2022

The Musashi RNA Binding Proteins Are Regulators of Alternative Splicing and Protein Expression in Photoreceptor Cells

Fatimah Kh. Matalkah
West Virginia University, fmatalkah@mix.wvu.edu

The Musashi RNA Binding Proteins Are Regulators of Alternative Splicing and Protein Expression in Photoreceptor Cells.

Fatimah KH. Matalkah

Dissertation submitted to the
School of Medicine
at West Virginia University
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Biochemistry and Molecular Biology

Committee Members
Peter Stoilov, Ph.D., Chair
Visvanathan Ramamurthy, Ph.D.
Maxim Sokolov, Ph.D.
Roberta Leonardi, Ph.D.
Ariel Agmon, Ph.D.

Graduate Program in Biochemistry
West Virginia University School of Medicine
Morgantown, West Virginia
2022

Keywords: RNA splicing, Retina, Photoreceptor, Musashi, PROM1

1. Abstract

The Musashi RNA Binding Proteins Are Regulators of Alternative Splicing and Protein Expression in Photoreceptor Cells.

Fatimah Kh Matalkah

The Musashi (Msi) family of RNA binding proteins consists of two paralogs, Msi1 and Msi2, that are highly conserved across species. The two paralogs have emerged as factors that promote stem cell proliferation by post-transcriptionally regulating translation. In addition to their expression in stem cells, the Musashi proteins are also expressed in postmitotic neurons, including the photoreceptor cells. The Musashi proteins have been observed to maintain high expression levels in the postmitotic photoreceptors within the eye of both invertebrates and vertebrates. These observations suggest an additional role in the maintenance of terminally differentiated neurons.

Building upon these observations, we investigated the role of Musashi individually and in combination in mature photoreceptors. Using a tamoxifen-inducible mouse model, I generated single and combined deletion of *Msi1* and *Msi2* in mature photoreceptor cells. Our results show that the Musashi proteins are required for the function and viability of mature photoreceptors. Global analysis of the Msi1 targets in the retina showed binding to UAG motifs predominantly located in introns and 3'-UTRs. Using RNA-sequencing and proteomics analysis, with the incorporation of the publicly available single-cell RNA seq, we found that in mature photoreceptors, the Musashi enhance the expression of proteins in high demand. Among these targets are proteins needed for the daily regeneration of the light sensory organelle of the photoreceptors. Collectively, the data provide new insights on the targets, possible molecular mechanisms, and function of the Musashi in mature photoreceptors. The results support a model of the Musashi proteins acting as a posttranscriptional activator for protein expression in mature photoreceptors.

In the course of our work, an unusual behavior of the 13A4 antibody to prominin-1 (Prom1) prompted us to analyze its epitope. Prom1 is a transmembrane protein with a role in the morphogenesis of photoreceptor outer segment disk membranes. Mutations in the *Prom1* gene have resulted in various forms of retinal degeneration affecting rods and cones. Scanning deletion mutagenesis and structural modeling demonstrated that mAB 13A4 recognizes a structural epitope that is affected by the inclusion of the alternative exon 19 during photoreceptor maturation. Consequently, the reactivity of mAB 13A4 towards the photoreceptor specific isoform of PROM1 is significantly reduced on a Western blot leading to gross underestimation of PROM1 protein levels in the retina.

2. Dedication

I dedicate this work to my parents, whose love and support gave me a chance for a better life.

Also, to Malik and Razan, I am genuinely thankful for having you in my life.

3. Acknowledgments

First of all, I would like to express my deepest appreciation to my advisor Dr. Peter Stoilov for his assistance at every stage of this project and for the financial support he provided for this work. His expertise was invaluable during my Ph.D. study. I am also extremely grateful to my committee members Drs. Visvanathan Ramamurthy, Maxim Sokolov, Roberta Leonardi, and Ariel Agmon, for their guidance and continued support. Their immense knowledge and expertise have helped me sail smoothly throughout this project. With special thanks to Dr. Ramamurthy's and Dr. Sokolov's lab members (Urikhan, Rawaa, Hunter, Daniella, Alexis, Faezah, Scott, Thamaraiselvi, Celine, and Dr. Douglas Kolson) for all the help and technical assistance on my study. I am also very grateful for my lab pal Bohye Joeng for all the work and help she provided during this study. I must also acknowledge the many fellow students, staff, and faculty in the department of biochemistry for their kind help and support.

Lastly, my family deserves immense gratitude. To my brother (Khaldoon), a lifelong friend and supporter, thank you for standing by my side as an iron shield supporting me all the way. Without your tremendous encouragement and support, I would not have made it to where I am today. To my mother and all my family members, thank you for the endless support and encouragement through my studies. The final thank you goes to the sweetest gift in my life, Malik and Razan; your presence makes me stronger and more fulfilled.

4. Table of Contents

5. List of Figures

1. Chapter 1

1.1 Diverse functions of RNA-binding proteins (RBPs)

RNA-binding proteins (RBPs) function in every step of the mRNA lifecycle. They associate with mRNAs by binding to specific sequences or secondary structures to regulate processes that include pre-mRNA splicing, polyadenylation, transport, translation, and stability (Figure 1) [1,2]. A single RBP can regulate multiple mRNA targets, while one mRNA can be controlled by various RBPs. Proper regulation of these processes is crucial for gene expression, and their perturbation often leads to disease [3].

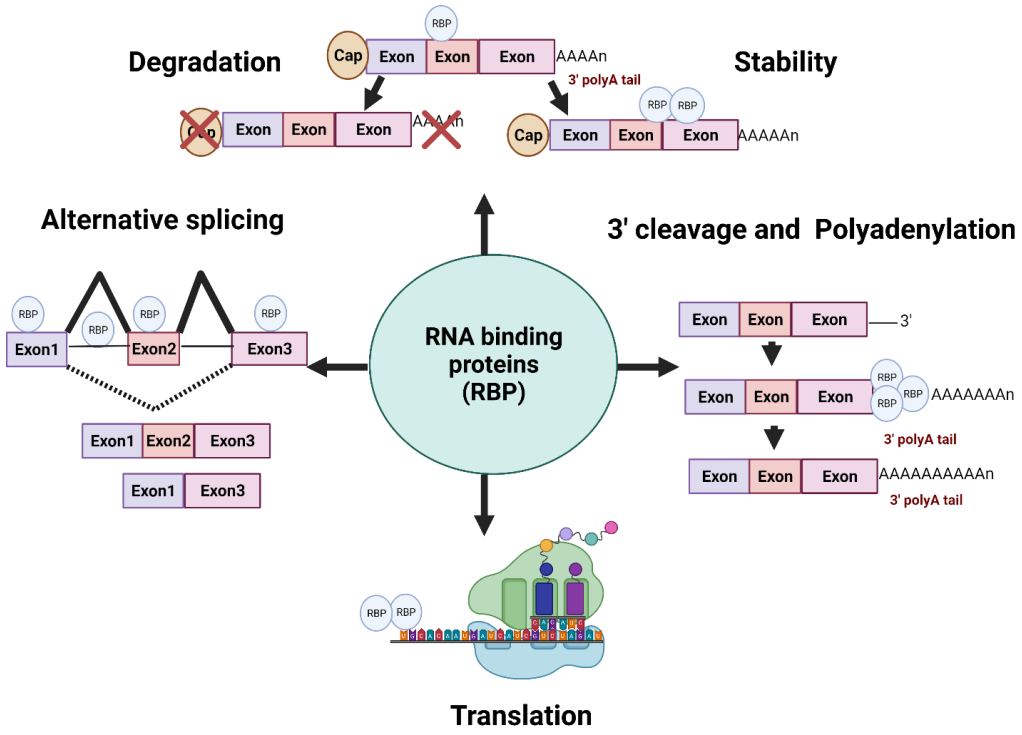


Figure 1-1 Posttranscriptional regulations mediated by the RNA binding proteins

Figure1: Posttranscriptional regulations mediated by the RNA binding proteins: Schematic summarizing various functions for the RNA binding proteins in the mRNA lifecycle. Figure created using Biorender.com.

