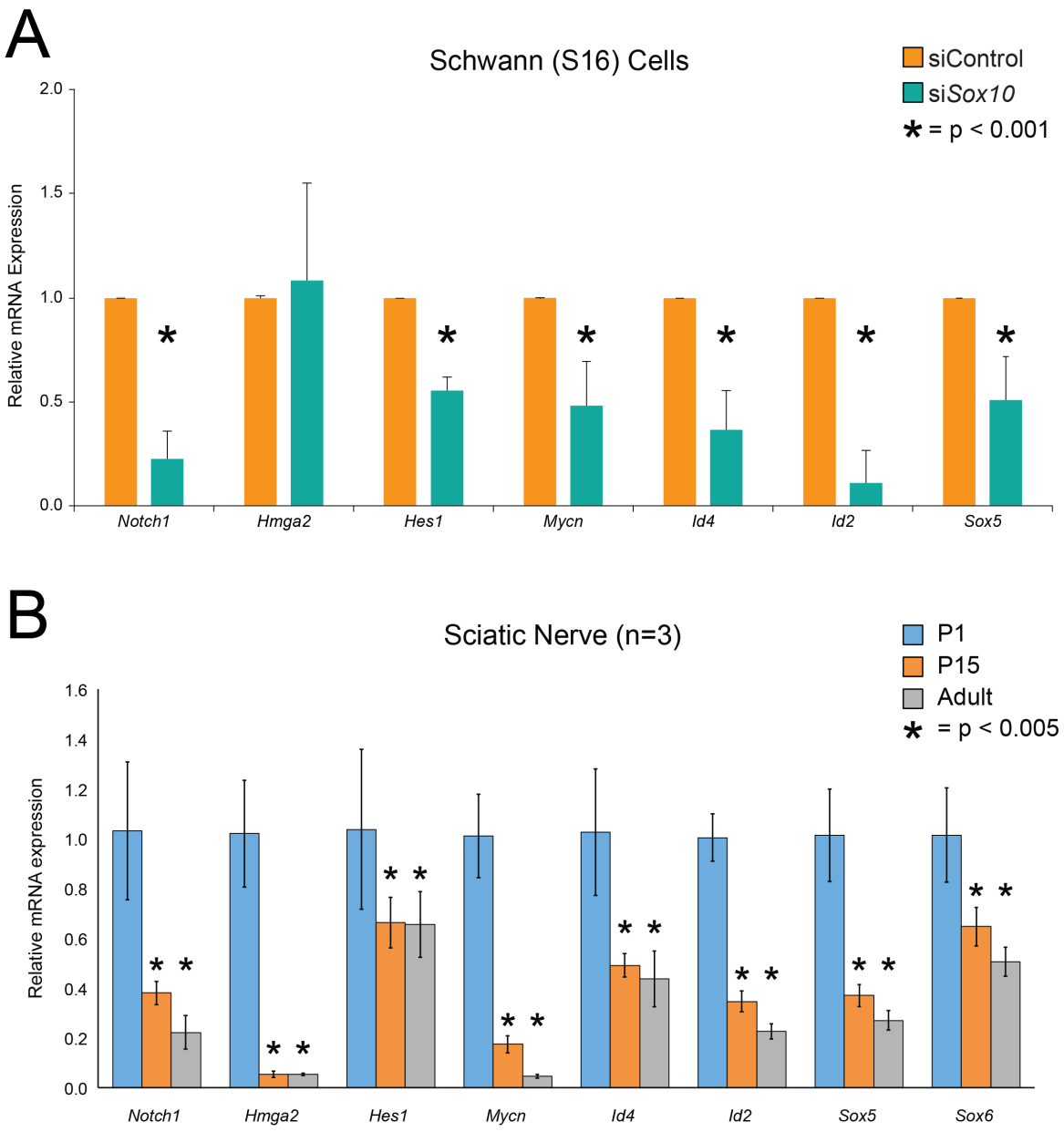


Figure 2.10. Inhibitors of glial cell differentiation are developmentally regulated (A) Rat Schwann (S16) cells were treated with a control siRNA or a siRNA targeted against Sox10. Quantitative RT-PCR was used to measure expression levels of each indicated gene. Asterisks indicate a p-value smaller than 0.001 and error bars indicate standard deviations. (B) RNA was purified from three independent rat sciatic nerves at the P1, P15, and adult timepoints. Quantitative RT-PCR was used to measure expression levels of each indicated gene with values expressed relative to expression levels at P1. Asterisks indicate a p-value smaller than 0.005 and error bars indicate standard deviations.

While the strategies presented here also are unable to fully capture all SOX10 binding sites, the combination of multiple datasets and methodologies generally yields a stronger predictive power for identifying regulatory regions compared to any one individual method (Kwasnieski et al., 2014). Indeed, when we combined our DNase-seq data with previously generated SOX10 ChIP-seq data (Srinivasan et al., 2012), we were able to quickly prioritize and validate novel SOX10 response elements near genes with a known role in myelination. While previous efforts to identify SOX10 response elements focused on the required function of SOX10 in cultured cells or tissues at specific developmental stages, our computational approach utilizes sequence conservation to identify putative SOX10 regulatory regions throughout the genome in a tissue-independent manner. The combination of our less-biased (albeit less biologically relevant) computational approach with DNase-seq and ChIP-seq is likely the reason that we were able to identify specific repressors of myelination as putative SOX10 target genes (*e.g.*, these would not have been identified in myelinating Schwann cells). As such, we feel that the datasets generated here will be useful to investigators studying comparative genomics, SOX protein function, and Schwann cell biology. First, the conserved sequences we identified could be used to similarly prioritize consensus sequences for other transcription factors important for vertebrate development. Second, the SOX10 consensus sequences we identified could be used to prioritize putative binding sites in other SOX10-positive cells including oligodendrocytes, melanocytes, and developing enteric nervous system neurons (Kelsh, 2006). Finally, our DNase-seq data from rat Schwann (S16) cells will be useful for anyone studying transcriptional regulatory elements, highly expressed genes, or any other nuclear structure characterized by open chromatin in myelinating Schwann cells; S16 cells express many myelin-related genes (*e.g.*, *PMP22*, *MPZ*, *MBP*, and *MAG*) and transcription

factors (*e.g.*, SOX10 and EGR2) and are biochemically similar to myelinating Schwann cells (Hai et al., 2002). The analyses performed here may also provide insight into the nucleotide requirements for a SOX10 response element. We functionally evaluated a total of 62 conserved, genomic segments that harbor a predicted dimeric SOX10 binding site. Interestingly, simply cloning a highly conserved dimeric SOX10 consensus sequence upstream of a minimal promoter was not enough to enact regulatory activity in cultured Schwann cells as evidenced by the 50 genomic segments with no (or very low) regulatory activity in Figure 2.1. While there are many possible explanations for this finding, one is that there are nucleotide-specific requirements for the intervening sequences (*i.e.*, the nucleotides between the head-to-head monomeric sites). To assess this, we compared the length and GC content of all 62 dimeric sites to those dimeric sites that were both active and required for the observed regulatory activity (n=7). While there was no significant difference in the average intervening sequence length between the two groups, the seven active sites all had intervening sequence lengths between five and eight nucleotides consistent with previous reports that six basepairs provides the ideal spacing between monomers (Peirano and Wegner, 2000). Interestingly, there was a marked difference in the GC content when comparing the total population of intervening sequences (GC content = 35%) to the intervening sequences in the active dimeric sites (GC content = 61%). These data are consistent with the high GC content of the intervening sequences within previously validated dimeric SOX10 binding sites(Antonellis et al., 2008; Brewer et al., 2014; Gokey et al., 2012; Hodonsky et al., 2012; Peirano et al., 2000) and with a 'G' nucleotide being the most commonly observed nucleotide after the core motif (Srinivasan et al., 2012). Thus, future predictions of dimeric SOX10 binding sites should allow for high GC content and five to eight basepairs between the head-to-head consensus sequences.

Our efforts predicted eight putative SOX10 target loci with a known role in repressing glial cell differentiation: *NOTCH1*, *HMGA2*, *HES1*, *MYCN*, *ID2*, *ID4, SOX5, and SOX6*. These findings were unexpected due to the known role of SOX10 in regulating the expression of genes that encode myelin proteins (*e.g.*, *MBP*, *MPZ*, and *PMP22*)(Jones et al., 2011; 2012; LeBlanc et al., 2006; Li et al., 2007; Peirano et al., 2000; Wei and Miskimins, 2004). We showed that all eight loci are developmentally regulated during myelination *in vivo* in a manner consistent with a role in inhibiting glial cell differentiation. We were also able to functionally validate a SOX10 binding site at seven of the eight loci. We identified a SOX10 ChIP-seq peak at *HES1* and luciferase assays demonstrated that this genomic segment has strong enhancer activity (Fig. 2.9A). However, deletion of the predicted SOX10 binding sites in *Hes1-R1* (Table 3) did not reduce luciferase activity. Further mutagenesis of this genomic segment will be required to identify sequences necessary for the observed activity, which may reveal a degenerate SOX10 consensus sequence. When we depleted SOX10 activity in Schwann cells *in vitro* and *in vivo* seven of the eight loci were down-regulated; while *HMGA2* harbors a validated SOX10 response element (Fig. 2.10A), depletion of SOX10 activity did not reduce *Hmga2* expression. Further analysis will be required to determine if the SOX10 response element at *Hmga2* regulates an adjacent locus or if depletion of SOX10 at specific developmental timepoints results in reduced *Hmga2* expression. Consistent with our findings, previous global analyses of SOX10 function revealed that two of the above eight loci are downstream of SOX10: *Id2* and *Notch1*(Srinivasan et al., 2012); our analysis now localizes at least some of the SOX10-dependent enhancers responsible for the regulation of these two loci.

SOX5 and SOX6 are members of the SOXD family of transcription factors and act as negative regulators of myelination in the central nervous system (Stolt et al., 2006); these proteins, which do not have transactivation or transrepression domains (Lefebvre, 2010), inhibit the expression of SOX10 target genes (*e.g.*, *MBP*) in oligodendrocytes by competing with SOX10 for DNA binding at sites within cis-acting regulatory elements. To allow oligodendrocyte differentiation and myelin production, *SOX6* mRNA is targeted for degradation by two microRNAs (miR) in these cells: miR-219 and miR-338 (Zhao et al., 2010). It was recently reported that SOX13 (the third and final member of the SOXD subgroup) also has an antagonistic effect on the ability of SOX10 to activate the expression of myelin genes in the central nervous system (Baroti et al., 2015). Indeed, *SOX13* is among the group of 191 loci at which we identified a highly confident SOX10 binding site: a single, conserved genomic segment within SOX10 ChIP-seq and DNase-seq peaks ~62 kb upstream of *Sox13* (rn5 coordinates chr13:55425486-55425636) was identified. Interestingly, a relationship between SOXD and SOXE (SOX8, SOX9, and SOX10) transcription factors has been proposed since ablation of SOX8 or SOX9 (but not SOX10) reduces *Sox6*, but not *Sox5*, expression in the developing spinal cord (Stolt et al., 2006).

In addition to genes encoding SOXD proteins, our studies predict that *NOTCH1*, *HES1*, *MYCN*, *ID2*, and *ID4* are SOX10 target genes. NOTCH1 is a transmembrane receptor that regulates Schwann cell proliferation and inhibits Schwann cell differentiation in perinatal nerves, and facilitates dedifferentiation of Schwann cells after nerve injury (Woodhoo et al., 2009). HES1 is an effector of NOTCH signaling, acts as a transcriptional repressor (Jarriault et al., 1995; Sasai et al., 1992), and is highly expressed during early stages of Schwann cell development (Woodhoo et al., 2009). In cultured mouse oligodendrocytes, HES1 maintains cells in an immature state and

overexpression of HES1 results in reduced expression of myelin related genes (*Mbp* and *Plp*) (Ogata et al., 2011). MYCN is a proto-oncogene and is known to inhibit astrocyte differentiation from neural precursor cells (Sanosaka et al., 2008); however, the role of MYCN during Schwann cell myelination has not been studied. Inhibitors of differentiation 2 and 4 (ID2 and ID4) are known to inhibit oligodendrocyte differentiation and the lack of both proteins results in premature oligodendrocyte differentiation (Kondo and Raff, 2000; Marin-Husstege et al., 2006; Wang et al., 2001). Furthermore, *Id2* and *Id4* expression declines in Schwann cell development and ID2 limits induction of *myelin protein zero* expression in primary Schwann cells(Mager et al., 2008; Stewart et al., 1997). Consistent with our findings, RNA-seq of oligodendrocytes isolated at various stages of mouse brain development (Zhang et al., 2014) show that *Sox5*, *Sox6*, *Notch1*, *Hes1*, *Mycn*, *Id2*, and *Id4* are developmentally regulated in the central nervous system—*Hmga2* does not appear to be expressed in the cells assessed in that study. Therefore, these genes likely play a role in preventing premature glial cell differentiation in both the central and peripheral nervous systems.

Combined with previous findings, our data predict a model (Figure 2.11) where SOX10 activates the expression of genes that inhibit Schwann cell differentiation, possibly during early stages of Schwann cell development, thus preventing the precocious expression of myelin proteins. Subsequently, EGR2, NAB, and microRNAs are known to inhibit the expression of the negative regulators of myelination (*e.g.*, SOXD proteins), which would allow the expression of myelin proteins. In addition to the data presented in this study, previous reports support specific aspects of this model. For example, EGR2 likely represses the expression of many of the eight loci reported here. EGR2 and NAB repress *ID2* and *ID4* before myelination via NAB binding to

CHD4 (Mager et al., 2008) [conditional ablation of CHD4 in Schwann cells causes increased expression of immature Schwann cell genes including *ID2* and delayed myelination, radial sorting defects, hypomyelination, and the persistence of promyelinating Schwann cells in conditional knockout mice (Hung et al., 2012)]. Furthermore, a comparison of SOX10 and EGR2 binding with expression profiles in Schwann cells treated with siRNA for SOX10 and EGR2-deficent peripheral nerves (Srinivasan et al., 2012) revealed that *NOTCH1* and *ID2* are SOX10-activated and EGR2-repressed, and *ID2*, *HMGA2*, *SOX5*, and *ID4* remain high in peripheral nerves from *Egr2- or Nab-deficient* mice(Le et al., 2005; Mager et al., 2008). Finally, SOX10 directly regulates the expression of *EGR2* (Ghislain and Charnay, 2006) and miR-338 (Gokey et al., 2012). In sum, our findings suggest that SOX10 has a role in maintaining a premyelinating state during non-myelinating stages of Schwann cell development.

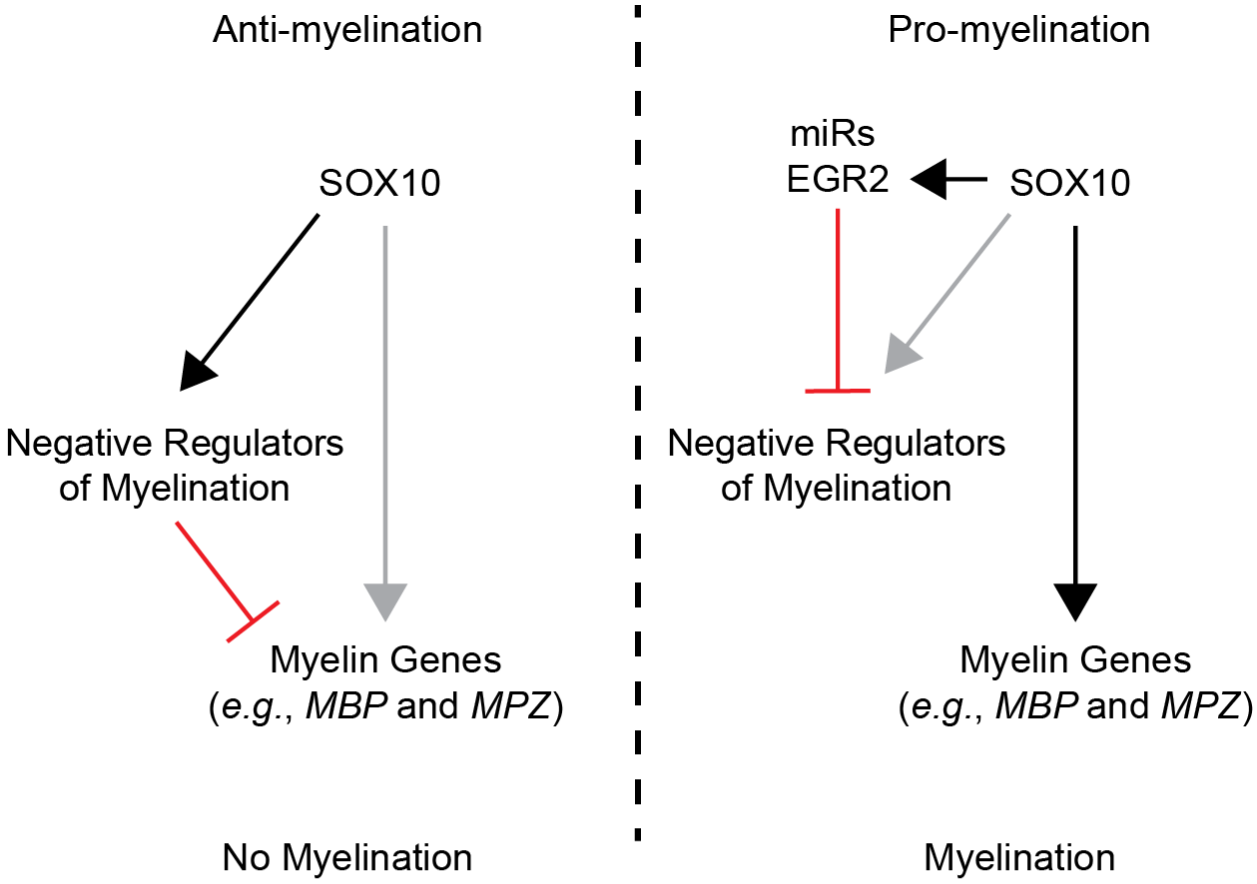


Figure 2.11. A model for the role of SOX10 in maintaining a pre-myelinating state. Previous to myelination (anti-myelination; left side), SOX10 activates the expression of negative regulators of myelination, which inhibit the expression of myelin genes such as MBP and MPZ. During activation of the myelination program (pro-myelination; right side), EGR2 and micro RNAs (miRs) inhibit the expression of negative regulators of myelination, which allows SOX10 (and EGR2) to positively regulate the expression of myelin genes.

Chapter 3

Exploring the Role of SOX6 in Peripheral Nerve Myelination

3.1 Introduction

The myelin sheath accelerates the propagation of electric impulses along the nerves of both the peripheral and central nervous systems. In the peripheral nervous system (PNS), this function is mediated by Schwann cells, which wrap their membranes around axonal segments to form compact myelin. Schwann cells originate from the neural crest cells, which are a multipotent, transient cell population that differentiate into various cell types including melanocytes, sensory neurons, enteric nervous system ganglia. Migratory neural crest cells associate with individual axons and differentiate into Schwann cell precursors. Schwann cell precursors differentiate into an intermediate immature state before transitioning into myelinating and non-myelinating Schwann cells. Schwann cells are remarkably plastic and can demyelinate and dedifferentiate after nerve trauma (Jessen and Mirsky, 2005).

Precise control of the transcriptional programs important for Schwann cell development ensures the correct timing of myelination in the peripheral nerve. Dysregulation of these pathways can

lead to hypomyelination or hypermyelination and are known to be associated with pathology of demyelinating diseases (Noseda et al., 2013; Svaren and Meijer, 2008; Triolo et al., 2006). SOX10, a transcription factor critical for every stage of Schwann cell development, is known to direct the expression of positive regulators of myelination, including OCT6, EGR2/KROX20, MBP, MPZ (Britsch et al., 2001; Kuhlbrodt et al., 1998; Svaren and Meijer, 2008). Conditional deletion of *Sox10* in mice at any stage of Schwann cell development causes severe demyelination (Bremer et al., 2011; Britsch et al., 2001; Finzsch et al., 2010). Published data on SOX10's role in Schwann cells present a paradoxical observation that SOX10 is expressed throughout the Schwann cell lineage but that pro-myelinating SOX10 target genes are not expressed during early periods of development. Negative regulators of myelination inhibit the expression of key myelin genes during the early stages of Schwann cell development and direct the dedifferentiation program following nerve injury. The transcriptional regulation of negative regulators of myelination is less understood and the mechanisms by which they suppress myelination are not well defined. Interestingly, the data presented in chapter 2 provides evidence for a previously underappreciated role for SOX10 in activating the expression of myelin inhibitors such as *SOX5*, *SOX6*, *NOTCH1*, *HMGA2*, *HES1*, *MYCN*, *ID4*, and *ID2* (Gopinath et al., 2016). On further examining the SOX10 responsive enhancers at each of these loci we found that SOX10-CCS-13 within the *SOX6* locus is directly upstream of a non-coding exon and we hypothesized that the SOX10-responsive intronic enhancer may be acting as an alternative promoter. To address this, I focused on characterizing the intronic SOX10 response element at *SOX6* and on defining the role of SOX6 in Schwann cells. In oligodendrocytes, SOX6 has been reported to compete with SOX10 for binding at *MBP* thereby inhibiting premature myelination in the central nervous system (Stolt et al., 2006). Specifically, in *Sox6* null mice *Mbp* was

prematurely expressed causing precocious myelination (Stolt et al., 2006). The function of SOX6 has not yet been characterized in Schwann cells and, to date, *MBP* is the only known target of SOX6; however, only *in vitro* studies have been performed to show that SOX6 directly represses *MBP* expression. Identifying additional SOX6 target genes will give us novel insights into the regulatory pathways that control peripheral nerve myelination.

In this chapter, I provide evidence that the intronic SOX10 response element at *SOX6* (SOX10-CCS-13) resides directly upstream of a previously unreported, promoter region. To understand the role of SOX6 in Schwann cells we performed RNA sequencing analysis after overexpressing GFP-tagged SOX6 in an *in vitro* cell culture model of myelinating Schwann cells. Comprehensive analysis of differentially expressed genes revealed a putative role for SOX6 in regulating Schwann cell proliferation during development and Schwann cell dedifferentiation after peripheral nerve injury.

The author performed all the experiments presented in this chapter except the Svaren laboratory performed the quantitative RT-PCR. For the RNA-seq experiment, flow cytometry was performed at the University of Michigan Flow Cytometry core. cDNA library preparation and sequencing was performed by the University of Michigan Sequencing core.

3.2 Methods

3.2.1 Standard and quantitative RT-PCR

$1x10^5$ cells were plated in a 6-well plate (USA Scientific cat # CC7682-7506) and incubated overnight at 37°C in 5% CO_2. Transfections were performed using the following protocol for

each well: 1mL Opti-MEM (Thermo Fisher Scientific cat # 31985062) was mixed with 10μL Lipofectamine 2000 and this mixture was incubated for 10 minutes at room temperature (Mix 1). 4μg of wild-type or E189X SOX10 (Inoue et al., 2004) 1mL of Opti-MEM (Mix 2). After the 10-minute incubation 1mL of Mix 1 was added to Mix 2, mixed well, and incubated at room temperature for 20 minutes. The cells were washed with 2mL 1XPBS and 2mL of transfection mixture was added to each well. Mock transfections were performed in the absence of DNA. Cells were incubated at 37°C for 4 hours and then grown in 3mL standard growth media. After 72 hours, total RNA was isolated from the transfected cells using the RNeasy kit (Qiagen cat # 74104). Subsequently, cDNA was synthesized using 1μg of total RNA and the High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific 4368814). The following protocol was used for cDNA synthesis: 3.2μL ultrapure water, 2μL 10X random primers, 2μL 10X RT buffer, 0.8μL 25X dNTPs, 1μL RNase inhibitor, 1μL reverse transcriptase, and 1μg RNA in 10μL ultrapure water. The cycler conditions are as follows: 25°C for 10 minutes, 37°C for two hours, and 85°C for five minutes. For standard RT-PCR was performed on isolated cDNA using gene specific primers. A PCR for *β-actin* served as a positive control. All PCR products were subjected to DNA sequencing to confirm specificity. The Svaren laboratory at the University of Wisconsin, Madison, performed the qRT-PCR assay.

3.2.2 5' Rapid amplification of cDNA ends (RACE)

5'RACE (Thermo Fischer Scientific cat # 18374058) was performed using the manufacturer's protocol. First strand cDNA libraries were synthesized using total RNA isolated from S16 cells and a primer designed within exon 5 (GSP1) of *Sox6*. To generate gene-specific cDNA library the following protocol was used: 1.25μL rnSOX6 GSP1 (2 μM), 6.3μL (~5 μg) S16 RNA, and

7.95μL ultrapure water were added and the reaction was incubated at 70°C for 10 minutes, chilled on ice for one minute, briefly centrifuged, and the following components were added (in order): 2.5μL 10X PCR buffer, 2.5μL 25mM $MgCl_2$, 1μL 10mM dNTP mix, and 2.5μL 0.1 M DTT. The reaction was gently mixed, centrifuged, incubated for one minute at 42°C, and 1μL of SuperScript II RT was added. The reaction was then incubated for 50 minutes at 42°C, and terminated by incubating at 70°C for 15 minutes. The mixture was briefly centrifuged and placed at 37°C for 30 minutes. The cDNA was then purified using S.N.A.P. column purification. First, 120μL of binding solution (6M NaI) was added to the cDNA mixture (see above). Next, the entire mixture was transferred to a S.N.A.P. column, centrifuged at 13,000g for 20 seconds, and the flow through was saved until the recovery of the cDNA was ensured. The cDNA was washed by adding 0.4 mL of 1X wash buffer to the column, centrifuged at 13,000g for 20 seconds, and the flow through was discarded. The wash step was repeated three additional times. After washing, 400μL of 70% ethanol was added, centrifuged at 13,000g for 20 seconds, and the flow through was discarded. The ethanol wash step was repeated one additional time (two ethanol washes in total). Finally, the column was transferred to a clean recovery tube, 50μL of 65°C ultrapure water was added, and the column was centrifuged at 13,000g for 20 seconds. The cDNA was TdT-tailed by adding 6.5μL ultrapure water, 5μL 5X tailing buffer, 2.5μL 2mM dCTP, and 10μL S.N.A.P.-purified cDNA. No additional quantification was performed on the S.N.A.P.-purified cDNA prior to the TdT-tailing reaction. The reaction was incubated for three minutes at 94°C, chilled on ice for one minute, and briefly centrifuged. Next, 1μL TDT was added, gently mixed, and incubated at 37°C for 10 minutes. The reaction was incubated at 65°C for 10 minutes.

Two nested PCRs were performed using a reverse primer designed within exon 4 (GSP2) of *Sox6*: 31.5µL ultrapure water, 5µL 10X PCR buffer, 3µL 25mM $MgCl_2$, 1µL 10mM dNTPmix, 2µL 10µM rnSOX6 GSP2, 2µL Abridged anchor primer, 5µL dC-tailed cDNA, and 0.5µL Taq polymerase (NEB Cat no. M0273S). A second nested PCR was performed using a reverse primer designed within exon 3 (GSP3) of *Sox6*: 33.5µL ultrapure water, 5µL 10X PCR buffer, 3µL 25mM $MgCl_2$, 1µL 10mM dNTP mix, 1µL 10µM rnSOX6 GSP3, 1µL 10µM AUAP primer, 5µL of PCR product from first nested reaction, and 0.5µL Taq polymerase (NEB cat # M0273S). The nested PCR products were separated on a 1% agarose gel, the bands of the right sizes were gel excised, and purified using the QIAquick gel extraction kit (Qiagen Cat no. 28704). Gel purified PCR products were TA cloned (Invitrogen Cat no. 450071): 1µL 5' RACE PCR product, 1µL salt solution (1.2 M $NaCl_2$ and 0.06 M $MgCl_2$), 1µL (10ng) pCR4 TOPO vector, and 3µL ultrapure water. The resulting plasmids were transformed into *E. coli*, plated on kanamycin selective plates, colonies were picked and grown overnight at 37°C. Plasmid DNA was isolated using Qiagen miniprep kit. 48 clones were subjected to Sanger sequencing.

3.2.3 Cloning SOX6 isoforms

Primers containing attB1 and attB2 Gateway cloning (Thermo Fisher Scientific) sequences were designed to amplify the mouse SOX6 open reading frame (ORF) and were synthesized by IDT. The primers were resuspended using ultrapure water (Thermo Fisher Scientific cat # 10977023) to make 200µM stock solutions and diluted 1:10 prior to PCR. Adult mouse brain cDNA was used as template DNA. The PCR product was cloned into pDONR221 using BP clonase as previously described except the reaction was incubated at room temperature overnight, 1µL Proteinase K was added the next day and the reaction was incubated at 37°C for ten minutes and

transformed into Top10 One shot *E.coli* cells as described in chapter 2. Eight individual colonies were picked and grown in 7mL of kanamycin selective media. Plasmid DNA was isolated using the Qiagen miniprep kit and was subjected to BsrG1 digestion to ensure presence of the insert as described in chapter 2. All the eight plasmids were subjected to Sanger sequencing and the PCR product contained two isoforms of SOX6: SOX6-full length (SOX6-FL), which contains exon 9 and SOX6-ΔE9 among the eight plasmids. SOX6-FL and SOX6-ΔE9 were independently cloned into pDEST53 (Thermo Fischer Scientific), which contains an N-terminal GFP tag using LR clonase. The following protocol was used: 1µL pDEST53 plasmid, 1µL pDONR221 construct (150ng/ul), 6µL TE buffer, and 1µL LR clonase. The reaction was incubated at room temperature for one hour. To terminate the reaction, 1µL Proteinase K was added and the reaction was incubated at 37°C for ten minutes. 3µL of above reaction was mixed with 12.5µL Top10 One shot *E.coli* cells (Thermo Fisher Scientific cat # C404003) and incubated on ice for 15 minutes. The bacteria were heat shocked at 42°C for 45 seconds and were recovered by adding 62.5µL SOC growth media (Thermo Fisher Scientific cat # C404003) and incubating at 37°C with shaking. The transformation reaction was plated on 100mg/mL ampicillin selective plates. The plates were incubated overnight at 37°C. Individual colonies were picked and grown in 7mL of ampicillin selective media. Plasmid DNA was isolated using the Qiagen miniprep kit and was subjected to BsrG1 digestion. Two independent plasmids for each isoform were submitted for Sanger sequencing to ensure that GFP tag was in-frame with SOX6 ORFs. Both the isoforms were cloned in a similar manner into a vector that contains an N-terminal FLAG tag, pEZY-FLAG.

