**SOX10-Regulated Promoter Use Defines Isoform-Specific
Gene Expression in Schwann Cells**

by

Elizabeth A. Fogarty

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Neuroscience)
in the University of Michigan
2019

Doctoral Committee:

        Professor Anthony Antonellis, Chair
        Professor Roman Giger
        Associate Professor Shigeki Iwase
        Assistant Professor Jacob Kitzman
        Professor Brian Pierchala, Indiana University

Elizabeth Fogarty

elizafo@umich.edu

ORCID: 0000-0002-0661-7568

Dedication

To my most important person, James.

Tony would brush off as a given, but I know that they are not a given and I have not taken them for granted in my time in his lab.

I also need to thank the members of my thesis committee—Drs. Giger, Iwase, Kitzman, and Pierchala—for the valuable input and time they have given to me over the course of my graduate career. Committee meetings have been nothing but productive and fruitful, and I am grateful for the enthusiasm and generosity with which they have contributed. Dr. Jacob Kitzman and the members of his laboratory, especially Isaac Jia and Bala Burugula, deserve special thanks for the contributions they made to my understanding and execution of next-generation sequencing-based analyses. His willingness to share purified Tn5 enzyme made the Tn5Prime library preparations possible, and for this I am extremely grateful. Similarly, I have to acknowledge Dr. Brian Pierchala and Jen Shadrach for their help with acquiring the rat sciatic nerves used in these analyses. Drs. Jacob Kitzman and Miriam Meisler—and all the members of their laboratories over the years—have provided critical feedback at joint lab meetings, and the project benefited greatly from their ongoing input. Members of the Kalantry, Kidd, Moran, Meisler, and Kitzman labs have been wonderfully collaborative colleagues, always willing to allow us to borrow reagents or equipment in times of crisis. I need to thank the staff of the Neuroscience Graduate Program and the Human Genetics department for personal and administrative support over the years. These include Drs. Ed Stuenkel, Audrey Seasholtz, and Les Satin, as well as Valerie Smith, Rachel Flaten, Carol Skala, and others.

The members of the Antonellis lab have been the most wonderful people to work beside for the last six years. Early on, Bill Law and Laurie Griffin helped me get going in the lab and were incredibly influential during my early years in graduate school. They really helped me understand the trajectory of a PhD through their example as senior students. Chetna Gopinath

and I entered the lab at the same time and worked through much of graduate school side by side, and I am very grateful for the companionship that she provided for the majority of my time. Stephanie Oprescu was an awesome colleague and contributed so much to the lab scientifically and personally. Rebecca Meyer has been a great bay-mate for the last 4 years; I have really enjoyed and will miss the scientific, personal, and pop-culture discussions that we've had while spinning around in our chairs, procrastinating on something else we should be doing. Molly Kuo has been such a wonderful addition to the lab, especially for her kind and thoughtful consideration of everyone around her at all times; she has added immensely to the lab culture and the relationships that we all have. Overall, I am just so thankful to have spent my PhD years in a lab environment where people connect with one another, look out for one another, and genuinely want what is best for one another. It has been great to have not just work colleagues, but friends.

I would like to thank my parents, Kent and Michelle, for their support throughout my life. They raised my sisters and I to be compassionate and instilled in us the importance of giving our best effort to whatever it is we decide to do. Their efforts in raising us have served me well throughout my life and during graduate school. My decisions to move to Michigan and then to have a family have not been easy for them, and I am grateful for their ongoing support. I also want to acknowledge my mother-in-law, Julie, for the help that she has provided us over the last six years, including the many days of watching sick children so that we could go to work. That has been indispensable to help alleviate some of the stress of being a working parent.

I want to thank my daughters. Although they don't realize it, they have also contributed to my graduate career in critical ways. They are the perfect antidote for my tendency to obsess and stress about things happening at school or in the lab. When I get home at the end of every day, I have no choice but to let go of whatever is going on with the science and engage with

work that is much more important: raising them. They have given me a sense of self-worth, as their mother, that keeps me from doubting my worth based on the result of an experiment or funding decision. They dictate a set of non-negotiable priorities, where science is not at the top of the list. All of these things are incredibly freeing and have insulated me from some of the worst stressors of graduate school. It is not easy to be a parent and be a graduate student, but there are things about being a parent that put everything else into perspective, and that is a good thing.

Finally, I have to thank James. When anyone starts graduate school, they don't know what the next five to seven years holds. We certainly had no idea what was in store for us six years ago. We have done (put ourselves through?) so much: worked through being newlyweds (still ongoing?), had a baby, secured a new job, sold a house, bought another house, moved, did home renovations (yay D.I.T.!), had another baby, did more home renovations, had another baby, and we're still doing MORE home renovations, all while I've been working on my PhD. I think it's clear that we have a tendency to over-do things but I love the life that we've built and are building together. James is the most hard-working person I know, a wonderful and thoughtful father and husband, and I can't imagine doing life with anyone else. I love you.

Table of Contents

List of Tables

List of Figures

Abstract

To support the development of diverse cell types, multicellular organisms have adopted mechanisms to tailor gene expression profiles to specific spatial and temporal contexts. These mechanisms—including alternative promoter use—mediate the expression of gene isoforms that are functionally required in a given cell, at a given time. Moreover, the investigation of isoform-specific gene expression profiles in a cell type of interest can provide insights into the biology of the cells and the functions of the associated isoforms. Schwann cells—the myelinating cells of the peripheral nervous system—are an example of a highly specialized cell type with a similarly specialized gene expression profile. SOX10 is a transcriptional activator that is required for the development and maintenance of these cells, and SOX10 target genes are known to be critical for Schwann cell function. Interestingly, SOX10-regulated promoter elements have been characterized and a subset of these regulatory elements dictate the expression of unique gene isoforms. These findings implicate isoform-specific expression of the associated loci in Schwann cell function. This dissertation explores the concept of SOX10-regulated promoter use as a critical factor for isoform-specific gene expression in Schwann cells. These efforts were predicted to identify gene products that contribute to Schwann cell function based on their SOX10-dependent expression in these cells.

To interrogate the function of SOX10 at promoter elements, we employ multiple strategies. First, we describe a sequence conservation-based analysis that identified a SOX10-regulated promoter at a known demyelinating disease gene, *MTMR2*. The SOX10-regulated promoter directs expression of transcripts encoding a unique MTMR2 protein isoform,

suggesting that this isoform may be particularly important for Schwann cell function. Subsequently, we took a broad, unbiased, genome-wide approach to identify SOX10-regulated promoters in Schwann cells. These studies utilize transcription start site (TSS) mapping to assess the activity of SOX10-bound promoters in Schwann cells in an intact adult peripheral nerve, in differentiating primary Schwann cells, and upon ablation of SOX10 expression *in vitro*. The integration of these data provides a comprehensive assessment of SOX10-dependent promoter activity as it relates to multiple models and stages of Schwann cell development. Based on these findings we explore characteristics of SOX10-dependent TSSs, which supports the characterization of these regulatory elements as *bona fide* SOX10-dependent Schwann cell promoters. Finally, we validate and expand upon our findings at four previously unreported SOX10 target genes in Schwann cells, each of which is a multi-TSS locus with a SOX10-regulated promoter: *ARPC1A*, *CHN2*, *DDR1*, and *GAS7*. For each gene we confirm the SOX10-mediated regulatory activity of the promoter and explore the expression of transcript and protein isoforms in Schwann cells. Importantly, these findings suggest roles for the SOX10-regulated gene products in peripheral myelination and—with further study—will likely reveal new insights into Schwann cell biology. As a whole, this dissertation contributes to our understanding of SOX10-regulated isoform expression in Schwann cells and confirms the utility of our strategy toward the identification of high-confidence candidate gene products for further study.

Chapter 1

An Overview of Genomic Complexity and Schwann Cell Biology

Cellular function—and by extension tissue, organ, and organismal function—depends on the ability of cells to transmit the information contained in stretches of deoxyribonucleic acid (DNA) into the functional building blocks of cellular life. This process is described by the central dogma of molecular biology, which clarifies the directionality of information flow from DNA into ribonucleic acid (RNA) and from RNA into protein [1]. Since the discovery of DNA structure [2, 3], decades of research have clarified many of the core components and processes cells use to generate RNA and protein products from gene units. One of the key findings from these decades of study is that mammalian genomes are extremely complex, meaning that many more unique RNA and protein products have been identified than there are genes in the genome. This complexity—the generation of multiple isoforms from a single gene locus—is realized and controlled through a number of regulatory mechanisms; is critical for the specialized gene expression profiles seen across various cell types; and is an important focus of this dissertation.

Schwann cells—the myelinating cells of the peripheral nervous system—provide conductive and metabolic support for motor and sensory axons innervating the limbs and peripheral organs. Importantly, these cells are relatively understudied and much remains to be learned about the genes and gene products that are critical for their function, as underscored by the lack of treatment options for myelin-related diseases. In this dissertation, I will discuss my efforts to identify genes and gene isoforms that are important in Schwann cells by leveraging the

1

critical role of SOX10, a transcription factor that is required for Schwann cell development and function. In this chapter I will begin by discussing the key concepts surrounding the complexity of mammalian genomes and cell type-specific gene expression profiles. I will then describe what is known regarding the development, function, and transcriptional regulation of Schwann cells. Finally, I will emphasize important questions that need to be addressed to improve our understanding of these critical cells.

Basics of Gene Expression

Gene Structure and Regulation

A gene consists of a stretch of double-stranded DNA that is transcribed by an RNA polymerase enzyme to generate a complementary, single-stranded RNA molecule[4]. As it is transcribed, this RNA may undergo splicing, where stretches of the sequence are excised by splicing machinery and the adjacent ends joined together. Regions of the sequence that are included in the RNA after processing are called exons and regions that are removed are called introns. Upon additional processing, which may include capping and polyadenylation, the RNA may perform some functional role as an untranslated (non-coding) RNA and/or be translated by ribosomes into a protein product. In the case of protein-coding transcripts, it is important to note that regions at the beginning and end of the transcript, referred to as 5' and 3' untranslated regions, do not contribute to the protein coding sequence but can have profound impacts on the stability, trafficking, and translation efficiency of the transcript.

Importantly, there are a number of regulatory DNA sequence elements within and surrounding a gene that mediate its transcription (Figure 1.1). The promoter element, an understanding of which is highly relevant for this dissertation, resides upstream of and immediately surrounding the point where RNA polymerase begins transcribing the DNA

2

template, referred to as the transcription start site (TSS). A promoter is critical for the recruitment of RNA polymerase and other basal transcriptional machinery to the proper location on the DNA template for initiation of transcription. A number of particular sequence motifs have been described that contribute to mammalian promoter function. Many of the first viral and eukaryotic promoters sequenced were found to harbor TATA box elements 25 to 30 bases upstream of the TSS[5-7] and this was immediately recognized for a resemblance to the prokaryotic RNA polymerase binding sequence identified by Pribnow in *E. coli*[8]. Indeed, further study confirmed that the TATA box is an important element for the positioning of the transcriptional machinery in relation to the TSS[9]. However, contrary to initial characterizations of the TATA box as a universal promoter element, comprehensive sequencing of mammalian genomes revealed that a minority of mammalian promoters harbor a TATA box. In the human genome, for example, only about 20% of promoters contain a TATA element in the functionally appropriate spacing relative to the TSS[10]. Subsequently, a number of other sequence motifs have been identified to contribute to eukaryotic promoter architecture at various frequencies. Examples include the initiator (INR) sequence located at the TSS[11] and the downstream promoter element (DPE) that resides 30 base pairs downstream of the TSS[12], though each of these are similarly present at a low fraction of mammalian promoters.

Early in gene sequencing efforts it was appreciated that CpG dinucleotides are present at a higher frequency surrounding mammalian TSSs than at other locations across gene bodies[13]. Regions of high CpG frequency came to be called 'CpG islands' (CGIs), and subsequent analysis of the fully sequenced human genome revealed that about 70% of promoters are associated with a CGI[14]. These islands have been associated with a more accessible DNA conformation and are thought to allow the transcriptional machinery to associate with the promoter region more
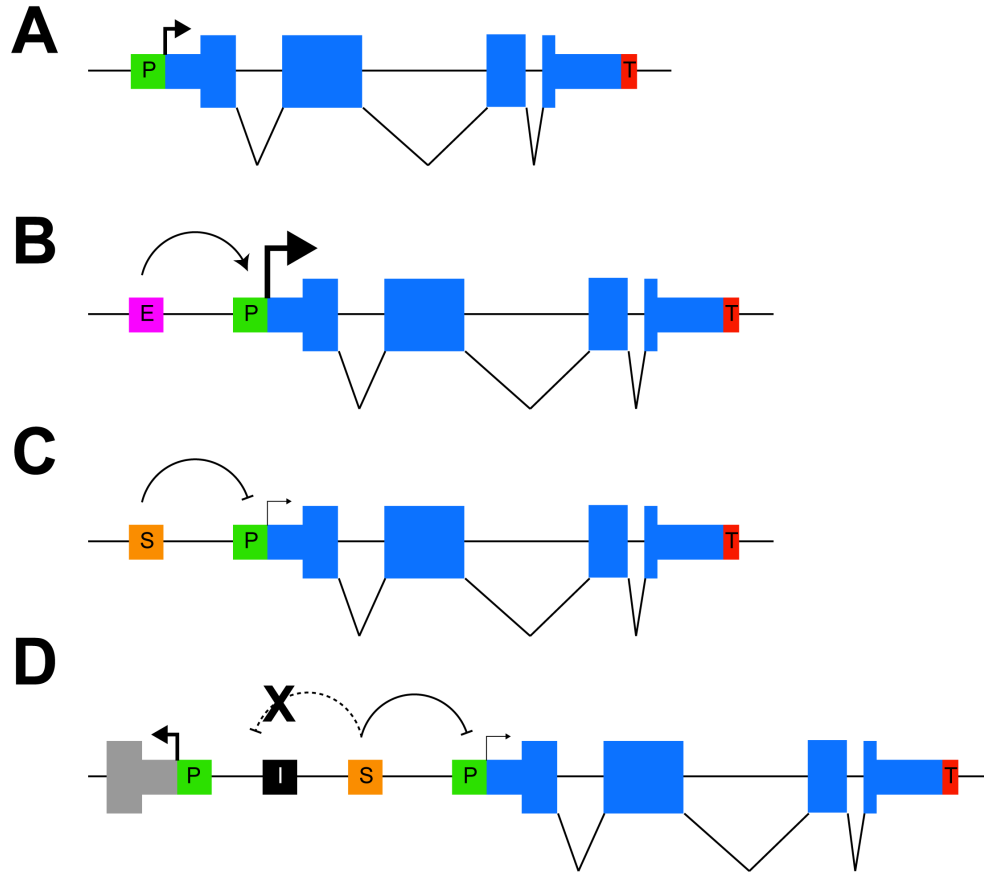
Figure 1.1 Gene structure and regulatory elements. (A) A cartoon schematic of a gene locus. A promoter element (P, green) initiates transcription at a basal level (arrow). Blue bars indicate exons and thin black lines indicate splicing of introns. Thick blue bars indicate protein coding sequence. The terminator (T, red) mediates transcript release from RNA Polymerase II. (B) The presence of an enhancer (E, magenta) induces increased transcription. (C) The presence of a silencer (S, orange) inhibits transcription. (D) Insulator elements (I, black) prevent the regulatory elements for one gene from inappropriately regulating expression of adjacent loci. Adapted from Noonan and McCallion[15].

readily[16]. Overall, perhaps the most interesting finding from decades of efforts to characterize mammalian promoters is their striking diversity.

Apart from the promoter, sequence elements that are important for proper gene transcription include a terminator signal at the end of the gene that induces cleavage of the transcribed RNA strand, releasing it from the polymerase complex[17]. Additionally, *cis*-regulatory elements (CREs) can be located within, nearby, or very far from the gene body and can exert varied effects on gene expression[15]. Enhancers and repressors are two varieties of CREs that increase and decrease gene transcription, respectively (Figure 1.1B and C). These elements are generally understood to function by serving as binding sites for transcription factors that can induce changes to DNA shape and/or regulatory state to increase or decrease the likelihood of transcription. Finally, insulator elements are thought to function as genomic barriers to prevent the CREs and/or regulatory state of one gene or genomic region to inappropriately regulate an adjacent locus or region (Figure 1.1D).

Transcriptional Mechanics

In addition to efforts made toward describing DNA sequence elements that contribute to gene structure and expression, a great deal of study has elucidated the complex and highly regulated processes that mediate assembly of the transcriptional machinery at the promoter, the initiation of transcription, and the elongation of an RNA molecule. First, the RNA polymerase II enzyme, which is responsible for the transcription of protein-coding genes in eukaryotic genomes, must assemble with at least five general transcription factors—TFIIA, -B, -E, -F, and –H—to form the pre-initiation complex (PIC)[18]. At this stage the general transcription factors, particularly TFIID, interact with double-stranded promoter DNA to anchor RNA pol II to the region. TFIIH then acts as a helicase, unwinding approximately 10 base pairs of DNA in an

ATP-dependent manner to form the open complex state. RNA pol II then directly interacts with the single-stranded DNA template and locates the TSS through a mechanism that is not fully understood but is thought to involve a scanning function by RNA pol II and/or other PIC components. Once the PIC is assembled and engaged with the single-stranded DNA template in the open complex state, transcription is initiated and a short nascent RNA strand is formed. It is worthwhile noting that nearly all of the studies describing these early stages of machinery assembly and localization at the promoter have been done at TATA box-containing promoters with TATA Binding Protein as a component of the PIC, as TATA-less promoters have proven difficult to assemble with transcriptional machinery *in vitro* and often do not act as robust substrates for transcriptional initiation in biochemical assays.

One of the seminal findings in the field of transcription biology has been that, after the initiation of transcription, the Pol II complex pauses promoter-proximally[19, 20]; moreover, it has been established that the return of a paused complex into a processive state to transcribe across the gene body—referred to as elongation—is an important and highly regulated aspect of gene expression[21]. Mechanistically, pausing occurs in part due to the interactions of Spt5 and the negative elongation factor (NELF) complex with RNA polymerase II[21]. After dissociation from the general transcription factors, RNA Pol II becomes accessible for Spt5 to bind near the RNA exit channel; this interaction mediates the capping of the nascent RNA, which at this point is about 20 bases long. Subsequently, the NELF complex recognizes and binds to the Pol II-Spt5 interface and mediates the pausing of Pol II using multiple structural mechanisms: restricting nucleotide access to the active site; restraining mobile Pol II domains; and blocking the elongation-promoting RNA pol II-TFIIS interaction[22]. Importantly, the release of a paused transcriptional complex and subsequent elongation into the gene body requires the coordinated

recruitment of positive transcription elongation factor-b (P-TEFb), often in the context of a large super-elongation complex[21]. This recruitment may be accomplished by transcription factors, the Mediator complex, and/or other coactivators. The most critical aspect of P-TEFb recruitment for elongation is that it phosphorylates Spt5, which induces the release of NELF from the complex and thereby allows a return of RNA pol II to the active state and the resumption of gene transcription[23].

Chromatin Structure and Regulation

The double-stranded DNA that makes up the eukaryotic genome does not exist in isolation in the nucleus. Rather, the genome is organized into units called nucleosomes, each of which comprises approximately 147 base pairs of DNA wrapped around a histone octamer scaffold, with 10-80 base pairs acting as a linker between adjacent nucleosomes[24]. After this structural arrangement was first described[25, 26], it came to be appreciated through histone deletion experiments that chromatin structure impacts gene expression[27]. The subsequent identifications of the SWI/SNF chromatin remodeling complex[28-30] and the yeast histone acetyltransferase Gcn5p as the first histone modifying protein[31] ushered in a field of research aiming to understand the regulated dynamics of chromatin states and the effects of these dynamics on cellular function. Since that time, numerous chromatin and histone modifying processes have been described with important impacts not only for gene expression, but also for DNA replication and repair[32]. Briefly, certain post-translational modifications of histone proteins are associated with accessible, actively transcribed regions of the genome (*e.g.*, methylation of lysines 4, 36, and 79 of histone 3) while others are more often seen in condensed and less active regions (*e.g.*, methylation of lysines 9 and 27 of histone 3)[33]. Additionally, chromatin remodeling complexes, including SWI/SNF, are known to physically move

nucleosomes to increase or decrease accessibility of the underlying DNA. Importantly, sequence-specific transcription factors are thought to be important for targeting histone-modifying proteins to the proper regions of the genome to initiate and maintain chromatin states that mediate the proper cell type- and context-specific gene expression profiles.

Mechanisms of Genomic Complexity

One of the critical findings arising from comprehensive transcript sequencing efforts [34] and RNA sequencing experiments[35] is that the number of distinct transcripts is at least an order of magnitude greater than the number of genes in the mammalian genome[36]. This phenomenon is generally described as 'genomic complexity' and is thought to be an important contributor to the evolution of multicellular organisms with diversified cellular functions requiring specialized gene expression profiles. This complexity is realized through a number of mechanisms; a subset of these—particularly those relevant to this thesis—will be considered below.

Alternative Splicing

As noted above, eukaryotic genes harbor stretches of DNA sequence called introns that are transcribed by RNA polymerase II but are then excised from the resulting RNA molecule, usually in a co-transcriptional manner. These splicing events are mediated by five small nuclear ribonucleoprotein complexes that assemble into a functional spliceosome[37] and that are highly dependent on the nucleotide sequences at the exon-intron boundaries called the splice donor and acceptor sites. Additional sequences, termed splice enhancers or silencers, can further modulate the likelihood of a particular splicing event[38]. Importantly, nearly all human loci are capable of generating multiple transcript isoforms on the basis of altered splicing patterns[35]; these most commonly involve exon skipping, alternative splice site selection, intron inclusion, or the use of

mutually exclusive exons[38]. Moreover, context-dependent regulation of RNA splicing is

understood as an important contributor to cell type-specific gene expression profiles[39]. Indeed,

a number of RNA binding factors that are expressed in a cell type- or tissue-restricted profile

have been functionally implicated in alternative splicing [40-42]. Interestingly, transcription and

chromatin state are also known to influence alternative splicing, as exemplified by seminal

experiments showing that promoter variation can impact splicing of a downstream exon[43] (see

'Alternative Promoters' below). Two non-mutually exclusive models have been proposed to

explain the connection between transcription and splicing: one on the basis of recruitment of

splicing factors by the RNA polymerase II transcriptional complex, and the other due to the

speed of elongation that may limit the amount of time for alternative splice sites to be recognized

and engaged by the spliceosome[44].

Alternative Polyadenylation

Another mechanism for the generation of unique transcript isoforms is through

alternative polyadenylation (APA). This arises when multiple cleavage and polyadenylation sites

may be included in the transcripts from a single locus[45]. The most straightforward of these

possibilities is tandem 3' UTR APA, which occurs when there are multiple polyadenylation

signals in the 3'UTR region and the transcripts resulting from APA differ only in the length of

the 3'UTR. In these cases, variation in the 3'UTR is expected to alter the stability, targeting,

and/or translation efficiency of the transcript isoforms. Other examples of APA occur in

conjunction with alternative splicing events and change not only the 3'UTR but also the coding

sequence of the resulting transcripts[46, 47]. In these cases, APA results from: (*i*) splicing into an

alternative terminal exon that harbors a stop codon and polyadenylation sequence; (*ii*) failure to

splice an intronic region that harbors a stop codon and polyadenylation sequence; or (*iii*) the

inclusion of a polyadenylation signal in exonic or intronic sequences that are not preceded by a stop codon and therefore produce a transcript without a 3'UTR.

A number of models have been proposed to explain the mechanisms of polyadenylation site selection[45] and multiple factors contribute. First and most directly, the polyadenylation signal (PAS) was described in the early days of gene structure studies[48]. The strength of the site (*i.e.*, how closely it matches the AAUAAA consensus) affects recognition by the cleavage and polyadenylation specificity factor (CPSF) complex, and alternative polyadenylation sites often harbor PAS variant sequences[49]. Additionally, other *cis* sequence motifs upstream[50] and downstream[51] of the PAS have been shown to modulate site recognition and selection. Furthermore, in humans there are more than 80 proteins that contribute to the cleavage and polyadenylation machinery, and it has been shown that changes in the levels of these factors can influence APA[52, 53]. Other factors thought to contribute to APA include transcriptional speed, splicing factors, and RNA-binding proteins[45].

Alternative Promoter Use

A critical contributor to the complexity of mammalian genomes is alternative promoter use, or the use of multiple transcription start sites—each associated with a distinct promoter element— at a single gene locus. One of the earliest reports supporting the idea of alternative promoter use for a mammalian gene was at the mouse α-amylase locus[54]. Since then, it has become well-appreciated that many loci employ the use of multiple promoters to generate unique transcript and/or protein isoforms from a single transcriptional unit[55], and that these elements can have profound impacts on the biological function of the locus. The use of an alternative promoter necessarily impacts the sequence of the resulting transcript at the 5'UTR and may cause a change to the protein coding sequence as well (Figure 1.2). Further, the activity of
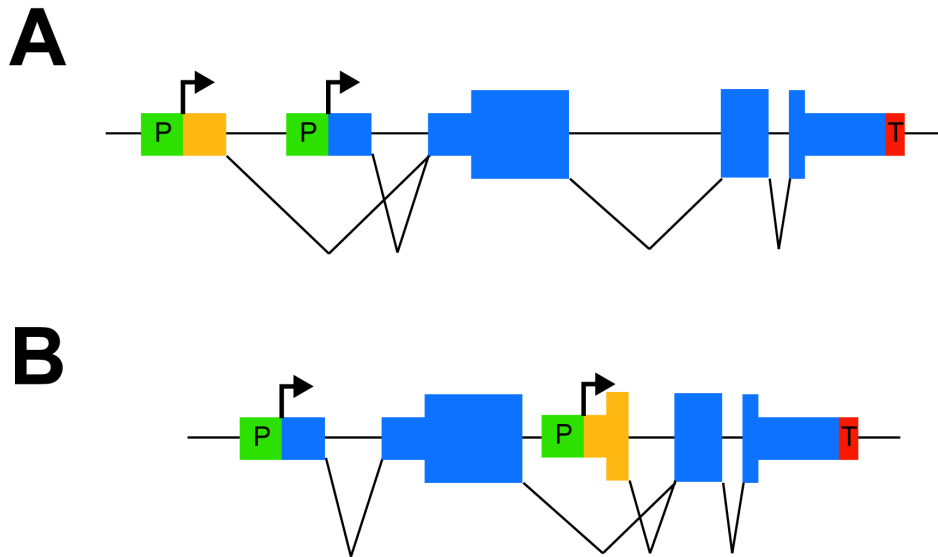
Figure 1.2 Alternative promoter use dictates transcript architecture. (A) A locus with two promoters (P, green) where promoter use dictates variant 5' untranslated regions of the resulting transcripts. (B) A gene with two promoters where promoter use alters 5' untranslated regions and protein coding sequences of the resulting transcripts.

alternative promoter elements may induce higher or lower levels of gene expression and may be associated with unique expression patterns in particular tissues or cell types (see below).

Interestingly, data suggest a connection between alternative promoter use and downstream alternative splicing events. For example, at the *NOS1*, *CASP2*, and *CFTR* loci there are relationships between promoter choice and the inclusion or exclusion of downstream alternatively spliced exons[56-58]. Moreover, large scale assessments of promoter use and downstream splicing also support a relationship between promoter selection and alternative splicing on a genome-wide level[59]. These effects are thought to occur through two possible mechanisms. First, it is possible that promoter elements recruit factors with dual roles in transcription and splicing, and that these factors will associate with the transcriptional machinery as it processes down the length of the gene body, then modulate downstream splicing events. This type of process has been described for a handful of factors including PGC-1[60]. Second, some evidence suggests that a promoter element can dictate splicing indirectly through modulating RNA pol II elongation rates; promoter elements that support faster elongation rates are thought to promote the exclusion of exons with weak splice site sequences by reducing the time window for their recognition by the splicing machinery. In support of this 'kinetic' model, the expression of an RNA Pol II mutant with a slower elongation rate was shown to affect alternative splicing patterns in human cells[61]. However, it is worthwhile to note that a definitive link between promoter function and downstream splicing events has yet to be proven, and that the rate of transcriptional elongation was first noted to affect splicing in the context of changes to the RNA secondary structure[62]. Because alternative promoters confer changes to the sequence content of the RNA, it will be important for future studies to disentangle the

possible confounding effects on splicing that could be conferred not by transcriptional dynamics but by changes to the structure of the RNA as a result of a novel 5' sequence[63].

Alternative Promoter Use and Cell Type-Specific Gene Expression

Genomic complexity is critical for the development and function of diverse and specialized cell types in multicellular organisms. Each of the mechanisms outlined above that mediate genomic complexity can be applied in particular cellular and/or developmental contexts to modulate cell type-specific gene expression profiles. For example, cell type-specific components of the basal transcriptional machinery have been shown to modulate gene expression profiles in differentiating germ cells[64]. Similarly, cell type-specific expression of transcription factors, splicing factors, and subunits of chromatin modifying complexes are all known to contribute to the specialization of gene expression profiles in specific cellular and developmental contexts[65-67].

Importantly, with the advent of high-throughput transcript sequencing approaches, alternative promoter use has come to be appreciated as a widespread and critical contributor to cell type-specific gene expression profiles. Indeed, large-scale assessments of transcriptional diversity indicate that alternative promoter use contributes more toward the transcriptional diversity of mammalian genomes than does alternative splicing[68-70]. Moreover, TSS mapping techniques have elucidated the wide extent of alternative promoter use, with each human gene harboring an average of four transcription start sites[71]. Comparisons of TSS use across many tissues indicate that the vast majority of these TSSs (80%) are utilized in a restricted (*i.e.*, non-ubiquitous) manner, supporting the idea that alternative promoter use is an important contributor to cell type-specific gene expression. Indeed, this dissertation seeks to better understand how

13

promoter use contributes to the cell type-specific expression profiles that are important in one understudied cell type and thereby provide novel insights into the biology of these cells.

Schwann Cell Biology

As the myelinating cells of the peripheral nervous system, Schwann cells mediate the fast conduction of action potentials down the lengths of peripheral axons and are necessary for sensory and motor function in the limbs and peripheral organs. Importantly, Schwann cells are a relatively understudied cell type and much remains to be learned about the development and function of these cells, as well as how Schwann cell dysfunction plays a role in human disease.

Schwann Cell Developmental Lineage

Schwann cells are developmentally derived from the neural crest, a population of migratory and multipotent cells that delaminate from the neural tube early in vertebrate development and migrate throughout the body to generate many diverse cell types[72]. In particular, Schwann cells are generated from trunk neural crest cells that migrate along the ventral pathway, travelling through the anterior sclerotome after lamination. During migration, neural crest cells integrate a number of intrinsic and extrinsic cues that inform their specification. Once the specification of Schwann cell precursors occurs (Figure 1.3A), these cells continue to migrate and proliferate. At this stage, the precursor cells depend on developing peripheral axons for survival and migration cues as they travel along them[73]. Interestingly, it has been shown that peripheral neurons are similarly reliant on Schwann cell precursors for survival, consistent with mutually supportive trophic signaling[74].

Schwann cell precursors next undergo a transition into the immature Schwann cell stage (Figure 1.3B). At this point, the cells stop migrating and survival becomes dependent on autocrine signaling rather than axonal cues[75]. The molecular mechanisms governing the
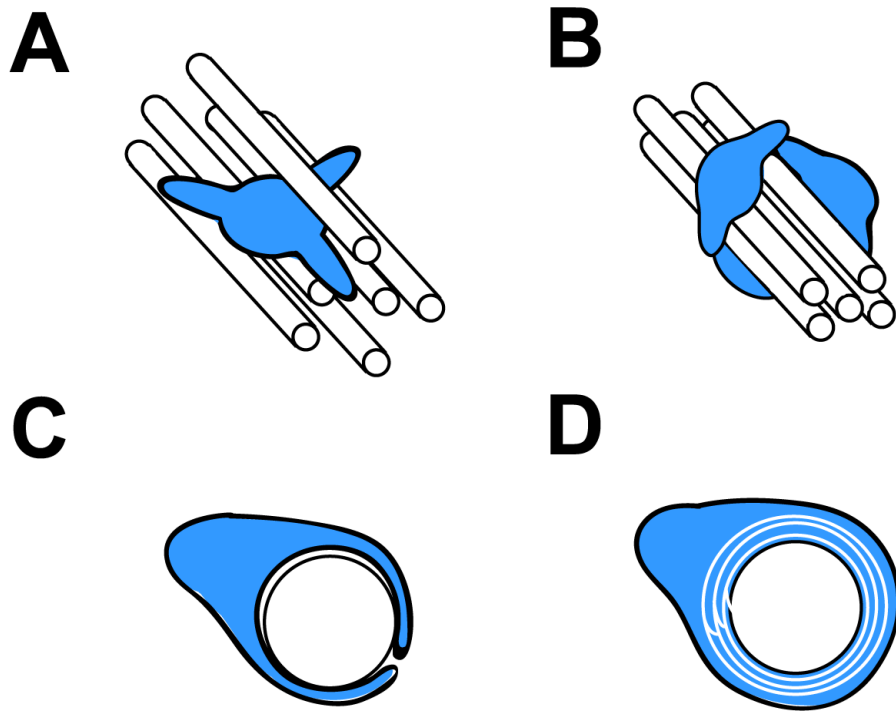
Figure 1.3 Schwann cell development. (A) Schwann cell precursors migrate and proliferate alongside developing peripheral axons. (B) Immature Schwann cells bundle axons and extend processes to perform radial sorting of large-caliber axons. (C) A promyelinating Schwann cell comes into a one-to-one relationship with an axon and subsequently ensheathes the axon with a multi-layered, compact myelin sheath (D). Adapted from Jessen and Mirsky[76].

transition into the next developmental stage—immature Schwann cells—are not well-defined, but it has been shown that modulating Notch signaling alters the generation of immature Schwann cells *in vivo*[77]. Immature Schwann cells play critical roles for peripheral nerve development. One role is the production of signaling molecules that drive the differentiation of surrounding mesenchymal cells to form the arteries and peri- and epineurial sheaths of the mature nerve[78, 79]. Another critical function of immature Schwann cells is to perform radial sorting of the developing axons such that axons that will be myelinated are physically separated from those that will not; this distinction is made based on axon diameter[80]. A seminal report by Webster and colleagues describing the morphology of nerves during this developmental time frame greatly influenced our understanding of the radial sorting process[81]. Briefly, immature Schwann cells move into the nerve and a group of 3-8 cells surrounds a bundle of axons. The Schwann cells extend large lamellipodia-like processes into the bundle, which interact with large axons and move them toward the periphery. It is thought that the immature Schwann cell then divides and the daughter cell remains associated with the axon as it defasciculates from the bundle. It is critical that this process brings Schwann cells and the large caliber axons into a 1:1 relationship for subsequent myelination. The small axons that remain bundled after sorting remain unmyelinated and are ensheathed by non-myelinating Schwann cells in Remak bundles. Cellular components that are required for radial sorting include extracellular matrix molecules and the associated receptors, as well as proteins involved in cytoskeleton regulation, cellular polarity, and cell cycle exit[82].

When a developing Schwann cell establishes a 1:1 relationship with a large caliber axon, it is referred to as a promyelinating Schwann cell (Figure 1.3C). Although the cell is properly positioned and poised for myelination at this stage, a number of signaling processes are required

for the commencement and completion of myelination (Figure 1.3D). One of the first of these processes to be described is axonal neuregulin-1 (NRG1) type III-mediated activation of Schwann cell ErbB2/3 receptors. Although this signaling axis is important for Schwann cells at many stages of development[83], it has been shown that loss of this signaling specifically in promyelinating Schwann cells reduces myelination[84] while neuronal overexpression of NRG1 induces hypermyelination[85]. Other events important for myelination are activation of neuronal ADAM22 by Schwann cell-secreted LGI4[86]; the downregulation of NOTCH-1 by the transcription factor EGR2[77] (see more below); and activation of G-protein coupled receptors including GPR126[87], GPR44[88], and LPA1[89]. These processes activate a number of downstream signaling cascades in Schwann cells including those associated with PI-3 kinase, MAPK, cAMP, and calcineurin[82]. Thus, many distinct extracellular and intracellular events are important for the proper development of Schwann cells and peripheral myelination, and it is likely that still more remain to be described.

Functions of Myelinating Cells

The primary function of Schwann cells and oligodendrocytes—the latter being the functionally related, myelinating cells in the central nervous system—is to mediate the rapid propagation of action potentials along axonal segments. Apart from the evolutionary advantage that is conferred by the ability to integrate a signal from the environment and respond to that information quickly, efforts have been made to understand if there are additional benefits to the evolution of myelin in jawed vertebrates. Indeed, recent work suggests that there are multiple ways, apart from increased nerve conduction speed, in which myelinating cells contribute to nervous system and organismal function.

It has been posited that an additional benefit of myelination is tied to conservation of energy[90, 91]. Indeed, the maintenance of the neuronal resting potential through Na^+/K^+ exchange is an ATP-dependent process that is energetically costly. By limiting the extent of sodium influx needed for action potential propagation to only intermittent sites (*i.e.*, the nodes of Ranvier), it follows that myelination is energetically favorable. However, an analysis of energy dynamics in the central nervous system found that when the energetic costs of oligodendrocyte development and maintenance are taken into account, myelination as a whole does not save the organism energy[92]. It is worthwhile to consider that these analyses were specific to the central nervous system and may not be applicable to the energetics of myelination in the peripheral nervous system. Nonetheless, these findings argue against an energetic advantage driving the evolution of myelin.

Recently, an important focus of the field of myelin biology has been to understand how—and even if—myelinating cells provide trophic support to the underlying axons[93]. This is considered a likely possibility, given: (*i*) the extreme lengths separating some axons from the neuronal cell body, especially in the peripheral nervous system; and (*ii*) the restriction of large portions of the axonal membrane from access to the extracellular environment due to occlusion by the myelin sheath[94]. Myelinating glia have been posited to provide a number of different molecules to axons in the contexts of steady state maintenance and post-injury recovery; these include metabolites necessary to support the energetic demands of the axon[95], ribosomes and mRNAs to support local protein expression[96], and growth factors[97]. Mechanisms have also been suggested for routes by which supportive factors may be delivered from myelinating glia to the axons they ensheath. One of the most intriguing possibilities is based on descriptions of Schwann cells[96, 98] and oligodendrocytes[99, 100] releasing cargo-carrying extracellular

vesicles that are taken up by axons in certain contexts. Though these findings have not been definitely proven and the relevance is not yet clear, this is certain to remain an active area of investigation for the foreseeable future.

Schwann Cells in Human Disease

Given the importance of myelination for nervous system function, it comes as no surprise that Schwann cell dysfunction causes human disease phenotypes that involve reduced sensory and/or motor function in the periphery. Peripheral neuropathies can result from: genetic causes (see below); complications of acquired disease (most commonly type 2 diabetes); systemic disorders related to metabolism, inflammation, or autoimmunity; and toxic exposures including chemotherapy agents, alcohol, and heavy metals[101]. Examples of inherited disease of the peripheral nerve include hereditary motor and sensory neuropathies (HMSN; also called Charcot-Marie-Tooth [CMT] disease), hereditary sensory and autonomic neuropathies (HSAN), hereditary motor neuropathies (HMN), and hereditary recurrent focal neuropathies[101]. HMSN (or CMT disease) is an umbrella term for a genetically and clinically heterogeneous group of disorders and is the most common of the inherited neuropathies, with reported frequencies that vary geographically from 1 in 2,500 to 1 in 12,500 when all CMT subtypes are included[102]. CMT disease can be broadly divided into two classes: axonal CMT disease, which arises from a primary defect in the sensory and motor neurons, and demyelinating CMT disease, which arises from a primary defect in the Schwann cells.

Of the two classes of CMT disease, demyelinating disease accounts for the majority of cases[103]. Affected individuals typically present with slowly progressive symptoms including muscle wasting in the distal extremities, abnormal gait, reduced reflexes, and sensory loss. Clinical assessment reveals reduced nerve conduction velocities (below 38 meters per second)

and characteristic signs of demyelination and/or hypertrophic remyelination from nerve biopsy[104]. Demyelinating CMT disease can follow dominant, recessive, and X-linked inheritance patterns. Advances in the identification of genetic causes mean that 60 to 90 percent of individuals with demyelinating CMT disease achieve a genetic diagnosis[103, 105, 106]. Genes implicated in disease include those encoding myelin constituent proteins (*e.g.*, *PMP22*, *MPZ*), regulators of intracellular vesicle trafficking (e.g., *LITAF*, *MTMR2*, *MTMR5*, *MTMR13*, *FIG4*), and other proteins that have specific, critical functions in Schwann cells (*e.g.*, *EGR2*, *GJB1*)[107]. Importantly, the identification of demyelinating disease genes and subsequent locus-specific functional studies have been an important catalyst toward understanding Schwann cell biology. However, there are no effective treatments for patients with Schwann cell-related disease. The pursuit of a more complete knowledge of Schwann cell development and function may contribute to the future development of therapeutic avenues.

The Role of SOX10 in Schwann Cells

Given the need to better understand the biology of Schwann cells, the identification of genes and proteins that contribute to the development and function of these cells has been an ongoing effort. Importantly, studies spanning the last 30 years have identified a number of transcriptional regulators that are important for the development of these cells. Among these, SOX10 is a transcription factor that has been shown to be essential for Schwann cell development and maintenance, and SOX10 function is the focus of this dissertation.

Basics of SOX10 Biology

SOX10 is a member of the SOX (Sry-related HMG-box) transcription factor family, the members of which have broad roles in embryonic development and cell fate specification[108]. The SOX family of transcription factors is defined by the presence of a conserved, high-mobility

group (HMG) DNA binding domain that interacts with the minor groove of DNA in a sequence-specific manner, bending it toward the major groove[109]. Vertebrate genomes harbor approximately 20 SOX family members; these are subdivided into groups based on HMG domain sequence similarity[110]. SOX10 belongs to the SOXE group, along with SOX8 and SOX9. SOXE factors exhibit shared expression profiles in certain tissues, particularly the embryonic neural crest and the peripheral and central nervous systems, and have some overlapping functions but are not functionally redundant[111].

SOX10 was first identified through the RT-PCR-mediated amplification of the HMG domain in RNA isolated from 11.5 days post coitum mouse embryos[112]. Subsequently the early embryonic expression of *Sox10* was localized to the premigratory neural crest [113, 114]. This expression is important for neural crest cell identity and survival, as exemplified by the apoptotic loss of neural crest cells in a *Sox10* mutant mouse line[115]. SOX10 expression persists in migrating neural crest cells, where SOX10 is important for the maintenance of the multipotent state and subsequent specification of the neural crest into various cellular lineages[116], including peripheral neurons, glia, and melanocytes[117, 118]. Expression of SOX10 is also detected in the developing central nervous system, specifically in oligodendrocyte precursors where it plays an important role in the differentiation of those cells[113, 119]. In the adult, SOX10 expression is restricted to the myelinating cells of the central and peripheral nervous systems[113] and melanocytes in the skin[120].

The SOX10 protein structure includes a number of annotated functional domains (Figure 1.4A). The dimerization domain mediates DNA-dependent dimerization of SOX10 at sites with two properly spaced motif sequences arranged in a head-to-head manner[121] (Figure 1.4B). Based on *in vitro* experiments, this dimerization is thought to increase the affinity and/or

21
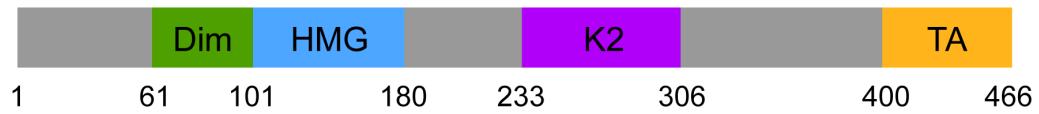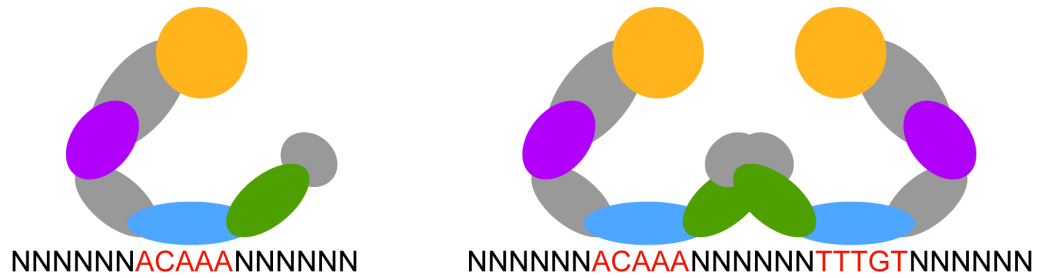
Figure 1.4 SOX10 protein structure and DNA binding. (A) The SOX10 protein sequence includes a dimerization (Dim, green) domain, high mobility group (HMG, blue) DNA-binding domain, and two transactivation domains (K2, purple and TA, orange). The numbers along the bottom indicate amino acid positions. (B) SOX10 binds to consensus motifs (red text) as a monomer (left), or when two motifs are arranged in a head-to-head manner, as a dimer (right).

duration of the SOX10-DNA interaction and induce a greater degree of DNA bending relative to monomeric binding[121]. Further, there is evidence that the dimerization domain supports not only the formation of SOX10 homodimers but also heterodimerization with other SOXE group members[122, 123]. As mentioned above, the HMG domain mediates sequence-specific DNA binding, with a consensus binding sequence of 5'(A/T)AACAA(T/A)3' that is common to all HMG domains[124]. The SOX10 K2 domain was defined in SOX8 as a transactivation domain[125] and has been shown to be important for SOX10 function *in vivo*, albeit in a cell type-specific manner[126]. Finally, the C-terminal region of SOX10 encodes a critical transactivation domain[127]. The K2 and C-terminal transactivation domains have recently been proposed to interact synergistically[128].

Mechanisms of Transactivation by SOX10

Interestingly, SOX10 and other SOXE group members have been shown to be functionally diverse transcriptional effectors, and can modulate expression of target genes via myriad mechanisms[129]. First, the bending of DNA by the HMG domain is likely an important architectural contributor to the gene regulatory activity of SOX factors, especially in the context of binding at distal regulatory regions where DNA looping would be important for bringing regulatory factors in close proximity to target promoter regions[129, 130]. Second, it is worth noting that the SOXE proteins harbor transactivation domains and not repressor domains; based on this and the overwhelming majority of functional studies to date, SOXE factors are thought to function primarily as transcriptional activators. However, this does not definitively exclude the possibility that the factors could function as repressors in as-yet unspecific contexts.

Nonetheless, for the purpose of this dissertation we will assume that SOX10 functions to positively regulate target gene transcription.

Many transactivation events induced by SOX10 are thought to rely on interactions with other transcription factors; these interactions are thought to synergize the activities of transactivation provided by each factor but also to play an important role in refining the spatial and temporal regulation of SOX10 target genes[129]. Further, many of these interactions have been mapped to the SOX10 C-terminal transactivation domain (see below for discussion of Schwann cell-specific examples). Although the structures and activities of the SOXE C-terminal transactivation domains are not well-defined biochemically, it is anticipated that the domain exhibits profound functional flexibility to mediate the wide variety of interactions described to date[129]. Interestingly, it has been reported that in addition to mediating DNA binding and bending, the SOX10 HMG domain is another site of cofactor interactions. Indeed, this domain is reported to mediate weak protein-protein interactions between SOXE proteins and numerous other transcription factors[131]. The C-terminal region of the HMG domain is solvent-accessible even during DNA binding[132], and this is thought to mediate the availability of this region for protein interactions. Further study is needed to clarify the relevance of these findings, but the authors suggest that this contributes to the selection of context- and cofactor-specific target genes and that it could also reflect a mechanism for sequestration of other DNA-binding proteins away from their targets (these interactions invariably occurred at the DNA binding domain of the interacting factor)[131].

In addition to interactions with transcriptional cofactors, SOX10 also modulates transcription by interacting with the general transcriptional machinery. For example, both the dimerization/HMG and C-terminal transactivation regions of SOX10 bind the regulatory Med12

subunit of the RNA polymerase II Mediator complex[133, 134]. This interaction suggests an ability of SOX10 to regulate the assembly of pre-initiation complexes at target promoters through recruitment of Mediator. However, Mediator has been shown to regulate many aspects of transcription and chromatin biology so multiple mechanisms may be at play[134, 135]. Additionally, SOX10 interacts with the transcriptional elongation factor P-TEFb in Schwann cells[136]. This interaction was mapped to the SOX10 C-terminal transactivation domain and implicates SOX10 in the regulation of not just transcription initiation but also elongation. Importantly, the interactions of SOX10 with the Mediator complex and P-TEFb support the ability of SOX10 to directly influence transcription initiation and elongation to promote expression of SOX10 target genes.

Finally, SOXE factors have been implicated in the regulation of epigenetic states through the recruitment of chromatin modifying complexes and enzymes. For example, SOX10 interacts with the Baf60a subunit of the SWI/SNF chromatin remodeling complex[137]. This interaction is mediated by the dimerization/HMG domain of SOX10 and follow-up studies confirmed the recruitment of the catalytic core of the SWI/SNF complex (Brg1) to genomic sites of SOX10 binding. Additionally, SOX10 interacts with histone-modifying enzymes including histone deacetylases (HDAC1 and 2); HDAC2 directly influences expression of SOX10 target genes and synergistically enhances expression of targets with SOX10[138], though this interaction has not been mapped to a specific SOX10 domain. Thus, a wealth of evidence supports the idea that SOX10 regulates gene expression using different mechanisms, ranging from interactions with sequence-specific or general transcription factors to modulation of transcriptional elongation and the recruitment of chromatin-modifying proteins.

Evidence for SOX10's Importance in Schwann Cells

There are many lines of evidence to support the idea that SOX10 plays a critical role in Schwann cells. First, human mutations in *SOX10* have been implicated in demyelinating disease phenotypes. Specifically, a subset of *SOX10* mutations have been associated with a complex dominant disorder wherein patients present with compromised myelination in the **p**eripheral and **c**entral nervous systems in addition to exhibiting symptoms of **W**aardenburg syndrome (*i.e.*, hearing loss and pigmentation defects) and **H**irschsprung disease (*i.e.*, enteric aganglionosis); this syndrome is abbreviated PCWH [139, 140]. Importantly, this constellation of phenotypes is consistent with the cellular expression profile of SOX10 and the function it has in mediating the specification of neural crest-derived cell types. In contrast to the PCWH presentation, other *SOX10* mutations are associated with more restricted dominant phenotypes that are consistent with Waardenburg syndrome and/or Hirschsprung disease but do not affect myelination[141]. An analysis of the genotype-phenotype correlations informed a theory to explain the molecular mechanisms of pathogenic *SOX10* mutations[140], most of which induce premature stop codons. Namely, a premature stop codon that results in nonsense-mediated decay of the *SOX10* transcript functions as a null allele, and haploinsufficiency for *SOX10* results in the restricted Waardenburg- and Hirschsprung-related phenotypes. However, the authors noted that many of the mutations associated with the broader PCWH phenotype fall in the last exon and are likely to escape nonsense-mediated decay. Therefore, these alleles are thought to generate truncated SOX10 protein products that exert a dominant-negative effect to cause the expanded PCWH syndrome.

While this is an attractive theory to explain the mechanisms of SOX10-related disease, additional lines of evidence suggest that this is not an absolute rule and that other factors may

modulate the phenotypes associated with *SOX10* mutations. First, there are examples of human frameshift mutations that reside in the last exon of *SOX10* that are not associated with neurological symptoms indicative of myelination defects[142]. Additionally, it has been observed that the severity of phenotypes associated with *Sox10* mutations in mouse models varies depending on the genetic background, with genetic modifiers of pigmentation and aganglionosis identified[143-145]. Further, a mouse model was recently generated to constitutively express one of the PCWH-associated truncation mutations (Q377X) in the last coding exon of *Sox10*[146]. In this study the authors report that heterozygous Q377X mice do not exhibit any prominent myelination defects, though it is worth noting that this has not been tested in other genetic backgrounds where SOX10-associated phenotypes can be more severe. As a whole, these data suggest that genetic and/or environmental factors contribute to the manifested phenotypes associated with *SOX10* mutations and further study is needed to fully elucidate mechanisms and modulators of SOX10-related disease.

Importantly, substantial effort has been devoted to the study the specific role(s) of SOX10 in Schwann cells. As noted above, SOX10 is expressed in neural crest cells as they migrate and begin to specify, and this expression is required for the survival of these cells[115]. In the majority of the resulting lineages, SOX10 expression is subsequently lost. However, in a few cell types, including Schwann cells, SOX10 expression persists through specification and differentiation and is maintained in adulthood[113]. This expression profile is suggestive of developmental and post-developmental roles for SOX10 in the function of these cells. Subsequently, mouse models have confirmed the critical role for SOX10 in this cell lineage. In a series of experiments wherein *Sox10* expression was conditionally ablated at different developmental time points of the Schwann cell lineage, Michael Wegner's group established that

SOX10 is required at every stage of Schwann cell development. Namely, gene inactivation in immature Schwann cells results in an arrest of cell development and failure of radial sorting[147], loss in pro-myelinating Schwann cells induces hypomyelination[148], and *Sox10* ablation in fully mature myelinating cells results in a loss of the differentiated cellular identity and demyelination[149]. Taken together, the human and mouse genetic studies provide convincing evidence regarding the critical nature of SOX10 expression for the entirety of the Schwann cell lineage.

Role for SOX10 in the Transcriptional Regulation of Schwann Cells

In conjunction with the lines of evidence that confirm the requirement for SOX10 in Schwann cells, progress has been made toward describing the molecular mechanisms of SOX10 function in peripheral myelination. As SOX10 is expressed and necessary at every stage of Schwann cell development, it can be expected that SOX10 induces expression of target genes that contribute to lineage survival and/or progression at each stage. Indeed, SOX10 sits atop a well-described transcriptional hierarchy in Schwann cells, inducing the expression of factors at each stage that drive the progression of the developmental trajectory. After specification from the neural crest, SOX10 induces the expression of targets that mediate the survival and proliferation of Schwann cell precursors. These targets include *Erbb3*, a receptor for neuregulin signaling[74, 150]. Subsequently, in the immature Schwann cell stage SOX10 mediates the induction of the transcription factor OCT6[151]; this is important for lineage progression, as reduced OCT6 expression is associated with a delay in the transition into the promyelinating stage[152]. Indeed, SOX10 and OCT6 act cooperatively to activate the expression of the critical transcription factor EGR2 in promyelinating Schwann cells[153]. Finally, SOX10 and EGR2 cooperate at myelin-related target genes and are each required for the terminal differentiation of Schwann cells and

the maintenance of peripheral myelin[149, 154-157]. Loci co-regulated by SOX10 and EGR2 in myelinating Schwann cells include genes encoding constituent myelin proteins such as myelin basic protein (*MBP*) and myelin protein zero (*MPZ*)[158, 159].

Importantly, SOX10 target genes that have been characterized in Schwann cells include a number of genes that have themselves been implicated in demyelinating disease. These include *MBP*, *MPZ*, *GJB1*, *EGR2*, *PMP2*, *PMP22*, and *SH3TC2*[160-166]. The disease associations held by each of these loci implicate their gene products as critical for the proper development and/or function of Schwann cells, and the regulation of these genes by SOX10 provides a leverage point toward a broader understanding of Schwann cell biology. Namely, based on the knowledge that SOX10 regulates many genes that are critical for myelination, it can be hypothesized that efforts to identify additional, novel SOX10 target genes in these cells will provide insights into pathways that contribute to Schwann cell biology. Thus, ongoing efforts in the field have been geared toward identifying SOX10 target genes and in defining the roles of these loci in peripheral myelination.

SOX10 Activity at Promoter Elements

SOX10 is known to regulate expression of Schwann cell loci by binding directly to a promoter element; examples of loci with SOX10-regulated promoters include *MBP*, *MPZ*, *GJB1*, *S100B, CNTF, PMP2,* and *SH3TC2*[160-162, 165-168]. Indeed, upon the generation of a dataset describing the genome-wide binding profile for SOX10 in rat sciatic nerve, the authors noted the preponderance of SOX10 binding at promoter elements[154]. Specifically, among loci that exhibit reduced expression upon *Sox10* knockdown *in vitro* and that harbor a SOX10 ChIP-Seq peak within 100 kilobases, 40% show SOX10 binding within the 2 kilobases upstream of the transcription start site. An analogous dataset and analysis for EGR2, on the other hand, indicated

promoter binding by that factor at less than 5% of target loci. These findings suggest a prevalent promoter-proximal regulatory function for SOX10 in Schwann cells. Moreover, the fact that SOX10 is known to regulate the promoters of critical myelin-related genes and demyelinating disease genes further supports the idea that the broader identification of SOX10-regulated promoters throughout the genome is likely to provide insights into new genes and transcripts that are important for Schwann cell biology.

Even more intriguing, however, are data showing that SOX10 induces the activity of specific promoters at loci with multiple TSSs; here SOX10 directs the expression of specific gene isoforms in Schwann cells. Examples include *SOX6*, *SH3KBP1*, *GJB1*, and *AATK*[169-171]. The *AATK* locus is a particularly interesting case wherein SOX10 regulates an alternative promoter that maps upstream of an intron harboring microRNA-338[171]. This microRNA (miRNA) was previously implicated in central nervous system myelination, where it is responsible for downregulating inhibitors of myelination as oligodendrocytes differentiate[172]. Thus, it is likely that the SOX10-regulated promoter at *AATK* mediates the expression of this miRNA in myelinating cells, though it is unclear if the unique *AATK* mRNA and protein products generated from this promoter also have functional properties that are important for Schwann cell function. Nonetheless, at each locus where SOX10 regulates a specific TSS, interesting questions are raised by the SOX10-regulated activity of alternative promoters: Is there isoform specificity for the target locus in Schwann cells? Are there functional attributes of the RNA and protein products produced from this promoter that make them particularly important in Schwann cells? How is promoter use regulated across the course of Schwann cell development? These are key questions driving the work presented in this dissertation.

Clinical Utility of Characterizing SOX10-Responsive Regulatory Elements in Schwann Cells

It is worthwhile noting that the identification of a SOX10-regulated promoter or enhancer element at a locus that is important for Schwann cell function is not strictly relevant only to academic endeavors or toward an understanding of basic biology. Indeed, the careful characterization of a critical non-coding regulatory region at a functionally important gene, and especially at a disease-associated gene, expands the genomic 'real estate' that may be screened for genetic variation mediating human disease. Namely, if an individual presents clinically with a phenotype consistent with variation at a hypothetical *geneA* and yet genome sequencing does not identify any protein-coding variation at the locus, relevant regulatory regions—including the *geneA* promoter region—may be screened for potentially deleterious variation affecting gene expression.

Strikingly, this is not purely a theoretical point but rather has already been proven clinically relevant in the realm of SOX10 biology and demyelinating disease. The *GJB1* gene encodes Connexin 32 and mutations at this locus cause X-linked demyelinating CMT disease[173]. Since the identification of *GJB1* as a causative disease gene, multiple groups have described families presenting with X-linked demyelinating disease wherein no Connexin 32 coding mutations were identified. Upon further sequencing, mutations and deletions in the SOX10-regulated promoter region of this locus—in one case a mutation directly affecting a SOX10 binding motif—were identified in these families as segregating with disease and functional assays revealed reduced activity of the mutant promoters[174-177]. These findings emphasize the potential for the characterization of non-coding regulatory elements to prove valuable not only for research applications, but also in clinical genetics.