1.1.1 Pre-mRNA splicing

Pre-mRNA splicing is an essential step of the gene expression process in eukaryotes. The process includes removing the introns from the pre-mRNA and joining the exons to produce mature mRNAs [4,5]. Furthermore, since during splicing, each exon is recognized as an individual module, the exon recognition process can be controlled so that individual genes produce multiple transcript isoforms differing in their exon composition [6,7]. This process is called alternative splicing and provides a major mechanism to expand the protein repertoire [7].

Splicing is carried out by a large RNA-protein complex known as the spliceosome [4]. Two parallel spliceosomes exist in eukaryotes, a major spliceosome that is responsible for the excision of most introns and a minor spliceosome that excises approximately 0.3% of introns [8]. The major core spliceosome consists of five small nuclear ribonucleoprotein particles (snRNPs), U1, U2, U4, U5, and U6 [9]. In the minor spliceosome U1 and U2 are replaced by the U11 and U12 snRNPs. In addition to the five core snRNPs, a host of proteins is associated with the spliceosome and is required to achieve efficient splicing [9]. The splicing reaction requires the spliceosome to assemble in a stepwise manner onto each intron. The first step of this process is the recognition of the sequence elements that define the intron/exon boundaries [10]. These sequence elements include the 5' splice site (5'SS), the branch point, and the 3' splice site (3'SS) [7]. The differences in the composition of the major and minor spliceosomes results in the two machineries recognizing different splice site and branch point sequences [10]. Hence, the introns spliced by the major spliceosome are called the U2-type, and the ones spliced by the minor spliceosome are called the U12-type introns [10].

The initial assembly is undertaken by the recognition of the dinucleotide GU at the 5’ SS by U1 snRNPs and the branch point (BP) by U2 snRNPs (major spliceosome), or the recognition of AU at the 5’ SS and the BP by the U11/ U12 (minor spliceosome), respectively (Figure 2) [4,9]. Furthermore, the polypyrimidine tract (PPT) found in the U2-type introns is recognized by the

U2AF1/U2AF2 heterodimer (major spliceosome) (Figure 2) [11]. However, the AC dinucleotide found at the 3'SS of U-12 type introns is recognized by a protein called ZRSR2 (minor spliceosome) (Figure 2). The subsequent interaction between the U1/U2 snRNPs (major spliceosome) or the U11/U12 (minor spliceosome) across the intron forms complex A, which brings the 5' and 3' SS together by looping out the intron (Figure 2). The assembly of complex A across the intron is known as the intron definition [10,12]. Once the A complex is formed the two major and minor spliceosomes follow a similar assembly path. The next major assembly step is binding the preassembled tri snRNPs U4/U6-U5 to complex A to form the precatalytic B complex [8]. This is followed by the rearrangement of complex B resulting in the removal of U1/U11 and U4 snRNPs and the formation of the active spliceosome (B*) [8]. Next, the active spliceosome carries the splicing reaction that includes two transesterification reactions, the results of which include the removal of the intron and the joining of the two adjacent exons [4,8].

The information contained in the consensus motifs of the 5' and 3' SS is typically insufficient to identify the correct boundaries of most introns. This problem is further exacerbated by the large size of metazoan introns that contain multiple sequences matching the splice site consensus. Hence, additional sequence elements within the exons and their nearby intronic regions are needed to correctly define the exon/intron boundaries. These additional cis-acting elements, collectively known as the splicing regulatory elements (SREs), are recognized by various RBPs [5]. The binding of the RBPs to SREs can either promote or block the assembly of the spliceosome depending on the nature of the RBP and the position of its binding relative to the adjacent splice site [13,14]. The SR family of RNA binding proteins are prototypical splicing activators that facilitate exon recognition and recruit the spliceosome components to the correct splice sites [15]. These proteins consist of one or two RRM-type RNA binding domains and extensive Serine/Arginine repeats from which the "SR"' name derives [16]. Another group of proteins, named heterogeneous nuclear ribonucleoproteins (hnRNP) due to their association with various RNAs,

was initially perceived as splicing inhibitors that block the use of cryptic splice sites [16]. It is important to understand that RNA binding proteins can act as activators and repressors. Recognizing the correct splice sites is a combinatorial process where RNA binding proteins cooperate and compete to recruit the spliceosome to the correct splice sites and block the recognition of cryptic splice sites. The balance of positive and negative regulators binding in the vicinity of the splice sites determines the outcome of the exon recognition process and can be controlled to cause the inclusion or skipping of exons depending on cell type or environmental cues.

Interestingly, the binding positions of the RBPs relative to splice sites can either activate or repress the splice site usage [17]. For example, the binding of Nova and Rbfox to the intronic region downstream of the exon results in its inclusion. However, the binding to the upstream intron results in exon skipping. The position-specific splicing regulation imposed by the RBPs on pre-mRNA can be elegantly visualized by the integration of the *in vivo* binding sites identified by the cross linking and immunoprecipitation (CLIP) and the events of alternative splicing identified by the RNA-seq [17]. Integration of both tools coupled with bioinformatics analysis has enabled the generation of "RNA splicing maps," a visualization tool that informs how the binding position relative to the splice sites affects the direction in which each protein affects splicing [17].

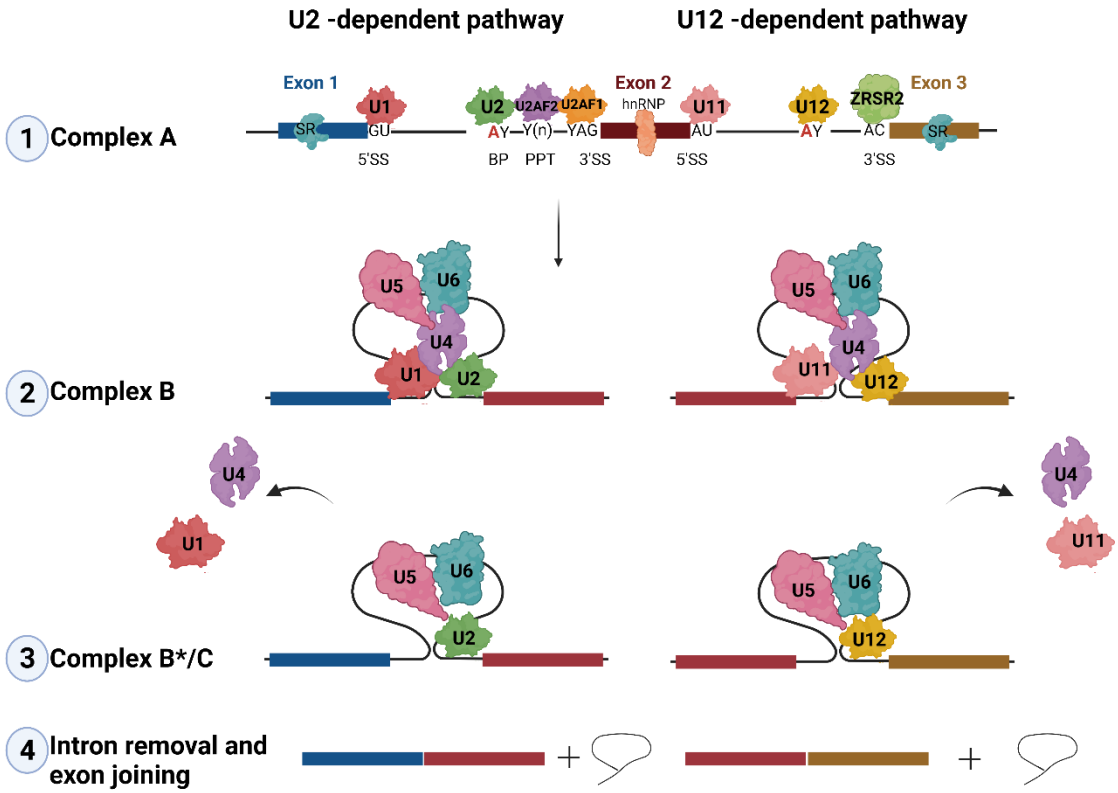


Figure 1-2 Pre-mRNA splicing by the major and minor spliceosomes

Figure 2: Pre-mRNA splicing by the major and minor spliceosomes: 1) Splice site sequences of major introns (U2-type) at the left side of the schematic and the minor introns (U12-type) at the right side and the assembly of complex A across the introns. **2)** The formation of complex B. **3)** The formation of complex B*/C. **4)** The removal of intron and exon joining. Figure created using Biorender.com.