3.2.4 Cell culture and transfections

HeLa cells were maintained in Dulbecco's modified Eagle's medium (Thermo Fisher Scientific cat # ILT12430054) supplemented with 10% fetal bovine serum (Thermo Fisher Scientific cat # 26140-079), 2mM L-glutamine (Thermo Fisher Scientific cat # ILT25030081), and 1X Penicillin-Streptomycin (Thermo Fisher Scientific cat # ILT15070063) at 37°C in 5% CO_2. $1x10^5$ cells were plated in a 6-well plate (USA Scientific cat # CC7682-7506) and incubated overnight at 37°C in 5% CO_2. To isolate total RNA and whole cell lysates, HeLa cells were transfected with molar equivalents of GFP-SOX6-FL (4.05µg) and GFP-SOX6-ΔE9 (4µg) using Lipofectamine 2000 as described above. 48 hours post transfection total RNA and whole cell lysates (WCL) were isolated independently. Total RNA was isolated as described above. After RNA isolation, the samples were treated with DNase to remove any residual plasmid DNA contamination. To DNase treat the RNA we used the Qiagen RNase-Free DNase set (Qiagen cat # 79254). The following protocol was used: 10µL RNA, 10µL RDD buffer, 2.5µL DNase, and 77.5µL ultrapure water. The reaction was incubated at room temperature for 10 minutes. Following DNase treatment, the RNA samples were cleaned up using RNeasy kit. cDNA was prepared as described above with (RT +) and without (RT -) reverse transcriptase enzyme. Image J was used to quantify the relative transcript abundance (Schneider et al., 2012). Whole cell lysates were isolated 48 hours after HeLa cells were transfected with tagged SOX6-FL and SOX6-ΔE9 using Radioimmunoprecipitation assay (RIPA) buffer (Thermo Fischer Scientific cat # 89900). The following protocol was used to isolate whole cell lysates: 2µL of 100X Halt Protease Inhibitor (Thermo Fischer Scientific cat # 78437) was added to 198µL of RIPA buffer. HeLa cells were harvested and 200µL of the above mixture was added to the cells, incubated at 4°C for 30 minutes on a shaker, and centrifuged at 15,000rpm for 30 minutes at 4°C. The

supernatant was transferred to a new tube and the lysates were stored at -80°C. The lysates were quantified using the BCA assay (Thermo Fischer Scientific cat # 23227).

3.2.5 Western blot analysis

20-40µg of whole cell lysates was combined with 25µL 2X Novex Tris-Glycine SDS sample buffer (Thermo Fischer Scientific cat # LC2676) and 2µL β-mercaptoethanol (MP Biomedicals cat #194834). The samples were heated at 99°C for five minutes. The samples were loaded on 4-20% Novex Tris-Glycine gels (Thermo Fischer Scientific cat # XP04200BOX) and electrophoresis was performed using 1X Novex Tris-Glycine SDS Running buffer (Thermo Fischer Scientific ILTLC2675) at 150 volts for 1 hour and 15 minutes. PVDF membranes were prepared by incubating in methanol for one minute and then incubating in 1X Transfer buffer (Thermo Fischer Scientific cat # LC3675) which contains 20% Methanol (Thermo Fischer Scientific cat # A452-1) for 15 minutes. Protein samples were transferred onto the PVDF membrane at 25 volts for 1 hour and 30 minutes. 2% blocking solution was prepared using 1XTBST and non-fat milk powder (Dot Scientific cat # DSM17200-500). Membranes were blocked in blocking solution overnight at 4°C. Primary antibodies were prepared in the blocking solution and incubated at room temperature for one hour. The antibody concentrations used were as follows: rabbit anti-GFP (1:5000) (Sigma cat # G1544), rabbit anti-actin (1:5000) (Sigma cat #A5060), rabbit anti-KDM5C (1:250) (a kind gift from the Iwase lab) and mouse anti-flag (1:5000) (Sigma, F3165) Membranes were washed three times in 1XTBST. HRP-conjugated anti-rabbit antibody (1:10,000) and anti-mouse antibody (1:10,000) (Millipore cat #AP182P, AP192P) was added to blocking solution and incubated in secondary antibody solution for one hour at room temperature. Membranes were washed three times with 1X TBST, incubated with

SuperSignal West Dura substrate (Thermo Fischer Scientific cat # 34075). Western blot analyses on whole cell lysates isolated from HeLa cells transfected with FLAG- SOX6-FL and FLAG-SOX6-ΔE9 was performed using the same protocol except the lysates were run on a 6% tris-glycine SDS gel (Thermo Fischer Scientific cat # XP00060BOX) and were transferred overnight at 4°C.

3.2.6 Epoxomicin treatment of HeLa cells

HeLa cells were transfected with molar equivalents of GFP-SOX6-FL (4.05μg) and GFP-SOX6-ΔE9 (4μg) as previously described. 48 hours later, transfected HeLa cells were either treated with DMSO (Sigma cat #D2650) or 20μM epoxomicin (Millipore Corporation cat #324801) for 2 hours. Whole cell lysates were isolated and western blot analysis was performed as described above.

3.2.7 SOX6 protein localization studies

For localization studies, ~25,000 S16 cells were plated in each well of a 4-well slide (Thermo Fischer Scientific cat # 08-774-25). The cells were grown overnight prior transfections. 1.5μg GFP-SOX6-ΔE9 was transfected using the protocol described above. 72 hours post transfections, the cells were washed with 500μL 1X PBS (Thermo Fischer Scientific cat # 70011044) for five minutes and fixed using 4% paraformaldehyde in 1XPBS for 10 minutes. The cells were then washed three times with 1X PBS and permeabilized with 0.1% Triton X-100 (MP Biomedical cat # 807426) in 1X PBS for 10 minutes. All cells were stained with DAPI (Molecular Probes cat # 21490) to visualize the nuclei and were washed with 1X PBS for three times. ProLong Gold anti-fade was applied to each well and the slides were covered with glass coverslips (Fischer

brand cat # 12-545-F). Cells were imaged using Olympus 1X71 inverted microscope using the 40X objective and the Olympus cellSens image software.

3.2.8 RNA-sequencing of Schwann cells upon overexpressing SOX6

30-40% confluent T-75 flask of S16 cells were transfected with 10μg GFP-SOX6-ΔE9 using the protocol described in chapter 2. 72 hours post transfections the cells were prepared for flow cytometry. The cells were harvested and spun at 2000rpm for two minutes. The media was discarded and the cells were resuspended in 1X PBS so the final concentration was 3-4 million cells per mL. The cells were mixed well and passed through a 40μm cell strainer (Fischer Brand cat # 22363547). The cells were pipetted into a new 5mL falcon tube (Falcon cat # 352063) and flow cytometry was performed to collect GFP positive and GFP negative S16 cell populations at the University of Michigan Flow Cytometry core. Untransfected S16 cells served as the mock control and were used to establish base line florescence. RNA was isolated from GFP positive and negative populations using RNeasy kit and was treated with the DNase using the protocol described above. Three independent transfections using independently prepared GFP-SOX6-ΔE9 were preformed and served as biological replicates. The RNA samples were submitted to the University of Michigan sequencing core for library preparation. mRNA libraries were generated using the TruSeq kit (Illumina cat # 20020594) and the three biological replicates were pooled and subjected to 50 base pair single end sequencing on a single lane on the HiSeq 4000 platform.

FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to assess the quality of reads. Rat (Rn5) ensembl genome was downloaded from Illumina (https://support.illumina.com/sequencing/sequencing_software/igenome.html) and the reads were

mapped using STAR (Dobin et al., 2013). Subsequently, number of reads were counted using HT-Seq and default parameters were used (Anders et al., 2015). Differential expression analysis between GFP positive and GFP negative samples was performed using DESeq2 (Love et al., 2014). All programs were run on the flux servers at the University of Michigan.

3.3 Results

3.3.1 SOX10-CCS-13 is a previously unreported, alternative *Sox6* promoter

The SOX10 consensus sequence within SOX10-CCS-13 is conserved between human and zebrafish (Fig. 3.1A), further suggesting an important role for this SOX10 response element in jawed vertebrates. Additionally, analysis of SOX10 ChIP-seq data generated from rat sciatic nerve (Srinivasan et al., 2012) and our DNase-seq data generated from S16 cells revealed evidence for SOX10 binding and open chromatin at SOX10-CCS-13 (Fig. 3.1B) Examination of SOX10-CCS-13 on the UCSC Genome Browser (Kent et al., 2002) revealed seven unique *SOX6* mRNA isoforms in human, mouse, and rat, distinguished by alternative, non-coding first exons. Interestingly, SOX10-CCS-13 maps directly upstream of the 3'-most alternative first exon, which we named *SOX6* exon 1G (Fig. 3.2). We therefore hypothesized that SOX10-CCS-13 acts as an alternative promoter at *SOX6*. To test this, we performed 5'-rapid amplification of cDNA ends (5'-RACE). Briefly, a cDNA library was generated using RNA isolated from cultured rat Schwann (S16) cells and a reverse primer in exon 5 of the rat *Sox6* gene. Subsequently, nested PCR was performed using reverse primers in exon 4 and then exon 3 of *Sox6*. The PCR products were cloned, sequenced, and aligned to the rat *Sox6* locus. These analyses revealed the presence of five unique *Sox6* transcription start sites in cultured Schwann cells with 14 of the 44 *Sox6*-specific sequences mapping directly downstream of SOX10-CCS-13 (Fig. 3.2).

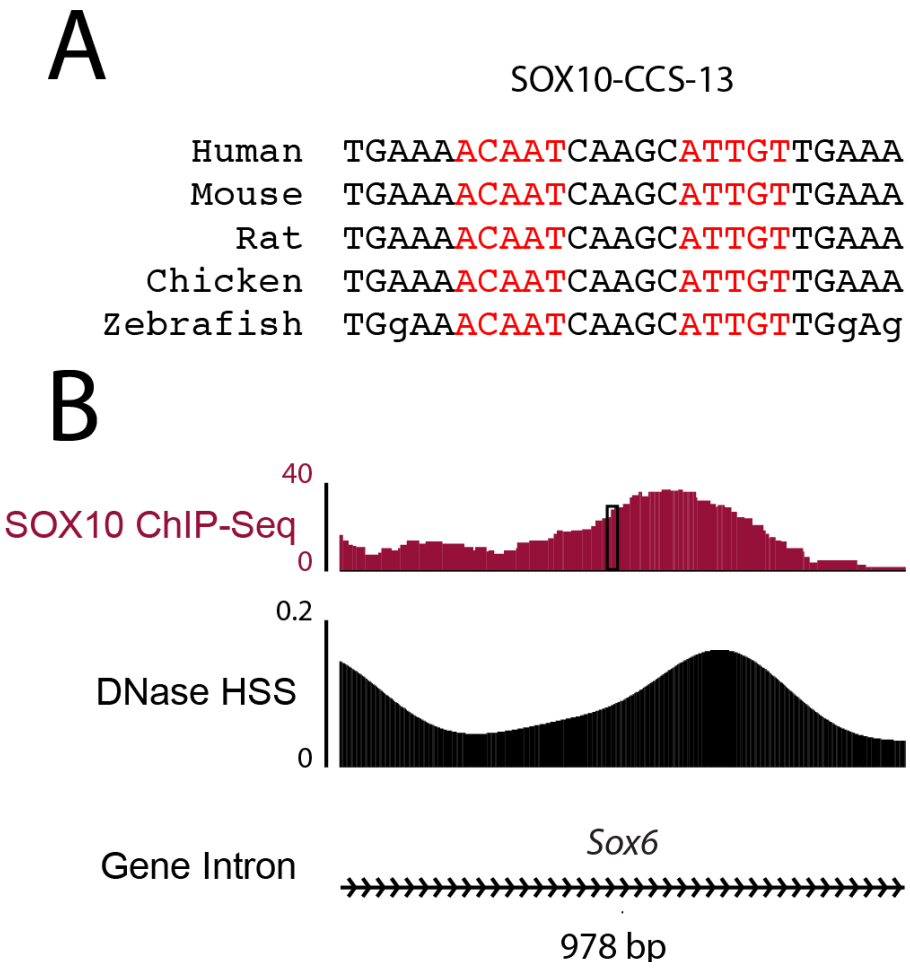


Figure 3.1. SOX10 binding and chromatin accessibility at SOX10-CCS-13. (A) Conservation of SOX10 consensus sequence (red text) and surrounding genomic region in mammals (human, mouse, and rat) and vertebrates (chicken and zebrafish). Uppercase letters indicate conserved bases and lowercase letters indicate non- conserved nucleotides. (B) SOX10 ChIP-seq and DNase-seq peaks are shown at SOX10-CCS-13. The y-axes represent normalized read depth (SOX10-ChIP-seq) and F-seq score (DNase-seq). The black box indicates the position of the SOX10 consensus sequence. The length of the genomic region we tested in the luciferase assay in indicated at the bottom and the arrows indicate the direction of transcription.

Analysis of RNA-seq data generated in S16 cells (Law et al., in review) also revealed reads that map to *Sox6* exon 1G (Fig. 3.2), with split reads into downstream exons, but no split reads into upstream exons (data not shown). Additionally, we were able to amplify and sequence-verify a full length *Sox6* mRNA that originates at exon 1G in S16 cells (Fig. 3.2). To assess the *in vivo* relevance of SOX10-CCS-13 we analyzed recently published SOX10 ChIP-seq data performed on nuclei isolated from rat spinal cord and sciatic nerve (Lopez-Anido et al., 2015; Srinivasan et al., 2012). Furthermore, to establish that this genomic segment resides in open chromatin we performed DNase-seq analysis on nuclei isolated from cultured rat Schwann (S16) cells. The SOX10 ChIP-seq analyses revealed that SOX10 binds to SOX10-CCS-13 in relevant tissues *in vivo* and the DNase-seq experiment revealed that this genomic segment resides in open chromatin in cultured Schwann cells (Fig. 3.2, green box). Combined, these data support our conclusion that SOX10-CCS-13 is a SOX10 response element in Schwann cells. Interestingly, these genome-wide functional studies revealed additional SOX10 binding sites at *Sox6 in vivo* further supporting the notion that *Sox6* is a SOX10 target gene (*e.g.*, brown highlighted region in Fig. 3.2). Combined, our data indicate that SOX10-CCS-13 represents an alternative, SOX10-responsive promoter at *Sox6*.

3.3.2 SOX10 is necessary and sufficient for the expression of *Sox6* transcripts harboring exon 1G

To determine if SOX10 is sufficient to direct the expression of *Sox6* transcript containing exon 1G, we performed RT-PCR using primers designed in *Sox6* exon 1G and exon 2 in regions

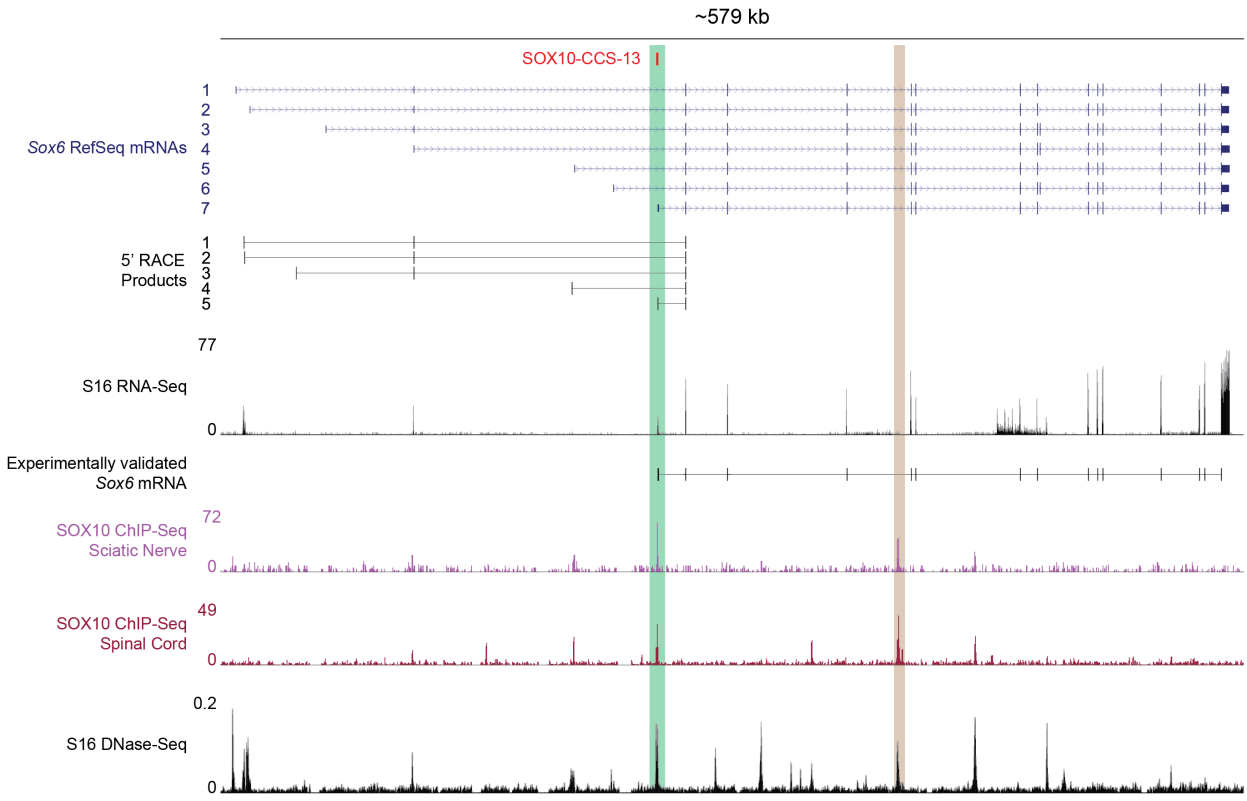


Figure 3.2. SOX10-CCS-13 is an alternative promoter at the Sox6 locus. The ~579 kb rat Sox6 locus is shown on the UCSC Rat Genome Browser. SOX10-CCS-13 is indicated in red along with the seven human, mouse, and rat SOX6 RefSeq mRNAs (blue). Sox6-specific 5' RACE was performed on RNA from S16 cells and the five distinct Sox6 sequences were mapped to the rat genome. Please note that SOX10-CCS-13 maps to both the 5' end of the seventh Sox6 mRNA and the fifth unique 5' RACE-generated sequence. RNA-Seq data from S16 cells were mapped to Sox6 (the y-axis indicates sequence read depth) as was a PCR-amplified, full-length mRNA that contains Sox6 exon 1G. Genome-wide regulatory marks were also mapped to Sox6 with the Y-axes indicating normalized sequence read depths (both SOX10 ChIP-seq data sets) and F-Seq scores (DNase-seq). The green highlighted region marks the general position of SOX10-CCS-13 across all data sets and the brown highlighted region marks another potential SOX10 response element at Sox6.

conserved between rat and mouse. While these primers amplify *Sox6* transcripts containing exon 1G from a cDNA library generated from RNA isolated from immortalized rat (S16) Schwann cells we were not able to amplify these transcripts from a cDNA library generated from cultured mouse motor neurons (MN1 cells), which do not express endogenous SOX10 (Fig. 3.3A). However, when MN1 cells were transfected with a construct to express wild-type SOX10, *Sox6* transcripts containing exon 1G were detected and verified by DNA sequence analysis. Mock transfection or transfection with a construct to express a non-functional mutant version of SOX10 (E189X) (Inoue et al., 2004) did not allow amplification of *Sox6* transcripts containing exon 1G (Fig. 3.3A). Thus, SOX10 is sufficient to activate the expression of *Sox6* transcripts harboring exon 1G in MN1 cells, which do not express endogenous SOX10 (Hodonsky et al., 2012). To determine if SOX10 is necessary for the expression of *Sox6* transcripts containing exon 1G in Schwann cells, we treated S16 cells with a previously validated siRNA against *Sox10* (Gokey et al., 2012; Lopez-Anido et al., 2015) and tested for an effect on total *Sox6* mRNA levels and for an effect on the level of transcripts containing exon 1G. This analysis revealed a ~70% decrease in both total *Sox6* expression and in the expression of transcripts containing exon 1G (Fig. 3.3B), consistent with SOX10 regulating the promoter activity of SOX10-CCS-13 which is responsible for regulating exon 1G containing *SOX6* transcript in Schwann cells. Combined, our data indicate that SOX10 is both necessary and sufficient for the expression of *Sox6* mRNA isoform 7 (Fig. 3.3) in our *in vitro* cell culture model systems.

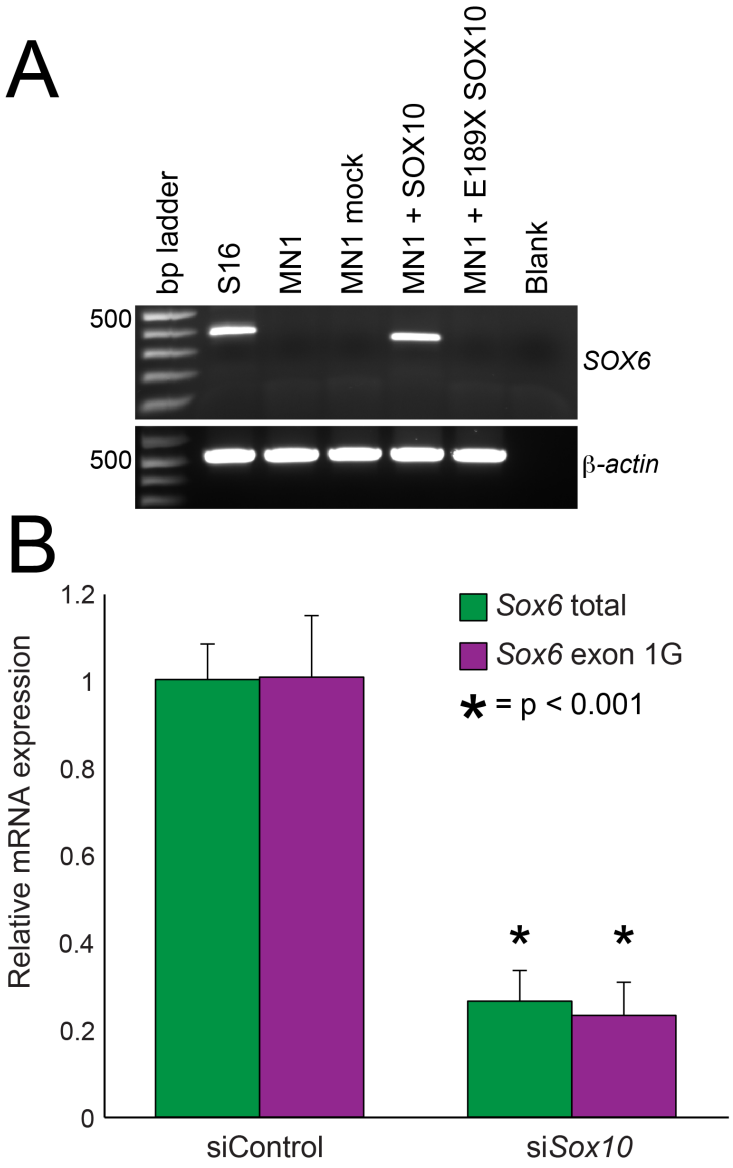


Figure 3.3. SOX10 is necessary and sufficient for SOX6 expression. (A) RT-PCR was performed to detect the expression of Sox6 transcripts harboring exon 1G using cDNA isolated from S16 cells, MN1 cells, or MN1 cells transfected with no expression construct (mock) or a construct to express wild-type or dominant-negative (E189X) SOX10. Base pair (bp) ladders are indicated on the left. RT-PCR for β-actin and samples including no cDNA ('Blank') were employed as positive and negative controls, respectively. Please note that while the same primers were used for each reaction, the rat (S16) PCR product was 402 base pairs and the mouse (MN1) PCR product was 349 bp; the rat genome harbors a 53 base pair rat-specific insertion, which we confirmed via DNA sequence analysis. (B) Rat Schwann (S16) cells were treated with a control siRNA (left side) or a siRNA targeted against Sox10 (right side). Quantitative RT-PCR was used to measure expression levels of total Sox6 (green bars) or Sox6 exon 1G-containing (purple bars) transcripts. Error bars indicate standard deviations.

3.3.3 *SOX6* mRNA undergoes alternative splicing and exon 9 is conserved among mammals

SOX6 belongs to the SOXD family of transcription factors along with SOX5 and SOX13 (Lefebvre, 2010). The SOXD transcription factors inhibit myelination in oligodendrocytes (Baroti et al., 2015; Stolt et al., 2006) and the role of SOX6 in Schwann cells is not yet known. *SOX6* mRNA undergoes alternative splicing to produce isoforms that either include (SOX6-FL) or exclude (SOX6-ΔE9) exon 9 (indicated by red arrow in Fig. 3.4A). I performed multiple species conservation analysis and interestingly, exon 9 is conserved only in mammals (red box in Fig. 3.4A). To test for the presence of exon 9 in mammals and lower vertebrates, we performed RT-PCR analysis using cDNA libraries prepared from mouse and zebrafish tissues. Forward and reverse primers were designed in mouse/zebrafish exon 8 and exon 10 respectively. cDNA libraries were generated from RNA isolated from brains of embryonic (E) 16.5 embryos, post-natal day (P) 3 and adult mice. We were able to amplify transcripts with (293bp indicated by an arrow in Fig. 3.4B) and without exon 9 (170bp Fig. 3.4B) in mice. cDNA libraries were prepared from RNA was isolated from zebrafish embryos at 24 hours post fertilization (hpf), 48 hpf, 5 days post fertilization (dpf) and from brain of adult zebrafish. RT-PCR analysis revealed the presence of transcript without exon 9 at all developmental time points (148bp, Fig. 3.4C). Taken together, these data indicate exon 9 is conserved only in mammals.

3.3.4 Overexpressing both isoforms of SOX6 indicates lower expression of SOX6-FL

SOX6 contains two coiled-coil (CC) domains, which is specific to the SOXD group and facilitates homo- and hetero-dimerization (Lefebvre, 2010). The HMG-domain is present towards the C-terminal of the protein. Interestingly, SOXD proteins do not contain transactivation or transrepression domain so their regulatory activity is dictated by the co-factors

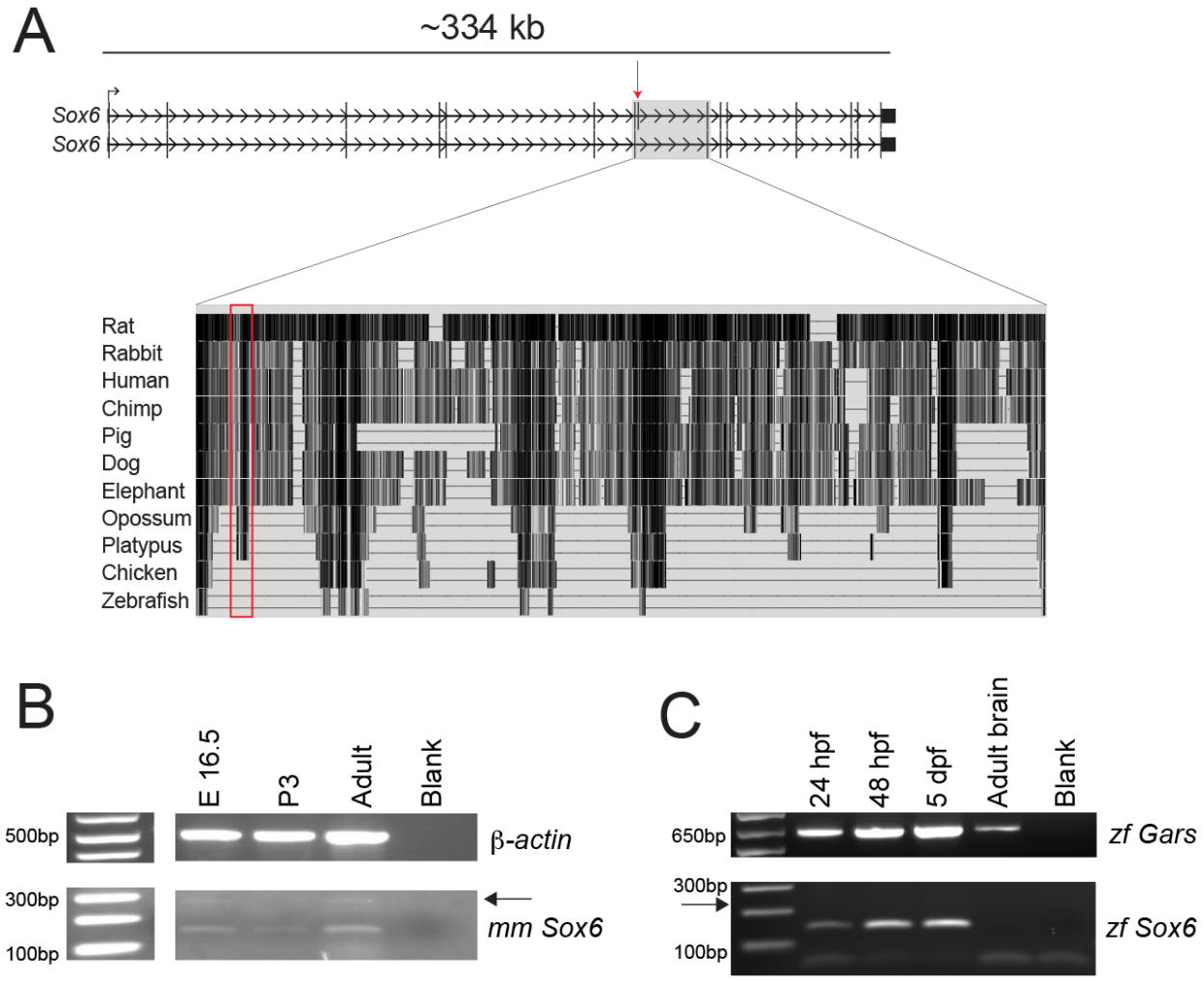
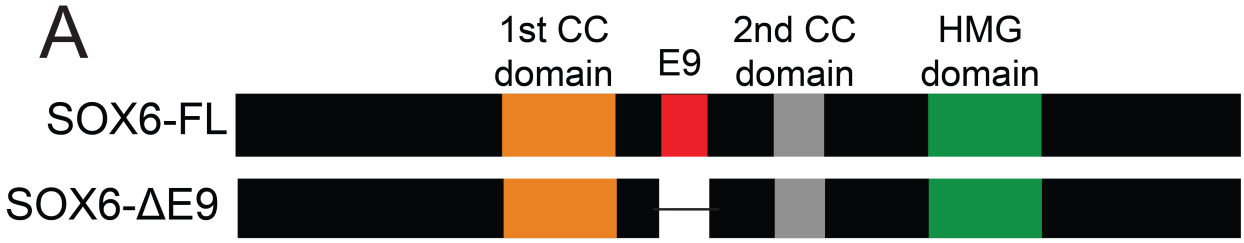


Figure 3.4. SOX6 mRNA undergoes alternative splicing. (A) ~334kb mouse Sox6 locus showing two annotated RefSeq mRNA isoforms. The black arrow indicates the translational start site and the red arrow indicates exon 9. The region from exon 8 to exon 10 is highlighted in grey and a zoomed in view displaying the conservation of the region in shown. Exon 9 region is depicted in the red box. (B) RT-PCR analysis was performed to detect Sox6 transcripts with and without exon 9 using cDNA isolated from embryonic (E) day 16.5, post-natal (P) 3, and Adult brains from mice. RT-PCR for β-actin and samples including no cDNA ('Blank') were used as positive and negative controls, respectively. The arrow indicates the presence of exon 9 containing transcripts. (C) RT-PCR assay was performed to detect Sox6 transcripts with and without exon 9 using cDNA isolated from 24 hours post fertilization (hpf), 48hpf, 5 days post fertilization (dpf) zebrafish embryos and adult zebrafish brain. RT-PCR for zfGars and samples including no cDNA ('Blank') were used as positive and negative controls, respectively. The arrow indicates the absence of exon 9 containing transcripts. Base pair (bp) ladder for both the gels are indicated on the left.

they interact with (Lefebvre, 2010) (Fig 3.5A). The two protein isoforms contain all the known functional domains both isoforms are known to display similar transcriptional activities (Lefebvre et al., 1998). The functional differences between the two isoforms are currently unknown and exon 9 does not code for any known functional domains (Cohen-Barak et al., 2001). Amino acid sequence for exon 9 is highly conserved among mammals (Fig. 3.5B). We unable to identify any putative protein domain encoded by exon 9 using Pfam analysis (Finn et al., 2016). Lysine residues are targets for posttranslational modification and are often ubiquitinated (Zencheck et al., 2012). Exon 9 contains three highly conserve lysine resides and we used UbPred software to identify potential ubiquitination sites (Radivojac et al., 2010). Interestingly, the last lysine within exon 9 (lysine residue shown in red in Fig. 3.5B) was predicted as an ubiquitination site suggesting that SOX6-FL may be targeted for proteosomal degradation.