Summary

Gene expression is a complex, dynamic, and highly-regulated process in mammalian genomes. It can be modulated by many mechanisms, including through chromatin and histone modifications, actions of transcription factors that alter the initiation and elongation activities of the transcriptional complex, via the use of alternative promoters and termination sites, as well as by alternative splicing. This complexity is critical for the development and function of highly specialized cell types, and in particular alternative promoter use has been found to be an important contributor to cell type-specific gene expression. Schwann cells are the myelinating cells of the peripheral nervous system and dysfunction of these cells is associated with motor and sensory defects as seen in inherited demyelinating neuropathies. SOX10 is a transcription factor that is critical for Schwann cell development and maintenance, and SOX10 target genes include demyelinating disease genes. Therefore, the identification of novel SOX10 target genes is likely to provide insights into additional loci that are important for myelination. Further, SOX10 is known to induce expression at target genes through binding and activation of promoter elements, in some cases directing expression of particular gene products through regulation of an alternative promoter.

In this dissertation, I leverage the known importance of SOX10 in Schwann cells and the previously-described activity of this transcription factor at promoter elements to identify novel SOX10 target genes and gene products that are important for myelination. I have a particular interest in the SOX10-regulated activity of alternative promoters, given the intriguing experimental questions regarding isoform specificity that arise from these regulatory events. In Chapter 2 I discuss our work to characterize the transcriptional regulation of a known demyelinating disease gene, *MTMR2,* which resulted in the identification of a SOX10-regulated

alternative promoter element at this locus. I present data confirming the expression of an alternative transcript from this promoter, which codes for a truncated MTMR2 protein isoform and discuss the implications of these findings for *MTMR2* function in Schwann cells. In Chapter 3 I outline my efforts toward the identification of SOX10-regulated promoter elements genome-wide. These studies include TSS mapping and quantification in Schwann cells by applying the Tn5Prime library preparation method to multiple models: intact adult sciatic nerve, differentiating primary Schwann cells, and immortalized Schwann cells with and without SOX10 expression. I also discuss the insights that these studies provide regarding the characteristics of SOX10-regulated promoter elements. Chapters 4 through 7 describe locus-specific validation and follow-up studies that I completed at four novel SOX10 target genes identified by the genome-wide studies: *ARPC1A*, *CHN2*, *DDR1*, and *GAS7*. In each case I present the TSS mapping data supporting the identification of a SOX10-regulated promoter at the locus, our *in vitro* assays to validate the element, and efforts to confirm the transcript and protein expression profiles for these genes in Schwann cells. Insights into the roles that these loci and their SOX10-regulated gene products may play in Schwann cells are discussed. Finally, Chapter 8 summarizes the relevance and implications of the presented data and suggests areas of interest for further study.

Chapter 2

SOX10 Regulates an Alternative Promoter at the Charcot-Marie-Tooth

Disease Locus *MTMR2*

Introduction

Charcot-Marie-Tooth (CMT) disease is a heterogeneous group of disorders that comprises the most common class of inherited peripheral neuropathies, and affected individuals show varying degrees of muscle weakness and sensory loss in the extremities [178]. CMT disease is broadly categorized into two types; demyelinating CMT arises from a primary defect in Schwann cells, whereas axonal CMT is caused by a primary defect in peripheral neurons [104, 179]. Schwann cells—the myelinating cells of the peripheral nervous system—produce a myelin sheath that mediates saltatory conduction along axons. Consistent with a defect in these cells, patients with CMT1 typically present in clinic with slow nerve conduction velocities [180]. CMT1 often manifests early in life, with onset by 10 years of age in more than 60% of patients[180], and can be debilitating to the point of rendering affected individuals wheelchair-bound. To date, over 20 genes have been implicated in demyelinating CMT disease [181]. These include genes encoding myelin proteins (*e.g.*, *MPZ*) [182] and others with known functions in Schwann cells (*e.g.*, *EGR2*) [183], but many have poorly understood roles in Schwann cell biology. Furthermore, there is no effective treatment for CMT disease, which will likely require gene- and mutation-specific therapies. Thus, characterizing the disease genes will provide a better understanding of Schwann cell physiology and the pathology of demyelinating neuropathies.

Loss-of-function mutations in the myotubularin-related protein 2 (*MTMR2*) gene cause CMT4B1, an autosomal recessive, demyelinating CMT disease distinguished by myelin outfoldings in the peripheral nerves [184]. CMT4B1 is particularly severe, as both proximal and distal limbs are affected and symptoms onset at ~34 months of age [185]. Consistent with the Schwann cell-specific phenotype in CMT4B1 patients, studies have shown an important role for MTMR2 in Schwann cell biology. In mice, conditional loss of *Mtmr2* in Schwann cells causes neuropathy with myelin outfoldings reminiscent of CMT4B1, whereas loss of the gene in motor neurons does not cause axonal or myelination phenotypes [186]. Furthermore, *Mtmr2* expression is developmentally regulated during peripheral myelination [187], and knock-down of *Mtmr2* in cultured Schwann cells decreases proliferation and increases cell death [188]. However, there is evidence that *Mtmr2* is important for neurons *in vivo*, based on genetic interactions between *Mtmr2* and *Fig4* in mouse models of CMT4B1 and CMT4J, respectively [189].

MTMR2 encodes a ubiquitously expressed lipid phosphatase, which converts phosphatidylinositol 3-phosphate (PI(3)P) and PI(3,5)P2 to PI and PI(5)P, respectively [190]. The protein contains a PH-GRAM domain that confers substrate specificity; a PTP catalytic phosphatase domain; a SET-interacting domain (SID); a coiled-coil domain that mediates oligomerization; and a PDZ binding domain for protein-protein interactions [191]. As such, it is thought that MTMR2 plays a role in membrane trafficking and cell signaling [192]. Interestingly, MTMR2 localizes to the nucleus of myelinating and non-myelinating Schwann cells co-cultured with sensory neurons [193]. However, finer details on the cytoplasmic or nuclear function(s) of MTMR2 remain unclear.

Currently, little is known about the transcriptional regulation of *MTMR2*. Deciphering the regulation of this locus will be important for understanding the function of MTMR2 and will be

particularly salient in determining how loss of a ubiquitously expressed gene causes a Schwann cell-specific phenotype. In this chapter, I describe our computational and functional analyses of the *MTMR2* locus, which revealed a SOX10-dependent alternative promoter; SOX10 is a transcription factor that is essential for all stages of the Schwann cell lineage [194]. Activity at this promoter directs the expression of a previously unreported, alternate transcript that encodes an N-terminally truncated protein isoform. Collectively, our findings reveal a SOX10-dependent *MTMR2* gene product and suggest an isoform-specific aspect to MTMR2 biology that may be particularly important for Schwann cells and CMT4B1 pathogenesis.

Please note that I performed the work presented in this chapter with the following exceptions: Dr. Megan Brewer cloned the candidate regulatory elements and deleted the SOX10 binding sequence from *MTMR2* Prom 2, as well as performed 5'RACE to identify *Mtmr2* TSSs in S16 cells. Dr. William Law generated the RNA-Seq data from S16 cells and analyzed the presence of split-reads at the 5' end of the *Mtmr2* locus. Dr. José Rodriguez-Molina in Dr. John Svaren's laboratory at the University of Wisconsin performed the qRT-PCR for *Mtmr2* transcripts in developing sciatic nerve and tested the expression of the transcripts in S16 and primary Schwann cells with *Sox10* siRNA. Dr. Joseph Ma, also in the Svaren laboratory, performed ChIP-Seq experiments for H3K4me3 in rat sciatic nerve. Finally, I wrote the entirety of the manuscript describing these findings, which was published in *Human Molecular Genetics*[195].

Materials and Methods

Comparative sequence analysis at MTMR2

To identify non-coding multiple species conserved sequences (MCSs) at *MTMR2* we scrutinized the human *MTMR* locus on chromosome 11 including genomic sequences up to the

two flanking genes using the UCSC Human Genome Browser [196]: *CEP57* (centromeric) and

MAML2 (telomeric). This resulted in the assessment of a 143,901 base pair genomic region

(chr11:95,565,857-95,709,757; hg19). Next, we visually examined the "Multiz Alignments of

100 Vertebrates" track [196] for non-coding genomic sequences that aligned between human,

mouse, and chicken—three vertebrate species relevant for the study of peripheral nervous system

myelination. These efforts revealed three regions of interest (MSC1, MCS2, and MCS3; Table

2.1 and Figure 2.1A); however, we also included the *MTMR2* proximal promoter (Table 2.1 and

Figure 2.1A) in our functional studies.

Generation of reporter gene and expression constructs

 Oligonucleotide primers containing attB1 and attB2 Gateway cloning sequences

(Invitrogen Life Technologies, Carlsbad, CA) were designed for PCR-based amplification of the

MTMR2 Promoter and each MCS (Table 2.1) and of the two *Mtmr2* open reading frames

(ORFs). Putative regulatory sequences were amplified from human genomic DNA and the

Mtmr2 ORFs were amplified from a cDNA library generated from RNA extracted from rat S16

cells (see below) using Phusion High-Fidelity DNA Polymerase (New England Biolabs Inc.,

Ipswich, MA). Subsequent to PCR amplification and purification, each genomic segment was

cloned into the pDONR221 vector using BP Clonase (Invitrogen). Resulting constructs were

genotyped by digestion with *Bsr*GI and subjected to DNA sequence analysis to ensure the

integrity of the insert. For luciferase assays, each resulting pDONR221 construct was

recombined with an expression construct (pE1B-luciferase) [197] using LR Clonase (Invitrogen)

to clone each region upstream of a minimal promoter directing expression of a luciferase reporter

gene. For localization studies, each resulting pDONR221 construct was recombined with an

expression construct using LR Clonase to clone each ORF in frame with an N-terminal GFP

(pcDNA-pDEST53, Invitrogen) or Flag tag [pEZYflag, Addgene plasmid #18700 [198]]. In each case, successful recombination was confirmed via digestion of DNA with *Bsr*G1 (New England Biolabs Inc.). DNA sequencing analysis was employed to confirm that the ORFs were in-frame with the tag.

Site-directed mutagenesis was performed using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc., Santa Clara, CA). Mutagenesis primers were designed to delete the dimeric SOX10 binding site within *MTMR2*-Prom2 and produce the minor alleles of the two SNPs identified in *MTMR2*-Prom2. Mutagenesis was performed in pDONR221 constructs (see above) and DNA from each resulting clone underwent sequence analysis to verify that only the desired mutation was produced. Verified clones were then recombined into pE1B-luciferase using LR Clonase (Invitrogen).

Cell culture, transfection, and luciferase assays

Immortalized rat Schwann cells (S16) [199], mouse spinal motor neurons (MN1) [200], and HeLa cells were grown under standard conditions in DMEM with 10% fetal bovine serum, 2mM L-glutamine, 50 U/mL penicillin, and 50 g/mL streptomycin. For luciferase assays, ~1x10^4 S16 or MN-1 cells were plated in each well of a 96-well plate. For RT-PCR and Western blot experiments, ~3.5x10^5 S16, MN-1, C2C12, or HeLa cells were plated in each well of a 6-well plate. For localization studies, ~5x10^4 HeLa cells were plated in each chamber of a 4-chamber slide. For all experiments, cells were cultured for 24 hours under standard conditions prior to transfections.

Lipofectamine 2000 (Life Technologies) was diluted 1:100 in OptiMEM I reduced serum medium (Life Technologies) and incubated at room temperature for 10 minutes. Each DNA construct to be transfected was individually diluted in OptiMEM to a concentration of 8 ng/μL

(96-well plates), 4 ng/µL (6-well plates), or 12 ng/µL (4-chamber slides). For luciferase assays, an internal control renilla construct was added to the solution at 8 pg/µL. For co-expression in luciferase assays, 100 ng of a construct to express either wild-type or dominant-negative (E189X) SOX10 [201] was co-transfected in each well. One volume of lipofectamine solution was added to each DNA solution and allowed to sit for 20 minutes at room temperature. Cells were incubated with transfection solution for 4 hours under standard conditions and then the media was changed to standard growth medium.

For luciferase assays, S16 or MN-1 cells were washed with 1X PBS 48 hours after transfection and lysed for 1 hour shaking at room temperature using 20 µL 1X Passive Lysis Buffer (Promega, Madison, WI). 10 µL of lysate from each well was transferred into a white polystyrene 96-well plate (Corning Inc., Corning, NY). Luciferase and renilla activities were determined using the Dual Luciferase Reporter 1000 Assay System (Promega) and a Glomax Multi-Detection System (Promega). Each reaction was performed at least 24 times. The ratio of luciferase to renilla activity and the fold change in this ratio compared to a control luciferase expression vector with no genomic insert were calculated. The mean (bar height) and standard deviation (error bars) of the fold difference are represented in the figures.

Fluorescence Microscopy

For localization studies, cells were cultured under standard conditions for 24 hours after transfection. Cells were then washed with 1X PBS for 5 minutes, fixed with 4% paraformaldehyde in PBS for 10 minutes at room temperature, washed 3 times with PBS, permeabolized with 0.1% Triton X-100 in PBS for 10 minutes at room temperature, and washed 3 times again with PBS. Cells were blocked with 1% BSA in PBS for 2 hours, rocking at room temperature. To enhance GFP signal, GFP-MTMR2-transfected cells were incubated for 1.5

39

hours at 37°C with a monoclonal mouse GFP antibody (Clone 3E6, Invitrogen) diluted 1:200 in 1% BSA. Similarly, Flag-MTMR2-transfected cells were incubated under the same conditions with a monoclonal mouse Flag antibody (Clone M2, Sigma) diluted 1:2000. For colocalization experiments with Calnexin, transfected cells were incubated with a rabbit GFP antibody conjugated to AlexaFluor 488 (Invitrogen) diluted 1:1000 and a monoclonal mouse Calnexin antibody (EMD Millipore, C8.B6) diluted 1:250. After washing, cells were incubated for 1 hour at 37°C with AlexaFluor 488 or 555 goat anti-mouse secondary antibody (Invitrogen) diluted 1:1000 in 1% BSA. All cells were stained with DAPI to visualize nuclei. ProLong Gold anti-fade reagent (Invitrogen) was applied and slides were covered with glass coverslips. For saponin treatment, cells were incubated in 0.02% saponin (Sigma-Aldrich, St. Louis, MO) in 1X PBS for 2 minutes at room temperature, then washed 3 times with PBS prior to fixation. For standard fluorescent microscopy, cells were imaged with an Olympus IX71 inverted microscope using cellSens Standard image software (Olympus, Center Valley, PA). For confocal fluorescent microscopy, cells were imaged with a Leica Upright SP5X Confocal Microscope or a Nikon A-1 Confocal Microscope in the University of Michigan Microscopy & Image Analysis Core Laboratory. Cell counts to quantify puncta formation were done across three independent transfections, and were confirmed by additional counts from an observer naïve to the nature of the experiments.

RNA isolation, cDNA library preparation, and 5'RACE

RNA was isolated from S16, MN1, C2C12, or HeLa cells (immediately or 24 hours after transfection) or from fresh mouse sciatic nerve or whole brain from animals of either sex using the RNeasy Mini Kit (Qiagen, Vinlo, Limburg) according to manufacturer's instructions, eluted into 50 μL of RNase-free water, and stored at -80°C. RNA concentration and purity were

40

determined using a NanoDrop Lite (Thermo Scientific, Waltham, MA). A cDNA library was generated from each sample using 1 μg of RNA and the High-Capacity cDNA Reverse Transcription Kit (Life Technologies), with the provided random reverse-transcription primers. Human oligodendrocyte precursor cDNA was purchased from ScienCell.

For 5'RACE experiments, first-strand cDNA was generated from RNA isolated from S16 cells using an oligonucleotide primer designed in exon 5 of the rat *Mtmr2* locus. The cDNA sample was subsequently TdT-tailed using the 5'RACE System (Invitrogen) and sequential PCR reactions were performed using nested primers in exon 4 and exon 2 of *Mtmr2*. PCR products were size separated by gel electrophoresis, excised, purified using the QIAquick Gel Extraction Kit (Qiagen, Venlo, Netherlands), and subjected to TA cloning. A total of 28 DNA samples extracted from individual clones were subjected to DNA sequence analysis, 19 of which mapped to the rat *Mtmr2* locus.

RT-PCR

cDNA samples were analyzed by PCR using species- and gene-specific primers. For each reaction, 23 μL of PCR Supermix (Life Technologies) was combined with 0.5 μL of each 20 μM primer solution and 1 μL of cDNA. Reactions to amplify the full-length *Mtmr2-2* transcript from S16 cells and mouse sciatic nerve were done with Phusion High-Fidelity Polymerase (NEB). Blank (cDNA-negative) controls were included for each primer pair. Standard PCR conditions were used.

Western blot analysis

Cytosolic and nuclear protein fractions were isolated using the NE-PER kit (Thermo Scientific) according to the manufacturer's instructions. For sciatic nerve lysates, nerves were sonicated in 200 μL of lysis solution containing 25mM Tris-HCl/150mM NaCl/1% NP-40/1%

sodium deoxycholate/0.1% SDS (Thermo Scientific) and protease inhibitor cocktail (Thermo Scientific). Protein concentrations were quantified using the BCA Protein Assay Kit (Pierce). 20 μg samples from cellular fractions or 13.5 μg samples from nerve lysates were combined with SDS loading buffer (Life Technologies) and β-mercaptoethanol (MP Biomedicals, Santa Ana, CA) and incubated at 99°C for 5 minutes. The samples were electrophoresed on 4-20% Tris-Glycine gels (Life Technologies) at 125V for 3 hours in running buffer containing SDS (Life Technologies) at 4°C. Protein samples were transferred onto PVDF membranes at 25 V for 1.5 hours in Tris-Glycine Transfer Buffer (Life Technologies). Membranes were blocked in 2% non-fat milk in TBST overnight at 4°C. Primary antibodies included: rabbit anti-GFP (1:2000, Sigma), mouse anti-JUN (1:500, BD Biosciences, San Jose, CA), rabbit anti-RAB7 (1:1000, Sigma), rabbit anti-MTMR2 [1:250, [202]], mouse anti-MTMR2 (1:5000, Abnova), or rabbit anti-MTMR2 [1:1000, [203]] and were applied in blocking solution for 1 hour at room temperature. Membranes were washed 3 times in TBST, incubated for 1 hour at room temperature with an anti-mouse or anti-rabbit secondary antibody conjugated to horseradish peroxidase (Millipore, Billerica, MA), washed three times in TBST, and incubated with SuperSignal West Dura substrate (Thermo Scientific) according to manufacturer's instructions.

siRNA Transfections and qRT-PCR

Control siRNA (siControl 1, Ambion catalog number AM4611) or Sox10 siRNA (siSox10 1, Ambion catalog number s131239) were transfected into S16 or primary rat Schwann cells [204] 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. The relative amounts of each *Mtmr2* isoform were determined by comparative Ct method and normalized to 18S rRNA. To study the developmental regulation of

Mtmr2 mRNA isoforms, qRT-PCR assays were performed on cDNA samples generated from RNA isolated from P1 and P15 rat sciatic nerves.

Micrococcal nuclease (MNase) aided Chromatin immunoprecipitation (ChIP) in vivo

Snap-frozen sciatic nerves (P30) were ground and incubated in 1 ml of lysis buffer (50 mM Hepes–KOH, pH 7.5; 140 mM NaCl; 4 mM MgCl2; 10% Glycerol; 0.5% Igepal CA-630; 0.25% Triton X-100; Protease Inhibitor Cocktail) [205] rotating for 20 min at 4°C. After 15 strokes with a Dounce homogenizer, samples were centrifuged at 18,000 x rcf for 10 min at 4°C. Pellets were washed with 1 mL of MNase digestion buffer (0.32 M sucrose; 50 mM Tris-HCl, pH 7.5; 4 mM MgCl2; 1 mM CaCl2; Protease Inhibitor Cocktail) [206]. Insoluble material was pelleted by centrifugation, samples were resuspended in 200 μl of MNase digestion buffer, and were incubated with 1 μl (2000 gel units) of MNase (New England Biolabs, M0247) for 7 min at 37°C. Digestion was terminated by addition of EDTA to a final concentration of 0.05 M. Samples were centrifuged at 18,000 x rcf for 20 min at 4°C, and supernatants containing small fragments of chromatin were pooled. The majority of digested chromatin was about 150 base pairs in length (data not shown). ~1.5 μg of chromatin DNA was mixed with 5 μg of anti-H3K4me3 antibody (Millipore 04-745) in ChIP incubation buffer (50 mM NaCl; 50 mM Tris-HCl, pH 7.5; Protease Inhibitor Cocktail; 5 mM EDTA) with a final volume of 1 mL. Samples were allowed to incubate for 12-16 h at 4°C on rotator. 60 μl of Dynabeads Protein G (Invitrogen, 10004D) slurry was washed twice with 0.5% BSA in PBS and then incubated with each ChIP sample for 4 h at 4°C on rotator. ChIP samples were washed with washing buffer 1 (WB1, 50 mM Tris-HCl, pH7.5; 10 mM EDTA; 125 mM NaCl) once, WB2 (50 mM Tris-HCl, pH7.5; 10 mM EDTA; 250 mM NaCl) once, and WB3 (50 mM Tris-HCl, pH7.5; 10 mM EDTA; 500 mM NaCl) twice. The samples were eluted at 65°C with elution buffer (50 mM NaCl; 50

mM Tris-HCl, pH 7.5; 5mM EDTA; 1% SDS) for 15 min. DNA was purified by phenol

chloroform extraction and subjected to sequencing.

Results

The MTMR2 locus harbors four putative transcriptional regulatory elements

Multiple-species comparative sequence analysis is a powerful tool for predicting cis-

acting transcriptional regulatory elements [207]. To identify evolutionarily conserved sequences

at *MTMR2*, we scrutinized aligned genomic sequences spanning *MTMR2* and extending to the

flanking loci (*MAML2* and *CEP57*) from human, mouse, and chicken using the UCSC Human

Genome Browser (http://genome.ucsc.edu/) [196]. Specifically, we identified non-coding, non-

repetitive genomic sequences that are conserved among these three species. This analysis

revealed three <u>m</u>ultiple-species <u>c</u>onserved <u>s</u>equences (MCSs)—two upstream of the annotated

transcription start site (TSS) of *MTMR2* and one within the first intron of the major *MTMR2*

transcript (Table 2.1 and Figure 2.1A). We also included the putative proximal promoter

('Prom1' in Table 2.1 and Figure 2.1A) and considered these four genomic segments to be

candidate *MTMR2* regulatory elements.

MTMR2-MCS3 displays strong enhancer activity in cultured Schwann cells

To determine if the four candidate regulatory elements described above (Table 2.1 and

Figure 2.1A) are important for *MTMR2* expression in Schwann cells, we tested each for

regulatory potential in S16 cells—a rat immortalized Schwann cell line that expresses myelin-

related genes (*e.g.*, *PMP22*, *MPZ*, and *SOX10*) [199, 208, 209]. Briefly, we cloned each genomic

segment upstream of a minimal promoter directing the expression of a firefly luciferase reporter

gene [197]. Next, we transfected the resulting constructs into cultured S16 cells along with a

44
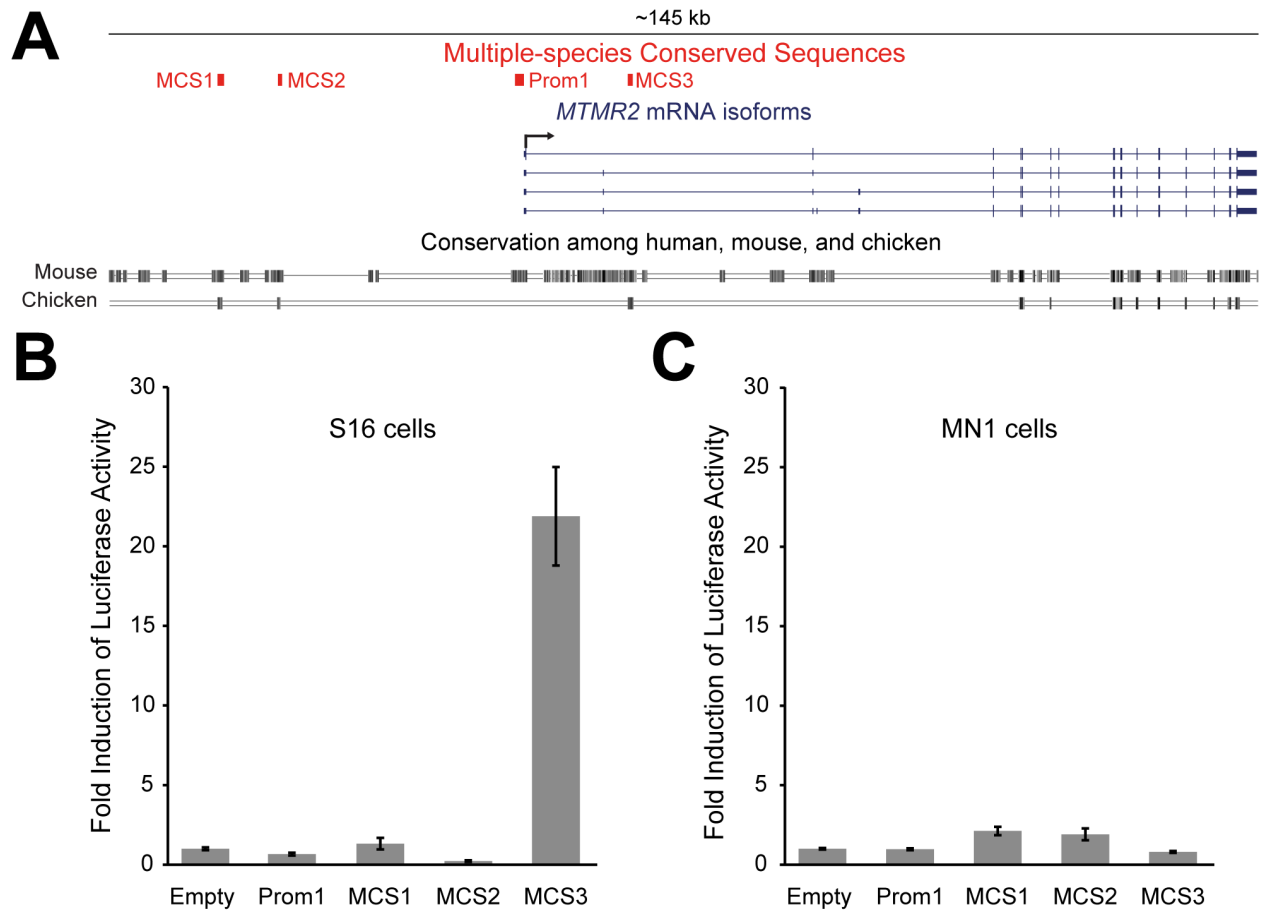
Figure 2.1 *MTMR2*-MCS3 is an active regulatory element in immortalized rat Schwann cells. (A) The ~145 kb *MTMR2* locus was analyzed for non-coding regions conserved between human, mouse, and chicken genomes (indicated along the bottom in black). The locations of the three <u>m</u>ultiple-species <u>c</u>onserved <u>s</u>equences (MCS1-3) and the proximal promoter (Prom1) are indicated in red along with the four described *MTMR2* mRNA isoforms (indicated in blue, see text for details), which are transcribed from left to right (black arrow). (B and C) The four elements identified in (A) were cloned upstream of a luciferase reporter gene and tested for induction of luciferase activity in S16 cells (B) and MN-1 cells (C) relative to an 'empty' control vector with no genomic insert. Error bars indicate standard deviation.

Element ID	Genomic Coordinates (hg19)	Size (bp)
MTMR2-Prom1	chr11:95,657,371-95,658,478	1,108
MTMR2-MCS1	chr11:95,694,768-95,695,577	810
MTMR2-MCS2	chr11:95,687,566-95,688,075	510
MTMR2-MCS3 (Prom2)	chr11:95,643,795-95,644,453	659

Table 2.1 Putative transcriptional regulatory elements at MTMR2. Coordinates are from the February 2009 UCSC Human Genome Browser (hg19).

renilla luciferase expression construct to control for variable cell viability and transfection efficiency. These experiments revealed that *MTMR2*-MCS3 has strong enhancer activity as indicated by a ~22-fold increase in luciferase activity compared to a control vector with no genomic insert (Figure 2.1B). In contrast, the promoter (Prom1), MCS1, and MCS2 did not dramatically increase luciferase activity above the control expression vector. To determine the specificity of these findings to Schwann cells, we repeated the above experiments in a non-glial cell line (the mouse motor neuron-derived cell line MN-1) [200] that does not express transcription factors important for myelination (*e.g.*, SOX10) [209]. While *MTMR2*-MCS1 and *MTMR2*-MCS2 showed some activity, none of the elements directed luciferase activity above a 3-fold increase and *MTMR2*-MCS3 displayed less activity than the empty vector (Figure 2.1C). In sum, these data are consistent with *MTMR2*-MCS3 being important for the transcription of this locus in Schwann cells.

MTMR2-MCS3 is a previously unreported, alternative MTMR2 promoter

Upon scrutinizing *MTMR2*-MCS3 on the UCSC Rat Genome Browser we noted three spliced ESTs (two from brain, one from E11-12 embryos) with 5' ends that overlap the last 75 base pairs of MCS3 (Figure 2.2A). These findings suggested that *MTMR2*-MCS3 may act as a promoter to express alternative *MTMR2* transcripts. To test this, we performed 5'-rapid amplification of cDNA ends (5'-RACE). Briefly, *Mtmr2* cDNA was generated using RNA isolated from cultured rat Schwann (S16) cells and a reverse primer in exon 5. Subsequently, 5'-RACE was performed using reverse primers in exons 4 and 2 of rat *Mtmr2*. The resulting PCR products were cloned and sequenced, and a total of 19 individual clones mapped to the rat *Mtmr2* locus. These studies revealed the presence of two *Mtmr2*-related transcription start sites (Fig. 2.2A). One of these represents the known exon 1 of *Mtmr2* (7 of the 19 sequences map to this

47

first exon; Figure 2.2A) and the other matches the ESTs that map directly adjacent to, and downstream of, MCS3 (12 of the 19 sequences map to this alternative first exon; Figure 2.2A). Importantly, this region does not map to any annotated exons for *MTMR2*, including the previously reported, alternatively spliced exons at the 5' end of the locus [184, 210]. We also scrutinized RNA-seq data generated using five unique passages of S16 cells[211]. This revealed abundant reads that map to either the annotated exon 1 or the newly identified exon just downstream of MCS3. An average of 124 reads per sample map to exon 1 (S.D. ± 21) including an average of 32 split reads to exon 2 (S.D. ± 9). An average of 60 reads per sample map to the region downstream of MCS3 (S.D. ± 33) including an average of 11 split reads to exon 2 (S.D. ± 7). Moreover, no split reads mapped to the 5' end of the newly identified exon, indicating that this is a transcription start site, and not an alternatively spliced exon; these findings are consistent with the above 5'-RACE data. Combined, our findings indicate that *MTMR2*-MCS3 represents a previously unreported alternative promoter that directs expression of a newly identified first exon at *MTMR2*. Therefore, we renamed this regulatory element *MTMR2*-Prom2 and the adjacent transcribed region *MTMR2* exon 1B.

The MTMR2-2 mRNA is expressed in Schwann cells

To define the expression of *MTMR2* transcripts with respect to the associated CMT phenotype, we performed first-exon-specific RT-PCR on cDNA samples from cells relevant to the peripheral nervous system and CMT disease: mouse sciatic nerve (mSN), MN-1 cells, and immortalized mouse muscle cells (C2C12) [212]. Importantly, the majority of mRNA from sciatic nerve is from Schwann cells, which provides information about Schwann cell gene expression *in vivo*. While all of the tested samples express exon 1A-containing *Mtmr2* transcripts, only sciatic nerve expresses *Mtmr2* exon 1B-containing transcripts (Figure 2.2B);
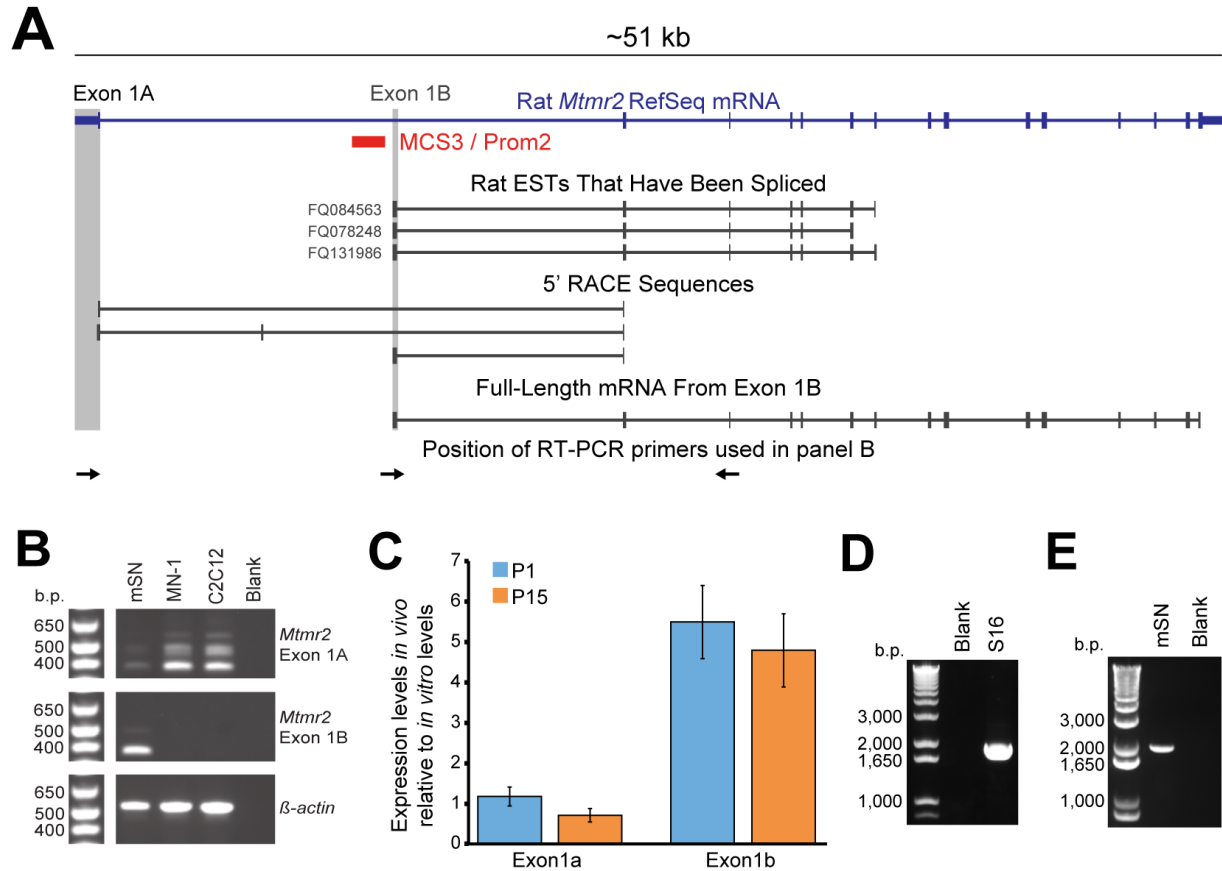
Figure 2.2 *MTMR2*-MCS3 is an alternative promoter that is active in Schwann cells. (A) The ~51 kb rat *Mtmr2* transcriptional unit (indicated in blue) from the UCSC Genome Browser is shown along with MCS3 (red), rat expressed sequence tags (ESTs), three distinct *Mtmr2* mRNA 5' ends identified by 5' RACE, and the full-length *Mtmr2* mRNA generated from exon 1B. The grey shading indicates the position of exons 1A and 1B. (B) RT-PCR assay to test for the presence of exon 1A- and exon 1B-containing transcripts in mouse sciatic nerve (mSN), immortalized motor neurons (MN-1), and immortalized muscle cells (C2C12). Negative controls without cDNA (Blank) were included for each primer pair and primers for β-actin were used as a positive control. Base pair (b.p.) sizes of markers are provided on the left. (C) Quantitative RT-PCR was performed to determine the expression levels of exon 1A- and 1B-containing *Mtmr2* transcripts in P1 and P15 rat sciatic nerve relative to the levels in the S16 Schwann cell line, after normalization to 18S rRNA. Average and standard deviations are shown for 3 biological replicates. (D) A full-length mRNA was amplified from Schwann (S16) cell cDNA using long-range RT-PCR and primers designed in exon 1B and in the last coding exon of *Mtmr2*. A negative control without cDNA was included (Blank). (E) A full-length mRNA was amplified from mouse sciatic nerve cDNA using long-range RT-PCR and primers designed in exon 1B and in the last coding exon of *Mtmr2*. A negative control without cDNA was included (Blank).

note that the original detection of transcripts harboring exon 1B was performed on cDNA from immortalized rat Schwann (S16) cells (see above and Figure 2.2A). In addition, we routinely detected multiple bands in *Mtmr2* RT-PCR products, reflecting mRNA isoforms produced by the alternative inclusion of multiple 'second exons'; this was confirmed by sequencing each PCR product (data not shown). These findings raised the question of whether the two *Mtmr2* isoforms, both expressed in Schwann cells *in vivo*, are differentially regulated during myelination. To address this, we performed quantitative RT-PCR experiments on sciatic nerve samples from P1 and P15 rats with first-exon specific primers. These efforts did not reveal any difference between the expression of exon 1A-containing transcripts and exon 1B-containing transcripts at either time point (Figure 2.2C).

To ensure that exon 1B-containing transcripts represent a full-length *MTMR2* mRNA, we performed long-range RT-PCR on cDNA isolated from S16 cells with a forward primer designed in *Mtmr2* exon 1B and a reverse primer designed in the last coding exon of *Mtmr2*. This revealed a single PCR product of the expected size of 1.8 kilobases (Figure 2.2D). We then cloned the above PCR product and subjected it to DNA sequence analysis, which revealed the presence of all protein-coding exons downstream of exon 1B (Figure 2.2A). To investigate whether this full transcript is also expressed in Schwann cells *in vivo*, we repeated the long-range RT-PCR with cDNA isolated from mouse sciatic nerve and primers specific to mouse *Mtmr2* exon 1B and the last coding exon. Similarly, this generated a single PCR product (Figure 2.2E) that was subject to DNA sequence analysis, again revealing the presence of all protein-coding exons. These data confirmed that the *MTMR2*-2 mRNA is expressed in Schwann cells *in vivo* and encodes a shorter, N-terminally truncated protein isoform (see below and Figure 2.3).
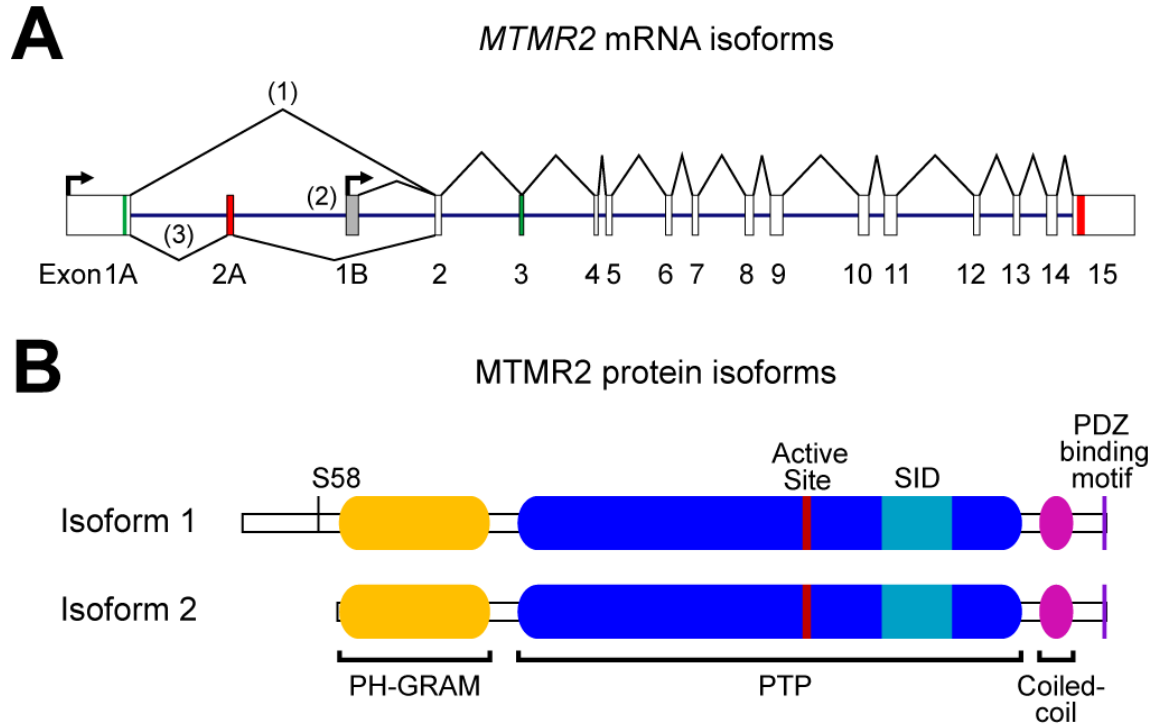
Figure 2.3 A new model of transcriptional activity at the *MTMR2* locus. (A) Schematic of the *MTMR2* locus including transcription start sites (arrows), translation start codons (green), and stop codons (red). Exon numbers are provided along the bottom. Multiple mRNA isoforms (*MTMR2-1*, *MTMR2-2*, and *MTMR2-3*) are produced via alternative splicing and alternative promoter use. (B) MTMR2 protein isoforms are shown indicating the pleckstrin homology/glucosyltransferases, Rab-like GTPase activators, and myotubularins (PH-GRAM; yellow), protein tyrosine phosphatase (PTP; blue), active site (red), SET interaction (SID; aqua), coiled coil (magenta), and PSD-95, Discs-large, ZO-1 binding (PDZ; purple) domains. Please note that isoform 2 lacks 72 amino acids at the N terminus, including the phosphorylation site Serine 58 (S58).

MTMR2 transcripts initiated from exon 1B (Figure 2.3A) are predicted to encode a shorter protein isoform (MTMR2-2), which lacks the first 72 amino acids at the N-terminus (Figure 2.3B). This protein isoform has been noted in the literature previously [184], and is thought to originate from exon 1A-containing transcripts when the alternatively spliced exon 2A is included. Here, full-length MTMR2 is not produced due to an in-frame stop codon in exon 2A, and it has been proposed that leaky ribosomal scanning mediates translation beginning in exon 3 [210]. Importantly, this truncated isoform retains all of the known functional domains of the protein, but lacks a phosphorylation site (serine 58; S58 in Figure 2.3B) proposed to regulate the localization of MTMR2-1 to endosomes [213]. Considering the enrichment of exon 1B-containing transcripts in sciatic nerve compared to motor neurons (Figure 2.2B), it is possible that there is also an enrichment of MTMR2-2 proteins in Schwann cells relative to motor neurons. Unfortunately, due to our inability to acquire a specific antibody against endogenous MTMR2, we were unable to address this (Figure 2.4A-C). However, our data show that the *MTMR2*-Prom2 is active in Schwann cells and directs the expression of an alternative *MTMR2* mRNA (Figure 2.3A) that is predicted to encode the MTMR2-2 protein isoform (Figure 2.3B).

SOX10 regulates MTMR2-Prom2 activity in Schwann cells

SOX10 is an important regulator of gene expression in Schwann cells, and many SOX10 target genes are known to be important for Schwann cell function. To investigate if SOX10 plays a role in the activity of *MTMR2*-Prom2 in Schwann cells, we analyzed this region for evidence of SOX10 binding *in vivo* using a previously published data set. ChIP-seq analyses from rat sciatic nerve, which is enriched for Schwann cell nuclei, contained evidence of SOX10 binding at this promoter (Figure 2.5A), indicating that SOX10 binds to this element *in vivo* [214]. Furthermore, this region is recognized by an antibody against H3K27Ac [214], which marks active regulatory
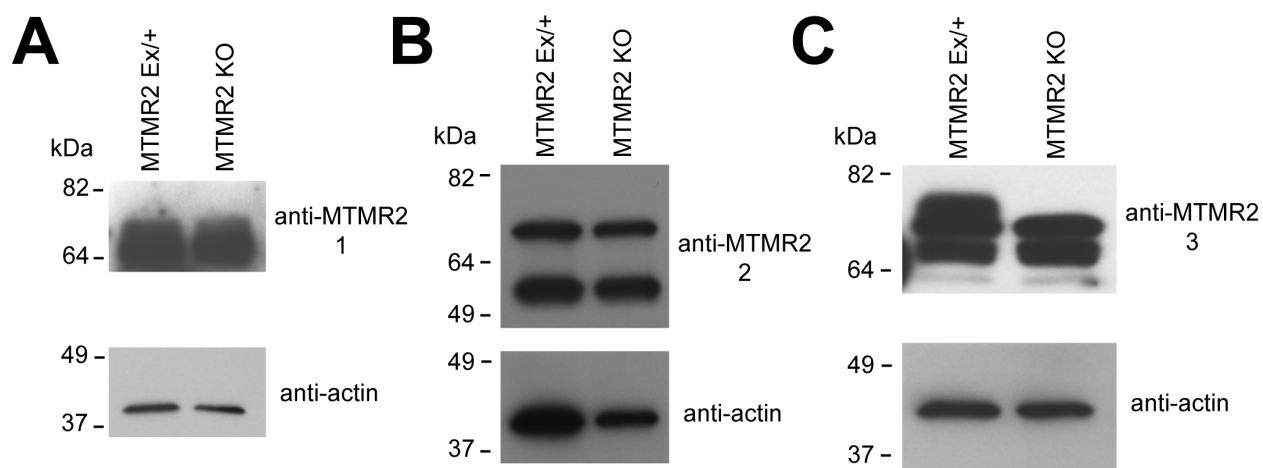
Figure 2.4 Available antibodies raised against endogenous MTMR2 are unable to specifically detect the protein or distinguish between protein isoforms. (A-C) 13.5 μg of protein lysates from sciatic nerves of MTMR2 heterozygous (Ex/+) or knockout (KO) mice were used to test the specificity of three different antibodies raised against MTMR2. Anti-actin was used as a protein loading control. Numbered dashes to the left of each blot indicate the position of protein size markers in kilodaltons (kDa). The predicted sizes of the MTMR2 protein isoforms are 73 and 66 kDa. (A) Antibody 1 originally reported by Li et al. (B) Antibody 2 obtained commercially from Abnova. (C) Antibody 3 originally reported by Ng et al. Please note that in this panel, the upper-most band appears to be specifically lost in the knockout lysate. This band was previously interpreted as MTMR2 isoform 1. However, due to the nonspecific bands below, it is impossible to establish specificity for MTMR2 isoform 2. Therefore, this antibody cannot be used to draw conclusions about isoform expression or localization.

elements, and we found that this region is recognized by an antibody against H3K4me3, indicative of an active promoter element (Figure 2.5A). These findings led us to analyze this region for binding by EGR2, which is a transcription factor that works synergistically with SOX10 at adjacent binding sites to regulate gene expression during myelination [215, 216]. This did not reveal any evidence of EGR2 binding within the alternative promoter or anywhere else at *MTMR2*, indicating that *MTMR2* is only regulated by SOX10 (data not shown). Interestingly, these findings may explain our data showing that *MTMR2* is not developmentally regulated during myelination. In summary, our data support the conclusions that *MTMR2*-Prom2 is an active promoter and bound by SOX10 in rat sciatic nerve.

SOX10 binds to a well-characterized consensus sequence [217] as a monomer or as a dimer when two consensus sequences are arranged in a head-to-head fashion [218]. We therefore examined the *MTMR2*-Prom2 element for conserved SOX10 consensus sequences, which revealed a dimeric SOX10 consensus site that is conserved between human and chicken (Figure 2.5B; note that 'ACAAT' is also a high confidence SOX10 consensus sequence). To determine if the dimeric consensus sequence is important for the regulatory activity of *MTMR2*-Prom2 in Schwann cells, we generated a luciferase reporter construct containing *MTMR2*-Prom2 with the dimeric SOX10 consensus sites and the intervening sequences deleted (ΔSOX10). When this construct was transfected into S16 cells, the regulatory activity was reduced by ~90% compared to wild-type *MTMR2*-Prom2 (Figure 2.5C), indicating that the dimeric SOX10 consensus sequence is critical for the full activity of the alternative promoter.

To determine the sufficiency of SOX10 in activating *MTMR2*-Prom2, wild-type and ΔSOX10 *MTMR2*-Prom2 constructs were transfected into MN-1 cells, which lack endogenous SOX10 [209]. In agreement with earlier results (Figure 2.1C), both reporter constructs show very
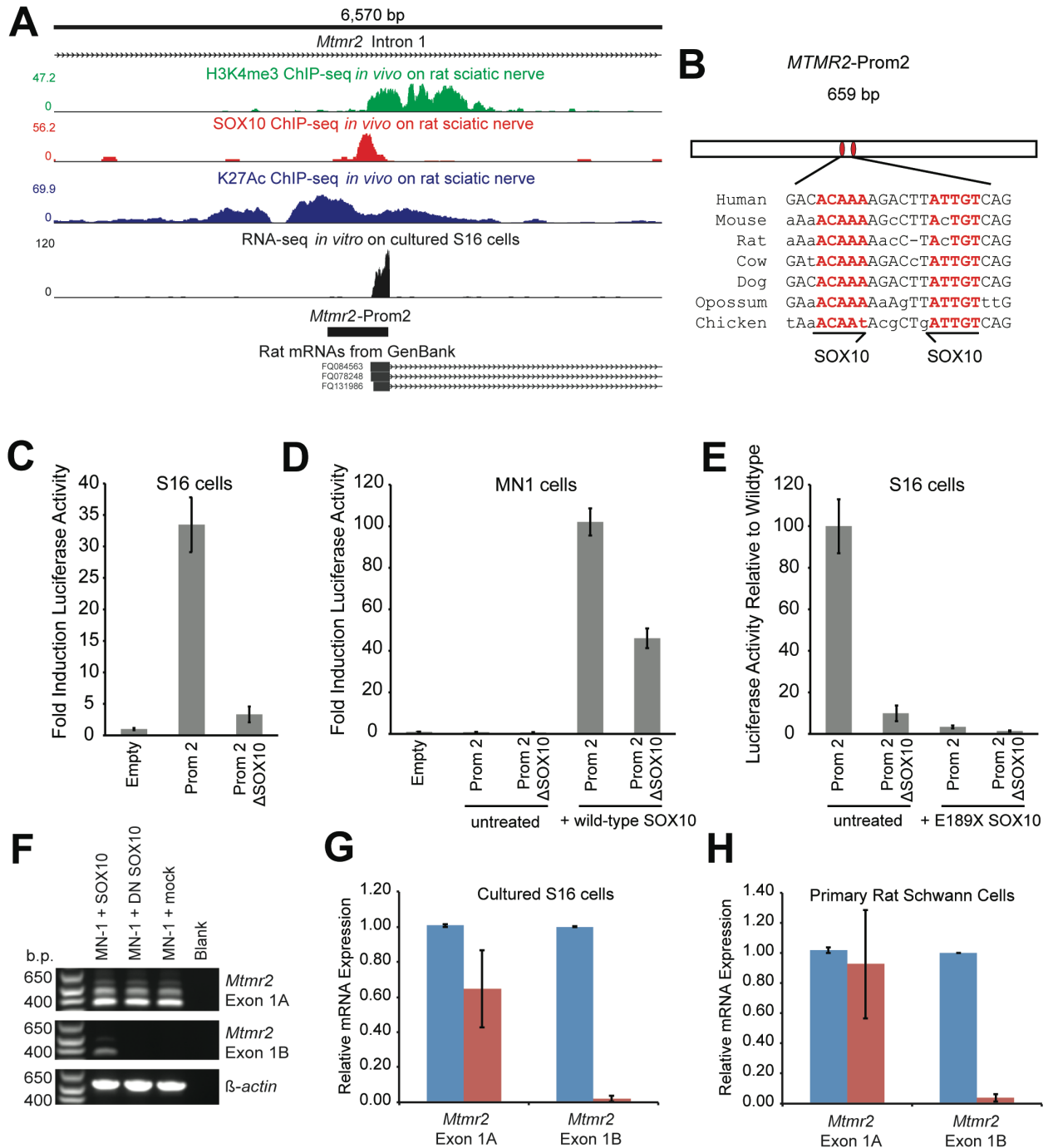
Figure 2.5 SOX10 regulates the activity of *MTMR2*-Prom2 and the expression of the *MTMR2*-2 mRNA. (A) A ~6.6 kb region of *MTMR2* intron 1 is shown along with ChIP-Seq data for histone 3 lysine 4 trimethylation (H3K4me3; green), SOX10 (red), and histone 3 lysine 27 acetylation (H3K27Ac; blue) all performed on rat sciatic nerve. RNA-Sequencing reads, showing per-base read depth across 2 samples of S16 cells, *MTMR2*-Prom2, and spliced rat *MTMR2* ESTs are also indicated. (B) The 659 base pair *MTMR2*-Prom2 is shown along with the position of the two SOX10 monomeric consensus sequences (red ovals and red text). The seven species utilized for comparative sequence analysis are shown on the left, with lower-case letters

55

indicating bases not conserved among all seven species. Half arrows indicate the position and orientation of the two monomeric SOX10 consensus sequences. (C) *MTMR2*-Prom2 with or without (ΔSOX10) the dimeric SOX10 sequence was cloned upstream of a luciferase reporter gene, transfected into cultured Schwann (S16) cells, and tested for activity in luciferase assays compared to an empty vector containing no genomic insert. The fold induction of luciferase activity is indicated along the y axis and error bars indicate standard deviations. (D) Wild-type and ΔSOX10 *MTMR2*-Prom2 were evaluated for regulatory activity in MN1 cells using luciferase assays as in (C) in the presence and absence of a construct to express SOX10. (E) Wild-type and ΔSOX10 *MTMR2*-Prom2 were evaluated for regulatory activity in S16 cells using luciferase assays as in (C) in the presence and absence of a construct to express dominant-negative (E189X) SOX10. (F) RT-PCR assays were performed as in Fig. 2B using cDNA prepared from MN-1 cells transfected with a construct to express either wild-type or dominant-negative SOX10, or mock transfected without a SOX10 expression construct. Note that transfection of the wild-type SOX10 expression construct was sufficient to allow detection of endogenous *Mtmr2* exon 1B-harboring transcripts. (G and H) S16 cells and rat Schwann cells were treated with siRNA targeted against SOX10 (red) or control siRNA (blue). Quantitative RT-PCR was used to measure expression levels of *Mtmr2* exon 1A- and exon 1B-containing transcripts. Error bars indicate standard deviations.

little regulatory activity in MN-1 cells (Figure 2.5D). However, upon overexpression of exogenous SOX10 in these cells, *MTMR2*-Prom2 directs luciferase expression ~100-fold higher than *MTMR2*-Prom2 in the absence of SOX10 (Figure 2.5D). Furthermore, deletion of the dimeric SOX10 binding site reduces the responsiveness to SOX10 by ~60% (Figure 2.5D). Notably, the *MTMR2*-Prom2 element contains other conserved and non-conserved SOX10 consensus sequences (Figure 2.6A), which may explain the remaining 40% of activity associated with ΔSOX10 *MTMR2*-Prom2 upon overexpression of SOX10 in MN-1 cells. We also found that *MTMR2*-Prom2 can be stimulated by expression of SOX8 and SOX9, which bind to similar sequences as SOX10 [219]; however, SOX10 has the largest effect on regulatory activity (Figure 2.6B). Upon examining *MTMR2*-Prom2 at the UCSC Human Genome Browser, we identified two annotated SNPs within *MTMR2*-Prom2; however, neither reside within the dimeric SOX10 binding site and neither affect the regulatory activity of *MTMR2*-Prom2 (data not shown). Combined, these data indicate that SOX10 is sufficient to activate *MTMR2*-Prom2 in MN-1 cells.

To investigate the necessity of SOX10 for *MTMR2*-Prom2 activity in Schwann cells, wild-type and ΔSOX10 Prom2 were each transfected into S16 cells along with a construct to express dominant-negative (E189X) SOX10, which interferes with the function of endogenous SOX10 [201]. The levels of luciferase induction from each construct alone were consistent with our previous experiments (Figure 2.5C and E). In contrast, expression of the dominant-negative SOX10 protein nearly ablates the regulatory activity of wild-type *MTMR2*-Prom2 (Figure 2.5E). These data indicate that SOX10 is necessary for the activity of this element in Schwann cells.

The data presented so far raise the possibility that *MTMR2*-Prom2 activity and exon 1B-containing transcripts are specific to myelinating cells. Oligodendrocytes—the myelinating cells
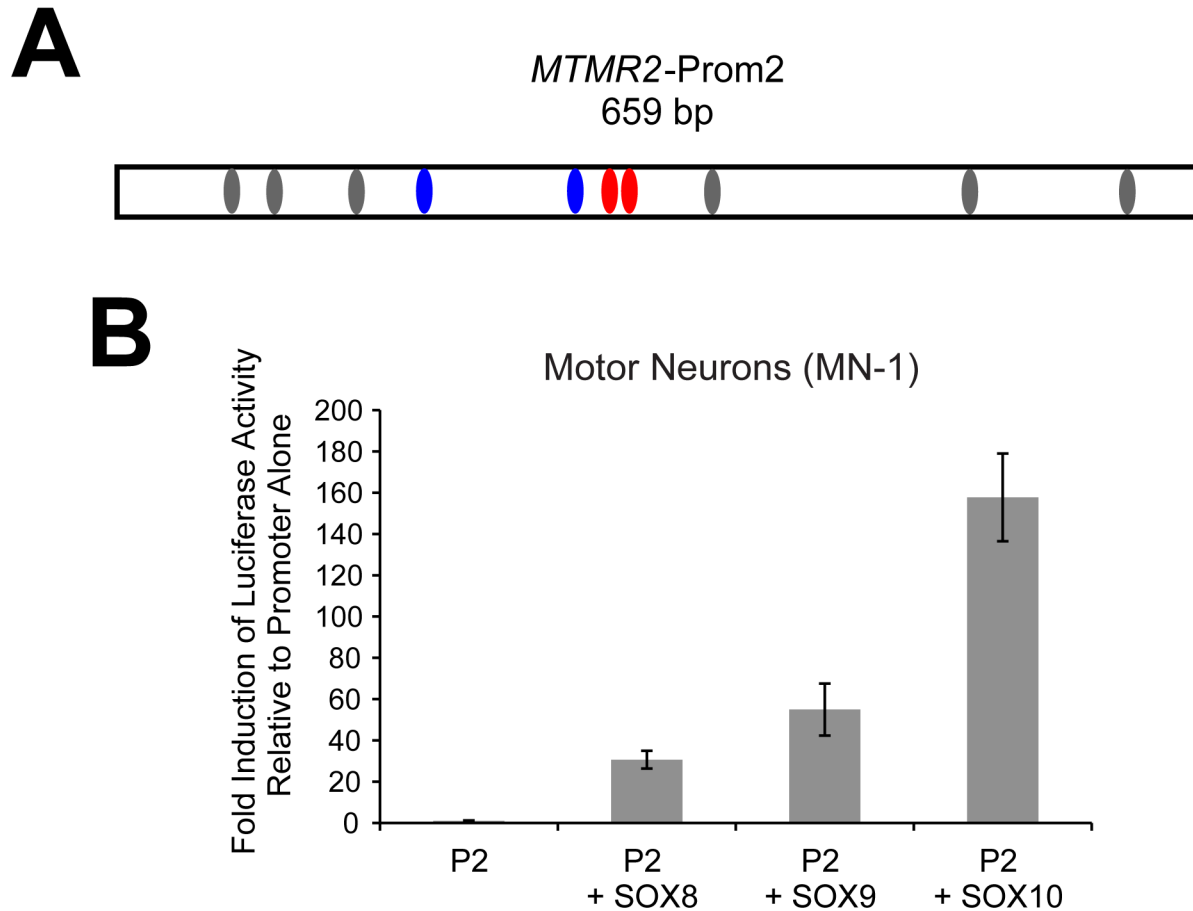
Figure 2.6 *MTMR2*-Prom2 harbors multiple predicted SOX10 binding sites and shows differential responses to members of the SOXE group of transcription factors. (A) Similarly to Figure 2.5B, the 659 base pair *MTMR2*-Prom2 is shown along with the position of predicted SOX10 monomeric consensus sequences. Red ovals indicate the dimeric, conserved SOX10 binding site as before. Here, blue ovals indicate other conserved SOX10 monomeric consensus sequences, while grey ovals indicate non-conserved SOX10 monomeric consensus sequences. (B) Similarly to Figure 2.5D, wild-type *MTMR2*-Prom2 was evaluated for regulatory activity in MN1 cells using luciferase assays in the presence and absence of constructs to express SOX8, SOX9, or SOX10. The induction of luciferase activity, relative to the promoter alone, is indicated along the y axis and error bars indicate standard deviations.

of the central nervous system (CNS)—also express SOX10 and, indeed, SOX10 expression in the CNS is mainly restricted to oligodendrocytes and oligodendrocyte precursor cells [220]. Consistent with this, SOX10 binds *Mtmr2*-Prom2 in the spinal cord [Figure 2.7A; [214]] and exon 1B-containing transcripts are expressed in the CNS (Figure 2.7B). These results suggest that the alternative MTMR2 protein isoform may also be important for myelination in the CNS.