1.1.2 Pre-mRNA cleavage and polyadenylation

Polyadenylation of nascent transcript is another key step in the gene expression process that, similar to splicing, is guided and regulated by RNA binding proteins. All mRNAs, except for histone transcripts, are modified with a poly (A) tail at their 3' end in a process called polyadenylation [18]. Adding the poly(A) tail facilitates the transport of the mRNA from the nucleus to the cytoplasm and protects it from degradation. The length of the poly(A) tail affects the mRNA's stability and rate of translation. While pre-mRNA polyadenylation is initiated in the nucleus, mRNA polyadenylation can also occur in the cytoplasm. Nuclear polyadenylation is a two-step process in which the nascent transcript is first cleaved, and a poly(A) tail is added by a specialized enzyme [6]. The cleavage and polyadenylation reactions are carried out by a multiprotein complex that binds to regulatory cis-elements found on the 3' UTR of the pre-mRNA [18]. The first regulatory cis-element is the poly(A) signal (PAS) (AAUAAA) located 20-30 nucleotides from the cleavage site [18]. The second regulatory sequence is a guanine and uracil G/U rich element located 30 nucleotides downstream of the cleavage site, thereby known as the downstream sequence element (DSE) [18]. The third cis element is the dinucleotide of CA or UA at the cleavage site located 20 to 30 nucleotides downstream of the PAS [18]. The first step of the polyadenylation process is the cleavage of the primary transcript that is carried by the multiprotein complex composed of more than 80 RNA-binding proteins [18]. These include the cleavage and polyadenylation specificity factors (CPSFs), the cleavage stimulation factors (CstFs), and the cleavage factor I and II (CFI, and CFII) [19]. The cleavage is initiated by the binding of the CPSFs to the PAS and the CstFs to the DSE [18]. The interaction between the CPSF and CstF leads to the recruitment of additional factors, including the CFI, CFII, and the Poly (A) polymerase, which join around the cleavage site [18]. As the polyadenylation complex assembles, the pre-mRNA is cleaved between the PAS and DSE elements. Then, the poly (A) polymerase is tethered to the cleaved site of the pre-mRNA and starts adding the poly(A) tail to the 3' end [20]. Once in the

cytoplasm, the poly (A) tail of the transcript is mostly coated with the cytoplasmic poly (A) binding protein (PABP) [21]. The binding of the PABP either enhances or represses the transcript translation by recruiting polyadenylating or deadenylating proteins, respectively [19].

Cytoplasmic polyadenylation is a mechanism the cell employs to activate the translation of dormant mRNAs via the elongation of poly(A) tails in the cytoplasm [22]. This form of translational control is utilized during multiple biological processes. For example, it is essential during early development, oocyte maturation, and in the adult brain [23]. The common theme among these processes is that transcription is either silenced or at a physiological distance from the translation machinery [22]. Taking oocyte maturation as an example, generally, in the immature oocyte, the dormant mRNAs would have a relatively short poly(A) tail consisting of about 20 nucleotides or less [23]. However, during oocyte maturation, several mRNAs are recruited to the polysome in a sequence-specific manner, and upon which the poly(A) tail on these mRNAs grows to about 80 to 150 nucleotides. Increasing the length of the poly(A) tail promotes the physical interaction between the 5'-end and 3'-end of the mRNA leading to its cyclization, encouraging its association with multiple ribosomes, and enhancing its translation [24]. Cytoplasmic polyadenylation is mediated by the Germ-Line-Development-Defective-2 poly(A) polymerase (GLD-2) and its orthologs. However, GLD-2 is a noncanonical cytoplasmic poly(A) polymerase that lacks the RNA binding domain and is dependent on its protein partners to be recruited to target mRNAs in a sequence-specific manner [25].

Not all mRNA are clients for cytoplasmic polyadenylation. As such, only RNA that contains the cytoplasmic polyadenylation elements (CPEs) at their 3' end are targets for this process. The CPE is a uracil-rich element with variable sequences; however, the most general consensus sequence is the (UUUUUAU) [26]. In the cytoplasm, mRNA harboring the CPEs is associated with several factors that control the poly (A) tail length. These include the CPE binding protein (CPEB), the cytoplasmic form of the cleavage and polyadenylation specificity factor

(CPSF), the scaffolding protein Symplekin, and the cytoplasmic poly(A) polymerase Gld-2 [24]. In addition to CPEB, several other proteins, including GLD-3, RNP-8, Musashi, Pumilio, and Nanos, have been shown to recruit GLD-2 and control cytoplasmic polyadenylation [27].

1.1.3 RNA turnover

The mRNA turnover is another process that the RBPs mediate. It is a highly regulated process in the cytoplasm and plays a crucial role in regulating the transcript level [28]. The RBPs regulate mRNA turnover by binding to cis-elements found in the 3' UTR of the transcript; the best-studied one is the adenylate and uridylate-rich element (AU-rich elements (AREs)). Interestingly, the binding of different RBPs to the same cis element has been shown to either stimulate or prevent mRNA decay. For example, Human antigen R (HuR) and tristetraprolin (TTP) are among the proteins that play a role in regulating the mRNA turnover [29]. These two proteins compete for the same AREs on the 3'UTR of a transcript. For example, it has been shown that the binding of HuR, a member of the ELAV family of RBPs, to the ARE inhibits the rapid mRNA turnover. The stabilizing effects of HuR are due to the competition with destabilizing proteins for the binding ARE on the mRNA. However, the binding of the TTP, a zinc finger protein, to the ARE promotes mRNA decay by stimulating the deadenylation of the transcript [29].

In eukaryotes, there are two pathways of mRNA decay, both of which are initiated by the removal of the poly (A) tail (deadenylation) [28]. In the first pathway, removing the poly (A) tail stimulates the hydrolysis of the mRNA cap structure and the subsequent degradation of the mRNA by the 5' exonucleases [28,30]. In the second pathway, the mRNA is degraded by a complex of 3′ exonucleases with no need for the hydrolyzation of the cap structure [28,30]. Degradation of mRNA occurs at specific loci within the cytoplasm known as the processing bodies (P-bodies). The P-body is where the mRNA decay factors are localized [28]. These include the deadenylase, decapping enzymes, and exonucleases that drive the mRNA degradation.

Similarly, multiple microRNAs (miRNA) can also regulate mRNA turnover. These small RNAs bind to regulatory elements in the 3' UTR of their target RNA with an imperfect match [31]. More than 60% of the mammalian genes are estimated to have at least one miRNA-binding site [32]. The binding of the miRNA to the target mRNA recruits Argonaute (Ago) proteins and GW182/TNRC6 to form the RNA-induced silencing complex (RISC). This is followed by the recruitment of several factors and enzymes that stimulate mRNA cleavage and degradation, including major deadenylase, decapping enzymes, 3′ and 5′ exonucleases, and endonucleases [33]. The consequence of this is either an increased rate of mRNA decay or translational repression.

1.1.4 Regulation of translation

The translation process is divided into three phases: initiation, elongation, and termination. The initiation phase consists of multiple steps that include the association of the eukaryotic initiation factors (eIFs) with the 5'cap and the subsequent recruitment of the 40S ribosomal subunit to form the scanning complex (Figure 3) [34]. The first step of initiation includes binding the eIF-1, eIF-1A, and eIF-3 to the 40s ribosomal subunit (Figure 3) [35]. Then the initiator methionyl tRNA is brought to the ribosome by eIF-2 (in complex with GTP) [35]. The recruitment of the mRNA to the 40s ribosome then follows, mediated via the cap binding complex eIF-4E (Figure 3). Another factor, the eIF-4G, is then recruited. The eIF-4G can bind both the eIF-4E at the 5' end and the PABP located at the 3' end, a step that promotes mRNA circulation and ribosomal recycling, promoting efficient translation [35]. The complex then starts scanning the 5' end of the transcript for the first AUG codon in the correct context (Figure 3). Once the start codon is recognized, this is followed by recruitment of the 60s subunit and the subsequent assembly of the 80s ribosome [35]. Then, the elongation factors are recruited, and translation is initiated [35]. The translation proceeds until a stop codon is reached and recognized by the termination factors [34].