To investigate the functional differences between the two isoforms we cloned them into a plasmid, which contains an N-terminal GFP tag. To detect the expression levels of our fusion proteins, we transiently transfected HeLa cells, which were chosen because of their high transfection efficiency. HeLa cells were transfected with GFP-SOX6-FL, GFP-SOX6-ΔE9, MTMR2-1 (Fogarty et al., 2016), or a plasmid to express GFP, we isolated whole cell lysates and performed western blot analyses using anti-GFP antibody. MTMR2-1 is one of the isoforms of myotubularin-related protein 2 (MTMR2), which has been cloned into the same N-terminal GFP-plasmid as the two SOX6 isoforms and serves as a positive control.



B

EXON 9

```
Human      KGHVSHPQINQRLKGLSDRFGRNLDPSEHGGGHSYNHKQIE
Rhesus     KGHVSHPQINPRLKGLSDRFGRNLDTFEHGGGHSYNHKQNE
Mouse      KGHVSHPQINPRLKGISDRFGRNLDPSEHGGGHSYNHRQIE
Dog        KGHVSHPQINPRLKGLSDRFGRSLDTFEHGGGHSYNHKQIE
Elephant   KGHVSHPQINPRLKGLSDRFGRNLDTFEHDGGHSYNHKQIE
```

Figure 3.5. Schematic representing the two SOX6 protein isoforms. (A) SOX6- full length (SOX6-FL) and SOX6-ΔE9. The two coiled-coil domain (1st CC in orange and 2nd CC in grey) and HMG (green) domain are indicated. Exon 9 in SOX6-FL is shown in red. (B) Amino acid residues that exon 9 codes for is shown for different species. Residues that are different among the species are depicted as green amino acids. Conserved lysine (K) residue shown in red is the predicted site for ubiquitination.

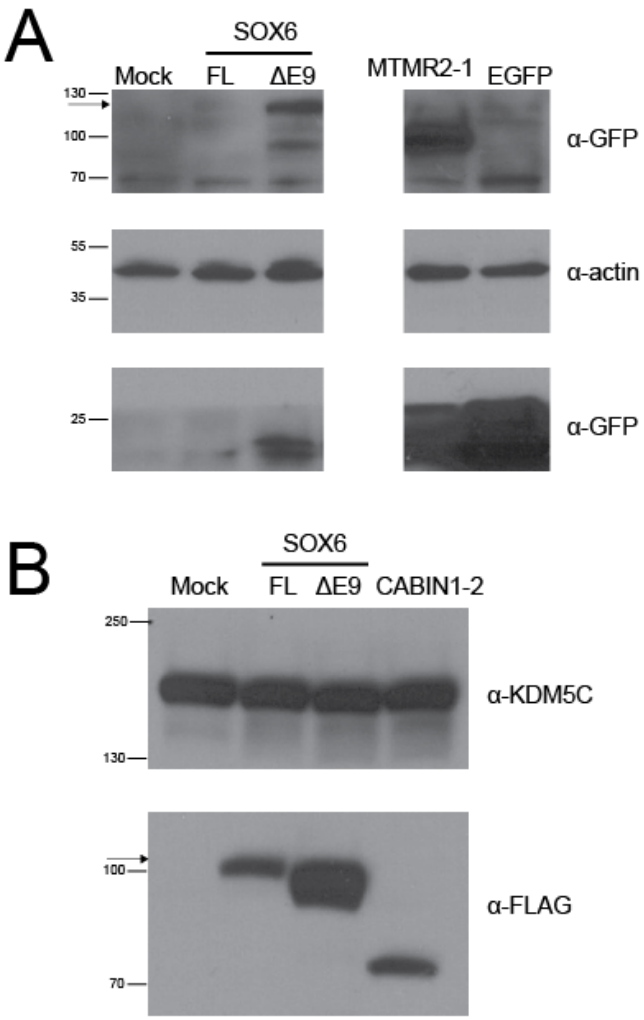


Figure 3.6. SOX6-ΔE9 is detected at higher levels upon overexpression *in vitro*. Western blot analyses using whole cell lysates isolated from HeLa cells transfected (A) with no expression construct (mock) or GFP-tagged SOX6-FL, SOX6-ΔE9, MTMR2-1 or plasmid expressing EGFP protein was performed using anti-GFP and anti-actin antibodies. (B) with no expression construct (mock) or FLAG-tagged SOX6-FL, SOX6-ΔE9, and CABIN1-2 was performed using anti-FLAG and ant-KDM5C antibodies. A protein ladder in kilodalton is shown on the left. The arrow indicates expected products for both SOX6 isoforms.

We detected appropriately sized bands for SOX6-ΔE9 (~114kDa), MTMR2-1(~100kDa), and GFP (~27kDa), and we did not detect fusion protein expression in mock-transfected cells (Fig 3.6A). Interestingly, we were unable to detect the GFP-SOX6-FL product (~119kDa, indicated by the arrow in Fig 3.5B). To rule out the possibility of the GFP tag having an effect on protein expression, we performed similar analysis using SOX6-FL and SOX6-ΔE9 cloned into an expression construct with an N-terminal FLAG-tag. Here, we observed an appropriately sized product for SOX6-FL (~92kDa), SOX6-ΔE9 (~88kDa), CABIN1-2 (~70kDa) (Fig 3.6B). CABIN1-2 is an isoform of CABIN1 that has been cloned into the same N-terminal FLAG-plasmid as the SOX6 isoforms and serves as a positive control (Fogarty et al., in preparation). Consistent with our expression studies using a GFP tag, we observed significantly lower levels of FLAG-SOX6-FL expression compared to FLAG-SOX6-ΔE9. There are multiple explanations for reduced expression of SOX6-FL in our transient transfection assays, including: (i) variation in transfection efficiency between the two plasmids; (ii) specific degradation of the SOX6-FL transcript produced from the expression plasmid; and/or (iii) specific degradation of the SOX6-FL protein produced from the expression plasmid.

To test the mRNA expression levels from the plasmid we performed RT-PCR using cDNA library prepared from RNA isolated from HeLa cells that were either mock transfected, or transfected with constructs to express either GFP-SOX6-FL or GFP-SOX6-ΔE9. RT-PCR primers were designed within the coding sequence of GFP and within exon one of SOX6. We were indeed able to amplify the right size product (136bp) from cells transfected with GFP-SOX6-FL and GFP-SOX6-ΔE9 and unable to amplify a product from mock transfected HeLa cDNA (Fig 3.7).

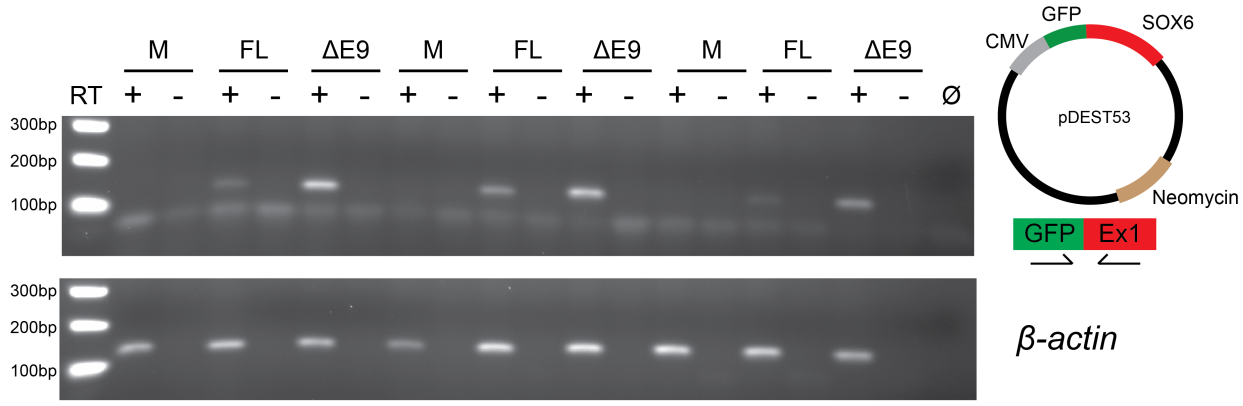


Figure 3.7. GFP-SOX6-FL mRNA is expressed at lower levels *in vitro*. RT-PCR was performed using cDNA library prepared from HeLa cells that were either mock (M) transfected or transfected with GFP-tagged SOX6-FL (FL) or GFP-tagged SOX6-ΔE9 (ΔE9) with (+) or without (-) reverse transcriptase (RT) enzyme and Φ does not contain a template. The experiment has been performed using cDNA libraries prepared from three independent transfections. Molecular weight ladder is shown on the left. A cartoon of pDEST53 plasmid, which was used to clone the SOX6 isoforms, is shown on the right. This plasmid contains an N-terminal GFP tag and a CMV promoter drives the expression of the fusion protein. Forward primer was designed towards the C-terminal of GFP open reading frame and reverse primer was designed within exon 1 of SOX6. β-actin served as a positive control. Base pair (bp) ladder is indicated on the left.

We were unable to amplify a product from cDNA libraries prepared without a reverse transcriptase (RT- in Fig. 3.7), ensuring that we were not amplifying products from plasmid DNA. To test the reproducibility of our data, we performed this experiment three times using cDNA libraries prepared from independent transfections. Amplification of endogenous β-actin served as a control to ensure the integrity of cDNA libraries. These data suggest that GFP-SOX6-FL transcript is expressed at lower levels compared to GFP-SOX6-ΔE9 (Fig 3.7).

To test the relative expression of GFP-SOX6-FL and GFP-SOX6-ΔE9 transcript to another plasmid encoded transcript we performed RT-PCR analysis to amplify a transcript encoded by a selection marker in the expression plasmid (Fig 3.8). The pDEST53 plasmid contains a neomycin selection cassette and primers were designed within this cassette and primers were designed within exon 13 and 14 of SOX6. We performed RT-PCR analysis using cDNA samples generated from HeLa cells that were either mock transfected, or transfected with GFP-SOX6-FL or GFP-SOX6-ΔE9 and using three different PCR amplification cycles: 26, 28, and 30. Following RT-PCR, the products were separated on agarose gels and the intensities of the bands were quantified using ImageJ software (Schneider et al., 2012). We were able to amplify the right size products for GFP-SOX6-FL, GFP-SOX6-ΔE9 (182bp) and neomycin (134bp) (Fig. 3.8) for all the three PCR cycles. The relative abundance of FL and ΔE9 transcripts was calculated by normalizing the intensity of the SOX6 bands to the neomycin bands. The expression of SOX6-FL and SOX6-ΔE9 is comparable to their respective expression of neomycin (Fig. 3.8). However, GFP-SOX6-FL transcript is expressed at lower levels compared to GFP-SOX6-ΔE9. The difference we observe may be due to reduced transfection efficiency of GFP-SOX6-FL or due to post-translational modifications.

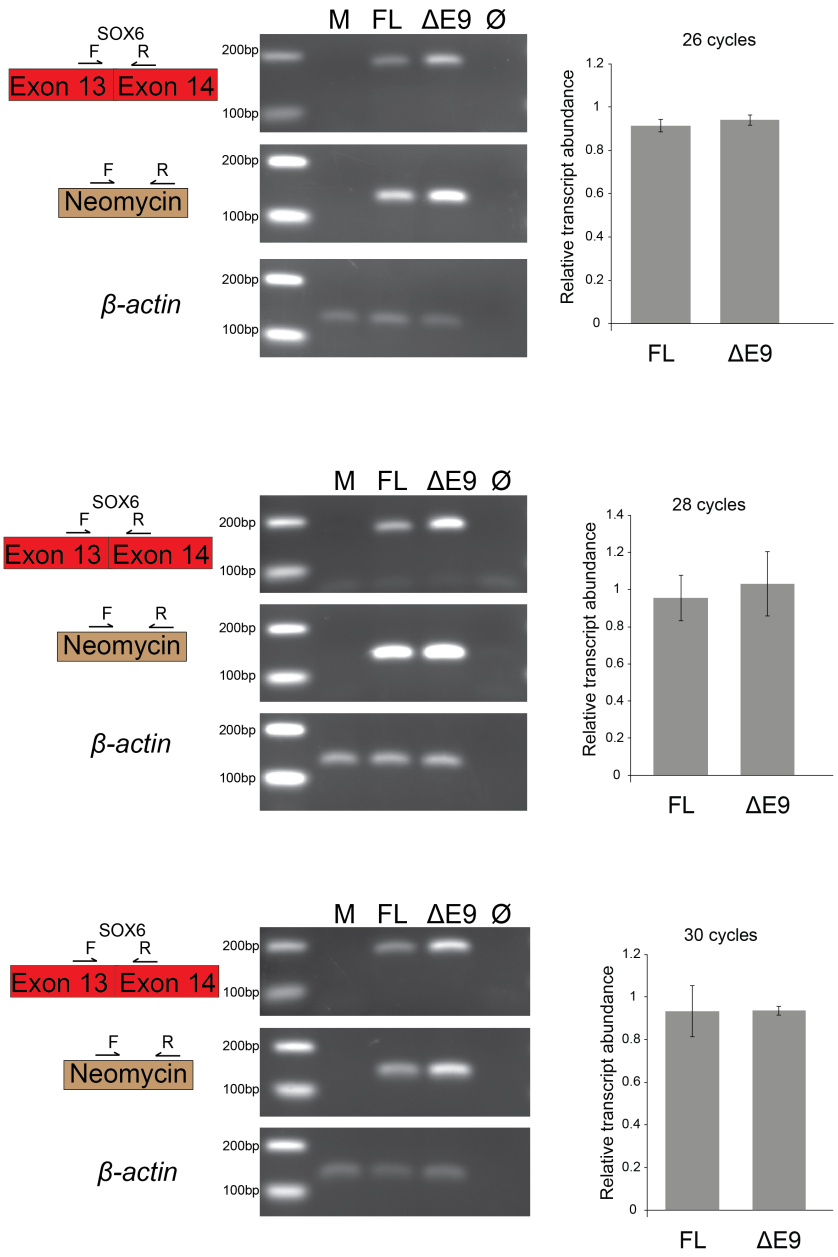


Figure 3.8. GFP-SOX6-ΔE9 transcript is expressed at higher levels *in vitro*. RT-PCR was performed using cDNA libraries prepared from mock (M), GFP-tagged SOX6-FL or GFP-tagged SOX6-ΔE9 (ΔE9) transfected HeLa cells and Φ does not contain a DNA template. The assay was performed at three different PCR cycles (26, 28, and 30). The position of the forward (F) and reverse (R) primers is indicated on the left. The ImageJ quantification is shown on the right for the three cycle numbers. The x-axis on the graph indicates the two Sox6 mRNA (FL and ΔE9) isoforms and the y-axis indicates relative abundance of the two transcripts. Base pair (bp) ladder is indicated on the left. β-actin served as a positive control.

3.3.5 SOX6-FL protein expressed from the plasmid is not stabilized upon epoxomicin treatment

Protein degradation in eukaryotes is mediated by the proteasome through the ubiquitin pathway. Once proteins have been ubiquitinated, they are targeted by the 26S proteasome for degradation (Varshavsky, 2017). Lysine residues within proteins are known to be ubiquitinated (Zencheck et al., 2012) and exon 9 contains a lysine residue predicted to undergo ubiquitination (Fig. 3.5B). To test if GFP-SOX6-FL is targeted by the proteasome, we treated HeLa cells with epoxomicin which inhibits *in vivo* proteasome function (Meng et al., 1999). Briefly, HeLa cells were mock transfected or transfected with GFP-SOX6-FL or GFP-SOX6-ΔE9. 48 hours post transfections, HeLa cells were treated with DMSO or 20μM epoxomicin for 2 hours and whole cell lysates were extracted as previously described and western blot analysis was performed using anti-GFP, anti-p53, and anti-Rab5 antibodies (Fig. 3.9). Epoxomicin treatment is known to stabilize p53 in HeLa cells and serves as our positive control to ensure inhibition of the proteasome (Meng et al., 1999). Treating the cells with DMSO did not affect the expression of GFP-SOX6-FL or GFP-SOX6-ΔE9 and these data are consistent with our earlier findings and we observe lower p53 levels in the DMSO treated cells. Indeed, we observe accumulation of p53 (53kDa) upon epoxomicin treatment however; we do not observe the right size band for GFP-SOX6-FL (indicated by the arrow in Fig. 3.9) and the level of GFP-SOX6-ΔE9 (114kDa) is unchanged upon epoxomicin treatment. Rab-5 (24kDa) serves as our loading control. These data indicate that GFP-SOX6-FL is not targeted by the proteasome and transfection efficiency of the two plasmids accounts for the observed expression difference.

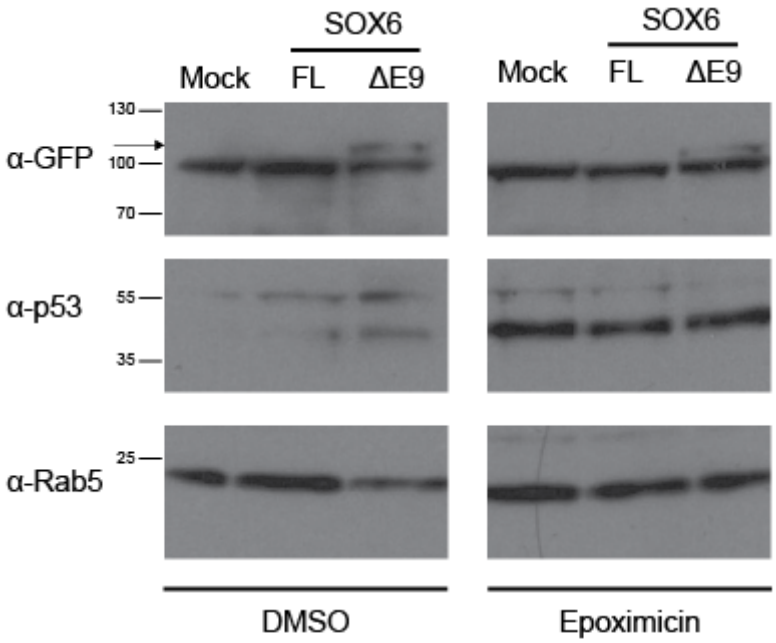


Figure 3.9. SOX6-ΔFL expression is unaffected by epoximicin treatment. Western blot analyses using whole cell lysates isolated from HeLa cells transfected with no expression construct (mock) or GFP-tagged SOX6-FL, SOX6-ΔE9 and treated either with DMSO or epoximicin using anti-GFP, anti-p53 and anti-Rab5 antibodies. A protein ladder in kilodalton is shown on the left. The arrow indicates expected products for both SOX6-FL isoform.

3.3.6 The SOX6-GFP fusion protein localizes to the nucleus in cultured Schwann cells

GFP-SOX6-ΔE9 transfects better that GFP-SOX6-FL so we used the former for functional studies. SOX6 is a transcription factor and it has been shown to localize to the nucleus (Ohe et al., 2002). To ensure that the GFP-tag does not affect localization of SOX6, we transfected S16 cells with GFP-SOX6-ΔE9. Our fluorescence microscopy analysis revealed that SOX6 fusion protein overlaps with nuclear DAPI staining (indicated by white arrows in Fig 3.10) suggesting that our fusion protein is able to translocate to the nucleus. These images were not taken on a confocal microscope and need to be validated. GFP-SOX6-ΔE9 forms puncta in the nucleus and the data are consistent with published data that suggest that endogenous SOX6 forms nuclear speckles and co-localizes with splicing factors (Ohe et al., 2002).

3.3.7 Global transcriptome analysis suggests a role for SOX6 in Schwann cell dedifferentiation

SOX6 is a negative regulator of myelination in oligodendrocytes and may act in a similar fashion in Schwann cells. Specifically, ablating *Sox6* in the mouse CNS resulted in premature myelination and *in vitro* studies suggested that SOX6 represses the expression of *Mbp* by outcompeting SOX10 for promoter occupancy (Stolt et al., 2006). We therefore tested the hypothesis that SOX6 globally represses loci important for myelination in Schwann cells by over-expressing GFP-SOX6-ΔE9 in immortalized rat Schwann (S16) cells and performing RNA-seq analysis.

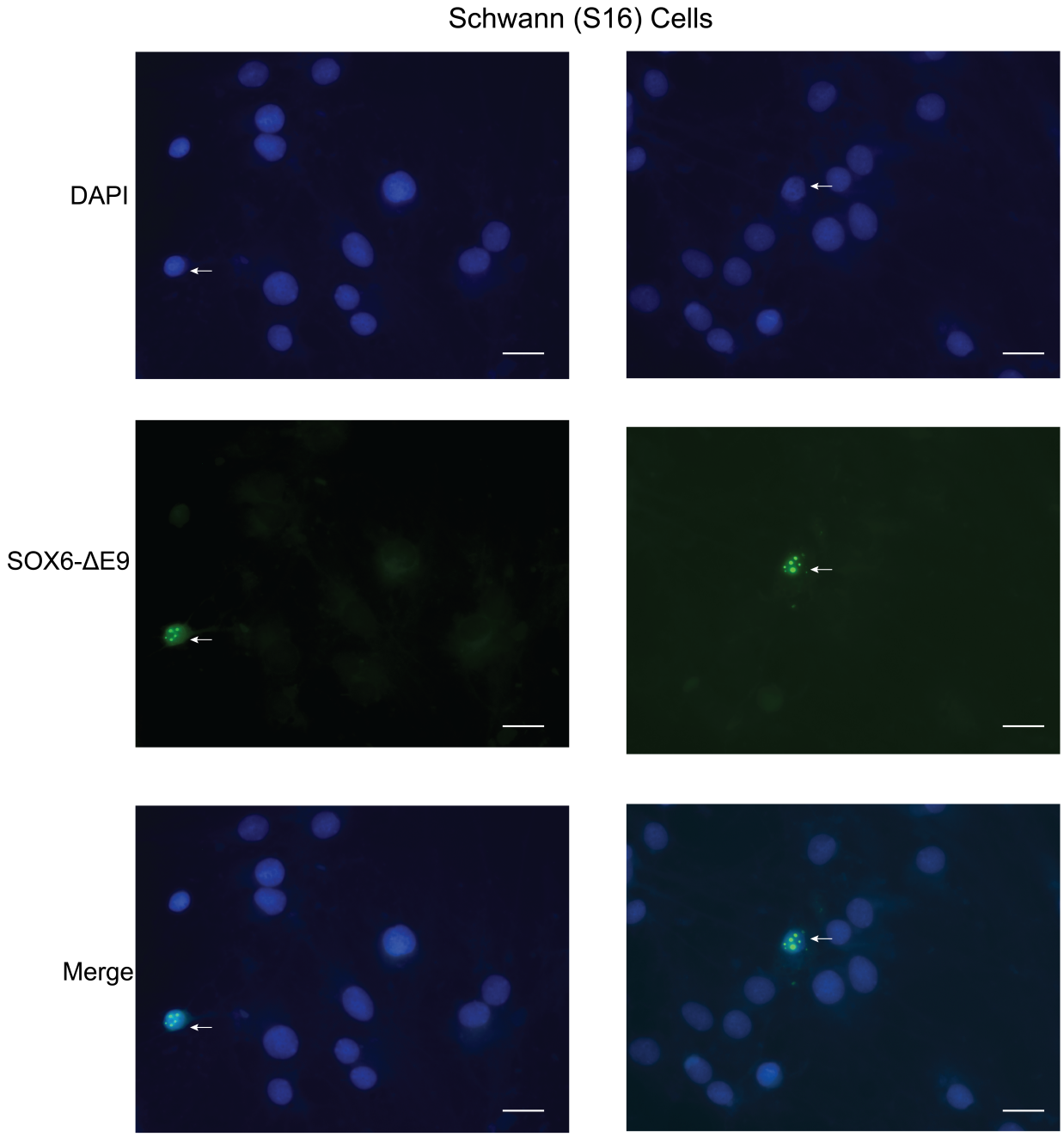


Figure 3.10 SOX6-ΔE9 fusion protein localizes to the nucleus. S16 cells were transfected with GFP-tagged SOX6-ΔE9. Cells were fixed and treated with DAPI (blue) to stain the nucleus. The images form the first two panels were merged (Merge) to show co-localization. The white arrow indicates cells that express the fusion protein, which is localized to the nucleus. Scale bars, 20um.

S16 cells express high levels of key myelin genes including *Sox10*, *Mbp*, and *Mpz* and are a relevant model to test our hypothesis (Hodonsky et al., 2012; Toda et al., 1994) (Law et al., under review). After transfection, flow cytometry was performed to collect GFP-positive and GFP-negative cells, RNA was isolated from both these cell populations, and we performed RNA seq analysis. RNA was isolated from three independent GFP-positive and GFP-negative S16 cells. Stranded libraries were prepared using the Truseq Illumina kit and all samples were pooled and sequenced on a single lane of a HiSeq4000. About 45-80 million high quality reads (FastQC quality score >38) were obtained for each sample and, of these, ~80% of the reads mapped to the rat genome (Table 3.1). To assess the variability between our biological replicates we performed principal component analysis (PCA) (Fig. 3.11A). Majority of the variance (84%) is explained by the first principal component (PC1), which separates the GFP-positive and GFP-negative samples. The second principal component (PC2) explains the variance among biological replicates of the two experimental conditions (GFP-positive and GFP-negative) (Fig. 3.11A). We used DeSeq2 to identify differentially expressed (DE) genes by comparing the transcriptomes of GFP-positive and GFP-negative cells (Fig. 3.11B). A total of 1,117 genes were differentially expressed between the two experimental conditions (680 up- and 437 down-regulated; adj P-value <0.01).

To gain insights into the function of SOX6 in Schwann cells, we performed gene ontology (GO) enrichment analysis using the overrepresentation test for biological processes using the 1,117 DE genes. GO analysis of up- regulated and down-regulated genes, separately, resulted in an enrichment of 188 and 111 GO terms, respectively (Table 3.2 and 3.3).