SOX10 is required for expression of endogenous MTMR2 transcripts harboring exon 1B

Based on our findings that exon 1B-containing transcripts are not expressed in MN-1 cells (Figure 2.2B) and that overexpression of SOX10 induces the activity of *MTMR2*-Prom2 in MN-1 cells (Figure 2.3D), we tested if SOX10 is sufficient to induce the expression of endogenous exon 1B-containing transcripts in these same cells. MN-1 cells were transfected with constructs to express SOX10 or dominant-negative (E189X) SOX10, or were mock transfected in the absence of a SOX10 expression construct. Remarkably, exon-specific RT-PCR assays showed that the expression of exogenous SOX10 is sufficient to induce endogenous *Mtmr2-2* mRNA expression in cultured MN-1 cells (Figure 2.5F; the specificity of PCR products was confirmed by DNA sequence analysis, not shown).

To assess the requirement of SOX10 for the expression of each *MTMR2* transcript (*MTMR2*-1 vs. *MTMR2*-2), we performed siRNA knockdown of SOX10 in S16 cells and in cultured primary rat Schwann cells. The expression levels of endogenous rat *Mtmr2* exon 1A- and exon 1B-containing transcripts were measured using quantitative RT-PCR and primers specific to each first exon. In S16 cells, exon 1A-containing transcripts were ~40% reduced with SOX10 knockdown compared to control siRNA. However, exon 1B- containing transcripts were almost completely eliminated with SOX10 knockdown (Figure 2.5G). Similarly, in cultured rat Schwann cells, exon 1A-containing transcripts were not significantly affected by SOX10
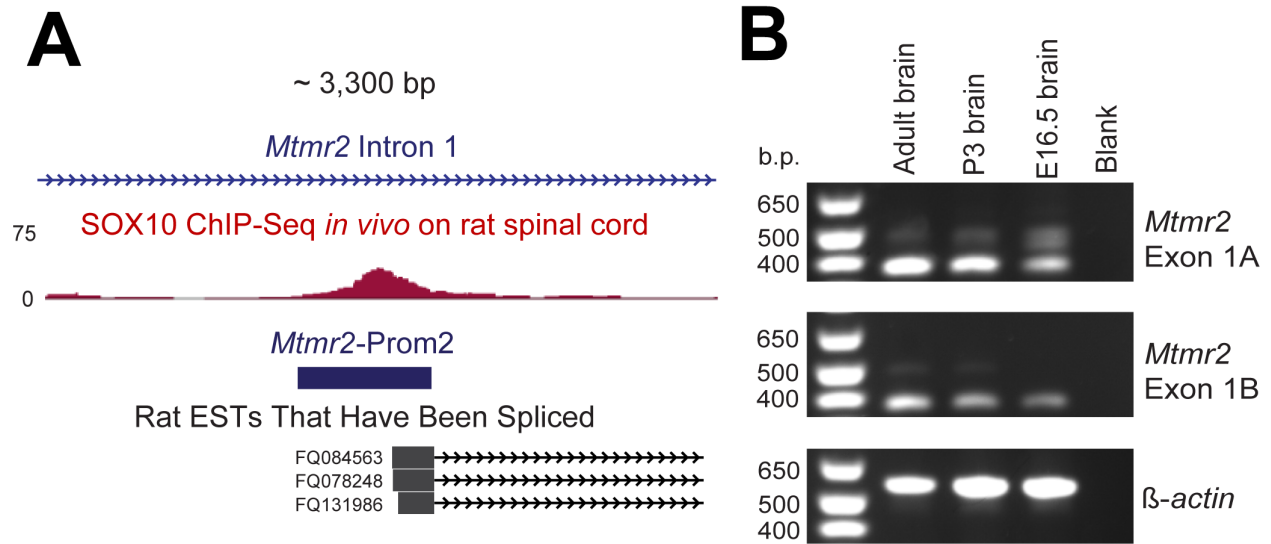
Figure 2.7 *Mtmr2*-Prom2 is bound by SOX10 and exon 1B-containing transcripts are expressed in the central nervous system. (A) As in Figure 2.5A, a ~3.3 kb region of *MTMR2* intron 1 is shown along with ChIP-Seq data for SOX10 (dark red) performed on rat spinal cord. *MTMR2*-Prom2 and spliced rat *MTMR2* ESTs are also indicated. (B) Similarly to Figure 2.2B, an RT-PCR assay was used to test for the presence of exon 1A- and exon 1B-containing transcripts in mouse brain at adult, postnatal day 3 (P3), and embryonic day 16.5 (E16.5) stages. Negative controls without cDNA (Blank) were included for each primer pair and primers for *β-actin* were used as a positive control. Base pair (b.p.) sizes of markers are provided on the left.

knockdown, while exon 1B-containing transcripts show near-complete abolishment (Figure 2.5H). Taken together, our data indicate that SOX10 is necessary and sufficient for the expression of *Mtmr2-2*.

MTMR2 protein isoforms differentially localize to the nucleus

Our data suggest an important role of the shorter MTMR2 protein isoform (MTMR2-2) in Schwann cells. To investigate the functional differences between the two MTMR2 protein isoforms, we transiently transfected HeLa cells with constructs to express each isoform with an N-terminal GFP tag and performed fluorescence microscopy. We found that both proteins localize diffusely in the cytoplasm as well as to discrete puncta, and that these puncta are more clearly visualized upon treating the cells with saponin to deplete cytosolic proteins (Figure 2.8A). Interestingly, cells expressing GFP-MTMR2-2 exhibited puncta more frequently than those expressing MTMR2-1 (Figure 2.8B). Approximately 20% of GFP-positive cells expressing GFP-MTMR2-1 have at least five distinct cellular puncta. In contrast, ~75% of cells expressing GFP-MTMR2-2 met this criteria (GFP-MTMR2-1 n=542, GFP-MTMR2-2 n=554, test for two-sample proportions, z=18.709, p < 0.0001).

Interestingly, our fluorescence microscopy analyses revealed that the MTMR2 puncta frequently overlap with nuclear DAPI staining (Figure 2.8A). To assess for nuclear localization of MTMR2 protein isoforms, we isolated cytosolic and nuclear protein fractions from cells transiently transfected with constructs to express GFP-MTMR2-1 or GFP-MTMR2-2. We found that both MTMR2 isoforms localize to the cytosolic and nuclear fractions, but that a larger fraction of GFP-MTMR2-2 is present in the nucleus compared to GFP-MTMR2-1 (Figure 2.8C). To confirm that the MTMR2 puncta are nuclear, we performed confocal fluorescent microscopy on HeLa cells expressing either GFP-MTMR2-1 or GFP-MTMR2-2. This revealed that both
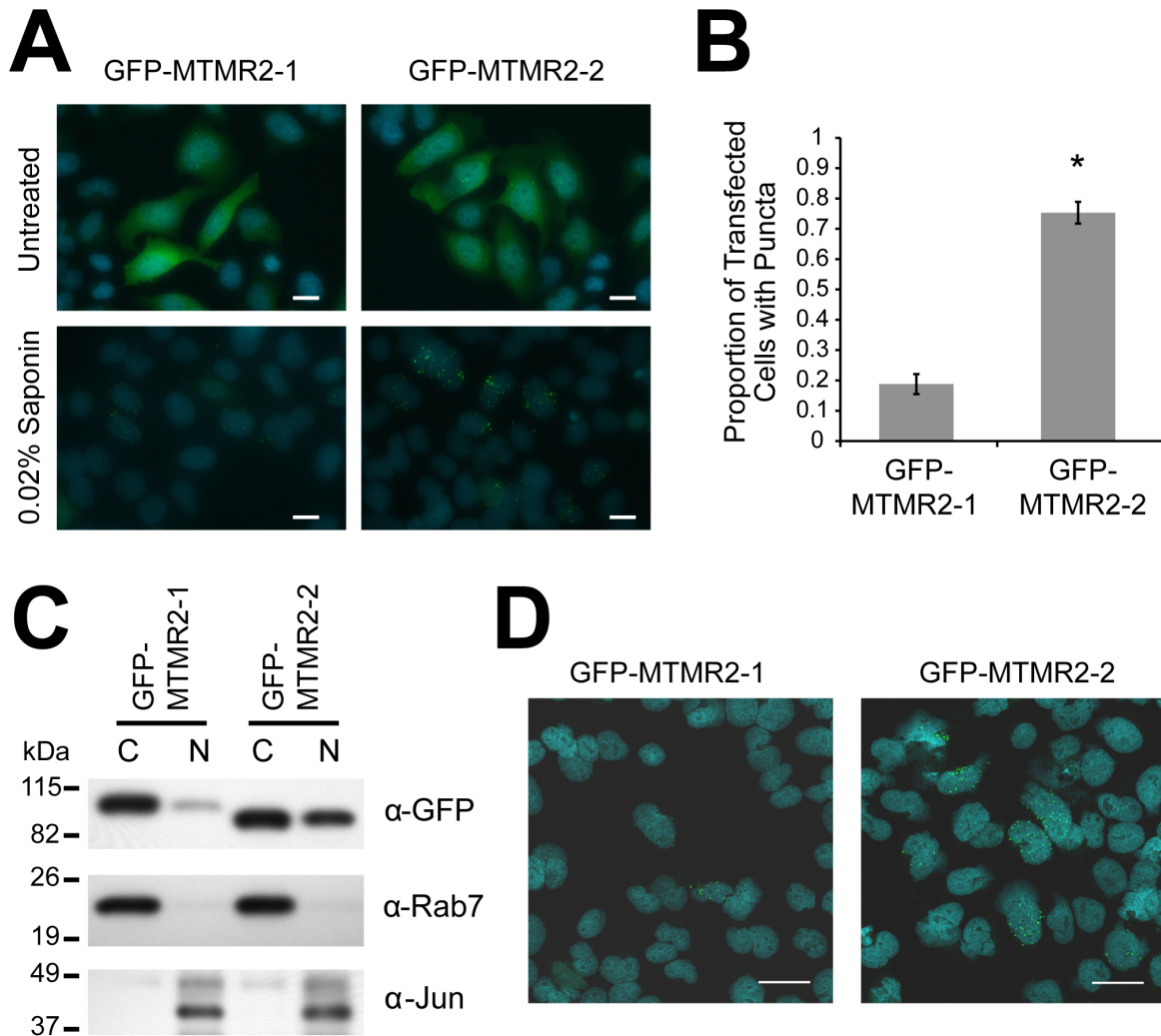
Figure 2.8 MTMR2 protein isoforms differentially localize to subcellular puncta and to the cell nucleus. (A) HeLa cells were transfected with constructs to express GFP-MTMR2 isoforms and imaged with standard fluorescence microscopy. Cells were untreated or treated with 0.02% saponin to remove cytosolic proteins and stained with DAPI to visualize nuclei prior to fixation. Scale bars, 20 μm. (B) Quantification of puncta formed by the two MTMR2 isoforms. Statistical analysis: significance test for two-sample proportions (MTMR2-1 n=542, MTMR2-2 n=554, p < 0.0001). Error bars represent margin of error for 95% confidence interval. (C) Western blot using an anti-GFP antibody and 20 μg of cytosolic (C) or nuclear protein (N) fractions isolated from HeLa cells transfected with constructs to express either GFP-MTMR2 isoform 1 or isoform 2. Rab7 and cJun antibodies were used on the same blot to assess protein loading and fraction purity. A protein marker in kilodaltons (kDa) is indicated on the left. (D) HeLa cells were transfected with constructs to express either GFP-MTMR2 isoform, treated with 0.02% saponin to clear cytosolic proteins and DAPI to visualize nuclei (blue), and imaged with a confocal fluorescence microscope. Each image was taken from a single 0.3 μm-depth slice in the z-stack. Scale bars, 30 μm.

MTMR2 isoforms localize to nuclear puncta (Figure 2.8D), again with a larger percentage

identified in cells expressing GFP-MTMR2-2. The perinuclear localization of some MTMR2

puncta raised the question of whether these proteins localize to the endoplasmic reticulum (ER).

To assess this possibility, we visualized the ER in GFP-MTMR2-transfected HeLa cells with an

antibody against the ER resident protein Calnexin, and found no consistent indications of co-

localization (Figure 2.9). To rule out a GFP-specific effect in localization, the above protein

studies were repeated with constructs expressing N-terminally Flag-tagged MTMR2 protein

isoforms, with results completely consistent with GFP-tagged proteins (Figure 2.10). As

mentioned above, the lack of a specific MTMR2 antibody precluded us from determining the

nuclear localization of either endogenous protein isoform *in vitro* or *in vivo*. Taken together, our

data suggest that MTMR2 has a nuclear function; indeed, MTMR2 has been previously reported

to localize to the nucleus [193]. Furthermore, given that Schwann cells specifically express a

transcript that encodes a protein isoform with greater nuclear localization, this function may be

particularly important for Schwann cells and relevant to CMT4B1 disease pathogenesis.

Discussion

Loss-of-function mutations in *MTMR2* cause autosomal recessive CMT4B1, a Schwann cell-

specific demyelinating peripheral neuropathy; however, *MTMR2* is a ubiquitously expressed

gene (Laporte et al., 1998). To investigate this discrepancy and clarify any specific role of

MTMR2 in Schwann cells, we used evolutionary conservation to identify candidate

transcriptional elements regulating the expression of *MTMR2*. These studies identified an

alternative *MTMR2* promoter that is active in immortalized Schwann cells and that mediates the

expression of a novel mRNA transcript; this mRNA is enriched in Schwann cells compared to

motor neurons. Closer analysis of the promoter element revealed a predicted dimeric SOX10
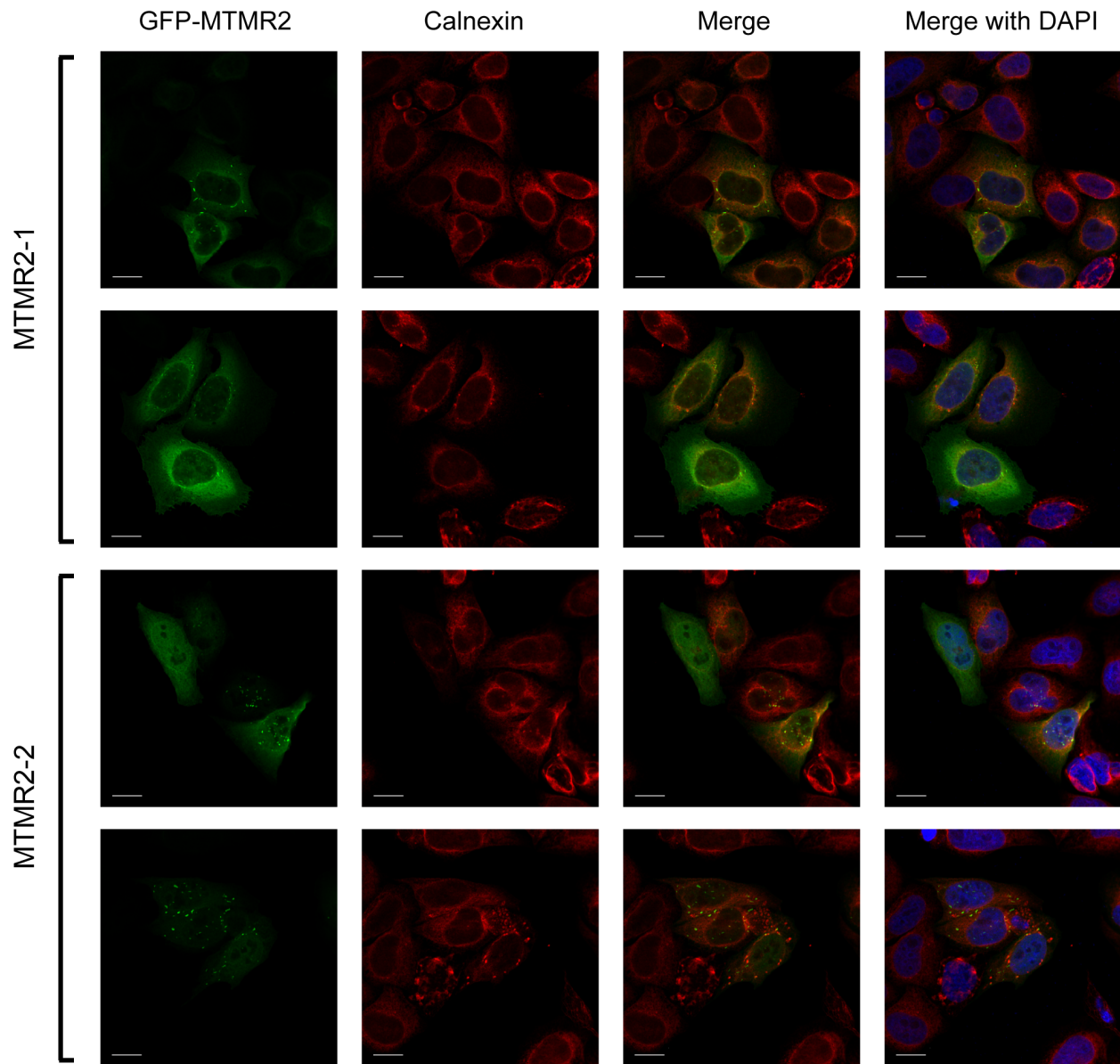
Figure 2.9 GFP-tagged MTMR2 puncta do not colocalize with an endoplasmic reticulum marker.(A) HeLa cells were transfected with constructs to express GFP-MTMR2 isoforms and endoplasmic reticulum (ER) was visualized via immunofluorescence against the ER resident protein, Calnexin. Cells were stained with DAPI to visualize nuclei and were imaged on a confocal fluorescence microscope. Scale bars, 20 μm.
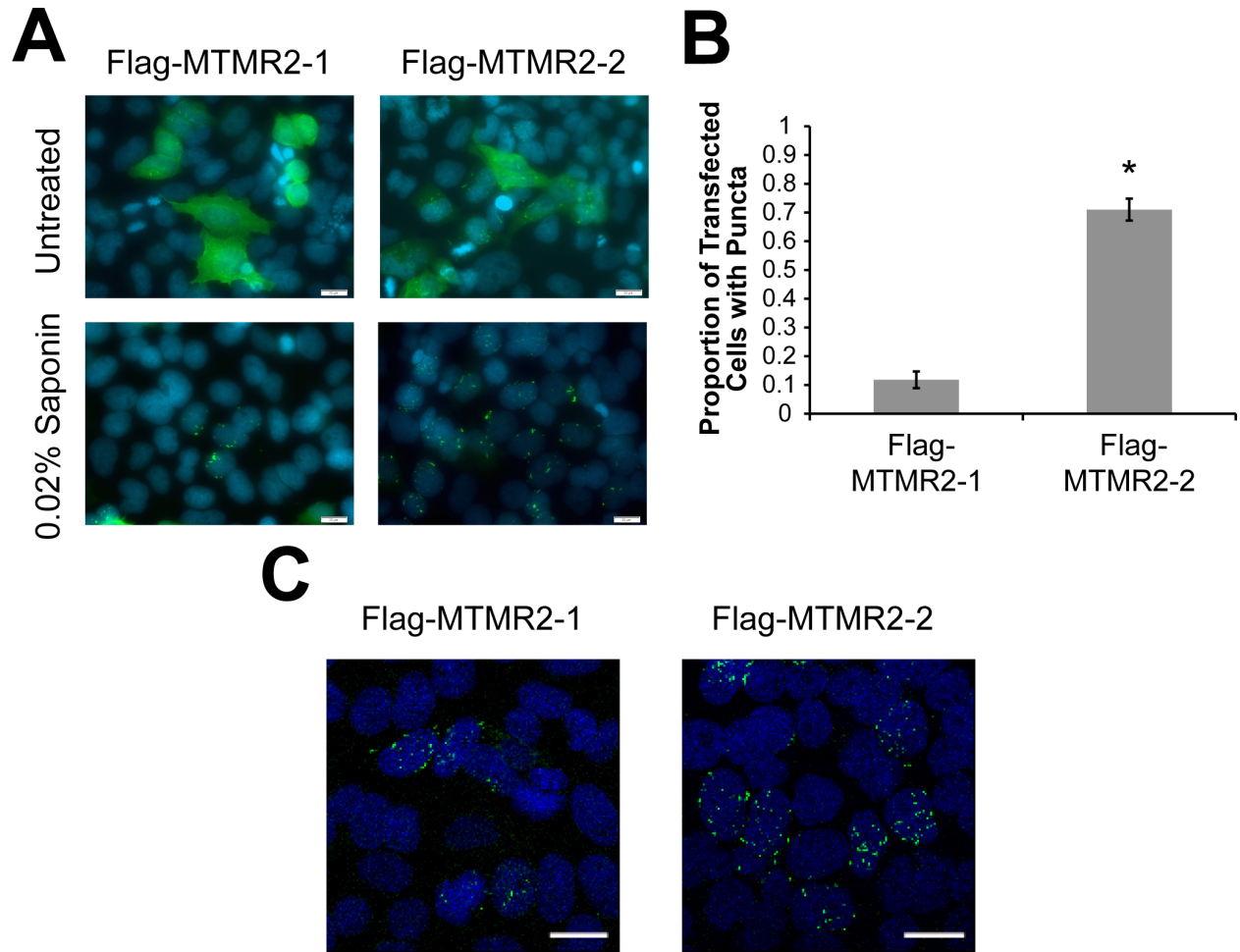
Figure 2.10 Flag-tagged MTMR2 protein isoforms replicate the findings reported with GFP-tagged isoforms.(A) HeLa cells were transfected with constructs to express Flag-MTMR2 isoforms and imaged with standard fluorescence microscopy. Cells were untreated or treated with 0.02% saponin to remove cytosolic proteins prior to fixation and stained with DAPI to visualize nuclei. Scale bars, 20 μm. (B) Quantification of puncta formed by Flag-MTMR2 isoforms. Statistical analysis: significance test for two-sample proportions (MTMR2-1 n=473, MTMR2-2 n=541, p < 0.0001). Error bars represent margin of error for 95% confidence interval. **(C)** HeLa cells were transfected with constructs to express either Flag-MTMR2 isoform, treated with 0.02% saponin to clear cytosolic proteins and DAPI to visualize nuclei (blue), and imaged with a confocal fluorescence microscope. Each image was taken from a single 0.3 μm-depth slice in the z-stack. Scale bars, 25 μm.

binding site that is evolutionarily conserved. SOX10 is essential for long-term Schwann cell function and directly regulates many genes important for myelination in the peripheral nerve [219, 221]. Our functional studies revealed that the SOX10 consensus sequence is responsible for a large portion of the regulatory activity of this alternative promoter. Furthermore, we showed that the regulatory activity and the generation of the associated mRNA transcript are reduced with disruption of SOX10 function in Schwann cells and induced with the expression of exogenous SOX10 in motor neurons. These data indicate that SOX10 is both necessary and sufficient for the activity of a Schwann cell promoter at *MTMR2* that directs the expression of a previously unreported transcript. Interestingly, we found that the *Mtmr2* isoforms are not differentially expressed across the postnatal myelination period in rat sciatic nerves. However, as SOX10 is expressed throughout the Schwann cell lineage, beginning in migrating neural crest cells and persisting in the mature myelinating Schwann cell, we cannot rule out differential expression at other stages in Schwann cell development.

Mutations in non-coding transcriptional regulatory sequences can cause or modify human disease and the identification of regulatory elements at disease-associated loci increases the genomic space to screen for disease-associated mutations. Indeed, multiple CMT-associated loci harbor non-coding mutations that cause or modify CMT disease, including *GJB1*, *MPZ*, and *SH3TC2* [222-224]. Because CMT4B1 is a recessive disease caused by loss-of-function *MTMR2* mutations, we predict that patients with a CMT4B1 phenotype but who do not carry *MTMR2* protein-coding mutations (or who carry only one coding mutation) may carry non-coding mutations at this locus. Therefore, regulatory elements important for *MTMR2* expression, including the alternative promoter reported here, are excellent candidate regions for disease-causing mutations.

The findings presented here support a new model of transcriptional regulation at the

MTMR2 locus, with at least three distinct mRNA products produced from two promoters (Figure

2.3A), which allow variable expression of two protein isoforms (Figure 2.3B). The first (i) and

second (ii) transcripts are both expressed from *MTMR2*-Prom1 and include exon 1A. However,

the exclusion (i) or inclusion (ii) of exon 2A, which harbors an in-frame stop codon, functionally

distinguishes the two. The first (i) encodes the full-length protein (MTMR2 isoform 1, MTMR2-

1), while the second (ii) is predicted to encode isoform 2 (MTMR2-2). The novel mRNA

described here (iii) is generated from *MTMR2*-Prom2, includes exon 1B, and only encodes the

shorter protein isoform MTMR2-2.

Our transcriptional regulatory data suggest that the MTMR2-2 protein plays an important

role in Schwann cells. To investigate this possibility, we analyzed the localization of each protein

isoform when overexpressed in HeLa cells. These studies revealed that both isoforms localize

both diffusely in the cytoplasm and to discrete puncta. However, MTMR2-2 shows a greater

propensity toward puncta formation than MTMR2-1. Furthermore, we found that each isoform

localizes to puncta within the cell nucleus, and that MTMR2-2 shows greater nuclear localization

compared to MTMR2-1. These results support the notion that there are physiological differences

between the two protein isoforms and provide evidence for a nuclear function of MTMR2 (see

below). However, it is important to note that the localization studies of tagged MTMR2 isoforms

in HeLa cells are limited by the use of a cell type that is not relevant to CMT disease; our

attempts to carry out these experiments in immortalized rat Schwann cells were unsuccessful.

Identifying a SOX10-responsive *MTMR2* promoter is a critical development in

understanding the etiology of CMT4B1. Indeed, the data presented here suggest a particular

requirement for MTMR2-2 in Schwann cells and it may be that the specific loss of MTMR2-2

mediates the development of disease. Consistent with this idea, all of the mutations identified in patients with CMT4B1 fall within the coding sequence of MTMR2-2; mutations that would strictly cause MTMR2-1 loss-of-function have not been discovered [225]. However, further experiments are needed to directly test this hypothesis. For example, because total MTMR2 knockout mice recapitulate the myelin outfoldings characteristic of CMT4B [186], generating animal models with targeted disruptions of each MTMR2 isoform in isolation (*e.g.*, alternatively deleting exon 1A or exon 1B) will likely be informative.

To better understand CMT4B1 pathogenesis, the role of MTMR2 in Schwann cell biology should be further clarified. Most critically, the development of antibodies specific to MTMR2 and that also distinguish between the two protein isoforms will be necessary to define the physiological relevance of our protein localization findings *in vivo*. Future studies should also define any functional differences between MTMR2-1 and MTMR2-2 at the molecular level. Interestingly, a recent study reports that the MTMR2-2 protein isoform is better able to rescue myotubularin (MTM1) loss in yeast and mouse models of myotubular myopathy compared to the full-length MTMR2-1[226]. The improved rescue was associated with a stronger restoration of PI(3)P and PI(5)P levels, indicating that MTMR2-2 functions as a more effective phosphatase than the full-length protein. These apparent differences in the enzymatic activities of these isoforms may be of important relevance to the regulation of phosphoinositide signaling in Schwann cells. Furthermore, future studies should include investigations into the relative ability of each isoform to bind known interacting proteins (*e.g.*, MTMR5 and MTMR13) [227, 228] and unbiased screens for novel and/or isoform-specific binding partners. As such, it will be important to determine the nature of the MTMR2 puncta we observed to gain insight into a possible nuclear function of MTMR2 (in addition to the lack of colocalization with the ER [Figure 2.9] tagged

MTMR2 puncta in the nucleus do not co-localize with Cajal bodies or paraspeckles; data not shown).

The nuclear localization of overexpressed MTMR2 is somewhat unexpected given that the protein lacks a canonical nuclear localization sequence. However, some evidence suggests that the nuclear localization reported here and elsewhere [193] could be mediated by the SET interaction domain (SID). This domain is specific to the myotubularin family of proteins [229] and was characterized based on the finding that MTMR5 and myotubularin are able to bind SET domain-containing nuclear proteins *in vitro* [230]. These previous studies implicate the myotubularin family in the regulation of chromatin structure and transcription via interactions with histone methyltransferase proteins in the nucleus. However, the role of the SID in mediating MTMR2 nuclear localization and the ability of MTMR2 isoforms to interact with SET domain-containing proteins have yet to be investigated. Additionally, decades of research have established important roles for nuclear phosphoinositide signaling in chromatin and gene regulation[231]. Indeed, PI(5)P—a product of MTMR2 enzymatic activity—has been described as an important signaling molecule for a number of chromatin modifying and transcriptional complexes[232, 233]. These studies suggest that nuclear MTMR2 activity could contribute to the regulation of gene expression. All told, further work is needed to clarify the relevance of MTMR2 localization in the nucleus and how this is related to Schwann cell biology.

The data presented in this chapter provide strong evidence for a novel transcriptional regulation mechanism at the *MTMR2* locus in Schwann cells and indicate that there may be a significant degree of isoform specificity in the pathophysiology of CMT4B1. These findings may resolve the apparent discrepancy between the ubiquitous expression previously reported for *MTMR2* and the restricted peripheral myelination phenotype seen in patients with CMT4B1.

Therefore, future studies should be designed with greater consideration for the relative roles of each MTMR2 protein isoform in Schwann cells. Broadly, these findings underscore the importance of careful study of the transcriptional regulation of genes that are important in a cell type of interest. Indeed, close investigation of a locus in the proper cellular context and with the nuances of cell type-specific regulatory mechanisms in mind may provide insights that have important functional implications. And finally, this work provokes the question of genome-wide relevance for SOX10-regulated promoter use and the possibility that this mechanism mediates isoform-specific gene expression at additional loci in Schwann cells. Indeed, this question will be explored in the next chapter of this dissertation.

Chapter 3

SOX10-Regulated Promoter Use Defines Isoform-Specific Gene Expression in Schwann Cells

Introduction

The complexity of the mammalian genome—in particular the generation of multiple, unique products from a single gene locus—is critical for the development and function of specialized cell types. This complexity is realized through a variety of mechanisms, including alternative splicing, alternative start codon usage, alternative polyadenylation, and alternative promoter use[234]. Comprehensive transcription start site (TSS) mapping across many cell types and tissues shows that each human gene harbors an average of four distinct TSSs[71]. Importantly, 80% of TSSs are utilized in a restricted (*i.e.*, non-ubiquitous) manner, indicating that regulation of TSS use is an important contributor to cell type-specific gene expression. Indeed, it has been reported that differential use of transcription start and termination sites, rather than alternative splicing, accounts for the majority of isoform diversity in mammalian genomes[68-70].

As the myelinating cells of the peripheral nervous system, Schwann cells mediate the saltatory conduction of action potentials down the lengths of peripheral axons. Additionally, these cells provide critical trophic and metabolic support for axons through development and in adulthood. Deficits of Schwann cell function, including those associated with inherited demyelinating peripheral neuropathies, can impair sensory and motor function to the point of rendering affected individuals wheelchair-bound[235]. While 60-90% of individuals with

demyelinating peripheral neuropathy achieve a genetic diagnosis[103, 105, 106], much remains to be learned about Schwann cell biology. In fact, these cells are relatively understudied, which is underscored by the lack of treatment options for disorders of myelination. Therefore, further study of Schwann cell biology and characterization of the genes that are important for their function is necessary toward a more complete understanding of Schwann cell function and related human diseases.

The SRY-box transcription factor 10 (SOX10) is a transcriptional activator that is critical for the development and maintenance of Schwann cells. SOX10 interacts with DNA through a high-mobility group (HMG) DNA-binding domain that recognizes the 5'(A/T)AACAA(T/A)3' consensus sequence[217, 236], and can bind DNA as a monomer or dimer[121]. Further, SOX10 is known to facilitate gene expression through a number of mechanisms including interacting with the mediator complex[134] and transcriptional elongation factors[136], and through the recruitment of chromatin remodelers[137] and histone modifying enzymes[138]. SOX10 is expressed early in the Schwann cell lineage, beginning in the migrating neural crest, with expression persisting in fully differentiated myelinating cells[113]. Loss of SOX10 expression in Schwann cells—even after the completion of developmental myelination—leads to demyelination[149]. Further, SOX10 target genes that have been characterized in Schwann cells to date include genes that are critical for myelination and that have been implicated in demyelinating disease (*e.g.*, *EGR2*, *PMP22*, *MPZ*)[161, 163, 164]. Therefore, the identification of novel SOX10 target genes in Schwann cells can be considered a viable strategy toward gaining new insights into peripheral nerve myelination.

SOX10 is known to function at distal enhancer elements to induce expression of numerous target genes[151, 153, 164]. However, SOX10 ChIP-Seq data generated from rat

72

sciatic nerve revealed that SOX10 binds more frequently at the proximal promoter regions of target genes compared to other myelin-related transcription factors. For example, among loci that are repressed in response to SOX10 knockdown *in vitro* and that harbor a SOX10 ChIP-Seq peak within 100 kilobases, 40% show evidence of SOX10 binding in the 2 kilobases upstream of the TSS[154]. In contrast, EGR2—a promyelinating transcription factor that is also critical for Schwann cell function—was found to bind immediately upstream of the TSS at less than 5% of target loci[154]. Consistent with the importance of SOX10 in activating gene promoters, multiple studies have described SOX10 binding directly at the promoters of myelin-related genes [119, 161, 162, 165]. Furthermore, a number of studies—including the work presented in Chapter 2 of this dissertation—have reported on loci harboring SOX10-regulated alternative promoters that direct expression of unique transcript and protein isoforms[169, 170, 195]. These findings raise important questions about isoform specificity at these loci and the roles of the SOX10-regulated gene products in Schwann cell function. To date, the relationship between SOX10 function and variable TSS usage has not been comprehensively explored at the genome level.

Given the importance of regulated TSS use for cell type-specific gene expression[70], the critical role of SOX10 for Schwann cell function, and the identification of SOX10-regulated promoters driving isoform-specific gene expression at multiple loci, I sought to define the extent of SOX10-regulated promoter use driving isoform-specific gene expression in Schwann cells genome-wide. In this chapter, I report on the identification of SOX10-regulated promoters by assessing TSS use: (***i***) in the context of a mature nerve; (***ii***) in differentiating primary Schwann cells; and (***iii***) upon loss of SOX10. I also provide an assessment of the characteristics of SOX10-bound promoter elements that exhibit SOX10-dependent expression, and discuss the potential for these studies to inform our understanding of isoform-specific expression at SOX10

target genes. In sum, these data hold promise toward a more complete understanding of the transcriptional biology of Schwann cells and will catalyze future efforts to study the nuances of isoform-specific gene expression with relevance for myelination and myelin-related disease.

Note that I completed all of the work presented in this chapter, with the following exception: next-generation sequencing of Tn5Prime libraries was performed by the University of Michigan DNA Sequencing Core.

Materials and Methods

Primary Schwann cell differentiation assay

Differentiation assays were completed as described by Paula Monje[237]. Primary rat Schwann cells (Kerafast, Boston, MA) were maintained under standard growth conditions in complete Schwann cell (SC) medium: DMEM supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, 50 U/mL penicillin, 50 g/mL streptomycin, 25 μg/mL gentamicin, 10 nM neuregulin EGF domain, and 2 μM forskolin. For differentiation assays, culture dishes (6-well plates or T-25 flasks) were coated with poly-L-lysine (PLL) and laminin prior to cell plating. 1 mL 0.01% PLL solution (Sigma, St. Louis, MO) per 25 cm^2 of surface area was applied to wells or flasks and brought up to sufficient volume to cover the bottom of the chamber with sterile water. The chamber and solution were allowed to incubate for 5 minutes at room temperature, then the solution was aspirated and the chamber was allowed to dry completely. Laminin derived from human fibroblasts (Sigma) was diluted in Hank's Basic Saline Solution (HBSS) without calcium and magnesium, and applied to dishes at 1 μg laminin per cm^2 of surface area. Dishes were incubated for 1 hour at room temperature. The laminin solution was then removed and dishes were washed with sterile water and allowed to dry completely before plating cells.

On day 1, primary Schwann cells were plated at ~215,000 cells per well in 6-well plates or ~600,000 cells per flask in T-25 flasks in complete SC medium. On day 2, medium was removed and replaced with D10 medium: DMEM supplemented with 10% FBS, 2 mM L-glutamine, 50 U/mL penicillin, 50 g/mL streptomycin, and 25 µg/mL gentamicin. On day 3, medium was removed and replaced with D5 medium: DMEM supplemented with 5% FBS, 2 mM L-glutamine, 50 U/mL penicillin, 50 g/mL streptomycin, and 25 µg/mL gentamicin. On day 4, medium was removed and replaced with D5 medium supplemented with 250 µM CPT-cAMP (Axxora, Farmingdale, NY) or vehicle. Subsequently, condition-specific medium was replenished each day. On day 7, cells were washed with HBSS and incubated in 0.15% trypsin to dissociate cells from the culture dish. Trypsin solution was quenched with D10 medium upon cellular detachment. The resulting cell suspension was centrifuged at 200 x g for 10 minutes at 4°C for downstream RNA or protein isolation.

Generation of S16 ΔSOX10 cell lines

Guide RNAs were designed against the first coding exon at the rat *Sox10* locus, such that Cas9-mediated cutting was predicted to occur ~261 bases (guide 1) or ~273 bases (guide 2) downstream of the *Sox10* start codon. Guides were cloned by restriction enzyme digestion and ligation into the *BbsI* site of the PX459 plasmid for co-expression with Cas9 and a puromycin resistance gene[238]. Proper insertion was confirmed with Sanger sequencing. S16 cells[199] were grown under standard conditions and plated at 30,000 cells/well in a 6-well plate in standard growth medium: DMEM supplemented with 10% FBS, 2 mM L-glutamine, 50 U/mL penicillin, and 50 g/mL streptomycin. The next day, guide RNA-containing PX459 plasmids (6 µg DNA per well) were transfected individually using Lipofectamine 2000 (ThermoFisher Scientific, Waltham, MA) in Opti-Mem (ThermoFisher Scientific) according to manufacturer's

protocol. Four hours after transfection, transfection medium was removed and replaced with standard growth medium. Twenty-four hours after transfection, cells were treated with growth medium containing 5 μg/mL puromycin. Puromycin-containing medium was replenished the following day. After 48 hours of puromycin treatment, cells were returned to standard growth medium. Surviving cells were grown to confluency and a bulk DNA isolation was performed on a portion of the cells. PCR was performed using primers flanking the guide cut sites and PCR Supermix (ThermoFisher Scientific; note that all primers and oligos were ordered from IDT, Coralville, IA; see Appendix for sequences); resulting amplicons were cloned and Sanger sequenced to verify the presence of indels as confirmation of editing activity. Subsequently, edited populations were collected by incubation in 2.5% trypsin. Upon cellular detachment, trypsin solution was quenched using standard growth media. The resulting cell suspensions were centrifuged at 800 x g for 2 minutes, resuspended in PBS, and sorted using a FACS machine to isolate single cells in individual wells of 96-well plates for clonal expansion. Resulting clones were subjected to crude DNA isolation using QuickExtract (Epicentre Technologies, Madison, WI) and targeted PCR surrounding the editing site. Amplicons were Tn5-tagmented and the resulting samples were subjected to PCR with barcoded primers; pooled PCR products were sequenced on an Illumina MiSeq sequencer. Tn5 enzyme loaded with Tn5ME-A/R and Tn5ME-B/R adapters and barcoded PCR primers were a kind gift of Dr. Jacob Kitzman. Clones with exclusively frameshift-bearing alleles (*i.e.*, no detection of unedited alleles or frame-preserving indels) were expanded via standard culture conditions for RNA and protein isolation.

Protein isolation and western blots

Western blotting analysis was performed to validate the differentiation of primary Schwann cells and the ablation of SOX10 expression in S16 ΔSOX10 clones. Cells were

incubated in 0.15% (primary Schwann cells) or 2.5% (S16 cells) trypsin as described above, collected, and centrifuged at 200 x g for 10 minutes (primary Schwann cells) or 800 x g for 2 minutes (S16 cells). Subsequently, medium was removed, cell pellets were resuspended in PBS to remove any remaining medium, then centrifuged again using the same conditions. PBS was then removed and cell pellets were resuspended in RIPA buffer (Pierce/ThermoFisher Scientific) supplemented with protease inhibitor cocktail (ThermoFisher Scientific). Cell suspensions were mixed by rocking for 30 minutes at 4°C then centrifuged at 16,000 x g at 4°C for 30 minutes. Lysates were moved into a clean tube and stored at -20°C. Protein yield was measured with a BCA Protein Assay (ThermoFisher Scientific).

For western blotting, each 10-µg sample of protein was supplemented with 2X SDS sample buffer (ThermoFisher Scientific) and beta-mercaptoethanol, incubated at 99°C for 5 minutes, then electrophoresed on a 4-20% gradient Tris-Glycine polyacrylamide gel (ThermoFisher Scientific) at 150V for 1.5 hours at room temperature. Protein was transferred to an Immobilon PVDF membrane (ThermoFisher Scientific) in Tris-glycine transfer buffer (ThermoFisher Scientific) containing 10% methanol for ~18 hours at room temperature and 0.03 A. Membranes were washed briefly in TBST, then transferred to 2% milk in TBST overnight, rocking at 4°C. The next day, membranes were transferred to primary antibody dilutions in 2% milk and incubated at 4°C overnight with rocking. Primary antibodies included: anti-MPZ (rabbit; 1:5000; EMD Millipore, Burlington, MA), anti-cJun (rabbit; 1:5000; Cell Signaling Technology, Danvers, MA), anti-IARS (rabbit; 1:5000; GeneTex, Irvine, CA), and anti-SOX10 (guinea pig; 1:2000; kind gift from Dr. Michael Wegner). Membranes were washed three times with TBST. Secondary antibodies conjugated to horse radish peroxidase were diluted in 2% milk and incubated with membranes for one hour rocking at room temperature. Antibodies included

anti-rabbit HRP (donkey; 1:5000; EMD Millipore, Burlington, MA), anti-mouse HRP (goat; 1:2000; ThermoFisher Scientific), and anti-guinea pig HRP (goat; 1:5000; kind gift from Dr. Miriam Meisler). After three washes with 1X TBST, membranes were incubated with West Dura HRP substrate (ThermoFisher Scientific) for four minutes, then membranes were drained and exposed to X-ray film for between one second and three minutes.

RNA isolation and RT-PCR

RNA was isolated from rat sciatic nerve, primary Schwann cells, and S16 cells using the RNeasy RNA isolation kit (Qiagen USA, Germantown, MD) according to the manufacturer's protocol, with the addition of on-column RNase-free DNase treatment (Qiagen USA). Each RNA sample was eluted in 30 µl water and the concentration was assessed using a Nanodrop Lite (ThermoFisher Scientific).

To validate the loss of *Sox10* transcript expression in S16 ΔSOX10 clones, cDNA was generated using 1 µg of RNA from parental S16 cells and each knockout clone with the High Capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific) according to the manufacturer's protocol. Two independent PCR primer sets flanking the boundaries between *Sox10* exons 1 and 2 and between exons 2 and 3 were designed to test for the presence of *Sox10* transcripts. PCR was performed using PCR Supermix (ThermoFisher Scientific) and each individual *Sox10* primer set, as well as positive control primers to amplify *Actb* transcripts.

Generation of Tn5Prime next-generation sequencing libraries

Tn5Prime libraries were prepared as described by Cole and colleagues[239]. RNA samples were subjected to reverse transcription using the SMARTscribe Reverse Transcriptase kit (Clontech/Takara Bio USA, Moutain View, CA). 5 ng (2 µl) total RNA was incubated with 1

μl Oligo-dT-smartseq2 primer (20 μM) and 1 μl water at 72°C for three minutes, then cooled on ice. Upon the addition of 2 μl 5X First Strand Buffer, 1 μl 10 mM dNTPs (Invitrogen/ThermoFisher Scientific), 1 μl 20 mM DTT, and 1 μl Nextera A TSO1 (20 mM), the solution was pipette-mixed and 1 μl SMARTscribe Reverse Transcriptase was added. The reaction was then pipette-mixed and incubated at 42°C for 60 minutes. The enzyme was inactivated at 70°C for 15 minutes. 1 μl each of 1:10 dilutions of RNase A (ThermoFisher Scientific) and Lambda Exonuclease (New England Biolabs, Ipswich, MA) were added and the reaction was pipette-mixed and then incubated at 37°C for 30 minutes.

Subsequently, PCR was performed using the KAPA Biosystems Hifi HotStart Readymix PCR kit (ThermoFisher Scientific). The reaction included 10 μl of water, 12.5 μl 2X KAPA Hifi Readymix, 0.75 μl IPSCR primer (10 μM), 0.75 μl Nextera A Index (N501-N508) primer (10 μM), and 1 μl cDNA from the reverse transcription described above. The reaction was cycled as follows: 95°C for three minutes, 15 cycles of: 98°C for 20 seconds, 67°C for 15 seconds, and 72°C for 4 minutes. Final extension time was 72°C for 5 minutes.

Samples were "tagmented" with Tn5 enzyme that had been pre-loaded with Tn5ME-B/R adapters. To prepare the adapters, Tn5ME-B and Tn5ME-R oligos were suspended at 1 mM in LoTE buffer (10 mL LoTE: 100 μl 1 M Tris-HCl pH 8, 2 μl 0.5 M EDTA, bring to 10 mL with water). 5 μl of each oligo were combined with 25 μl fresh 2X annealing buffer (1.5 mL 2X annealing buffer: 30 μl 1 M Tris-HCl pH 8, 6 μl 5 M NaCl, bring to 1.5 mL with water) and 15 μl water. The solution was incubated in a thermocycler at 95°C for 10 minutes, then the temperature was programmed to decrease by 0.5°C per minute until it reached 25°C, where it stayed for 10 minutes. Annealed oligos were placed on ice and stored at -20°C. 15.7 μl purified recombinant Tn5 enzyme (kind gift from Dr. Jacob Kitzman) was mixed with 2.5 μl annealed

Tn5ME-B/R oligos and incubated at room temperature for 60 minutes. The solution was then moved to ice and stored at -20°C. Tagmentation reactions included 5 µl indexed PCR reaction, 1 µl oligo-loaded Tn5 enzyme, 10 µl water, and 4 µl 5X TAPS-PEG buffer (50 mM TAPS-NaOH pH 8.5, 25 mM $MgCl_2$, 40% PEG-4000; stored at -20°C). The reaction was incubated at 55°C for 5 minutes. 5 µl 0.2% SDS solution was added, pipette-mixed, and the reaction was incubated at 55°C for another 5 minutes.

An additional PCR incorporated full Nextera adapter sequences using the KAPA Hifi Polymerase PCR kit (ThermoFisher Scientific). The reaction included 12.25 µl water, 5 µl 5X KAPA Hifi Fidelity Buffer, 0.75 µl 10 mM KAPA dNTP mix, 0.75 µl Nextera B Primer (10 µM; N701 index used for all samples), 0.75 µl Nextera A Universal Primer (10 µM), 5 µl tagmented product, and 0.5 µl KAPA Hifi DNA Polymerase. The reaction was cycled as follows: 72°C for 6 minutes (nick-translation), 98°C for 30 seconds, 13 cycles of: 98°C for 10 seconds, 63°C for 30 seconds, 72°C for 2 minutes. Final extension was 72°C for 5 minutes. Note: electrophoresis of 5 µl of the above reaction on a 1% TBE agarose gel should give a visible smear. If no smear is detected, tagmentation may be insufficient or SDS carry-over from the tagmentation reaction may be inhibiting polymerase activity in the final PCR. Reducing the volume of the tagmentation reaction as template in final PCR to 2 µl may improve yield. To collect sufficient product for next-generation sequencing, two PCRs per sample were electrophoresed on a 0.8% low-melt agarose TAE gel. The gel was visualized on a Safe Imager Blue Light Transilluminator (ThermoFisher Scientific) and the 400 bp to 1 kb size region was excised. The gel fragments from a single sample were combined and DNA was isolated using the Qiagen Gel Extraction Kit (Qiagen USA) according to manufacturer's protocol. Samples were eluted in 30 µl 65°C water

that was allowed to sit on the column for 2 hours. Sample yield was measured using the Qubit Broad Range kit for double-stranded DNA (ThermoFisher Scientific).

Two Tn5Prime libraries were generated using RNA samples from two independent adult rat sciatic nerves (ages 6-9 months). Six libraries were generated using RNA isolated from independent populations of primary Schwann cells, three each that were CPT-cAMP- or control-treated. Two libraries were generated using two independent RNA samples from parental S16 cells. Four libraries were generated using RNA samples from each of the four independent ΔSOX10 S16 clonal cell lines. Sequencing data from all samples were included in downstream analyses.

Sequencing and analysis of Tn5Prime libraries

Libraries were subjected to next-generation sequencing on Illumina HiSeq 4000 (Sciatic Nerve and S16 libraries) or NovaSeq (primary Schwann cell libraries) sequencers. Quality of data was assessed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapter sequences were trimmed with Cutadapt[240], trimmed reads were mapped to the rat genome (rn5) with STAR[241], and bam files were generated with SAMtools[242]. Reads were organized by start site using the Make_CTSS script from Takahashi and colleagues[243] and the resulting start site counts were clustered into defined transcription start sites (TSSs) using Paraclu[244]. Read counts per TSS per sample were generated using featureCounts[245] and statistical analysis was done using edgeR[246].

Software and datasets employed for computational analyses

Genomic coordinates for rat RefSeq (rn5) genes were extracted from the UCSC Genome Browser[196]. TSSs were assigned to genes using BEDTools[247] and requiring the TSS to map within 1 kilobase and on the same strand as the gene. To account for the poor gene annotation in

81

the rat genome, TSSs that did not map to a rat gene were converted to the orthologous mouse

coordinates (mm10) using the liftOver executable from the UCSC Genome Browser[196] and

mapped to mouse RefSeq genes in the same manner. H3K4me3 ChIP-Seq peaks from rat scatic

nerve[248], SOX10 ChIP-Seq peaks from rat sciatic nerve[154], and TSSs defined by Tn5Prime

in sciatic nerve were intersected using BEDTools to define TSSs mapping within 1 kilobase of

H3K4me3 and/or SOX10 ChIP-Seq peaks. Human RefSeq transcript annotations (hg38) for loci

associated with SOX10-bound promoters were downloaded through the UCSC Genome

Browser[196] and manually curated for transcription start site and coding sequence diversity.

Empirical cumulative distribution function curves for TSS expression levels were generated

using the plot() and ecdf() functions in R[249]. Aggregate analysis of SOX10 ChIP-Seq data

surrounding TSSs was performed using metagene[250] and similaRpeak

(https://github.com/adeschen/similaRpeak). Gene ontology analyses were performed using

geneontology.org[251, 252]. Heatmaps organized by gene ontology terms were generated using

Heatmapper[253].

Genomic sequences of regions surrounding TSSs were extracted using the UCSC

Genome Browser[196]. SOX10, TATA Box, and Initiator motifs were identified using custom

pearl scripts in Bioperl[254]. Conservation scores for SOX10 motif sequences were extracted

from the UCSC Genome Browser rn5 13-way PhastCons data file[196] using a publicly

available custom script written by Dr. Ian Donaldson, University of Manchester. GC content and

CpG islands were defined by the EMBOSS freak and cpgplot tools, respectively[255]. CpG

islands were defined using cpgplot default parameters. CAGE data from 11 mouse tissues

including cortex, spinal cord, skin, lung, heart, colon, thymus, stomach, liver, ovary, and testis

were downloaded from FANTOM5[256]. TSS genomic coordinates were converted to the

orthologous mouse (mm10) coordinates using liftOver from the UCSC Genome Browser[196] and featureCounts was used to quantify read counts in each tissue. TSS-specific Tau scores were calculated as described by Yanai and colleagues[257].

Results

SOX10 binds to promoters in Schwann cells in vivo

We and others have reported on SOX10 activity at the promoter elements of critical myelin-related genes in Schwann cells[119, 161, 162, 195]. This suggests that genome-wide characterization of SOX10 activity at promoter elements represents a strategy for the unbiased identification of gene products that play important roles in peripheral nerve myelination. To characterize the prevalence of SOX10 binding at promoter elements in Schwann cells *in vivo*, I performed Tn5Prime library preparations[239] using RNA isolated from adult rat sciatic nerves (age 6-9 months) to define transcription start sites (TSSs) in this tissue. Briefly, this method anchors a sequencing adaptor to the 5' ends of transcripts through the use of a template-switching oligo during reverse transcription and employs next-generation sequencing to develop a dataset of mRNA 5' sequences that can be mapped to the genome to infer TSS positions. After generating and mapping the TSS data, I intersected them with published SOX10 and H3K4me3 ChIP-seq datasets [248, 258] from rat sciatic nerve to account for SOX10 binding and active promoters, respectively. I found that 4,993 of the 39,706 TSSs (12.6%) expressed in sciatic nerve reside within one kilobase of an H3K4me3 ChIP-Seq peak that overlaps a SOX10 ChIP-Seq peak; this dataset represents candidate TSSs proximally regulated by SOX10 (Figure 3.1A). These 4,993 TSSs map to 2,993 unique loci, including previously characterized SOX10 target genes (*e.g.*, *Mpz, Mbp,* and *Pmp22*). Ontology analysis of these loci shows an enrichment for gene products associated with 'regulation of myelination' (GO:0031641; FDR-corrected p-value

= 0.0324). In addition, 7,455 of the 39,706 TSSs (18.8%) expressed in sciatic nerve map to an H3K4me3 ChIP-Seq peak but not a SOX10 ChIP-Seq peak, 431 (1.1%) map to a SOX10 peak but not an H3K4me3 peak, and the remaining 26,827 (67.6%) TSSs do not map to an H3K4me3 or SOX10 ChIP-Seq peak (Figure 3.1A). The large number of TSSs defined by Tn5Prime that are not associated with an H3K4me3 ChIP-Seq peak is consistent with other TSS-mapping approaches which have identified extensive low-level 'exonic promoter' activity throughout gene bodies[34]. In support of this notion, we find that TSSs that are not associated with an H3K4me3 mark: (*i*) are expressed at lower levels compared to those associated with promoter marks, with greater than 80% of these TSSs detected at fewer than 5 reads per million (Figure 3.1B); and (*ii*) map to gene bodies at the same rate (greater than 90%) as those TSSs associated with promoter marks (Table 3.1).

To assess the specificity of SOX10 binding at TSSs associated with SOX10 ChIP-Seq peaks and to investigate the spatial relationship between SOX10 binding and the TSS, I performed a metagene analysis[250] on the aggregate SOX10 ChIP-Seq signal[154] in a 2-kilobase window surrounding each class of TSS. For those TSSs residing near an H3K4me3 peak, SOX10 ChIP-Seq signal is concentrated directly over the TSS, in the -500 to +500 base pair region (Figure 3.1C). This finding is consistent with SOX10 proximally regulating transcript expression at these promoters. Moreover, TSSs associated with SOX10 and H3K4me3 ChIP-Seq peaks exhibit stronger SOX10 ChIP-Seq signal than those associated with only an H3K4me3 peak, as expected; this supports prioritization of these elements as highly confident candidate SOX10-regulated promoters. Therefore, I considered the 4,993 TSSs associated with both H3K4me3 and SOX10 ChIP-Seq peaks as candidates of interest for further study.
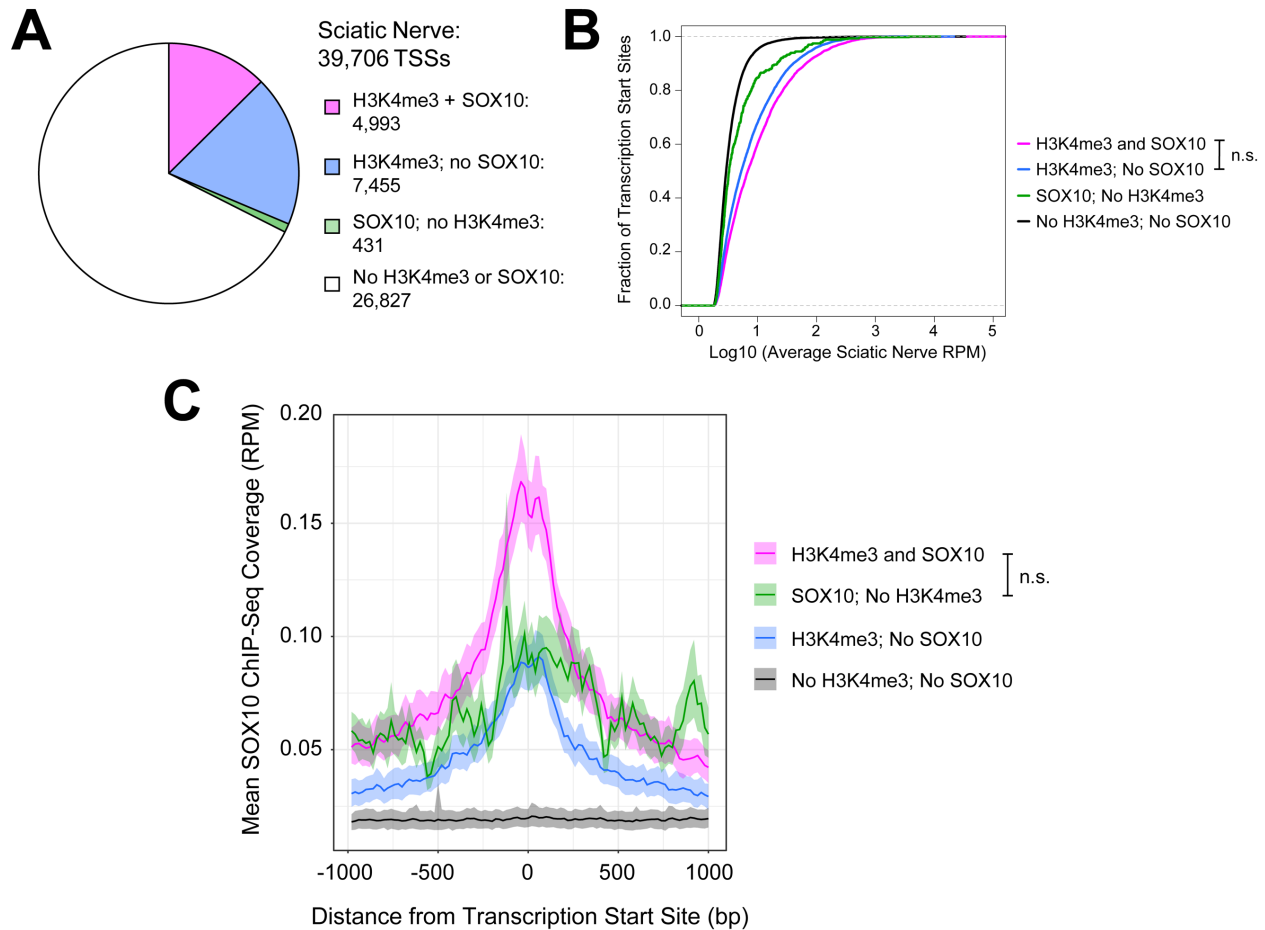
Figure 3.1 SOX10 binds to promoter elements in Schwann cells *in vivo*. (A) Tn5Prime-defined transcription start sites (TSSs) in sciatic nerve were intersected with SOX10 and H3K4me3 ChIP-Seq peaks to define TSSs mapping within 1 kilobase of these marks in combination or isolation. (B) Expression levels of TSSs associated with each combination of ChIP-Seq peaks as shown in panel A were compared using empirical cumulative distribution functions. The x-axis indicates the averaged expression level (reads per million, RPM) across the two sciatic nerve libraries, log10 transformed. The y-axis indicates the cumulative fraction of TSSs. Comparisons between groups were performed using the Mann-Whitney U Test. Every comparison was found to be statistically significant (p < 0.00001) with the exception of the TSSs mapping near H3K4me3 and SOX10, compared to those mapping near H3K4me3 but not SOX10 (n.s., not significant; p=0.15). (C) Aggregate SOX10 ChIP-Seq data were analyzed in the 2 kilobase region surrounding TSSs associated with each combination of SOX10 and H3K4me3 ChIP-Seq peaks as shown in panel A. The y-axis indicates average SOX10 ChIP-Seq signal (RPM, reads per million) at each class of TSS. The x-axis indicates genomic distance from the TSS, which is centered at position 0 (bp, base pairs). Regions surrounding negative-strand TSSs were reversed to orient all regions toward transcription running left to right. Comparisons between groups were performed using the area under the curve and permutation-based analysis; each comparison was found to be statistically significant (p<0.01) with the exception of the TSSs mapping near H3K4me3 and SOX10, compared to those mapping near SOX10 but not H3K4me3 (n.s., not significant; p=0.09).