The RNA-binding proteins can modulate the translation level at any of the three steps [35]. One of these mechanisms includes binding the RBPs to the preinitiation complex and subsequently blocking the circularization step of the mRNA to repress translation [35]. The Musashi proteins are among the RBPs that have been reported to regulate the translation of its target transcripts by interfering with the preinitiation complex assembly. For example, Msi1 competes with eIF-4G for the PABP, blocking the circularization step of the mRNA [36]. Another mode of regulation is to prevent the binding of the 60s ribosomal subunit to the 40s ribosomal subunit [37]. The heterogenous nucleoproteins K and E (hnRNPK and hnRNPE1) are among the RNA-binding proteins that have been shown to intervene with the translation by blocking the assembly of the two subunits [37]. By binding to a CU-rich element found in the 3'UTR of their target transcript, hnRNPK and hnRNPE1 target the transcript to ribonucleoproteins (RNPs) from which the 60s is excluded [35]. RBPs can also interfere with the elongation step. In this case, the binding of the RBPs to the coding sequence will stall the 80s ribosome on the target mRNA [35]. An example of this mechanism is the fragile X mental retardation protein (FMRP), a neuronal RNA binding protein known to negatively regulate translation by binding to the coding sequence stalling the ribosomal translocation [35]. Collectively, by binding to cis elements found at the 5' and 3' UTRs or the coding sequence of the mRNA, the RBPs can control the rate of translation and subsequently direct protein expression required in response to the cellular demand.

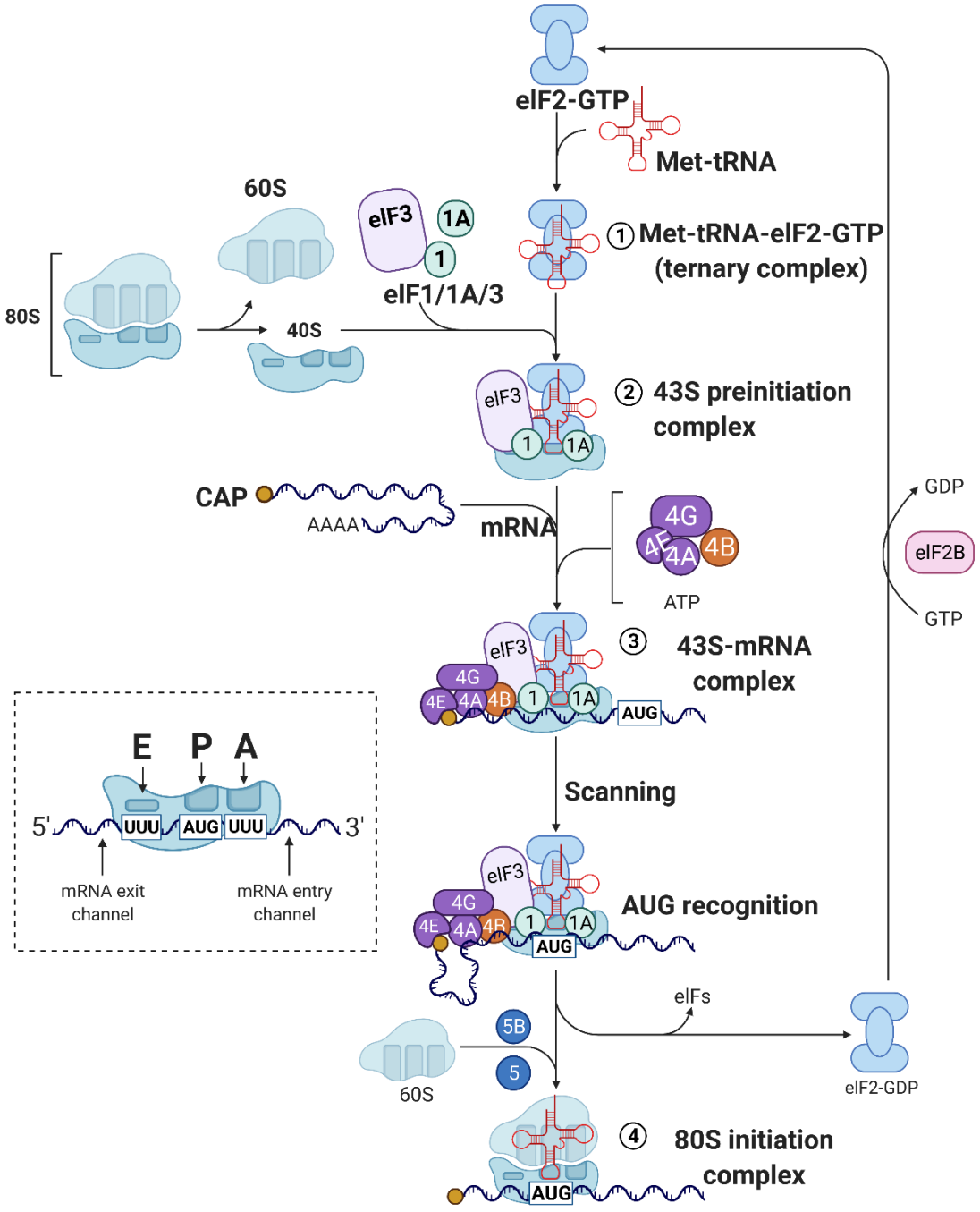


Figure 1-3 The translation initiation phase in eukaryotes

Figure 3: The translation initiation phase in eukaryotes: The schematic outlines the initiation, which starts by binding eIF-3, -1, and -1A to the 40s ribosome. The eIF-2 brings the initiator methionyl tRNA to the ribosome, and the mRNA is brought by eIF-4E (which binds to the 5´ cap). The eIF-4G then binds both the eIF-4E complex and the PABP. The ribosome then starts scanning the mRNA for the AUG initiation codon. Figure created using Biorender.com

1.2 Photoreceptor cells

1.2.1 Structure and function

Vertebrates rely on the photoreceptor cells, the light-capturing cells within the retina, for image-forming vision [38]. There are two types of photoreceptor cells: rods and cones. The rods specialize in night vision as they are extremely sensitive to light and can detect a single photon, while cones mediate daylight and color vision, respectively [38]. Cones can be further classified into different subtypes depending on the absorption maxima of the visual pigment they contain. These include the S-type, the M-type, and the L-type, which are highly sensitive to the blue, green, and red colors [38]. Both cones and rods are compartmentalized cells consisting of four structural regions: The outer segment (OS), the inner segment (IS), the cell body, and the synapse terminal. The OS function is to capture light and, in response, trigger electrochemical signals in a process known as phototransduction. Structurally, the OS is a primary cilium with an elaborate stack of membrane disks [39]. The disks are densely packed with the visual pigment, rhodopsin or cone opsin, and other proteins that are necessary for the phototransduction cascade (Figure. 4-section 1.2.2) [38]. The dense stacking of the disks increases the efficiency of photon capture [5]. The OS is constantly regenerated throughout the lifetime of photoreceptor cells to ensure maximum photosensitivity [39,40]. The process of regeneration occurs daily, with 10-15% of the OS at the distal end being removed by phagocytosis by the retinal pigmented epithelial cells (RPE), and new membrane stacks filled with proteins are continuously added to the base of the OS to compensate for the loss at the tip [39,40].