	Total number of reads	Mapped unique (%)
S16 SOX6 ΔE9 GFP positive- 1	81,837,766	80.7
S16 SOX6 ΔE9 GFP negative- 1	88,615,458	83
S16 SOX6 ΔE9 GFP positive- 2	87,861,686	80.9
S16 SOX6 ΔE9 GFP negative- 2	43,342,890	81.1
S16 SOX6 ΔE9 GFP positive- 3	46,662,646	81.2
S16 SOX6 ΔE9 GFP negative- 3	47,821,905	83.3

Table 3.1 Summary of the RNA-seq data. Total number for reads for each of sample and percentage (%)of the reads that mapped to rat ensemble (RN5) genome is shown

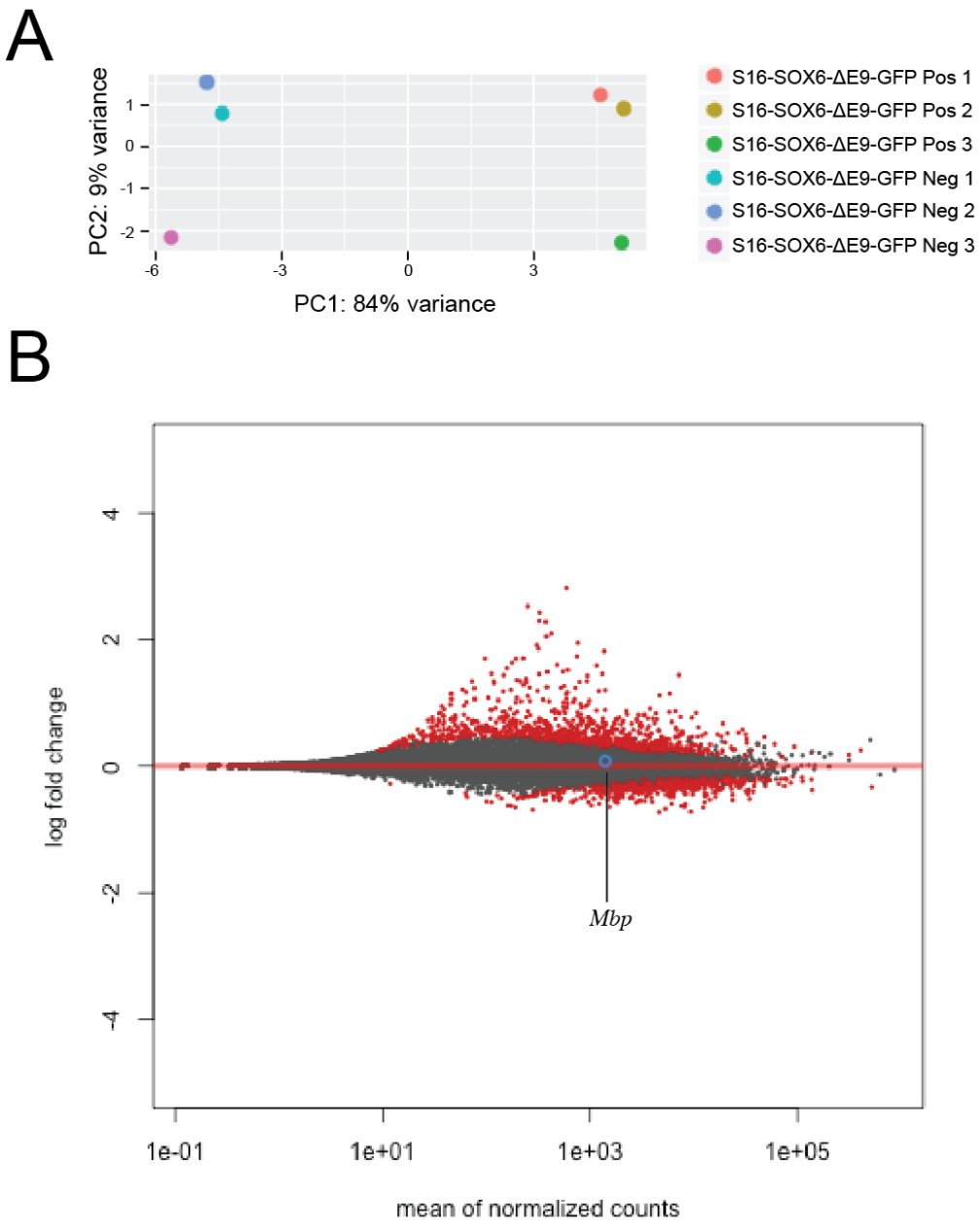


Figure 3.11. Global transcriptomic changes upon SOX6 overexpression in Schwann cells. (A) Principal component analysis performed to calculate the variability between the biological replicates. PC1 (x-axis) displays the majority of variance (84%) and PC2 (y-axis) displays the residual variance (9%). Different colored dots are associated with different biological replicate as shown on the right (B) MA plot showing the mean expression change of every gene (x-axis) plotted against log2 fold change (y-axis). The red line at 0 indicates genes no change in expression upon SOX6 overexpression. The dots about and below the red line are up- and down-regulated, respectively. The red dots indicate genes that are significantly differentially expressed (p-value<0.05). *Mbp* is indicated by the blue circle

Consistent with the role of SOX6 in regulating cell proliferation during erythroid development we indeed observe a GO term (p-value= $9.04x10^{-9}$) associated with this specific function of SOX6 in our data (Dumitriu et al., 2006). Some of the GO terms associated with up-regulated genes important for Schwann cell biology are glial cell development, glial cell differentiation, gliogenesis, and positive regulation cell migration. The glial cell development GO term is associated with genes that are up-regulated and the list contains genes such as *SIRT2*, *MMP14*, *MXRA8*, *DAG1*, *GFAP*, *NDRG1*, *SOD1, GSN*, *CD9*, and *LRP1*. These genes are either implicated in facilitating Schwann cell regeneration following repair or promoting Schwann cell proliferation and migration. We identified 111 GO terms associated with the down-regulated genes in our RNA-seq analysis. SOX6 is known to inhibit neuronal differentiation (Lee et al., 2014) and indeed we observe a GO term associated with neuron development that is enriched in our data. Interestingly, we did not observe any decrease in *Mbp* expression (Fig. 3.9) upon GFP-SOX6-ΔE9 overexpression in S16 cells, suggesting that SOX6 may not directly or indirectly inhibit *Mbp* expression; however, additional analyses are required to confirm our findings and we cannot rule out that *Mbp* repression is mediated by full-length (FL) SOX6. Taken together our GO analysis reveals putative role for SOX6 during Schwann cell development and after nerve injury.

3.3.8 *MBP* promoter activity is unaffected by SOX6 overexpression

Previous studies showed that SOX6 is important for repressing premature myelination in the CNS. Moreover, it was shown that this process may involve SOX6 outcompeting SOX10 at myelin loci, specifically the *Mbp* promoter (Stolt et al., 2006). Interestingly, our RNA-seq data in S16 cells revealed no significant change in *Mbp* upon overexpression of SOX6-ΔE9. To explain

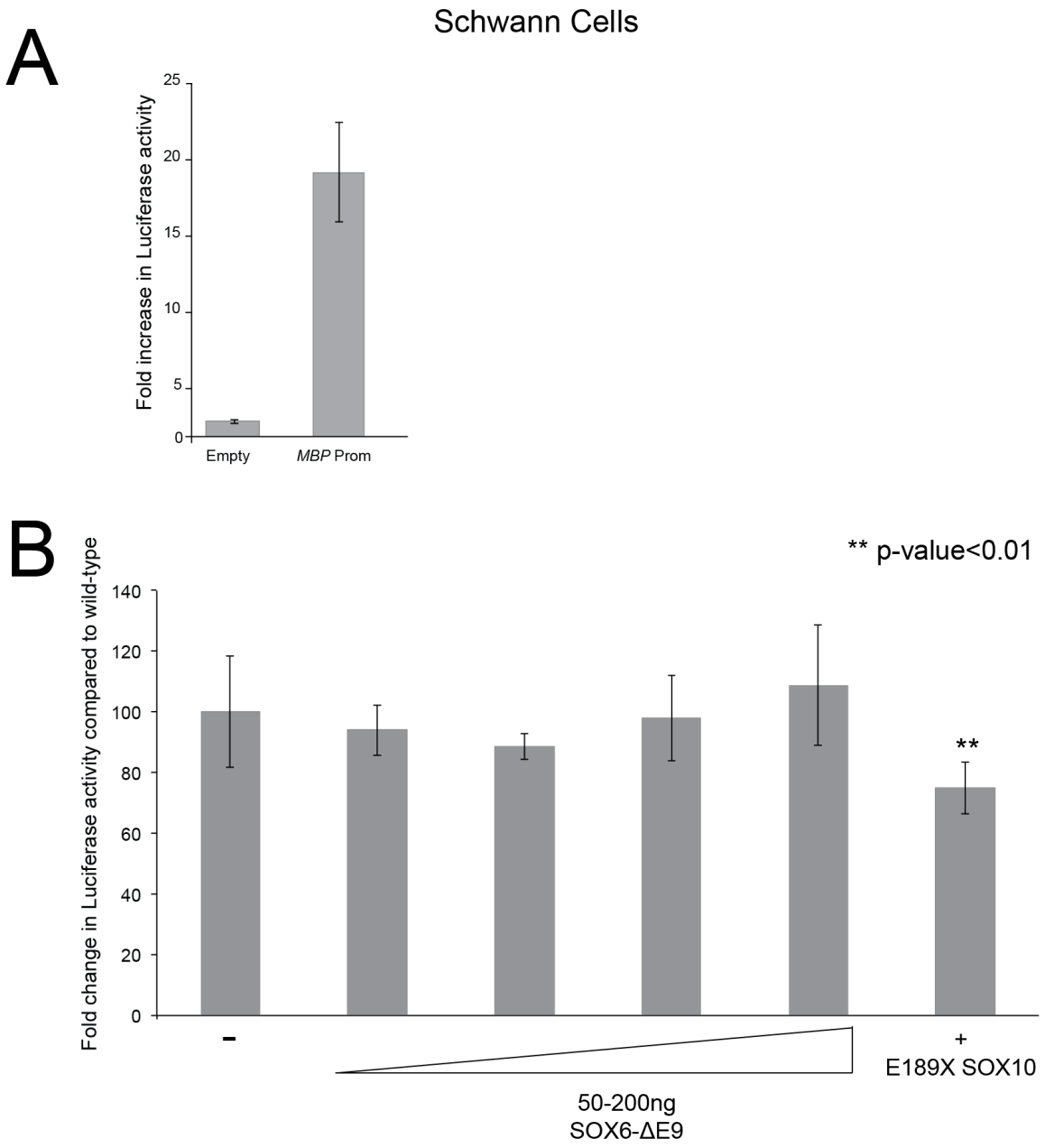


Figure 3.12. SOX6 overexpression does not reduce MBP promoter activity in Schwann cells. MBP promoter was cloned upstream of a c-fos minimal promoter driving luciferase reporter gene and luciferase assays were performed in S16 cells. (A) The luciferase activity of the MBP promoter was compared to an empty vector containing no genomic insert and y-axis shows the fold increase in luciferase expression. (B) MBP promoter activity without (-), in presence of increasing concentrations of SOX6 expression (50,100,150, and 200 ng) and in the presence of dominant-negative (E189XSOX10). The y-axis shows fold change in luciferase activity compared to untreated MBP promoter activity set to 100%. Errors bars in both the panels indicate standard deviations. P-value was calculated using two-tailed student t-test.

these apparently discrepant findings, we performed luciferase assays to test the effect of GFP-SOX6-ΔE9 overexpression on *MBP* promoter activity. We cloned the *Mbp* promoter (same region as previously published (Stolt et al., 2006)) upstream of a *c-fos* minimal promoter driving luciferase reporter gene expression, and performed luciferase assays in S16 cells. The *Mbp* promoter we cloned harbors a conserved SOX10 monomeric site. This region displayed ~25-fold increase in luciferase activity compared to the empty vector, which does not contain a genomic insert (Fig. 3.10A) suggesting this genomic segment possesses regulatory activity in S16 cells. To test if SOX6 is able to repress *MBP* promoter activity *in vitro*, we co-transfected a construct to express GFP-SOX6-ΔE9 or one to express a dominant-negative form of SOX10 (E189X SOX10), which is known to repress luciferase activity of SOX10 genes (Brewer et al., 2014; Fogarty et al., 2016; Gopinath et al., 2016). With increasing concentrations of GFP-SOX6-ΔE9 we did not observe a significant decrease in luciferase activity of the *MBP* promoter. Indeed, overexpression of E189X SOX10 causes a (p-value<0.01) reduction of 26% in the luciferase activity of the *Mbp* promoter. Combined, our *in vitro* data do not support a role for SOX6 in repressing *Mbp* expression in Schwann cells.

3.4 Discussion

Schwann cells develop from the neural crest cells and after their specification go through several developmental stages to become myelinating Schwann cells that ensheath the axons of the peripheral nerves. Gene regulatory networks govern the transition between each stage of Schwann cell development. The transcription factor SOX10 plays a central role in Schwann cell development by regulating genes in a stage specific manner (Svaren and Meijer, 2008). To date, SOX10 has been known to regulate a myelination promoting transcriptional program; however,

to observation that SOX10 is expressed during all stages of Schwann cell development has raised questions on how the myelination program is repressed during non-myelinating stages of cell development. Interestingly, we identified SOX10-responsive elements at loci known to inhibit glial cell differentiation (Chapter 2). SOX10-CCS-13 resides within an intron of the *SOX6* locus and is directly upstream of a non-coding exon (exon 1G) suggesting that this intronic enhancer may be acting as an alternative promoter. Using 5'RACE we were able to identify a TSS directly downstream of SOX10-CCS-13 and we were able to identify reads that mapped to this exon and split reads that go into the downstream exon. Combined, these data indicate that the intronic enhancer is a SOX10-responsive promoter at the *SOX6* locus (Fig. 3.2). Indeed, SOX10 is known to regulate other loci including genes that are mutated in demyelinating peripheral neuropathy, via alternative promoters (Brewer et al., 2014; Fogarty et al., 2016; Hodonsky et al., 2012). These data suggest that SOX10 may regulate loci to produce alternative protein products important for Schwann cell function. SOX10 is selectively expressed in other neural crest derived cells such as melanocytes. Interestingly, *Sox6* expression is reduced (> 6-fold) in melanocytes derived from *Sox10* haploinsufficient mice suggesting that SOX10 may regulate SOX6 expression in melanocytes (Fufa et al., 2015). Currently, the role of SOX6 in melanocytes has not yet been characterized.

Alternative spicing is a post-transcriptional process, which increases the transcript diversity in eukaryotic organisms and recent data suggest that transcription elongation and splicing are coupled (Dujardin et al., 2013). Alternative splicing of *SOX6* mRNA produces isoforms with and without exon 9 and our *in vitro* data indicate that *SOX6* is regulated via an SOX10-reponsive alternative promoter. Interestingly, alternative promoters tend to influence alternative splicing

(Xin et al., 2008). SOX10 interacts with positive transcription elongation factor b (pTEFb) in Schwann cells and could play a role in regulating alternative promoters and alternative splicing. Among chordates, exon 9 is highly conserved among mammals (Fig. 3.4). This is not surprising as there are over 7,000 genes that contain mammalian specific exons (Merkin et al., 2015). Interestingly, *MBP* undergoes alternative splicing to produce at least seven different isoforms and exon 2 is conserved in mammals and absent in non-mammals (Gould et al., 2008; Schweigreiter et al., 2006). MBP isoforms with and without exon 2 display different functional properties. Exon 2 containing protein localized to the cytoplasm whereas the isoform without exon 2 localized primarily to the plasma membrane (Allinquant et al., 1991). Function of exon 9 of SOX6 is currently unknown. To test the importance of exon 9, transgenes expressing the isoforms can be generated separately and the phenotype of the *Sox6* mice (Smits et al., 2001) can be rescued by introducing isoform-specific transgene separately.

We independently cloned both isoforms of mouse SOX6 (SOX6-FL and SOX6-ΔE9) into plasmids, which contain an N-terminal GFP or FLAG-tags. Human and mouse SOX6 proteins display 94.3% sequence identity and the functional domains are 100% conserved (Cohen-Barak et al., 2001). Both SOX6 isoforms display similar DNA-binding and transactivation activities and an N-terminal FLAG tag did not affect the activity of SOX6 (Lefebvre et al., 1998). We transfected HeLa cells with GFP-SOX6-FL and GFP-SOX6-ΔE9 and performed western blot analyses to confirm the presence of the fusion protein product. Our data revealed that SOX6-FL is expressed at lower levels compared to SOX6-ΔE9 (Fig. 3.6A). To ensure that the GFP tag is not having an effect on the expression of SOX6-FL and SOX6-ΔE9, we transiently transfected HeLa cells with FLAG-SOX6-FL and FLAG-SOX6-ΔE9 and performed western blot analyses

(Fig. 3.6B). Consistent with previous observations FLAG-SOX6-FL is expressed at lower levels compared to FLAG-SOX6-ΔE9. There are multiple explanations for reduced expression of SOX6-FL in our transient transfection assays, including: (i) specific degradation of the SOX6-FL transcript produced from the expression plasmid; (ii) specific degradation of the SOX6-FL protein produced from the expression plasmid and/or (iii) variation in transfection efficiency between the two plasmids.

To test the expression of GFP-SOX6-FL and GFP-SOX6-ΔE9 transcripts produced from the plasmids we performed RT-PCR using cDNA libraries prepared from HeLa cells transfected with GFP-SOX6-FL and GFP-SOX6-ΔE9. We were able to detect the presence of both GFP-SOX6-FL and GFP-SOX6-ΔE9 however GFP-SOX6-FL was expressed at lower levels (Fig. 3.7). We next wanted to test the transcript levels of GFP-SOX6-FL and GFP-SOX6-ΔE9 relative to another transcript produced from the expression plasmid. These data revealed lower expression levels of GFP-SOX6-FL suggesting the observed difference could be due to post-translational modifications of GFP-SOX6-FL or variation in transfection efficiency (Fig. 3.8).

SOX6-FL contains exon 9 and UbPred software predicted that the last lysine within exon 9 (residue shown in red in Fig. 3.5B) could be ubiquitinated. Ubiquitination of lysine residues targets proteins for proteosomal degradation (Zencheck et al., 2012) and we hypothesized that GFP-SOX6-FL is targeted by the proteasome leading to degradation of the protein. To test this hypothesis we transiently transfected HeLa cells with GFP-SOX6-FL and GFP-SOX6-ΔE9 and two days after transfection we treated the cells with epoxomicin, a drug known to irresversibly inhibit the proteasome (Meng et al., 1999). We performed western blot analysis using lysates

collected from DMSO and epoxomicin treated cells. If GFP-SOX6-FL is targeted by the proteasome, then we would observe stabilization of GFP-SOX6-FL upon epoxomicin treatment. However, we did not observe stabilization of GFP-SOX6-FL and the levels of GFP-SOX6-ΔE9 were unchanged (Fig. 3.9) indicating that the difference we observe is due to reduced transfection efficiency of GFP-SOX6-FL.

Previous reports suggest SOX6 inhibits myelination in the central nervous system and specifically deleting *Sox6* in developing oligodendrocytes in mice caused premature myelination (Stolt et al., 2006), we hypothesized that SOX6 may compete with SOX10 for binding at myelin loci and inhibit myelination in Schwann cells. To test this hypothesis, we overexpressed GFP-SOX6-ΔE9 in S16 cells. These cells were chosen because they express high levels of myelin genes such as *Mbp* and *Mpz*, and are thus a relevant model to test our hypothesis (Toda et al., 1994) (Law et al., in review). Upon overexpressing GFP-SOX6-ΔE9 we expect to observe a decrease in expression of SOX10 target genes such as *MBP* and *MPZ*. After transfecting S16 cells with GFP- SOX6-ΔE9, flow cytometry was performed to collect GFP-positive and GFP-negative cells. We performed RNA-seq on both populations and identified differentially expressed genes 1,117 genes that were differentially expressed. Surprisingly, we did not observe a significant change in expression of myelin genes; expression of *SOX10, EGR2, MBP, MPZ*, and *PMP22*. We explored the relevance of these findings by performing luciferase assays to test the effect of overexpressing SOX6 on MBP promoter activity. Consistent with our RNA-seq data, we did not observe a significant change on promoter activity with increasing concentration of SOX6. There are several caveats to our RNA-seq experiment. First, S16 cells are an immortalized cell line and may not be amenable to changing their cellular identity upon SOX6

overexpression. Knocking down SOX6 in primary Schwann cells and performing RNA-seq will complement our data. Second, SOX6 is known to heterodimerize with other members of the SOXD family (Lefebvre, 2010) and the factors required for repressing SOX10 target genes are not be expressed at high levels in S16 cells (Law et al., in review).

Our RNA-seq experiments identified several putative SOX6 candidate target genes. Gene ontology analysis revealed several interesting direct or indirect targets of SOX6. Several GO terms important for Schwann cell development were enriched in our data set. This list contains genes such as *SIRT2*, *MXRA8*, *DAG1*, *GFAP*, *NDRG1*, *GSN, SOD1*, *CD9*, and *LRP1* (Table 3.2). Sir-two-homolog 2 (SIRT2) is a NAD(+)-dependent deacetylase involved in myelination and remyelination after injury (Beirowski et al., 2011). Dystroglycan (DAG1) expression is reduced after injury during the axonal degeneration, however, it is upregulated at the onset of the regeneration process and interestingly there is evidence of SOX10 occupancy in rat sciatic nerve at this locus (Srinivasan et al., 2012). Dystroglacan interacts with laminin-2 to aid Schwann cells in wrapping around axons and facilitating remyelination (Masaki et al., 2000). Glial fibrillary acidic protein (GFAP) is a constituent of the cytoskeleton, which is upregulated during nerve damage and regeneration of Schwann cells post injury is impaired in GFAP-null mice. Schwann cell regeneration is delayed in GFAP-null mice and is due to reduced Schwann cell proliferation (Triolo et al., 2006). N-myc downstream-regulated gene 1 (NDRG1) plays a role in terminal differentiation of Schwann cells during the nerve regeneration process post injury and SOX10 ChIP-seq data from rat sciatic nerve reveals SOX10 occupancy at *NDRG1* (Hirata et al., 2004; Srinivasan et al., 2012). Gelsolin (GSN) recruits macrophages to the injury site to facilitates Schwann cell remyelination following nerve damage (Gonçalves et al., 2010). CD9 is a cell

surface glycoprotein which interacts with integrins to promoter Schwann cell migration *in vitro* (Anton et al., 1995). LDL receptor-related protein (LRP1) is predominantly expressed after Schwann cell injury and promotes Schwann cell migration to allow for regeneration (Mantuano et al., 2010). Recent data identified changes in H3K27ac enhancer marks post nerve injury using nerve crush injury in rats. Interestingly, these data reveal H3K27ac peaks at the *SOX6* locus indicating that its expression may be induced post nerve injury (Hung et al., 2015). Our GO analysis indicates up regulation of genes involved in nerve regeneration post injury suggesting a role for SOX6 in Schwann cell dedifferentiation. Based on these findings we propose a model for putative role of SOX6 in the non-myelinating stages of Schwann cell development (Fig. 3.13). Schwann cells are in a proliferative state during the early stages of development and SOX10 may activate the expression of *SOX6* to up-regulate expression of genes important for proliferation and migration. Negative regulators of myelination inhibit the expression of myelin genes at this stage (chapter 2). Once immature Schwann cells exit the cell cycle and become promyelinating, then EGR2 and microRNAs (miRs) inhibit the expression of SOX6 and SOX10 activates expression of myelin genes. Following nerve injury, SOX10 may activate the expression of *SOX6*, which then activates expression of genes that are important for dedifferentiation of Schwann cells. These genes facilitate Schwann cell regeneration by down regulating the myelination program, recruiting macrophages to the injury site, and then remyelinating the axons (Fig. 3.13).

Table 3.4 and 3.5 show the top 10 significantly down- and up-regulated genes. Among these genes some of them have a known role in myelination. COUP transcription factor 1 (*NR2F1*) is the most significantly down-regulated gene. NR2F1 is a nuclear hormone receptor and a

transcription factor. Interestingly eight out of the top 10 down regulated genes have evidence for SOX10 binding in rat sciatic nerve (Srinivasan et al., 2012) these data suggest that SOX6 may inhibit the expression of SOX10 targets. The roles of the top 10 up regulated genes in Schwann cells are currently unknown.

In summary, we identified a previously unreported SOX10-responsive promoter at the *SOX6* locus. Following nerve injury, Schwann cells undergo dedifferentiation followed by proliferation to mediate axonal regrowth. Our RNA-seq data indicates a putative role for SOX6 during development and activating genes involved in Schwann cell regeneration. Better understanding the regulatory mechanisms of SOX6 in Schwann cells will assist efforts toward developing therapeutics for nerve injury and other myelin-related diseases.

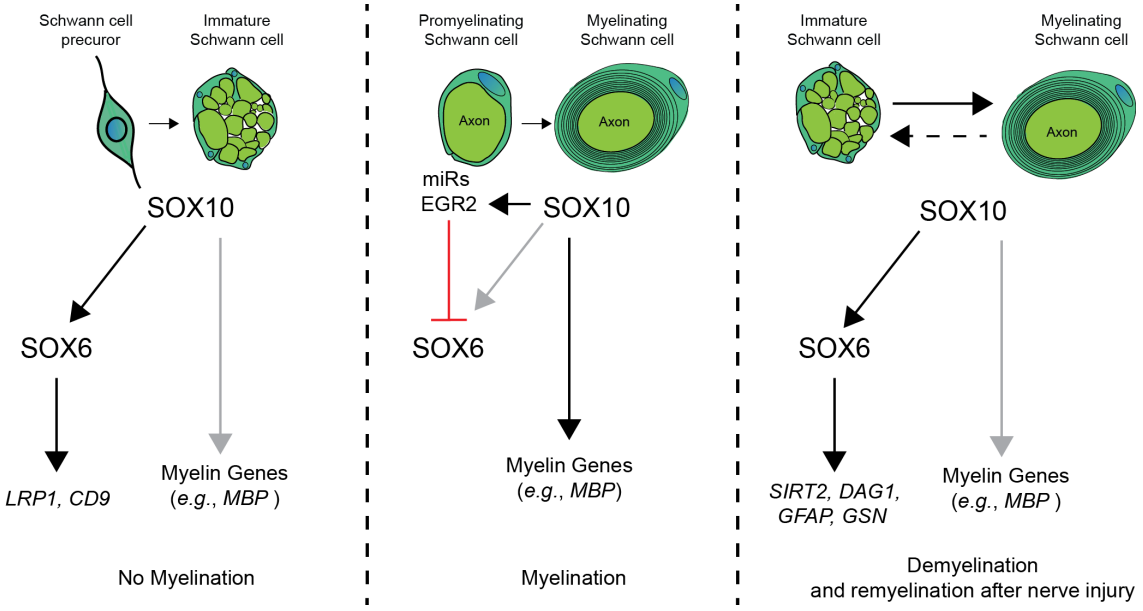


Figure 3.13. A model for the role of SOX6 during non-myelinating stages of Schwann cell development. During early stages of Schwann cell development, SOX10 activates the expression of SOX6, which activates expression of genes important for proliferation and migration (left panel). During activation of the myelination program (pro-myelination; middle panel), EGR2 and micro RNAs (miRs) inhibit the expression of SOX6, which allows SOX10 (and EGR2) to positively regulate the expression of myelin genes. Following nerve trauma, SOX10 activates expression of SOX6 to up-regulate expression of genes critical for demyelination and remyelination of axons of the peripheral nerves. Adapted from Jessen and Mirsky, 2005.

GO biological process	Homo sapiens	Our List	Expected	P-value
Localization	5501	263	159.99	1.42E-15
Negative regulation of biological process	4937	239	143.59	5.58E-14
Negative regulation of cellular process	4448	220	129.37	2.63E-13
Response to stimulus	8033	335	233.64	6.90E-13
Positive regulation of biological process	5607	255	163.08	5.90E-12
Transport	4411	214	128.29	8.19E-12
Regulation of cell differentiation	1681	110	48.89	1.20E-11
Establishment of localization	4520	217	131.46	1.42E-11
Regulation of cell death	1627	107	47.32	2.31E-11
Developmental process	5511	249	160.29	3.97E-11
Positive regulation of cellular process	4984	231	144.96	5.13E-11
System development	4202	204	122.21	5.59E-11
Anatomical structure development	5146	235	149.67	1.36E-10
Regulation of apoptotic process	1495	99	43.48	2.04E-10
Regulation of developmental process	2375	135	69.08	2.05E-10
Regulation of programmed cell death	1510	99	43.92	3.73E-10
Regulation of response to stimulus	4038	195	117.44	5.37E-10
Multicellular organism development	4791	220	139.34	1.00E-09
Response to organic substance	2782	148	80.91	1.55E-09
Regulation of molecular function	3386	170	98.48	1.86E-09
Regulation of localization	2605	141	75.77	1.95E-09
Regulation of signal transduction	3076	158	89.46	3.06E-09
Regulation of multicellular organismal process	2796	147	81.32	4.80E-09
Negative regulation of cell death	945	71	27.49	5.48E-09
Regulation of cell migration	724	60	21.06	7.87E-09
Regulation of cell communication	3363	167	97.81	8.46E-09
Regulation of cell proliferation	1594	99	46.36	9.04E-09
Regulation of signaling	3413	168	99.27	1.46E-08
Tissue development	1659	101	48.25	1.64E-08
Secretion	1067	75	31.03	2.77E-08
Negative regulation of apoptotic process	851	65	24.75	2.82E-08
Cellular response to stimulus	6360	266	184.98	2.84E-08
Response to chemical	4100	191	119.25	3.17E-08
Regulation of biological process	11402	415	331.62	3.98E-08
Regulation of cell motility	778	61	22.63	4.84E-08
Negative regulation of programmed cell death	865	65	25.16	5.63E-08
Regulation of locomotion	848	64	24.66	6.93E-08

Secretion by cell	961	69	27.95	8.90E-08
Cellular process	15042	509	437.49	9.81E-08
Biological regulation	12019	430	349.57	1.24E-07
Biological process	17481	565	508.43	1.52E-07
Regulation of biological quality	3557	169	103.45	2.21E-07
Anatomical structure morphogenesis	2015	112	58.61	2.28E-07
Regulation of cellular component movement	852	63	24.78	2.35E-07
Negative regulation of multicellular organismal process	1053	72	30.63	2.70E-07
Response to external stimulus	1870	106	54.39	2.86E-07
Regulation of cellular process	10750	392	312.66	5.79E-07
Leukocyte activation	881	63	25.62	8.96E-07
Neutrophil degranulation	483	43	14.05	1.84E-06
Cell communication	5414	228	157.46	2.01E-06
Leukocyte degranulation	504	44	14.66	2.05E-06
Neutrophil activation involved in immune response	485	43	14.11	2.08E-06
Regulated exocytosis	686	53	19.95	2.17E-06
Response to endogenous stimulus	1294	80	37.64	2.30E-06
Response to organic cyclic compound	882	62	25.65	2.48E-06
Animal organ development	2991	145	86.99	2.58E-06
Myeloid leukocyte activation	567	47	16.49	2.60E-06
Response to organonitrogen compound	884	62	25.71	2.71E-06
Cell activation involved in immune response	612	49	17.8	3.35E-06
Cellular response to chemical stimulus	2590	130	75.33	3.36E-06
Neutrophil activation	493	43	14.34	3.39E-06
Multicellular organismal process	6703	268	194.95	3.69E-06
Myeloid cell activation involved in immune response	514	44	14.95	3.72E-06
Cell activation	1027	68	29.87	3.80E-06
Granulocyte activation	496	43	14.43	4.06E-06
Neutrophil mediated immunity	497	43	14.46	4.31E-06
Positive regulation of multicellular organismal process	1507	88	43.83	4.33E-06
Regulation of cell adhesion	679	52	19.75	4.35E-06
Myeloid leukocyte mediated immunity	517	44	15.04	4.44E-06
Neurogenesis	1512	88	43.98	5.09E-06
Exocytosis	770	56	22.4	5.79E-06
Regulation of multicellular	1822	100	52.99	6.45E-06

organismal development				
Response to oxygen-containing compound	1446	85	42.06	6.57E-06
Signal transduction	4932	210	143.45	6.72E-06
Leukocyte activation involved in immune response	608	48	17.68	7.98E-06
Negative regulation of signal transduction	1165	73	33.88	8.67E-06
Negative regulation of response to stimulus	1456	85	42.35	9.07E-06
Vesicle-mediated transport	1787	98	51.97	1.03E-05
Signaling	5323	222	154.82	1.04E-05
Response to hormone	784	56	22.8	1.09E-05
Cellular developmental process	3599	164	104.68	1.16E-05
Cardiovascular system development	496	42	14.43	1.28E-05
Cell differentiation	3525	161	102.52	1.50E-05
Blood vessel development	462	40	13.44	1.73E-05
Response to nitrogen compound	997	65	29	1.74E-05
Response to stress	3336	154	97.03	1.90E-05
Vasculature development	486	41	14.14	2.24E-05
Negative regulation of cell communication	1265	76	36.79	2.33E-05
Negative regulation of signaling	1269	76	36.91	2.66E-05
Regulation of catalytic activity	2511	123	73.03	4.83E-05
Regulation of intracellular signal transduction	1793	96	52.15	5.15E-05
Positive regulation of developmental process	1232	73	35.83	8.71E-05
Anatomical structure formation involved in morphogenesis	837	56	24.34	1.04E-04
Cell development	1493	83	43.42	1.30E-04
Blood vessel morphogenesis	380	34	11.05	1.31E-04
Generation of neurons	1418	80	41.24	1.31E-04
Circulatory system development	807	54	23.47	1.91E-04
Negative regulation of metabolic process	2790	130	81.15	3.07E-04
Regulation of cellular protein metabolic process	2522	120	73.35	3.91E-04
Positive regulation of cell communication	1612	86	46.88	4.24E-04
Nervous system development	2230	109	64.86	4.95E-04
Positive regulation of signaling	1619	86	47.09	5.12E-04
Negative regulation of protein metabolic process	1120	66	32.57	5.80E-04
Positive regulation of cell migration	425	35	12.36	5.92E-04