Group	Number of TSSs	Number Mapping Within 1kb of Gene on Same Strand	Percent Mapping Within 1kb of Gene on Same Strand	Number of Unique Genes
H3K4 and SOX10	4,993	4,698	94.09	2,993
H3K4; No SOX10	7,455	6,918	92.80	4,741
SOX10; No H3K4	431	392	90.95	225
No H3K4 and No SOX10	26,827	24,867	92.69	7,088
Total	39,706	36,875	92.87	10,000

Table 3.1 Each class of transcription start site map to gene bodies at a high rate.
Transcription start sites (TSSs) were assigned to RefSeq gene loci, requiring a TSS to map
within 1 kilobase of a gene on the same strand.

To characterize the potential for the 4,993 candidate TSSs to inform our understanding of isoform-specific gene expression in Schwann cells, I assessed the complexity of human RefSeq transcripts annotated for the 2,993 loci that harbor candidate SOX10-regulated promoters. I found that 739 of the 2,993 loci (25%) are annotated with at least two RefSeq transcripts originating from unique transcription start sites. Furthermore, 525 of the 2,993 loci (17.5%) include at least two TSSs associated with transcripts that harbor unique protein-coding sequences. Taken as a whole, these data support the need to functionally characterize candidate SOX10 target promoters to delineate isoform-specific expression profiles relevant for myelinating Schwann cells.

Defining SOX10 function at promoters in differentiating Schwann cells

The Tn5Prime data collected from adult sciatic nerve provide important insight into TSS use in the context of mature Schwann cells. However, these analyses are limited by the cellular heterogeneity of the sciatic nerve and the lack of developmental expression profiles for transcripts of interest. SOX10 is known to induce target genes that are expressed in a temporally-specific manner and that are associated with stage-specific functions during Schwann cell development[111]. To address these considerations, I assayed TSS use in a well-described *in vitro* model of primary Schwann cell differentiation[237]. Briefly, as cells are treated with CPT-cAMP in culture, they exhibit a shift toward a more differentiated state. Indeed, by culturing cells with or without CPT-cAMP, isolating protein, and assessing expression of established marker proteins, I was able to replicate the induction of myelin proteins (MPZ) and reduced expression of immature Schwann cell markers (c-Jun) in CPT-cAMP-treated cells (Figure 3.2). To measure TSS expression in this paradigm we generated and sequenced Tn5Prime libraries using RNA isolated from cAMP- and control-treated cells, and focus here specifically on the
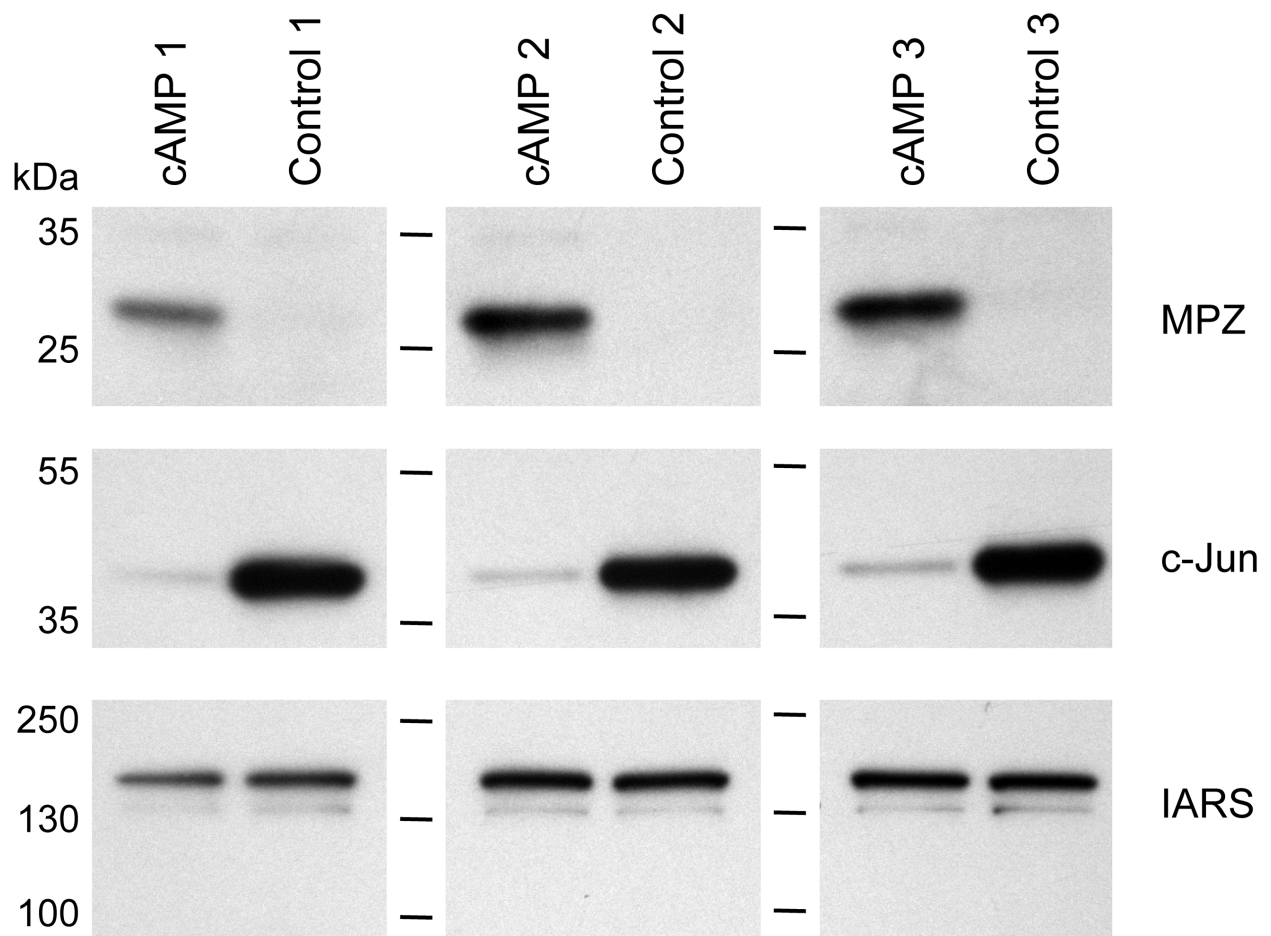
Figure 3.2 Validation of CPT-cAMP-induced differentiation of primary Schwann cells. 10 µg of protein lysates from independent populations of primary Schwann cells treated with CPT-cAMP (cAMP) or vehicle (Control) were used to test the expression of markers associated with differentiation. Anti-MPZ was used a positive marker of differentiation, while anti-cJun serves as a negative marker of differentiation. Anti-IARS was used as a protein loading control. Numbered dashes between blots indicate the position of protein size markers in kilodaltons (kDa).

expression profiles of the 4,993 candidate TSSs associated with SOX10-bound promoters in sciatic nerve (see above).

First, 4,044 of the 4,993 candidate TSSs identified in sciatic nerve (81%) were expressed in control and/or cAMP-treated primary Schwann cells (Figure 3.3A). Notably, those TSSs that were not expressed by primary Schwann cells *in vitro* were among the most lowly-expressed TSSs in the sciatic nerve with greater than 80% detected at less than 5 reads per million *in vivo* (Figure 3.3B). Of the 4,044 TSSs expressed in this model of a developing Schwann cell, 465 (11.5%) exhibited increased expression with CPT-cAMP treatment while 401 (9.9%) showed reduced expression with treatment (Figure 3.3 A). These data provide cellular differentiation expression profiles for more than 4,000 TSSs that are associated with SOX10-bound promoters and are especially relevant for the 866 TSSs with altered expression upon differentiation.

To assess the functional implications associated with these findings, I performed a gene ontology enrichment analysis with the upregulated and downregulated groups. The 465 upregulated TSSs map to 355 unique gene loci, and these genes are enriched for functional roles related to lipid synthesis, myelination, protein translation, and targeting of proteins to the endoplasmic reticulum and membrane (Table 3.2 and Figure 3.3C). These results not only support the validity of the differentiation model but also emphasize the presence of SOX10-bound promoters at broadly-expressed, "housekeeping" genes with functions that are particularly critical during active myelination such as lipid synthesis and protein targeting to the membrane. Conversely, the 401 downregulated TSSs map to 372 unique loci and are enriched for functions related to intracellular vesicle sorting and transport, cortical cytoskeleton regulation, and the unfolded protein response (Table 3.2 and Figure 3.3C). These findings suggest that the candidate SOX10 target transcripts associated with these functions may be most relevant for Schwann cell

89
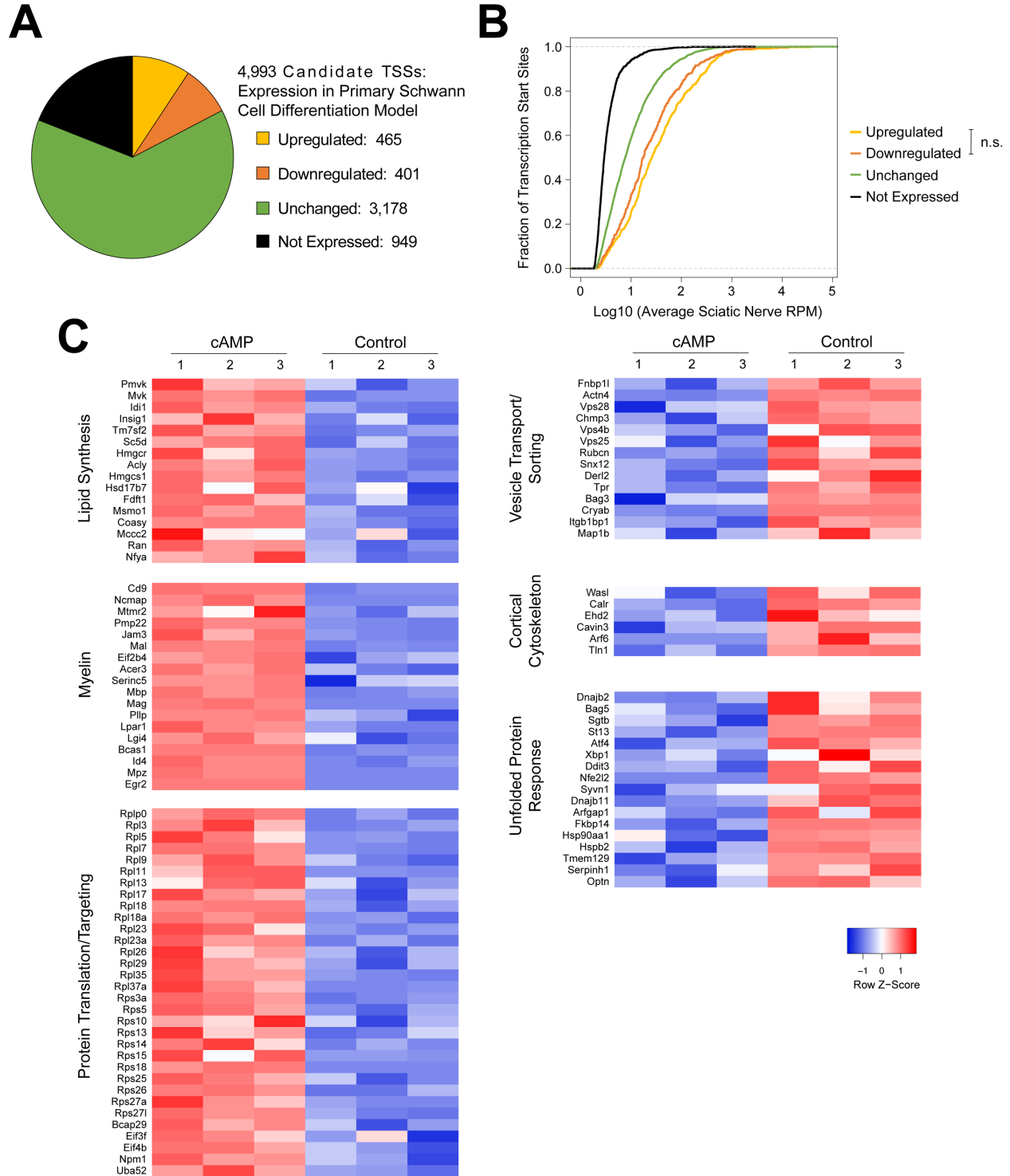
Figure 3.3 Assessment of transcription start sites associated with SOX10-bound promoters in differentiating primary Schwann cells. (A) Expression of transcription start sites (TSSs) associated with SOX10-bound promoters in sciatic nerve were assessed for expression in vehicle- and cAMP-treated primary Schwann cells. TSSs were grouped by those upregulated by cAMP (yellow), downregulated by cAMP (orange), unchanged (green), or not expressed in

either condition (black). (B) *In vivo* expression levels of TSSs belonging to each group as shown in panel A were compared using empirical cumulative distribution functions. The x-axis indicates the averaged expression level (reads per million, RPM) across the two sciatic nerve libraries, log10 transformed. The y-axis indicates the cumulative fraction of TSSs. Comparisons between groups were performed using the Mann-Whitney U Test. Every comparison was found to be statistically significant ($p < 0.03$) with the exception of the TSSs that were upregulated by cAMP, compared to those that were downregulated (n.s., not significant; $p=0.10$). (C) Gene ontology analysis was completed on loci associated with upregulated and downregulated TSSs. Expression levels of TSSs associated with the resulting terms were compiled into a heatmap to visualize consistency of expression difference across replicates and identify gene loci associated with each ontology term. Expression values were scaled per TSS (Z-score).

Gene Ontology Term	Fold Enrichment	FDR-Corrected P-Value
Upregulated Upon Differentiation		
isopentenyl diphosphate biosynthetic process (GO:0009240)	36.73	1.56E-02
isopentenyl diphosphate metabolic process (GO:0046490)	36.73	1.55E-02
ribosomal small subunit export from nucleus (GO:0000056)	26.23	2.99E-02
cholesterol biosynthetic process (GO:0006695)	17.49	1.55E-08
secondary alcohol biosynthetic process (GO:1902653)	17.08	1.87E-08
SRP-dependent cotranslational protein targeting to membrane (GO:0006614)	17.04	2.45E-19
coenzyme A metabolic process (GO:0015936)	16.32	1.52E-02
cotranslational protein targeting to membrane (GO:0006613)	16.2	5.64E-19
ribosomal small subunit assembly (GO:0000028)	15.3	3.80E-03
sterol biosynthetic process (GO:0016126)	15.3	4.86E-08
protein targeting to ER (GO:0045047)	15.02	2.04E-18
viral transcription (GO:0019083)	14.9	5.06E-19
establishment of protein localization to endoplasmic reticulum (GO:0072599)	14.5	3.04E-18
regulation of cholesterol biosynthetic process (GO:0045540)	14.24	2.22E-06
regulation of sterol biosynthetic process (GO:0106118)	14.24	2.19E-06
viral gene expression (GO:0019080)	14.24	2.34E-19
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	13.77	7.89E-18
isoprenoid biosynthetic process (GO:0008299)	13.12	1.47E-03
translational initiation (GO:0006413)	13	2.20E-19
protein localization to endoplasmic reticulum (GO:0070972)	12.96	1.83E-18
myelin assembly (GO:0032288)	12.89	2.98E-02
nucleotide-excision repair, DNA duplex unwinding (GO:0000717)	11.13	4.41E-02
regulation of cholesterol metabolic process (GO:0090181)	10.93	1.50E-05
protein targeting to membrane (GO:0006612)	10.65	4.14E-16
glucose 6-phosphate metabolic process (GO:0051156)	10.65	4.94E-02
myelination (GO:0042552)	10.59	4.51E-10
ensheathment of neurons (GO:0007272)	10.39	5.84E-10
axon ensheathment (GO:0008366)	10.39	5.70E-10
ribosome assembly (GO:0042255)	10.05	8.16E-06
cytoplasmic translation (GO:0002181)	9.62	1.17E-05
Downregulated Upon Differentiation		
vesicle transport along actin filament (GO:0030050)	33.96	2.44E-02
regulation of protein folding (GO:1903332)	22.64	8.45E-03
activation of MAPKKK activity (GO:0000185)	20.58	1.08E-02
response to laminar fluid shear stress (GO:0034616)	18.86	2.87E-03
establishment of protein localization to mitochondrial membrane (GO:0090151)	13.32	3.33E-02
membrane protein intracellular domain proteolysis (GO:0031293)	12.58	3.83E-02
membrane protein proteolysis (GO:0033619)	10.71	1.80E-03
response to fluid shear stress (GO:0034405)	9.99	7.03E-03
multivesicular body sorting pathway (GO:0071985)	8.84	3.32E-02
positive regulation of transcription from RNA polymerase II promoter in response to stress (GO:0036003)	8.84	3.30E-02
interleukin-12-mediated signaling pathway (GO:0035722)	8.61	4.59E-03
cortical actin cytoskeleton organization (GO:0030866)	8.49	1.36E-02
cellular response to interleukin-12 (GO:0071349)	8.25	5.46E-03
response to interleukin-12 (GO:0070671)	8.08	6.06E-03
ER-nucleus signaling pathway (GO:0006984)	7.86	4.74E-02
cortical cytoskeleton organization (GO:0030865)	7.55	2.25E-02
response to epidermal growth factor (GO:0070849)	6.93	3.03E-02
response to unfolded protein (GO:0006986)	6.41	2.32E-06
endoplasmic reticulum unfolded protein response (GO:0030968)	6.23	8.34E-04
negative regulation of intracellular transport (GO:0032387)	6.19	2.13E-02
IRE1-mediated unfolded protein response (GO:0036498)	6.17	4.57E-02
cellular response to unfolded protein (GO:0034620)	6.03	2.09E-04
intracellular steroid hormone receptor signaling pathway (GO:0030518)	5.74	2.91E-02
response to topologically incorrect protein (GO:0035966)	5.6	9.46E-06
positive regulation of axonogenesis (GO:0050772)	5.33	2.01E-02
regulation of transcription from RNA polymerase II promoter in response to stress (GO:0043618)	5.23	2.69E-03
cellular response to topologically incorrect protein (GO:0035967)	5.14	8.06E-04
cell redox homeostasis (GO:0045454)	5.14	4.57E-02
endosome organization (GO:0007032)	5.14	4.55E-02
Rho protein signal transduction (GO:0007266)	5.08	4.76E-02

Table 3.2 Gene ontology results for loci associated with transcription start sites that are upregulated and downregulated upon primary Schwann cell differentiation. The top 30 terms (ranked by fold enrichment) from each analysis are shown.

function at earlier developmental time points rather than in the myelinating stage.

Identification of SOX10-associated TSSs that are responsive to loss of SOX10

The presence of a SOX10 ChIP-Seq signal near a TSS does not necessarily indicate that SOX10 is required for expression of the TSS[259-261]. To prioritize the subset of 4,993 SOX10-associated TSSs (see above) toward identifying *bona fide* SOX10-dependent target transcripts, I measured the expression of each TSS upon ablation of SOX10. Here, I employed an immortalized rat myelinating Schwann cell line (S16 cells)[262] that expresses many markers of myelinating Schwann cells[263]. I generated S16 cells that lack SOX10 (ΔSOX10 S16 cells) via CRISPR/Cas9-mediated genome editing[264] with guide RNAs designed in the first coding exon of the rat *Sox10* locus (Figure 3.4A). The delivery of guide RNAs and Cas9 induced frame-shifting indels at the *Sox10* locus, and I recovered four clonal cell lines derived from 2 independent guide RNAs that each exhibit little or no *Sox10* transcript expression via RT-PCR (Figure 3.4B) and no detectable SOX10 protein expression via Western blot (Figure 3.4C). Therefore, I established a tractable model to test the effects of SOX10 ablation in a Schwann cell-like context *in vitro*.

To assay the effect of SOX10 ablation on the expression of candidate SOX10-dependent TSSs, Tn5Prime libraries were generated using RNA collected from ΔSOX10 S16 clones and compared to those prepared from unmodified, parental S16 cells. For the purposes of this study I focused my analysis on the 4,993 candidate TSSs associated with SOX10-bound promoters *in vivo*. These data revealed that 265 candidate TSSs were downregulated in ΔSOX10 S16 cells compared to controls, consistent with activation by SOX10 (Figure 3.5A). Of the remaining TSSs, 100 were upregulated with loss of SOX10, 3,766 were unchanged, and 862 TSSs were not expressed in the S16 model (Figure 3.5A). Similar to the primary Schwann cell model (Figure
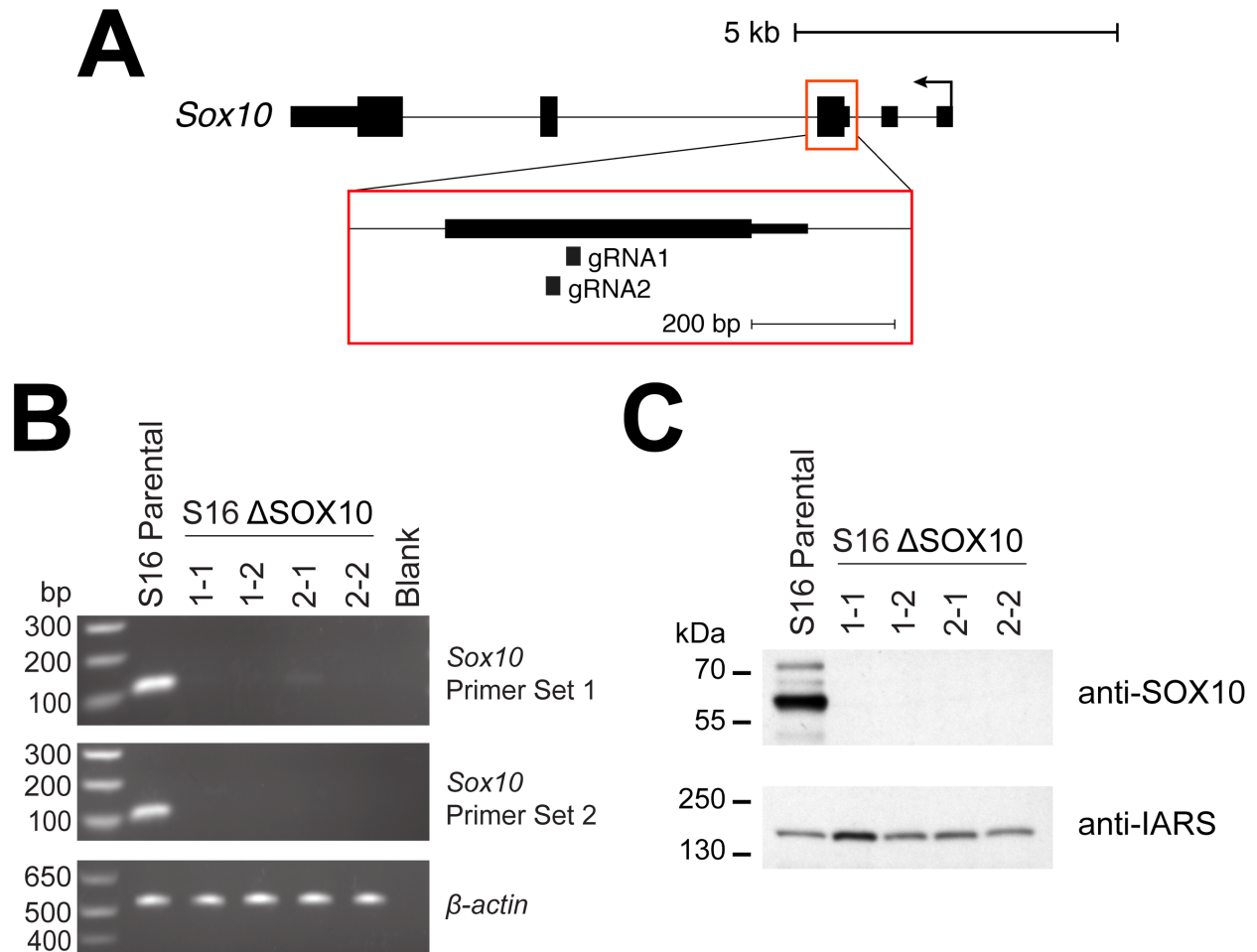
Figure 3.4 Generation of ΔSOX10 S16 cell model. (A) Guide RNAs were designed against the first coding exon of the rat *Sox10* locus to induce indel formation less than 300 bases downstream of the ATG start codon. (B) RT-PCR was performed to assay *Sox10* transcript expression using RNA isolated from unmodified, parental S16 cells and each individual ΔSOX10 S16 clone. Two independent primer sets for *Sox10* were used, along with primers for *Actb* as a positive control. Blank reactions (no cDNA) were included for each primer pair. Sizes of DNA ladder are shown to the left in base pairs (bp). (C) Western blot analysis was performed to assay SOX10 protein expression using protein lysates from unmodified, parental S16 cells and each individual ΔSOX10 S16 clone. Primary antibodies included anti-SOX10 and anti-IARS (loading control). Dashed lines indicate positions of protein ladder fragments with sizes labeled in kiloDaltons (kDa).

3.3B), the candidate TSSs that were not expressed in the S16 cell model included those that were most lowly expressed *in vivo,* again with greater than 80% detected at less than 5 reads per million in nerve (Figure 3.5B). Overall the data from this model support the identification of 265 *bona fide* TSSs that are dependent on SOX10 in Schwann cells and that are regulated in a promoter-proximal manner.

To assess the functional roles associated with the SOX10-dependent TSSs I performed a gene ontology enrichment analysis for this group. The downregulated TSSs, which map to 169 unique genes, reside at loci associated with myelination and glial differentiation, as is expected for SOX10 target genes (Figure 3.5C and Table 3.3). Importantly, the enrichment for known SOX10-regulated pathways supports the validity of the model; in contrast, the above terms are not enriched among genes associated with upregulated or unchanged TSSs (data not shown). Interestingly, the 169 genes associated with SOX10-dependent TSSs are additionally enriched for functions related to cell adhesion, extracellular structure, cytoskeleton organization, and cellular projection (Figure 3.5C and Table 3.3). Knockout of SOX10 in developing Schwann cells *in vivo* has been reported to induce excessive deposition of extracellular matrix components[147] and SOX10-ablated primary Schwann cells exhibit greater cell adhesion *in vitro*[265], but the transcriptional mechanisms and target genes mediating these effects were not investigated. Therefore, further study of the SOX10-regulated TSSs identified at genes in these functional groups will likely provide further insight into these processes in Schwann cells.

SOX10-regulated TSSs are associated with high-affinity SOX10 binding at conserved motifs

As noted above, the Tn5Prime data from the ΔSOX10 S16 cells identified 265 TSSs as high-confidence transcripts that are dependent on SOX10 and that are likely regulated by a SOX10-responsive promoter. Given that 4,993 TSSs reside near marks of SOX10-bound
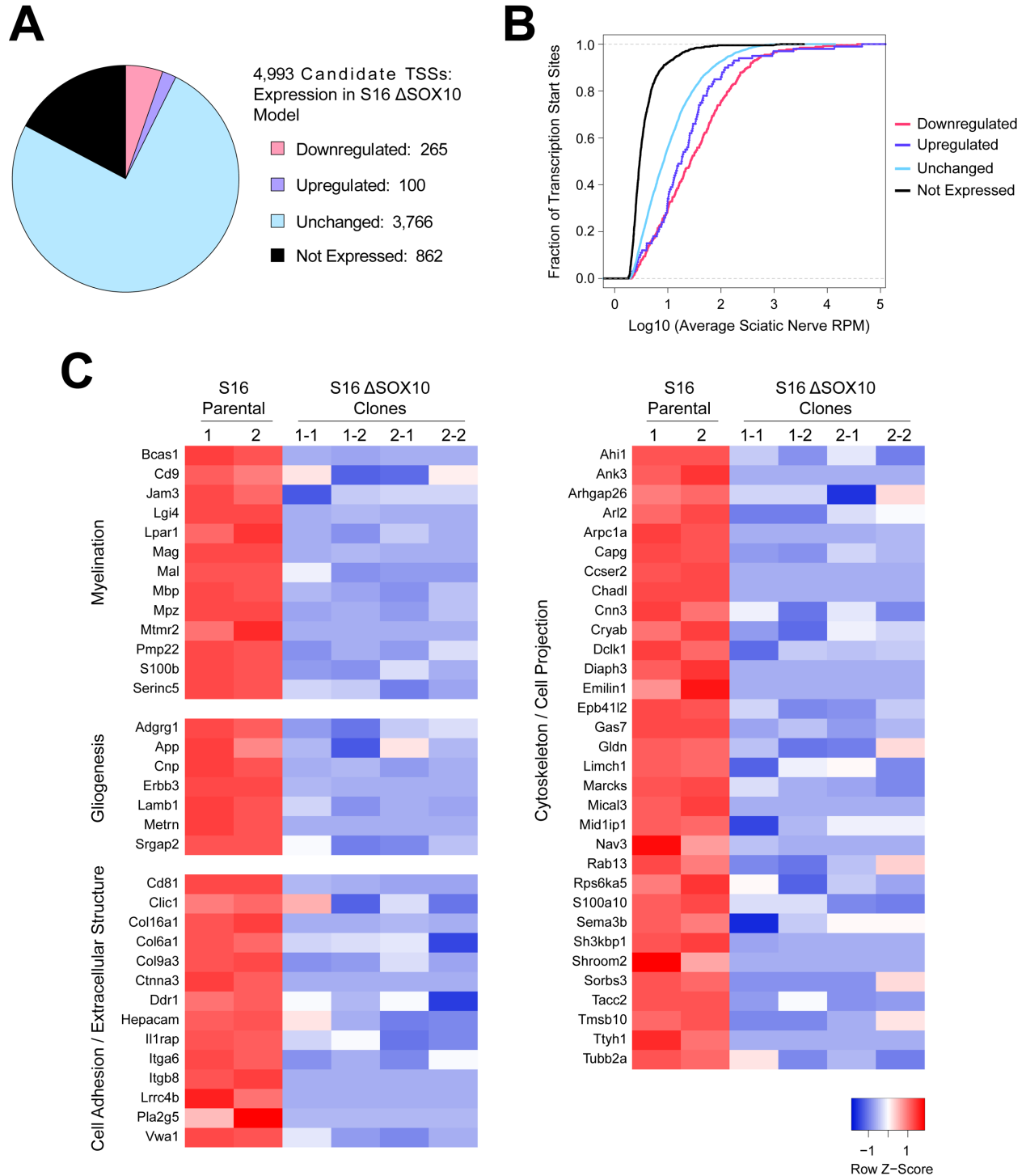
Figure 3.5 Assessment of transcription start sites associated with SOX10-bound promoters in ΔSOX10 S16 cells. (A) Expression of transcription start sites (TSSs) associated with SOX10-bound promoters in sciatic nerve were assessed for expression in ΔSOX10 S16 cells compared to unmodified S16s. TSSs were grouped by those downregulated by loss of SOX10 (pink), upregulated by loss of SOX10 (purple), unchanged (blue), or not expressed in either condition (black). (B) *In vivo* expression levels of TSSs belonging to each group as shown in panel A were

compared using empirical cumulative distribution functions. The x-axis indicates the averaged expression level (reads per million, RPM) across the two sciatic nerve libraries, log10 transformed. The y-axis indicates the cumulative fraction of TSSs. Comparisons between groups were performed using the Mann-Whitney U Test. Every comparison was found to be statistically significant ($p < 0.05$). (C) Gene ontology analysis was completed on loci associated with downregulated TSSs. Expression levels of TSSs associated with the resulting terms were compiled into a heatmap to visualize consistency of expression difference across replicates and identify gene loci associated with each ontology term. Expression values were scaled per TSS (Z-score).

Gene Ontology Term	Fold Enrichment	FDR-Corrected P-Value
Downregulated with Loss of SOX10		
regulation of myelination (GO:0031641)	17.54	1.41E-02
myelination (GO:0042552)	15.35	1.86E-07
ensheathment of neurons (GO:0007272)	15.06	1.16E-07
axon ensheathment (GO:0008366)	15.06	7.73E-08
substantia nigra development (GO:0021762)	13.35	3.88E-02
gliogenesis (GO:0042063)	7.53	1.51E-04
glial cell differentiation (GO:0010001)	7.44	2.38E-03
regulation of supramolecular fiber organization (GO:1902903)	4.84	2.53E-03
extracellular structure organization (GO:0043062)	3.87	4.98E-02
biological adhesion (GO:0022610)	3.19	1.60E-03
cell adhesion (GO:0007155)	3.07	2.72E-03
neuron development (GO:0048666)	2.9	2.82E-02
regulation of cellular component movement (GO:0051270)	2.79	1.48E-02
cytoskeleton organization (GO:0007010)	2.65	1.81E-02
plasma membrane bounded cell projection organization (GO:0120036)	2.5	4.96E-02

Table 3.3 Gene ontology results for loci associated with transcription start sites that are downregulated in ΔSOX10 S16 cells. The top 15 terms (ranked by fold enrichment) are shown.

promoter elements *in vivo,* it is notable that only 265 TSSs (5%) exhibit the expected

downregulation upon loss of SOX10 *in vitro*. To define the distinctions between TSSs that are

SOX10 dependent and those that are not, I first asked whether there are differences in the

strength of interactions between SOX10 and promoter sequences at these elements. To address

this I performed a metagene analysis using sciatic nerve SOX10 ChIP-Seq data[154, 250], this

time assessing the aggregate signal in a 2-kilobase window surrounding TSSs that were

downregulated, upregulated, and unchanged in the S16 ΔSOX10 model. This revealed that

SOX10-dependent TSSs are associated with higher affinity for SOX10 compared to upregulated

or unchanged TSSs, as reflected by an increase in SOX10 ChIP-Seq signal (Figure 3.6A).

To investigate the mechanism of increased SOX10 ChIP-Seq signal at SOX10-dependent

TSSs, we computationally tested for differences in the presence and/or composition of SOX10

binding motifs associated with these TSSs. Using an inclusive algorithm (see methods) that

allowed a one-base mismatch from a high-confidence SOX10 binding motif[154, 236], the DNA

sequence underlying each SOX10 ChIP-Seq peak was assessed for predicted SOX10 binding

sites. SOX10 ChIP-Seq peaks associated with SOX10-dependent TSSs exhibited a modest

increase in the number of SOX10 binding motifs when compared to those associated with

upregulated or unchanged TSSs; this was the case for both monomeric (Figure 3.6B) and dimeric

(Figure 3.6C) predicted motifs. However, I found no differences in the sequence composition of

the motifs associated with each group (data not shown).

Evolutionary sequence conservation can serve as a measure of functional significance of

non-coding gene regulatory sequences[207]. Therefore, I analyzed the conservation of the

SOX10 binding motifs identified above using phastCons evolutionary conservation scores[266]

across 13 vertebrate species. This revealed that SOX10 binding motifs associated with SOX10-
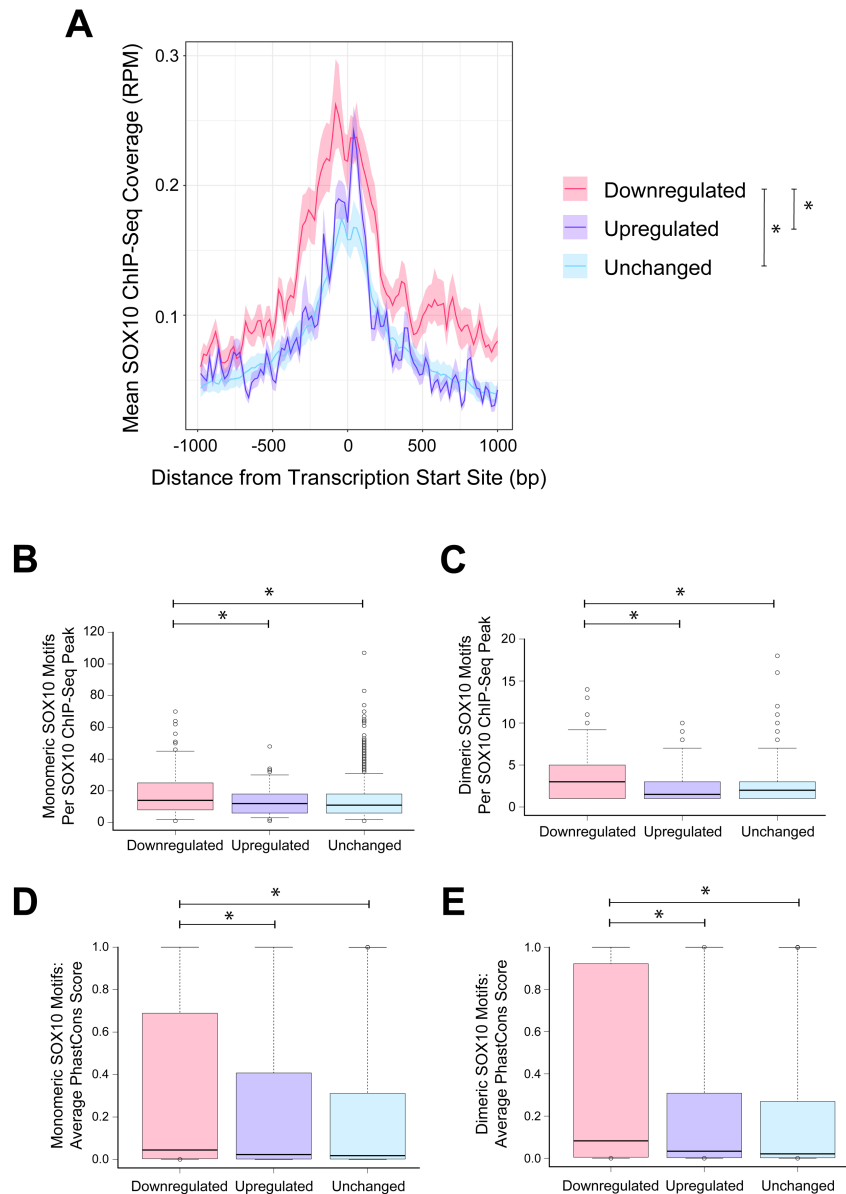
Figure 3.6 SOX10-dependent transcription start sites are associated with high-affinity SOX10 binding at conserved motifs. (A) Aggregate SOX10 ChIP-Seq data were analyzed in the 2 kilobase region surrounding TSSs that were downregulated, upregulated, or unchanged in ΔSOX10 S16 cells as shown in Figure 3.5A. The y-axis indicates average SOX10 ChIP-Seq signal (RPM, reads per million) at each class of TSS. The x-axis indicates genomic distance from the TSS, which is centered at position 0 (bp, base pairs). Regions surrounding negative-strand TSSs were reversed to orient all regions toward transcription running left to right. Comparisons between groups were performed using the area under the curve and permuatation-based analysis; each comparison was found to be statistically significant (asterisks, $p<0.001$) with the exception of the upregulated TSSs compared to those that were unchanged ($p=0.35$). (B and C) Genomic sequences underlying SOX10 ChIP-Seq peaks associated with each class of TSS were computationally analyzed for the presence of monomeric (B) or dimeric (C) SOX10 binding motifs. Boxplots show the distribution of number of motifs per peak (y-axis). Whiskers extend to

the 5th and 95th percentile of the data. Statistical analysis was done using the Mann-Whitney U test. (D and E) SOX10 binding motifs that were computationally identified and quantified in panels B and C were assessed for conservation using base-wise phastCons scores that were averaged across the bases comprising the motif. Boxplots show the distribution of scores (y-axis). Whiskers extend to the 5th and 95th percentile of the data. Statistical analysis was done using the Mann-Whitney U test. In all panels, asterisks indicate statistically significant differences ($p < 0.01$).

the 5th and 95th percentile of the data. Statistical analysis was done using the Mann-Whitney U test. (D and E) SOX10 binding motifs that were computationally identified and quantified in panels B and C were assessed for conservation using base-wise phastCons scores that were averaged across the bases comprising the motif. Boxplots show the distribution of scores (y-axis). Whiskers extend to the 5th and 95th percentile of the data. Statistical analysis was done using the Mann-Whitney U test. In all panels, asterisks indicate statistically significant differences ($p < 0.01$).

dependent TSSs had higher conservation scores compared to motifs associated with TSSs that were not dependent on SOX10 for expression (Figure 3.6D and E). Importantly, this shift toward higher sequence conservation supports the idea that these motifs carry greater functional importance. As a whole, the increased SOX10 ChIP-Seq signal, increased number of SOX10 binding motifs, and higher conservation scores for motifs associated with SOX10-dependent TSSs are consistent with the regulation of these TSSs by *bona fide* SOX10-responsive promoters.

SOX10-regulated TSSs are associated with characteristics of cell type-specific regulation

In addition to SOX10 binding sequences, other features of promoter elements could explain the differences between SOX10-bound TSSs that are downregulated (265), upregulated (100), or unchanged (3,766) upon SOX10 deletion. Algorithm-based sequence analysis (see methods) revealed no enrichment for TATA box[267] or initiator motifs[11] at SOX10-dependent TSSs (data not shown) and these TSSs exhibit no difference in GC content (Figure 3.7A). However, the 265 SOX10-dependent TSSs are depleted of CpG islands compared to the other two groups. Specifically, only 35% of the 265 SOX10-dependent TSSs reside within a CpG island, while ~70% of the TSSs that were upregulated or unchanged in ΔSOX10 S16 are associated with a CpG island at the TSS (Figure 3.7B); this latter value is consistent with the genome-wide assessments of CpG island content[14]. CpG islands contribute to nucleosome depletion at promoter regions[268] and are more frequently present at the promoters of ubiquitously expressed transcripts[269]. Thus, the depletion of CpG islands at SOX10-responsive promoters suggests that the downregulated TSSs may be enriched for transcripts that exhibit a restricted expression pattern; indeed, this is consistent with what is known about SOX10 as an activator of cell type-specific myelin genes in Schwann cells. To address this

possibility we analyzed the expression of TSSs that were downregulated, upregulated, or unchanged in the ΔSOX10 S16 cells across 11 mouse tissues using publicly available CAGE datasets[256]. All TSSs in each group were assigned a tissue specificity (Tau) score[257] based on the breadth and strength of expression in the 11 tissues, where a score of 0 indicates ubiquitous expression and 1 corresponds to highly restricted (*e.g.*, tissue-specific) expression. Interestingly, the 265 SOX10-dependent TSSs exhibit a distribution that is shifted toward greater Tau scores compared to the TSSs that were upregulated and unchanged upon loss of SOX10 (Figure 3.8A), indicating more restricted expression patterns for the SOX10-dependent TSSs. Moreover, these analyses revealed that the expression of SOX10-dependent TSSs is generally highest in the spinal cord, skin, and cortex (Figure 3.8B). It is notable that each of these three tissues harbor SOX10-positive cell types including myelinating oligodendrocytes in the central nervous system, and Schwann cells and melanocytes in the skin, suggesting that TSSs identified in this analysis are relevant in other SOX10-positive cell types.

Developmental and isoform-specific targets of SOX10 in Schwann cells

By virtue of the data collected from the three models studied here, we have identified transcripts that may play important roles in peripheral myelination based on: (***i***) expression and association with SOX10-bound promoter marks *in vivo*; (***ii***) expression in differentiating primary Schwann cells *in vitro*; and (***iii***) SOX10-dependent expression in a knockout cell model. As SOX10 is known to induce the expression of developmentally-regulated genes in the Schwann cell lineage, I next asked to what extent SOX10-dependent TSSs in the S16 model exhibit regulated expression during primary Schwann cell differentiation. A simple comparison of these datasets revealed that of the 265 TSSs expressed in sciatic nerve and downregulated in the ΔSOX10 S16 model, 132 (50%) were upregulated upon primary cell differentiation, supporting
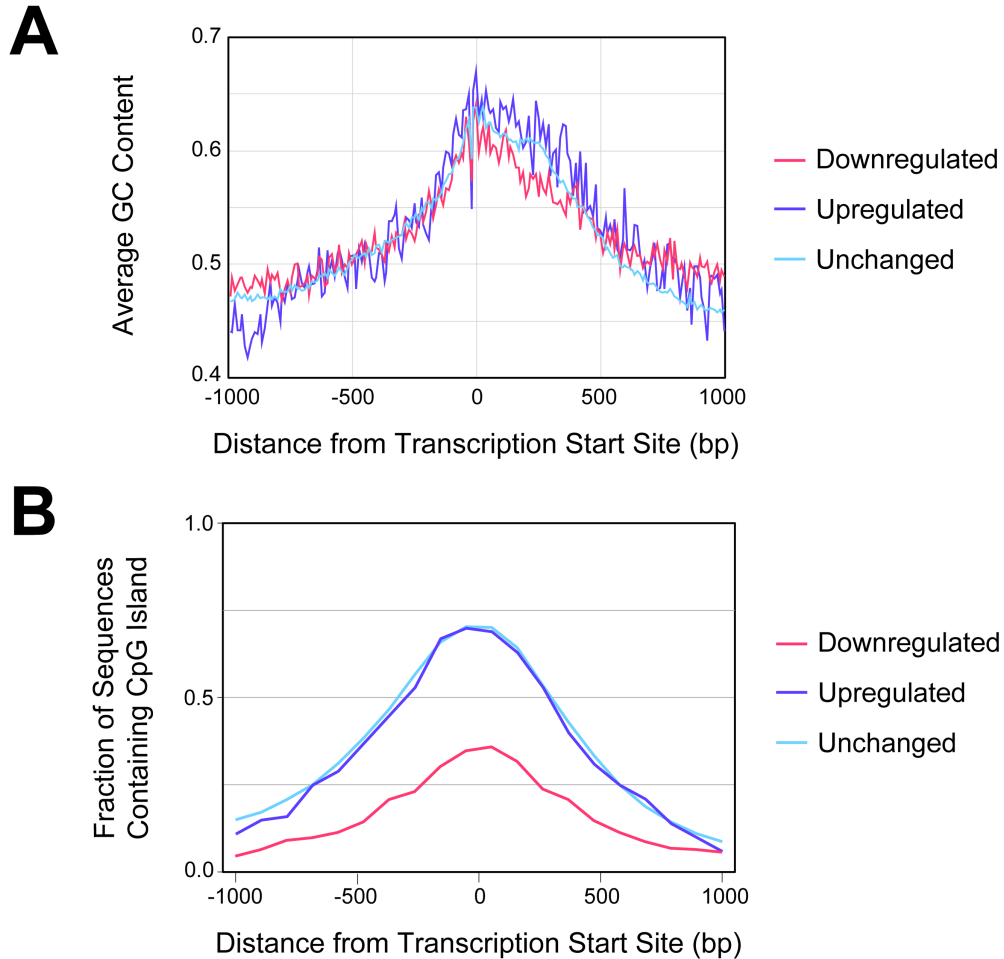
Figure 3.7 SOX10-dependent transcription start sites are depleted of CpG islands. (A) 2 kilobase regions surrounding transcription start sites (TSSs) that were downregulated, upregulated, or unchanged in ΔSOX10 S16 cells as shown in Figure 3.5A were computationally assessed for GC content in 10 base pair bins and the per-bin averages (y-axis) were plotted as a function of distance from the TSS (x-axis), which is positioned at 0. (B) CpG islands were identified computationally in the 2 kilobase regions surrounding the downregulated, upregulated, and unchanged TSSs in the ΔSOX10 S16 model. The fraction of regions harboring a CpG island (y-axis) was calculated in 100 base pair bins and plotted as a function of distance from the TSS (x-axis), which is positioned at 0.
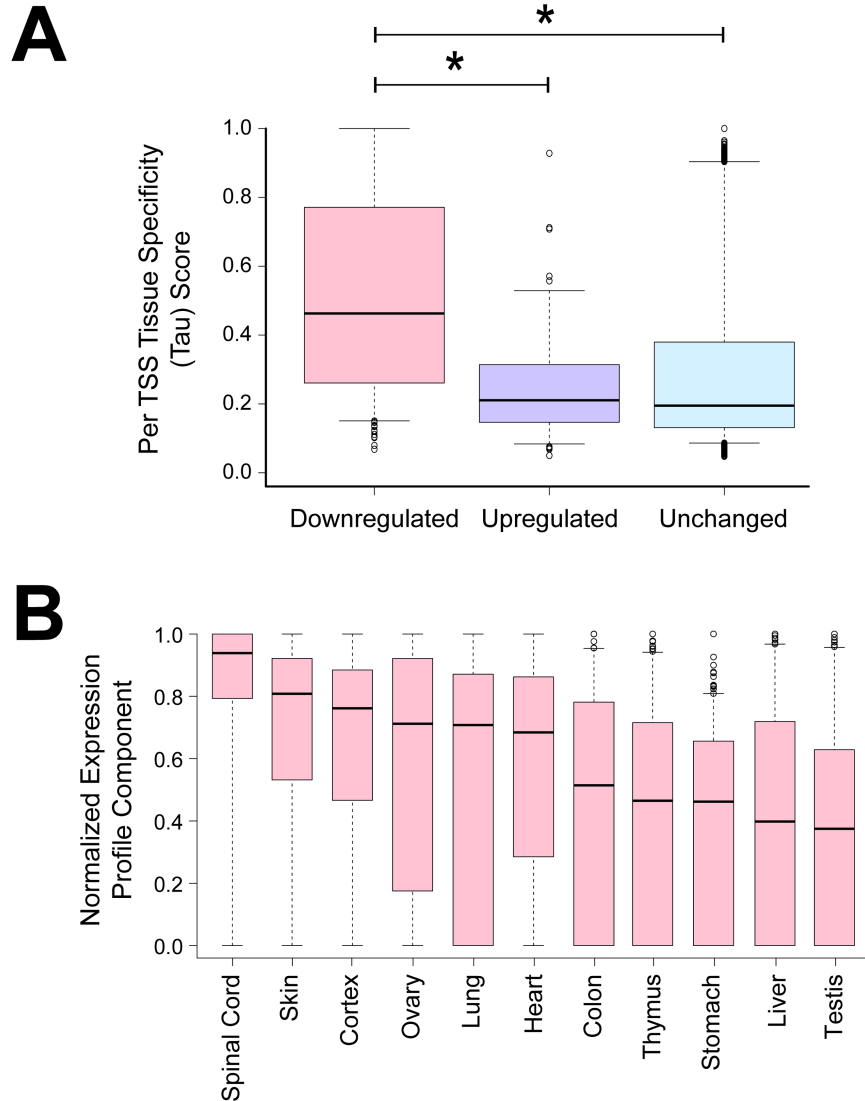
Figure 3.8 SOX10-dependent transcription start sites exhibit restricted expression profiles consistent that are consistent with expression in SOX10-positive cell types. (A) Expression of transcription start sites that were downregulated, upregulated, and unchanged in the ΔSOX10 S16 model were analyzed across 11 mouse tissues using publicly available CAGE data. For each TSS, a Tau score was calculated as described by Yanai and colleagues[257] to quantify restricted (Tau=1) or ubiquitous (Tau=0) expression. Boxplots show the distribution of Tau scores (y-axis) per class of TSS. Whiskers extend to the 5th and 95th percentile of data. Statistical analysis was done using the Mann-Whitney U test (asterisks, p<0.001). (B) To compute Tau scores, a normalized expression profile component is calculated per tissue for each TSS. The normalized expression profile component score is 1 for the tissue where the TSS is most highly expressed. Expression values in other tissues are normalized to this value. Per-tissue distributions of normalized expression profile components are displayed for the SOX10-dependent TSSs by boxplots. Whiskers extend to the 5th and 95th percentile of the data.

that these SOX10-regulated transcripts are relevant for actively myelinating cells. Moreover, 24 SOX10-dependent TSSs (9%) were downregulated with differentiation, indicating that these SOX10 targets may function during earlier developmental stages. Finally, 87 SOX10-dependent TSSs (33%) were unchanged during differentiation, suggesting that these SOX10 target transcripts may be important at multiple stages of Schwann cell development, and 22 (8%) were not expressed in primary Schwann cells.

As noted above, TSS-specific analyses mediate a more nuanced understanding of gene expression in an isoform-specific manner. Indeed, the loci associated with SOX10-dependent TSSs in these studies may harbor multiple TSSs, and this knowledge will drive a better understanding of isoform-specific gene function in Schwann cells. Therefore, we next assessed the complexity of annotated transcripts in the human RefSeq database for each of the loci containing SOX10-dependent TSSs [270]. The 265 SOX10-dependent TSSs identified in this analysis map to 169 unique gene loci, and 76 of these loci (45%) are annotated with multiple TSSs in the human RefSeq database. At 55 loci (32%) the alternative TSSs confer unique protein coding sequences to the resulting transcripts. Therefore, the identification of SOX10-regulated TSSs at a substantial fraction of these loci will, with locus-specific follow-up studies, likely provide insight into specific transcript and protein isoforms that play important roles in Schwann cell function.

Discussion

In this chapter, I presented our efforts to identify SOX10-regulated promoters in Schwann cells genome-wide. We leveraged existing datasets defining SOX10 binding and active promoter elements in sciatic nerve[154, 248] with global TSS mapping in the same tissue to catalog TSSs associated with SOX10-bound promoters. Importantly, this represents a

comprehensive and unbiased analysis using datasets that are each derived from the same *in vivo* context, and resulted in the identification of nearly 5,000 candidate TSSs that may be expressed from SOX10-regulated promoters (Figure 3.1A). Aggregate analysis of SOX10 ChIP-Seq data surrounding TSSs that are and are not associated with SOX10-bound promoters in sciatic nerve confirmed enrichment of SOX10 binding directly over the TSS and confirmed greater intensity signal at the 4,993 TSSs associated with called SOX10 ChIP-Seq peaks (Figure 3.1C). However, it is interesting that SOX10 binding is concentrated over the TSS even at elements that are associated with an H3K4me3-marked promoter element but are not associated with a SOX10 ChIP-Seq peak. This signal likely reflects relatively transient and/or weak interactions with these largely accessible H3K4me3-enriched regions. However, whether there could be any functional significance to this signal is unclear.

Although the Tn5Prime data collected from sciatic nerve provide important insight into TSS use in this tissue, these data have a number of important limitations with regard to a comprehensive understanding of SOX10-regulated TSS use in Schwann cells. First, sciatic nerve is a heterogeneous tissue composed of many cell types besides Schwann cells including axons, perineurial cells, and fibroblasts. Therefore, the cellular origin of transcripts detected in this tissue are unclear. Furthermore, these data describe TSS use strictly in adult sciatic nerve, providing only a snapshot of gene expression without an understanding of how expression changes developmentally. Because SOX10 is expressed in each stage of the Schwann cell lineage, understanding how the expression of a target gene changes as cells differentiate is of particular interest. To address these limitations, I performed an *in vitro* differentiation assay using primary Schwann cells and quantified global TSS use in differentiated cells compared to control. These data not only provide evidence for Schwann cell-specific expression of TSSs

detected in sciatic nerve, but also serve as a model for developmental expression changes. Thus, the 465 TSSs that exhibit increased expression in differentiated cells are likely to play important roles in the later stages of Schwann cell development, while the 401 with reduced expression in the differentiated state may be more important in early stages of the lineage (Figure 3.3A).

Another important limitation of the association of TSSs in sciatic nerve with nearby SOX10 and H3K4me3 ChIP-Seq peaks is that the presence of a SOX10 ChIP-Seq signal does not necessarily reflect functional regulation or strictly SOX10-dependent expression. To address this limitation, we turned to the S16 model of myelinating Schwann cells, wherein I ablated SOX10 expression and assayed the associated changes in TSS expression. These studies identified 265 TSSs that are associated with SOX10-bound promoters *in vivo* and that also exhibit reduced expression with loss of SOX10 in S16 cells (Figure 3.5A), consistent with dependence on SOX10. Our finding that a very small minority (5%) of the TSSs that are associated with a SOX10-bound promoter element *in vivo* exhibit decreased expression upon the loss of SOX10 *in vitro* could suggest that SOX10 activity in S16 cells does not faithfully recapitulate its activity *in vivo*. However, this finding is consistent with other studies showing that a minority of targets near ChIP-Seq peaks for a factor of interest exhibit altered expression with modulation of the factor[259-261]. Thus, it may be that once the gene expression profile of a fully differentiated myelinating cell has been established (as modeled by high myelin gene expression in unmodified S16 cells[263]) it is relatively stable, and though SOX10 binds widely throughout the genome it is not required for maintained expression at a majority of loci. It is notable that the loss of SOX10 in fully mature myelinating Schwann cells *in vivo* is known to induce demyelination and loss of the differentiated state[149]. Our data show that myelin genes *Mbp*, *Mpz*, and *Pmp22* are among the severely downregulated genes in ΔSOX10 S16 cells, so it

may also be that misregulated expression of a few critical transcripts is sufficient to radically affect the myelinating phenotype of differentiated cells upon loss of SOX10.

Our analyses of characteristics that distinguish *bona fide* SOX10-dependent TSSs from those that are unaltered by loss of SOX10 revealed an increase of SOX10 ChIP-Seq signal intensity, a modest enrichment of SOX10 binding motifs, but a substantial increase in the evolutionary conservation of SOX10 motif sequences associated with SOX10-dependent promoters (Figure 3.6A-E). These lines of evidence provide support for the functional relevance of SOX10 binding at these elements and support the utility of sequence conservation-based prioritization of SOX10 binding motifs[169]. Moreover, we found that CpG islands are depleted at SOX10-dependent TSSs and they exhibit restricted expression profiles across a variety of mouse tissues (Figures 3.7 and 3.8). Indeed, the SOX10-dependent TSSs exhibited highest expression in the central nervous system and skin samples from mouse (Figure 3.8B), each of which harbor SOX10-positive cell types including oligodendrocytes, Schwann cells, and melanocytes. These data are consistent with the identification of these TSSs as targets of a cell type-specific transcription factor and suggest that further study of the associated transcripts is likely to provide new insights into the biology of Schwann cells as well as other SOX10-positive cell types.