The OS is connected to the inner segment via a narrow bridge-like structure called the connecting cilia [40]. The inner segment is the place where protein and lipid synthesis take place. All the proteins that are residents of the OS are synthesized in the IS and then transported via the connecting cilia to reach their final destination in the OS [40]. In the mammalian retina, the nuclei of photoreceptor cells are stacked on top of each other, forming the outer nuclear layer (ONL)

[40]. The photoreceptors connect to the downstream bipolar and horizontal neurons via ribbon-type synapse that is densely packed with synaptic vesicles containing the neurotransmitter glutamate. The synaptic ribbon is a proteinaceous structure anchored to the plasma membrane at the active zone and extends to the cytoplasm [41]. The main function of the synaptic ribbon is to anchor a large number of the readily releasable vesicles at the presynaptic zone to increase the release rate upon the signal arrival [41]. The signal is then passed to the inner neurons within the retina for further processing and subsequently into the midbrain for visual reflexes, and to the thalamus and visual cortex within the central nervous system for additional processing and for conscious vision.

1.2.2 Phototransduction in rod photoreceptor cells

In the vertebrate rod photoreceptor neurons, image-forming vision starts with the absorption of the light photons by the retinal chromophore covalently attached to the rhodopsin molecule [38]. This results in the isomerization of the chromophore from 11-*cis* to all-*trans* isomer, causing a conformational change in the C-terminal domain of the rhodopsin molecule allowing it to interact and activate transducin, a heterotrimeric G protein composed of 3 subunits G$\boldsymbol{\alpha}$, Gβ and Gγ(Figure 4) [38,39]. Upon the binding of rhodopsin to transducin, it catalyzes the exchange of GDP to GTP on the G$\boldsymbol{\alpha}$ subunit and subsequently activates it [38,39]. Once activated, the G$\boldsymbol{\alpha}$ dissociates from its native partner Gβγ and diffuses in the disc membrane until it encounters the phosphodiesterase enzyme (PDE) (Figure 4). PDE is a tetrameric protein consisting of two catalytic subunits $\boldsymbol{\alpha}$ and β, and two identical inhibitory γ subunits. In the dark the two γ subunits bind to the catalytic domain of $\boldsymbol{\alpha}$ and β subunits inhibiting the hydrolysis of cGMP (Figure 4). In the light, once G$\boldsymbol{\alpha}$-GTP encounters PDEγ subunits it sterically displaces it from the catalytic site relieving the inhibition and permitting the hydrolysis of cGMP to GMP by the PDE$\boldsymbol{\alpha}$ and PDEβ subunits [38]. The decline in cGMP level upon hydrolysis leads to the rapid closure of the cGMP-gated channel (CNG), a nonselective cation channel found in the plasma membrane of the

photoreceptor cell [37,38]. The closure of the channel upon illumination ceases the Na^{+} and Ca^{2+} influx resulting in the hyperpolarization of the rod photoreceptor cell and subsequently inhibits the release of glutamate at the synaptic terminal [38]. The channel closure is the final step in the phototransduction cascade, after which the photoreceptor cell must reset and return to the dark state. The recovery step starts with the phosphorylation of rhodopsin at multiple serine/threonine residues found at the C-terminal by the G-protein coupled receptor kinase (GRK1) [38]. The phosphorylation of rhodopsin decreases its activity, creating binding sites for a protein known as arrestin [38,39]. The binding of arrestin further reduces rhodopsin activity and subsequently abolishes its ability to activate transducin (Figure 4). Then PDE is returned to its dark state when the PDE© subunits bind to the catalytic domain blocking the entry to the active site on the PDE***α*** and PDEβ [38]. In addition, the cGMP level rises again following the activation of the transmembrane protein guanylate cyclase (GC) through the Ca^{+2} sensor guanylate-cyclase-activating proteins (GCAPs) (Figure 4). This occurs upon illumination when the GCAPs sense the decline in Ca^{+2} concentration and subsequently activate GC to mediate the synthesis of cGMP [39]. The rise in the level of cGMP leads to the opening of the CNG channel and the subsequent depolarization of the photoreceptor cells (Figure 4). Finally, to regenerate rhodopsin, the all-*trans* retinal must be converted back to all-*cis* [38,39]. The regeneration process is initiated in the photoreceptor cells and then completed in the RPE via a biochemical pathway called the visual cycle [42]. The initial reaction occurs in the photoreceptor cell, where all- *trans*-retinal is converted to all-*trans*-retinol before being transferred into the adjacent RPE [43]. In the RPE, after an esterification step, the all-*trans*-retinol is converted to all-*cis*-retinol by the isomerhydrolase RPE65 [42]. Then it is oxidized back to 11-*cis*-retinal before being sent back to the photoreceptor neurons to be combined with opsin [42].

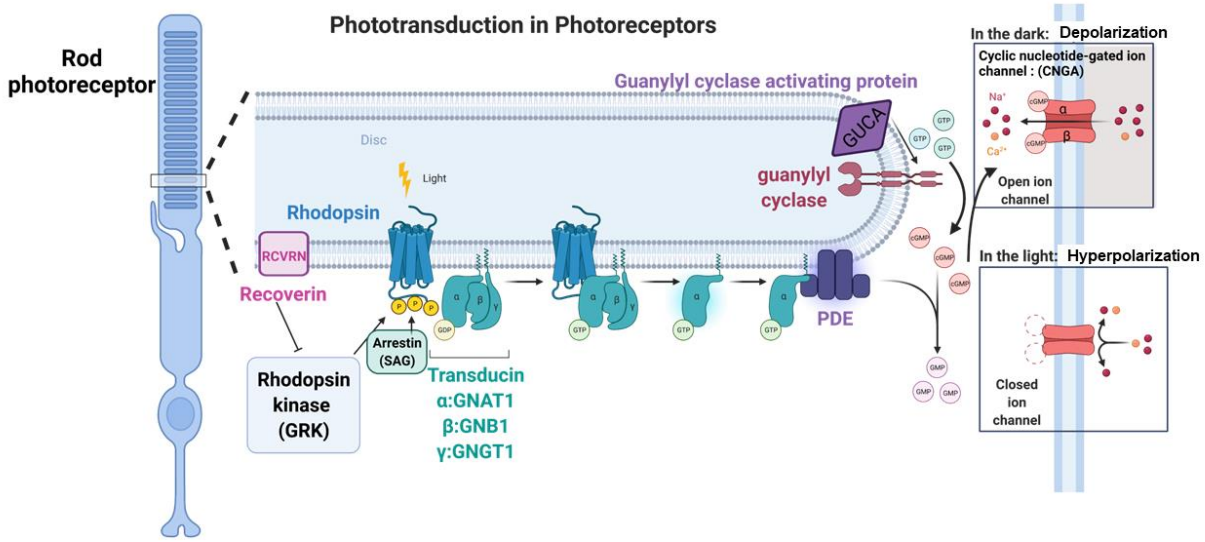


Figure 1-4 Phototransduction cascade in rod photoreceptor cells

Figure 4: Phototransduction cascade in rod photoreceptor cells: Schematic outlines the phototransduction cascade in the vertebrate's rod photoreceptor cells. The cascade is initiated by the absorption of light photons by the chromophore within the rhodopsin molecule leading to its isomerization and the activation of the phototransduction cascade resulting in the closure of the cGMP-gated channel and the subsequent hyperpolarization of the photoreceptor cells. Figure created using Biorender.com

1.3 The Musashi RNA binding proteins

The Musashi (MSI) family is an evolutionary conserved group of RNA-binding proteins in metazoans. Two Musashi paralogs, MSI1 and MSI2, exist in vertebrates originating from gene duplication events [44]. Thus, the invertebrates contain only the gene of the first member (MSI1). Prokaryotes, plants, and single-cell organisms lack the Musashi gene, suggesting that the Musashi are required in metazoans [45]. The characterization of the first member of the Musashi family (MSI1) was in *Drosophila* in 1994. The *Drosophila* Musashi (d-MSI) is highly expressed in various precursor cells of the embryonic CNS and was found to be essential for the asymmetric cell division of the ectodermal sensory hair-like structure known as the sensory organ precursors (SOPs) [46]. The sensory organ precursor cell normally goes through two successive asymmetric cell divisions giving rise to a bristle, a hair-like structure composed of a shaft, a socket, a glia, and a neuronal cell. However, a mutation affecting the d-MSI leads to a double shaft and double socket phenotype at the expense of the glia and neuronal cells [46]. This results from the loss of the asymmetric division of the SOPs, causing both secondary precursor cells to assume the same fate [46]. The resemblance of the double shaft phenotype to the great samurai Miyamoto Musashi's two swords leads to the gene's name after the samurai [46].