Negative regulation of cellular metabolic process	2514	119	73.12	5.97E-04
Movement of cell or subcellular component	1471	80	42.78	6.08E-04
Regulation of protein metabolic process	2768	128	80.51	6.20E-04
Cellular response to organic substance	2139	105	62.21	7.44E-04
Small molecule metabolic process	1817	93	52.85	7.45E-04
Angiogenesis	295	28	8.58	7.61E-04
Positive regulation of signal transduction	1481	80	43.07	8.02E-04
Lipid metabolic process	1229	70	35.75	8.10E-04
Regulation of neurogenesis	710	48	20.65	8.90E-04
Locomotion	1216	69	35.37	1.16E-03
Immune effector process	1028	61	29.9	1.50E-03
Positive regulation of cell motility	443	35	12.88	1.56E-03
Positive regulation of metabolic process	3278	144	95.34	1.65E-03
Immune system process	2566	119	74.63	1.76E-03
Regeneration	153	19	4.45	1.88E-03
Negative regulation of nitrogen compound metabolic process	2347	111	68.26	1.92E-03
Positive regulation of cellular metabolic process	3056	136	88.88	1.97E-03
Organonitrogen compound metabolic process	5533	218	160.93	2.04E-03
Regulation of nervous system development	801	51	23.3	2.13E-03
Regulation of cell development	827	52	24.05	2.36E-03
Regulation of phosphorus metabolic process	1735	88	50.46	2.58E-03
Positive regulation of locomotion	475	36	13.82	2.76E-03
Positive regulation of cellular component movement	455	35	13.23	2.88E-03
Response to peptide hormone	375	31	10.91	3.08E-03
Positive regulation of response to stimulus	2153	103	62.62	3.41E-03
Regulation of phosphate metabolic process	1721	87	50.05	3.49E-03
Positive regulation of cell differentiation	886	54	25.77	3.57E-03
Aging	266	25	7.74	4.43E-03
Cellular component organization	5306	209	154.32	4.47E-03
Response to peptide	444	34	12.91	4.69E-03

Response to abiotic stimulus	1115	63	32.43	4.90E-03
Negative regulation of cell proliferation	665	44	19.34	5.42E-03
Biological adhesion	875	53	25.45	5.52E-03
Negative regulation of intracellular signal transduction	491	36	14.28	5.93E-03
Leukocyte mediated immunity	738	47	21.46	6.38E-03
Intracellular signal transduction	1616	82	47	6.81E-03
Negative regulation of developmental process	864	52	25.13	8.51E-03
Response to inorganic substance	478	35	13.9	8.73E-03
Negative regulation of cellular protein metabolic process	1061	60	30.86	9.03E-03
Negative regulation of macromolecule metabolic process	2538	115	73.82	9.32E-03
Cell adhesion	870	52	25.3	1.04E-02
Positive regulation of cell proliferation	876	52	25.48	1.26E-02
Regulation of cellular metabolic process	6179	234	179.71	1.35E-02
Gliogenesis	193	20	5.61	1.44E-02
Regulation of response to stress	1436	74	41.77	1.45E-02
Regulation of metabolic process	6667	249	193.91	1.45E-02
Regulation of muscle cell differentiation	176	19	5.12	1.46E-02
Response to oxidative stress	364	29	10.59	1.50E-02
Animal organ morphogenesis	882	52	25.65	1.54E-02
Neuron projection development	624	41	18.15	1.56E-02
Regulation of cellular response to growth factor stimulus	249	23	7.24	1.67E-02
Regulation of sodium ion transport	85	13	2.47	1.68E-02
Cell surface receptor signaling pathway	2289	105	66.57	1.77E-02
Cellular lipid metabolic process	985	56	28.65	1.80E-02
Negative regulation of cell differentiation	652	42	18.96	1.90E-02
Positive regulation of intracellular signal transduction	940	54	27.34	2.05E-02
Regulation of transforming growth factor beta receptor signaling pathway	101	14	2.94	2.11E-02
Cell morphogenesis	632	41	18.38	2.12E-02
Formation of primary germ layer	116	15	3.37	2.15E-02
Glial cell development	87	13	2.53	2.16E-02
Glial cell differentiation	148	17	4.3	2.27E-02

Epithelial cell development	182	19	5.29	2.35E-02
Cellular component organization or biogenesis	5528	212	160.78	2.54E-02
Organic acid metabolic process	997	56	29	2.57E-02
Regulation of cellular response to transforming growth factor beta stimulus	103	14	3	2.64E-02
Chemotaxis	503	35	14.63	2.66E-02
Taxis	503	35	14.63	2.66E-02
Cell migration	875	51	25.45	2.66E-02
Regulation of protein phosphorylation	1356	70	39.44	2.68E-02
Positive regulation of nitrogen compound metabolic process	2945	127	85.65	2.82E-02
Positive regulation of macromolecule metabolic process	3034	130	88.24	2.89E-02
Regulation of transport	1870	89	54.39	2.96E-02
Oxoacid metabolic process	980	55	28.5	3.24E-02
Cell proliferation	691	43	20.1	3.38E-02
Neuron development	763	46	22.19	3.51E-02
Negative regulation of phosphate metabolic process	577	38	16.78	3.54E-02
Negative regulation of phosphorus metabolic process	578	38	16.81	3.68E-02
Regulation of leukocyte migration	171	18	4.97	3.82E-02
Negative regulation of catalytic activity	892	51	25.94	4.50E-02
Cellular localization	2282	103	66.37	4.54E-02
Regulation of phosphorylation	1484	74	43.16	4.58E-02
Regulation of anatomical structure morphogenesis	968	54	28.15	4.72E-02
Gland development	408	30	11.87	4.76E-02
Negative regulation of molecular function	1172	62	34.09	4.88E-02

Table 3.2. Gene ontology analysis of up-regulated genes**.** Overrepresentation analysis of up-regulated genes (adj p-value<0.01) for biological process is shown. The numbers in the column labeled Homo sapiens, Our list, and Expected are all the genes within the human genome associated with the GO term, up-regulated genes from our DeSeq analysis associated with that term, and genes from our list expected by chance for that GO term, respectively.

GO biological process	Homo sapiens	Our List	Expected	P-value
Developmental process	5511	162	97.43	3.14E-09
Anatomical structure morphogenesis	2015	82	35.62	6.71E-09
Response to organic substance	2782	100	49.18	1.77E-08
Response to chemical	4100	129	72.48	4.03E-08
Anatomical structure development	5146	149	90.98	2.07E-07
Response to endoplasmic reticulum stress	247	24	4.37	2.96E-07
Cellular response to organic substance	2139	81	37.82	3.53E-07
Cellular protein metabolic process	3747	117	66.24	1.13E-06
Multicellular organism development	4791	138	84.7	2.89E-06
Response to stimulus	8033	201	142.01	3.60E-06
Circulatory system development	807	42	14.27	6.00E-06
Cell morphogenesis involved in differentiation	500	32	8.84	6.01E-06
Cellular amino acid metabolic process	314	25	5.55	6.75E-06
System development	4202	124	74.29	9.07E-06
Cell morphogenesis	632	36	11.17	1.08E-05
Organonitrogen compound metabolic process	5533	151	97.82	1.25E-05
Macromolecule modification	3440	107	60.82	1.29E-05
Biological process	17481	348	309.05	1.40E-05
Cellular response to chemical stimulus	2590	87	45.79	2.23E-05
Enzyme linked receptor protein signaling pathway	716	38	12.66	2.46E-05
Protein metabolic process	4464	128	78.92	2.62E-05
Animal organ development	2991	96	52.88	2.64E-05
Cellular component morphogenesis	726	38	12.83	3.56E-05
Renal system development	271	22	4.79	5.07E-05
Cellular component organization	5306	144	93.8	6.11E-05
Axon guidance	227	20	4.01	6.66E-05
Nervous system development	2230	77	39.42	7.46E-05
Neuron projection guidance	229	20	4.05	7.68E-05
Urogenital system	305	23	5.39	8.78E-05

development				
Cellular protein modification process	3158	98	55.83	9.60E-05
Protein modification process	3158	98	55.83	9.60E-05
Neuron development	763	38	13.49	1.31E-04
Cellular response to stimulus	6360	163	112.44	1.89E-04
Cell morphogenesis involved in neuron differentiation	403	26	7.12	2.15E-04
Axonogenesis	348	24	6.15	2.22E-04
Axon development	378	25	6.68	2.51E-04
Cellular component organization or biogenesis	5528	146	97.73	2.82E-04
Cellular process	15042	311	265.93	3.01E-04
Neuron projection development	624	33	11.03	3.12E-04
Cell development	1493	57	26.39	3.55E-04
Response to topologically incorrect protein	181	17	3.2	3.64E-04
Kidney development	254	20	4.49	4.10E-04
Response to unfolded protein	161	16	2.85	4.27E-04
Heart development	477	28	8.43	4.36E-04
Nitrogen compound metabolic process	8629	204	152.55	4.55E-04
Neuron projection morphogenesis	454	27	8.03	5.88E-04
Plasma membrane bounded cell projection morphogenesis	458	27	8.1	6.99E-04
Locomotion	1216	49	21.5	7.38E-04
Blood vessel development	462	27	8.17	8.30E-04
Cell projection morphogenesis	462	27	8.17	8.30E-04
Cellular macromolecule metabolic process	7017	173	124.05	8.58E-04
Cardiovascular system development	496	28	8.77	9.61E-04
IRE1-mediated unfolded protein response	58	10	1.03	1.10E-03
Neuron differentiation	949	41	16.78	1.57E-03
Primary metabolic process	9151	211	161.78	1.74E-03
Endoplasmic reticulum unfolded protein response	115	13	2.03	1.85E-03
Cell part morphogenesis	483	27	8.54	1.97E-03
Regulation of developmental process	2375	76	41.99	2.14E-03
Vasculature development	486	27	8.59	2.22E-03

Transmembrane receptor protein tyrosine kinase signaling pathway	519	28	9.18	2.37E-03
Cellular metabolic process	9022	208	159.5	2.53E-03
Macromolecule metabolic process	7766	185	137.29	2.65E-03
Cellular response to topologically incorrect protein	140	14	2.48	2.73E-03
Regulation of cellular protein metabolic process	2522	79	44.59	2.75E-03
Movement of cell or subcellular component	1471	54	26.01	2.91E-03
Head development	726	34	12.83	3.18E-03
Negative regulation of cellular process	4448	120	78.64	3.44E-03
Cellular response to unfolded protein	122	13	2.16	3.57E-03
Organic substance metabolic process	9511	216	168.14	3.74E-03
Regulation of protein metabolic process	2768	84	48.94	3.95E-03
Chemotaxis	503	27	8.89	4.28E-03
Taxis	503	27	8.89	4.28E-03
Plasma membrane bounded cell projection organization	1039	42	18.37	6.23E-03
Localization	5501	140	97.25	6.68E-03
Cellular response to stress	1602	56	28.32	7.96E-03
Regulation of GTPase activity	722	33	12.76	8.07E-03
Brain development	688	32	12.16	8.24E-03
ER to Golgi vesicle-mediated transport	178	15	3.15	8.66E-03
Negative regulation of biological process	4937	128	87.28	1.03E-02
Regulation of hydrolase activity	1413	51	24.98	1.05E-02
Response to endogenous stimulus	1294	48	22.88	1.07E-02
Cell projection organization	1066	42	18.85	1.19E-02
Cellular response to endogenous stimulus	1067	42	18.86	1.22E-02
Response to growth factor	501	26	8.86	1.29E-02
Regulation of cellular metabolic process	6179	152	109.24	1.29E-02
Multicellular organismal process	6703	162	118.5	1.31E-02

Regulation of cell communication	3363	95	59.45	1.37E-02
Protein localization	1927	63	34.07	1.42E-02
Cellular response to growth factor stimulus	472	25	8.34	1.45E-02
Cell-substrate adhesion	164	14	2.9	1.71E-02
Response to stress	3336	94	58.98	1.78E-02
Positive regulation of hydrolase activity	976	39	17.25	2.10E-02
Nephron development	122	12	2.16	2.24E-02
Glycoprotein metabolic process	391	22	6.91	2.38E-02
Regulation of macromolecule metabolic process	6139	150	108.53	2.44E-02
Metabolic process	9949	220	175.89	2.54E-02
Regulation of metabolic process	6667	160	117.87	2.57E-02
Regulation of signaling	3413	95	60.34	2.60E-02
Cellular protein localization	1338	48	23.65	2.64E-02
Vesicle coating	65	9	1.15	2.76E-02
Nitrogen compound transport	1671	56	29.54	2.85E-02
tRNA aminoacylation for protein translation	49	8	0.87	2.98E-02
Tube development	560	27	9.9	3.13E-02
cellular macromolecule localization (GO:0070727)	1348	48	23.83	3.23E-02
Macromolecule localization	2239	69	39.58	3.24E-02
Positive regulation of GTPase activity	666	30	11.77	3.41E-02
Vesicle targeting, to, from or within Golgi	68	9	1.2	3.96E-02
Regulation of biological process	11402	244	201.58	4.51E-02
tRNA aminoacylation	52	8	0.92	4.58E-02
Branched-chain amino acid metabolic process	24	6	0.42	4.71E-02
Cardiac chamber development	155	13	2.74	4.78E-02

Table 3.3. Gene ontology analysis of down-regulated genes. Overrepresentation analysis of down-regulated genes (adj p-value<0.01) for biological process with GO ids (in parenthesis) is shown. The numbers in the column labeled Homo sapiens, Our list, and Expected are all the genes within the human genome associated with the GO term, down-regulated genes from our DeSeq analysis associated with that term, and genes from our list expected by chance for that GO term, respectively.

Gene Name	log2FoldChange	adj P-value	SOX10 ChIP-seq peak (-/+) 10Kb
COUP transcription factor 1 (Nr2f1)	-0.643495256	1.50E-17	chr2:5,547,324-5,547,758
Transcription factor AP-2-beta (Tfap2b)	-0.59948108	2.06E-14	-
Potassium voltage-gated channel subfamily D member 3 (Kcnd3)	-0.677286517	1.53E-11	chr2:227,554,246-227,554,782
WW domain-containing protein 2 (Wwc2)	-0.402951757	3.36E-10	chr16:47,091,452-47,092,740
ETS translocation variant 6 (Etv6)	-0.577489698	8.94E-10	chr4:230,430,522-230,431,508
Chromodomain-helicase-DNA-binding protein 6 (Chd6)	-0.549825307	9.41E-10	chr3:163,324,522-163,325,039
Cysteine-rich angiogenic inducer 61 (Cyr61)	-0.516195683	1.03E-09	chr2:270,058,231-270,058,627
CUGBP Elav-like family member 2 (Celf2)	-0.563180159	1.37E-09	chr17:76,894,056-76,894,589
Regulatory factor X 5 (Rfx5)	-0.447530029	1.68E-09	-

Table 3.4 Top 10 genes significantly down-regulated upon SOX6 overexpression in S16 cells**.** log2fold change is calculated relative to GFP negative cells and a negative values correspond to lower expression in GFP positive cells. Adjusted (adj) P-value calculated by DeSeq using Benjamini-Hochberg method for multiple testing. The last column indicates the rn5 co-ordinates for SOX10 ChIP-seq peak in rat sciatic nerve (Srinivasan et al., 2012)

Gene Name	log2FoldChange	adj P-value	SOX10 ChIP-seq peak (-/+) 10Kb
Lysyl Oxidase Like 4 (Loxl4)	1.944365168	1.53E-56	-
Glutathione S-transferase P(Gstp1)	1.437891421	1.50E-52	chr1:226,164,673-226,165,173
Wnt7a	2.425242064	1.19E-51	-
Interleukin-23 subunit alpha(Il23a)	1.814388837	3.07E-51	-
Rearrange during transfection (Ret)	2.287918881	6.80E-43	chr4:216,155,302-216,155,832
Cysteine-rich secretory protein LCCL domain-containing 2 (Crispld2)	2.273841049	1.26E-40	-
Slit homolog 1 protein (Slit1)	1.586861198	2.82E-40	-
Rab-like protein 2A (Rabl2a)	1.426088347	1.18E-37	-
Dual specificity protein phosphatase 1 (Dusp1)	1.139397765	8.11E-35	-
Keratin, type II cytoskeletal 7 (Krt7)	2.097125454	8.11E-35	-

Table 3.5 Top 10 genes significantly up-regulated upon SOX6 overexpression in S16 cells. log2fold change is calculated relative to GFP negative cells and a positive values correspond to higher expression in GFP positive cells. Adjusted (adj) P-value calculated by DeSeq2 using Benjamini-Hochberg method for multiple testing. The last column indicates the rn5 co-ordinates for SOX10 ChIP-seq peak in rat sciatic nerve (Srinivasan et al., 2012)

Chapter 4

Conclusions and Future Directions

4.1 Summary

Compact myelin is a lipid rich layer formed by wrapping of glial cells around the axons of all jawed vertebrates. Schwann cells and oligodendrocytes myelinate axonal segments in the peripheral and the central nervous systems, respectively. The presence of myelin was first observed in ancient gnathostomes (jawed vertebrates), the placoderms (Zalc et al., 2008). Interestingly, neural crest cells are unique to vertebrates and differentiate into diverse cell types including Schwann cells and craniofacial structures such as the jaws (Sauka-Spengler and Bronner-Fraser, 2008) and co-evolution of myelin and jaws has been speculated (Zalc et al., 2008). Myelination results in faster responses to stimulus due to increased propagation of action potentials and has significantly contributed to the rapid predatory and escape behaviors in jawed vertebrates (Sauka-Spengler and Bronner-Fraser, 2008; Schweigreiter et al., 2006; Zalc, 2016). During early vertebrate evolution, the lamprey lineage (jawless vertebrates) split from the other vertebrates before the emergence of hinged jaws. Lampreys express migratory neural crest cells and Schwann cells wrap their axons but they do not form compact myelin (Gelman et al., 2009; Medeiros, 2013). Interestingly, sequencing data from the lamprey genome identified the presence of various myelin proteins (Smith et al., 2013).

The SOXE family of transcription factors comprising SOX8, SOX9, and SOX10 are important for neural crest development and specification and interestingly, lampreys express three paralogs of SoxE genes (*SoxE1*, *SoxE2*, and *SoxE3*) (McCauley and Bronner-Fraser, 2006). However, only *SoxE3* shares sequence homology with vertebrate *SOX9* and there is little sequence similarity between the other two SOXE paralogs and *SOX8* and *SOX10* (McCauley and Bronner-Fraser, 2006). Neofunctionalization of the SOXE proteins and appearances of cis-acting regulatory sequences may have contributed to the expression of myelin genes important for formation of compact myelin in jawed vertebrates. In this thesis, I have explored the gene regulatory network that controls formation of compact myelin in vertebrates.

Schwann cells arise from trunk neural crest cells and myelinate motor and sensory axons of the peripheral nervous system. The formation of myelin around axons is a multistep process in which the neural crest cells differentiate to form Schwann cell precursors. These cells transition into an immature Schwann cell stage before becoming either myelinating or non-myelinating Schwann cells. Schwann cells that are associated with single large axons form a layer of compact myelin (Jessen and Mirsky, 2005). The importance of Schwann cells is underscored by their involvement in demyelinating diseases and identifying regulatory pathways governing development and function of Schwann cells will aid in a better understanding of pathology of these diseases (Svaren and Meijer, 2008). Positive and negative transcriptional regulators complement each other to ensure that formation of compact myelin occurs in a timely manner. Misregulation of these pathways can lead to hypomyelination (Svaren and Meijer, 2008) or hypermyelination (Noseda et al., 2013; Triolo et al., 2012) and are associated with the pathology of demyelinating diseases.

The SRY-related HMG box 10 (SOX10) is a transcription factor critical for Schwann cell development, function, and maintenance. SOX10 expression starts as the neural crest cells start migrating and continues to be expressed through the entirety of the Schwann cell lineage (Bremer et al., 2011; Kuhlbrodt et al., 1998). So far, SOX10 is known to direct a promyelinating cascade in Schwann cells by activating expression of genes important for myelin formation (Svaren and Meijer, 2008). Importantly, human mutations in *SOX10* and SOX10 target genes have been implicated in demyelinating peripheral neuropathies (Inoue et al., 2004; Kulshrestha et al., 2017; Roa et al., 1996; Warner et al., 1998; Weterman et al., 2010). The main focus of my dissertation is to expand our current knowledge about the role of SOX10 in peripheral nerve myelination by identifying novel SOX10 target genes, which will help dissect the regulatory network controlled by SOX10 in Schwann cells.

Previously, a number of direct SOX10 response elements were identified on a gene-by-gene basis. However, genome-wide approaches were needed to discover additional SOX10 gene targets in a less-biased and a high-throughput manner. Moreover, these studies would also facilitate identification of novel genes with no known role in myelination and studying the function of these novel genes in myelination will vastly improve our understanding of gene regulatory networks important for Schwann cell biology.

In chapter 2, I presented a computational pipeline using sequence conservation to identify putative SOX10 consensus sequences in the human genome. We used this approach for two reasons: (i) if non-coding regions are conserved among diverse species, then they could harbor important biological function. This technique has been successfully implemented to identify

functional regulatory elements (Antonellis et al., 2008; Hardison et al., 1997); and (ii) SOX10 binding sites conserved in human, mouse, and chicken have been identified and functionally validated (Antonellis et al., 2008; Gokey et al., 2012; Hodonsky et al., 2012). SOX10 binds to well-characterized consensus sequences ('ACACA', 'ACAAG','ACAAA', or 'ACAAT') as a monomer or as a dimer when two consensus sequences are oriented in a head-to-head manner (Peirano and Wegner, 2000; Srinivasan et al., 2012). To identify conserved SOX10 consensus sequences we generated two datasets: (i) we compared the genomes of human, mouse, and chicken, and identified over two million conserved regions that are five base pairs or longer; and (ii) we identified all SOX10 consensus sequences in the human genome (33 million monomeric and ~549,000 dimeric consensus sequences). To prioritize elements for functional validation and to identify candidate genes, we focused on conserved dimeric SOX10 consensus sequences that resided within the gene body (intronic) or proximal promoter elements, and this revealed 238 elements within 160 genes. The full panel of genes regulated by SOX10 at every developmental stage of myelination is not completely understood so we further prioritized by selecting loci with a known or predicted role in myelination and we identified 57 genomic elements for functional studies. We cloned the genomic regions surrounding the SOX10 consensus sequence for each element upstream of a minimal promoter driving luciferase reporter gene expression (Antonellis et al., 2008) and performed luciferase assays using cultured rat Schwann (S16) cells, which express endogenous SOX10 (Hodonsky et al., 2012). Seven out of 57 regions displayed greater than 2.5-fold increase in luciferase activity compared to an empty vector, which does not contain a genomic insert. Surprisingly, only ~12% of the regions we tested were active *in vitro* in Schwann cells and there could be multiple possibilities for this result. There are inherent drawbacks to using luciferase assays to test enhancer activity, which could contribute to a limited

set of regions being active. Although the regions we amplified and cloned were chosen based on general conservation, we may need to clone longer fragments for some of the regions to show greater enhancer activity and some genomic elements we tested may harbor sequences, which allow binding of a repressor, that could neutralize the effect of SOX10 thus resulting in low luciferase activity. Alternatively, SOX10 could be repressing the luciferase activity of some of the regions (Cruz-Solis et al., 2009). Another possibility could be the GC content of the intervening sequence within a dimeric SOX10 site. When we compared the GC content of the intervening sequence of an active and an inactive dimeric site, we found high GC content correlated with high luciferase activity and these data are consistent with previous active dimeric SOX10 response elements (Antonellis et al., 2008; Brewer et al., 2014; Gokey et al., 2012; Hodonsky et al., 2012; Peirano et al., 2000). Therefore, the GC content of the intervening sequence can be used as a criterion for future predictions of dimeric SOX10 sites.

Our computational strategy combined with functional validation revealed SOX10-responsive intronic enhancers at *SOX5*, *SOX6,* and *NFIB*. To further utilize our computational predictions we compared our conserved SOX10 consensus sequences to (i) SOX10 ChIP-seq data generated from sciatic nerve (Srinivasan et al., 2012), which became available to us as we were developing our computational pipeline; and (ii) we generated genome-wide chromatin accessibility by identifying DNase hypersensitive sites using cultured rat Schwann cells (S16 cells). This analysis revealed 214 genomic segments that reside within 191 loci. We performed gene ontology on these loci and interestingly three out of the top 10 biological processes were related to negative regulation of glial cells. Upon further functional validation, our data indicate that SOX10 regulates the expression of *NOTCH1*, *HMGA2*, *HES1*, *MYCN*, *ID2*, *ID4, SOX5, and SOX6*.

These data revealed a surprising role for SOX10 where it activates negative regulators of myelination and based on our data we proposed a model for SOX10 in maintaining a premyelinating state (Gopinath et al., 2016).

There are caveats to identifying SOX10 response elements based on sequence analysis alone and genome-wide approaches such as ChIP-seq have identified putative SOX10 targets in a developmental stage specific manner (Srinivasan et al., 2012). The work I have presented indicates that the best way to identify functional SOX10 elements is by combining these diverse data sets. This increased robustness and has led us to identify *bona fide* SOX10 targets. The comparative genomic datasets we developed will be beneficial to the scientific community. The genomic segments conserved in human, mouse, and chicken can be used to identify putative consensus sequences for other transcription factors important for other cell types. Investigators can also use these data to identify SOX10 target genes in other SOX10-positive cells types such as oligodendrocytes, melanocytes, and enteric nervous system neurons.

In chapter 3, we validated that the SOX10-responsive intronic enhancer at the *SOX6* locus resides at an alternative promoter (Gopinath et al., 2016). SOX10 is known to regulate other loci via alternative promoters suggesting that the alternative protein products may be important for Schwann cell function (Brewer et al., 2014; Fogarty et al., 2016; Hodonsky et al., 2012). Interestingly, *SOX6* mRNA undergoes alternative splicing and produces isoforms with (SOX6-FL) and without exon 9 (SOX6-ΔE9), and this exon is conserved only in mammals. To understand the role of both these isoforms in Schwann cells, we cloned SOX6-FL and SOX6-ΔE9 into a plasmid that contains an N-terminal GFP tag. We overexpressed both these isoforms

in HeLa cells to detect the expression of the fusion protein. HeLa cells were chosen as they have high transfection efficiency. Our *in vitro* data indicate that GFP-tagged SOX6-ΔE9 displays better transfection efficiency and we chose this isoform for follow-up functional studies and both isoforms are known to display similar transcriptional activities (Lefebvre et al., 1998).

The function of SOX6 has not yet been studied in Schwann cells; however, SOX6 was reported to inhibit myelination in oligodendrocytes. The SOX family of transcription factors bind similar consensus motifs (Prior and Walter, 1996) and these data suggest a scenario where SOX6 may bind SOX motifs at the myelin loci prior myelination to inhibit their expression and at the onset of myelination SOX6 gets displaced from these motifs allowing SOX10 to bind and activate these genes. The counter acting SOX proteins would ensure timely activation of myelination. Our original hypothesis was that SOX6 inhibits expression of myelin genes by competing with SOX10 and to further explore its role in Schwann cells and identify a broader panel of genes regulated by SOX6 in Schwann cells we performed RNA sequencing.

S16 cells express high levels of SOX10 and other myelin genes such as *MBP* and *MPZ* (Toda et al., 1994)(Law et al., in review). We overexpressed GFP-SOX6-ΔE9 in S16 cells to identify potential SOX6 target genes; we chose this isoform because it is more robustly expressed compared to SOX6-FL and both isoforms are known to display similar transcriptional activities (Lefebvre et al., 1998). After transfecting the S16 cells with GFP-SOX6-ΔE9, flow cytometry was performed to collect GFP-positive and GFP-negative cell populations, and RNA was isolated from three independent experiments. Libraries were prepared from these samples and sequenced to identify differentially regulated transcripts.

Interestingly, our RNA-seq data revealed no change in *MBP* expression upon SOX6 overexpression and we performed luciferase assays to test if overexpressing SOX6 would cause reduced regulatory activity of the *MBP* promoter. We also did not observe significant changes in expression of other SOX10-target genes known to be important for myelination such as *SOX10*, *EGR2*, *OCT6*, *MPZ*, *GJB1*. These data suggest that SOX6 does not directly or indirectly inhibit the expression of myelin genes in S16 cells. Some potential pitfalls of our approach are that, while S16 cells express high levels of *Sox10*, *Egr2*, *Mpz*, *Mbp* (Law et al., in review), it is an immortalized cell line and may not accurately represent Schwann cells *in vivo*. Additionally, S16 cells do not have axonal contact and may have lost their ability to revert to an immature stage. Interestingly, our gene ontology enrichment analysis provided insights into putative roles for SOX6 in Schwann cells. Specifically, our data indicate that SOX6 may up-regulate the expression of genes known to play critical role in Schwann cell proliferation, migration and regeneration such as *SIRT2*, *MXRA8*, *DAG1*, *GFAP*, *NDRG1*, *SOD1, CD9*, and *LRP1*.