Interestingly, nearly half of the loci that we defined as harboring a SOX10-regulated promoter have been annotated with multiple TSSs in the human genome, and a third of the genes have alternative TSSs that confer protein-coding changes to the resulting transcripts. Therefore, these studies support the importance of SOX10-mediated isoform-specific gene expression at a number of loci in Schwann cells. We suggest that functional studies at these loci should be designed with isoform-specific biology in mind, as these findings may reflect a particular

physiological relevance of the SOX10-regulated isoforms for peripheral myelination. Indeed, I will explore these concepts at four loci of interest in the following chapters of this dissertation.

Finally, it is worth noting that the TSS expression data described in this chapter will serve as an important resource for understanding gene expression in Schwann cells. The analyses presented here were tailored and limited to the analysis of SOX10-bound promoters. However, broader analyses will be meaningful toward understanding promoter use in Schwann cells that is mediated by other factors and mechanisms.

Chapter 4

A SOX10-Regulated Promoter and Novel Transcription

Start Site at *ARPC1A* in Schwann Cells

Introduction

The *ARPC1A* locus encodes subunit 1A of the actin related protein 2/3 (Arp2/3) complex; this complex is critical for the polymerization, organization, and recycling of the actin cytoskeleton[271]. Indeed, the Arp2/3 complex catalyzes the formation of branched actin networks and has been implicated in a variety of cellular processes including intracellular trafficking, cell migration, endocytosis, phagocytosis, and the formation of cellular projections[272]. Importantly, Arp2/3 is not active without the influence of nucleation promoting factors (NPFs) such as WASP, N-WASP, SCAR/WAVE, and WASH, which stimulate the complex by inducing a conformational change and increasing affinity between Arp2/3 and the actin mother filament[272]. This activity is mediated by a VCA domain that is common to a number of NPFs [273].

The Arp2/3 complex was originally described as comprising seven subunits, though recent work has identified diverse, 'hybrid' arrangements of the complex with various functional impacts[272]. In particular, mammalian genomes contain two ARPC1 subunits, ARPC1A and ARPC1B, which interact with the complex in a mutually-exclusive manner[274]. Moreover, some evidence suggests that the differential inclusion of ARPC1 subunits impacts actin dynamics, as ARPC1A-containing complexes are apparently less efficient actin nucleators than their ARPC1B-containing counterparts in certain contexts[275]. Based on targeted mutation

111

studies in yeast, where there is only one ARPC1 protein, this subunit is thought to perform multiple critical functions in the regulation of Arp2/3 complex activity[276]. Namely, ARPC1 contributes to the suppression of nucleation in the absence of NPFs, serves as an interaction site for the VCA domain of WASP, and transmits activation signals to the rest of the complex upon WASP binding[276]. Moreover, the critical nature of the ARPC1 subunit is underscored by the fact that this is the only Arp2/3 subunit that is essential for yeast viability[277].

Although the Arp2/3 complex is understandably critical for cellular function across all tissue types, Arp2/3 and Arp2/3 regulators have been experimentally shown to be required also for the proper function of myelinating cells in the central and peripheral nervous systems[278-281]. The deletion of N-WASP in Schwann cell precursors, for example, has little effect on radial sorting but severely reduces both the number of myelinated axons and the internode length of the few myelin sheaths that are formed[280]. Thus, it is thought that N-WASP mediates spiral membrane wrapping and the longitudinal extension of the Schwann cell as myelination begins. These effects are consistent with: (*i*) an important role for Arp2/3 activity in extending the lamellipodia-like projections of developing Schwann cells; and (*ii*) previous work describing the essential contributions of Arp2/3 activity to lamellipodia formation in other cellular contexts[274, 282].

In this chapter I discuss the identification of a SOX10-regulated and previously undescribed TSS at *Arpc1a* that was identified by the genome-wide efforts described in Chapter 3 of this dissertation. I present my work toward validating the regulation of this element by SOX10 and my efforts to characterize the *Arpc1a* gene products generated from this promoter in Schwann cells. Finally, I will discuss the implications of these findings and outstanding

questions regarding the potential role(s) of this novel gene product in peripheral nerve myelination.

Note that I completed the entirety of the work described in this chapter, with the exception of Sanger sequencing performed by the University of Michigan DNA Sequencing Core.

Materials and Methods

Generation of luciferase reporter gene constructs

Oligonucleotide primers containing attB1 and attB2 Gateway cloning sequences (Invitrogen Life Technologies, Carlsbad, CA) were designed for PCR-based amplification of *ARPC1A* Prom 2 (hg38 coordinates chr7:99,357,845-99,358,748; primer sequences available in Appendix). The region was amplified from human genomic DNA using PCR Supermix (ThermoFisher Scientific). Subsequent to PCR amplification and purification, each genomic segment was cloned into the pDONR221 vector using BP Clonase (Invitrogen). Resulting constructs were genotyped by digestion with *Bsr*GI (New England Biolabs) and subjected to DNA sequence analysis to ensure the integrity of the insert. The resulting pDONR221 construct was recombined with an expression construct (pE1B-luciferase) [197] using LR Clonase (Invitrogen) to clone each region upstream of a minimal promoter directing expression of a luciferase reporter gene. Successful recombination was confirmed via digestion of DNA with *Bsr*GI.

Site-directed mutagenesis was performed using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc., Santa Clara, CA). Mutagenesis primers were designed to delete each of the dimeric SOX10 binding sites within *ARPC1A* Prom 2. Mutagenesis was performed in pDONR221 constructs and DNA from each resulting clone

113

underwent sequence analysis to verify that only the desired mutation was produced. Verified clones were then recombined into pE1B-luciferase using LR Clonase (Invitrogen).

Cell culture, transfection, and luciferase assays

Unmodified parental S16 cells [199] and ΔSOX10 S16 cells were grown under standard conditions in DMEM with 10% fetal bovine serum, 2mM L-glutamine, 50 U/mL penicillin, and 50 g/mL streptomycin. For luciferase assays, ~1,000 cells were plated in each well of a 96-well plate. Cells were cultured for 24 hours under standard conditions prior to transfections.

Lipofectamine 2000 (ThermoFisher Scientific) was diluted 1:100 in OptiMEM I reduced serum medium (ThermoFisher Scientific) and incubated at room temperature for 10 minutes. Each DNA construct to be transfected was individually diluted in OptiMEM to a concentration of 8 ng/μL. An internal control renilla construct was added to the solution at 8 pg/μL. One volume of lipofectamine solution was added to each DNA solution and allowed to sit for 20 minutes at room temperature. Cells were incubated with transfection solution for 4 hours under standard conditions and then the medium was changed to standard growth medium.

Cells were washed with 1X PBS 48 hours after transfection and lysed for 1 hour shaking at room temperature using 20 μL 1X Passive Lysis Buffer (Promega, Madison, WI). 10 μL of lysate from each well was transferred into a white polystyrene 96-well plate (Corning Inc., Corning, NY). Luciferase and renilla activities were determined using the Dual Luciferase Reporter 1000 Assay System (Promega) and a Glomax Multi-Detection System (Promega). Each reaction was performed at least 24 times. The ratio of luciferase to renilla activity and the fold change in this ratio compared to a control luciferase expression vector with no genomic insert were calculated. The mean (bar height) and standard deviation (error bars) of each fold difference are represented in the figures.

RNA isolation and RT-PCR

RNA was isolated from rat sciatic nerve using the RNeasy Mini Kit (Qiagen USA) according to manufacturer's instructions. RNA samples were eluted into 30 µL of RNase-free water and stored at -80°C prior to experimentation. RNA concentration and purity were determined using a NanoDrop Lite (Thermo Scientific, Waltham, MA). A cDNA library was generated using 1 µg of RNA and the High-Capacity cDNA Reverse Transcription Kit (Life Technologies), with the provided random reverse-transcription primers and according to manufacturer's protocol. cDNA samples were analyzed by PCR using PCR Supermix (Life Technologies), 0.5 µL of each 20 µM primer solution, and 1 µL of cDNA. A blank (cDNA-negative) control was included and standard PCR conditions were used.

Results

SOX10 mediates expression of a novel, intronic TSS at Arpc1a

There are two *ARPC1A* RefSeq isoforms annotated in the human genome that differ based on the use of alternative splice acceptor sites upstream of exon 3 (Figure 4.1A); these transcripts encode ARPC1A protein isoforms 1 and 2, which are largely identical with the exception of isoform-specific N-terminal sequences. In addition to transcripts arising from the annotated TSS at exon 1A (data not shown), our Tn5Prime data from multiple Schwann cell models identified a TSS in the seventh intron of the rat *Arpc1a* locus (Figure 4.1B). The expression of this TSS: (***i***) is associated with SOX10 and H3K4me3 ChIP-Seq peaks in sciatic nerve; (***ii***) exhibits a 50% downregulation upon differentiation in primary Schwann cells (FDR-corrected p-value = 4.20×10^{-20}); and (***iii***) is largely abolished upon ablation of SOX10 in S16 cells (FDR-corrected p-value = 1.18×10^{-43}) (Figure 4.1B). These data support the identification

115
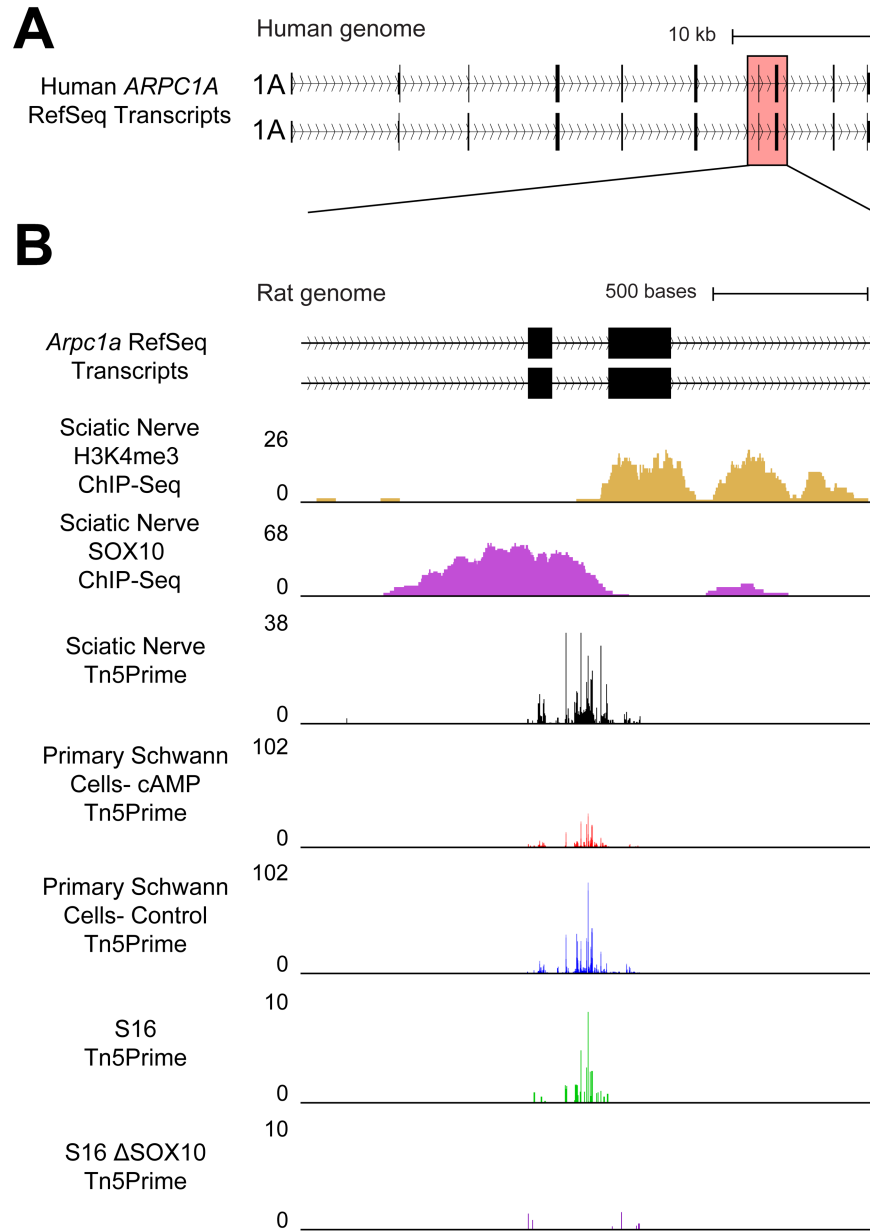
Figure 4.1 SOX10 regulates expression of a novel transcription start site at *Arpc1a*. (A) The human *ARPC1A* locus is annotated with two RefSeq transcript isoforms, both originating at exon 1A ('1A' in panel). (B) The genomic region surrounding exons 7 and 8 of the rat *Arpc1a* locus. This region is recognized by antibodies against H3K4me3 and SOX10 in sciatic nerve. Y-axes for ChIP-Seq data indicate the fold enrichment of sequencing reads above chromatin input. Tn5Prime data at this region are shown from rat sciatic nerve, CPT-cAMP- (cAMP) and vehicle-treated (Control) primary Schwann cells, and unmodified and ΔSOX10 S16 cells. Y-axes for Tn5Prime data indicate the number of transcript 5'ends mapped per base, in reads per million.

of a novel TSS and SOX10-regulated promoter at *Arpc1a* and suggest a role for this TSS (and an associated protein product, if the transcript is translated) in early Schwann cell development.

SOX10 regulates ARPC1A Prom 2 via two dimeric SOX10 binding motifs

To investigate the mechanism by which SOX10 regulates the above *Arpc1a* TSS, we analyzed the genomic sequence underlying the associated SOX10 ChIP-Seq peak and found two predicted dimeric SOX10 binding motifs that are located less than 500 bases upstream of the TSS (SOX10-1 and SOX10-2 in Figure 4.2A). These motifs show variable evolutionary conservation among mammals (SOX10-1: 1/5 and 5/5 bases conserved to opossum; SOX10-2: 2/5 and 4/5 bases conserved to opossum; Figure 4.2A). To test the regulatory activity of this region, we amplified a 904 base pair fragment (referred to as *ARPC1A* Prom 2) from the orthologous region of the human genome (see methods), cloned the element upstream of a minimal E1B promoter and firefly luciferase reporter gene[197], and tested regulatory activity using a dual luciferase reporter assay. In unmodified S16 cells where SOX10 is highly expressed, a construct harboring *ARPC1A* Prom 2 exhibits a 45-fold induction of luciferase activity relative to an empty control vector (Figure 4.2B, p-value=9×10^{-11}), consistent with strong regulatory activity of this genomic segment. To test the necessity of the predicted SOX10 binding motifs in the regulatory activity of *APRC1A* Prom 2, we mutagenized Prom 2 to delete each motif in isolation and together. Deletion of the SOX10-1 motif reduced the activity of the element by ~30% (p-value=1.9×10^{-4}) while deletion of SOX10-2 reduced activity by ~20% (p-value=0.02; Figure 4.2B). Consistent with an additive effect of these sequences, deletion of both motifs reduced the activity of Prom 2 by ~50% compared to the wild-type element (Figure 4.2B, p-value=2.7×10^{-6}). To further investigate the necessity of SOX10 for Prom 2 activity, we performed a similar assay in the ΔSOX10 S16 cell model. In this assay, the activity of wild-type

117
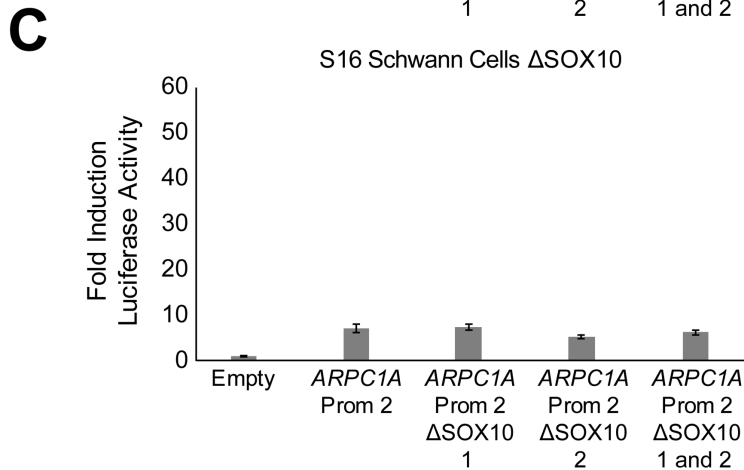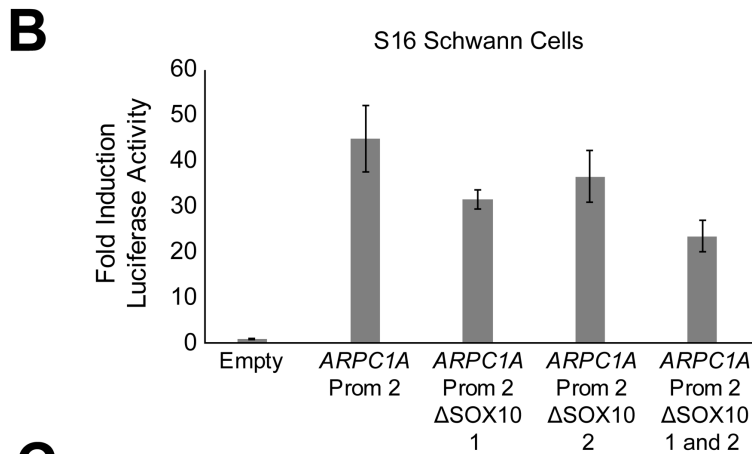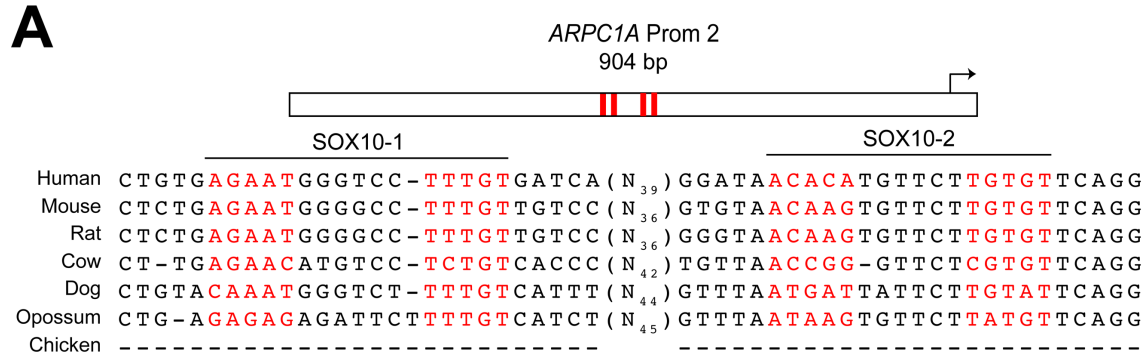
Figure 4.2 SOX10 regulates *ARPC1A* Prom 2 via two dimeric binding motifs. (A) The 904 base pair *ARPC1A* Prom 2 is shown along with the position of the two SOX10 dimeric consensus sequences (red bars and red text). The seven species utilized for comparative sequence analysis are shown on the left. (B and C) *ARPC1A* Prom 2 with or without the dimeric SOX10 sequences as indicated was cloned upstream of a luciferase reporter gene, transfected into cultured Schwann (S16) cells (B) or ΔSOX10 S16 cells (C), and tested for activity in luciferase assays compared to an empty vector containing no genomic insert. The fold induction of luciferase activity is indicated along the y axis and error bars indicate standard deviations.

ARPC1A Prom 2 was only 7-fold induced relative to the empty control (Figure 4.2C, p-value=8.6x10^{-11}). Moreover, deletion of the SOX10 binding motifs rendered no or little change to activity of the element in this context (Figure 4.2C; ΔSOX10-1, p-value=0.56; ΔSOX10-2, p-value=3.4x10^{-4}; and ΔSOX10-1 and 2, p-value=0.07). Taken together, these results support the validation of *APRC1A* Prom 2 as an active, SOX10-responsive regulatory element in Schwann cells.

Arpc1a Prom 2 directs expression of a spliced transcript isoform in sciatic nerve

Given the validation of *ARPC1A* Prom 2 activity in Schwann cells, we sought to understand more about the *ARPC1A* gene products that are generated from this promoter. To do this we first characterized the *Arpc1a* transcript expressed from Prom 2 by performing RT-PCR using cDNA generated from rat sciatic nerve RNA. For this experiment the forward primer was anchored in what we refer to as *Arpc1a* exon 1B (corresponding to the annotated *APRC1A* exon 8 with inclusion of upstream intronic sequences) and a reverse primer was designed against the *Arpc1a* 3'UTR (Figure 4.3B) This reaction produced a ~500bp fragment (Figure 4.3A). We next subjected the transcript amplified from rat sciatic nerve to Sanger sequencing. Because the rat genome assembly omits *Arpc1a* exon 9 due to a gap in the genome sequence (Figure 4.3B), we mapped the data to the mouse genome to confirm the expected transcript architecture; indeed, this confirmed that the transcript originating from *Arpc1a* Prom 2 includes exons 1B, 9, and 10 (Figure 4.3B). Based on the location of the first in-frame ATG codon, this transcript isoform is predicted to encode a 101 amino acid protein with a predicted molecular weight of 12 kDa. However, this ATG codon does not reside in a favorable Kozak context, as there is neither a purine in the -3 position nor a guanine in the +1 position (Figure 4.4). Therefore, we anticipate that the transcript may be more efficiently translated from one of multiple downstream ATG

119
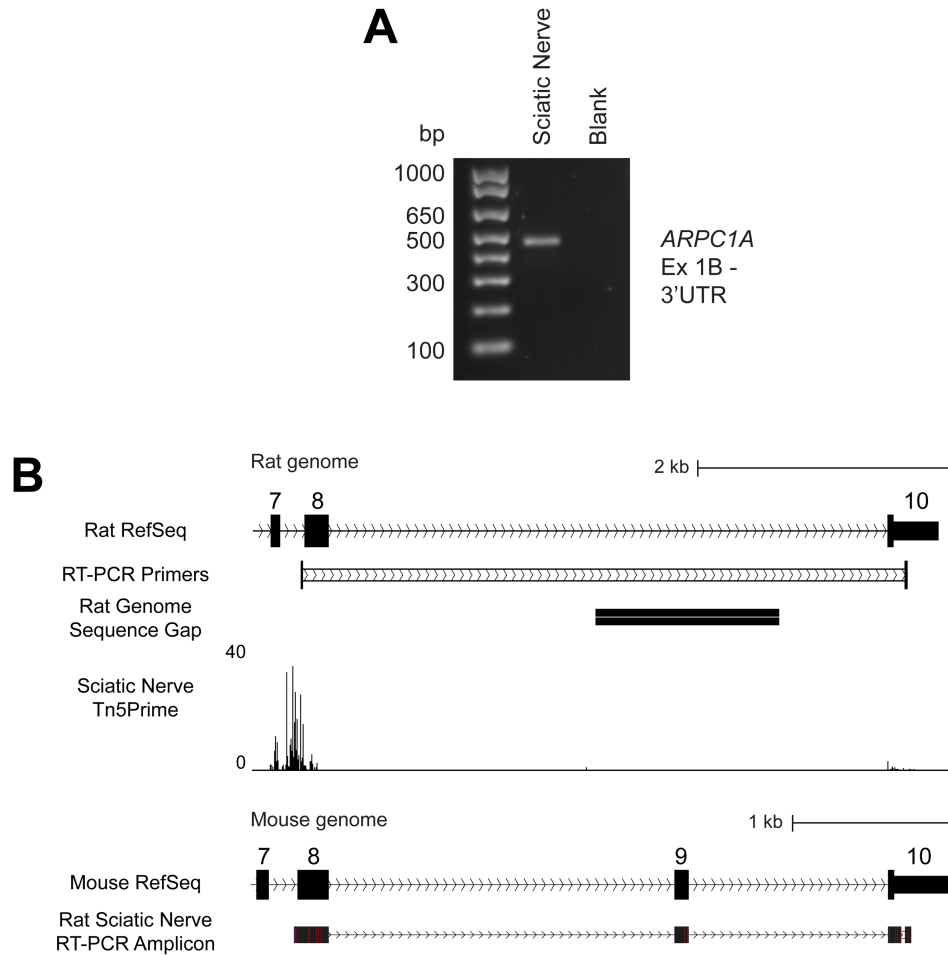
Figure 4.3 *Arpc1a* Prom 2 directs expression of a spliced *Arpc1a* transcript *in vivo*. (A) RT-PCR was used to validate the expression of a spliced *Arpc1a* transcript with the expected architecture using cDNA from rat sciatic nerve. A blank reaction (no cDNA) was used as a negative control. Sizes of DNA ladder markers are indicated to the left in base pairs (bp). (B) The rat *Arpc1a* locus is shown above, with exons 7, 8, and 10 indicated. The locations of RT-PCR primers used in panel A are shown by vertical black bars. The sequence gap in the rat genome is shown by the thick black horizontal bar. Tn5Prime data from rat sciatic nerve is shown as reference for the location of the SOX10-dependent TSS; y-axis indicates number of transcript 5' ends per base in reads per million. The mouse *Arpc1a* locus is shown below, with exons 7, 8, 9, and 10 indicated. The rat sciatic nerve-derived transcript sequence mapped to the mouse genome as shown at the bottom.
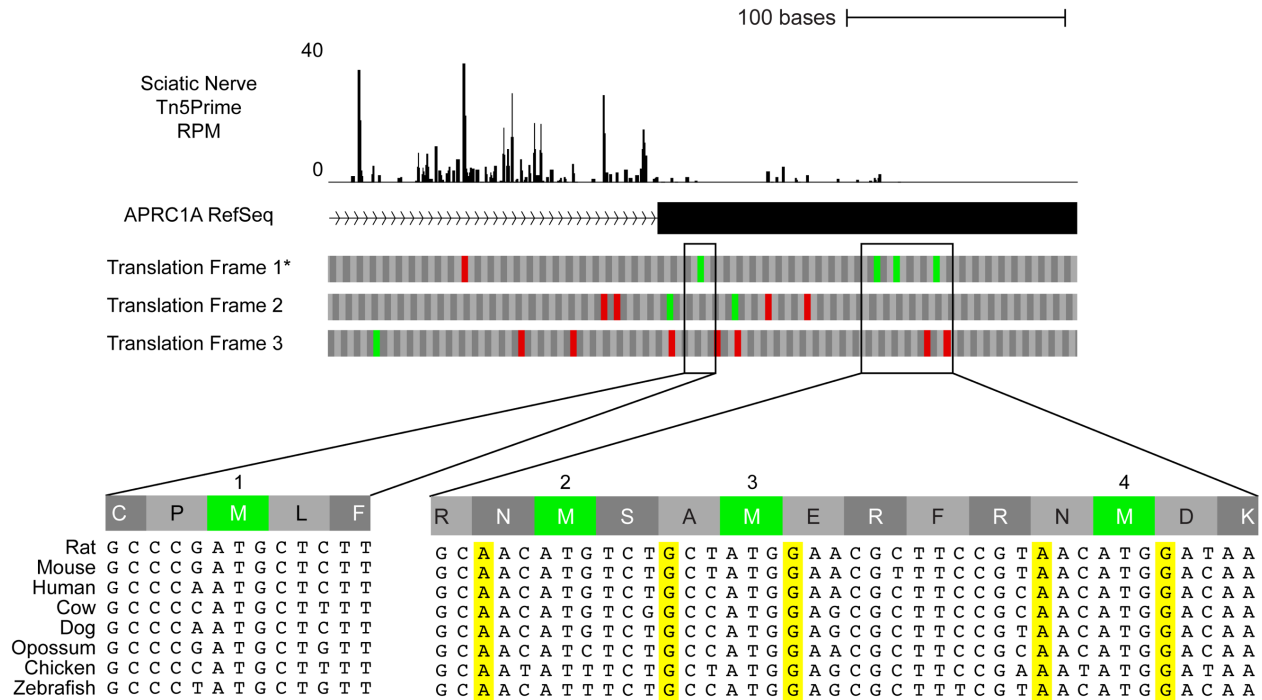
Figure 4.4 Kozak sequences surrounding in-frame ATG codons in the SOX10-regulated *Arpc1a* transcript. Sciatic nerve-derived Tn5Prime data are shown in relation to the annotated exon 8 of *Arpc1a*. The y-axis indicates number of transcript 5' ends mapped per base, in reads per million (RPM). The three possible translation frames of the transcript are shown below, with start codons indiciated in green and stop codons indicated in red. The asterisk by Translation Frame 1 indicates that this is the translation frame for full-length ARPC1A. DNA sequences surrounding each in-frame ATG codon are shown below, for four in-frame methionines (1-4). Yellow bars highlight bases that contribute to strong Kozak sequences (purine base in the -3 position and guanine in the +1 position).

codons with stronger Kozak sequences (Figure 4.4). From the latter ATG codons, we predict the

expression of a 71 or 74 amino acid protein with a molecular weight of ~8 kDa that lacks 5 of the

6 WD40 repeat regions of full-length ARPC1A (Figure 4.5). Unfortunately, our attempts to

validate the expression of a stable APRC1A protein arising from this transcript were

unsuccessful, as we were unable to identify an antibody directed against the C-terminus of the

protein that performed well in a Western blot assay. In sum, our data support the SOX10-

regulated expression of a novel *Arpc1a* transcript isoform in Schwann cells that encodes an N-

terminally truncated protein isoform.

Discussion

The *ARPC1A* gene is likely to play an important role in nearly every cellular context due

to it being a critical subunit of the Arp2/3 actin remodeling complex. Indeed, studies in yeast

confirm the essential nature of the gene and support a model wherein ARPC1 is responsible for

inducing or suppressing activity of the complex in the presence and absence of a nucleation

promoting factor, respectively[276, 277]. Here we present data supporting the identification of a

previously unreported SOX10-regulated TSS at the *Arpc1a* locus. We show that the promoter

element harbors two dimeric SOX10 binding motifs, that these motifs contribute to the

regulatory activity of the element in Schwann cells *in vitro*, and confirm the *in vivo* expression of

a spliced *Arpc1a* transcript that is predicted to encode a protein isoform that does not contain a

significant portion of the protein sequence present in longer isoforms. Based on our findings, we

speculate that this SOX10-regulated ARPC1A isoform plays a role in actin cytoskeleton

dynamics known to be important for Schwann cells[278-280].

To characterize the mechanism by which SOX10 regulates the identified *Arpc1a* TSS

(Figure 4.1B), we computationally assessed the genomic region upstream of the TSS and
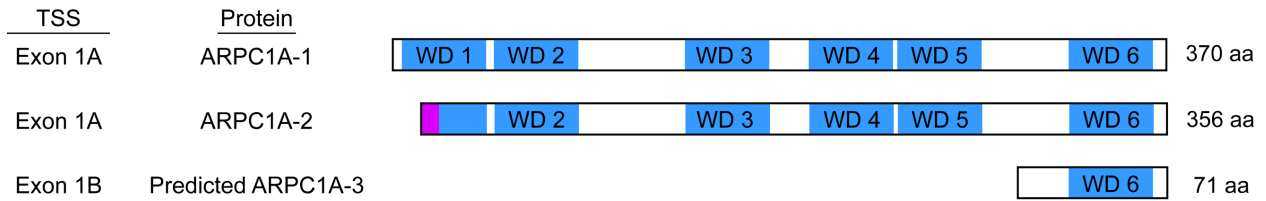
Figure 4.5 ARPC1A protein isoforms. ARPC1A isoforms 1 and 2 originate from exon 1A and are distinguished by isoform-specific N-terminal sequences (purple). ARPC1A-1 contains 6 WD40 repeats (WD, blue). The predicted protein product of the SOX10-regulated *Arpc1a* transcript includes one of these repeats.

identified two dimeric SOX10 binding motifs. Luciferase assays confirmed that these sequences contribute to the activity of the element in the presence of SOX10 (Figure 4.2B), but the activity of *ARPC1A* Prom 2 was much more severely reduced in the ΔSOX10 S16 context than it had been by deletion of the SOX10 binding motifs (Figure 4.2C). This result could be explained by the presence of additional SOX10 binding motifs at *ARPC1A* Prom 2 that were not tested by deletion and conferred residual activity in the unmodified S16 cells. However, it may also reflect additional, indirect regulation of the *ARPC1A* Prom 2 element by SOX10 via EGR2. The *EGR2* locus is regulated by SOX10 and encodes a promyelinating transcription factor[153]. EGR2 functions at regulatory elements co-occupied by SOX10[157]. Interestingly, ChIP-Seq data[154] indicate that EGR2 binds to the *Arpc1a* Prom 2 genomic region in sciatic nerve and our bulk RNA-Seq data from the ΔSOX10 S16 cells confirm that EGR2 is downregulated in this model (FDR-corrected p-value=0.02; data not shown). Thus, EGR2 may contribute to the patterns of activity exhibited by *ARPC1A* Prom 2 we observed in luciferase assays and, more broadly, may co-regulate the expression of this transcript in Schwann cells *in vivo*.

The *Arpc1a* transcript that we identified encodes a protein isoform that lacks a majority of the sequence of full-length ARPC1A protein isoforms (Figure 4.5). With the caveat that expression of this protein remains to be verified, it is worthwhile to consider the possible functions of this protein isoform in Schwann cells. The crystal structure of ARPC1 in the Arp2/3 complex has been resolved[283] and together with genetic studies of ARPC1 in yeast[277] provides important insights into ARPC1 function. The ARPC1 subunit adopts a 7-bladed β-propeller conformation and makes multiple contacts with other subunits in the complex[283]. For this reason, it seems unlikely that the protein isoform we describe here is capable of assembling into the Arp2/3 complex. However, a region of the protein near the C-terminus was

identified as extending away from the β-propeller formation and adopting an alpha-helical structure that is likely solvent-accessible and referred to as the ARPC1 'arm'[276]. Moreover, this arm region of the protein is thought to interact directly with the VCA domain of WASP (and likely other nucleation promoting factors) to mediate the stimulation of the complex[276]. Importantly, the arm region is included in the ARPC1A protein isoform predicted to arise from the SOX10-regulated transcript. Therefore, if this protein isoform is expressed and capable of interacting with WASP and other NPFs, we anticipate that it could perform a dominant-negative function by sequestering NPFs away from the intact Arp2/3 complex.

TSS expression data from differentiating Schwann cells indicate that the SOX10-regulated *Arpc1a* TSS is downregulated as cells differentiate (Figure 4.1B). Conditional deletion of N-WASP in developing Schwann cells indicates that activity of this regulator—and by extension, the Arp2/3 complex—is required as myelination commences. Considering these observations and the structural and functional lines of evidence described above, we propose that the SOX10-regulated ARPC1A protein isoform acts as a "brake" on precocious Arp2/3 activity during early Schwann cell development. At early stages of the lineage, when extensive cellular projection is not required, the SOX10-regulated ARPC1A protein may interact with NPFs to limit the stimulation of Arp2/3 nucleating activity. As cells differentiate and enter the promyelinating stage, some undefined regulatory factor(s) would then reduce activity of *ARPC1A* Prom 2 to release this brake and mediate the extension of Schwann cell projections to support myelination.

In sum, the data presented in this chapter validate the identification of a SOX10-regulated promoter element at *Arpc1a* and suggest isoform-specific functions of this gene in Schwann cells. We confirmed the expression of a processed transcript arising from this promoter *in vivo*,

and report on the expected protein product that may result. Although further study is needed to confirm the expression of this isoform and to investigate our proposed model for the function of this protein isoform, our data support a novel mechanism of transcriptional regulation at *ARPC1A* and will catalyze further consideration of this locus in Schwann cell biology.

Chapter 5

SOX10-Regulated, Isoform-Specific Expression of β-chimaerin

in Differentiating Schwann Cells

Introduction

β-chimaerin, which is encoded by the *CHN2* gene, localizes to the cell membrane and

functions as a GTPase activating protein (GAP) for the Rho family small GTPase Rac1[284].

The GTPase-promoting activity of β-chimaerin mediates the conversion of active, GTP-bound

Rac1 to the inactive, GDP-bound form; in this way β-chimaerin acts as a negative regulator of

Rac1 activity. The membrane localization of β-chimaerin is contingent upon activation by the

lipid second messenger diacylglycerol (DAG); thus, β-chimaerin activity is induced via signaling

cues from cell surface receptors and activation of phospholipase C[285]. Rac1, in turn, operates

in an array of cellular functions including cell cycle progression, cytoskeleton regulation, and

cellular motility[286].

Importantly, multiple lines of evidence implicate Rac1 in Schwann cell function. Early in

Schwann cell development, Rac1 activation contributes to the migration of Schwann cell

precursors[287]. Moreover, it appears that Rac1 must be carefully regulated during development

to mediate the establishment of Schwann cell-axon interactions, as excessively high Rac1

activity has been associated with a failure to establish these connections[288]. Multiple *in vivo*

models of Schwann cell-specific Rac1 ablation have been generated, each supporting an

important role for this GTPase in peripheral myelination. First, the conditional loss of Rac1 in

Schwann cell precursors results in a failure of radial sorting and subsequent hypomyelination; this was shown to be associated with a defect in the extension and stabilization of cellular processes by developing Schwann cells[289, 290]. Moreover, loss of Rac1 slightly later in Schwann cell development causes a delay in radial sorting but profound impairment in subsequent myelination[291]. Therefore, although *CHN2* has not previously been implicated in myelination, the β-chimaerin protein—as a regulator of Rac1 activity—is an excellent candidate for a role in Schwann cell biology.

In this chapter I describe my efforts to characterize a SOX10-regulated promoter at *Chn2* that was identified through the studies described in Chapter 3. I demonstrate the mechanism by which SOX10 regulates the *Chn2* promoter and reveal the *Chn2* transcript generated from this promoter in Schwann cells. Finally, I discuss the functional implications of these findings related to isoform-specific biology of β-chimaerin proteins and outstanding questions regarding the possible contribution of this locus to Schwann cell biology.

Note that I completed the entirety of the work described in this chapter, with the exception of Sanger sequencing performed by the University of Michigan DNA Sequencing Core.

Materials and Methods

Generation of luciferase reporter gene constructs

Oligonucleotide primers containing attB1 and attB2 Gateway cloning sequences (Invitrogen Life Technologies/ThermoFisher Scientific) were designed for PCR-based amplification of *CHN2* Prom 4 (hg38 coordinates chr7:29,479,363-29,480,206; primer sequences available in Appendix). The region was amplified from human genomic DNA using PCR Supermix (ThermoFisher Scientific). Subsequent to PCR amplification and purification, each

genomic segment was cloned into the pDONR221 vector using BP Clonase (Invitrogen). Resulting constructs were genotyped by digestion with *Bsr*GI (New England Biolabs, Ipswich, MA) and subjected to DNA sequence analysis to ensure the integrity of the insert. The resulting pDONR221 construct was recombined with an expression construct (pE1B-luciferase) [197] using LR Clonase (Invitrogen) to clone each region upstream of a minimal promoter directing expression of a luciferase reporter gene. Successful recombination was confirmed via digestion of DNA with *Bsr*G1.

Site-directed mutagenesis was performed using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc., Santa Clara, CA). Mutagenesis primers were designed to delete the dimeric SOX10 binding site within *CHN2* Prom 4. Mutagenesis was performed in the pDONR221 construct and DNA from each resulting clone underwent sequence analysis to verify that only the desired mutation was produced. Verified clones were then recombined into pE1B-luciferase using LR Clonase (Invitrogen).

Cell culture, transfection, and luciferase assays

Unmodified parental S16 cells [199] and ΔSOX10 S16 cells were grown under standard conditions in DMEM with 10% fetal bovine serum, 2mM L-glutamine, 50 U/mL penicillin, and 50 g/mL streptomycin. For luciferase assays, ~1,000 cells were plated in each well of a 96-well plate. Cells were cultured for 24 hours under standard conditions prior to transfections.

Lipofectamine 2000 (ThermoFisher Scientific) was diluted 1:100 in OptiMEM I reduced serum medium (ThermoFisher Scientific) and incubated at room temperature for 10 minutes. Each DNA construct to be transfected was individually diluted in OptiMEM to a concentration of 8 ng/μL. An internal control renilla construct was added to the solution at 8 pg/μL. One volume of lipofectamine solution was added to each DNA solution and allowed to sit for 20

minutes at room temperature. Cells were incubated with transfection solution for 4 hours under standard conditions and then the medium was changed to standard growth medium.

Cells were washed with 1X PBS 48 hours after transfection and lysed for 1 hour shaking at room temperature using 20 µL 1X Passive Lysis Buffer (Promega, Madison, WI). 10 µL of lysate from each well was transferred into a white polystyrene 96-well plate (Corning Inc., Corning, NY). Luciferase and renilla activities were determined using the Dual Luciferase Reporter 1000 Assay System (Promega) and a Glomax Multi-Detection System (Promega). Each reaction was performed at least 24 times. The ratio of luciferase to renilla activity and the fold change in this ratio compared to a control luciferase expression vector with no genomic insert were calculated. The mean (bar height) and standard deviation (error bars) of the fold difference are represented in the figures.

RNA isolation and RT-PCR

RNA was isolated from rat sciatic nerve using the RNeasy Mini Kit (Qiagen USA) according to manufacturer's instructions, eluted into 30 µL of RNase-free water, and stored at -80°C. RNA concentration and purity were determined using a NanoDrop Lite (Thermo Scientific, Waltham, MA). A cDNA library was generated using 1 µg of RNA and the High-Capacity cDNA Reverse Transcription Kit (Life Technologies), with the provided random reverse-transcription primers and according to manufacturer's protocol. cDNA samples were analyzed by PCR using Phusion High-Fidelity PCR kit (New England Biolabs) supplemented with 2.5% DMSO, 0.4 µL of each 20 µM primer solution, and 1 µL of cDNA in a 25 µL reaction. A blank (cDNA-negative) control was included and standard PCR conditions were used.

Results

Chn2 harbors a SOX10-dependent transcription start site

The *CHN2* locus encodes the β-chimaerin protein from a complex transcriptional unit; in the human RefSeq database there are five *CHN2* TSSs defined and multiple downstream alternative splicing events, such that the locus is annotated with 14 distinct transcript isoforms (Figure 5.1A and data not shown). Sciatic nerve-derived Tn5Prime data at *Chn2* revealed predominant expression from a single TSS at exon 1D that maps within H3K4me3 and SOX10 ChIP-Seq peaks (Figure 5.1B). Tn5Prime data from differentiating primary Schwann cells revealed that the TSS at exon 1D exhibits little expression in untreated cells and is induced 24-fold upon differentiation (FDR-corrected p-value = 0.000592) (Figure 5.1B). Furthermore, this TSS is expressed in unmodified S16 Schwann cells but is lost with deletion of SOX10 (FDR-corrected p-value = 0.0000165) (Figure 5.1B). Therefore, the *Chn2* TSS at exon 1D was identified through our integrated analysis as a high confidence, SOX10-dependent transcript that may be relevant for differentiating Schwann cells.

SOX10 regulates CHN2 Prom 4 through a dimeric binding motif

To characterize the candidate SOX10-regulated promoter associated with this TSS, we visually analyzed the genomic region underlying the associated SOX10 ChIP-Seq peak for predicted SOX10 binding motifs. This revealed the presence of a dimeric SOX10 binding motif that maps ~150 bases upstream of the TSS; one of the monomeric binding motifs is perfectly conserved from human to chicken (5/5 bases), while the other is less conserved (4/5 bases conserved between human and each species shown; Figure 5.2A). To test the regulatory activity of this element and the contribution of this dimeric motif, we next PCR-amplified, cloned, and sequence-verified an orthologous 844 base pair region from the human genome (see methods)
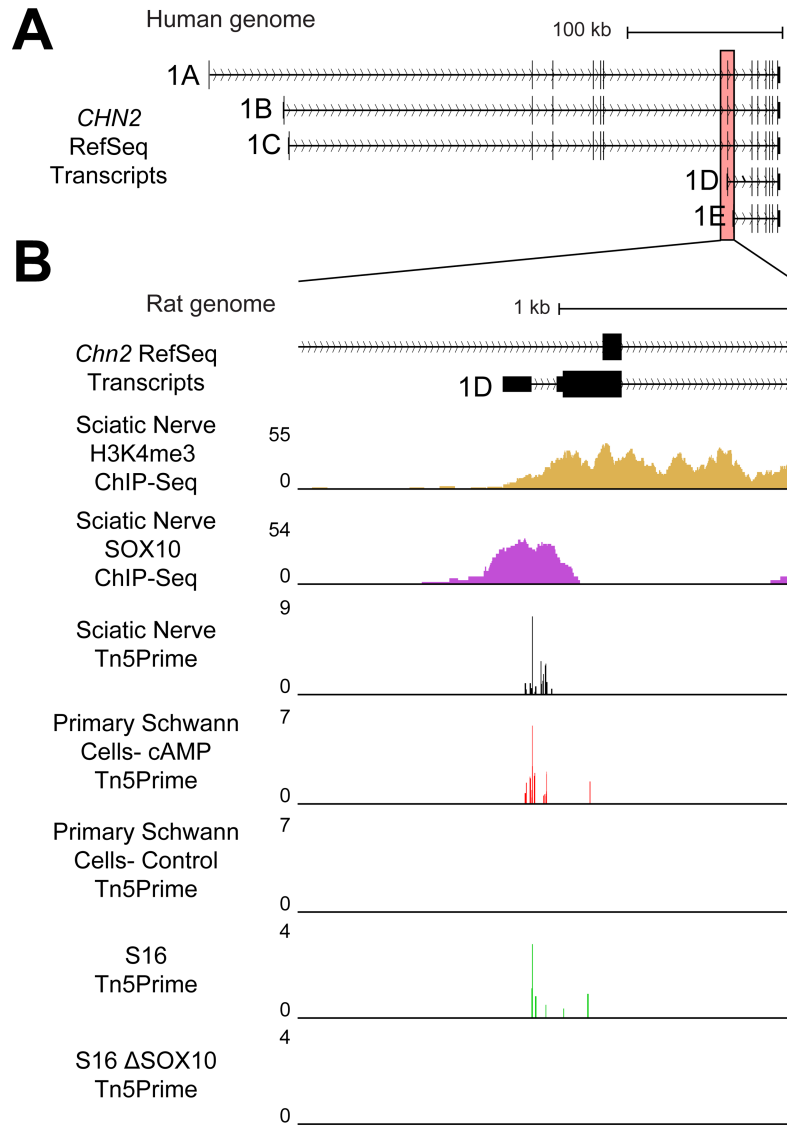
Figure 5.1 SOX10-dependent expression of *Chn2* transcripts originating at exon 1D. (A) The human *CHN2* locus is annotated with five RefSeq transcript start sites, originating at exons 1A through 1E ('1A' through '1E' in panel). (B) The genomic region surrounding exon 1B at the rat *Chn2* locus. This region is recognized by antibodies against H3K4me3 and SOX10 in sciatic nerve. Y-axes for ChIP-Seq data indicate the fold enrichment of sequencing reads above chromatin input. Tn5Prime data at this region are shown from rat sciatic nerve, CPT-cAMP- (cAMP) and vehicle-treated (Control) primary Schwann cells, and unmodified and ΔSOX10 S16 cells. Y-axes for Tn5Prime data indicate the number of transcript 5'ends mapped per base, in reads per million.

132
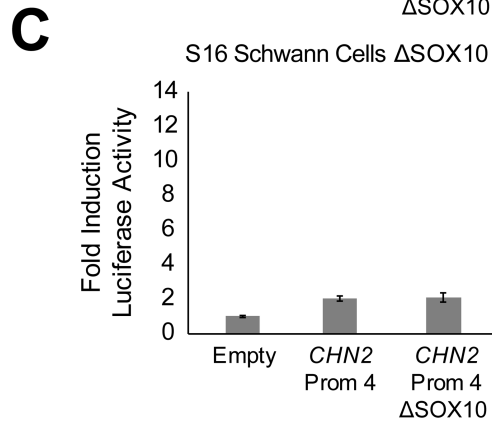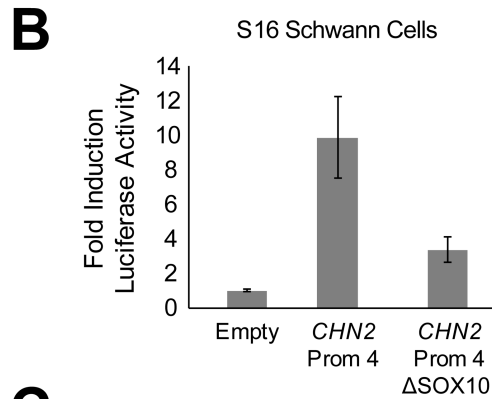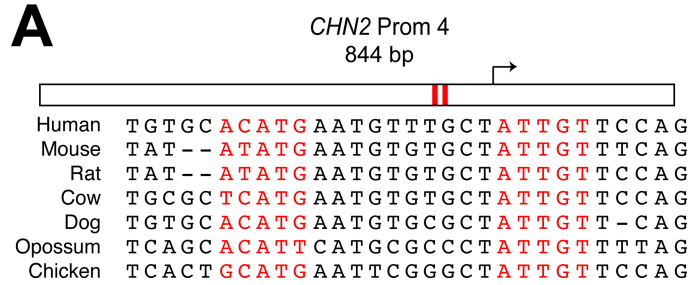
Figure 5.2 SOX10 regulates *CHN2* Prom 4 via a dimeric binding motif. (A) The 844 base pair *CHN2* Prom 4 is shown along with the position of the SOX10 dimeric consensus sequence (red bars and red text). The seven species utilized for comparative sequence analysis are shown on the left. (B and C) *CHN2* Prom 4 with or without the dimeric SOX10 sequence as indicated was cloned upstream of a luciferase reporter gene, transfected into cultured Schwann (S16) cells (B) or ΔSOX10 S16 cells (C), and tested for activity in luciferase assays compared to an empty vector containing no genomic insert. The fold induction of luciferase activity is indicated along the y axis and error bars indicate standard deviations.

for luciferase reporter assays. When cloned upstream of a minimal promoter and luciferase reporter gene, this element—referred to as *CHN2* Prom 4—induced luciferase activity 10-fold greater than the empty control in unmodified S16 cells (p-value=5.2×10^{-10}, Figure 5.2B), consistent with regulatory activity in this context. Deletion of the predicted SOX10 dimeric motif resulted in 70% reduced activity (p-value=2.9×10^{-7}, Figure 5.2B). We next tested each of these constructs in the absence of SOX10 using the ΔSOX10 S16 model, where the wild-type *CHN2* Prom 4 element exhibited only a 2-fold induction of luciferase activity relative to the empty control (p-value=2.9×10^{-11}, Figure 5.2C) and deletion of the SOX10 binding motif had no effect on activity in this context (p-value=0.62). These findings support an important role for SOX10 and the dimeric motif sequence in mediating the activity of *CHN2* Prom 4 in Schwann cells.

Chn2 Prom 4 mediates expression of transcripts encoding β1-chimaerin in sciatic nerve

Due to variation in downstream splicing events, *CHN2* RefSeq transcripts arising from exon 1D are reported to encode a variety of β-chimaerin protein isoforms. Therefore, we next sought to validate the identity of the *Chn2* transcript(s) generated by this promoter in Schwann cells *in vivo*. We designed a forward primer against the 5'UTR of exon 1D and a reverse primer in the 3'UTR (Figure 5.3B) and then performed an RT-PCR using cDNA generated from rat sciatic nerve RNA. This reaction produced a single predominant fragment with an apparent size greater than one kilobase (Figure 5.3A). The reaction was then purified and subjected to Sanger sequencing for identification; the sequenced fragment included all seven of the annotated coding exons for *Chn2* downstream of exon 1D (Figure 5.3B) with no peaks on peaks suggestive of alternative splice isoforms (data not shown). This result confirmed that activity of this promoter induces expression of transcripts coding for the β1-chimaerin isoform in sciatic nerve. Unfortunately, we were unable to acquire an antibody against β-chimaerin that recognizes the β1
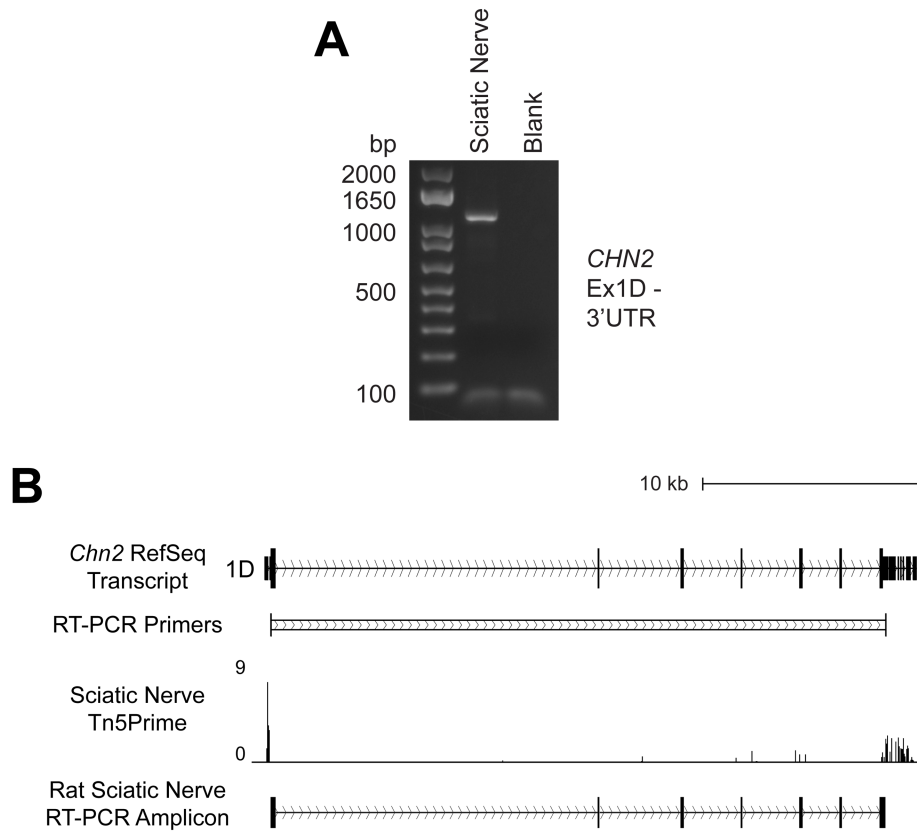
Figure 5.3 Exon 1D-containing *Chn2* transcripts are expressed in sciatic nerve. (A) RT-PCR was used to validate the expression of a spliced *Chn2* transcript with the expected architecture using cDNA from rat sciatic nerve. A blank reaction (no cDNA) was used as a negative control. Sizes of DNA ladder markers are indicated to the left in base pairs (bp). (B) The rat *Chn2* locus is shown, with exon 1D indicated. The locations of RT-PCR primers used in panel A are shown by vertical black bars. Tn5Prime data from rat sciatic nerve is shown as reference for the location of the SOX10-dependent TSS; y-axis indicates number of transcript 5' ends per base in reads per million. The rat sciatic nerve-derived transcript sequence mapped to the rat genome as shown at the bottom of the panel.

isoform and that performs well in a Western blot. Therefore, we were unable to confirm the expression of this protein in Schwann cells. However, the data as a whole support the SOX10-mediated expression of β1-chimaerin in Schwann cells and suggest a role for this protein isoform in mature, myelinating cells.

Discussion

In this chapter I describe transcriptional regulation of *Chn2* in Schwann cells, where transcripts are expressed from a single, predominant TSS at exon 1D. This TSS exhibits increased expression in cAMP-treated primary Schwann cells and SOX10-dependent expression in S16 cells. Regulation by SOX10 is mediated through a dimeric SOX10 binding motif located less than 200 bases upstream of the TSS. I further confirmed that the activity of this promoter in sciatic nerve drives expression of *Chn2* transcripts that encode the β1-chimaerin protein isoform, though the limitations of antibody availability prevented me from confirming the expression of this protein in Schwann cells. β1-chimaerin is one of three β-chimaerin protein isoforms that have been described and studied in the literature, along with β2- and β3-chimaerin[292]. Interestingly, previous studies have established that β1-chimaerin exhibits isoform-specific functional characteristics due to the lack of the N-terminal SH2 domain (Figure 5.4) that acts as an auto-inhibitory module in the β2- and β3-chimaerin isoforms[293]. Indeed, β1-chimaerin localizes to the cellular membrane in response to a 60-fold lower dose of DAG mimic than that required for β2-chimaerin[293]. With the caveat that β1-chimaerin protein expression remains to be confirmed in Schwann cells, the above reports and my own findings suggest that SOX10-mediated expression of β1-chimaerin may reflect a particular requirement for carefully titrated Rac1 activity in differentiating Schwann cells.
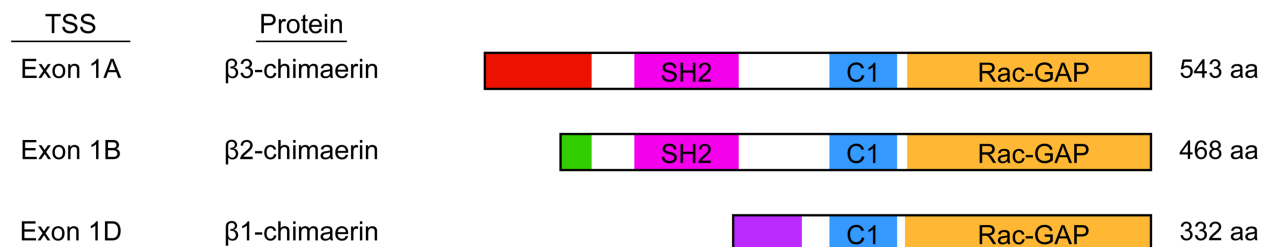
Figure 5.4 β-chimaerin protein isoforms. β-chimaerin isoforms 1, 2, and 3 originate from exons 1D, 1B, and 1A, respectively. β2-and β3-chimaerins contain Src-homology 2 (SH2, magenta), diacylglycerol binding (C1, blue), and Rac-GTPase activating (Rac-GAP, orange) domains, and are distinguished by isoform-specific N-terminal sequences (red and green). β1-chimaerin lacks the SH2 domain and includes an isoform-specific N-terminus sequence (purple).

Interestingly, previous work supports a profound relationship between cellular migration, cellular projections, and precise levels of Rac1 activity *in vitro*. Namely, low levels of Rac1 activity mediate intrinsic, directionally-persistent cell migration, while higher levels promote random, multi-directional migration[294]. Moreover, Laura Feltri and colleagues have applied this model to the role of Rac1 signaling in Schwann cells *in vivo*[291]. They argue that at early stages of Schwann cell development, as cells have come into contact with axons, Rac1 activity is low; this mediates the formation of axial lamellae and elongation of the cells along axons. This stage is analogous to the directionally-persistent cell migration seen in cultured cells with low levels of Rac1 activity[294]. However, in this model Rac1 activity increases as the nerve develops to induce the formation of large, multidirectional radial Schwann cell lamellae that mediate radial sorting and myelination.

Based on the model proposed by Feltri and colleagues, one would predict that β1-chimaerin, as a potent inhibitor of Rac1 at the membrane, would be required at early stages of Schwann cell development to minimize Rac1 activity and may be downregulated upon the onset of radial sorting and myelination. However, our data indicate that β1-chimaerin-encoding transcripts are essentially not expressed in immature primary Schwann cells and are upregulated in the differentiated condition. When considering these findings, it is notable that our primary Schwann cell data are derived from Schwann cells cultured in isolation and therefore do not provide a full impression of developmental gene expression. Therefore, although these conditions mediate a transition toward a differentiated phenotype in a pure population of Schwann cells, it is possible that there are axonal or other cues that modulate the expression of *Chn2* transcripts during development *in vivo*. To better understand how the expression of β1-chimaerin may contribute to the model of Rac1 activity during nerve development, I propose an

138

assessment of β-chimaerin transcript and protein expression at various stages of sciatic nerve development or in a Schwann cell/neuron co-culture context *in vitro.* These studies may provide important insights into the timing of *Chn2* expression as Schwann cells form initial contacts with axons, migrate, elongate, ensheath, and myelinate.