The Musashi protein was implicated in the function of the stem cells of a more primitive invertebrate such as the ascidians [47]. Similarly, in vertebrates, the Musashi are highly enriched in the developing central nervous system (CNS) and an essential element in controlling the stem and progenitor cell function. Notably, the expression of MSI1 and MSI2 was mainly detected in the proliferating precursor cells localized within the ventricular zone (VZ) of the neural tube in the embryos and the neurogenic sites within the postnatal brain, including the subventricular zone (SVZ), olfactory bulb, and the ependyma [48–51]. The Musashi proteins were also identified in other embryonic and adult stem cells such as the eye, intestine, skin, breast, and hematopoietic cells [52–57]. With their preferential expression in neuronal stem cells, the Musashi proteins have

emerged as developmental regulators maintaining the self-renewal capacity of stem cells. AlthoughF Musashi was first described in stem cells and in proliferative neural progenitor cells, it is also expressed in postmitotic neurons [51]. For example, the Musashi proteins have been observed to maintain high expression levels in the postmitotic photoreceptors within the eye of both invertebrates and vertebrates [52–54,58].

1.3.1 Musashi structure

The Musashi proteins share high sequence similarity at the primary and tertiary structures, with a 68% amino acid identity on the overall structure (Figure 5). The Musashi family belongs to the heterogeneous nuclear ribonuclear class of proteins (hnRNPA/B) characterized by having two copies of RNA recognition motifs that mediate the binding to mRNA. Both MSI1 and MSI2 contain two RNA recognition motifs (RRM1 and RRM2) tandemly positioned at the N-terminus. The RRM is one of the highly conserved folds that has been shown to interact with nucleotide sequences ranging from two to eight nucleotides in length [59]. Like other RRMs, each of the Musashi RRMs consists of four antiparallel β-sheets wrapped by two $\boldsymbol{\alpha}$-helices and contain the highly conserved ribonucleoprotein type-1 (RNP1) and ribonucleoprotein type-1 (RNP2) consensus sequences (Figure 5).The consensus sequences of the RNP1/2 are characterized by the presence of aliphatic and aromatic amino acids (Figure 5, yellow highlight) which provides the surface for the interaction and play an important role in substrate recognition [60]. The MSI1 binding motif was determined to include a consensus sequence of (G/A) UnAGU by *in vitro* selection tool using a pool of degenerate 50-mer sequences [61]. Further structural and biochemical studies demonstrated the preference of Musashi binding to a core motif of r(GUAG) and r(UAG) to be mediated by RRM1 and RRM2, respectively [59,60]. Mutation analysis has shown that nucleotides outside the UAG core motif have little impact on the binding energy [59].

Aromatic base stacking and intermolecular hydrogen bonding between the core nucleotides r(UAG) and the conserved aromatic amino acids within the RNP1/2 region are the

two types of interactions that initiate substrate recognition [60,62]. Notably, the important residues for RNA-binding are highly conserved in MSI1 and MSI2 proteins (Figure 5-highlighted in yellow). This raises the possibility that both MSI1 and MSI2 may recognize the same targets and explains the frequently reported functional redundancy between the two paralogs, a possibility that is supported by the observation that the knockdown of Msi*2* in the *Msi1*$^{-/-}$ neuronal stem cells significantly decreases neurosphere formation [51]. The C-terminals of MSI1 and MSI2 are different in length, and the amino acid sequence share only 56% identity. Three protein-interacting domains were identified in the C-terminal of the MSI1, including the PolyA- binding protein domain (PABPD), the LIN-28 binding domain (LIN28-D), and the GLD2 binding domain [25,36,63]. The PABP and GLD2 bind directly to amino acids 190-240 and 190-220 located at the C-terminal region of MSI1, respectively (Figure 5). However, a search for the binding region for LIN-28 identified 13 amino acids 246-265 as a putative binding site for LIN-28 in MSI1 [63]. Interestingly, the primary sequence of MSI2 lacks the LIN-28 domain (Figure 5), suggesting that the regulation of miRNA biogenesis is probably a unique function of MSI1 [63]. Similarly, only MSI1 and not MSI2 can interact with GLD2 [25]. As per their subcellular localization, the Musashi proteins are present in the cytoplasm and the nuclear compartment. Two putative nuclear localization signal (NLS) sequences are present in the RRMs of the two paralogues (Figure 5). A classical NLS and a peptide-like NLS (Figure 5, red color). Nevertheless, it was postulated that when the Musashi proteins bind their target in the cytoplasm, the NLS is blocked due to their location within the RRMs and subsequently trapping it in the cytoplasm.

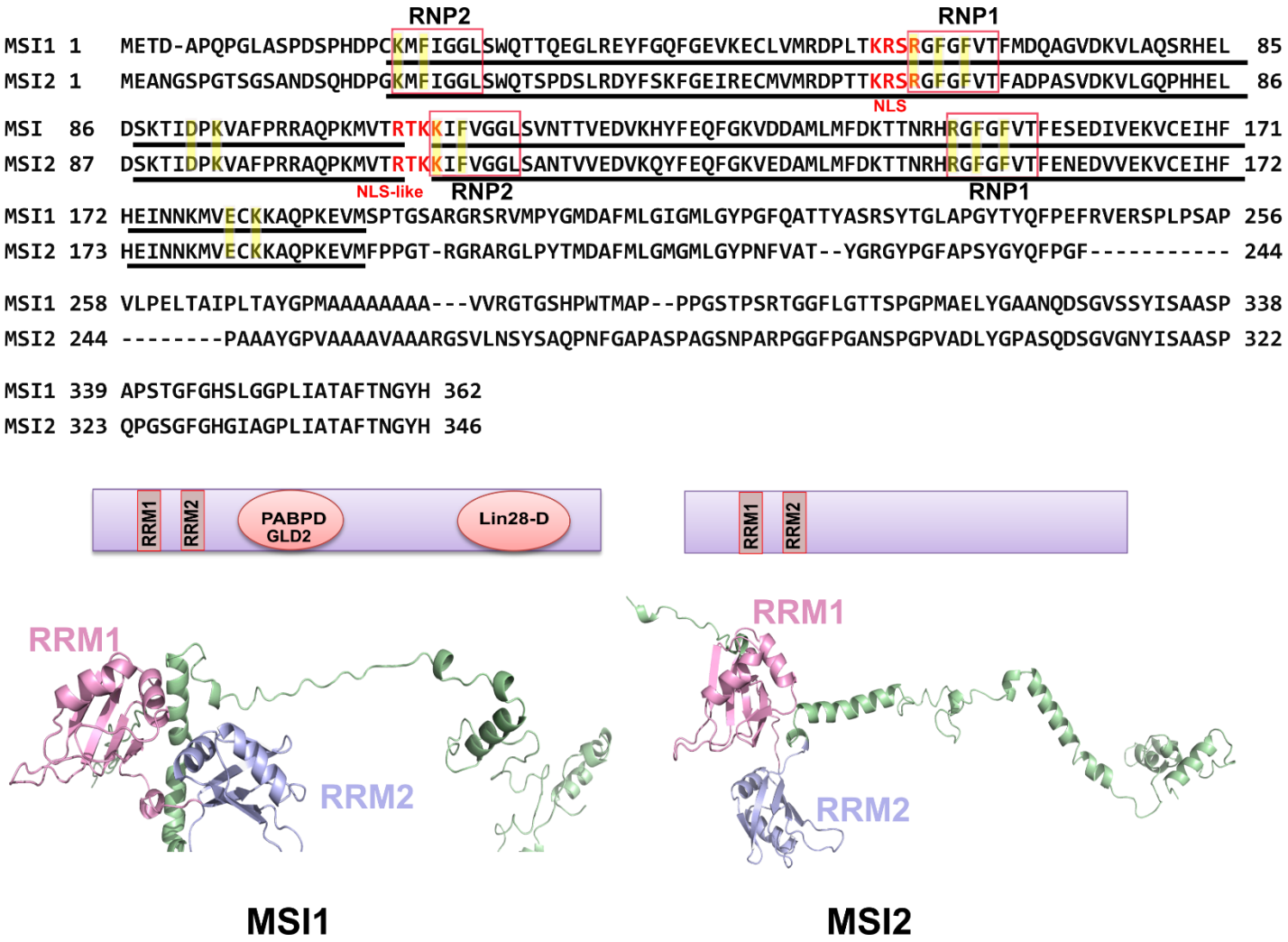


Figure 1-5 Sequence alignment of the mouse MSI1 and MSI2 and the 3D structural prediction by Robetta