4.2 Future Directions

While my dissertation has greatly improved our understanding of Schwann cell biology by revealing a novel role for SOX10 in maintaining Schwann cells in a premyelinating state and identifying SOX6 as a putative regulator during the non-myelinating stages of Schwann cell development, there are many questions that need to be addressed.

4.2.1 Better defining the SOX10 consensus sequences

During my thesis research, we developed a computational approach to identity conserved SOX10 consensus sequences and we tested the regulatory activity of a subset of prioritized elements

using luciferase reporter assays. Although each element harbored a highly conserved dimeric SOX10 site, only ~12% of the regions were active in cultured Schwann cells. A lot of factors could be contributing to this result and one possibility could be the intervening sequence within the dimeric site. We observed that the GC content of the intervening sequence of an active dimeric SOX10 site was high (~60%) compared to the inactive sites. This is consistent with an analysis that was performed on ~400 ChIP-seq data generated for 119 transcription factors, which revealed that the majority transcription factors bind GC-rich sequences (Wang et al., 2012). To test the necessity of high GC content of the intervening sequence, it would be worth introducing the intervening sequence from an active site to an inactive site and vice versa and performing luciferase reporter assays. If high GC content indeed plays a role in determining the activity of the SOX10 site then we would observe an increase in the luciferase activity when we increase the GC content of a previously inactive site. This will improve our understanding of the requirements of a dimeric SOX10 site.

4.2.2 The role of SOX10 in peripheral nerve injury

Schwann cells display remarkable plasticity and following nerve injury they dedifferentiate into a state that closely resembles the immature Schwann cell state (Quintes and Brinkmann, 2017). Schwann cells promote repair by: (i) downregulating the myelination program; (ii) upregulating negative inhibitors of myelination; (iii) recruiting macrophages to the injury site; and (iv) clearing myelin debris and regenerating tracks for axonal regrowth followed by remyelination (Jessen and Mirsky, 2016). The transcription factors c-jun and SOX2 are known to reprogram Schwann cells and facilitate repair post injury (Arthur-Farraj et al., 2012; Roberts et al., 2017b). SOX10 expression is unchanged following nerve injury (Arthur-Farraj et al., 2012). The

transcriptional program driving the dedifferentiation process is not fully characterized and the role of SOX10 in regulating genes post nerve injury has not yet been explored. Interestingly, *NOTCH1*, one of the genes we identified as SOX10 target, is important for Schwann cell demyelination after injury (Gopinath et al., 2016; Jessen and Mirsky, 2016). In the data set generated in chapter 2, we found evidence for SOX10 binding and the presence of conserved SOX10 consensus sequences at additional loci known to be important for Schwann cell regeneration following injury: *Ets1*, *Megf10*, *Ataxin10* (*Atxn10),* and *SOX2-OT*. ETS1 is a transcription factor induced following Schwann cell injury and is important for Schwann cell survival (Hung et al., 2015; Parkinson et al., 2002). *Megf10* is up regulated during injury and is involved in phagocytosis of myelin debris to facilitate repair (Brosius Lutz et al., 2017). *Atxn1* is up regulated in rat sciatic nerve after injury and may be involved in Schwann cell proliferation (Cheng et al., 2013). *SOX2-OT* is a multi-exon long non-coding (lnc) RNA which contains the *SOX2* gene within one intron and SOX2-OT expression is up regulated in response to Schwann cell injury (Arthur-Farraj et al., 2017). These data suggest a role for SOX10 in activating the expression of these genes post nerve injury. Knockout *Sox10* in vitro in S16 cells and test for reduced expression of *Ets1*, *Megf10*, *Ataxin10* (*Atxn10),* and *SOX2-OT*. Luciferase assays can be performed by cloning the genomic segments containing the conserved SOX10 consensus sequence upstream of a luciferase reporter gene and test for their regulatory activity in Schwann cells. Delete each SOX10 binding site and test for a decrease in enhancer activity in cultured Schwann cells using luciferase assays. These studies will lead to the identification of a novel role for SOX10 in nerve regeneration following injury.

4.2.3 Identifying target genes of transcription factors that inhibit myelination

In addition to SOX6 we identified other transcription factors such as SOX5, SOX6, NOTCH1, HMGA2, HES1, MYCN, ID4, and ID2 and identifying their gene targets will provide a comprehensive list of factors that control each stage of Schwann cell development. Interestingly SOX5 competes with SOX10 to inhibit melanocyte differentiation and melanocytes are neural crest derived cells (Stolt et al., 2008). SOX5 is expressed in migratory neural crest cells (Perez-Alcala et al., 2004) and continues to be expressed during the early stages of melanocyte development (Stolt et al., 2008). SOX5 binds to SOX10 response elements at the *MITF* and *DCT* promoters and recruits co-repressor CtBP2 to inhibit their activity. MITF and DCT are both important for differentiation of melanocytes (Stolt et al., 2008). SOX5 is also known inhibit myelination in the central nervous system (Stolt et al., 2006) and SOX5 may act in a similar manner in Schwann cells. Our qRT-PCR data indicates that SOX5 is developmentally regulated in Schwann cells. To identify SOX5 target loci, SOX5 can be knocked out in primary Schwann cells and RNA-seq can be performed. Performing SOX5 ChIP-seq using a developing sciatic nerve will complement these studies. We expect SOX5 to inhibit myelination by inhibiting expression of SOX10 target genes such as *EGR2, MBP* and *MPZ*.

4.2.4 The role of SOX6 in Schwann cells

Negative regulators of myelination are up regulated post nerve injury and facilitate dedifferentiation of Schwann cells (Jessen and Mirsky, 2016). Based on our RNA-seq data SOX6 could be involved in the dedifferentiation process. A recent study identified global changes in H3K27ac enhancer marks post nerve injury and there are two H3K27ac peaks in intron 9 of *Sox6* suggesting that expression may be induced post injury (Hung et al., 2015). To

directly test if SOX6 is up regulated following nerve injury, we could perform qRT-PCR using a cDNA library prepared from RNA isolated from sham and injured rodent sciatic nerve. We would expect increased *Sox6* expression upon injury. Additionally, we could specifically test if SOX10-regulated *SOX6* transcript containing exon 1G is up regulated post injury. These data may suggest that SOX10 activates the expression of *SOX6* following never damage.

If indeed we observe up regulation of SOX6 expression post nerve injury then we can identify genome-wide occupancy of SOX6, performing SOX6 ChIP-seq using injured sciatic nerve from a rodent model. These data would complement our RNA-seq data and will allow identification of direct SOX6 target genes.

Homozygous *Sox6* null mice are born alive but die within hours after birth and heterozygous mice were identical to their wild-type littermates (Smits et al., 2001). Deleting *Sox6* specifically in oligodendrocytes causes premature expression of *Mbp* (Stolt et al., 2006). Terminal differentiation of oligodendrocyte precursor cells starts at the end of embryogenesis in mice (E18.5), however *Sox6*$^{-/-}$ mice undergo premature differentiation of OLs (E15.5) (Stolt et al., 2006). Our data suggest that SOX10 regulates *SOX6* transcript containing exon 1G and the alternative promoter contains a highly conserved SOX10 binding site (chapter 3). To further characterize the role of SOX6 in Schwann cells and to test the importance of the SOX10-regulated *SOX6* transcript, it would be interesting to knock out the *Sox6* transcript containing exon 1G in mice using CRISPR-Cas9 approaches. Guide RNAs can be designed to target the conserved SOX10 binding site. Based on our *in vitro* data deleting SOX10-regulated *Sox6* transcript may cause defects in Schwann cells and other SOX10-positive cells such as

melanocytes and oligodendrocytes. If we don't observe any defects, it could be due to SOX5 compensating for the loss of SOX6. SOX5 can effectively compensate for the loss of SOX6 in chondrocytes (Liu and Lefebvre, 2015; Smits et al., 2001).

A conditional *Sox6* mouse model can be generated which harbors lox P sites flanking exon 1G (*Sox6* exon1G$^{fl/fl}$). This mouse model can be crossed to various Cre lines to specifically delete *Sox6* in a temporal manner during Schwann cells development. Crossing *Sox6* exon1G$^{fl/fl}$ with *Dhh*::Cre, will specifically delete *Sox6* exon1G transcript in the Schwann cell precursor stage. Cre recombinase under the control of *Dhh* regulatory element will express around E12.5 in mice (Jaegle et al., 2003). If SOX6 is a negative regulator of Schwann cell myelination then specifically deleting its expression in Schwann cell precursors would result in precocious myelination and we would observe ectopic expression of *Mbp*. If these mutant mice develop normally and we do not observe any phenotype during development then these mutant mice can also be used to test the importance of SOX6 post nerve injury. Nerve crush injuries can be performed in wild type and mutant mice and regeneration defects might be observed in *Sox6* mutant mice.

To study the role of SOX6 during neural crest development *Sox6* exon1G$^{fl/fl}$ can be crossed with SOX10-MCS4 Cre (S4F Cre) (Stine et al., 2009). SOX10-MCS4 is a distal enhancer ~28kb upstream of *SOX10* and is known to direct reporter gene expression in neural crest cells in zebrafish and mice (Antonellis et al., 2008). Using these mutant mice we can ask several exciting questions such as: is SOX6 important for neural crest specification, does it play a role in

migration of neural crest cells, how does loss of SOX6 in neural crest cells affect Schwann cell development?

SOX10 expression starts in the premigratory neural crest cells and persists throughout the life of Schwann cells, which would suggest expression of SOX6 in all stages of Schwann cell development. Our data indicate that *SOX6* expression reduces as myelination occurs (Gopinath et al., 2016). Therefore, it is necessary to identify the mechanism by which SOX6 is repressed to allow myelination to occur. In the CNS, *SOX6* mRNA is targeted for degradation by microRNAs (including miR-338-3p) to allow oligodendrocyte differentiation. Interestingly SOX10 in known to regulate the expression of miR-338-3p in Schwann cells (Gokey et al., 2012). miR-338-3p plays an important role in Schwann cells (Dugas and Notterpek, 2011; Yun et al., 2010) and could be repressing *SOX6* expression for Schwann cell myelination. Clone the 3'UTR of *SOX6* downstream of a luciferase reporter mRNA, into S16 cells and test for reduced luciferase activity. Delete the miR-338-3p site from the 3'UTR of *SOX6* and perform luciferase assays to test for increased activity.

4.2.5 *ZEB2* regulation by SOX10

Recently, the transcription factor zinc finger E-box-binding homeo-box 2 (ZEB2) has been implicated in inhibiting the expression of myelin inhibitors during Schwann cell development and remyelination post injury. It has been called the 'inhibitor of inhibitors' and could be the switch that controls the transition from premyelination to myelination. ZEB2 is highly expressed prior to myelination and is re-expressed following nerve injury (Quintes et al., 2016; Wu et al., 2016). However, the transcriptional regulation of *ZEB2* expression in Schwann cells had not been studied. In chapter 2 we tested the regulatory activity of five regions within the *ZEB2* locus

(SOX10-CCS-2,3,4,5,6, and 7) and none of the regions displayed greater than 2.5-fold activity and none of these regions showed evidence of SOX10 occupancy in sciatic nerve (Srinivasan et al., 2012). However, the regions we identified by comparing our computational predictions with SOX10 ChIP-seq (Srinivasan et al., 2012) and our DNase-seq revealed evidence for SOX10 binding and conserved SOX10 site at the *ZEB2* locus, suggesting that SOX10 may control the expression of *ZEB2* during Schwann cell myelination. The genomic fragment harboring the conserved SOX10 site can be cloned upstream of a luciferase reporter gene and tested for enhancer activity in Schwann cells and the importance of the SOX10 site for the activity can be assessed by deleting the SOX10 site and performing luciferase assays. We would expect the region to be active in Schwann cells and the activity to be primarily driven by SOX10. qRT-PCR can be performed to test for decreased *ZEB2* expression upon SOX10 knockdown in Schwann cells. RNA-seq data generated from oligodendrocytes at various developmental stages (Zhang et al., 2014) indicates that *Zeb2* is highly expressed in newly forming oligodendrocytes and these consistent with observations made during Schwann cell development (Quintes et al., 2016). To identify target loci RNA-seq can be performed using sciatic nerve from mice with specific deletion of *Zeb2* in Schwann cells. We would expect increased expression of myelin inhibitors. To identify genome-wide occupancy of ZEB2, ChIP-seq can be performed using primary Schwann cells. We would expect ZEB2 binding at inhibitors of myelination.

4.3 Concluding Remarks

Previously, SOX10 response elements were identified on a gene-by gene basis. Interestingly, our in-depth computational and functional analysis—in addition to recent ChIP-seq data sets published by our collaborator Dr. John Svaren—has revealed a novel role for SOX10 in

Schwann cells. We hope that the genome-wide data sets we have generated will be useful to the scientific community to uncover novel roles for SOX10 not only in Schwann cells but also in other cell types such as oligodendrocytes and melanocytes. We identified *SOX6* as previously unreported target of SOX10 and our efforts revealed a putative role for SOX6 in Schwann cell repair and regeneration. Further understanding the role of SOX6 will provide key regulatory pathways that control Schwann cell dedifferentiation.

References

Allinquant, B., Staugaitis, S.M., D'Urso, D., and Colman, D.R. (1991). The ectopic expression of myelin basic protein isoforms in Shiverer oligodendrocytes: implications for myelinogenesis. J. Cell Biol. *113*, 393–403.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Anderson, D.J. (1993). Cell and molecular biology of neural crest cell lineage diversification. Current Opinion in Neurobiology.

Anton, E.S., Hadjiargyrou, M., Patterson, P.H., and Matthew, W.D. (1995). CD9 plays a role in Schwann cell migration in vitro. J. Neurosci. *15*, 584–595.

Antonellis, A., and Green, E.D. (2008). Inter-Species Comparative Sequence Analysis: Applications in Human Disease Research and Genomic Medicine. In *Genomic and Personalized Medicine*, H.F. Willard, and G.S. Ginsburg, eds. (Salt Lake City: Academic Press), pp. 120–130.

Antonellis, A., Bennett, W.R., Menheniott, T.R., Prasad, A.B., Lee-Lin, S.-Q., NISC Comparative Sequencing Program, Green, E.D., Paisley, D., Kelsh, R.N., Pavan, W.J., et al. (2006). Deletion of long-range sequences at Sox10 compromises developmental expression in a mouse model of Waardenburg-Shah (WS4) syndrome. Hum. Mol. Genet. *15*, 259–271.

Antonellis, A., Huynh, J.L., Lee-Lin, S.-Q., Vinton, R.M., Renaud, G., Loftus, S.K., Elliot, G., Wolfsberg, T.G., Green, E.D., McCallion, A.S., et al. (2008). Identification of neural crest and glial enhancers at the mouse Sox10 locus through transgenesis in zebrafish. PLoS Genet. *4*, e1000174.

Arter, J., and Wegner, M. (2015). Transcription factors Sox10 and Sox2 functionally interact with positive transcription elongation factor b in Schwann cells. J. Neurochem. *132*, 384–393.

Arthur-Farraj, P.J., Latouche, M., Wilton, D.K., Quintes, S., Chabrol, E., Banerjee, A., Woodhoo, A., Jenkins, B., Rahman, M., Turmaine, M., et al. (2012). c-Jun reprograms Schwann cells of injured nerves to generate a repair cell essential for regeneration. Neuron *75*, 633–647.

Arthur-Farraj, P.J., Morgan, C.C., Adamowicz, M., Gomez-Sanchez, J.A., Fazal, S.V., Beucher, A., Razzaghi, B., Mirsky, R., Jessen, K.R., and Aitman, T.J. (2017). Changes in the Coding and

Non-coding Transcriptome and DNA Methylome that Define the Schwann Cell Repair Phenotype after Nerve Injury. Cell Rep *20*, 2719–2734.

Bai, Q., Sun, M., Stolz, D.B., and Burton, E.A. (2011). Major isoform of zebrafish P0 is a 23.5 kDa myelin glycoprotein expressed in selected white matter tracts of the central nervous system. J. Comp. Neurol. *519*, 1580–1596.

Banerji, J., Olson, L., and Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299–308.

Baroti, T., Schillinger, A., and Wegner, M. (2015). Sox13 functionally complements the related Sox5 and Sox6 as important developmental modulators in mouse spinal cord oligodendrocytes. Journal of … *136*, 316–328.

Beirowski, B., Gustin, J., Armour, S.M., Yamamoto, H., Viader, A., North, B.J., Michán, S., Baloh, R.H., Golden, J.P., Schmidt, R.E., et al. (2011). Sir-two-homolog 2 (Sirt2) modulates peripheral myelination through polarity protein Par-3/atypical protein kinase C (aPKC) signaling. Proc. Natl. Acad. Sci. U.S.a. *108*, E952–E961.

Bergoffen, J., Scherer, S.S., and Wang, S. (1993). Connexin mutations in X-linked Charcot-Marie-Tooth disease.

Bird, A. (1999). DNA methylation de novo. Science.

Bondurand, N., Girard, M., Pingault, V., Lemort, N., Dubourg, O., and Goossens, M. (2001). Human Connexin 32, a gap junction protein altered in the X-linked form of Charcot-Marie-Tooth disease, is directly regulated by the transcription factor SOX10. Hum. Mol. Genet. *10*, 2783–2795.

Boullerne, A.I. (2016). The history of myelin. Exp. Neurol. *283*, 431–445.

Bowles, J., Schepers, G., and Koopman, P. (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. Dev. Biol. *227*, 239–255.

Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S., and Crawford, G.E. (2008a). High-resolution mapping and characterization of open chromatin across the genome. Cell *132*, 311–322.

Boyle, A.P., Guinney, J., Crawford, G.E., and Furey, T.S. (2008b). F-Seq: a feature density estimator for high-throughput sequence tags. Bioinformatics *24*, 2537–2538.

Bremer, M., Fröb, F., Kichko, T., Reeh, P., Tamm, E.R., Suter, U., and Wegner, M. (2011). Sox10 is required for Schwann-cell homeostasis and myelin maintenance in the adult peripheral nerve. Glia *59*, 1022–1032.

Brennan, K.M., Bai, Y., and Shy, M.E. (2015). Demyelinating CMT–what's known, what's new and what's in store? Neurosci. Lett.

Brewer, M.H., Ma, K.H., Beecham, G.W., Gopinath, C., the Inherited Neuropathy Consortium (INC), Baas, F., Choi, B.-O., Reilly, M.M., Shy, M.E., Züchner, S., et al. (2014). Haplotype-Specific Modulation of a SOX10/CREB Response Element at the Charcot-Marie-Tooth Disease Type 4C Locus SH3TC2. Hum. Mol. Genet. *23*, 5171–5187.

Britsch, S., Goerich, D.E., Riethmacher, D., Peirano, R.I., Rossner, M., Nave, K.A., Birchmeier, C., and Wegner, M. (2001). The transcription factor Sox10 is a key regulator of peripheral glial development. Genes Dev. *15*, 66–78.

Britsch, S., Li, L., Kirchhoff, S., and Theuring, F. (1998). The ErbB2 and ErbB3 receptors and their ligand, neuregulin-1, are essential for development of the sympathetic nervous system.

Brosius Lutz, A., Chung, W.-S., Sloan, S.A., Carson, G.A., Zhou, L., Lovelett, E., Posada, S., Zuchero, J.B., and Barres, B.A. (2017). Schwann cells use TAM receptor-mediated phagocytosis in addition to autophagy to clear myelin in a mouse model of nerve injury. Proc. Natl. Acad. Sci. U.S.a. *114*, E8072–E8080.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Bullock, T.H., Moore, J.K., and Fields, R.D. (1984). Evolution of myelin sheaths: both lamprey and hagfish lack myelin. Neurosci. Lett. *48*, 145–148.

Bunge, R.P., Bunge, M.B., and Bates, M. (1989). Movements of the Schwann cell nucleus implicate progression of the inner (axon-related) Schwann cell process during myelination. J. Cell Biol. *109*, 273–284.

Bürgisser, P., Matthieu, J.M., and Jeserich, G. (1986). Myelin lipids: a phylogenetic study. Neurochemical ….

Campagnoni, A.T. (1988). Molecular biology of myelin proteins from the central nervous system. J. Neurochem.

Cardone, B., and Roots, B.I. (1991). Comparative studies of myelin-like membranes in annelids and arthropods.

Cardone, B., and Roots, B.I. (1996). Monoclonal antibodies to proteins of the myelin-like sheath of earthworm giant axons show cross-reactivity to crayfish CNS glia: an immunogold electron microscopy study. Neurochem Res *21*, 505–510.

Chambon, P. (1975). Eukaryotic Nuclear RNA Polymerases. Annu. Rev. Biochem. *44*, 613–638.

Cheng, X., Gan, L., Zhao, J., Chen, M., Liu, Y., and Wang, Y. (2013). Changes in Ataxin-10 expression after sciatic nerve crush in adult rats. Neurochem Res *38*, 1013–1021.

Cheung, M., and Briscoe, J. (2003). Neural crest development is regulated by the transcription factor Sox9. Development *130*, 5681–5693.

Cohen-Barak, O., Hagiwara, N., Arlt, M.F., Horton, J.P., and Brilliant, M.H. (2001). Cloning, characterization and chromosome mapping of the human SOX6 gene. Gene *265*, 157–164.

Corden, J.L. (1990). Tails of RNA polymerase II. Trends in Biochemical Sciences.

Corden, J., Wasylyk, B., and Buchwalder, A. (1980). Promoter sequences of eukaryotic protein-coding genes.

Cravioto, H. (1965). The role of Schwann cells in the development of human peripheral nerves: An electron microscopic study. Journal of Ultrastructure Research.

Crawford, G.E., Davis, S., Scacheri, P.C., Renaud, G., Halawi, M.J., Erdos, M.R., Green, R., Meltzer, P.S., Wolfsberg, T.G., and Collins, F.S. (2006). DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. Nat. Methods *3*, 503–509.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. U.S.a. *107*, 21931–21936.

Cruz-Solis, I., Zepeda, R.C., Ortiz, S., Aguilera, J., López-Bayghen, E., and Ortega, A. (2009). Glutamate-dependent transcriptional control in Bergmann glia: Sox10 as a repressor. J. Neurochem. *109*, 899–910.

D'Este, E., Kamin, D., Balzarotti, F., and Hell, S.W. (2017). Ultrastructural anatomy of nodes of Ranvier in the peripheral nervous system as revealed by STED microscopy. Proc. Natl. Acad. Sci. U.S.a. *114*, E191–E199.

Davis, A.D., Weatherby, T.M., Hartline, D.K., and Lenz, P.H. (1999). Myelin-like sheaths in copepod axons. Nature *398*, 571–571.

Decker, L., Desmarquet-Trin-Dinh, C., Taillebourg, E., Ghislain, J., Vallat, J.-M., and Charnay, P. (2006). Peripheral myelin maintenance is a dynamic process requiring constant Krox20 expression. J. Neurosci. *26*, 9771–9779.

del Rio, H.P. (1921). La glia de escasas radiaciónes (oligodendroglia) (Arch Neurobiol (Madr)).

Denarier, E., Forghani, R., Farhadi, H.F., Dib, S., Dionne, N., Friedman, H.C., Lepage, P., Hudson, T.J., Drouin, R., and Peterson, A. (2005). Functional organization of a Schwann cell enhancer. J. Neurosci. *25*, 11210–11217.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Douarin, N.L., Dulac, C., Dupin, E., and Curry, P.C. (1991). Glial cell lineages in the neural crest. Glia.

Duflocq, A., Le Bras, B., Bullier, E., Couraud, F., and Davenne, M. (2008). Nav1.1 is predominantly expressed in nodes of Ranvier and axon initial segments. Mol. Cell. Neurosci. *39*, 180–192.

Dugas, J.C., and Notterpek, L. (2011). MicroRNAs in oligodendrocyte and Schwann cell differentiation. Dev. Neurosci. *33*, 14–20.

Dujardin, G., Lafaille, C., Petrillo, E., Buggiano, V., Gómez Acuña, L.I., Fiszbein, A., Godoy Herz, M.A., Nieto Moreno, N., Muñoz, M.J., Alló, M., et al. (2013). Transcriptional elongation and alternative splicing. Biochim. Biophys. Acta *1829*, 134–140.

Dumitriu, B., Patrick, M.R., Petschek, J.P., Cherukuri, S., Klingmuller, U., Fox, P.L., and Lefebvre, V. (2006). Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. Blood *108*, 1198–1207.

Dutton, K.A., Pauliny, A., Lopes, S.S., Elworthy, S., Carney, T.J., Rauch, J., Geisler, R., Haffter, P., and Kelsh, R.N. (2001). Zebrafish colourless encodes sox10 and specifies non-ectomesenchymal neural crest fates. Development *128*, 4113–4125.

Emery, B. (2013). Playing the field: Sox10 recruits different partners to drive central and peripheral myelination. PLoS Genet. *9*, e1003918.

ENCODE Project Consortium, Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

Erickson, C.A., and Reedy, M.V. (1998). Neural crest development: the interplay between morphogenesis and cell differentiation. Curr. Top. Dev. Biol. *40*, 177–209.

Falls, D.L. (2003). Neuregulins: functions, forms, and signaling strategies. Exp. Cell Res. *284*, 14–30.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., et al. (2016). The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. *44*, D279–D285.

Finzsch, M., Schreiner, S., Kichko, T., Reeh, P., Tamm, E.R., Bösl, M.R., Meijer, D., and Wegner, M. (2010). Sox10 is required for Schwann cell identity and progression beyond the immature Schwann cell stage. J. Cell Biol. *189*, 701–712.

Fogarty, E.A., Brewer, M.H., Rodriguez-Molina, J.F., Law, W.D., Ma, K.H., Steinberg, N.M., Svaren, J., and Antonellis, A. (2016). SOX10 Regulates an Alternative Promoter at the Charcot-Marie-Tooth Disease Locus MTMR2. Hum. Mol. Genet. *00*, 1–12.

FOLCH, J., and LEES, M. (1951). Proteolipides, a new type of tissue lipoproteins; their isolation from brain. J. Biol. Chem. *191*, 807–817.

Friede, R.L., and Bischhausen, R. (1982). How are sheath dimensions affected by axon caliber and internode length? Brain Res. *235*, 335–350.

Fröb, F., Bremer, M., Finzsch, M., Kichko, T., Reeh, P., Tamm, E.R., Charnay, P., and Wegner, M. (2012). Establishment of myelinating Schwann cells and barrier integrity between central and peripheral nervous systems depend on Sox10. Glia *60*, 806–819.

Fufa, T.D., Harris, M.L., Watkins-Chow, D.E., Levy, D., Gorkin, D.U., Gildea, D.E., Song, L., Safi, A., Crawford, G.E., Sviderskaya, E.V., et al. (2015). Genomic analysis reveals distinct mechanisms and functional classes of SOX10-regulated genes in melanocytes. Hum. Mol. Genet. *24*, 5433–5450.

Garbay, B., Heape, A.M., Sargueil, F., and Cassagne, C. (2000). Myelin synthesis in the peripheral nervous system. Prog. Neurobiol. *61*, 267–304.

Gelman, S., Cohen, A.H., and Sanovich, E. (2009). Developmental changes in the ultrastructure of the lamprey lateral line nerve during metamorphosis. J. Morphol. *270*, 815–824.

Geren, B.B. (1954). The formation from the schwann cell surface of myelin in the peripheral nerves of chick embryos. Exp. Cell Res.

Ghazvini, M., Mandemakers, W., Jaegle, M., Piirsoo, M., Driegen, S., Koutsourakis, M., Smit, X., Grosveld, F., and Meijer, D. (2002). A cell type-specific allele of the POU gene Oct-6 reveals Schwann cell autonomous function in nerve development and regeneration. Embo J. *21*, 4612–4620.

Ghislain, J., and Charnay, P. (2006). Control of myelination in Schwann cells: a Krox20 cis-regulatory element integrates Oct6, Brn2 and Sox10 activities. EMBO Rep. *7*, 52–58.

Giese, K.P., Martini, R., Lemke, G., Soriano, P., and Schachner, M. (1992). Mouse P0 gene disruption leads to hypomyelination, abnormal expression of recognition molecules, and degeneration of myelin and axons. Cell *71*, 565–576.

Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R., and Lieb, J.D. (2007). FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res. *17*, 877–885.

Givogri, M.I., Bongarzone, E.R., Schonmann, V., and Campagnoni, A.T. (2001). Expression and regulation of golli products of myelin basic protein gene during in vitro development of oligodendrocytes. J. Neurosci. Res. *66*, 679–690.

Goda, S., Hammer, J., Kobiler, D., and Quarles, R.H. (1991). Expression of the myelin-associated glycoprotein in cultures of immortalized Schwann cells. J. Neurochem. *56*, 1354–1361.

Gokey, N.G., Srinivasan, R., Lopez-Anido, C., Krueger, C., and Svaren, J. (2012). Developmental regulation of microRNA expression in Schwann cells. Mol. Cell. Biol. *32*, 558–568.

Gonçalves, A.F., Dias, N.G., Moransard, M., Correia, R., Pereira, J.A., Witke, W., Suter, U., and Relvas, J.B. (2010). Gelsolin is required for macrophage recruitment during remyelination of the peripheral nervous system. Glia *58*, 706–715.

Gopinath, C., Law, W.D., Rodriguez-Molina, J.F., Prasad, A.B., Song, L., Crawford, G.E., Mullikin, J.C., Svaren, J., and Antonellis, A. (2016). Stringent comparative sequence analysis reveals SOX10 as a putative inhibitor of glial cell differentiation. BMC Genomics *17*, 887.

Gould, R.M., Oakley, T., Goldstone, J.V., Dugas, J.C., Brady, S.T., and Gow, A. (2008). Myelin sheaths are formed with proteins that originated in vertebrate lineages. Neuron Glia Biol. *4*, 137–152.

Greenfield, S., Brostoff, S., Eylar, E.H., and Morell, P. (1973). PROTEIN COMPOSITION OF MYELIN OF THE PERIPHERAL NERVOUS SYSTEM. J. Neurochem. *20*, 1207–1216.

Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature *346*, 245–250.

Guth, S.I.E., and Wegner, M. (2008). Having it both ways: Sox protein function between conservation and innovation. Cell. Mol. Life Sci. *65*, 3000–3018.

Günther, J. (1976). Impulse conduction in the myelinated giant fibers of the earthworm. Structure and function of the dorsal nodes in the median giant fiber. J. Comp. Neurol. *168*, 505–531.

Hai, M., Muja, N., DeVries, G.H., Quarles, R.H., and Patel, P.I. (2002). Comparative analysis of Schwann cell lines as model systems for myelin gene transcription studies. J. Neurosci. Res. *69*, 497–508.