In sum, the data presented in this chapter implicate β1-chimaerin as an isoform-specific SOX10 target gene product and suggest a role for this protein in Schwann cell development through regulation of Rac1. Importantly, this work underscores a need for further study of *CHN2* transcriptional regulation during peripheral nerve development to further elucidate the functional implications of these findings for Schwann cell biology.

Chapter 6

A Novel, SOX10-Regulated Transcription Start Site Contributes to DDR1 Expression in Schwann Cells

Introduction

The *DDR1* locus encodes the discoidin domain receptor tyrosine kinase 1, a member of the DDR family of collagen-activated transmembrane receptors[295, 296]. Like other receptor tyrosine kinases, DDR receptors autophosphorylate upon ligand binding; these phosphorylation events activate the intracellular kinase domain, inducing phosphorylation of binding proteins and downstream cellular signaling. However, DDR receptors exhibit unusually slow and sustained activation kinetics upon ligand stimulation relative to other receptor tyrosine kinases[295, 296]. DDR1 (and another DDR family member, DDR2) are both responsive to many collagen species, including the fibrillar collagen types I-III; however, binding to non-fibrillar collagen appears to be receptor-specific, as only DDR1 binds and is activated by collagen type IV[295, 297]. DDR1 has been implicated in cellular functions related to adhesion and migration as well as extracellular matrix dynamics[298, 299].

Interestingly, biochemical and receptor activation studies have elucidated complex mechanisms regulating DDR1 signaling and the extracellular matrix. For example, it is known that the extracellular regions of DDR proteins are subject to cleavage through the action of membrane-anchored metalloproteinases[298, 300, 301]. This likely acts as a regulatory event to limit DDR signaling. However, this ectodomain 'shedding' may have broader implications for extracellular matrix homeostasis, as the soluble extracellular domains are reported to inhibit

140

collagen fibrillogenesis *in vitro*[302]. Moreover, it has been reported that smooth muscle cells lacking DDR1 expression exhibit increased expression and deposition of fibrillar collagens, but reduced expression of non-fibrillar collagen type IV[303]. Thus, interactions relating DDR1 signaling to the extracellular matrix are complex, may be modulated by post-translational cleavage, and vary depending on the relevant collagen types.

Importantly, collagens are components of the peripheral nerve extracellular matrix and play critical structural and signaling roles during Schwann cell development and function[304]. Indeed, the extracellular matrix molecules that make up the basal lamina—including collagens and laminins—and their cellular receptors are required for Schwann cell survival, radial sorting of axons, and subsequent myelination[82, 305]. Type IV collagen has been well-studied as a constituent of the Schwann cell-deposited basal lamina. Recently, mouse models harboring mutations in *COL4A1* confirmed the importance of collagen type IV in peripheral myelination, as these mice exhibit impaired radial sorting and hypomyelination[306]. Mechanistically, type IV collagen activates the G-protein coupled receptor GPR126[307], which itself is required for Schwann cell myelination[87]. However, other functions of collagen type IV in Schwann cell biology have not been clearly defined.

Based on the ability of DDR1 to bind collagen type IV, it is a plausible candidate to contribute to Schwann cell-basal lamina interactions. In this chapter I describe our identification of a SOX10-dependent TSS at *Ddr1*, the characterization of the SOX10-regulated promoter associated with this TSS, and the expression of DDR1 gene products in Schwann cells. I conclude with a discussion of the implications of these findings for peripheral nerve biology, as well as the possibility of a broader relevance in other SOX10-positive cell types.

Please note that I completed the entirety of this work, with the exception of Sanger sequencing completed by the University of Michigan DNA Sequencing Core.

Materials and Methods

Generation of luciferase reporter gene constructs

Oligonucleotide primers containing attB1 and attB2 Gateway cloning sequences (Invitrogen Life Technologies, Carlsbad, CA) were designed for PCR-based amplification of *DDR1* Prom 5 (hg38 coordinates chr6:30,885,974-30,886,685; primer sequences available in Appendix). The region was amplified from human genomic DNA using Phusion High-Fidelity Polymerase (New England Biotechnology) supplemented with 2.5% DMSO. Subsequent to PCR amplification and purification, each genomic segment was cloned into the pDONR221 vector using BP Clonase (Invitrogen). Resulting constructs were genotyped by digestion with *Bsr*GI (New England Biolabs) and subjected to DNA sequence analysis to ensure the integrity of the insert. The resulting pDONR221 construct was recombined with an expression construct (pE1B-luciferase) [197] using LR Clonase (Invitrogen) to clone each region upstream of a minimal promoter directing expression of a luciferase reporter gene. Successful recombination was confirmed via digestion of DNA with *Bsr*GI.

Site-directed mutagenesis was performed using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc., Santa Clara, CA). Mutagenesis primers were designed to delete the dimeric SOX10 binding site within *DDR1* Prom 5. Mutagenesis was performed in pDONR221 constructs and DNA from each resulting clone underwent sequence analysis to verify that only the desired mutation was produced. Verified clones were then recombined into pE1B-luciferase using LR Clonase (Invitrogen).

Cell culture, transfection, and luciferase assays

Unmodified parental S16 cells [199] and ΔSOX10 S16 cells were grown under standard conditions in DMEM with 10% fetal bovine serum, 2mM L-glutamine, 50 U/mL penicillin, and 50 g/mL streptomycin. For luciferase assays, ~1,000 cells were plated in each well of a 96-well plate. Cells were cultured for 24 hours under standard conditions prior to transfections.

Lipofectamine 2000 (ThermoFisher Scientific) was diluted 1:100 in OptiMEM I reduced serum medium (ThermoFisher Scientific) and incubated at room temperature for 10 minutes. Each DNA construct to be transfected was individually diluted in OptiMEM to a concentration of 8 ng/μL. An internal control renilla construct was added to the solution at 8 pg/μL. One volume of lipofectamine solution was added to each DNA solution and allowed to sit for 20 minutes at room temperature. Cells were incubated with transfection solution for 4 hours under standard conditions and then the medium was changed to standard growth medium.

Cells were washed with 1X PBS 48 hours after transfection and lysed for 1 hour shaking at room temperature using 20 μL 1X Passive Lysis Buffer (Promega, Madison, WI). 10 μL of lysate from each well was transferred into a white polystyrene 96-well plate (Corning Inc., Corning, NY). Luciferase and renilla activities were determined using the Dual Luciferase Reporter 1000 Assay System (Promega) and a Glomax Multi-Detection System (Promega). Each reaction was performed at least 24 times. The ratio of luciferase to renilla activity and the fold change in this ratio compared to a control luciferase expression vector with no genomic insert were calculated. The mean (bar height) and standard deviation (error bars) of the fold difference are represented in the figures.

RNA isolation and RT-PCR

RNA was isolated from rat sciatic nerve using the RNeasy Mini Kit (Qiagen USA) according to manufacturer's instructions. RNA samples were eluted into 30 μL of RNase-free water and stored at -80°C prior to experimentation. RNA concentration and purity were determined using a NanoDrop Lite (Thermo Scientific, Waltham, MA). A cDNA library was generated using 1 μg of RNA and the High-Capacity cDNA Reverse Transcription Kit (Life Technologies), with the provided random reverse-transcription primers and according to manufacturer's protocol. cDNA samples were analyzed by PCR using Phusion High-Fidelity Polymerase (New England Biotechnology) supplemented with 2.5% DMSO, 0.4 μL of each 20 μM primer solution, and 1 μL of cDNA. A blank (cDNA-negative) control was included and standard PCR conditions were used.

Protein isolation and Western blots

To confirm the expression of DDR1 in Schwann cells, protein lysates were collected from rat sciatic nerve, primary Schwann cells (cAMP- or control-treated), unmodified S16 cells, and ΔSOX10 S16 cells. An adult sciatic nerve (age 6-9 months) was sonicated in 200 uL of RIPA buffer (ThermoFisher Scientific) supplemented with protease inhibitor cocktail (ThermoFisher Scientific). Primary Schwann cells and S16 cells were incubated in 0.15% (primary Schwann cells) or 2.5% (S16 cells) trypsin solution to dissociate cells from culture vessels and trypsinization was quenched with D10 (primary Schwann cells) or standard growth (S16 cells) medium. The cell suspensions were centrifuged at 200 x g for 10 minutes (primary Schwann cells), or 800 x g for 2 minutes (S16 cells). Growth medium was removed, cell pellets were suspended in PBS, then centrifuged as above. PBS was removed and cell pellets were suspended in RIPA buffer (Pierce/ThermoFisher Scientific) supplemented with protease inhibitor

144

cocktail (ThermoFisher Scientific). Nerve and cell suspensions were allowed to rock for 30 minutes at 4°C then spun down at 16,000 x g at 4°C for 30 minutes. Lysates were moved into a clean tube and stored at -20°C. Protein yield was measured with a BCA Protein Assay (ThermoFisher Scientific).

For Western blot analysis, 50 μg (sciatic nerve) or 10 μg (primary Schwann and S16 cells) of protein was supplemented with 2X SDS sample buffer (ThermoFisher Scientific) and beta-mercaptoethanol, incubated at 99°C for 5 minutes, then electrophoresed on a 4-20% gradient Tris-Glycine polyacrylamide gel (ThermoFisher Scientific) at 150V for 1.5 hours at room temperature. Protein was transferred to an Immobilon PVDF membrane (ThermoFisher Scientific) in Tris-glycine transfer buffer (ThermoFisher Scientific) containing 10% methanol for ~18 hours at room temperature, running at 0.03 A. Membranes were washed briefly in TBST, then moved into 2% milk in TBST overnight, rocking at 4°C. After 24 hours, membranes were moved into primary antibody dilutions in 2% milk and incubated at 4°C overnight with rocking. Primary antibodies included: anti-DDR1 (rabbit; 1:1000; Cell Signaling Technology, Danvers, MA) and anti-actin (rabbit; 1:5000; Sigma-Aldrich, St. Louis, MO). Membranes were washed three times with TBST. Secondary antibodies (anti-rabbit HRP [donkey; 1:5000; EMD Millipore, Burlington, MA]) conjugated to horseradish peroxidase were diluted in 2% milk and incubated on membranes for one hour rocking at room temperature. After three washes with TBST, membranes were incubated with West Dura HRP substrate (ThermoFisher Scientific) for four minutes. Solutions were then removed and membranes were exposed to X-ray film.

Results

Ddr1 harbors an unannotated, SOX10-dependent transcription start site

        The human *DDR1* locus is a complex transcriptional unit, with five distinct TSSs

annotated in RefSeq (Figure 6.1A), along with three alternatively spliced exons and two

alternative splice sites, resulting in a total of nine *DDR1* transcripts that encode six DDR1

protein isoforms. Note that Figure 6.1A depicts one transcript for each TSS but not all of the

annotated splicing variations. In addition to transcripts arising from the annotated TSSs (data not

shown), our Tn5Prime data from rat sciatic nerve identified a TSS in the first intron of the *Ddr1*

locus that overlaps H3K4me3 and SOX10 ChIP-Seq peaks and that is not annotated in RefSeq

(Figure 6.1B). The TSS identified in sciatic nerve was also detected in cultured primary Schwann

cells with no change in expression between control and cAMP-treated cells (Figure 6.1B).

Further, utilization of this TSS was dependent on SOX10 in the S16 model, with greater than

90% reduced expression in ΔSOX10 S16 cells compared to control (FDR-corrected p-value =

0.027) (Figure 6.1B). These data support the identification of an unannotated, SOX10-dependent

TSS at *Ddr1* that is utilized in Schwann cells and that may be relevant across multiple stages of

Schwann cell development, both before and after differentiation.

DDR1 Prom 5 harbors a dimeric SOX10 binding site

        To characterize the promoter element and SOX10 regulation associated with the above

TSS at *Ddr1*, we visually analyzed the genomic sequence surrounding the TSS and found a

predicted dimeric SOX10 binding motif located ~100 bases upstream of the TSS (Figure 6.2A);

conservation analysis revealed that this dimeric motif is perfectly conserved among mammals

(10/10 bases conserved between human and opossum) and one monomeric motif is further
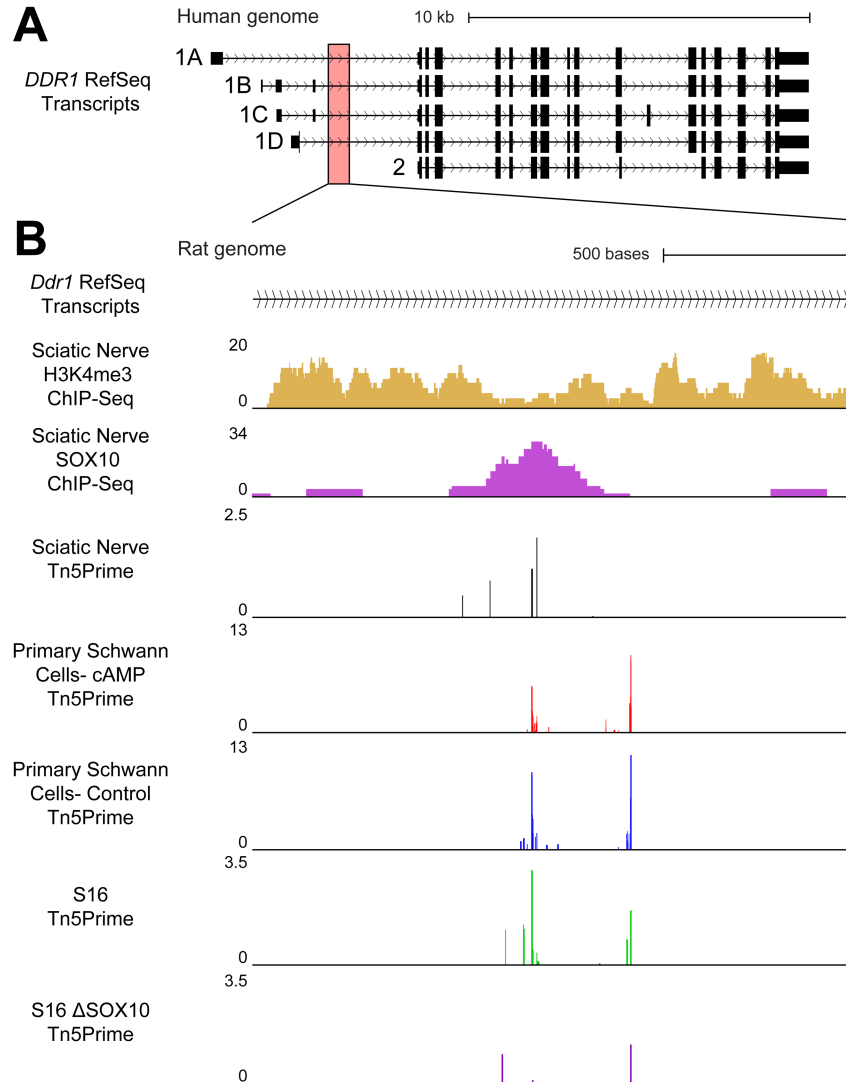
Figure 6.1 SOX10 regulates expression of a novel *Ddr1* transcription start sites in Schwann cells. (A) The human *DDR1* locus is annotated with five RefSeq transcript start sites, originating at exons 1A through 2 ('1A' through '2' in panel). (B) A genomic region in *DDR1* intron 1 at the rat *Ddr1* locus. This region is recognized by antibodies against H3K4me3 and SOX10 in sciatic nerve. Y-axes for ChIP-Seq data indicate the fold enrichment of sequencing reads above chromatin input. Tn5Prime data at this region are shown from rat sciatic nerve, CPT-cAMP-(cAMP) and vehicle-treated (Control) primary Schwann cells, and unmodified and ΔSOX10 S16 cells. Y-axes for Tn5Prime data indicate the number of transcript 5'ends mapped per base, in reads per million.
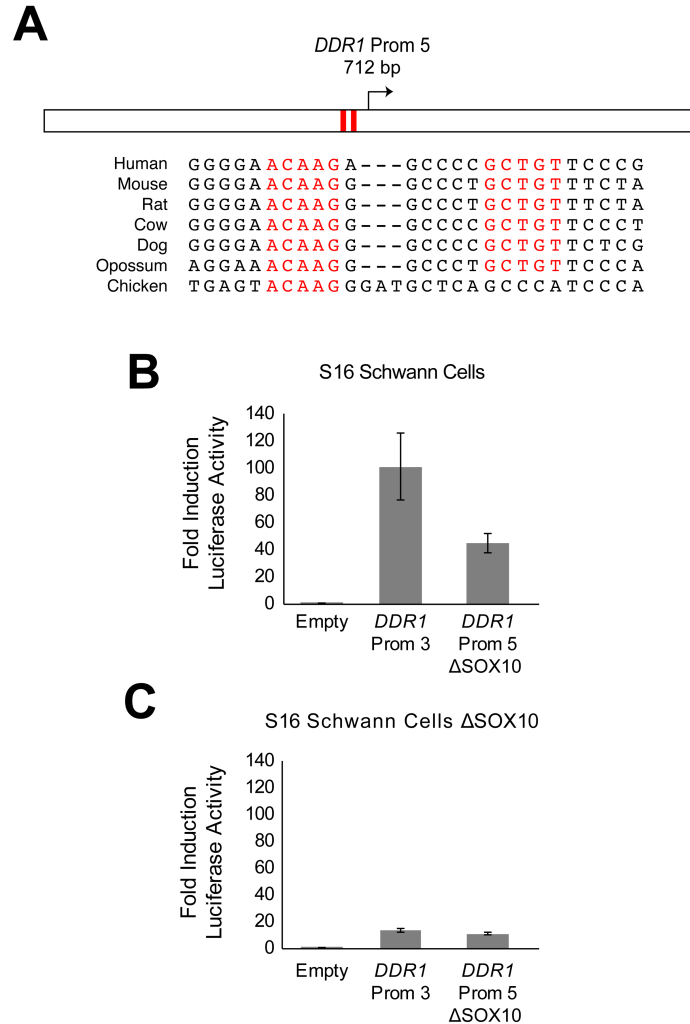
Figure 6.2 SOX10 regulates *DDR1* Prom 5 via a dimeric binding motif. (A) The 712 base pair *DDR1* Prom 5 is shown along with the position of the SOX10 dimeric consensus sequence (red bars and red text). The seven species utilized for comparative sequence analysis are shown on the left. (B and C) *DDR1* Prom 5 with or without the dimeric SOX10 sequence as indicated was cloned upstream of a luciferase reporter gene, transfected into cultured Schwann (S16) cells (B) or ΔSOX10 S16 cells (C), and tested for activity in luciferase assays compared to an empty vector containing no genomic insert. The fold induction of luciferase activity is indicated along the y axis and error bars indicate standard deviations.

conserved among vertebrates (5/5 bases conserved between human and chicken). To test the activity of the promoter and the role of this SOX10 binding motif in promoter activity, we amplified a 712 base pair fragment surrounding the TSS from the human genome (see methods) and tested for regulatory activity in luciferase assays. This element, referred to as *DDR1* Prom 5, exhibited high regulatory activity in S16 cells with 100-fold induction of luciferase activity relative to the empty control vector (Figure 6.2B, p-value=2×10^{-8}). Deletion of the dimeric SOX10 motif sequence resulted in a 60% reduction in activity (Figure 6.2B, p-value=2×10^{-5}), consistent with an important role for this motif in mediating regulatory activity in this context. Next we tested each of the above constructs in the absence of SOX10 using the ΔSOX10 S16 model and found that activity of the wild-type *DDR1* Prom 5 element is severely reduced with only 14-fold induced luciferase activity relative to the empty control (Figure 6.2C, p-value=1×10^{-12}). Moreover, in the absence of SOX10 the deletion of the dimeric SOX10 binding sequence confers only a 20% reduction in the element's activity (Figure 6.2C, p-value=4×10^{-4}). As a whole, these data validate the regulatory activity of the *DDR1* Prom 5 element and identify a dimeric SOX10 binding motif that is directly upstream of the TSS and that mediates a large portion of promoter activity in these assays.

Ddr1 Prom 5 drives transcript expression in sciatic nerve

To understand the functional implications of the SOX10-regulated promoter at *Ddr1*, we next sought to characterize the *Ddr1* transcript and protein products arising from the Tn5Prime-defined TSS. First, we performed an RT-PCR using a forward primer directly downstream of the TSS (exon 1E) and a reverse primer in the 3'UTR sequence of *Ddr1* (Figure 6.3B) with cDNA generated from rat sciatic nerve RNA as template. This resulted in the amplification of a fragment approximately 3 kilobases in size (Figure 6.3A). Sanger sequencing of this amplicon
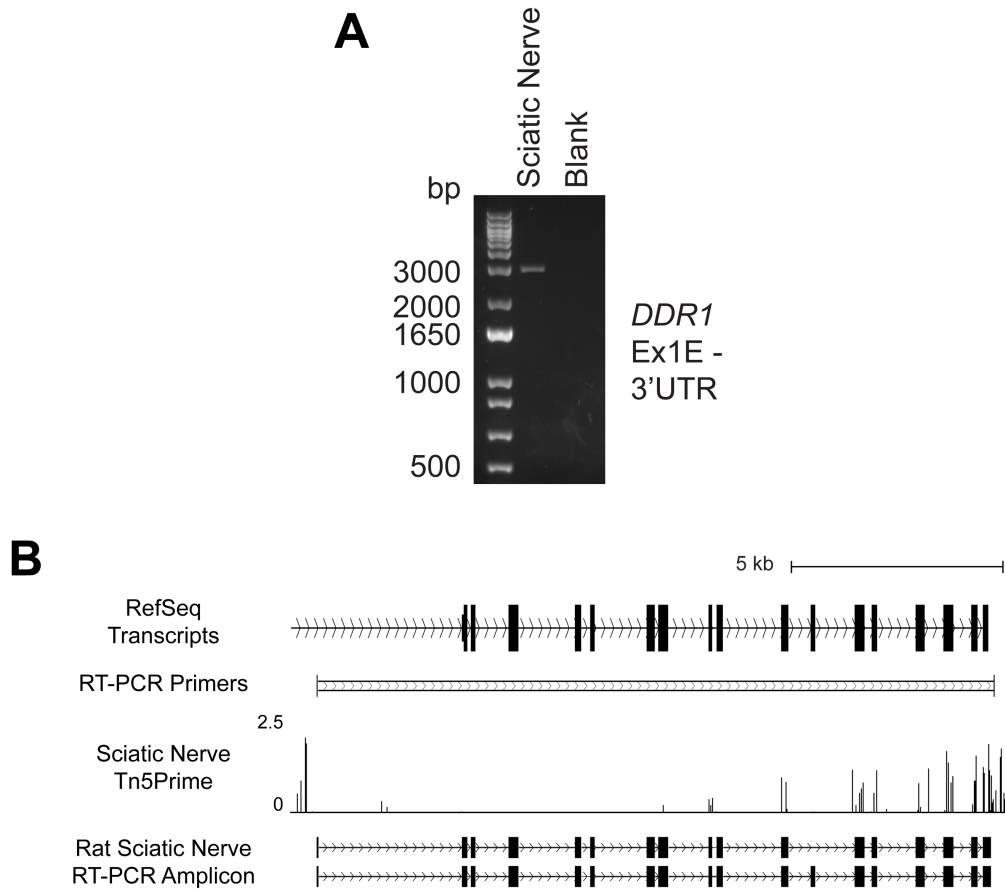
Figure 6.3 *Ddr1* Prom 5 directs expression of *Ddr1* transcripts in sciatic nerve. (A) RT-PCR was used to validate the expression of a spliced *Ddr1* transcript with the expected architecture using cDNA from rat sciatic nerve. A blank reaction (no cDNA) was used as a negative control. Sizes of DNA ladder markers are indicated to the left in base pairs (bp). (B) The rat *Ddr1* locus is shown. The locations of RT-PCR primers used in panel A are shown by vertical black bars. Tn5Prime data from rat sciatic nerve is shown as reference for the location of the SOX10-dependent TSS; y-axis indicates number of transcript 5' ends per base in reads per million. The rat sciatic nerve-derived transcript sequences mapped to the rat genome as shown at the bottom of the panel.

confirmed the expected transcript architecture with the presence of peaks-on-peaks beginning at the boundary between exons 11 and 12 (Figure 6.3B and data not shown). This suggested the presence of alternatively spliced transcripts, as has been described for exon 12 at the human locus[308]. To define the splice isoforms arising from the SOX10-regulated TSS, we performed a subsequent PCR with primers designed against exons 10 and 15 using the 3 kilobase exon 1E-containing RT-PCR fragment (Figure 6.3A) as template. Cloning and Sanger sequencing of the resulting products confirmed the presence of exon 1E-containing transcripts that include or exclude *Ddr1* exon 12 (Figure 6.3B). This alternative splicing event gives rise to the two most abundant DDR1 protein isoforms, DDR1a (encoded by exclusion of exon 12) and DDR1b (encoded by inclusion of exon 12) (Figure 6.4). Note that Figure 6.4 includes only the three autophosphorylation-competent DDR1 protein isoforms that have been described in the literature[308].

DDR1 proteins are expressed in Schwann cells

To further confirm the relevance of the transcript-based studies above, we next assessed the expression of DDR1 protein products in Schwann cells via Western blot analyses. First, it is noteworthy that Tn5Prime data from each of the Schwann cell models support the utilization of multiple *Ddr1* TSSs. Therefore, we anticipate the expression of DDR1 protein products in Schwann cells that do and do not arise from the SOX10-regulated TSS at exon 1E. First, to confirm DDR1 expression *in vivo* we analyzed protein lysates from rat sciatic nerve with an antibody recognizing the C-terminus of DDR1. Indeed, this revealed the expression of protein product(s) with an apparent size of approximately 120 kDa, consistent with the expression of full-length DDR1 in sciatic nerve (Figure 6.5A). Note that DDR1 protein isoforms a, b, and c are not well size-resolved under these conditions. Moreover, a smaller fragment of approximately 60
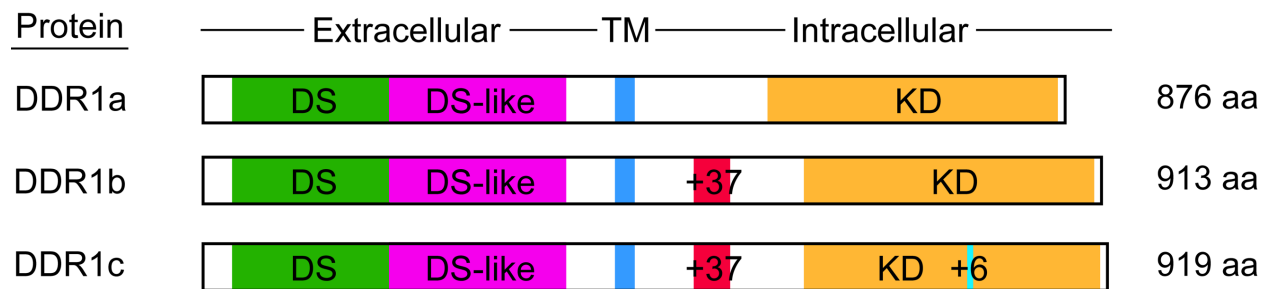
Figure 6.4 DDR1 protein isoforms. DDR1 isoforms a, b, and c contain discoidin (DS, green), discoidin-like (DS-like, magenta), transmembrane (TM, blue) and kinase (KD, orange) domains, and are distinguished by insertions in the intracellular sequences (red and light blue).
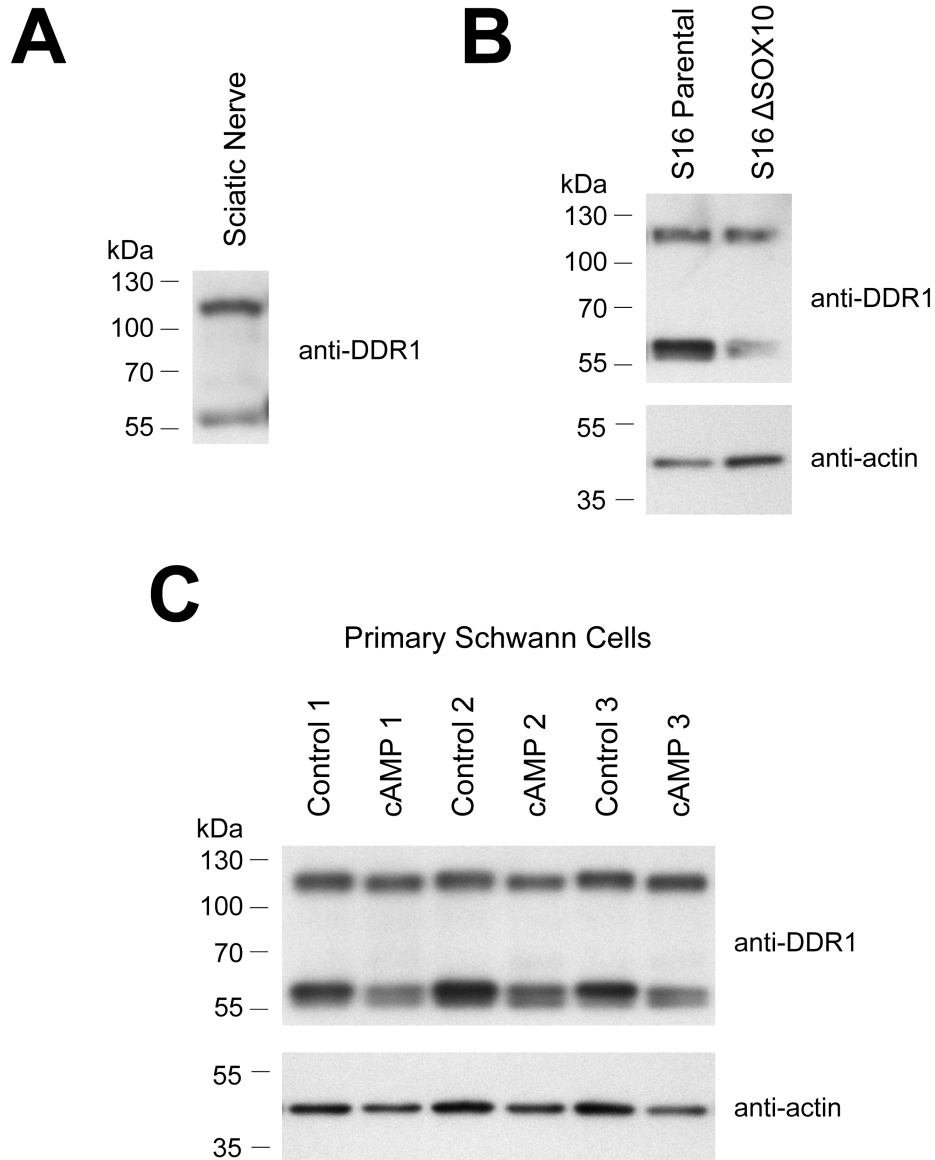
Figure 6.5 DDR1 proteins are expressed in Schwann cells. (A) 50 µg of protein lysate from rat sciatic nerve was used to test the expression of DDR1 *in vivo*. (B) 10 µg of protein lysate from unmodified and ΔSOX10 S16 cells was used to test the expression of DDR1 with loss of SOX10. (C) 10 µg of protein lysate from independent populations of vehicle- (Control) and CPT-cAMP-treated (cAMP) primary Schwann cells was used to test the expression of DDR1 upon cellular differentiation. Anti-actin was used as a protein loading control. Numbered dashes to the left of each blot indicate the position of protein size markers in kilodaltons (kDa).

kDa was also detected in sciatic nerve (Figure 6.5A). This is consistent with previous reports describing metalloproteinase-mediated shedding of the DDR1 extracellular domain, leaving a ~62 kDa membrane-bound intracellular fragment[298, 300, 301]. This finding for the first time suggests that ectodomain cleavage regulates DDR1 signaling in the context of the peripheral nerve.

Next we sought to confirm DDR1 protein expression in the S16 model. To do this we repeated the Western blot using protein lysates from unmodified parental S16 cells and ΔSOX10 S16 cells. This revealed expression of full-length and cleaved DDR1 protein in both conditions (Figure 6.5B); note that the apparent reduction in the cleaved DDR1 fragment in the ΔSOX10 S16 cells in Figure 6.5B has not been observed consistently across experiments and can be considered artifactual to this blot. This finding was not unexpected, given that additional *Ddr1* TSSs are utilized in S16 cells and are not affected by the loss of SOX10 as measured by Tn5Prime (data not shown). Nonetheless, these findings support the expression of DDR1 in Schwann cells and suggest that the expression of Schwann cell-derived metalloproteinases can mediate the cleavage of the ectodomain. Finally, to assess DDR1 expression in primary cells and determine whether expression or cleavage patterns are altered by Schwann cell differentiation, we next performed Western blot analysis with protein lysates collected from cAMP- or control-treated primary Schwann cells (three replicates each). This confirmed expression of full-length and cleaved DDR1 in control- and cAMP-treated primary Schwann cells (Figure 6.5C). Interestingly, there appears to be a slight but consistent size difference in full-length DDR1 between conditions, with an apparently smaller protein detected in cAMP-treated cells compared to controls. Moreover, cAMP-treated primary Schwann cells exhibit a slightly but consistently decreased signal corresponding to the cleaved DDR1 fragment (Figure 6.5C). These findings

suggest that there are differences in splicing, post-translational modifications, and/or cleavage dynamics regulating DDR1 expression in differentiated versus immature Schwann cells. While not conclusive, these findings certainly warrant further investigation. In sum, our data support the expression and cleavage of DDR1 proteins in Schwann cells and are suggestive of dynamic regulation of DDR1 during Schwann cell differentiation.

Discussion

In this chapter I presented our work to identify and characterize a SOX10-regulated promoter at the *Ddr1* locus. The regulatory activity of this element is mediated in part by a dimeric SOX10 binding motif directly upstream of the TSS. Moreover, this promoter directs expression of transcripts encoding DDR1 protein isoforms a and b in sciatic nerve, which are distinguished by the presence or absence of a 37 amino-acid insertion in the intracellular juxtamembrane region (Figure 6.4). This insertion harbors an NPXY motif that mediates the localization of membrane proteins to clathrin-coated pits for endocytosis[309, 310]; the effect of this insertion on trafficking of DDR1b as compared to DDR1a has not been delineated. Notably, preliminary analyses of SOX10-associated changes to *Ddr1* splicing patterns using bulk RNA-Seq data from ΔSOX10 S16 cells showed no striking changes compared to control, suggesting that there is not a relationship between SOX10-regulated promoter use and downstream splicing at this locus.

Despite the expression of multiple *Ddr1* splice isoforms in sciatic nerve, it is noteworthy that the distinguishing characteristic of exon 1E-containing transcripts arising from the SOX10-regulated promoter as opposed to transcripts originating at other *DDR1* TSSs, is not a change to the DDR1 protein-coding sequence but the presence of a unique 5'UTR. Importantly, DDR1 expression has been described in other SOX10-positive cell types, including oligodendrocytes,

melanocytes, and supporting cells of the organ of Corti in the inner ear[311-313], raising the possibility that the activity of this SOX10-regulated promoter mediates DDR1 expression in these cell types as well. As 5'UTR sequence variation is known to affect transcript targeting, stability, and efficiency of translation[314], the use of this alternative promoter may have important implications for the post-transcriptional regulation of *DDR1* in Schwann cells and other SOX10-positive cells. Future studies into the unique characteristics of *DDR1* transcripts derived from exon 1E will be paramount for understanding the functional implications of SOX10 activity at the locus.

The expression of DDR1 in Schwann cells is likely to contribute to the critical interactions of Schwann cells with the basal lamina. DDR1 is activated by collagen type IV, an important basal lamina component that is known to be required for proper myelination[297, 306]. In the case of melanocytes and supporting cells of the organ of Corti, DDR1 has been functionally associated with the adhesion of these cells to basement membranes[312, 313]. This supports the idea that SOX10-mediated expression of DDR1 contributes to basal lamina adhesion and that there may be a similar function for DDR1 in Schwann cells. Indeed, the group describing inner ear defects in *Ddr1* knockout mice briefly noted that myelination of spiral ganglion cells in these mice is perturbed[313]; comprehensive phenotyping of the peripheral nerves of *Ddr1* knockout mice will be required to properly assess this model for a myelin-related phenotype.

Besides playing a structural role in cellular adhesion to the basement membrane, DDR1 signaling was previously implicated in regulating the composition of the extracellular matrix. Given that *Ddr1* ablation in smooth muscle cells resulted in a decrease of collagen type IV expression[303], it can be hypothesized that DDR1 expression in Schwann cells may act to

positively regulate deposition of this basal lamina component. Moreover, our primary Schwann cell data indicate that *Ddr1* transcripts are expressed before and after differentiation; this finding suggests that DDR1 may be important at multiple stages of Schwann cell development. Indeed, basal lamina deposition begins at the immature Schwann cell stage and mice carrying *Col4a1* mutations exhibit an impairment in radial sorting[82, 306], suggesting that *Ddr1* expression at early time points may be functionally important for the developmental maturation of the basal lamina.

Finally, DDR receptors have been shown to increase collagen-mediated integrin signaling *in vitro*[315, 316]. Integrins are well-described extracellular matrix receptors that are important for peripheral myelination, as conditional ablation of the β1 integrin subunit in the embryonic Schwann cell lineage causes defects in radial sorting and subsequent hypomyelination[317]. Although integrins are capable of binding collagens and laminins, laminin-mediated integrin signaling is thought to be the more important contributor to Schwann cell biology[317]. Therefore, the relevance of cooperative interactions between DDR1 and integrins is not clear in the context of the peripheral nerve, where: (***i***) laminin-mediated integrin activation appears to be more important than that derived from collagen, and (***ii***) expression of integrin subunits vary across development timepoints[318]. Nonetheless, further study should test for a link between these receptors to investigate the possibility that DDR1 modulates integrin signaling in Schwann cells.

In sum, the studies described in this chapter demonstrate that SOX10 directs the expression of *Ddr1* in Schwann cells and implicate DDR1 in Schwann cell biology; thus justifying studies to assess for a functional link between DDR1 and myelination. Additionally, the *Ddr1* locus serves as an example of cell type-specific alternative promoter use that mediates

expression of a unique transcript rather than a unique protein isoform. Further study will be required to elucidate the mechanism(s) by which activity of this promoter informs *DDR1* expression and function in Schwann cells, as well as in other SOX10-positive cell types.

Chapter 7

SOX10 Mediates Isoform-Specific Expression

of *GAS7* in Schwann Cells

Introduction

The *Gas7* locus encodes growth arrest specific 7, a cytoskeleton regulator protein that localizes to the cellular membrane and has been functionally implicated in actin assembly, microtubule bundling, and membrane outgrowth *in vitro*[319, 320]. A number of reports have implicated GAS7 in the development of the neural crest, cartilage, and bone[321-324]. However, GAS7 has been studied predominantly in neurons; in this context it is known to induce the formation of neurite projections, including filopodia and lamellipodia, and mediate neural migration[325-328]. Importantly, a mouse model with reduced expression of GAS7 exhibits behavioral motor dysfunction and motor neuron loss upon aging, consistent with a role for this protein in peripheral nervous system function[329].

The regulation of cytoskeleton dynamics during peripheral nerve development has been a prominent focus of groups studying Schwann cell biology in recent years. Indeed, the large cytoplasmic extensions that immature and promyelinating Schwann cells use to sort and wrap axons have been compared to large lamellipodiae[291] and much of the same cellular machinery that mediates lamellipodia formation in cell culture models has been functionally implicated in the early stages of peripheral myelination[330]. Therefore, the ability of GAS7 to mediate actin assembly and induce the formation of cellular protrusions—especially as it relates to lamellipodia formation—serves as a plausible link between GAS7 and Schwann cell function.

159

In this chapter I describe the SOX10-dependent and isoform-specific expression of GAS7 in Schwann cells, mediated by the activation of one promoter at this multi-TSS locus. I present data that validate the regulatory activity of the associated promoter element and that implicate two dimeric SOX10 binding motifs in promoter activity. Additionally, I provide data supporting the expression of *Gas7* transcript and protein isoforms in Schwann cells, with possible implications for context-dependent protein translation. Finally, I discuss the significance of these findings as they relate to Schwann cell development and maintenance.

Note that I completed the entirety of the work described in this chapter, with the exception of Sanger sequencing, which was performed by the University of Michigan DNA Sequencing Core.

Materials and Methods

Generation of luciferase reporter gene constructs

Oligonucleotide primers containing attB1 and attB2 Gateway cloning sequences (Invitrogen Life Technologies, Carlsbad, CA) were designed for PCR-based amplification of *GAS7* Prom 2 (hg38 coordinates chr17:10,036,393-10,037,227; primer sequences available in Appendix). The region was amplified from human genomic DNA using Phusion High-Fidelity Polymerase (New England Biotechnology) supplemented with 2.5% DMSO. Subsequent to PCR amplification and purification, each genomic segment was cloned into the pDONR221 vector using BP Clonase (Invitrogen). Resulting constructs were genotyped by digestion with *Bsr*GI (New England Biolabs) and subjected to DNA sequence analysis to ensure the integrity of the insert. The resulting pDONR221 construct was recombined with an expression construct (pE1B-luciferase) [197] using LR Clonase (Invitrogen) to clone each region upstream of a minimal

promoter directing expression of a luciferase reporter gene. Successful recombination was confirmed via digestion of DNA with *Bsr*GI.

Site-directed mutagenesis was performed using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc., Santa Clara, CA). Mutagenesis primers were designed to delete each of the dimeric SOX10 binding sites within *GAS7* Prom 2. Mutagenesis was performed in pDONR221 constructs and DNA from each resulting clone underwent sequence analysis to verify that only the desired mutation was produced. Verified clones were then recombined into pE1B-luciferase using LR Clonase (Invitrogen).

Cell culture, transfection, and luciferase assays

Unmodified parental S16 cells [199] and ΔSOX10 S16 cells were grown under standard conditions in DMEM with 10% fetal bovine serum, 2mM L-glutamine, 50 U/mL penicillin, and 50 g/mL streptomycin. For luciferase assays, ~1,000 cells were plated in each well of a 96-well plate. Cells were cultured for 24 hours under standard conditions prior to transfections.

Lipofectamine 2000 (ThermoFisher Scientific) was diluted 1:100 in OptiMEM I reduced serum medium (ThermoFisher Scientific) and incubated at room temperature for 10 minutes. Each DNA construct to be transfected was individually diluted in OptiMEM to a concentration of 8 ng/μL. An internal control renilla construct was added to the solution at 8 pg/μL. One volume of lipofectamine solution was added to each DNA solution and allowed to sit for 20 minutes at room temperature. Cells were incubated with transfection solution for 4 hours under standard conditions and then the medium was changed to standard growth medium.

Cells were washed with 1X PBS 48 hours after transfection and lysed for 1 hour shaking at room temperature using 20 μL 1X Passive Lysis Buffer (Promega, Madison, WI). 10 μL of lysate from each well was transferred into a white polystyrene 96-well plate (Corning Inc.,

Corning, NY). Luciferase and renilla activities were determined using the Dual Luciferase Reporter 1000 Assay System (Promega) and a Glomax Multi-Detection System (Promega). Each reaction was performed at least 24 times. The ratio of luciferase to renilla activity and the fold change in this ratio compared to a control luciferase expression vector with no genomic insert were calculated. The mean (bar height) and standard deviation (error bars) of each fold difference are represented in the figures.

RNA isolation and RT-PCR

RNA was isolated from rat sciatic nerve using the RNeasy Mini Kit (Qiagen USA) according to manufacturer's instructions. RNA samples were eluted into 30 μL of RNase-free water and stored at -80°C prior to experimentations. RNA concentration and purity were determined using a NanoDrop Lite (Thermo Scientific, Waltham, MA). A cDNA library was generated using 1 μg of RNA and the High-Capacity cDNA Reverse Transcription Kit (Life Technologies), with the provided random reverse-transcription primers and according to manufacturer's protocol. cDNA samples were analyzed by PCR using PCR Supermix (ThermoFisher Scientific) supplemented with 3% DMSO, 0.5 μL of each 20 μM primer solution, and 1 μL of cDNA. A blank (cDNA-negative) control was included and standard PCR conditions were used.

Protein isolation and Western blots

To confirm the expression of DDR1 in Schwann cells, protein lysates were collected from rat sciatic nerve, primary Schwann cells (cAMP- or control-treated), unmodified S16 cells, and ΔSOX10 S16 cells. An adult sciatic nerve (age 6-9 months) was sonicated in 200 uL of RIPA buffer (ThermoFisher Scientific) supplemented with protease inhibitor cocktail

(ThermoFisher Scientific). Primary Schwann cells and S16 cells were incubated in 0.15%

(primary Schwann cells) or 2.5% (S16 cells) trypsin solution to dissociate cells from culture

vessels and trypsin solution was quenched with D10 (primary Schwann cells) or standard growth

(S16 cells) medium. The cell suspensions were centrifuged at 200 x g for 10 minutes (primary

Schwann cells), or 800 x g for 2 minutes (S16 cells). Growth medium was removed, cell pellets

were suspended in PBS, then centrifuged as above. PBS was removed and cell pellets were

suspended in RIPA buffer (Pierce/ThermoFisher Scientific) supplemented with protease inhibitor

cocktail (ThermoFisher Scientific). Nerve and cell suspensions were allowed to rock for 30

minutes at 4°C then spun down at 16,000 x g at 4°C for 30 minutes. Lysates were moved into a

clean tube and stored at -20°C. Protein yield was measured with a BCA Protein Assay

(ThermoFisher Scientific).

For Western blot analysis, 50 μg (sciatic nerve) or 10 μg (primary Schwann and S16

cells) of protein was supplemented with 2X SDS sample buffer (ThermoFisher Scientific) and

beta-mercaptoethanol, incubated at 99°C for 5 minutes, then electrophoresed on a 4-20%

gradient Tris-Glycine polyacrylamide gel (ThermoFisher Scientific) at 150V for 1.5 hours at

room temperature. Protein was transferred to an Immobilon PVDF membrane (ThermoFisher

Scientific) in Tris-glycine transfer buffer (ThermoFisher Scientific) containing 10% methanol for

~18 hours at room temperature, running at 0.03 A. Membranes were washed briefly in TBST,

then moved into 2% milk in TBST overnight, rocking at 4°C. After 24 hours, membranes were

moved into primary antibody dilutions in 2% milk and incubated at 4°C overnight with rocking.

Primary antibodies included: anti-GAS7 (mouse; 1:100; Santa Cruz Biotechnology, Dallas, TX)

and anti-AARS (rabbit; 1:2000; Bethyl Laboratories, Montgomery, TX). Membranes were

washed three times with TBST. Secondary antibodies conjugated to horse radish peroxidase

were diluted in 2% milk and incubated on membranes for one hour rocking at room temperature. Antibodies included anti-rabbit HRP (donkey; 1:5000; EMD Millipore, Burlington, MA) and anti-mouse HRP (goat; 1:2000; ThermoFisher Scientific). After three washes with TBST, membranes were incubated with West Dura HRP substrate (ThermoFisher Scientific) for four minutes, then membranes were drained and exposed to film.

Results

Gas7 transcripts originating at exon 1B are SOX10-dependent in Schwann cells

The human *GAS7* locus harbors four RefSeq-annotated TSSs (Figure 7.1A) that each encode distinct protein isoforms. Tn5Prime data from rat sciatic nerve identified a single predominant TSS at the *Gas7* locus that maps to exon 1B (Figure 7.1B). This TSS maps near H3K4me3 and SOX10 ChIP-Seq peaks, is expressed similarly in control and differentiated primary Schwann cells, and exhibits expression in S16 cells that is largely lost upon deletion of SOX10 (FDR-adjusted p-value = 2.15×10^{-165}) (Figure 7.1B). Taken together these data support the identification of a candidate SOX10-regulated promoter and associated TSS at *Gas7* in Schwann cells.

GAS7 Prom 2 is regulated by SOX10 through two dimeric binding motifs

To assess the mechanism by which this element is regulated by SOX10, we visually scanned the genomic region upstream of the Tn5Prime-defined *Gas7* TSS for predicted SOX10 binding motifs. This revealed the presence of two dimeric binding motifs (SOX10-1 and SOX10-2) that are separated by approximately 30 base pairs and reside less than 350 bases upstream of the TSS (Figure 7.2A). These motifs exhibit perfect or near-perfect sequence conservation
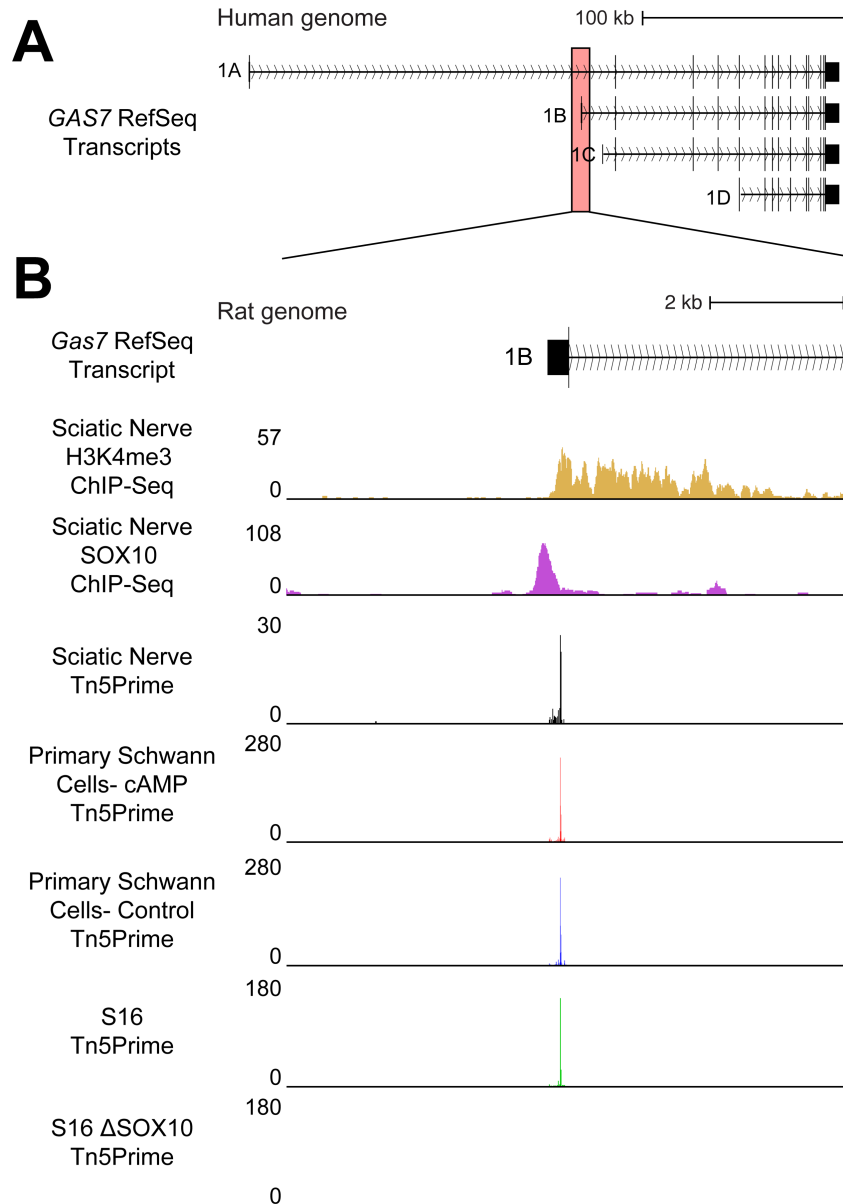
Figure 7.1 SOX10-dependent expression of a *Gas7* transcription start site at exon 1B. (A) The human *GAS7* locus is annotated with four RefSeq transcript start sites, originating at exons 1A through 1D ('1A' through '1D' in panel). (B) The genomic region surrounding *GAS7* exon 1B at the rat *Gas7* locus. This region is recognized by antibodies against H3K4me3 and SOX10 in sciatic nerve. Y-axes for ChIP-Seq data indicate the fold enrichment of sequencing reads above chromatin input. Tn5Prime data at this region are shown from rat sciatic nerve, CPT-cAMP- (cAMP) and vehicle-treated (Control) primary Schwann cells, and unmodified and ΔSOX10 S16 cells. Y-axes for Tn5Prime data indicate the number of transcript 5'ends mapped per base, in reads per million.
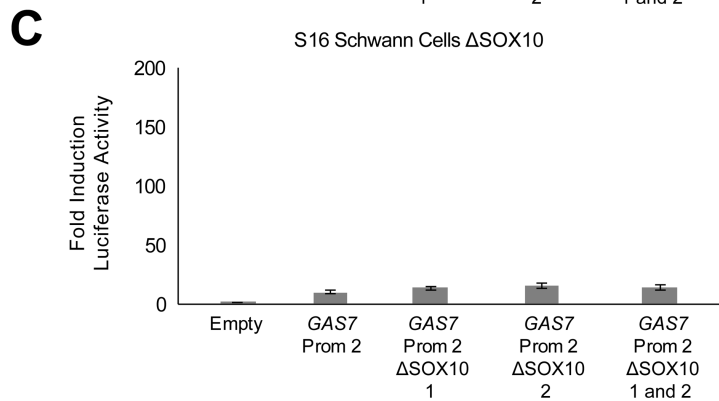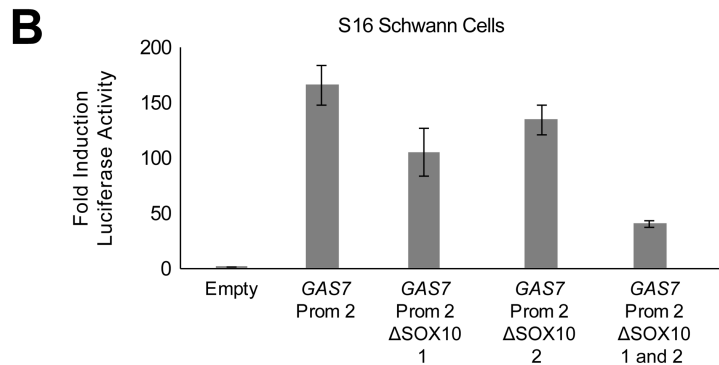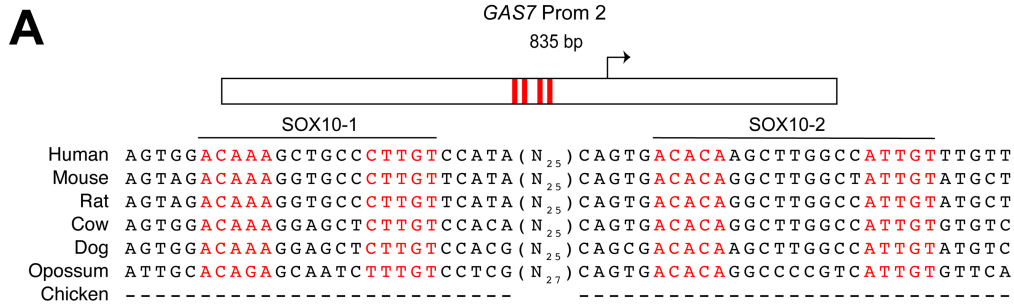
Figure 7.2 SOX10 regulates GAS7 Prom 2 via two dimeric binding motifs. (A) The 835 base pair *GAS7* Prom 2 is shown along with the position of the SOX10 dimeric consensus sequences (red bars and red text). The seven species utilized for comparative sequence analysis are shown on the left. (B and C) *GAS7* Prom 2 with or without the dimeric SOX10 sequences as indicated was cloned upstream of a luciferase reporter gene, transfected into cultured Schwann (S16) cells (B) or ΔSOX10 S16 cells (C), and tested for activity in luciferase assays compared to an empty vector containing no genomic insert. The fold induction of luciferase activity is indicated along the y axis and error bars indicate standard deviations.

among mammals (SOX10-1: 8/10 bases conserved between human and opossum; SOX10-2: 10/10 bases conserved between human and opossum; Figure 7.2A). To test the regulatory activity of this promoter region and to determine the role of the dimeric SOX10 binding sites in promoter activity, an 835 base pair fragment containing the orthologous sequence from the human genome (see methods) was PCR-amplified and cloned upstream of a minimal promoter and luciferase reporter gene. Assaying the activity of this element—referred to as *GAS7* Prom 2—in S16 cells revealed that it is highly active in this cellular context, with greater than 160-fold induced luciferase activity relative to a control vector with no genomic insert (Figure 7.2B, p-value=3×10^{-13}). To measure the contributions of the SOX10-1 and SOX10-2 dimeric motifs to the regulatory activity, each motif was deleted individually; this resulted in 30% (p-value=2×10^{-5}) and 20% (p-value=1×10^{-3}) reduced activity relative to the wild-type construct, respectively (Figure 7.2B). However, when both dimeric motifs were deleted together the activity of *GAS7* Prom 2 was reduced by approximately 75% (p-value=2×10^{-11}), indicating that these sequences together confer a large portion of activity in this assay (Figure 7.2B). Subsequently, we tested the wild-type and mutagenized *GAS7* Prom 2 constructs for their regulatory activity in the absence of SOX10 using the ΔSOX10 S16 cells, where wild-type *GAS7* Prom 2 exhibited only a 10-fold induction of luciferase relative to the empty control (Figure 7.2C, p-value=5×10^{-10}). Moreover, in this context deletion of the SOX10 binding motifs individually or together did not reduce the activity of the element, and even marginally increased activity relative to the wild-type construct (ΔSOX10-1, p-value=4×10^{-4}; ΔSOX10-2, p-value=1×10^{-4}; and ΔSOX10-1 and 2, p-value=6×10^{-4}). As a whole, these studies support the identification of a TSS at *Gas7* that is proximally regulated by SOX10 via two dimeric binding motifs.

Gas7 Prom 2 drives full-length transcript expression in sciatic nerve

To better understand the impact of this SOX10-regulated promoter and to identify the

GAS7 gene products that are expressed in Schwann cells, we next sought to characterize the

transcripts originating from *Gas7* Prom 2 in peripheral nerve. We performed an RT-PCR using a

forward primer recognizing exon 1B and a reverse primer in the *Gas7* 3'UTR (Figure 7.3B) with

sciatic nerve cDNA as template. This reaction generated a single predominant amplicon with an

apparent size greater than 1,650 base pairs (Figure 7.3A). Purification and Sanger sequencing

confirmed that this exon 1B-containing fragment exhibits the expected transcript architecture and

gave no indication of alternative splicing as evidenced by clean signal without peaks-on-peaks

(Figure 7.3B and data not shown). These results confirmed that the SOX10-dependent TSS

identified by Tn5Prime is associated with the expression of transcripts encoding the GAS7-b

protein isoform (Figure 7.4).