Figure 5: Sequence alignment of the mouse MSI1 and MSI2 and the 3D structural prediction by Robetta. Identical residues within the RNP1 and RNP2 of Msi1 and MSI2 are highlighted in yellow. In the 3D structure, the MSI1/2 RRM1 and RRM2 domains are indicated in the pink and blue colors, respectively. Although NMR has resolved the RRMs structure, the full-length 3D structure is still to be elucidated. Here Robetta was used to predict the overall 3D structure of the mouse MSI1 and MSI2.

1.3.2 Mechanisms of posttranscriptional regulation by Musashi proteins

The Musashi proteins are localized to the cytoplasm and nucleus, supporting the notion that they regulate various processes in mRNA metabolism. Consistent with their expression in the cytoplasm, the Musashi proteins have emerged as master regulators of translation. It has been shown that the Musashi can positively or negatively regulate the translation of their target transcripts without affecting their mRNA level. For example, while MSI1 acts as a translational repressor in the mammalian cells, it switches to an activator of translation in the context of the Xenopus oocytes [36,64]. This was shown elegantly when the Xenopus and mouse MSI1 was expressed in mammalian cells; they repress translation. However, they act as an activator of translation when expressed in Xenopus oocytes, suggesting that the switch is not due to species-specific differences in the protein but instead is context-dependent [64]. The proposed mechanism of translational repression in mammalian cells involves the direct binding of MSI1 via a region within its C-terminal half to the poly-A binding protein (PABP) [36]. Previous work by Kawahara et al. showed that the binding of MSI1 to the 3' end of *m-Numb* mRNA, which encodes Notch antagonist, inhibits its translation at the initiation step. The proposed model includes the simultaneous binding of MSI1 via its RRM to the 3' end of its target transcript and to the Poly (A) binding protein (PABP) via its C-terminal domain, competing with the eIF4G for the PABP and the subsequent inhibition of 80S ribosome complex formation [36].

In contrast, Cragle and MacNicol reported in the context of the *Xenopus* oocytes that MSI1 interacting with the Poly (A) binding protein (PABP) stimulates translation of its target transcript, acting as an activator rather than an inhibitor of translation [65]. However, the mechanism of activation proposed by Cragle and McNicol is not fully understood. One of the proposed mechanisms for the bifunctional behavior of Musashi is either through posttranslational modification of the Musashi protein and/or the association with different protein co-factors [64,65]. The Musashi has also been linked to the cytoplasmic adenylation apparatus. Previous work by

Cragle et al. showed that MSI1 could interact with the cytoplasmic poly(A) polymerase GLD-2 [25]. The researchers mapped the interacting domain to 31-amino acids within the C-terminal region of MSI1 (amino acids 190–220), which lies directly outside the two N-terminal RNA recognition motifs of MSI1 (Figure 5). Therefore, the authors proposed that MSI1, through its interaction with GLD-2, serves as the guide to selectively recruit GLD-2 to specific mRNA transcripts, promoting their cytoplasmic adenylation and translation, and subsequently linking the MSI1 protein to the polyadenylation apparatus via its interaction with GLD-2.

In the nucleus, the Musashi proteins have been shown to regulate alternative splicing and to be involved in the miRNA biogenesis. The binding of MSI1 to the LIN-28 blocks neuronal differentiation by inhibiting the maturation of the let-7 family of miRNA miR98 at the cropping step [63]. Alternatively, our lab, among others, demonstrated a role for the Musashi proteins in regulating alternative splicing in photoreceptor cells [52,66–68]. We show that the Musashi regulate the splicing of a set of pre-mRNA, many of which belong to cilia-related genes, to produce photoreceptor-specific isoforms [52,68].

1.3.3 Musashi in photoreceptor cells

In addition to their predominant expression in neuronal stem and progenitor cells, the Musashi proteins have been shown to be expressed in a small population of postmitotic neurons, including the photoreceptor cells [51]. The expression of the Musashi proteins in mature photoreceptors has been validated in animal models, including *Xenopus Leavis*, mouse retina, and newt retina [53,69–71]. In the developing mouse retina, the subcellular localization of MSI1 changes from being exclusively cytoplasmic in retinal progenitor cells to transiently become nuclear in differentiating cells and then shifting predominately perinuclear and cytoplasmic in the mature retina [69]. A critical role for MSI1 in retinal development and survival has been established in both Drosophila and mammals [52,54,58]. In Drosophila, loss of function in d-MSI resulted in weak abnormalities in the photoreceptor differentiation. However, the phenotype significantly

increased in the background of seven absentia (Sina) knockouts [58]. It has been shown that in the developing Drosophila eye, both the d-MSI and Sina function redundantly by controlling the expression level of Tramtrack (Ttk), a transcriptional regulator of neuronal cell fate [58]. As such, loss of function in d-MSI and Sina resulted in ectopic expression of Ttk that was accompanied by a failure in photoreceptor specification [58]. In mammals, the functional role of Musashi proteins in photoreceptors has been investigated following constitutive and conditional germline deletions. For example, Susaki et al. [9] show that global embryonic deletion of the *Msi1* gene results in the eventual death of photoreceptors. They attributed the photoreceptor cell death to a non-cell autonomous effect caused by a malfunction in the retinal pigmented epithelium (RPE). They proposed that the failure of RPE65, a visual cycle protein, to localize to the microvilli of RPE has influenced photoreceptor survival. Building upon this work, our previous work showed the combined deletion of *Msi1* and *Msi2* in photoreceptor progenitor cells caused severe morphological defects in the outer segment (OS) that was accompanied by progressive degeneration, which ultimately led to the loss of vision [52]. In addition to its cytoplasmic expression, our published work shows the Musashi proteins are also localized in the photoreceptor nuclei [66]. We show that besides being a translational regulator, the Musashi proteins function as splicing regulators promoting the inclusion of a set of alternatively splice exons to generate a photoreceptor-specific isoforms [52,66]. The localization of the Musashi proteins to the nuclei and cytoplasm of photoreceptor cells suggests that the Musashi proteins might have a compound role where they regulate mRNA splicing in the nucleus and its translation in the cytoplasm.

1.4 Summary and aims of the thesis

The Musashi proteins are widely known to support stem cell maintenance and self-renewal. By being highly expressed in stem cells, most of the previous work had focused on characterizing the Musashi protein's mechanistic role in stem cells of various cellular systems. Few, if any, had investigated their role in postmitotic tissues. Germline and targeted embryonic deletion of *Msi1/2* established a role for Musashi proteins in photoreceptor function and survival and supported functional redundancy between the two paralogs [52,54,58].

Interestingly, our recently published results showed that the Musashi proteins maintain an exceptionally high protein level in the mature retina [52]. We observed the expression of the two paralogs to be developmentally regulated, with MSI1 predominating early around postnatal day 4 (PN4) and MSI2 prevailing after PN13 [52]. Single deletion of *Msi1* or *Msi2* in committed rod progenitor cells showed that the two paralogs are partially redundant and appear to act differently at different time points of retinal development [52]. As such, the single deletion of *Msi1* resulted in an early visual defect that was observed at the time point of eye-opening in mice (PN16), while the removal of *Msi2* resulted in a visual defect that was detected at a later time point when photoreceptors were fully developed [52]. Building upon these findings, we hypothesize that MSI2 may have a unique role in maintaining mature photoreceptor cells, with MSI1 unable to compensate for MSI2 Loss. To test our hypothesis, we sought to investigate the functional role of the Musashi individually and in combination in mature photoreceptors.