Hardison, R.C., Oeltjen, J., and Miller, W. (1997). Long human-mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. Genome Res. *7*, 959–966.

Harley, V.R., Jackson, D.I., Hextall, P.J., Hawkins, J.R., Berkovitz, G.D., Sockanathan, S., Lovell-Badge, R., and Goodfellow, P.N. (1992). DNA binding activity of recombinant SRY from normal males and XY females. Science *255*, 453–456.

Harley, V.R., Lovell-Badge, R., and Goodfellow, P.N. (1994). Definition of a consensus DNA binding site for SRY. Nucleic Acids Res. *22*, 1500–1501.

Harris, M.L., Baxter, L.L., Loftus, S.K., and Pavan, W.J. (2010). Sox proteins in melanocyte development and melanoma. Pigment Cell Melanoma Res *23*, 496–513.

Hayasaka, K., Himoro, M., Sato, W., Takada, G., Uyemura, K., Shimizu, N., Bird, T.D., Conneally, P.M., and Chance, P.F. (1993). Charcot-Marie-Tooth neuropathy type 1B is associated with mutations of the myelin P0 gene. Nat. Genet. *5*, 31–34.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Heinz, S., Romanoski, C.E., Benner, C., and Glass, C.K. (2015). The selection and function of cell type-specific enhancers. Nat. Rev. Mol. Cell Biol. *16*, 144–154.

Herbarth, B., Pingault, V., Bondurand, N., Kuhlbrodt, K., Hermans-Borgmeyer, I., Puliti, A., Lemort, N., Goossens, M., and Wegner, M. (1998). Mutation of the Sry-related Sox10 gene in Dominant megacolon, a mouse model for human Hirschsprung disease. Proc. Natl. Acad. Sci. U.S.a. *95*, 5161–5165.

Heuser, J.E., and Doggenweiler, C.F. (1966). The fine structural organization of nerve fibers, sheaths, and glial cells in the prawn, Palaemonetes vulgaris. J. Cell Biol. *30*, 381–403.

Hirata, K., Masuda, K., Morikawa, W., He, J.-W., Kuraoka, A., Kuwano, M., and Kawabuchi, M. (2004). N-myc downstream-regulated gene 1 expression in injured sciatic nerves. Glia *47*, 325–334.

Hodonsky, C.J., Kleinbrink, E.L., Charney, K.N., Prasad, M., Bessling, S.L., Jones, E.A., Srinivasan, R., Svaren, J., McCallion, A.S., and Antonellis, A. (2012). SOX10 regulates expression of the SH3-domain kinase binding protein 1 (Sh3kbp1) locus in Schwann cells via an alternative promoter. Mol. Cell. Neurosci. *49*, 85–96.

Huang, Y.-H., Jankowski, A., Cheah, K.S.E., Prabhakar, S., and Jauch, R. (2015). SOXE transcription factors form selective dimers on non-compact DNA motifs through multifaceted interactions between dimerization and high-mobility group domains. Sci Rep *5*, 10398.

Hung, H.A., Sun, G., Keles, S., and Svaren, J. (2015). Dynamic regulation of Schwann cell enhancers after peripheral nerve injury. J. Biol. Chem. *290*, 6937–6950.

Hung, H., Kohnken, R., and Svaren, J. (2012). The nucleosome remodeling and deacetylase chromatin remodeling (NuRD) complex is required for peripheral nerve myelination. J. Neurosci. *32*, 1517–1527.

Hutton, E.J., Carty, L., Laurá, M., Houlden, H., Lunn, M.P.T., Brandner, S., Mirsky, R., Jessen, K., and Reilly, M.M. (2011). c-Jun expression in human neuropathies: a pilot study. J. Peripher. Nerv. Syst. *16*, 295–303.

Inoue, K., Tanabe, Y., and Lupski, J.R. (1999). Myelin deficiencies in both the central and the peripheral nervous systems associated with a SOX10 mutation. Ann. Neurol. *46*, 313–318.

Inoue, K., Khajavi, M., Ohyama, T., Hirabayashi, S.-I., Wilson, J., Reggin, J.D., Mancias, P.,

Butler, I.J., Wilkinson, M.F., Wegner, M., et al. (2004). Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. Nat. Genet. *36*, 361–369.

Inoue, K., Shilo, K., Boerkoel, C.F., Crowe, C., Sawady, J., Lupski, J.R., and Agamanolis, D.P. (2002). Congenital hypomyelinating neuropathy, central dysmyelination, and Waardenburg-Hirschsprung disease: phenotypes linked by SOX10 mutation. Ann. Neurol. *52*, 836–842.

Jacob, C., Christen, C.N., Pereira, J.A., Somandin, C., Baggiolini, A., Lötscher, P., Ozçelik, M., Tricaud, N., Meijer, D., Yamaguchi, T., et al. (2011). HDAC1 and HDAC2 control the transcriptional program of myelination and the survival of Schwann cells. Nat. Neurosci. *14*, 429–436.

Jaegle, M., Ghazvini, M., Mandemakers, W., Piirsoo, M., Driegen, S., Levavasseur, F., Raghoenath, S., Grosveld, F., and Meijer, D. (2003). The POU proteins Brn-2 and Oct-6 share important functions in Schwann cell development. Genes Dev. *17*, 1380–1391.

Jagalur, N.B., Ghazvini, M., Mandemakers, W., Driegen, S., Maas, A., Jones, E.A., Jaegle, M., Grosveld, F., Svaren, J., and Meijer, D. (2011). Functional dissection of the Oct6 Schwann cell enhancer reveals an essential role for dimeric Sox10 binding. J. Neurosci. *31*, 8585–8594.

Jang, S.-W., Srinivasan, R., Jones, E.A., Sun, G., Keles, S., Krueger, C., Chang, L.-W., Nagarajan, R., and Svaren, J. (2010). Locus-wide identification of Egr2/Krox20 regulatory targets in myelin genes. J. Neurochem. *115*, 1409–1420.

Jarriault, S., Brou, C., Logeat, F., Schroeter, E.H., Kopan, R., and Israel, A. (1995). Signalling downstream of activated mammalian Notch. Nature *377*, 355–358.

Jeserich, G., Müller, A., and Jacque, C. (1990). Developmental expression of myelin proteins by oligodendrocytes in the CNS of trout. Developmental Brain Research.

Jessen, K.R., and Mirsky, R. (2016). The repair Schwann cell and its function in regenerating nerves. J. Physiol. (Lond.) *594*, 3521–3531.

Jessen, K.R., Brennan, A., Morgan, L., Mirsky, R., and Kent, A. (1994). The Schwann cell precursor and its fate: a study of cell death and differentiation during gliogenesis in rat embryonic nerves. Neuron.

Jessen, K.R., and Mirsky, R. (2005). The origin and development of glial cells in peripheral nerves. Nat. Rev. Neurosci. *6*, 671–682.

Jessen, K.R., and Mirsky, R. (2008). Negative regulation of myelination: relevance for development, injury, and demyelinating disease. Glia *56*, 1552–1565.

Jiang, Y., Yan, M., and Gralla, J.D. (1996). A three-step pathway of transcription initiation leading to promoter clearance at an activation RNA polymerase II promoter. Mol. Cell. Biol. *16*, 1614–1621.

Jiao, Z., Mollaaghababa, R., Pavan, W.J., Antonellis, A., Green, E.D., and Hornyak, T.J. (2004).

Direct interaction of Sox10 with the promoter of murine Dopachrome Tautomerase (Dct) and synergistic activation of Dct expression with Mitf. Pigment Cell Res. *17*, 352–362.

Jones, E.A., Brewer, M.H., Srinivasan, R., Krueger, C., Sun, G., Charney, K.N., Keles, S., Antonellis, A., and Svaren, J. (2012). Distal enhancers upstream of the Charcot-Marie-Tooth type 1A disease gene PMP22. Hum. Mol. Genet. *21*, 1581–1591.

Jones, E.A., Jang, S.-W., Mager, G.M., Chang, L.-W., Srinivasan, R., Gokey, N.G., Ward, R.M., Nagarajan, R., and Svaren, J. (2007). Interactions of Sox10 and Egr2 in myelin gene regulation. Neuron Glia Biol. *3*, 377–387.

Jones, E.A., Lopez-Anido, C., Srinivasan, R., Krueger, C., Chang, L.-W., Nagarajan, R., and Svaren, J. (2011). Regulation of the PMP22 gene through an intronic enhancer. J. Neurosci. *31*, 4242–4250.

Kabzińska, D., Kotruchow, K., Ryniewicz, B., and Kochański, A. (2011). Two pathogenic mutations located within the 5'-regulatory sequence of the GJB1 gene affecting initiation of transcription and translation. Acta Biochim. Pol. *58*, 359–363.

Kamachi, Y., and Kondoh, H. (2013). Sox proteins: regulators of cell fate specification and differentiation. Development *140*, 4129–4144.

Karolchik, D., Hinrichs, A.S., and Furey, T.S. (2004). The UCSC Table Browser data retrieval tool.

Kellerer, S., Schreiner, S., Stolt, C.C., Scholz, S., Bösl, M.R., and Wegner, M. (2006). Replacement of the Sox10 transcription factor by Sox8 reveals incomplete functional equivalence. Development *133*, 2875–2886.

Kelsh, R.N., and Eisen, J.S. (2000). The zebrafish colourless gene regulates development of non-ectomesenchymal neural crest derivatives. Development.

Kelsh, R.N. (2006). Sorting out Sox10 functions in neural crest development. Bioessays *28*, 788–798.

Kent, W.J., Sugnet, C.W., Furey, T.S., and Roskin, K.M. (2002). The human genome browser at UCSC. Genome … *12*, 996–1006.

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J.-I., Sekine, M., Tsuritani, K., Wakaguri, H., et al. (2006). Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes.

Klein, D., Groh, J., Wettmarshausen, J., and Martini, R. (2014). Nonuniform molecular features of myelinating Schwann cells in models for CMT1: Distinct disease patterns are associated with NCAM and c□Jun upregulation. Glia.

Kondo, T., and Raff, M. (2000). The Id4 HLH protein and the timing of oligodendrocyte

differentiation. Embo J. *19*, 1998–2007.

Kondoh, H., and Kamachi, Y. (2010). SOX-partner code for cell specification: Regulatory target selection and underlying molecular mechanisms. Int. J. Biochem. Cell Biol. *42*, 391–399.

Kuhlbrodt, K., Herbarth, B., Sock, E., Hermans-Borgmeyer, I., and Wegner, M. (1998). Sox10, a novel transcriptional modulator in glial cells. J. Neurosci. *18*, 237–250.

Kulshrestha, R., Burton-Jones, S., and Antoniadi, T. (2017). Deletion of P2 promoter of GJB1 gene a cause of Charcot-Marie-Tooth disease.

Küspert, M., Hammer, A., Bösl, M.R., and Wegner, M. (2011). Olig2 regulates Sox10 expression in oligodendrocyte precursors through an evolutionary conserved distal enhancer. Nucleic Acids Res. *39*, 1280–1293.

Kwasnieski, J.C., Fiore, C., Chaudhari, H.G., and Cohen, B.A. (2014). High-throughput functional testing of ENCODE segmentation predictions. Genome Res. *24*, 1595–1602.

LaBonne, C., and Bronner-Fraser, M. (1999). Molecular mechanisms of neural crest formation. Annual Review of Cell and ….

Lakiza, O., Miller, S., Bunce, A., Lee, E.M.-J., and McCauley, D.W. (2011). SoxE gene duplication and development of the lamprey branchial skeleton: Insights into development and evolution of the neural crest. Dev. Biol. *359*, 149–161.

Landry, J.-R., Mager, D.L., and Wilhelm, B.T. (2003). Complex controls: the role of alternative promoters in mammalian genomes. Trends Genet. *19*, 640–648.

Lane, P.W., and Liu, H.M. (1984). Association of megacolon with a new dominant spotting gene (Dom) in the mouse. J. Hered. *75*, 435–439.

Lang, D., Chen, F., Milewski, R., Li, J., Lu, M.M., and Epstein, J.A. (2000). Pax3 is required for enteric ganglia formation and functions with Sox10 to modulate expression of c-ret. J. Clin. Invest. *106*, 963–971.

Langmead, B. (2010). Aligning short sequencing reads with Bowtie. Current Protocols in Bioinformatics.

Lanwert, C., and Jeserich, G. (2001). Structure, heterologous expression, and adhesive properties of the P0□like myelin glycoprotein IP1 of trout CNS. Microscopy Research and Technique.

Laudet, V., Stehelin, D., and Clevers, H. (1993). Ancestry and diversity of the HMG box superfamily. Nucleic Acids Res. *21*, 2493–2501.

Laurence, K. (2017). The Evolution of Myelin: Theories and Application to Human Disease. Journal of Evolutionary Medicine.

Le, N., Nagarajan, R., Wang, J.Y.T., Svaren, J., LaPash, C., Araki, T., Schmidt, R.E., and

Milbrandt, J. (2005). Nab proteins are essential for peripheral nervous system myelination. Nat. Neurosci. *8*, 932–940.

LeBlanc, S.E., Jang, S.-W., Ward, R.M., Wrabetz, L., and Svaren, J. (2006). Direct regulation of myelin protein zero expression by the Egr2 transactivator. J. Biol. Chem. *281*, 5453–5460.

Lee, K.E., Nam, S., Cho, E.-A., Seong, I., Limb, J.-K., Lee, S., and Kim, J. (2008). Identification of direct regulatory targets of the transcription factor Sox10 based on function and conservation. BMC Genomics *9*, 408.

Lee, K.E., Seo, J., Shin, J., Ji, E.H., Roh, J., Kim, J.Y., Sun, W., Muhr, J., Lee, S., and Kim, J. (2014). Positive feedback loop between Sox2 and Sox6 inhibits neuronal differentiation in the developing central nervous system. Proc. Natl. Acad. Sci. U.S.a. *111*, 2794–2799.

Lee, M.M., Badache, A., and DeVries, G.H. (1999). Phosphorylation of CREB in axon-induced Schwann cell proliferation. J. Neurosci. Res. *55*, 702–712.

Lee, M., Goodall, J., Verastegui, C., Ballotti, R., and Goding, C.R. (2000). Direct regulation of the Microphthalmia promoter by Sox10 links Waardenburg-Shah syndrome (WS4)-associated hypopigmentation and deafness to WS2. J. Biol. Chem. *275*, 37978–37983.

Lefebvre, V., Li, P., and de Crombrugghe, B. (1998). A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. Embo J. *17*, 5718–5733.

Lefebvre, V. (2010). The SoxD transcription factors--Sox5, Sox6, and Sox13--are key cell fate modulators. Int. J. Biochem. Cell Biol. *42*, 429–432.

Leimeroth, R., Lobsiger, C., Lüssi, A., Taylor, V., and Suter, U. (2002a). Membrane-bound neuregulin1 type III actively promotes Schwann cell differentiation of multipotent progenitor cells.

Leimeroth, R., Lobsiger, C., Lüssi, A., Taylor, V., Suter, U., and Sommer, L. (2002b). Membrane-bound neuregulin1 type III actively promotes Schwann cell differentiation of multipotent Progenitor cells. Dev. Biol. *246*, 245–258.

Lemke, G., and Axel, R. (1985). Isolation and sequence of a cDNA encoding the major structural protein of peripheral myelin. Cell *40*, 501–508.

Levi, A.D., Bunge, R.P., and Lofgren, J.A. (1995). The influence of heregulins on human Schwann cell proliferation.

Li, H., Lu, Y., Smith, H.K., and Richardson, W.D. (2007). Olig1 and Sox10 interact synergistically to drive myelin basic protein transcription in oligodendrocytes. J. Neurosci. *27*, 14375–14382.

Limpert, A.S., and Carter, B.D. (2010). Axonal neuregulin 1 type III activates NF-kappaB in Schwann cells during myelin formation. J. Biol. Chem. *285*, 16614–16622.

Liu, C.-F., and Lefebvre, V. (2015). The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. Nucleic Acids Res. *43*, 8183–8203.

Lopez-Anido, C., Sun, G., Koenning, M., Srinivasan, R., Hung, H.A., Emery, B., Keles, S., and Svaren, J. (2015). Differential Sox10 genomic occupancy in myelinating glia. Glia *63*, 1897–1914.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lupski, J.R., de Oca-Luna, R.M., Slaugenhaupt, S., Pentao, L., Guzzetta, V., Trask, B.J., Saucedo-Cardenas, O., Barker, D.F., Killian, J.M., Garcia, C.A., et al. (1991). DNA duplication associated with Charcot-Marie-Tooth disease type 1A. Cell *66*, 219–232.

Mager, G.M., Ward, R.M., Srinivasan, R., Jang, S.-W., Wrabetz, L., and Svaren, J. (2008). Active gene repression by the Egr2.NAB complex during peripheral nerve myelination. J. Biol. Chem. *283*, 18187–18197.

Mandemakers, W., Zwart, R., Jaegle, M., Walbeehm, E., Visser, P., Grosveld, F., and Meijer, D. (2000). A distal Schwann cell-specific enhancer mediates axonal regulation of the Oct-6 transcription factor during peripheral nerve development and regeneration. Embo J. *19*, 2992–3003.

Mantuano, E., Jo, M., Gonias, S.L., and Campana, W.M. (2010). Low density lipoprotein receptor-related protein (LRP1) regulates Rac1 and RhoA reciprocally to control Schwann cell adhesion and migration. J. Biol. Chem. *285*, 14259–14266.

Marin-Husstege, M., He, Y., Li, J., Kondo, T., Sablitzky, F., and Casaccia-Bonnefil, P. (2006). Multiple roles of Id4 in developmental myelination: predicted outcomes and unexpected findings. Glia *54*, 285–296.

Martini, R., Fischer, S., Vales, R.L., and David, S. (2008). Interactions between Schwann cells and macrophages in injury and inherited demyelinating disease. Glia.

Martini, R., Mohajeri, M.H., Kasper, S., Giese, K.P., and Schachner, M. (1995). Mice doubly deficient in the genes for P0 and myelin basic protein show that both proteins contribute to the formation of the major dense line in peripheral nerve myelin. J. Neurosci. *15*, 4488–4495.

Masaki, T., Matsumura, K., Saito, F., Sunada, Y., Shimizu, T., Yorifuji, H., Motoyoshi, K., and Kamakura, K. (2000). Expression of dystroglycan and laminin-2 in peripheral nerve under axonal degeneration and regeneration. Acta Neuropathol. *99*, 289–295.

McCauley, D.W., and Bronner-Fraser, M. (2006). Importance of SoxE in neural crest development and the evolution of the pharynx. Nature *441*, 750–752.

Medeiros, D.M. (2013). The evolution of the neural crest: new perspectives from lamprey and invertebrate neural crest-like cells. Wiley Interdiscip Rev Dev Biol *2*, 1–15.

Meng, L., Mohan, R., Kwok, B.H., Elofsson, M., Sin, N., and Crews, C.M. (1999). Epoxomicin, a potent and selective proteasome inhibitor, exhibits in vivo antiinflammatory activity. Proc. Natl. Acad. Sci. U.S.a. *96*, 10403–10408.

Merkin, J.J., Chen, P., Alexis, M.S., Hautaniemi, S.K., and Burge, C.B. (2015). Origins and impacts of new mammalian exons. Cell Rep *10*, 1992–2005.

Meyer, D., and Birchmeier, C. (1995). Multiple essential functions of neuregulin in development. Nature.

Michailov, G.V., Sereda, M.W., Brinkmann, B.G., Fischer, T.M., Haug, B., Birchmeier, C., Role, L., Lai, C., Schwab, M.H., and Nave, K.-A. (2004). Axonal neuregulin-1 regulates myelin sheath thickness. Science *304*, 700–703.

Mikol, D.D., and Stefansson, K. (1988). A phosphatidylinositol-linked peanut agglutinin-binding glycoprotein in central nervous system myelin and on oligodendrocytes. J. Cell Biol. *106*, 1273–1279.

Mollaaghababa, R., and Pavan, W.J. (2003). The importance of having your SOX on: role of SOX10 in the development of neural crest-derived melanocytes and glia. Oncogene *22*, 3024–3034.

Monuki, E.S., Kuhn, R., and Lemke, G. (1993). Repression of the myelin P0 gene by the POU transcription factor SCIP. Mech. Dev. *42*, 15–32.

Munson, L.M., and Reznikoff, W.S. (1981). Abortive initiation and long ribonucleic acid synthesis. Biochemistry *20*, 2081–2085.

Nawaz, S., Schweitzer, J., Jahn, O., and Werner, H.B. (2013). Molecular evolution of myelin basic protein, an abundant structural myelin component. Glia *61*, 1364–1377.

Nevins, J.R. (1983). The pathway of eukaryotic mRNA formation. Annu. Rev. Biochem.

Noonan, J.P., and McCallion, A.S. (2010). Genomics of long-range regulatory elements.

Noseda, R., Belin, S., Piguet, F., Vaccari, I., Scarlino, S., Brambilla, P., Martinelli Boneschi, F., Feltri, M.L., Wrabetz, L., Quattrini, A., et al. (2013). DDIT4/REDD1/RTP801 is a novel negative regulator of Schwann cell myelination. J. Neurosci. *33*, 15295–15305.

Ogata, T., Ueno, T., Hoshikawa, S., Ito, J., Okazaki, R., Hayakawa, K., Morioka, K., Yamamoto, S., Nakamura, K., Tanaka, S., et al. (2011). Hes1 functions downstream of growth factors to maintain oligodendrocyte lineage cells in the early progenitor stage. Neuroscience *176*, 132–141.

Ohe, K., Lalli, E., and Sassone-Corsi, P. (2002). A direct role of SRY and SOX proteins in pre-mRNA splicing. Proc. Natl. Acad. Sci. U.S.a. *99*, 1146–1151.

Omenn, G.S., and McKusick, V.A. (1979). The association of Waardenburg syndrome and Hirschsprung megacolon. Am. J. Med. Genet. *3*, 217–223.

Parkinson, D.B., Bhaskaran, A., Arthur-Farraj, P., Noon, L.A., Woodhoo, A., Lloyd, A.C., Feltri, M.L., Wrabetz, L., Behrens, A., Mirsky, R., et al. (2008). c-Jun is a negative regulator of myelination. J. Cell Biol. *181*, 625–637.

Parkinson, D.B., Langner, K., Namini, S.S., Jessen, K.R., and Mirsky, R. (2002). beta-Neuregulin and autocrine mediated survival of Schwann cells requires activity of Ets family transcription factors. Mol. Cell. Neurosci. *20*, 154–167.

Peirano, R.I., and Wegner, M. (2000). The glial transcription factor Sox10 binds to DNA both as monomer and dimer with different functional consequences. Nucleic Acids Res. *28*, 3047–3055.

Peirano, R.I., Goerich, D.E., Riethmacher, D., and Wegner, M. (2000). Protein zero gene expression is regulated by the glial transcription factor Sox10. Mol. Cell. Biol. *20*, 3198–3209.

Pereyra, P.M., and Roots, B.I. (1988). Isolation and initial characterization of myelin-like membrane fractions from the nerve cord of earthworms (Lumbricus terrestris L). Neurochem Res *13*, 893–901.

Perez-Alcala, S., Nieto, M.A., and Barbas, J.A. (2004). LSox5 regulates RhoB expression in the neural tube and promotes generation of the neural crest. Development *131*, 4455–4465.

Pérez-Cerdá, F., Sánchez-Gómez, M.V., and Matute, C. (2015). Pío del Río Hortega and the discovery of the oligodendrocytes. Front Neuroanat *9*, 92.

Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev.

Phelps, D.E., Hsiao, K.M., Li, Y., Hu, N., Franklin, D.S., Westphal, E., Lee, E.Y., and Xiong, Y. (1998). Coupled transcriptional and translational control of cyclin-dependent kinase inhibitor p18INK4c expression during myogenesis. Mol. Cell. Biol. *18*, 2334–2343.

Pingault, V., Bondurand, N., Kuhlbrodt, K., Goerich, D.E., Préhu, M.O., Puliti, A., Herbarth, B., Hermans-Borgmeyer, I., Legius, E., Matthijs, G., et al. (1998). SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nat. Genet. *18*, 171–173.

Prasad, M.K., Reed, X., Gorkin, D.U., Cronin, J.C., McAdow, A.R., Chain, K., Hodonsky, C.J., Jones, E.A., Svaren, J., Antonellis, A., et al. (2011). SOX10 directly modulates ERBB3 transcription via an intronic neural crest enhancer. BMC Dev. Biol. *11*, 40.

Prior, H.M., and Walter, M.A. (1996). SOX genes: architects of development. Mol. Med. *2*, 405–412.

Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. Cell.

Puckett, C., Hundson, L., and Ono, K. (1987). Myelin specific proteolipid protein is expressed in myelinating schwann cells but is not incorporated into myelin sheaths. Journal of ….

Pusch, C., Hustert, E., Pfeifer, D., Südbeck, P., Kist, R., Roe, B., Wang, Z., Balling, R., Blin, N., and Scherer, G. (1998). The SOX10/Sox10 gene from human and mouse: sequence, expression, and transactivation by the encoded HMG domain transcription factor. Hum. Genet. *103*, 115–123.

Quintes, S., and Brinkmann, B.G. (2017). Transcriptional inhibition in Schwann cell development and nerve regeneration. Neural Regen Res *12*, 1241–1246.

Quintes, S., Brinkmann, B.G., Ebert, M., Fröb, F., Kungl, T., Arlt, F.A., Tarabykin, V., Huylebroeck, D., Meijer, D., Suter, U., et al. (2016). Zeb2 is essential for Schwann cell differentiation, myelination and nerve repair. Nat. Neurosci. *19*, 1050–1059.

Radivojac, P., Vacic, V., Haynes, C., Cocklin, R.R., Mohan, A., Heyen, J.W., Goebl, M.G., and Iakoucheva, L.M. (2010). Identification, analysis, and prediction of protein ubiquitination sites. Proteins *78*, 365–380.

Razin, A., and Riggs, A.D. (1980). DNA methylation and gene function. Science.

Reiprich, S., Stolt, C.C., Schreiner, S., Parlato, R., and Wegner, M. (2008). SoxE proteins are differentially required in mouse adrenal gland development. Mol. Biol. Cell *19*, 1575–1586.

Riethmacher, D., and Sonnenberg-Riethmacher, E. (1997). Severe neuropathies in mice with targeted mutations in the ErbB3 receptor. Nature.

Roa, B.B., Warner, L.E., Garcia, C.A., Russo, D., Lovelace, R., Chance, P.F., and Lupski, J.R. (1996). Myelin protein zero (MPZ) gene mutations in nonduplication type 1 Charcot-Marie-Tooth disease. Hum. Mutat. *7*, 36–45.

Roach, A., Takahashi, N., Pravtcheva, D., Ruddle, F., and Hood, L. (1985). Chromosomal mapping of mouse myelin basic protein gene and structure and transcription of the partially deleted gene in shiverer mutant mice. Cell *42*, 149–155.

Roberts, S.L., Dun, X., Doddrell, R., and Mindos, T. (2017a). Sox2 expression in Schwann cells inhibits myelination in vivo and induces influx of macrophages to the nerve.

Roberts, S.L., Dun, X.-P., Doddrell, R.D.S., Mindos, T., Drake, L.K., Onaitis, M.W., Florio, F., Quattrini, A., Lloyd, A.C., D'Antonio, M., et al. (2017b). Sox2 expression in Schwann cells inhibits myelination in vivo and induces influx of macrophages to the nerve. Development *144*, 3114–3125.

Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. Trends in Biochemical Sciences.

Roots, B.I., and Lane, N.J. (1983). Myelinating glia of earthworm giant axons: thermally induced intramembranous changes. Tissue Cell *15*, 695–709.

Roots, B.I., Cardone, B., and Pereyra, P. (1991). Isolation and characterization of the myelin-like membranes ensheathing giant axons in the earthworm nerve cord. Ann. N. Y. Acad. Sci. *633*,

559–561.

Roots, B.I. (2008). The phylogeny of invertebrates and the evolution of myelin. Neuron Glia Biol. *4*, 101–109.

Rosenbluth, J. (1999). A brief history of myelinated nerve fibers: one hundred and fifty years of controversy. J. Neurocytol. *28*, 251–262.

Ryu, E.J., Wang, J.Y.T., Le, N., Baloh, R.H., Gustin, J.A., Schmidt, R.E., and Milbrandt, J. (2007). Misexpression of Pou3f1 results in peripheral nerve hypomyelination and axonal loss. J. Neurosci. *27*, 11552–11559.

Saavedra, R.A., Fors, L., Aebersold, R.H., and Arden, B. (1989). The myelin proteins of the shark brain are similar to the myelin proteins of the mammalian peripheral nervous system.

Sakuta, M. (2010). [One hundred books which built up neurology (48)--Roberto Remak "Observationes anatomicae et microscopicae de systematis nervosi structura"].

Salazar-Grueso, E.F., Kim, S., and Kim, H. (1991). Embryonic mouse spinal cord motor neuron hybrid cells. Neuroreport *2*, 505–508.

Samstein, R.M., Arvey, A., Josefowicz, S.Z., Peng, X., Reynolds, A., Sandstrom, R., Neph, S., Sabo, P., Kim, J.M., Liao, W., et al. (2012). Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. Cell *151*, 153–166.

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

Saporta, A., Sottile, S.L., and Miller, L.J. (2011). Charcot marie tooth disease subtypes and genetic testing strategies.

Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R., and Nakanishi, S. (1992). Two mammalian helix-loop-helix factors structurally related to Drosophila hairy and Enhancer of split. Genes Dev. *6*, 2620–2634.

Sauka-Spengler, T., and Bronner-Fraser, M. (2008). A gene regulatory network orchestrates neural crest formation. Nat. Rev. Mol. Cell Biol. *9*, 557–568.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl. Acad. Sci. U.S.a. *103*, 1412–1417.

Schepers, G.E., Bullejos, M., Hosking, B.M., and Koopman, P. (2000). Cloning and characterisation of the Sry-related transcription factor gene Sox8. Nucleic Acids Res. *28*, 1473–1480.

Schlierf, B., Ludwig, A., Klenovsek, K., and Wegner, M. (2002). Cooperative binding of Sox10

to DNA: requirements and consequences. Nucleic Acids Res. *30*, 5509–5516.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.-Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell *132*, 887–898.

Schreiner, S., Cossais, F., Fischer, K., Scholz, S., Bösl, M.R., Holtmann, B., Sendtner, M., and Wegner, M. (2007). Hypomorphic Sox10 alleles reveal novel protein functions and unravel developmental differences in glial lineages. Development *134*, 3271–3281.