GAS7 protein isoforms a and b are expressed in Schwann cells

To characterize the relevance of the above transcriptomic studies for protein isoform

expression, we next assessed GAS7 expression in Schwann cells via Western blot analyses. First,

we analyzed the expression of GAS7 in rat sciatic nerve protein lysates using an antibody

directed against the C-terminus of GAS7. This confirmed the expression of a protein with an

apparent size consistent with the GAS7-b isoform (~48 kDa) in sciatic nerve (Figure 7.5A). This

finding supports the notion of GAS7-b expression in Schwann cells *in vivo*, with the caveat that

the cellular origin of GAS7 protein from sciatic nerve lysate is unclear; as noted in Chapter 3 the

sciatic nerve is composed of a heterogeneous population of cells. Next, to confirm that GAS7

protein expression, like *Gas7* transcript expression, is dependent on SOX10 in Schwann cells, we

performed a similar Western blot assay using protein lysates from unmodified parental and
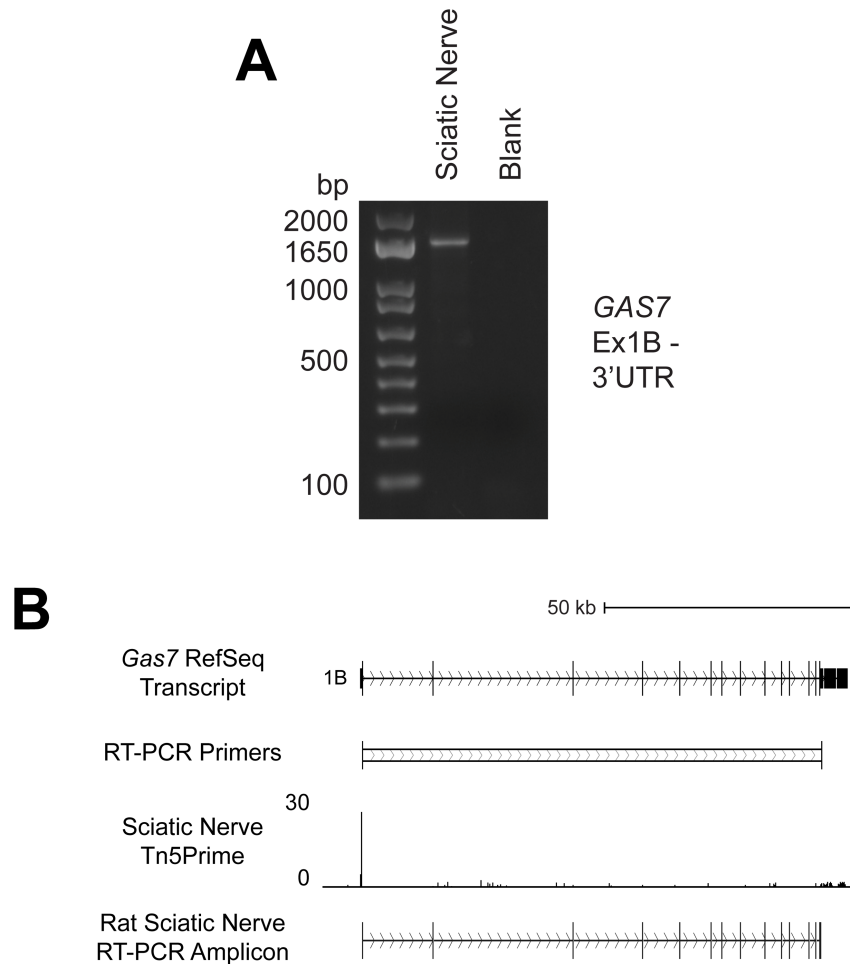
Figure 7.3 *Gas7* Prom 2 directs transcript expression in sciatic nerve. (A) RT-PCR was used to validate the expression of exon 1B-derived *Gas7* transcripts with the expected architecture using cDNA from rat sciatic nerve. A blank reaction (no cDNA) was used as a negative control. Sizes of DNA ladder markers are indicated to the left in base pairs (bp). (B) The rat *Gas7* locus is shown with the locations of RT-PCR primers used in panel A indicated by vertical black bars. Tn5Prime data from rat sciatic nerve is shown as reference for the location of the SOX10-dependent TSS; y-axis indicates number of transcript 5' ends per base in reads per million. The rat sciatic nerve-derived transcript sequence mapped to the rat genome as shown at the bottom of the panel.
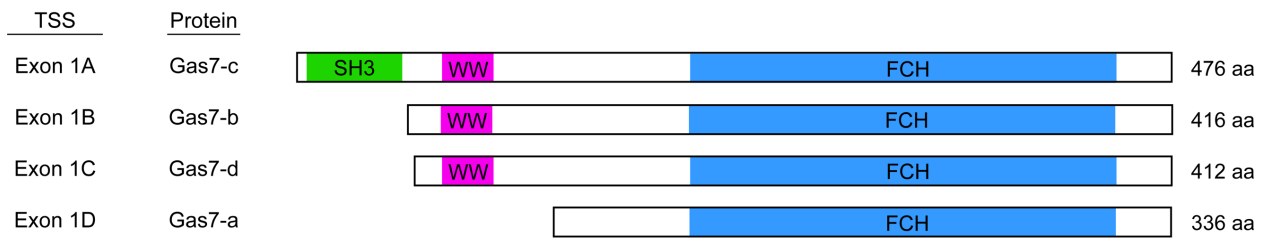
Figure 7.4 GAS7 protein isoforms. GAS7 isoforms a, b, c, and d are associated with transcripts originating from exons 1D, 1B, 1A, and 1C, respectively. Isoform Gas7-c contains Src homology 3 (SH3, green), WW (magenta), and Fes/Cip4 homology (FCH, blue) domains. Isoforms Gas7-b and Gas7-d lack the SH3 domain, and Gas7-a lacks both the SH3 and WW domains.
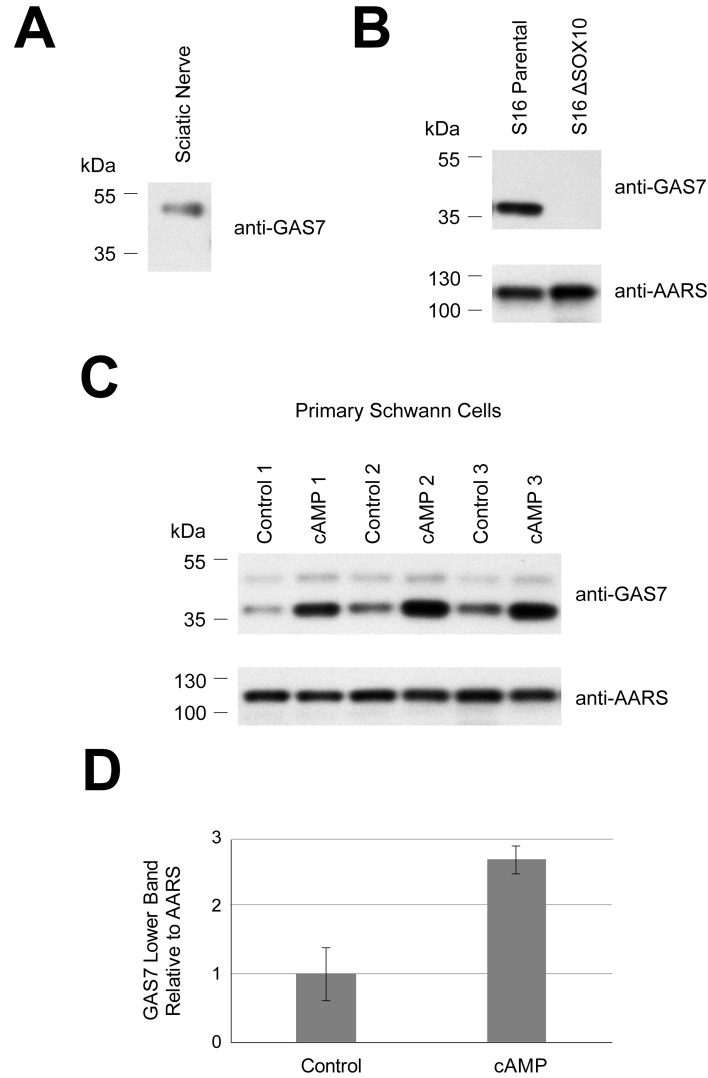
Figure 7.5 GAS7 protein isoforms are expressed in Schwann cells. (A) 50 μg of protein lysate from rat sciatic nerve was used to test the expression of GAS7 *in vivo*. (B) 10 μg of protein lysate from unmodified and ΔSOX10 S16 cells was used to test the expression of GAS7 with loss of SOX10. (C) 10 μg of protein lysate from independent populations of vehicle- (Control) and CPT-cAMP-treated (cAMP) primary Schwann cells was used to test the expression of GAS7 upon cellular differentiation. Anti-AARS was used as a protein loading control. Numbered dashes to the left of each blot indicate the position of protein size markers in kilodaltons (kDa). (D) The intensity of the lower GAS7 band was normalized to AARS signal for each primary Schwann cell sample. Ratios were normalized to the control condition, and the average across the three independent samples is indicated by the bar height. Error bars indicate standard deviation.

ΔSOX10 S16 cells. First, this revealed that parental S16 cells express a protein product recognized by the GAS7 antibody and that this protein product is absent in ΔSOX10 S16 cells (Figure 7.5B). However, this protein is substantially smaller than the protein expressed in the sciatic nerve. The smaller protein detected in S16 cells exhibits an apparent size consistent with the GAS7-a protein isoform (~39 kDa; Figure 7.4), despite the fact that our TSS mapping data provide no evidence for transcripts originating at the TSS associated with this isoform (exon 1D; data not shown). Importantly, it has been previously described that cells transfected with the GAS7-b-encoding cDNA sequence express multiple GAS7 protein isoforms with molecular weights consistent with the products detected in sciatic nerve and S16 cells[327, 331]. Based on the presence of in-frame ATG codons, it is plausible that the smaller protein product is derived from protein translation that initiates at the third in-frame ATG, resulting in a 330 amino acid protein with an expected molecular weight of approximately 39 kDa. Therefore, previously published data and the work presented here support the idea that generation of variable protein products is a property of exon 1B-derived *Gas7* transcripts and may result from a leaky ribosomal scanning mechanism. To further confirm GAS7 protein expression in Schwann cells and to determine if protein expression is altered upon differentiation, we repeated the GAS7 western using protein lysates from three replicates each of cAMP- and control-treated primary Schwann cells. This revealed that primary Schwann cells express both the larger and smaller GAS7 protein isoforms, with greater expression of the smaller product (Figure 7.5C). Moreover, cAMP-treated cells exhibited a consistent, ~2.5-fold upregulation of the smaller GAS7 protein relative to controls (Figure 7.5D), suggesting that GAS7 protein expression is dynamically regulated during Schwann cell differentiation despite the unchanged TSS expression level at exon 1B (Figure 7.1B). As a whole, our data support the expression of multiple GAS7 protein

172

products in Schwann cells, and suggest a role for dynamic and context-dependent expression of these proteins during Schwann cell differentiation.

Discussion

In this chapter I describe our efforts to implicate SOX10 in the expression of *GAS7* in Schwann cells and suggest a functional role for GAS7 in peripheral myelination. We found that Schwann cells express *Gas7* through the use of a single, SOX10-dependent TSS at exon 1B. Our studies of the associated promoter element support the conclusion that SOX10 regulates expression proximally to the TSS through two conserved dimeric SOX10 binding motifs. Notably, the regulatory activity of *GAS7* Prom 2 was more severely reduced in the ΔSOX10 S16 model than by the deletion of the dimeric SOX10 binding motifs in unmodified S16 cells; this pattern of activity is reminiscent of that seen for *ARPC1A* Prom 2 in Chapter 4. In a similar manner to that locus, EGR2 ChIP-Seq data from sciatic nerve[154] indicate that EGR2 binds to the *Gas7* Prom 2 region *in vivo*. This suggests that EGR2 may contribute to the regulatory activity of the promoter element that we observed in unmodified S16 cells, and that the loss of both SOX10 and EGR2 in the ΔSOX10 S16 model may contribute to the severely reduced activity in that context.

Given our characterization of a SOX10-regulated promoter at the *Gas7* locus, it is notable that GAS7 has also been studied in the context of chondrocyte differentiation. Indeed, *Gas7* expression is induced in an *in vitro* chondrocyte differentiation model[321]. Interestingly, SOX9—a SOXE group member that is closely related to SOX10—is known to be essential for cartilage formation[332]. Moreover, the induction of *Gas7* in the chondrocyte differentiation assay was abrogated by SOX9 knockdown[321]. SOXE group members are known to bind to similar motifs due to high sequence similarity in the HMG binding domain; indeed, it has been

173

previously reported that SOXE factors can bind and regulate the same element in their respective cellular contexts[333]. Thus, it is plausible that SOX9 mediates the expression of *Gas7* in chondrocytes in a similar manner to what we have observed for SOX10 regulation of the locus in Schwann cells. This idea could be tested by: (***i***) analysis of SOX9 ChIP-Seq data to assess binding at the *GAS7* Prom 2 region in chondrocytes[334], (***ii***) investigating the ability of SOX9 to induce regulatory activity of the *GAS7* Prom 2 element in luciferase reporter assays, and (***iii***) using the motif deletion constructs we generated to test the role of the same dimeric motif sequences in SOX9-mediated activity.

Our Tn5Prime data from primary Schwann cells indicate that the expression level of the SOX10-regulated *Gas7* TSS does not change as cells differentiate *in vitro*. This suggests that GAS7 may be important at multiple points in the Schwann cell lineage. Interestingly, there is some evidence to suggest that GAS7 may be important even earlier than Schwann cell specification. This idea is derived from neural crest-associated defects observed upon morpholino-mediated knockdown of *gas7* in single-cell zebrafish embryos[324]. In this model *gas7* deficiency was associated with a reduction in *sox10*-expressing neural crest cells and reduced numbers of derivative cell populations as measured by markers of neuronal, glial, and melanocyte lineages[324]. These findings were associated with an increase in apoptosis, leading the authors to posit a role for *gas7* in cell survival in the neural crest. Although these data are not definitive and require further evaluation in other models, they suggest that *GAS7* could be a transcriptional target of SOX10 in neural crest cells and may play a functional role early in this lineage.

Simple sequence analysis and assessment of the open-reading frame for the *Gas7* transcript arising from exon 1B indicate that this transcript encodes GAS7 protein isoform b.

Moreover, the Tn5Prime data from sciatic nerve, primary Schwann cells, and S16 cells support the predominant expression of transcripts arising only from exon 1B in each of these models and do not support the use of alternative *Gas7* TSSs in Schwann cells. Therefore, we anticipated expression of a single protein product corresponding in size to GAS7-b via Western blot. However, we found that the expression of GAS7 protein products varies in each of these models, where cells in the sciatic nerve express a protein of ~48 kDa consistent with GAS7-b, S16 cells express a smaller protein with an apparent size of ~38 kDa, and primary Schwann cells express both of these proteins. Interestingly, it has been previously noted in the literature that cells transfected with the cDNA sequence encoding GAS7-b exhibit expression of multiple GAS7 protein isoforms[327, 331]. This was hypothesized to result from leaky ribosomal scanning across the GAS7-b initiation codon to a downstream ATG. However, this has not been mechanistically investigated. Moreover, in the primary Schwann cells, cAMP treatment induces a specific increase in production of the smaller GAS7 protein isoform. This suggests that this protein product may have a particular role in differentiating Schwann cells, though the *in vivo* relevance of this finding is not clear, as the smaller GAS7 protein was not detected in adult sciatic nerve (see below). As a whole, these findings raise questions regarding the mechanism driving the production of multiple GAS7 protein products (*e.g.,* leaky ribosomal scanning as opposed to a protein cleavage event) and the contextual cues dictating variable protein expression between the models studied here. More broadly, they underscore the importance of integrating transcriptomic and protein expression studies to develop a complete understanding of the mechanisms regulating expression of gene products from a locus of interest.

Given the role of GAS7 in cytoskeleton regulation and extension of cellular projections[319, 325-327], the expression patterns we report for GAS7 proteins in Schwann cells

may have important implications for the cellular extensions that are required as these cells sort and ensheath axons during development. In neuronal cells, the expression of GAS7 isoform a has been associated with the induction of lamellipodia formation *in vitro*, while GAS7-b was associated with filopodia[327]. Moreover, under the leaky ribosomal scanning model we anticipate that the smaller GAS7 protein isoform, which exhibits increased expression in cAMP-treated primary Schwann cells, represents a protein product that is largely similar to GAS7 protein isoform a. Similarity of this cAMP-induced product to the lamellipodia-associated isoform a is consistent with the flattened morphology and cellular extensions exhibited by these cells as opposed to the bipolar, spindle shape of control-treated cells[237]. Thus, it may be that as immature Schwann cells differentiate, expression of a GAS7 protein that is similar to isoform a mediates the extension of the lamellipodia-like processes that sort and wrap axons[330]. The lack of this smaller protein expression in adult sciatic nerve could be explained by a developmental regulatory mechanism whereby the protein is transiently upregulated only during radial sorting and active myelination *in vivo*. This warrants further investigation and could be addressed by collecting protein lysates from sciatic nerve at a series of developmental time points and assaying for expression of the smaller GAS7 isoform via Western blots.

Animal models would be useful toward experimentally implicating GAS7 in peripheral nerve function. A mouse model with constitutively reduced—but not absent—GAS7 expression has been reported to exhibit motor impairments with age, and this behavioral phenotype was accompanied by loss of motor neuron cell bodies in the spinal cord as well as changes to the musculature[329]. This study included an assessment of GAS7 protein expression at neuromuscular junctions and found that in control mice GAS7 expression is largely localized to the terminal Schwann cells surrounding presynaptic axon terminals[329]; this provides further

evidence for GAS7 expression in Schwann cells *in vivo*. It is important to note, however, that the cellular origin of the phenotype in the above mouse model was not explored. The authors of the report focused primarily on the roles of GAS7 in neuron morphology and function but did not explore the possibility that GAS7 deficiency in Schwann cells may cause or contribute to the impaired motor function in this model. Indeed, they did not analyze the myelination status of GAS7-deficient mice. Therefore, assessments of nerve conduction velocities and nerve morphology before and after symptom onset will be critical toward understanding the extent to which loss of GAS7 in Schwann cells may contribute to the phenotype in this mouse model. Furthermore, the age-related onset of symptoms in the reported GAS7 mouse model might suggest that—if Schwann cell dysfunction contributes to the phenotype—GAS7 may be important for myelin maintenance rather than the initiation and completion of developmental myelination. However, it is not clear if a complete ablation of GAS7 expression would result in a more severe, developmental phenotype. Therefore, much remains to be learned regarding the role of GAS7 in peripheral nerve development and function.

In sum, the data presented in this chapter support the characterization of *GAS7* as a SOX10 target locus exhibiting isoform-specific expression in Schwann cells. Moreover, *GAS7* is an intriguing candidate gene that likely plays an important role in peripheral nerve biology. This locus provides an excellent example of the ways in which gene expression can be exquisitely regulated to achieve a variety of isoform-specific functional outputs tailored to specific biological contexts: through transcriptional, translational, spatial, and temporal mechanisms.

Chapter 8

Conclusions and Future Directions

Summary

As the myelinating cells of the peripheral nervous system, Schwann cells support peripheral nerve function in multiple ways; these include generating a compact myelin sheath that is required for rapid, saltatory conduction of action potentials and providing metabolic support for axonal segments that are isolated from neuronal cell bodies[93, 94]. Consistent with the highly specialized nature of these roles, Schwann cells require a unique gene expression profile to support cellular function. Indeed, substantial effort has been expended toward understanding mechanisms of transcriptional regulation that mediate the development and maintenance of this cell population[335]. In this dissertation, I have described my work toward characterizing regulatory events at loci that are important for the biology of these important and understudied cells.

In Chapter 2, I discussed our assessment of transcriptional regulation at the disease gene *MTMR2*. Recessively-inherited loss-of-function mutations at this locus are causal for CMT disease Type 4B, a particularly severe demyelinating neuropathy that often renders affected individuals wheelchair-bound[184]. Using a sequence conservation-based approach, we identified multiple putative noncoding *cis*-regulatory elements at the *MTMR2* locus and assessed the ability of these elements to exhibit Schwann cell-specific regulatory activity using luciferase reporter assays. These experiments identified a single intronic region, *MTMR2*-MCS3, as a candidate Schwann cell-specific regulatory element. However, closer interrogation of the

178

surrounding genomic region revealed that *MTMR2*-MCS3 resides directly upstream of annotated expressed sequence tags, suggesting transcriptional activity at this element. Based on this realization, we investigated the possibility that this element acts as an alternative promoter for *MTMR2*. Indeed, through the use of 5'RACE and RNA-Seq analyses, we identified *Mtmr2* transcripts originating downstream of MCS3 in Schwann cells, confirming that it functions as an alternative promoter (referred to thereafter as *MTMR2*-Prom2).

To understand the implications of *MTMR2*-Prom2 activity for Schwann cell biology, we first analyzed the cell type specificity of the *Mtmr2* transcripts expressed from this promoter. As opposed to *Mtmr2* transcripts originating at the previously annotated TSS (exon 1A), which were detected in sciatic nerve, motor neurons, and muscle cells, we only detected expression of the novel *Mtmr2* exon 1B-containing transcripts in sciatic nerve, consistent with cell type-specific expression in Schwann cells *in vivo*. Furthermore, amplification, purification, and Sanger-sequencing of the full-length exon 1B-containing transcript confirmed that this product encodes expression of an N-terminally truncated MTMR2 protein isoform. However, the lack of a specific MTMR2 antibody precluded us from confirming the expression of this MTMR2 protein isoform in Schwann cells.

Based on the cell type-specific regulatory activity of *MTMR2*-Prom2 in luciferase assays and the cell type-specific expression of the associated transcript, we sought to more carefully characterize the regulatory mechanisms dictating *MTMR2*-Prom2 activity. Due to the critical role of the transcriptional activator SOX10 in Schwann cell development and function, we analyzed previously published SOX10 ChIP-Seq data from rat sciatic nerve[154] to assess SOX10 binding at the *MTMR2*-Prom2 region. This revealed a SOX10 ChIP-Seq peak at the element, consistent with SOX10 binding and proximally regulating this promoter. Sequence analysis of the element

revealed an evolutionarily conserved dimeric SOX10 binding motif at the promoter, and luciferase assays in the presence and absence of SOX10 confirmed that this motif confers a large portion of the SOX10-responsive regulatory activity exhibited by *MTMR2*-Prom2 in Schwann cells. Similarly, we assayed the role of SOX10 in expression of *Mtmr2*-2 transcripts by assaying transcript expression in cellular contexts with and without SOX10; these studies confirmed that SOX10 is necessary and sufficient for expression of *Mtmr2*-2.

To investigate the functional impacts of *MTMR2*-Prom2 activity in Schwann cells at the protein level, we performed a series of *in vitro* protein overexpression studies. These revealed that the truncated MTMR2-2 protein isoform exhibits differential localization relative to the full-length protein isoform, with a greater propensity toward puncta formation and nuclear localization. These findings suggest that the activity of a SOX10-regulated promoter at *MTMR2* in Schwann cells directs the expression of a functionally distinct gene product and may have important implications for the role of this locus in Schwann cells. Importantly, insights into the biology of MTMR2 protein isoforms that are expressed in Schwann cells will be critical toward a clear understanding of CMT4B pathology and the future development of well-informed therapeutic approaches.

Importantly, in addition to our work characterizing a SOX10-regulated promoter at *MTMR2*[195], we and others previously described SOX10-regulated promoters at genes known to be important in Schwann cell biology as well as at other loci without known roles in myelination[160-162, 165-170]. In a number of these cases—and similar to our findings at *MTMR2*—SOX10 activates use of one promoter at a multi-TSS locus, and thereby induces transcript and/or protein isoform-specific gene expression[162, 169, 170]. These findings, paired with our understanding of the critical role of SOX10 and SOX10 target genes in Schwann cells,

inspired us to perform a genome-wide assessment of SOX10-regulated promoter use in Schwann cells. We reasoned that the identification of these elements throughout the genome would generate a prioritized dataset of genes and gene products that can be tested for a role in Schwann cell biology.

In Chapter 3, I described my efforts to identify SOX10-regulated promoters in Schwann cells genome-wide. To achieve this, I began by mapping TSS use in adult rat sciatic nerve as a readout of promoter use in Schwann cells *in vivo*. This was accomplished using the Tn5Prime library preparation method[239] and resulted in the identification of ~40,000 TSSs expressed in sciatic nerve. Integration of these data with previously published SOX10[154] and H3K4me3[336] ChIP-Seq datasets from the same tissue allowed me to prioritize TSSs based on proximity to marks of a SOX10-bound promoter. In this manner, 4,993 TSSs were associated with SOX10-bound promoter elements *in vivo* and were thus defined as candidates for further study.

Notably, there are limitations to the approach that I took to associate TSSs in sciatic nerve with SOX10-bound promoter regions. First, sciatic nerve is a heterogeneous tissue made up of many cell types. Although a large portion of sciatic nerve RNA is expected to arise from Schwann cells, the cellular origin of any given transcript is not known. Second, this analysis provides an assessment of TSS use in fully differentiated myelinating and non-myelinating Schwann cells, but provides no developmental context for the expression of those transcripts. This is an important point because SOX10 is expressed through every stage of the Schwann cell lineage[113]; a SOX10 target gene may play a functional role at one or many stages of Schwann cell development. To address this limitation, I employed an *in vitro* primary Schwann cell differentiation assay[237] and assessed TSS use from RNA samples of control and differentiated

cells. For subsequent analysis of these data, I focused on the expression profiles of the 4,993 candidate TSSs identified in sciatic nerve. I found that greater than 4,000 are expressed in primary Schwann cells and greater than 850 exhibit expression changes upon differentiation. These data lend support to the relevance of the candidate TSSs in Schwann cells and provide developmental context for many of them; this will inform a more complete understanding of their functional roles in the Schwann cell lineage.

A third limitation of the sciatic nerve data is that the association of a TSS with a SOX10-bound promoter as defined by a ChIP-Seq peak does not ensure that the activity of the promoter and expression of the transcript are functionally dependent on SOX10. To address this limitation, I developed a SOX10 loss-of-function model using the immortalized S16 Schwann cells[199]. By assaying TSS expression in S16 cells with and without SOX10, I determined that 265 of the 4,993 candidate TSSs are dependent on SOX10 for expression. Interestingly, analysis of the SOX10-dependent TSSs revealed that they are associated with: (**1**) greater intensity of SOX10 binding; (**2**) modestly increased numbers of SOX10 binding motifs; and (**3**) more highly conserved motifs relative to those TSSs that were unaffected by loss of SOX10. SOX10-dependent TSSs were also depleted of CpG islands and expressed in a more restricted manner than TSSs that were not dependent on SOX10, consistent with the characterization of these TSSs as targets of a cell type-specific transcription factor. In sum, these data support that our strategy successfully identified TSSs expressed from *bona fide* SOX10-regulated promoters.

The efforts described above have provided us with a rich dataset of TSSs and associated transcripts that remain to be carefully evaluated for functional roles in peripheral myelination. To begin addressing this I selected four SOX10-dependent TSSs for locus-specific validation

studies, each residing at a novel SOX10 target gene that is annotated with multiple transcript and protein isoforms.

In Chapter 4 I described our validation of an unannotated and SOX10-regulated TSS at the *ARPC1A* locus. This TSS was identified in our Tn5Prime datasets as expressed in sciatic nerve, downregulated upon differentiation in primary Schwann cells, and dependent on SOX10 in S16 cells. The promoter region upstream of and immediately surrounding the TSS exhibits positive regulatory activity in luciferase reporter assays in Schwann cells; this activity is mediated in part by two dimeric SOX10 binding motifs and is severely reduced in the absence of SOX10. I confirmed the expression of a spliced *Arpc1a* transcript originating from the SOX10-regulated promoter *in vivo*. The transcript is predicted to encode a dramatically truncated ARPC1A protein isoform that lacks most of the coding sequence of full-length ARPC1A. Importantly, ARPC1A functions as a subunit of the Arp2/3 actin remodeling complex. Although the expression of the protein isoform that is encoded by the SOX10-regulated transcript remains to be confirmed, I hypothesize that the SOX10-regulated ARPC1A protein isoform functions in a dominant-negative manner by binding and sequestering nucleation promoting factors away from intact Arp2/3 complexes, thereby limiting actin nucleation in developing Schwann cells.

In Chapter 5, I presented my validation of a SOX10-regulated promoter at the *CHN2* locus. *CHN2* encodes β-chimaerin proteins, which function as GTPase activating proteins for Rac1. TSS mapping revealed that a single TSS is utilized at *Chn2* in Schwann cells; the expression of this TSS is induced by differentiation of primary Schwann cells and is dependent on SOX10 in the S16 model. The promoter element associated with this TSS exhibits regulatory activity in Schwann cells that is mediated by a dimeric SOX10 binding motif and is reduced in the absence of SOX10. Finally, I confirmed expression of a *Chn2* transcript arising from this

promoter in sciatic nerve that encodes the β1-chimaerin protein isoform. This isoform has been previously characterized and functions as a more potent Rac1 inhibitor than other β-chimaerin proteins due to the lack of an autoinhibitory domain. Therefore, the SOX10-mediated expression of β1-chimaerin likely reflects a need for carefully titrated Rac1 activity during Schwann cell development.

In Chapter 6 I described my validation studies for a previously unannotated and SOX10-regulated TSS at the *DDR1* locus. The expression of this TSS was defined in sciatic nerve, unchanged during primary Schwann cell differentiation, and SOX10-dependent in S16 cells. Luciferase reporter assays confirmed that the associated promoter region exhibits high regulatory activity in Schwann cells that is reduced with the loss of SOX10, and a large portion of this activity is mediated by a dimeric SOX10 binding site. *Ddr1* transcript expression from this promoter was confirmed in sciatic nerve, with two splice variants identified. These variants encode DDR1 protein isoforms a and b, and expression of DDR1 in Schwann cells was confirmed using western blots. Thus, it is likely that DDR1 protein expression in Schwann cells is important for cellular adhesion and signaling with the basal lamina. However, this locus provides an example of a SOX10-regulated promoter element that generates transcripts distinguished strictly by changes to the 5'UTR sequence. For this reason, follow-up studies should be designed to investigate changes to transcript stability, localization, and/or translation efficiency that might be conferred by this 5'UTR.

Finally, in Chapter 7 I presented my work to validate a SOX10-regulated promoter at the *GAS7* locus. The SOX10-regulated TSS at this locus is expressed in sciatic nerve, does not change in expression level upon primary cell differentiation, and is dependent on SOX10 in the S16 model. The promoter element is highly active in Schwann cells; this activity is dependent on

SOX10 and two dimeric SOX10 binding motifs. Finally, transcripts derived from this promoter are expressed in sciatic nerve and encode GAS7 isoform b. However, western blot analysis of GAS7 expression in each of the Schwann cell models used here indicate that the transcript arising from the SOX10-regulated promoter may be translated from multiple ATG start codons in a context-dependent manner, thereby producing multiple GAS7 protein products; this is consistent with previous reports in the literature. Moreover, the expression of the smaller GAS7 protein is induced by cAMP in the primary cell differentiation model, indicating that differential expression of GAS7 protein products may contribute to Schwann cell development.

Future Directions

Naturally, each aspect of the work presented in this dissertation introduces questions that require further study. Our findings underscore the utility of identifying and characterizing SOX10-regulated promoters toward catalyzing a more complete understanding of peripheral myelination. However, we have much to learn regarding the mechanisms by which SOX10 regulates promoter elements and how these findings relate to other SOX10-positive cell types. Moreover, we have validated isoform-specific gene expression patterns at four novel SOX10 target genes in Schwann cells and further study will be required to fully interrogate the relevance of these expression patterns.

Mechanisms of promoter-proximal gene regulation by SOX10

In Chapter 3 of this dissertation I reported that SOX10 binding at promoters is enriched in the -/+ 500 base pairs surrounding the TSS. Based on what is known about the function of SOX10 as a multifaceted modulator of gene expression, this could reflect proximal gene regulation through a number of mechanisms. These include recruitment of chromatin- and/or histone-modifying factors[137, 138], and the promotion of transcriptional initiation or elongation

via interactions with Mediator or P-TEFb, respectively[134, 136]. It seems likely that the mechanism of action of SOX10 at gene promoters will vary across loci, involve a combination of mechanisms at any given element, and may vary across developmental time points. Thus, it will be important for future studies to pursue a clearer understanding of the mechanisms by which SOX10 acts at promoter elements of interest.

Importantly, the ΔSOX10 S16 cell model may provide an essential experimental context for such mechanistic studies, as these cells will allow investigations into the state of SOX10-regulated promoter elements in the absence of SOX10. For example, do the promoters exhibit changes to the chromatin state, consistent with an important role for SOX10 in recruiting remodelers or histone-modifying enzymes? This idea may be tested by performing ChIP-Seq experiments with antibodies detecting histone modifications of interest [337] or micrococcal nuclease treatments to measure nucleosome occupancy[338]. For example, histone deacetylase 2 (HDAC2) has been reported to interact with SOX10 and synergistically activate the expression of target genes during myelination[138]. Therefore, one might ask: (*i*) if HDAC2 is associated with SOX10-regulated promoter regions by performing HDAC2 ChIP-Seq analyses in sciatic nerve and intersecting the resulting data with SOX10 and H3K4me3 ChIP-Seq data; (*ii*) if SOX10 recruits HDAC2 to these regions by repeating this analysis in unmodified and ΔSOX10 S16 cells; and (*iii*) if loss of SOX10 affects the acetylation status of target promoters by performing ChIP-Seq analyses to measure acetylated histone variants in unmodified and ΔSOX10 S16 cells. If HDAC2 associates with these promoters, this localization is mediated by SOX10, and the loss of SOX10 mediates a change to the histone acetylation patterns at these regions, recruitment of HDAC2 may be an important contributor to SOX10-regulated promoter activity.

It is also possible that the loss of SOX10 directly affects the initiation and/or elongation of transcription. To test these possibilities, high resolution ChIP-based assays with antibodies recognizing RNA polymerase II, other components of the pre-initiation complex (PIC), and the Mediator complex may reveal how loss of SOX10 changes transcription dynamics at individual loci[339]. For example, one would expect that if SOX10 mediates assembly of the PIC at a promoter, signal for PIC components would be lost at that element in ΔSOX10 S16 cells relative to control. If SOX10 is instead required for transcriptional elongation of paused RNA polymerase II, one would expect a build-up of RNA Pol II ChIP-Seq reads in the 50 base pair region downstream of the TSS in the ΔSOX10 condition. Such studies are likely to provide fascinating and critical insights into the mechanisms that are important for SOX10-mediated promoter activation in Schwann cells.

Investigating SOX10-regulated promoter use in Schwann cells versus other cell types

In addition to Schwann cells, SOX10 is known to be important for the function of cell types including oligodendrocytes in the central nervous system and melanocytes in the skin. Therefore, another interesting avenue for follow-up studies based on the work presented in this dissertation would be to investigate if SOX10-regulated promoter use is a shared mechanism across SOX10-positive cell types. Our analysis of SOX10-dependent TSS expression across mouse tissues showed enriched expression in the central nervous system and skin, suggesting that these TSSs may be expressed in these other SOX10-positive cell types. However, a comprehensive, genome-wide analysis of SOX10-bound promoters in each context would provide a full picture of whether this is a common mechanism of SOX10 function.

Importantly, SOX10 ChIP-Seq data are available from spinal cord[214] and from an immortalized melanocyte cell line, melan-a[340], providing readouts of SOX10 binding in

oligodendrocytes *in vivo* and a melanocyte-like cellular context, respectively. Interestingly, the authors of each of these reports found little overlap between SOX10 binding and promoter regions, and therefore concluded that SOX10 does not frequently bind to promoter regions in oligodendrocytes and melanocytes. These findings are intriguing and warrant follow-up to explore the biological significance. If it is the case that SOX10 preferentially binds promoter elements in Schwann cells relative to other cellular contexts, are there Schwann cell-specific factors that mediate recruitment and binding of SOX10 to promoter regions? Does SOX10 interact with components of the transcriptional machinery that are preferentially expressed in Schwann cells, or is there some other chromatin modifier or transcription factor recruiting SOX10 to these elements?

It will also be important to assess which of the SOX10-dependent TSSs that we identified in Schwann cells are similarly regulated and expressed in oligodendrocytes and melanocytes. In the case of oligodendrocytes, which are functionally related to Schwann cells, this may implicate these targets in shared mechanisms mediating myelination in the central and peripheral nervous systems. These analyses may be mediated by the availability of RNA-Seq data from purified oligodendrocytes[341] and melanocytes[342], but could be carried out with greater confidence using the Tn5Prime library preparation method to ensure that the limitations of traditional RNA-Seq for coverage of transcript 5' ends do not confound the analysis. The identification of SOX10-dependent TSSs common to multiple cell types may then reveal shared biological pathways and functions across the variety of SOX10-expressing cells.

Investigating the roles of SOX10-regulated transcripts in Schwann cells

In Chapters 4 through 7, I presented data supporting the SOX10-regulated expression of transcript isoforms at four novel target genes: *ARPC1A*, *CHN2*, *DDR1*, and *GAS7*. In each case

the SOX10-dependent nature of transcript expression supports the idea that these transcripts are important for Schwann cell function. However, further work is needed to characterize functions associated with each SOX10-regulated isoform. Here I briefly propose initial studies for each locus.

ARPC1A

The full-length ARPC1A protein functions as a subunit of the Arp2/3 actin remodeling complex and has been implicated in regulation of complex activity by nucleation-promoting factors (NPFs; see Chapter 4)[277]. Based on the extent of coding sequence truncation in the SOX10-regulated transcript, I do not anticipate that this protein isoform would be incorporated into Arp2/3 complexes. Rather, because the truncated protein includes an extension region termed the 'arm' that has been shown to interact with NPFs, I anticipate that this isoform may act in a dominant-negative manner to sequester NPFs away from Arp2/3 complexes and thereby limit actin nucleation. To test these this possible function, it will first be critical to confirm the expression of the predicted SOX10-regulated protein isoform in Schwann cells; generation of an antibody that recognizes the C-terminus of ARPC1A and that performs well in a western blot will be required.

With expression of the protein confirmed, functional studies should be performed to determine the impact of this protein on actin nucleation activity *in vitro*. First, the model described above predicts that overexpression of this protein isoform will inhibit Arp2/3-mediated actin nucleation and branching. This can be tested by transfecting cells with overexpression constructs harboring the open-reading frame of the short ARPC1A isoform and then visualizing the cellular cytoskeleton with actin-staining dyes. One would expect that cells expressing the SOX10-regulated ARPC1A isoform will exhibit reduced or less complex cytoskeletal networks.

Moreover, this model dictates that this is a dosage-dependent effect, not just for the expression level of truncated ARPC1A, but also the expression of NPFs. Therefore, one could test this further by co-transfecting cells with WASP or other NPFs along with truncated ARPC1A and determine if this rescues the observed cellular phenotype(s). Additionally, it will be important to test for direct interactions between the SOX10-regulated ARPCIA isoform and the respective NPFs to support the dominant-negative model proposed here.

CHN2

The SOX10-regulated TSS at *Chn2* mediates the expression of transcripts that encode the β1-chimaerin protein isoform. Importantly, the functional characteristics of this isoform have been previously investigated. Due to the lack of the N-terminal SH2 domain that functions in an autoinhibitory manner, β1-chimaerin is a more active and potent Rac1 inhibitor than other β-chimaerin isoforms[293]. This suggests that Schwann cells require carefully regulated Rac1 activity. Indeed, the loss of Rac1 is known to interfere with the ability of Schwann cells to extend cellular projections *in vitro* and *in vivo*[291]. Therefore, it can be expected that modulating β1-chimaerin expression in Schwann cells will affect cellular projections as well. To test this, primary Schwann cells can be transfected with a β1-chimaerin expression construct, then fixed and assessed for a reduction in the formation of lamellipodia. Importantly, the effect of β1-chimaerin can be compared to that of the Rac1 inhibitor molecule NSC23766, as treatment with this compound is known to reduce lamellipodia formation by Schwann cells.

DDR1

The SOX10-regulated TSS at *Ddr1* is unique among these four loci, given that the TSS and associated transcript do not encode a different protein relative to other *Ddr1* TSSs. Moreover, other *Ddr1* TSSs that are not dependent on SOX10 are expressed in Schwann cells.

Therefore, while the expression of DDR1 in Schwann cells has important implications for adhesion and signaling of these cells with collagen in the extracellular matrix, the SOX10-regulated transcript may have some additional functional implications. Specifically, the unique 5'UTR conferred by the use of this TSS may have implications for the stability, localization, and translation dynamics for SOX10-regulated transcripts. Therefore, I propose that functional assessments of this gene product in Schwann cells should be centered around the characterization of this transcript as compared to *Ddr1* transcripts derived from other TSSs.

The idea that the SOX10-regulated transcript has altered stability, for example, could be tested by treating Schwann cells with a transcription inhibitor (examples include actinomycin D and 5,6-dichloro-1β-1-ribofuranosylbenzimidazole [DRB]) to stop the production of new mRNAs and measure decay of *Ddr1* transcripts in a TSS-specific manner. This could be accomplished through northern blots or quantitative PCR assays (*e.g.*, droplet digital PCR [ddPCR]). Furthermore, the localization of *Ddr1* transcripts could be assessed with a technique such as RNA fluorescence in situ hybridization (RNA-FISH) to establish if the SOX10-regulated transcript exhibits a unique cellular distribution. Finally, translation efficiency can be assessed by cloning each of the *Ddr1* 5'UTR sequences upstream of a luciferase reporter gene, transfecting Schwann cells with these constructs individually, and testing for differences in luciferase activity based on 5'UTR sequence variation. These assessments are likely to reveal any novel characteristics of the SOX10-regulated *Ddr1* transcript and thereby clarify the relevance of the SOX10-regulated promoter at this locus.

GAS7

The SOX10-regulated promoter at *GAS7* mediates the expression of transcripts that encode the GAS7-b protein isoform. Interestingly, what we observed when we set out to confirm

GAS7 protein expression in Schwann cells is expression of a protein product with a molecular weight consistent with GAS7-b in sciatic nerve, a smaller GAS7 protein in the S16 model, and both GAS7 proteins in primary Schwann cells. Indeed, the ability of GAS7-b-encoding cDNA to mediate expression of multiple isoforms has previously been reported[327, 331]. Moreover, we found that the expression of the smaller GAS7 protein is induced by differentiation in primary Schwann cells, despite the fact that TSS expression did not change in this model. Thus, it appears that the translation of this transcript may be developmentally regulated. Although both GAS7 protein products are expected to include the FCH domain that mediates actin binding and membrane localization[319], the full-length GAS7-b harbors a WW domain that has been implicated in interactions with N-WASP[343]. Therefore, it will be important for future studies to determine the regulatory mechanisms dictating translation of GAS7 isoforms from the SOX10-regulated transcript. One approach to address this might be an RNA-pull down assay to investigate the RNA binding proteins that interact with exon 1B-containing *Gas7* transcripts.

Another critical area of study will be to investigate the relevance of differential GAS7 isoform expression for cytoskeletal dynamics in Schwann cells. To better understand the distinctions between the two GAS7 protein isoforms that are expressed in Schwann cells, I would transfect cultured Schwann cells with constructs encoding each GAS7 protein isoform and assess: (*i*) the localization of the GAS7 proteins, (*ii*) the formation of cellular projections, and (*iii*) the cellular cytoskeleton as visualized by actin staining. Based on previous studies, I would expect the smaller GAS7 protein to induce lamellipodia and the larger protein to induce filopodia; however, those studies were performed in neuronal contexts and it will be interesting to see if GAS7 functions differently in Schwann cells. It should be considered that in these experiments, a construct with the open reading frame for GAS7-b will likely mediate the

192

expression of the smaller protein as well. Therefore, to isolate the effects of GAS7-b, the in-frame ATG codon that initiates the smaller open reading frame and/or the Kozak sequence surrounding it should be mutagenized to limit expression of the smaller protein in the GAS7-b condition. These studies may provide some initial insights into the relevance of the protein expression patterns that we have observed.

Assaying the roles of SOX10-regulated promoters in vivo

Ultimately, to experimentally implicate a SOX10 target transcript in peripheral myelination, an animal model must be used. In the case of a transcript expressed from a SOX10-regulated promoter at a multi-TSS locus, an appealing approach would be to generate a model wherein the SOX10-regulated promoter and associated first exon are ablated, leaving the remainder of the locus and other TSSs intact. An alternative, though risky, approach would be to delete only the SOX10 binding motif(s). With subsequent analysis of peripheral nervous system function and histology, such models would elegantly assay the importance of the SOX10-regulated transcript and protein isoform in Schwann cell development and function. This approach would likely be fruitful for each of the specific loci presented in this dissertation.

For example, a mouse model of *Mtmr2* loss of function recapitulates the peripheral neuropathy and myelin outfoldings seen in the nerves of patients with CMT4B[344]. However, in this model expression of both MTMR2 protein isoforms are abolished so the relative contributions of each isoform to the phenotype are unclear. The generation of a mouse with exon 1B and the associated promoter region deleted, however, would be expected to address the importance of the SOX10-regulated promoter and *Mtmr2* gene products for Schwann cell biology.

Similarly, a mouse model with reduced expression of GAS7 exhibits a motor phenotype that is associated with loss of motor neurons upon aging[329]. However, the extent to which GAS7 loss in neurons versus Schwann cells contribute to the phenotype are unknown. Moreover, this genetic lesion affects the expression of every GAS7 protein isoform. Therefore, a similar approach involving the deletion of the SOX10-regulated promoter and associated exon 1B will define the specific contribution of this *Gas7* transcript (and each of the protein isoforms it produces) to any resulting phenotypes.

An important consideration regarding these experiments is the cell-type specificity of the SOX10-regulated TSS. It may be that SOX10 regulates a promoter in Schwann cells that is also active and important in other cell types. In this case the constitutive deletion of the SOX10-regulated promoter and first exon may cause broader phenotypes arising from loss of the transcript in other SOX10-positive or –negative cells. Therefore, it may be worth considering the design of conditional alleles. By flanking a promoter and first exon with loxP sequences, deletion of the region could be performed in a ubiquitous or cell type-restricted manner, depending on the selection of Cre-expressing lines or Cre delivery methods. While the maintenance of the model is more involved in this system, it may be worth the effort to compare phenotypes from ubiquitous versus Schwann cell-specific deletions, and thereby determine definitively whether a defect arises from loss of expression in Schwann cells. This is likely to be particularly relevant for transcripts that are known to be expressed in peripheral neurons.

Concluding Remarks

Importantly, there are many more possible applications of the work presented in this dissertation. Most directly, there are many SOX10-regulated promoters identified by these efforts that remain to be validated and characterized. Additionally, the promoter elements that we

have identified could be subjected to high-throughput mutagenesis to assess the sequence elements surrounding SOX10 binding motifs that are required for regulatory activity or to identify motifs for other transcription factors that may co-regulate these promoters with SOX10. Finally, the elements that we have identified could be screened in patient populations to determine if sequence variation at SOX10-regulated promoters contributes to or modifies demyelinating disease.

The major aim of the work presented in this dissertation has been to identify novel genes and gene products that are important for Schwann cell biology. To achieve this, we have taken multiple approaches. These include a sequence conservation-based analysis of transcriptional regulation for a locus that is known to be important for Schwann cells. These efforts revealed the SOX10-mediated expression of a novel isoform that may provide important insights into Schwann cell function and disease. Subsequently, a genome-wide approach was employed to leverage the importance of SOX10 as a means of identifying and prioritizing gene products that may be important for Schwann cell biology. Importantly, these studies provide the field with a comprehensive assessment of TSS use in multiple experimental models relevant for Schwann cell biology. Based on those data, we validated SOX10-regulated promoters at four previously unreported SOX10 target genes that had not been implicated in myelination. Indeed, at each of these loci SOX10 mediates expression of particular gene isoforms and we anticipate that the functional characterization of these isoforms will catalyze important insights into peripheral nerve biology. It is the author's hope that with further study, these efforts may provide incremental but critical advances toward a more complete understanding of Schwann cell development, function, and disease.

References

1.   CRICK, F.H., *On protein synthesis.* Symp Soc Exp Biol, 1958. **12**: p. 138-63.

2.   WATSON, J.D. and F.H. CRICK, *Genetical implications of the structure of deoxyribonucleic acid.* Nature, 1953. **171**(4361): p. 964-7.

3.   WATSON, J.D. and F.H. CRICK, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-8.

4.   Nevins, J.R., *The pathway of eukaryotic mRNA formation.* Annu Rev Biochem, 1983. **52**: p. 441-66.

5.   Ziff, E.B. and R.M. Evans, *Coincidence of the promoter and capped 5' terminus of RNA from the adenovirus 2 major late transcription unit.* Cell, 1978. **15**(4): p. 1463-75.

6.   Konkel, D.A., S.M. Tilghman, and P. Leder, *The sequence of the chromosomal mouse beta-globin major gene: homologies in capping, splicing and poly(A) sites.* Cell, 1978. **15**(4): p. 1125-32.

7.   Tonegawa, S., et al., *Sequence of a mouse germ-line gene for a variable region of an immunoglobulin light chain.* Proc Natl Acad Sci U S A, 1978. **75**(3): p. 1485-9.

8.   Pribnow, D., *Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter.* Proc Natl Acad Sci U S A, 1975. **72**(3): p. 784-8.

9.   Wasylyk, B., *Transcription elements and factors of RNA polymerase B promoters of higher eukaryotes.* CRC Crit Rev Biochem, 1988. **23**(2): p. 77-120.

10.   Gershenzon, N.I. and I.P. Ioshikhes, *Synergy of human Pol II core promoter elements revealed by statistical sequence analysis.* Bioinformatics, 2005. **21**(8): p. 1295-300.

11.   Smale, S.T. and D. Baltimore, *The "initiator" as a transcription control element.* Cell, 1989. **57**(1): p. 103-13.

12.   Burke, T.W. and J.T. Kadonaga, *Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters.* Genes Dev, 1996. **10**(6): p. 711-24.

13.   McClelland, M. and R. Ivarie, *Asymmetrical distribution of CpG in an 'average' mammalian gene.* Nucleic Acids Res, 1982. **10**(23): p. 7865-77.

14.   Saxonov, S., P. Berg, and D.L. Brutlag, *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters.* Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1412-7.

15.   Noonan, J.P. and A.S. McCallion, *Genomics of long-range regulatory elements.* Annu Rev Genomics Hum Genet, 2010. **11**: p. 1-23.

16.   Ramirez-Carrozzi, V.R., et al., *A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling.* Cell, 2009. **138**(1): p. 114-28.

17.   Porrua, O., M. Boudvillain, and D. Libri, *Transcription Termination: Variations on Common Themes.* Trends Genet, 2016. **32**(8): p. 508-522.

18.   Grünberg, S. and S. Hahn, *Structural insights into transcription initiation by RNA polymerase II.* Trends Biochem Sci, 2013. **38**(12): p. 603-11.

19.   Gilmour, D.S. and J.T. Lis, *RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells.* Mol Cell Biol, 1986. **6**(11): p. 3984-9.

20.   Rougvie, A.E. and J.T. Lis, *The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged.* Cell, 1988. **54**(6): p. 795-804.

21.   Core, L. and K. Adelman, *Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation.* Genes Dev, 2019. **33**(15-16): p. 960-982.

22.   Vos, S.M., et al., *Structure of paused transcription complex Pol II-DSIF-NELF.* Nature, 2018. **560**(7720): p. 601-606.

23.   Yamada, T., et al., *P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation.* Mol Cell, 2006. **21**(2): p. 227-37.

24.   Hübner, M.R., M.A. Eckersley-Maslin, and D.L. Spector, *Chromatin organization and transcriptional regulation.* Curr Opin Genet Dev, 2013. **23**(2): p. 89-95.

25.   Kornberg, R.D. and J.O. Thomas, *Chromatin structure; oligomers of the histones.* Science, 1974. **184**(4139): p. 865-8.

26.   Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA.* Science, 1974. **184**(4139): p. 868-71.

27.   Han, M. and M. Grunstein, *Nucleosome loss activates yeast downstream promoters in vivo.* Cell, 1988. **55**(6): p. 1137-45.

28.   Côté, J., et al., *Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex.* Science, 1994. **265**(5168): p. 53-60.

29.   Imbalzano, A.N., et al., *Facilitated binding of TATA-binding protein to nucleosomal DNA.* Nature, 1994. **370**(6489): p. 481-5.

30.   Kwon, H., et al., *Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex.* Nature, 1994. **370**(6489): p. 477-81.

31.   Brownell, J.E., et al., *Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation.* Cell, 1996. **84**(6): p. 843-51.

32.   Tzelepis, I., M. Martino, and A. Göndör, *A Brief Introduction to Chromatin Regulation and Dynamics*, in *Chromatin Regulation and Dynamics*, A. Göndör, Editor. 2017 Elsevier: Amsterdam. p. 1-34.

33.   Bannister, A.J., A.M. Falcão, and G. Castelo-Branco, *Histone Modifications and Histone Variants in Pluripotency and Differentiation*, in *Chromatin Regulation and Dynamics*, A. Göndör, Editor. 2017 Elsevier: Amsterdam. p. 35-64.

34.   Carninci, P., et al., *The transcriptional landscape of the mammalian genome.* Science, 2005. **309**(5740): p. 1559-63.

35.   Pan, Q., et al., *Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing.* Nat Genet, 2008. **40**(12): p. 1413-5.

36.   Gustincich, S., et al., *The complexity of the mammalian transcriptome.* J Physiol, 2006. **575**(Pt 2): p. 321-32.

37.   Lee, Y. and D.C. Rio, *Mechanisms and Regulation of Alternative Pre-mRNA Splicing.* Annu Rev Biochem, 2015. **84**: p. 291-323.

38.   Braunschweig, U., et al., *Dynamic integration of splicing within gene regulatory pathways.* Cell, 2013. **152**(6): p. 1252-69.

39.   Nilsen, T.W. and B.R. Graveley, *Expansion of the eukaryotic proteome by alternative splicing.* Nature, 2010. **463**(7280): p. 457-63.

40.   Licatalosi, D.D., et al., *HITS-CLIP yields genome-wide insights into brain alternative RNA processing.* Nature, 2008. **456**(7221): p. 464-9.

41.   Wang, Z., et al., *iCLIP predicts the dual splicing effects of TIA-RNA interactions.* PLoS Biol, 2010. **8**(10): p. e1000530.

42.   Wang, E.T., et al., *Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins.* Cell, 2012. **150**(4): p. 710-24.

43.   Cramer, P., et al., *Functional association between promoter structure and transcript alternative splicing.* Proc Natl Acad Sci U S A, 1997. **94**(21): p. 11456-60.

44.   Saldi, T., et al., *Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing.* J Mol Biol, 2016. **428**(12): p. 2623-2635.

45.   Turner, R.E., A.D. Pattison, and T.H. Beilharz, *Alternative polyadenylation in the regulation and dysregulation of gene expression.* Semin Cell Dev Biol, 2018. **75**: p. 61-69.

46.   Li, W., et al., *Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation.* PLoS Genet, 2015. **11**(4): p. e1005166.

47.   Tian, B., Z. Pan, and J.Y. Lee, *Widespread mRNA polyadenylation events in introns indicate dynamic interplay between polyadenylation and splicing.* Genome Res, 2007. **17**(2): p. 156-65.

48.   Proudfoot, N.J. and G.G. Brownlee, *3' non-coding region sequences in eukaryotic messenger RNA.* Nature, 1976. **263**(5574): p. 211-4.

49.   Tian, B., et al., *A large-scale analysis of mRNA polyadenylation of human and mouse genes.* Nucleic Acids Res, 2005. **33**(1): p. 201-12.

50.   Hu, J., et al., *Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation.* RNA, 2005. **11**(10): p. 1485-93.

51.   Gil, A. and N.J. Proudfoot, *Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA 3' end formation.* Cell, 1987. **49**(3): p. 399-406.

52.   Takagaki, Y. and J.L. Manley, *Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation.* Mol Cell, 1998. **2**(6): p. 761-71.

53.   Kim Guisbert, K.S., H. Li, and C. Guthrie, *Alternative 3' pre-mRNA processing in Saccharomyces cerevisiae is modulated by Nab4/Hrp1 in vivo.* PLoS Biol, 2007. **5**(1): p. e6.

54.   Young, R.A., O. Hagenbüchle, and U. Schibler, *A single mouse alpha-amylase gene specifies two different tissue-specific mRNAs.* Cell, 1981. **23**(2): p. 451-8.

55.   Landry, J.R., D.L. Mager, and B.T. Wilhelm, *Complex controls: the role of alternative promoters in mammalian genomes.* Trends Genet, 2003. **19**(11): p. 640-8.

56.   Wang, Y., et al., *RNA diversity has profound effects on the translation of neuronal nitric oxide synthase.* Proc Natl Acad Sci U S A, 1999. **96**(21): p. 12150-5.

57.   Logette, E., et al., *The human caspase-2 gene: alternative promoters, pre-mRNA splicing and AUG usage direct isoform-specific expression.* Oncogene, 2003. **22**(6): p. 935-46.

58.   Pagani, F., et al., *Promoter architecture modulates CFTR exon 9 skipping.* J Biol Chem, 2003. **278**(3): p. 1511-7.

59.   Xin, D., L. Hu, and X. Kong, *Alternative promoters influence alternative splicing at the genomic level.* PLoS One, 2008. **3**(6): p. e2377.

60.   Monsalve, M., et al., *Direct coupling of transcription and mRNA processing through the thermogenic coactivator PGC-1.* Mol Cell, 2000. **6**(2): p. 307-16.

61.   de la Mata, M., et al., *A slow RNA polymerase II affects alternative splicing in vivo.* Mol Cell, 2003. **12**(2): p. 525-32.

62.   Eperon, L.P., et al., *Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase?* Cell, 1988. **54**(3): p. 393-401.

63.   Kornblihtt, A.R., *Promoter usage and alternative splicing.* Curr Opin Cell Biol, 2005. **17**(3): p. 262-8.

64.   Goodrich, J.A. and R. Tjian, *Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation.* Nat Rev Genet, 2010. **11**(8): p. 549-58.

65.   Heinz, S., et al., *The selection and function of cell type-specific enhancers.* Nat Rev Mol Cell Biol, 2015. **16**(3): p. 144-54.

66.   Porter, R.S., F. Jaamour, and S. Iwase, *Neuron-specific alternative splicing of transcriptional machineries: Implications for neurodevelopmental disorders.* Mol Cell Neurosci, 2018. **87**: p. 35-45.

67.   Meier, K. and A. Brehm, *Chromatin regulation: how complex does it get?* Epigenetics, 2014. **9**(11): p. 1485-95.

68.   Pal, S., et al., *Alternative transcription exceeds alternative splicing in generating the transcriptome diversity of cerebellar development.* Genome Res, 2011. **21**(8): p. 1260-72.

69.   Shabalina, S.A., et al., *Evolution at protein ends: major contribution of alternative transcription initiation and termination to the transcriptome and proteome diversity in mammals.* Nucleic Acids Res, 2014. **42**(11): p. 7132-44.

70.   Reyes, A. and W. Huber, *Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues.* Nucleic Acids Res, 2018. **46**(2): p. 582-592.

71.   Forrest, A.R., et al., *A promoter-level mammalian expression atlas.* Nature, 2014. **507**(7493): p. 462-70.

72.   Theveneau, E. and R. Mayor, *Neural crest delamination and migration: from epithelium-to-mesenchyme transition to collective cell migration.* Dev Biol, 2012. **366**(1): p. 34-54.

73.   Dong, Z., et al., *Neu differentiation factor is a neuron-glia signal and regulates survival, proliferation, and maturation of rat Schwann cell precursors.* Neuron, 1995. **15**(3): p. 585-96.

74.   Riethmacher, D., et al., *Severe neuropathies in mice with targeted mutations in the ErbB3 receptor.* Nature, 1997. **389**(6652): p. 725-30.

75.   Meier, C., et al., *Developing Schwann cells acquire the ability to survive without axons by establishing an autocrine circuit involving insulin-like growth factor, neurotrophin-3, and platelet-derived growth factor-BB.* J Neurosci, 1999. **19**(10): p. 3847-59.

76.   Jessen, K.R. and R. Mirsky, *The origin and development of glial cells in peripheral nerves.* Nat Rev Neurosci, 2005. **6**(9): p. 671-82.

77.   Woodhoo, A., et al., *Notch controls embryonic Schwann cell differentiation, postnatal myelination and adult plasticity.* Nat Neurosci, 2009. **12**(7): p. 839-47.

78.   Mukouyama, Y.S., et al., *Peripheral nerve-derived VEGF promotes arterial differentiation via neuropilin 1-mediated positive feedback.* Development, 2005. **132**(5): p. 941-52.

79.   Parmantier, E., et al., *Schwann cell-derived Desert hedgehog controls the development of peripheral nerve sheaths.* Neuron, 1999. **23**(4): p. 713-24.

80.   Friede, R.L., *Control of myelin formation by axon caliber (with a model of the control mechanism).* J Comp Neurol, 1972. **144**(2): p. 233-52.

81.   Webster, H.D., R. Martin, and M.F. O'Connell, *The relationships between interphase Schwann cells and axons before myelination: a quantitative electron microscopic study.* Dev Biol, 1973. **32**(2): p. 401-16.

82.   Monk, K.R., M.L. Feltri, and C. Taveggia, *New insights on Schwann cell development.* Glia, 2015. **63**(8): p. 1376-93.

83.   Newbern, J. and C. Birchmeier, *Nrg1/ErbB signaling networks in Schwann cell development and myelination.* Semin Cell Dev Biol, 2010. **21**(9): p. 922-8.

84.   Garratt, A.N., et al., *A dual role of erbB2 in myelination and in expansion of the schwann cell precursor pool.* J Cell Biol, 2000. **148**(5): p. 1035-46.

85.   Taveggia, C., et al., *Neuregulin-1 type III determines the ensheathment fate of axons.* Neuron, 2005. **47**(5): p. 681-94.

86.   Ozkaynak, E., et al., *Adam22 is a major neuronal receptor for Lgi4-mediated Schwann cell signaling.* J Neurosci, 2010. **30**(10): p. 3857-64.

87.   Monk, K.R., et al., *A G protein-coupled receptor is essential for Schwann cells to initiate myelination.* Science, 2009. **325**(5946): p. 1402-5.

88.   Trimarco, A., et al., *Prostaglandin D2 synthase/GPR44: a signaling axis in PNS myelination.* Nat Neurosci, 2014. **17**(12): p. 1682-92.

89.   Anliker, B., et al., *Lysophosphatidic acid (LPA) and its receptor, LPA1 , influence embryonic schwann cell migration, myelination, and cell-to-axon segregation.* Glia, 2013. **61**(12): p. 2009-22.

90.   Wang, S.S., et al., *Functional trade-offs in white matter axonal scaling.* J Neurosci, 2008. **28**(15): p. 4047-56.

91.   Hartline, D.K. and D.R. Colman, *Rapid conduction and the evolution of giant axons and myelinated fibers.* Curr Biol, 2007. **17**(1): p. R29-35.

92.   Harris, J.J. and D. Attwell, *The energetics of CNS white matter.* J Neurosci, 2012. **32**(1): p. 356-71.

93.   Nave, K.A. and H.B. Werner, *Myelination of the nervous system: mechanisms and functions.* Annu Rev Cell Dev Biol, 2014. **30**: p. 503-33.

94.   Nave, K.A., *Myelination and the trophic support of long axons.* Nat Rev Neurosci, 2010. **11**(4): p. 275-83.

95.   Beirowski, B., et al., *Metabolic regulator LKB1 is crucial for Schwann cell-mediated axon maintenance.* Nat Neurosci, 2014. **17**(10): p. 1351-61.

96.   Court, F.A., et al., *Schwann cell to axon transfer of ribosomes: toward a novel understanding of the role of glia in the nervous system.* J Neurosci, 2008. **28**(43): p. 11024-9.

97.   Wilkins, A., et al., *Oligodendrocytes promote neuronal survival and axonal length by distinct intracellular mechanisms: a novel role for oligodendrocyte-derived glial cell line-derived neurotrophic factor.* J Neurosci, 2003. **23**(12): p. 4967-74.

98.   Lopez-Verrilli, M.A., F. Picou, and F.A. Court, *Schwann cell-derived exosomes enhance axonal regeneration in the peripheral nervous system.* Glia, 2013. **61**(11): p. 1795-806.

99.   Frühbeis, C., et al., *Neurotransmitter-triggered transfer of exosomes mediates oligodendrocyte-neuron communication.* PLoS Biol, 2013. **11**(7): p. e1001604.

100.   Krämer-Albers, E.M., et al., *Oligodendrocytes secrete exosomes containing major myelin and stress-protective proteins: Trophic support for axons?* Proteomics Clin Appl, 2007. **1**(11): p. 1446-61.

101.   Katona, I. and J. Weis, *Diseases of the peripheral nerves.* Handb Clin Neurol, 2017. **145**: p. 453-474.

102.   Martyn, C.N. and R.A. Hughes, *Epidemiology of peripheral neuropathy.* J Neurol Neurosurg Psychiatry, 1997. **62**(4): p. 310-8.

103.   Saporta, A.S., et al., *Charcot-Marie-Tooth disease subtypes and genetic testing strategies.* Ann Neurol, 2011. **69**(1): p. 22-33.

104.   Dyck, P.J. and E.H. Lambert, *Lower motor and primary sensory neuron diseases with peroneal muscular atrophy. I. Neurologic, genetic, and electrophysiologic findings in hereditary polyneuropathies.* Archives of neurology, 1968. **18**(6): p. 603-18.

105.   Gess, B., et al., *Charcot-Marie-Tooth disease: frequency of genetic subtypes in a German neuromuscular center population.* Neuromuscul Disord, 2013. **23**(8): p. 647-51.

106.   Murphy, S.M., M. Laurá, and M.M. Reilly, *DNA testing in hereditary neuropathies.* Handb Clin Neurol, 2013. **115**: p. 213-32.

107.   Brennan, K.M., Y. Bai, and M.E. Shy, *Demyelinating CMT--what's known, what's new and what's in store?* Neurosci Lett, 2015. **596**: p. 14-26.

108.   Kamachi, Y. and H. Kondoh, *Sox proteins: regulators of cell fate specification and differentiation.* Development, 2013. **140**(20): p. 4129-44.

109.   Reményi, A., et al., *Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers.* Genes Dev, 2003. **17**(16): p. 2048-59.

110.   Bowles, J., G. Schepers, and P. Koopman, *Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators.* Dev Biol, 2000. **227**(2): p. 239-55.

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

112.   Wright, E.M., B. Snopek, and P. Koopman, *Seven new members of the Sox gene family expressed during mouse development.* Nucleic Acids Res, 1993. **21**(3): p. 744.

113.   Kuhlbrodt, K., et al., *Sox10, a novel transcriptional modulator in glial cells.* J Neurosci, 1998. **18**(1): p. 237-50.

114.   Cheung, M. and J. Briscoe, *Neural crest development is regulated by the transcription factor Sox9.* Development, 2003. **130**(23): p. 5681-93.

115.   Southard-Smith, E.M., L. Kos, and W.J. Pavan, *Sox10 mutation disrupts neural crest development in Dom Hirschsprung mouse model.* Nat Genet, 1998. **18**(1): p. 60-4.

116.   Kim, J., et al., *SOX10 maintains multipotency and inhibits neuronal differentiation of neural crest stem cells.* Neuron, 2003. **38**(1): p. 17-31.

117. Stolt, C.C. and M. Wegner, *SoxE function in vertebrate nervous system development.* Int J Biochem Cell Biol, 2010. **42**(3): p. 437-40.

118. Mollaaghababa, R. and W.J. Pavan, *The importance of having your SOX on: role of SOX10 in the development of neural crest-derived melanocytes and glia.* Oncogene, 2003. **22**(20): p. 3024-34.

119. Stolt, C.C., et al., *Terminal differentiation of myelin-forming oligodendrocytes depends on the transcription factor Sox10.* Genes Dev, 2002. **16**(2): p. 165-70.

120. Harris, M.L., et al., *A dual role for SOX10 in the maintenance of the postnatal melanocyte lineage and the differentiation of melanocyte stem cell progenitors.* PLoS Genet, 2013. **9**(7): p. e1003644.

121. Peirano, R.I. and M. Wegner, *The glial transcription factor Sox10 binds to DNA both as monomer and dimer with different functional consequences.* Nucleic Acids Res, 2000. **28**(16): p. 3047-55.

122. Stolt, C.C., et al., *Transcription factors Sox8 and Sox10 perform non-equivalent roles during oligodendrocyte development despite functional redundancy.* Development, 2004. **131**(10): p. 2349-58.

123. Huang, Y.H., et al., *SOXE transcription factors form selective dimers on non-compact DNA motifs through multifaceted interactions between dimerization and high-mobility group domains.* Sci Rep, 2015. **5**: p. 10398.

124. Harley, V.R. and P.N. Goodfellow, *The biochemical role of SRY in sex determination.* Mol Reprod Dev, 1994. **39**(2): p. 184-93.

125. Schepers, G.E., et al., *Cloning and characterisation of the Sry-related transcription factor gene Sox8.* Nucleic Acids Res, 2000. **28**(6): p. 1473-80.

126. Schreiner, S., et al., *Hypomorphic Sox10 alleles reveal novel protein functions and unravel developmental differences in glial lineages.* Development, 2007. **134**(18): p. 3271-81.

127. Pusch, C., et al., *The SOX10/Sox10 gene from human and mouse: sequence, expression, and transactivation by the encoded HMG domain transcription factor.* Hum Genet, 1998. **103**(2): p. 115-23.

128. Haseeb, A. and V. Lefebvre, *The SOXE transcription factors-SOX8, SOX9 and SOX10-share a bi-partite transactivation mechanism.* Nucleic Acids Res, 2019. **47**(13): p. 6917-6931.

129. Wegner, M., *All purpose Sox: The many roles of Sox proteins in gene expression.* Int J Biochem Cell Biol, 2010. **42**(3): p. 381-90.

130. Tolhuis, B., et al., *Looping and interaction between hypersensitive sites in the active beta-globin locus.* Mol Cell, 2002. **10**(6): p. 1453-65.

131. Wissmüller, S., et al., *The high-mobility-group domain of Sox proteins interacts with DNA-binding domains of many transcription factors.* Nucleic Acids Res, 2006. **34**(6): p. 1735-44.

132. Werner, M.H., et al., *Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex.* Cell, 1995. **81**(5): p. 705-14.

133. Zhou, R., et al., *SOX9 interacts with a component of the human thyroid hormone receptor-associated protein complex.* Nucleic Acids Res, 2002. **30**(14): p. 3245-52.

134. Vogl, M.R., et al., *Sox10 cooperates with the mediator subunit 12 during terminal differentiation of myelinating glia.* J Neurosci, 2013. **33**(15): p. 6679-90.

135. Malik, S. and R.G. Roeder, *The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation.* Nat Rev Genet, 2010. **11**(11): p. 761-72.

136. Arter, J. and M. Wegner, *Transcription factors Sox10 and Sox2 functionally interact with positive transcription elongation factor b in Schwann cells.* J Neurochem, 2015. **132**(4): p. 384-93.

137. Weider, M., et al., *Chromatin-remodeling factor Brg1 is required for Schwann cell differentiation and myelination.* Dev Cell, 2012. **23**(1): p. 193-201.

138. Jacob, C., et al., *HDAC1 and HDAC2 control the transcriptional program of myelination and the survival of Schwann cells.* Nat Neurosci, 2011. **14**(4): p. 429-36.

139. Inoue, K., Y. Tanabe, and J.R. Lupski, *Myelin deficiencies in both the central and the peripheral nervous systems associated with a SOX10 mutation.* Ann Neurol, 1999. **46**(3): p. 313-8.

140. Inoue, K., et al., *Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations.* Nat Genet, 2004. **36**(4): p. 361-9.

141. Pingault, V., et al., *SOX10 mutations in patients with Waardenburg-Hirschsprung disease.* Nat Genet, 1998. **18**(2): p. 171-3.

142. Pingault, V., et al., *Review and update of mutations causing Waardenburg syndrome.* Hum Mutat, 2010. **31**(4): p. 391-406.

143. Southard-Smith, E.M., et al., *The Sox10(Dom) mouse: modeling the genetic variation of Waardenburg-Shah (WS4) syndrome.* Genome Res, 1999. **9**(3): p. 215-25.

144.   Owens, S.E., et al., *Genome-wide linkage identifies novel modifier loci of aganglionosis in the Sox10Dom model of Hirschsprung disease.* Hum Mol Genet, 2005. **14**(11): p. 1549-58.

145.   Walters, L.C., et al., *Genetic background impacts developmental potential of enteric neural crest-derived progenitors in the Sox10Dom model of Hirschsprung disease.* Hum Mol Genet, 2010. **19**(22): p. 4353-72.

146.   Truch, K., et al., *Analysis of the human SOX10 mutation Q377X in mice and its implications for genotype-phenotype correlation in SOX10-related human disease.* Hum Mol Genet, 2018. **27**(6): p. 1078-1092.

147.   Finzsch, M., et al., *Sox10 is required for Schwann cell identity and progression beyond the immature Schwann cell stage.* J Cell Biol, 2010. **189**(4): p. 701-12.

148.   Fröb, F., et al., *Establishment of myelinating Schwann cells and barrier integrity between central and peripheral nervous systems depend on Sox10.* Glia, 2012. **60**(5): p. 806-19.

149.   Bremer, M., et al., *Sox10 is required for Schwann-cell homeostasis and myelin maintenance in the adult peripheral nerve.* Glia, 2011. **59**(7): p. 1022-32.

150.   Prasad, M.K., et al., *SOX10 directly modulates ERBB3 transcription via an intronic neural crest enhancer.* BMC Dev Biol, 2011. **11**: p. 40.

151.   Jagalur, N.B., et al., *Functional dissection of the Oct6 Schwann cell enhancer reveals an essential role for dimeric Sox10 binding.* J Neurosci, 2011. **31**(23): p. 8585-94.

152.   Ghazvini, M., et al., *A cell type-specific allele of the POU gene Oct-6 reveals Schwann cell autonomous function in nerve development and regeneration.* EMBO J, 2002. **21**(17): p. 4612-20.

153.   Ghislain, J. and P. Charnay, *Control of myelination in Schwann cells: a Krox20 cis-regulatory element integrates Oct6, Brn2 and Sox10 activities.* EMBO Rep, 2006. **7**(1): p. 52-8.

154.   Srinivasan, R., et al., *Genome-wide analysis of EGR2/SOX10 binding in myelinating peripheral nerve.* Nucleic Acids Res, 2012. **40**(14): p. 6449-60.

155.   Topilko, P., et al., *Krox-20 controls myelination in the peripheral nervous system.* Nature, 1994. **371**(6500): p. 796-9.

156.   Decker, L., et al., *Peripheral myelin maintenance is a dynamic process requiring constant Krox20 expression.* J Neurosci, 2006. **26**(38): p. 9771-9.

157.   Jones, E.A., et al., *Interactions of Sox10 and Egr2 in myelin gene regulation.* Neuron Glia Biol, 2007. **3**(4): p. 377-87.

158. Forghani, R., et al., *A distal upstream enhancer from the myelin basic protein gene regulates expression in myelin-forming schwann cells.* J Neurosci, 2001. **21**(11): p. 3780-7.

159. LeBlanc, S.E., R.M. Ward, and J. Svaren, *Neuropathy-associated Egr2 mutants disrupt cooperative activation of myelin protein zero by Egr2 and Sox10.* Mol Cell Biol, 2007. **27**(9): p. 3521-9.

160. Li, H., et al., *Olig1 and Sox10 interact synergistically to drive myelin basic protein transcription in oligodendrocytes.* J Neurosci, 2007. **27**(52): p. 14375-82.

161. Peirano, R.I., et al., *Protein zero gene expression is regulated by the glial transcription factor Sox10.* Mol Cell Biol, 2000. **20**(9): p. 3198-209.

162. Bondurand, N., et al., *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, 2001. **10**(24): p. 2783-95.

163. Reiprich, S., et al., *Activation of Krox20 gene expression by Sox10 in myelinating Schwann cells.* J Neurochem, 2010. **112**(3): p. 744-54.

164. Jones, E.A., et al., *Regulation of the PMP22 gene through an intronic enhancer.* J Neurosci, 2011. **31**(11): p. 4242-50.

165. Brewer, M.H., et al., *Haplotype-specific modulation of a SOX10/CREB response element at the Charcot-Marie-Tooth disease type 4C locus SH3TC2.* Hum Mol Genet, 2014. **23**(19): p. 5171-87.

166. Graf, S.A., et al., *The myelin protein PMP2 is regulated by SOX10 and drives melanoma cell invasion.* Pigment Cell Melanoma Res, 2019. **32**(3): p. 424-434.

167. Fujiwara, S., et al., *SOX10 transactivates S100B to suppress Schwann cell proliferation and to promote myelination.* PLoS One, 2014. **9**(12): p. e115400.

168. Ito, Y., et al., *Sox10 regulates ciliary neurotrophic factor gene expression in Schwann cells.* Proc Natl Acad Sci U S A, 2006. **103**(20): p. 7871-6.

169. Gopinath, C., et al., *Stringent comparative sequence analysis reveals SOX10 as a putative inhibitor of glial cell differentiation.* BMC Genomics, 2016. **17**(1): p. 887.

170. Hodonsky, C.J., et al., *SOX10 regulates expression of the SH3-domain kinase binding protein 1 (Sh3kbp1) locus in Schwann cells via an alternative promoter.* Mol Cell Neurosci, 2012. **49**(2): p. 85-96.

171. Gokey, N.G., et al., *Developmental regulation of microRNA expression in Schwann cells.* Mol Cell Biol, 2012. **32**(2): p. 558-68.

172.   Zhao, X., et al., *MicroRNA-mediated control of oligodendrocyte differentiation.* Neuron, 2010. **65**(5): p. 612-26.

173.   Ionasescu, V., C. Searby, and R. Ionasescu, *Point mutations of the connexin32 (GJB1) gene in X-linked dominant Charcot-Marie-Tooth neuropathy.* Hum Mol Genet, 1994. **3**(2): p. 355-8.

174.   Flagiello, L., et al., *Mutation in the nerve-specific 5'non-coding region of Cx32 gene and absence of specific mRNA in a CMTX1 Italian family. Mutations in brief no. 195. Online.* Hum Mutat, 1998. **12**(5): p. 361.

175.   Houlden, H., et al., *Connexin 32 promoter P2 mutations: a mechanism of peripheral nerve dysfunction.* Ann Neurol, 2004. **56**(5): p. 730-4.

176.   Kulshrestha, R., et al., *Deletion of P2 promoter of GJB1 gene a cause of Charcot-Marie-Tooth disease.* Neuromuscul Disord, 2017. **27**(8): p. 766-770.

177.   Wang, H.L., et al., *Point mutation associated with X-linked dominant Charcot-Marie-Tooth disease impairs the P2 promoter activity of human connexin-32 gene.* Brain Res Mol Brain Res, 2000. **78**(1-2): p. 146-53.

178.   Skre, H., *Genetic and clinical aspects of Charcot-Marie-Tooth's disease.* Clinical genetics, 1974. **6**(2): p. 98-118.

179.   Dyck, P.J. and E.H. Lambert, *Lower motor and primary sensory neuron diseases with peroneal muscular atrophy. II. Neurologic, genetic, and electrophysiologic findings in various neuronal degenerations.* Archives of neurology, 1968. **18**(6): p. 619-25.

180.   Harding, A.E. and P.K. Thomas, *The clinical features of hereditary motor and sensory neuropathy types I and II.* Brain : a journal of neurology, 1980. **103**(2): p. 259-80.

181.   Brennan, K.M., Y. Bai, and M.E. Shy, *Demyelinating CMT-what's known, what's new and what's in store?* Neuroscience letters, 2015. **596**: p. 14-26.

182.   Su, Y., et al., *Myelin protein zero gene mutated in Charcot-Marie-tooth type 1B patients.* Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(22): p. 10856-60.

183.   Warner, L.E., et al., *Mutations in the early growth response 2 (EGR2) gene are associated with hereditary myelinopathies.* Nature genetics, 1998. **18**(4): p. 382-4.

184.   Bolino, A., et al., *Charcot-Marie-Tooth type 4B is caused by mutations in the gene encoding myotubularin-related protein-2.* Nature genetics, 2000. **25**(1): p. 17-9.

185.   Gambardella, A., et al., *Autosomal recessive hereditary motor and sensory neuropathy with focally folded myelin sheaths (CMT4B).* Annals of the New York Academy of Sciences, 1999. **883**: p. 47-55.

186. Bolis, A., et al., *Loss of Mtmr2 phosphatase in Schwann cells but not in motor neurons causes Charcot-Marie-Tooth type 4B1 neuropathy with myelin outfoldings.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2005. **25**(37): p. 8567-77.

187. Patzig, J., et al., *Quantitative and integrative proteome analysis of peripheral nerve myelin identifies novel myelin proteins and candidate neuropathy loci.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2011. **31**(45): p. 16369-86.

188. Chojnowski, A., et al., *Silencing of the Charcot-Marie-Tooth associated MTMR2 gene decreases proliferation and enhances cell death in primary cultures of Schwann cells.* Neurobiology of disease, 2007. **26**(2): p. 323-31.

189. Vaccari, I., et al., *Genetic interaction between MTMR2 and FIG4 phospholipid phosphatases involved in Charcot-Marie-Tooth neuropathies.* PLoS genetics, 2011. **7**(10): p. e1002319.

190. Berger, P., et al., *Loss of phosphatase activity in myotubularin-related protein 2 is associated with Charcot-Marie-Tooth disease type 4B1.* Human molecular genetics, 2002. **11**(13): p. 1569-79.

191. Begley, M.J. and J.E. Dixon, *The structure and regulation of myotubularin phosphatases.* Current opinion in structural biology, 2005. **15**(6): p. 614-20.

192. Bolis, A., et al., *Myotubularin-related (MTMR) phospholipid phosphatase proteins in the peripheral nervous system.* Molecular neurobiology, 2007. **35**(3): p. 308-16.

193. Previtali, S.C., et al., *Myotubularin-related 2 protein phosphatase and neurofilament light chain protein, both mutated in CMT neuropathies, interact in peripheral nerve.* Human molecular genetics, 2003. **12**(14): p. 1713-23.

194. Bremer, M., et al., *Sox10 is required for Schwann-cell homeostasis and myelin maintenance in the adult peripheral nerve.* Glia, 2011. **59**(7): p. 1022-32.

195. Fogarty, E.A., et al., *SOX10 regulates an alternative promoter at the Charcot-Marie-Tooth disease locus MTMR2.* Hum Mol Genet, 2016. **25**(18): p. 3925-3936.

196. Kent, W.J., et al., *The human genome browser at UCSC.* Genome research, 2002. **12**(6): p. 996-1006.

197. Antonellis, A., et al., *Deletion of long-range sequences at Sox10 compromises developmental expression in a mouse model of Waardenburg-Shah (WS4) syndrome.* Human molecular genetics, 2006. **15**(2): p. 259-71.

198. Guo, F., et al., *An in vitro recombination method to convert restriction- and ligation-independent expression vectors.* Biotechnology journal, 2008. **3**(3): p. 370-7.

199.   Goda, S., et al., *Expression of the myelin-associated glycoprotein in cultures of immortalized Schwann cells.* Journal of neurochemistry, 1991. **56**(4): p. 1354-61.

200.   Salazar-Grueso, E.F., S. Kim, and H. Kim, *Embryonic mouse spinal cord motor neuron hybrid cells.* Neuroreport, 1991. **2**(9): p. 505-8.

201.   Inoue, K., et al., *Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations.* Nature genetics, 2004. **36**(4): p. 361-9.

202.   Li, J.C., et al., *Rat testicular myotubularin, a protein tyrosine phosphatase expressed by Sertoli and germ cells, is a potential marker for studying cell-cell interactions in the rat testis.* Journal of cellular physiology, 2000. **185**(3): p. 366-85.

203.   Ng, A.A., et al., *The CMT4B disease-causing phosphatases Mtmr2 and Mtmr13 localize to the Schwann cell cytoplasm and endomembrane compartments, where they depend upon each other to achieve wild-type levels of protein expression.* Human molecular genetics, 2013. **22**(8): p. 1493-506.

204.   Gokey, N.G., et al., *Developmental regulation of microRNA expression in Schwann cells.* Molecular and cellular biology, 2012. **32**(2): p. 558-68.

205.   Schmidt, D., et al., *ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions.* Methods, 2009. **48**(3): p. 240-8.

206.   Umlauf, D., Y. Goto, and R. Feil, *Site-specific analysis of histone methylation and acetylation.* Methods in molecular biology, 2004. **287**: p. 99-120.

207.   Antonellis, A. and E.D. Green, *Inter-Species Comparative Sequence Analysis: A Tool for Genomic Medicine.*, in *Genomic and Personalized Medicine*, H. Willard and G. Ginsburg, Editors. 2008, Academic Press: Salt Lake City. p. 120-130.

208.   Hai, M., et al., *Comparative analysis of Schwann cell lines as model systems for myelin gene transcription studies.* Journal of neuroscience research, 2002. **69**(4): p. 497-508.

209.   Hodonsky, C.J., et al., *SOX10 regulates expression of the SH3-domain kinase binding protein 1 (Sh3kbp1) locus in Schwann cells via an alternative promoter.* Molecular and cellular neurosciences, 2012. **49**(2): p. 85-96.

210.   Bolino, A., et al., *Molecular characterization and expression analysis of Mtmr2, mouse homologue of MTMR2, the Myotubularin-related 2 gene, mutated in CMT4B.* Gene, 2002. **283**(1-2): p. 17-26.

211.   Law, W.D., et al., *A genome-wide assessment of conserved SNP alleles reveals a panel of regulatory SNPs relevant to the peripheral nerve.* BMC Genomics, 2018. **19**(1): p. 311.

212.   Yaffe, D. and O. Saxel, *Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle.* Nature, 1977. **270**(5639): p. 725-7.

213. Franklin, N.E., G.S. Taylor, and P.O. Vacratsis, *Endosomal targeting of the phosphoinositide 3-phosphatase MTMR2 is regulated by an N-terminal phosphorylation site.* The Journal of biological chemistry, 2011. **286**(18): p. 15841-53.

214. Lopez-Anido, C., et al., *Differential Sox10 genomic occupancy in myelinating glia.* Glia, 2015. **63**(11): p. 1897-1914.

215. Jones, E.A., et al., *Interactions of Sox10 and Egr2 in myelin gene regulation.* Neuron glia biology, 2007. **3**(4): p. 377-87.

216. Srinivasan, R., et al., *Genome-wide analysis of EGR2/SOX10 binding in myelinating peripheral nerve.* Nucleic acids research, 2012. **40**(14): p. 6449-60.

217. Harley, V.R., R. Lovell-Badge, and P.N. Goodfellow, *Definition of a consensus DNA binding site for SRY.* Nucleic acids research, 1994. **22**(8): p. 1500-1.

218. Peirano, R.I. and M. Wegner, *The glial transcription factor Sox10 binds to DNA both as monomer and dimer with different functional consequences.* Nucleic acids research, 2000. **28**(16): p. 3047-55.

219. Stolt, C.C. and M. Wegner, *SoxE function in vertebrate nervous system development.* The international journal of biochemistry & cell biology, 2010. **42**(3): p. 437-40.

220. Kuhlbrodt, K., et al., *Sox10, a novel transcriptional modulator in glial cells.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 1998. **18**(1): p. 237-50.

221. Weider, M., S. Reiprich, and M. Wegner, *Sox appeal - Sox10 attracts epigenetic and transcriptional regulators in myelinating glia.* Biological chemistry, 2013. **394**(12): p. 1583-93.

222. Houlden, H., et al., *Connexin 32 promoter P2 mutations: a mechanism of peripheral nerve dysfunction.* Annals of Neurology, 2004. **56**(5): p. 730-4.

223. Antonellis, A., et al., *A rare myelin protein zero (MPZ) variant alters enhancer activity in vitro and in vivo.* PloS one, 2010. **5**(12): p. e14346.

224. Brewer, M.H., et al., *Haplotype-specific modulation of a SOX10/CREB response element at the Charcot-Marie-Tooth disease type 4C locus SH3TC2.* Human molecular genetics, 2014. **23**(19): p. 5171-87.

225. Luigetti, M., et al., *A novel homozygous mutation in the MTMR2 gene in two siblings with 'hypermyelinating neuropathy'.* Journal of the peripheral nervous system : JPNS, 2013. **18**(2): p. 192-4.

226. Raess, M.A., et al., *Expression of the neuropathy-associated MTMR2 gene rescues MTM1-associated myopathy.* Hum Mol Genet, 2017. **26**(19): p. 3736-3748.

227. Kim, S.A., et al., *Regulation of myotubularin-related (MTMR)2 phosphatidylinositol phosphatase by MTMR5, a catalytically inactive phosphatase.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(8): p. 4492-7.

228. Robinson, F.L. and J.E. Dixon, *The phosphoinositide-3-phosphatase MTMR2 associates with MTMR13, a membrane-associated pseudophosphatase also mutated in type 4B Charcot-Marie-Tooth disease.* The Journal of biological chemistry, 2005. **280**(36): p. 31699-707.

229. Laporte, J., et al., *The myotubularin family: from genetic disease to phosphoinositide metabolism.* Trends in genetics : TIG, 2001. **17**(4): p. 221-8.

230. Cui, X., et al., *Association of SET domain and myotubularin-related proteins modulates growth control.* Nature genetics, 1998. **18**(4): p. 331-7.

231. Shah, Z.H., et al., *Nuclear phosphoinositides and their impact on nuclear functions.* FEBS J, 2013. **280**(24): p. 6295-310.

232. Gozani, O., et al., *The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor.* Cell, 2003. **114**(1): p. 99-111.

233. Stijf-Bultsma, Y., et al., *The basal transcription complex component TAF3 transduces changes in nuclear phosphoinositides into transcriptional output.* Mol Cell, 2015. **58**(3): p. 453-67.

234. de Klerk, E. and P.A. 't Hoen, *Alternative mRNA transcription, processing, and translation: insights from RNA sequencing.* Trends Genet, 2015. **31**(3): p. 128-39.

235. Harding, A.E. and P.K. Thomas, *The clinical features of hereditary motor and sensory neuropathy types I and II.* Brain, 1980. **103**(2): p. 259-80.

236. Mertin, S., S.G. McDowall, and V.R. Harley, *The DNA-binding specificity of SOX9 and other SOX proteins.* Nucleic Acids Res, 1999. **27**(5): p. 1359-64.

237. Monje, P.V., *Scalable Differentiation and Dedifferentiation Assays Using Neuron-Free Schwann Cell Cultures.* Methods Mol Biol, 2018. **1739**: p. 213-232.

238. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.

239. Cole, C., et al., *Tn5Prime, a Tn5 based 5' capture method for single cell RNA-seq.* Nucleic Acids Res, 2018. **46**(10): p. e62.

240. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads.* EMBnet.journal, 2011. **17**(1): p. 10-12.

241. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.

242. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.

243. Takahashi, H., et al., *5' end-centered expression profiling using cap-analysis gene expression and next-generation sequencing.* Nat Protoc, 2012. **7**(3): p. 542-61.

244. Frith, M.C., et al., *A code for transcription initiation in mammalian genomes.* Genome Res, 2008. **18**(1): p. 1-12.

245. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.* Bioinformatics, 2014. **30**(7): p. 923-30.

246. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40.

247. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841-2.

248. Ma, K.H., H.A. Hung, and J. Svaren, *Epigenomic Regulation of Schwann Cell Reprogramming in Peripheral Nerve Injury.* J Neurosci, 2016. **36**(35): p. 9135-47.

249. Team, R.C., *R: A Language and Environment for Statistical Computing.* . 2018: Vienna, Austria.

250. Joly Beauparlant, C., et al., *metagene Profiles Analyses Reveal Regulatory Element's Factor-Specific Recruitment Patterns.* PLoS Comput Biol, 2016. **12**(8): p. e1004751.

251. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.

252. The Gene Ontology Consortium, *The Gene Ontology Resource: 20 years and still GOing strong.* Nucleic Acids Res, 2019. **47**(D1): p. D330-D338.

253. Babicki, S., et al., *Heatmapper: web-enabled heat mapping for all.* Nucleic Acids Res, 2016. **44**(W1): p. W147-53.

254. Stajich, J.E., et al., *The Bioperl toolkit: Perl modules for the life sciences.* Genome Res, 2002. **12**(10): p. 1611-8.

255. Rice, P., I. Longden, and A. Bleasby, *EMBOSS: the European Molecular Biology Open Software Suite.* Trends Genet, 2000. **16**(6): p. 276-7.

256. Lizio, M., et al., *Gateways to the FANTOM5 promoter level mammalian expression atlas.* Genome Biol, 2015. **16**: p. 22.

257. Yanai, I., et al., *Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification.* Bioinformatics, 2005. **21**(5): p. 650-9.

258. Lopez-Anido, C., et al., *Differential Sox10 genomic occupancy in myelinating glia.* Glia, 2015.

259. Scacheri, P.C., et al., *Genome-wide analysis of menin binding provides insights into MEN1 tumorigenesis.* PLoS Genet, 2006. **2**(4): p. e51.

260. Krig, S.R., et al., *Identification of genes directly regulated by the oncogene ZNF217 using chromatin immunoprecipitation (ChIP)-chip assays.* J Biol Chem, 2007. **282**(13): p. 9703-12.

261. Lizio, M., et al., *Mapping Mammalian Cell-type-specific Transcriptional Regulatory Networks Using KD-CAGE and ChIP-seq Data in the TC-YIK Cell Line.* Front Genet, 2015. **6**: p. 331.

262. Goda, S., et al., *Expression of the myelin-associated glycoprotein in cultures of immortalized Schwann cells.* J Neurochem, 1991. **56**(4): p. 1354-61.

263. Toda, K., et al., *Biochemical and cellular properties of three immortalized Schwann cell lines expressing different levels of the myelin-associated glycoprotein.* J Neurochem, 1994. **63**(5): p. 1646-57.

264. Lin, S., et al., *Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery.* Elife, 2014. **3**: p. e04766.

265. Doddrell, R.D., et al., *Loss of SOX10 function contributes to the phenotype of human Merlin-null schwannoma cells.* Brain, 2013. **136**(Pt 2): p. 549-63.

266. Siepel, A., et al., *Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes.* Genome Res, 2005. **15**(8): p. 1034-50.

267. Yang, C., et al., *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, 2007. **389**(1): p. 52-65.

268. Fenouil, R., et al., *CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters.* Genome Res, 2012. **22**(12): p. 2399-408.

269. Schug, J., et al., *Promoter features related to tissue specificity as measured by Shannon entropy.* Genome Biol, 2005. **6**(4): p. R33.

270. O'Leary, N.A., et al., *Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation.* Nucleic Acids Res, 2016. **44**(D1): p. D733-45.

271.   Goley, E.D. and M.D. Welch, *The ARP2/3 complex: an actin nucleator comes of age.* Nat Rev Mol Cell Biol, 2006. **7**(10): p. 713-26.

272.   Pizarro-Cerdá, J., et al., *The Diverse Family of Arp2/3 Complexes.* Trends Cell Biol, 2017. **27**(2): p. 93-100.

273.   Egile, C., et al., *Activation of the CDC42 effector N-WASP by the Shigella flexneri IcsA protein promotes actin nucleation by Arp2/3 complex and bacterial actin-based motility.* J Cell Biol, 1999. **146**(6): p. 1319-32.

274.   Machesky, L.M., et al., *Mammalian actin-related protein 2/3 complex localizes to regions of lamellipodial protrusion and is composed of evolutionarily conserved proteins.* Biochem J, 1997. **328 (Pt 1)**: p. 105-12.

275.   Abella, J.V., et al., *Isoform diversity in the Arp2/3 complex determines actin filament dynamics.* Nat Cell Biol, 2016. **18**(1): p. 76-86.

276.   Balcer, H.I., K. Daugherty-Clarke, and B.L. Goode, *The p40/ARPC1 subunit of Arp2/3 complex performs multiple essential roles in WASp-regulated actin nucleation.* J Biol Chem, 2010. **285**(11): p. 8481-91.

277.   Winter, D.C., E.Y. Choe, and R. Li, *Genetic dissection of the budding yeast Arp2/3 complex: a comparison of the in vivo and structural roles of individual subunits.* Proc Natl Acad Sci U S A, 1999. **96**(13): p. 7288-93.

278.   Bacon, C., et al., *N-WASP regulates extension of filopodia and processes by oligodendrocyte progenitors, oligodendrocytes, and Schwann cells-implications for axon ensheathment at myelination.* Glia, 2007. **55**(8): p. 844-58.

279.   Jin, F., et al., *N-WASp is required for Schwann cell cytoskeletal dynamics, normal myelin gene expression and peripheral nerve myelination.* Development, 2011. **138**(7): p. 1329-37.

280.   Novak, N., et al., *N-WASP is required for membrane wrapping and myelination by Schwann cells.* J Cell Biol, 2011. **192**(2): p. 243-50.

281.   Zuchero, J.B., et al., *CNS myelin wrapping is driven by actin disassembly.* Dev Cell, 2015. **34**(2): p. 152-67.

282.   Svitkina, T.M. and G.G. Borisy, *Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia.* J Cell Biol, 1999. **145**(5): p. 1009-26.

283.   Robinson, R.C., et al., *Crystal structure of Arp2/3 complex.* Science, 2001. **294**(5547): p. 1679-84.

284. Caloca, M.J., H. Wang, and M.G. Kazanietz, *Characterization of the Rac-GAP (Rac-GTPase-activating protein) activity of beta2-chimaerin, a 'non-protein kinase C' phorbol ester receptor.* Biochem J, 2003. **375**(Pt 2): p. 313-21.

285. Balla, T., *Putting G protein-coupled receptor-mediated activation of phospholipase C in the limelight.* J Gen Physiol, 2010. **135**(2): p. 77-80.

286. Bosco, E.E., J.C. Mulloy, and Y. Zheng, *Rac1 GTPase: a "Rac" of all trades.* Cell Mol Life Sci, 2009. **66**(3): p. 370-4.

287. Yamauchi, J., J.R. Chan, and E.M. Shooter, *Neurotrophin 3 activation of TrkC induces Schwann cell migration through the c-Jun N-terminal kinase pathway.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14421-6.

288. Nakai, Y., et al., *Temporal control of Rac in Schwann cell-axon interaction is disrupted in NF2-mutant schwannoma cells.* J Neurosci, 2006. **26**(13): p. 3390-5.

289. Benninger, Y., et al., *Essential and distinct roles for cdc42 and rac1 in the regulation of Schwann cell biology during peripheral nervous system development.* J Cell Biol, 2007. **177**(6): p. 1051-61.

290. Guo, L., et al., *Rac1 controls Schwann cell myelination through cAMP and NF2/merlin.* J Neurosci, 2012. **32**(48): p. 17251-61.

291. Nodari, A., et al., *Beta1 integrin activates Rac1 in Schwann cells to generate radial lamellae during axonal sorting and myelination.* J Cell Biol, 2007. **177**(6): p. 1063-75.

292. Zubeldia-Brenner, L., et al., *β3-chimaerin, a novel member of the chimaerin Rac-GAP family.* Mol Biol Rep, 2014. **41**(4): p. 2067-76.

293. Canagarajah, B., et al., *Structural mechanism for lipid activation of the Rac-specific GAP, beta2-chimaerin.* Cell, 2004. **119**(3): p. 407-18.

294. Pankov, R., et al., *A Rac switch regulates random versus directionally persistent cell migration.* J Cell Biol, 2005. **170**(5): p. 793-802.

295. Vogel, W., et al., *The discoidin domain receptor tyrosine kinases are activated by collagen.* Mol Cell, 1997. **1**(1): p. 13-23.

296. Shrivastava, A., et al., *An orphan receptor tyrosine kinase family whose members serve as nonintegrin collagen receptors.* Mol Cell, 1997. **1**(1): p. 25-34.

297. Xu, H., et al., *Collagen binding specificity of the discoidin domain receptors: binding sites on collagens II and III and molecular determinants for collagen IV recognition by DDR1.* Matrix Biol, 2011. **30**(1): p. 16-26.

298. Fu, H.L., et al., *Shedding of discoidin domain receptor 1 by membrane-type matrix metalloproteinases.* J Biol Chem, 2013. **288**(17): p. 12114-29.

299. Ferri, N., N.O. Carragher, and E.W. Raines, *Role of discoidin domain receptors 1 and 2 in human smooth muscle cell-mediated collagen remodeling: potential implications in atherosclerosis and lymphangioleiomyomatosis.* Am J Pathol, 2004. **164**(5): p. 1575-85.

300. Vogel, W.F., *Ligand-induced shedding of discoidin domain receptor 1.* FEBS Lett, 2002. **514**(2-3): p. 175-80.

301. Shitomi, Y., et al., *ADAM10 controls collagen signaling and cell migration on collagen by shedding the ectodomain of discoidin domain receptor 1 (DDR1).* Mol Biol Cell, 2015. **26**(4): p. 659-73.

302. Flynn, L.A., et al., *Inhibition of collagen fibrillogenesis by cells expressing soluble extracellular domains of DDR1 and DDR2.* J Mol Biol, 2010. **395**(3): p. 533-43.

303. Franco, C., et al., *Increased cell and matrix accumulation during atherogenesis in mice with vessel wall-specific deletion of discoidin domain receptor 1.* Circ Res, 2010. **106**(11): p. 1775-83.

304. Chen, P., M. Cescon, and P. Bonaldo, *The Role of Collagens in Peripheral Nerve Myelination and Function.* Mol Neurobiol, 2015. **52**(1): p. 216-25.

305. Bunge, R.P., M.B. Bunge, and C.F. Eldridge, *Linkage between axonal ensheathment and basal lamina production by Schwann cells.* Annu Rev Neurosci, 1986. **9**: p. 305-28.

306. Labelle-Dumais, C., et al., *COL4A1 Mutations Cause Neuromuscular Disease with Tissue-Specific Mechanistic Heterogeneity.* Am J Hum Genet, 2019. **104**(5): p. 847-860.

307. Paavola, K.J., et al., *Type IV collagen is an activating ligand for the adhesion G protein-coupled receptor GPR126.* Sci Signal, 2014. **7**(338): p. ra76.

308. Alves, F., et al., *Identification of two novel, kinase-deficient variants of discoidin domain receptor 1: differential expression in human colon cancer cell lines.* FASEB J, 2001. **15**(7): p. 1321-3.

309. Leitinger, B., *Discoidin domain receptor functions in physiological and pathological conditions.* Int Rev Cell Mol Biol, 2014. **310**: p. 39-87.

310. Bonifacino, J.S. and L.M. Traub, *Signals for sorting of transmembrane proteins to endosomes and lysosomes.* Annu Rev Biochem, 2003. **72**: p. 395-447.

311. Franco-Pons, N., et al., *Expression of discoidin domain receptor 1 during mouse brain development follows the progress of myelination.* Neuroscience, 2006. **140**(2): p. 463-75.

312. Fukunaga-Kalabis, M., et al., *CCN3 controls 3D spatial localization of melanocytes in the human skin through DDR1.* J Cell Biol, 2006. **175**(4): p. 563-9.

313. Meyer zum Gottesberge, A.M., et al., *Inner ear defects and hearing loss in mice lacking the collagen receptor DDR1.* Lab Invest, 2008. **88**(1): p. 27-37.

314. Hughes, T.A., *Regulation of gene expression by alternative untranslated regions.* Trends Genet, 2006. **22**(3): p. 119-22.

315. Xu, H., et al., *Discoidin domain receptors promote α1β1- and α2β1-integrin mediated cell adhesion to collagen by enhancing integrin activation.* PLoS One, 2012. **7**(12): p. e52209.

316. Staudinger, L.A., et al., *Interactions between the discoidin domain receptor 1 and β1 integrin regulate attachment to collagen.* Biol Open, 2013. **2**(11): p. 1148-59.

317. Feltri, M.L., et al., *Conditional disruption of beta 1 integrin in Schwann cells impedes interactions with axons.* J Cell Biol, 2002. **156**(1): p. 199-209.

318. Berti, C., et al., *Role of integrins in peripheral nerves and hereditary neuropathies.* Neuromolecular Med, 2006. **8**(1-2): p. 191-204.

319. She, B.R., G.G. Liou, and S. Lin-Chao, *Association of the growth-arrest-specific protein Gas7 with F-actin induces reorganization of microfilaments and promotes membrane outgrowth.* Exp Cell Res, 2002. **273**(1): p. 34-44.

320. Gotoh, A., et al., *Gas7b (growth arrest specific protein 7b) regulates neuronal cell morphology by enhancing microtubule and actin filament assembly.* J Biol Chem, 2013. **288**(48): p. 34699-706.

321. Chang, Y., et al., *Involvement of Gas7 along the ERK1/2 MAP kinase and SOX9 pathway in chondrogenesis of human marrow-derived mesenchymal stem cells.* Osteoarthritis Cartilage, 2008. **16**(11): p. 1403-12.

322. Chao, C.C., F.C. Hung, and J.J. Chao, *Gas7 is required for mesenchymal stem cell-derived bone development.* Stem Cells Int, 2013. **2013**: p. 137010.

323. Hung, F.C., et al., *Gas7 mediates the differentiation of human bone marrow-derived mesenchymal stem cells into functional osteoblasts by enhancing Runx2-dependent gene expression.* J Orthop Res, 2011. **29**(10): p. 1528-35.

324. Hung, F.C., et al., *Growth-Arrest-Specific 7 Gene Regulates Neural Crest Formation and Craniofacial Development in Zebrafish.* Stem Cells Dev, 2015. **24**(24): p. 2943-51.

325. Ju, Y.T., et al., *gas7: A gene expressed preferentially in growth-arrested fibroblasts and terminally differentiated Purkinje neurons affects neurite formation.* Proc Natl Acad Sci U S A, 1998. **95**(19): p. 11423-8.

326. Chao, C.C., et al., *Involvement of Gas7 in nerve growth factor-independent and dependent cell processes in PC12 cells.* J Neurosci Res, 2003. **74**(2): p. 248-54.

327. Chao, C.C., P.Y. Chang, and H.H. Lu, *Human Gas7 isoforms homologous to mouse transcripts differentially induce neurite outgrowth.* J Neurosci Res, 2005. **81**(2): p. 153-62.

328. Zhang, Z., et al., *Growth arrest specific gene 7 is associated with schizophrenia and regulates neuronal migration and morphogenesis.* Mol Brain, 2016. **9**(1): p. 54.

329. Huang, B.T., et al., *Gas7-deficient mouse reveals roles in motor function and muscle fiber composition during aging.* PLoS One, 2012. **7**(5): p. e37702.

330. Feltri, M.L., U. Suter, and J.B. Relvas, *The function of RhoGTPases in axon ensheathment and myelination.* Glia, 2008. **56**(14): p. 1508-17.

331. Chang, P.Y., et al., *Identification of rat Gas7 isoforms differentially expressed in brain and regulated following kainate-induced neuronal injury.* J Neurosci Res, 2005. **79**(6): p. 788-97.

332. Bi, W., et al., *Sox9 is required for cartilage formation.* Nat Genet, 1999. **22**(1): p. 85-9.

333. Betancur, P., T. Sauka-Spengler, and M. Bronner, *A Sox10 enhancer element common to the otic placode and neural crest is activated by tissue-specific paralogs.* Development, 2011. **138**(17): p. 3689-98.

334. Ohba, S., et al., *Distinct Transcriptional Programs Underlie Sox9 Regulation of the Mammalian Chondrocyte.* Cell Rep, 2015. **12**(2): p. 229-43.

335. Svaren, J. and D. Meijer, *The molecular machinery of myelin gene transcription in Schwann cells.* Glia, 2008. **56**(14): p. 1541-51.

336. Ma, K.H. and J. Svaren, *Epigenomic reprogramming in peripheral nerve injury.* Neural Regen Res, 2016. **11**(12): p. 1930-1931.

337. O'Geen, H., L. Echipare, and P.J. Farnham, *Using ChIP-seq technology to generate high-resolution profiles of histone modifications.* Methods Mol Biol, 2011. **791**: p. 265-86.

338. Bryant, G.O., *Measuring nucleosome occupancy in vivo by micrococcal nuclease.* Methods Mol Biol, 2012. **833**: p. 47-61.

339. Shao, W. and J. Zeitlinger, *Paused RNA polymerase II inhibits new transcriptional initiation.* Nat Genet, 2017. **49**(7): p. 1045-1051.

340. Fufa, T.D., et al., *Genomic analysis reveals distinct mechanisms and functional classes of SOX10-regulated genes in melanocytes.* Hum Mol Genet, 2015. **24**(19): p. 5433-50.

341. Zhang, Y., et al., *An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex.* J Neurosci, 2014. **34**(36): p. 11929-47.

342. Haltaufderhyde, K.D. and E. Oancea, *Genome-wide transcriptome analysis of human epidermal melanocytes.* Genomics, 2014. **104**(6 Pt B): p. 482-9.

343. You, J.J. and S. Lin-Chao, *Gas7 functions with N-WASP to regulate the neurite outgrowth of hippocampal neurons.* J Biol Chem, 2010. **285**(15): p. 11652-66.

344. Bolino, A., et al., *Disruption of Mtmr2 produces CMT4B1-like neuropathy with myelin outfolding and impaired spermatogenesis.* J Cell Biol, 2004. **167**(4): p. 711-21.

Appendix

Primer or Oligo Name	Sequence
MTMR2 MCS1 GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTAGGTCAGACCAAAGCAAAGC
MTMR2 MCS1 GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTAGGTACCAAAGCTCAACAGTTCAT
MTMR2 MCS2 GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGTCACATGAAGTGCTCGTG
MTMR2 MCS2 GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTCAACATGTCAGTTTTCAAAGCAC
MTMR2 MCS3 GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGCCAATTCAATTTTTCTTC
MTMR2 MCS3 GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTGGGAAATGTAATTAATAGACC
MTMR2 Prom GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGGTCTTGGGGAAAGGG
MTMR2 Prom GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTAAGTGCTACGAATGAACTGG
MTMR2 ORF1 Ntag GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGAGAAGAGCTCGAGCTGC
MTMR2 ORF2 Ntag GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGAGGAACCAGCCCTGCTTCC
MTMR2 ORF Ntag GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTTCATACAACAGTTTGGACAGGAGT
MTMR2 ORF Seq F 1	AAATCGGTGGTGCATCTAGC
MTMR2 ORF Seq F 2	GGATCCCAGTTTTATCATGG
MTMR2 ORF Seq F 3	TCTTCGGATTGCTGACAAGG
MTMR2 ORF Seq F 4	AGAAGACAGTGTCACTGTGG
MTMR2 ORF Seq R 1	TCGTACAGCACCAGTGAACG
MTMR2 P2 mut delSOX10 F	CAGTGTGTATAAAACATTTGAGACCAGAATATTTGCAGGATCCTCAAA
MTMR2 P2 mut delSOX10 R	TTTGAGGATCCTGCAAATATTCTGGTCTCAAATGTTTTATACACACTG
MTMR2 P2 rs565771 mut F	TTTATAATTTTTAATTATGTTTAAAAATGCT
MTMR2 P2 rs565771 mut R	AGCATTTTTAAACATAATTAAAAATTATAAA
MTMR2 P2 rs189276120 mut F	AGGAACGCAAATTTTCCTCTCAATACTTAAC
MTMR2 P2 rs189276120 mut R	GTTAAGTATTGAGAGGAAAATTTGCGTTCCT
MTMR2 RACE 1R	ACAGGCCATAAGAATTTTCACC
MTMR2 RACE 2R	ACGGACATATGTAAGTCACATCTTT
MTMR2 RACE 3R	TATGAACCGAGTTCTCTGAATGAG
rnMTMR2 RT Ex1A F 2	GCTCCTCCCTCTGCATCTCG
rnMTMR2 RT Ex1B F	GAAATGGATTTGACAGTGAGGA
rnMTMR2 RT Ex4 R	CACAGCACCAGTGAATGGAC
hsMTMR2 Exon1A F	ACTCCGCCCTCCGCTTCTCC
hsMTMR2 Exon1B F	AAAATTGAATTGGCAGCAAGAGTGA
hsMTMR2 Exon4 R	TACAGCGCCAGTGAATGGAC
rnMTMR2 3UTR R	AGCGGTTGGAAATTTCTCTCT
MTMR2 mRNA Seq F 1	GCAAGGATATCAGGAACTTGC
MTMR2 mRNA Seq F 2	TGTTATGAGAGAATCATTACG
mm rn bActin F	CGCGGGCGACGATGCTCC
mm rn bActin R	GTAGCCACGCTCGGTCAGG
qPCR 18S rRNA F	CGCCGCTAGAGGTGAAATTCT
qPCR 18S rRNA R	CCAACCTCCGACTTTCGTTCT
qPCR Mtmr2 Exon1A F	GAGGACTCACTGTCCAGCCAAA
qPCR Mtmr2 Exon1A R	TTTGTGACAGTCAGCGTCCCT
qPCR Mtmr2 Exon1B F	GAAATTGAATTGGCAGTGAAAA
qPCR Mtmr2 Exon1B R	TACAGCACCAGTGAACGGAC

Table A.1 Primers used for analyses at *MTMR2*. Primer sequences are provided 5' to 3'.

Primer or Oligo Name	Sequence
ΔSOX10 gRNA 1	CAGCGGCTACGACTGGACGC
ΔSOX10 gRNA 2	CCGTTGACCCGCACGGGCAT
ΔSOX10 gRNA 1 PX459 oligo F	CACCGCAGCGGCTACGACTGGACGC
ΔSOX10 gRNA 1 PX459 oligo R	AAACGCGTCCAGTCGTAGCCGCTGC
ΔSOX10 gRNA 2 PX459 oligo F	CACCGCCGTTGACCCGCACGGGCAT
ΔSOX10 gRNA 2 PX459 oligo R	AAACATGCCCGTGCGGGTCAACGGC
ΔSOX10 Screening Exon1 F	CCTCCTTCCCGTCTAGGTG
ΔSOX10 Screening Exon1 R	CTGGTCGGCTAACTTTCTGC
SOX10 RTPCR F1	CCACCGGCACCCAGAAGAAGG
SOX10 RTPCR R1	GACTGCAGCTCTGTCTTTGG
SOX10 RTPCR F2	CACAATGCTGAGCTCAGC
SOX10 RTPCR R2	TTTGTGCTGCATCCGGAGC
mm rn bActin F	CGCGGGCGACGATGCTCC
mm rn bActin R	GTAGCCACGCTCGGTCAGG
Oligo-dT-smartseq2	/5Me-isodC/AAGCAGTGGTATCAACGCAGAGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTVN
Nextera A TSO1	TCGTCGGCAGCGTCAGATGTGTATAAGAGArCAG rUGA ArU rUC TGGTrGrGrG
ISPCR	AAGCAGTGGTATCAACGCAGAGT
Nextera A Index N501	AATGATACGGCGACCACCGAGATCTACAC TAGATCGC TCGTCGGCAGCGTCAGATG
Nextera A Index N502	AATGATACGGCGACCACCGAGATCTACAC CTCTCTAT TCGTCGGCAGCGTCAGATG
Nextera A Index N503	AATGATACGGCGACCACCGAGATCTACAC TATCCTCT TCGTCGGCAGCGTCAGATG
Nextera A Index N504	AATGATACGGCGACCACCGAGATCTACAC AGAGTAGA TCGTCGGCAGCGTCAGATG
Nextera A Index N505	AATGATACGGCGACCACCGAGATCTACAC GTAAGGAG TCGTCGGCAGCGTCAGATG
Nextera A Index N506	AATGATACGGCGACCACCGAGATCTACAC ACTGCATA TCGTCGGCAGCGTCAGATG
Nextera A Index N507	AATGATACGGCGACCACCGAGATCTACAC AAGGAGTA TCGTCGGCAGCGTCAGATG
Nextera A Index N508	AATGATACGGCGACCACCGAGATCTACAC CTAAGCCT TCGTCGGCAGCGTCAGATG
Tn5ME-R	[phos]CTGTCTCTTATACACATCT
Tn5ME-B	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG
Nextera B Index N701	CAAGCAGAAGACGGCATACGAGAT TCGCCTTA GTCTCGTGGGCTCGGAGATGTGTAT
Nextera A Universal	AATGATACGGCGACCACCGAGATCTACAC

Table A.2 Primers used for genome-wide analyses. Primer sequences are provided 5' to 3'.

Primer or Oligo Name	Sequence
Human ARPC1A Prom 2 GWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTAGCTCCTAACACAGTGC
Human ARPC1A Prom 2 GWR	GGGGACCACTTTGTACAAGAAAGCTGGGTGACCAACATGGAGAAACC
Human ARPC1A Prom 2 delSOX10 1 mut F	ACAGACACAGATGATCCACAGGGGCTTCTGCA
Human ARPC1A Prom 2 delSOX10 1 mut R	TGCAGAAGCCCCTGTGGATCATCTGTGTCTGT
Human ARPC1A Prom 2 delSOX10 2 mut F	GTCGAGACCCTGATATCCCAAGATGCAACAGACTATTA
Human ARPC1A Prom 2 delSOX10 2 mut R	TAATAGTCTGTTGCATCTTGGGATATCAGGGTCTCGAC
ARPC1A Ex1B F	TAGTGCTGACCGATCTCC
ARPC1A 3UTR R	CAGTGTGACAGTGTTTCC

Table A.3 Primers used for analyses at *ARPC1A*. Primer sequences are provided 5' to 3'.

Primer or Oligo Name	Sequence
Human CHN2 Prom 4 GWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAGGAAGGACAGGAGAGACC
Human CHN2 Prom 4 GWR	GGGGACCACTTTGTACAAGAAAGCTGGGTTACGGCCAGCAGTTCTTGG
Human CHN2 Prom 4 delSOX10 mut F	ACGTGCGCCTGGAGCACAGCAACCCC
Human CHN2 Prom 4 delSOX10 mut R	GGGGTTGCTGTGCTCCAGGCGCACGT
CHN2 Ex1B F	ATGTTCTCTCAGGAATTGTGG
CHN2 3UTR R	TGCAATCCACTGTGGTCC

Table A.4 Primers used for analyses at *CHN2*. Primer sequences are provided 5' to 3'.

Primer or Oligo Name	Sequence
Human DDR1 Prom 5 GWR	GGGGACCACTTTGTACAAGAAAGCTGGGTTCCCACTTCCAGAAGGTGG
Human DDR1 Prom 5 GWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCAAGGTCTTCTGGGCTTGTCC
DDR1 Prom 5 delSOX10 mut R	CTCTCTCGGGATCCCCTAAGTAACCTTCATGC
DDR1 Prom 5 delSOX10 mut F	GCATGAAGGTTACTTAGGGGATCCCGAGAGAG
rn DDR1 Ex1D F	CAACTCAGTCCTCTCAGC
rn DDR1 3'UTR R	GCTTCTCTGGTTTTAGTGTCC
rn DDR1 Ex10 F	AACTGCCATCCTCATTGGCTGC
rn DDR1 Ex15 R	AATCATGCAGAGAGGGTCATCC

Table A.5 Primers used for analyses at *DDR1*. Primer sequences are provided 5' to 3'.

Primer or Oligo Name	Sequence
Human GAS7 Prom 2 GWF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGACTTTGTGATCCTCAGG
Human GAS7 Prom 2 GWR	GGGGACCACTTTGTACAAGAAAGCTGGGTACTGGAGCATACATCTGC
Human GAS7 Prom 2 delSOX10 1 mut F	CCTAAAACACAGTGGCCACAGGCCTGGGGG
Human GAS7 Prom 2 delSOX10 1 mut R	CCCCCAGGCCTGTGGCCACTGTGTTTTAGG
Human GAS7 Prom 2 delSOX10 2 mut F	AGGCATCTGCTCAGTGTTGTTGTGACCTCAGC
Human GAS7 Prom 2 delSOX10 2 mut R	GCTGAGGTCACAACAACACTGAGCAGATGCCT
Rat GAS7 Ex1B F	TCTTCCATAGCAACTCTCC
Rat GAS7 3UTR R	TGTGGGTCCATCATTATGG

Table A.6 Primers used for analyses at *GAS7*. Primer sequences are provided 5' to 3'.