Schweigreiter, R., Roots, B.I., Bandtlow, C.E., and Gould, R.M. (2006). Understanding myelination through studying its evolution. Int. Rev. Neurobiol. *73*, 219–273.

Sherman, D.L., and Brophy, P.J. (2005). Mechanisms of axon ensheathment and myelin growth. Nat. Rev. Neurosci. *6*, 683–690.

Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. Nat. Rev. Genet. *15*, 272–286.

Smith, J.J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M.S., Yandell, M.D., Manousaki, T., Meyer, A., Bloom, O.E., et al. (2013). Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. Nat. Genet. *45*, 415–21–421e1–2.

Smits, P., Li, P., Mandel, J., Zhang, Z., Deng, J.M., Behringer, R.R., de Crombrugghe, B., and Lefebvre, V. (2001). The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. Dev. Cell *1*, 277–290.

Snipes, G.J., Suter, U., Welcher, A.A., and Shooter, E.M. (1992). Characterization of a novel peripheral nervous system myelin protein (PMP-22/SR13). J. Cell Biol. *117*, 225–238.

Song, L., and Crawford, G.E. (2010). DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. Cold Spring Harb Protoc *2010*, pdb.prot5384.

Southard-Smith, E.M., Kos, L., and Pavan, W.J. (1998). Sox10 mutation disrupts neural crest development in Dom Hirschsprung mouse model. Nat. Genet. *18*, 60–64.

Srinivasan, R., Sun, G., Keles, S., Jones, E.A., Jang, S.-W., Krueger, C., Moran, J.J., and Svaren, J. (2012). Genome-wide analysis of EGR2/SOX10 binding in myelinating peripheral nerve. Nucleic Acids Res. *40*, 6449–6460.

Steinthorsdottir, V., Stefansson, H., Ghosh, S., Birgisdottir, B., Bjornsdottir, S., Fasquel, A.C., Olafsson, O., Stefansson, K., and Gulcher, J.R. (2004). Multiple novel transcription initiation sites for NRG1. Gene *342*, 97–105.

Stewart, H.J., Zoidl, G., Rossner, M., Brennan, A., Zoidl, C., Nave, K.A., Mirsky, R., and Jessen, K.R. (1997). Helix-loop-helix proteins in Schwann cells: a study of regulation and subcellular localization of Ids, REB, and E12/47 during embryonic and postnatal development. J. Neurosci. Res. *50*, 684–701.

Stine, Z.E., Huynh, J.L., Loftus, S.K., Gorkin, D.U., Salmasi, A.H., Novak, T., Purves, T., Miller, R.A., Antonellis, A., Gearhart, J.P., et al. (2009). Oligodendroglial and pan-neural crest expression of Cre recombinase directed by Sox10 enhancer. Genesis *47*, 765–770.

Stolt, C.C., and Wegner, M. (2010). SoxE function in vertebrate nervous system development. Int. J. Biochem. Cell Biol. *42*, 437–440.

Stolt, C.C., and Wegner, M. (2016). Schwann cells and their transcriptional network: Evolution of key regulators of peripheral myelination. Brain Res. *1641*, 101–110.

Stolt, C.C., Lommes, P., Hillgärtner, S., and Wegner, M. (2008). The transcription factor Sox5 modulates Sox10 function during melanocyte development. Nucleic Acids Res. *36*, 5427–5440.

Stolt, C.C., Rehberg, S., Ader, M., Lommes, P., Riethmacher, D., Schachner, M., Bartsch, U., and Wegner, M. (2002). Terminal differentiation of myelin-forming oligodendrocytes depends on the transcription factor Sox10. Genes Dev. *16*, 165–170.

Stolt, C.C., Schlierf, A., Lommes, P., Hillgärtner, S., Werner, T., Kosian, T., Sock, E., Kessaris, N., Richardson, W.D., Lefebvre, V., et al. (2006). SoxD proteins influence multiple stages of oligodendrocyte development and modulate SoxE protein function. Dev. Cell *11*, 697–709.

Svaren, J., and Meijer, D. (2008). The molecular machinery of myelin gene transcription in Schwann cells. Glia *56*, 1541–1551.

Tai, F.L., and Smith, R. (1984). Comparison of the major proteins of shark myelin with the proteins of higher vertebrates. J. Neurochem. *42*, 426–433.

Taveggia, C., Zanazzi, G., Petrylak, A., and Yano, H. (2005). Neuregulin-1 type III determines the ensheathment fate of axons. Neuron.

Toda, K., Small, J.A., Goda, S., and Quarles, R.H. (1994). Biochemical and cellular properties of three immortalized Schwann cell lines expressing different levels of the myelin-associated glycoprotein. J. Neurochem. *63*, 1646–1657.

Tomassy, G.S., Dershowitz, L.B., and Arlotta, P. (2016). Diversity Matters: A Revised Guide to Myelination. Trends Cell Biol. *26*, 135–147.

Topilko, P., Schneider-Maunoury, S., Levi, G., Baron-Van Evercooren, A., Chennoufi, A.B., Seitanidou, T., Babinet, C., and Charnay, P. (1994). Krox-20 controls myelination in the peripheral nervous system. Nature *371*, 796–799.

Travis, A., Amsterdam, A., and Belanger, C. (1991). LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function

[corrected]. Genes & ….

Triolo, D., Dina, G., Lorenzetti, I., Malaguti, M., Morana, P., Del Carro, U., Comi, G., Messing, A., Quattrini, A., and Previtali, S.C. (2006). Loss of glial fibrillary acidic protein (GFAP) impairs Schwann cell proliferation and delays nerve regeneration after damage. J. Cell. Sci. *119*, 3981–3993.

Triolo, D., Dina, G., Taveggia, C., Vaccari, I., Porrello, E., Rivellini, C., Domi, T., La Marca, R., Cerri, F., Bolino, A., et al. (2012). Vimentin regulates peripheral nerve myelination. Development *139*, 1359–1367.

Turnescu, T., Arter, J., Reiprich, S., Tamm, E.R., Waisman, A., and Wegner, M. (2017). Sox8 and Sox10 jointly maintain myelin gene expression in oligodendrocytes. Glia.

Vardimon, L., Kressmann, A., Cedar, H., Maechler, M., and Doerfler, W. (1982). Expression of a cloned adenovirus gene is inhibited by in vitro methylation. Proc. Natl. Acad. Sci. U.S.a. *79*, 1073–1077.

Varshavsky, A. (2017). The Ubiquitin System, Autophagy, and Regulated Protein Degradation. Annu. Rev. Biochem. *86*, 123–128.

Virchow, R. (1854). Ueber das ausgebreitete Vorkommen einer dem Nervenmark analogen Substanz in den thierischen Geweben. Virchows Archiv.

Vogl, M.R., Reiprich, S., Küspert, M., Kosian, T., Schrewe, H., Nave, K.-A., and Wegner, M. (2013). Sox10 cooperates with the mediator subunit 12 during terminal differentiation of myelinating glia. J. Neurosci. *33*, 6679–6690.

Waehneldt, T.V., and Jeserich, G. (1984). Biochemical characterization of the central nervous system myelin proteins of the rainbow trout, Salmo gairdneri. Brain Res. *309*, 127–134.

Waehneldt, T.V., Matthieu, J.M., and Jeserich, G. (1986). Appearance of Myelin proteins during vertebrate evolution. Neurochem. Int. *9*, 463–474.

Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. Genome Res. *22*, 1798–1812.

Wang, S., Sdrulla, A., Johnson, J.E., Yokota, Y., and Barres, B.A. (2001). A role for the helix-loop-helix protein Id2 in the control of oligodendrocyte development. Neuron *29*, 603–614.

Warner, L.E., Garcia, C.A., and Lupski, J.R. (1999). Hereditary peripheral neuropathies: clinical forms, genetics, and molecular mechanisms. Annu. Rev. Med. *50*, 263–275.

Warner, L.E., Mancias, P., Butler, I.J., McDonald, C.M., Keppen, L., Koob, K.G., and Lupski, J.R. (1998). Mutations in the early growth response 2 (EGR2) gene are associated with hereditary myelinopathies. Nat. Genet. *18*, 382–384.

Waterman, M.L., and Jones, K.A. (1990). Purification of TCF-1 alpha, a T-cell-specific transcription factor that activates the T-cell receptor C alpha gene enhancer in a context-dependent manner. The New Biologist.

Wei, Q., and Miskimins, W.K. (2004). Sox10 acts as a tissue specific transcription factor enhancing activation of the myelin basic protein gene promoter by p27Kip1 and Sp1.

Weider, M., and Wegner, M. (2017). SoxE factors: Transcriptional regulators of neural differentiation and nervous system development. Semin. Cell Dev. Biol. *63*, 35–42.

Weider, M., Küspert, M., Bischof, M., Vogl, M.R., Hornig, J., Loy, K., Kosian, T., Müller, J., Hillgärtner, S., Tamm, E.R., et al. (2012). Chromatin-remodeling factor Brg1 is required for Schwann cell differentiation and myelination. Dev. Cell *23*, 193–201.

Weterman, M.A.J., van Ruissen, F., de Wissel, M., Bordewijk, L., Samijn, J.P.A., van der Pol, W.L., Meggouh, F., and Baas, F. (2010). Copy number variation upstream of PMP22 in Charcot-Marie-Tooth disease. Eur. J. Hum. Genet. *18*, 421–428.

Woldeyesus, M.T., Britsch, S., and Riethmacher, D. (1999). Peripheral nervous system defects in erbB2 mutants following genetic rescue of heart development.

Wolpowitz, D., Mason, T., Dietrich, P., and Mendelsohn, M. (2000). Cysteine-rich domain isoforms of the neuregulin-1 gene are required for maintenance of peripheral synapses. Neuron.

Woodhoo, A., Alonso, M.B.D., Droggiti, A., Turmaine, M., D'Antonio, M., Parkinson, D.B., Wilton, D.K., Al-Shawi, R., Simons, P., Shen, J., et al. (2009). Notch controls embryonic Schwann cell differentiation, postnatal myelination and adult plasticity. Nat. Neurosci. *12*, 839–847.

Wright, E.M., Snopek, B., and Koopman, P. (1993). Seven new members of the Sox gene family expressed during mouse development. Nucleic Acids Res.

Wu, L.M.N., Wang, J., Conidi, A., Zhao, C., Wang, H., Ford, Z., Zhang, L., Zweier, C., Ayee, B.G., Maurel, P., et al. (2016). Zeb2 recruits HDAC-NuRD to inhibit Notch and controls Schwann cell differentiation and remyelination. Nat. Neurosci. *19*, 1060–1072.

Xin, D., Hu, L., and Kong, X. (2008). Alternative promoters influence alternative splicing at the genomic level. PLoS ONE *3*, e2377.

Yang, C., Bolotin, E., Jiang, T., Sladek, F.M., and Martinez, E. (2007). Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. Gene.

Yun, B., Anderegg, A., Menichella, D., Wrabetz, L., Feltri, M.L., and Awatramani, R. (2010). MicroRNA-deficient Schwann cells display congenital hypomyelination. J. Neurosci. *30*, 7722–7728.

Zalc, B. (2016). The acquisition of myelin: An evolutionary perspective. Brain Res. *1641*, 4–10.

Zalc, B., Goujet, D., and Colman, D. (2008). The origin of the myelination program in vertebrates. Curr. Biol. *18*, R511–R512.

Zencheck, W.D., Xiao, H., and Weiss, L.M. (2012). Lysine post-translational modifications and the cytoskeleton. Essays Biochem. *52*, 135–145.

Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res. *21*, 1273–1283.

Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keeffe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., et al. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J. Neurosci. *34*, 11929–11947.

Zhao, X., He, X., Han, X., Yu, Y., Ye, F., Chen, Y., Hoang, T., Xu, X., Mi, Q.-S., Xin, M., et al. (2010). MicroRNA-mediated control of oligodendrocyte differentiation. Neuron *65*, 612–626.

Zorick, T.S., and Lemke, G. (1996). Schwann cell differentiation. Current Opinion in Cell Biology.

Zorick, T.S., Syroid, D.E., Brown, A., Gridley, T., and Lemke, G. (1999). Krox-20 controls SCIP expression, cell cycle exit and susceptibility to apoptosis in developing myelinating Schwann cells. Development *126*, 1397–1406.

Appendix

Element ID	Locus	Coordinates (hg18)	Forward Primer	Reverse Primer	Size (bp)
SOX10-CCS-01	PAX7	chr1:18,854,300-18,855,055	ATTCCAGTTTCCACGGTCAG	GCTCAAGTCTGATGCTGCAA	756
SOX10-CCS-02	ZEB2	chr2:144,875,965-144,876,906	AGCCCAGTTTTTCCTGAGGT	GAGAACAGCTCATGTAAATTATTCCA	942
SOX10-CCS-03	ZEB2	chr2:144,900,863-144,901,802	GCTCAATGTGTGAAAATGAAACA	TGGCTATTTGGACCAAGAAACT	940
SOX10-CCS-04	ZEB2	chr2:144,944,023-144,944,960	GAGTTGCAAGCAACCTGTGA	TGAAGAACAACAGGCTTTGG	938
SOX10-CCS-05	ZEB2	chr2:144,973,970-144,974,917	TATGGCAGCATTTGTTCAGC	GAGCAATCCTTCCATTTCCA	948
SOX10-CCS-06	ZEB2	chr2:144,988,817-144,989,623	AAGCAATGGACAGGCTTGAT	TCCCCAAGATTCAGTTCAGG	807
SOX10-CCS-07	PAX3	chr2:222,845,095-222,845,899	CGCCACTGTTTATCCCAG	GAGAAACCTGGCAAGGG	805
SOX10-CCS-08	SLIT2	chr4:20,088,937-20,089,889	TGCCCCTTCATATGAGTAACC	CCTTGATGTCATGCAAATGG	953
SOX10-CCS-09	PPARGC1A	chr4:23,484,443-23,485,333	CAACAAGAAAGCTTGCCAGAG	AGTGTGTGCCTGTGTATGTG	891
SOX10-CCS-10	SOX6	chr11:16,099,680-16,100,466	CCAGTTTTCAGCTTACTTTGG	CTGGAAATAAGACAGGGTGG	787
SOX10-CCS-11	SOX6	chr11:16,268,203-16,269,201	CGGTTACTACCCTCAGAATGGA	TTATTGGTGGCCAAAGCACT	999
SOX10-CCS-12	SOX6	chr11:16,272,728-16,273,714	TGAAGTTGCCAGTTTTAATGC	GGGAGTTCTGTTTTGGGACA	987
SOX10-CCS-13	SOX6	chr11:16,334,301-16,335,278	TGACACCTTCCCAAATCACA	TTCGTGCCAATGATGACCT	978
SOX10-CCS-14	SOX6	chr11:16,382,713-16,383,658	CAACCAGGTTTCACCATCAA	CTGGCTGAGAGTGTTCTGGA	946
SOX10-CCS-15	SOX6	chr11:16,419,585-16,420,519	GGTCAGCACCTCTCCAACAT	TTCCAGAGGCAGGTTTCATT	935
SOX10-CCS-16	HTATIP2	chr11:20,350,779-20,351,724	TGTCTGTCCACATGGTTAGG	AGCAAGATTGATTGGAAGG	946

SOX10-CCS-17	NTM	chr11:130,816,285-130,817,036	ACAGCTCTTTTTGGTCATGCAG	TTTTCTCCAGGCCTCCAGTG	752
SOX10-CCS-18	SOX5	chr12:24,058,988-24,059,872	GACTCCTTAAATTCACAATCTGG	GGCCCTGCTACTTTATCAGC	885
SOX10-CCS-19	SOX5	chr12:24,059,397-24,060,164	AAGCGAGTGTCGCCTAGGTA	TCCTCCCTCTGTGCTGTCTT	768
SOX10-CCS-20	SOX5	chr12:24,064,421-24,065,373	CATTAACCAACCCCTGATGC	TCCATGCACTTCCTTTGTGT	953
SOX10-CCS-21	IGF1R	chr15:97,238,558-97,239,353	TTCCTGGTAAACAGTTCTGCTG	CCCCAGTACTGTGAGCAACA	796
SOX10-CCS-22	TCF4	chr18:51,243,509-51,244,258	TCTTAGCATGGGCCCTATC	GGGTTGTATCCATCTCAGAGC	750
SOX10-CCS-23	AKT3	chr1:241,943,244-241,944,344	ACATGAATAAGGGAGAGAAGAGGA	TGTGCCTTAACTTAGAAACACTCC	1101
SOX10-CCS-24	FOXP1	chr3:71,181,977-71,183,215	GCCACTCCCTTCCCAAACTC	CCTGGAGTCCTGTTGAGCAG	1239
SOX10-CCS-25	FOXP1	chr3:71,373,215-71,374,005	GTCTGACTTAGGGGCGAGTG	TGCTTGTTCGAGACAGGTCA	791
SOX10-CCS-26	FOXP1	chr3:71,441,201-71,441,519	ACACACTGTTGACTTCACAAGT	ACTGCATTGTGTAAATTTGCTGTG	319
SOX10-CCS-27	FOXP2	chr7:113,841,332-113,842,324	AGTCAGTTCTTGCAATAGGAGG	CTTTGGTGTGCAACGTGAGG	993
SOX10-CCS-28	FOXP2	chr7:113,852,613-113,853,735	CACAGCCAGGTTGTTTCTGC	CAAGATGTCCCTCTCTGCCA	1123
SOX10-CCS-29	FOXP2	chr7:113,859,912-113,860,503	AGAAATGGGAAAATGTGGCATCT	ATGGACTAGGACACAAATGCTCA	592
SOX10-CCS-30	FOXP2	chr7:113,929,584-113,930,458	AGAAACTGACAGTGTTTTGGAAGT	TGCTTGAGGAGAAAGGGGATC	875
SOX10-CCS-31	FOXP2	chr7:114,082,350-114,083,553	AGACATGTATCTTTTTGAATCTGACA	TGGCACATTCAGAACCCAGA	1204
SOX10-CCS-32	LRPPRC	chr2:44,053,183-44,053,867	TGTGGTTCCAAAACACTGGGT	TGGTCATTTTCTTTGTGGGCC	685
SOX10-CCS-33	NFIA	chr1:61,419,236-61,419,765	CGGGGCTGGCATATAAGAGC	TCCATCTTACAGACTTTCACAATGA	530
SOX10-CCS-34	NFIA	chr1:61,482,314-61,482,776	TGGGGTGTATGTGTATGCTGG	ACAGCTAAACCCCTAGCCCT	463
SOX10-CCS-35	NFIA	chr1:61,685,814-61,686,119	CACCCAGAAAATCCGGCAGT	TTCTGGAGCCGCTTATGACG	306
SOX10-CCS-36	ROBO2	chr3:77,703,835-77,704,225	CACTGAAGTGTGCAAGTGTGC	TGAAATAAGGCAACCAAGAGGC	391
SOX10-CCS-37	ST18	chr8:53,370,162-53,370,714	TACCTCTAAGGAGCCTGCCA	AGGGGGAAGTCAGAGATATGTCA	553
SOX10-CCS-38	TCF7L2	chr10:114,809,055-114,809,611	GCTTTCAAGGCTGGACCACT	AGGAGAAAACAATCTGCTCTTTTCC	557
SOX10-CCS-39	TCF7L2	chr10:114,894,980-114,895,808	TGAACATGAGCTTGTGACCCA	GGGGTGTCTGAATCCTCCTG	829
SOX10-CCS-40	ZFP536	chr19:35,553,294-35,554,399	ATCCAGGCAAAACAGAGGGG	ATACAGCAGGGAGGCAGATG	1106
SOX10-CCS-41	ZFP536	chr19:35,584,495-35,585,327	TTTGTGGGTGGTAGGTGTGT	GCTGGGAGAGGTAGAACAGG	833
SOX10-CCS-42	BCAS3	chr17:56,264,175-56,264,859	GTCATTTGTCAAACGAAGCAGC	GCACACTTTAAGATCCAAATTCTCC	685
SOX10-CCS-43	BCAS3	chr17:56,683,905-56,684,657	AGACTGCTAGGTTCCCAGCT	CTCTGGAGCCCTGGGTTATG	753
SOX10-CCS-44	CELF4	chr18:33,340,118-33,340,363	TGCCTTCGTGTCTTGAAGCC	CCATGGGCTTGACCTACAGG	246

SOX10-CCS-45	CNTLN	chr9:17,440,149-17,440,784	CAGATTGGCATTTTCAGACCCA	TCTGAAAAATCCACTGAGTTACTGC	636
SOX10-CCS-46	EHBP1	chr2:63,083,637-63,084,587	TCCTACAAGTTGCATTCTGAACT	CAGCATCAAGATGGTATTGTCTCAC	951
SOX10-CCS-47	EHBP1	chr2:63,114,958-63,115,973	ATGGCTTTCAATATTGTATGTCTTGAA	AGGACACATTACTCATTGCTTCAC	1016
SOX10-CCS-48	HAT1	chr2:172,528,652-172,529,540	TGTGAATGAGTTGCAAGGACTG	GTGACACAATTCTTACAGACCTGG	889
SOX10-CCS-49	LRBA	chr4:151,496,516-151,497,064	CCATGTAATACGGCCTTCTTCC	TGCTAAAGTAACTCAGATTCACTGC	549
SOX10-CCS-50	NFIB	chr9:14,293,449-14,293,925	CCCAAGAATCATTGGACGTCT	ATGTCTCCCTGCACTTCACC	477
SOX10-CCS-51	NFIB	chr9:14,299,332-14,299,796	GGAAGGAGTACATGTCCCATCC	GGAAGTGAGTTTCCAAAGCACA	465
SOX10-CCS-52	NFIB	chr9:14,301,911-14,302,364	CCAGCCGATGGGTAATATTAATGG	AAGTGTCAGCCAGTCTTGGG	454
SOX10-CCS-53	POLA1	chrX:24,774,505-24,775,184	CCCTGGTCCTTGTTGGTTCC	TGTGGCTGCTTCTTGGATGG	680
SOX10-CCS-54	SORBS2	chr4:186,930,490-186,930,820	TGCTTGCAATGTTCCCTTGG	GTTTGTAGCCGTGGGATCGA	331
SOX10-CCS-55	TLE4	chr9:81,473,416-81,473,843	TGACAGGCATGACGTTGAGG	ACAATCCTAAGCCAGGGAGAC	428
SOX10-CCS-56	ZFHX3	chr16:71,426,545-71,427,021	GGAGGGTGGGATGTTTGAGG	TTTCCCACCTGCTTCAGTGG	477
SOX10-CCS-57	MPP7	chr10:28,530,884-28,531,299	ATACAGAGCCAGCTCACCAC	TTGGCATGTTCCAGCTGTCA	416

Table A.1. Primers used to amplify the 57 regions from human genomic DNA. Primer sequences are shown in 5' to3' direction. The following Gateway sequences 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCT-3'and 5'-GGGGACCACTTTGTACAAGAAAGCTGGGT-3'were added to forward and reverse primers respectively. Length of regions amplified is shown in bp and do not include the Gateway sequences.

Element ID	Coordinates (rn5)	Forward Primer	Reverse Primer	Size (bp)
Notch1-R1	chr3:9,307,836-9,308,296	AGTCCAGACCTGTGCCTATC	CTTAGCCCCACCAGAAAAGG	461
Notch1-R2	chr3:9,308,175-9,309,096	ACACAGATACTAAGACACGAGC	TTGGTACTGCTGTCTGGGTG	922
Hmga2-R1	chr7:65,390,088-65,391,287	CGAGCCACCTTGATTCAGAG	GTCATCATTCGGTCCTGAGG	1200
Hmga2-R2	chr7:65,427,912-65,428,606	GAGGCCTACGATTGGTGATC	AGCAGGACAAGGGAGATCTC	695
Hes1-R1	chr11:77,415,315-77,415,779	CAACTTCTCGGGTTTTCTGC	AGCTGCTGGGAGGTGAGC	465
Mycn-R1	chr6:51,229,947-51,230,533	CAGCGAGCATTAGAGTCTGC	AATTCTCGTCTGGCATGCAG	587
Id4-R1	chr17:18,701,460-18,702,118	TCATTTGCCCCAGAGAAAGC	AAAGCAAAGAATGGGCCCAG	659
Id2-R1	chr6:53,090,794-53,091,254	Region synthesized by IDT		

Table A.2. Primers used to amplify genomic segment at loci with SOX10 occupancy and open chromatin. Primer sequences are shown in 5' to3' direction. The following Gateway sequences 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCT-3'and 5'-GGGGACCACTTTGTACAAGAAAGCTGGGT-3'were added to forward and reverse primers respectively. Length of regions amplified is shown in bp and do not include the Gateway sequences

Element ID	Forward Primer	Reverse Primer
SOX10-CCS-01	GCCTAATGAGTGTTGTTTAAAGAAGGGAGGGGGAACC	GGTTCCCCCTCCCTTCTTTAAACAACACTCATTAGGC
SOX10-CCS-13	TGCTTTTCTTTCTTTCATTTCACAAAAATTCAGGTTCCATAGACAGGAG	CTCCTGTCTATGGAACCTGAATTTTTGTGAAATGAAAGA AAGAAAAGCA
SOX10-CCS-18	GATCTGAACTGAGCCAAGGAATCCTCGCAGATGG	CCATCTGCGAGGATTCCTTGGCTCAGTTCAGATC
SOX10-CCS-19	TGGTGGTGTATTATGTGTCGCCTCTTGTTCCCC	GGGGAACAAGAGGCGACACATAATACACCACCA
SOX10-CCS-39	GTGTTTCTTCCCTTTGATCCTTCCTGTTCTGATCAAATGAGAAT	ATTCTCATTTGATCAGAACAGGAAGGATCAAAGGGAAG AAACAC
SOX10-CCS-43	CCATCAGAAACTGAACTCTAATGTAATCTGCCCTCTATGTTG	CAACATAGAGGGCAGATTACATTAGAGTTCAGTTTCTGA TGG
SOX10-CCS-51	GCAGGGCTGCCTTCCCGGTTTAACTCTTTCCATG	CATGGAAAGAGTTAAACCGGGAAGGCAGCCCTGC
Notch1-R1	CCACTGAGTTCTATCCGCTCACTCCGTAAGTC	GACTTACGGAGTGAGCGGATAGAACTCAGTGG
Notch1-R2	CTGTGAGCAGTTTATTGGGCCTCTGTCAGACACTAA	TTAGTGTCTGACAGAGGCCCAATAAACTGCTCACAG
Hmga2-R1	CTCCCGGCTACACTTAGGCACTTCATACTTCTCA	TGAGAAGTATGAAGTGCCTAAGTGTAGCCGGGAG
Hmga2-R2	GACAATTCCTGGGCTGGCAGCCTGGTGG	CCACCAGGCTGCCAGCCCAGGAATTGTC
Hes1-R1	GCCGCCGGCTGTGGAACACCGCCCC	GGGGCGGTGTTCCACAGCCGGCGGC
Mycn-R1	ATGGCCTCTTTCTACAGGGTCAGCATGAGGAATG	CATTCCTCATGCTGACCCTGTAGAAAGAGGCCAT
	CAGATATGTCAGCAGTGGACAGACAATGGTCAGCAT	ATGCTGACCATTGTCTGTCCACTGCTGACATATCTG
Id4-R1	TAAACCCTTTTAATCCATGTTTCCTCCGCTGCCCTAC	GTAGGGCAGCGGAGGAAACATGGATTAAAAGGGTTTA
	GTAGGGCAGCGGAGGAAAAACAGCAGTAAATGGAG	CTCCATTTACTGCTGTTTTTCCTCCGCTGCCCTAC
	CAGCAGTAAATGGAGGCCCATGGATTAAAAGGGTTT	AAACCCTTTTAATCCATGGGCCTCCATTTACTGCTG
Id2-R1	GGCAGCGCACTAGGGACTCCCAAGCCGG	CCGGCTTGGGAGTCCCTAGTGCGCTGCC

Table A.3. Mutagenic Primers used to delete the SOX10 consensus sequences. Primer sequences are shown in 5' to3' direction.

rn_SOX6_GSP1	GATTTCTCCAAGAAGTTCACTCG
rn_SOX6_GSP2	TCTCCTCCAGCTTCTTCTGC
rn_SOX6_GSP3	TGGGTCATTGTTTCCTCTCC
rnSOX6_RTPCR_Fwd	GTGCTGTATCTCCCCACAGG
rnSOX6_RTPCR_Rev	TGGGTCATTGTTTCCTCTCC
rnmmB-Actin RTPCR_F	CGCGGGCGACGATGCTCC
rnmmB-Actin RTPCR_R	GTAGCCACGCTCGGTCAGG
rnSox6_RTPCR_F3	GCCAGGAGTCTTCACTGCTCC
rnSox6_3'UTR_R2	GGGAGCGAAATGTCAGAGTG

Table A.4.5'RACE and RT-PCR primers. Primer sequences are shown in 5' to3' direction.

Element ID	Primer
mSox6 ORF F	GGGGACAAGTTTGTACAAAAAAGCAGGCT**TC** ATGTCTTCCAAGCAAGCCACC
mSox6 ORF R	GGGGACCACTTTGTACAAGAAAGCTGGGT**T** TCAGTTGGCACTGACAGGC
SOX6_Seq_F1	CCAACAGCAAGAACAGATCG
SOX6_Seq_R1	ACAGGGCAGGAGAGTTGAGA
SOX6_Seq_F2	GTGAAGTCCCCAACATCTCC
SOX6_Seq_R2	TCCTCCTTTCGTCCTTTGC
SOX6_Seq_F3	CGAAGATGATCCCAAATCAGA
SOX6_Seq_F4	CATCACCACAGGAACAGGTG
SOX6_Seq_R3	GACAGAACGCTGTCCCAGTC
SOX6_Seq_F5	CTGGAAATCCATGTCCAACC
SOX6_Seq_F6	TTGGGGAGTACAAGCAACTG
Sox6 qRT-PCR F3	CTCGGAAGATGCGAGAACAG
Sox6 qRT-PCR R3	GCCCAGTTTTCCATCTTCAC
Neomycin pDEST53 F	GGCTATTCGGCTATGACTGG
Neomycin pDEST53 R	CGTCCTGCAGTTCATTCAGG
hs b-actin qRTPCR F	GTACATGGCTGGGGTGTTG
pDEST53 GFP F	TACACATGGCATGGATGAGC
Sox6 Ex1 qRT-PCR R	TCATTGCTTCCTCTCCATCC

Table A.5. Primers used to amplify the SOX6 open reading frame and primers used for sequencing the construct. RT-PCR primers used for HeLa cell transfections. Primer sequences are shown in 5' to3' direction.