To understand the role of the Musashi proteins in mature photoreceptor cells, we utilized an inducible mouse model that allows for deleting Musashi proteins from mature photoreceptors. We generated single and combined deletion of *Msi1* and *Msi2*. Additionally, how the Musashi regulate their targets, what processes they regulate, or what targets they bind in photoreceptor cells is not fully understood. To answer these questions, I used the genome-wide sequencing

approach eCLIP, RNA sequencing, and proteomics analysis to define the mechanistic role of the Musashi in the retina. This is the first study that combines all three global wide analyses to delineate the mechanistic role of the Musashi proteins in the mouse retina.

Chapter 2 describes the functional role of MSI1 and MSI2 in mature photoreceptor cells and illustrates the binding interactions and potential targets of Musashi proteins in the retina as determined through eCLIP-seq, RNA sequencing, and proteomics analysis [72]. The results provide insights into the molecular mechanisms mediating the functional role of Musashi in photoreceptors [72]. Chapter 3 illustrates the efforts to map the epitope of the widely used rat monoclonal antibody mAB 13A4 to the mouse Prominin-1 protein [73]. We show that the mAB 13A4 recognizes a structural epitope that gets disrupted by the inclusion of the photoreceptor-specific exon (exon 19), interfering with Prominin-1 detection on a Western blot analysis [73]. In chapter 4, I discuss the results and conclude by suggesting experimentation that builds upon and extends this study's framework.

1.5 References

1. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nat Rev Genet. 2014;15: 829–845. doi:10.1038/nrg3813

2. Hentze MW, Castello A, Schwarzl T, Preiss T. A brave new world of RNA-binding proteins. Nat Rev Mol Cell Biol. 2018;19: 327–341. doi:10.1038/nrm.2017.130

3. Gebauer F, Schwarzl T, Valcárcel J, Hentze MW. RNA-binding proteins in human genetic disease. Nat Rev Genet. 2021;22: 185–198. doi:10.1038/s41576-020-00302-y

4. Wilkinson ME, Charenton C, Nagai K. RNA Splicing by the Spliceosome. Annu Rev Biochem. 2020;89: 359–388. doi:10.1146/annurev-biochem-091719-064225

5. Fu X-D, Ares M. Context-dependent control of alternative splicing by RNA-binding proteins. Nat Rev Genet. 2014;15: 689–701. doi:10.1038/nrg3778

6. Glisovic T, Bachorik JL, Yong J, Dreyfuss G. RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett. 2008;582: 1977–1986. doi:10.1016/j.febslet.2008.03.004

7. Black DL. Mechanisms of Alternative Pre-Messenger RNA Splicing. Annu Rev Biochem. 2003;72: 291–336. doi:10.1146/annurev.biochem.72.121801.161720

8. Patel AA, McCarthy M, Steitz JA. The splicing of U12-type introns can be a rate-limiting step in gene expression. EMBO J. 2002;21: 3804–3815. doi:10.1093/emboj/cdf297

9. Gehring NH, Roignant J-Y. Anything but Ordinary – Emerging Splicing Mechanisms in Eukaryotic Gene Regulation. Trends Genet. 2021;37: 355–372. doi:10.1016/j.tig.2020.10.008

10. Borao S, Ayté J, Hümmer S. Evolution of the Early Spliceosomal Complex—From Constitutive to Regulated Splicing. Int J Mol Sci. 2021;22: 12444. doi:10.3390/ijms222212444

11. Akinyi MV, Frilander MJ. At the Intersection of Major and Minor Spliceosomes: Crosstalk Mechanisms and Their Impact on Gene Expression. Front Genet. 2021;12. Available: https://www.frontiersin.org/article/10.3389/fgene.2021.700744

12. Matera AG, Wang Z. A day in the life of the spliceosome. Nat Rev Mol Cell Biol. 2014;15: 108–121. doi:10.1038/nrm3742

13. De Conti L, Baralle M, Buratti E. Exon and intron definition in pre-mRNA splicing. WIREs RNA. 2013;4: 49–60. doi:10.1002/wrna.1140

14. Will CL, Lührmann R. Spliceosome Structure and Function. Cold Spring Harb Perspect Biol. 2011;3: a003707. doi:10.1101/cshperspect.a003707

15. Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods. 2016;13: 508–514. doi:10.1038/nmeth.3810

16. Lee Y, Rio DC. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu Rev Biochem. 2015;84: 291–323. doi:10.1146/annurev-biochem-060614-034316

17. Yee BA, Pratt GA, Graveley BR, Nostrand ELV, Yeo GW. RBP-Maps enables robust generation of splicing regulatory maps. RNA. 2019;25: 193–204. doi:10.1261/rna.069237.118

18. Marsollier A-C, Joubert R, Mariot V, Dumonceaux J. Targeting the Polyadenylation Signal of Pre-mRNA: A New Gene Silencing Approach for Facioscapulohumeral Dystrophy. Int J Mol Sci. 2018;19: 1347. doi:10.3390/ijms19051347

19. Charlesworth A, Meijer HA, de Moor CH. Specificity factors in cytoplasmic polyadenylation. WIREs RNA. 2013;4: 437–461. doi:10.1002/wrna.1171

20. Moore KS, von Lindern M. RNA Binding Proteins and Regulation of mRNA Translation in Erythropoiesis. Front Physiol. 2018;9. Available: https://www.frontiersin.org/article/10.3389/fphys.2018.00910

21. Nicholson AL, Pasquinelli AE. Tales of Detailed Poly(A) Tails. Trends Cell Biol. 2019;29: 191–200. doi:10.1016/j.tcb.2018.11.002

22. Richter JD. Cytoplasmic Polyadenylation in Development and Beyond. Microbiol Mol Biol Rev. 1999;63: 446–456. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC98972/

23. Rouhana L, Wang L, Buter N, Kwak JE, Schiltz CA, Gonzalez T, et al. Vertebrate GLD2 poly(A) polymerases in the germline and the brain. RNA. 2005;11: 1117–1130. doi:10.1261/rna.2630205

24. Villalba A, Coll O, Gebauer F. Cytoplasmic polyadenylation and translational control. Curr Opin Genet Dev. 2011;21: 452–457. doi:10.1016/j.gde.2011.04.006

25. Cragle C, MacNicol AM. Musashi Protein-directed Translational Activation of Target mRNAs Is Mediated by the Poly(A) Polymerase, Germ Line Development Defective-2. J Biol Chem. 2014;289: 14239–14251. doi:10.1074/jbc.M114.548271

26. Mendez R, Richter JD. Translational control by CPEB: a means to the end. Nat Rev Mol Cell Biol. 2001;2: 521–529. doi:10.1038/35080081

27. Nakel K, Bonneau F, Basquin C, Habermann B, Eckmann CR, Conti E. Structural basis for the antagonistic roles of RNP-8 and GLD-3 in GLD-2 poly(A)-polymerase activity. RNA. 2016;22: 1139–1145. doi:10.1261/rna.056598.116

28. Łabno A, Tomecki R, Dziembowski A. Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim Biophys Acta BBA - Mol Cell Res. 2016;1863: 3125–3147. doi:10.1016/j.bbamcr.2016.09.023

29. García-Mauriño SM, Rivero-Rodríguez F, Velázquez-Cruz A, Hernández-Vellisca M, Díaz-Quintana A, De la Rosa MA, et al. RNA Binding Protein Regulation and Cross-Talk in the Control of AU-rich mRNA Fate. Front Mol Biosci. 2017;4. Available: https://www.frontiersin.org/article/10.3389/fmolb.2017.00071

30. Meyer S, Temme C, Wahle E. Messenger RNA Turnover in Eukaryotes: Pathways and Enzymes. Crit Rev Biochem Mol Biol. 2004;39: 197–216. doi:10.1080/10409230490513991

31. Wiederhold K, Passmore LA. Cytoplasmic deadenylation: Regulation of mRNA fate. Biochem Soc Trans. 2010;38: 1531–1536. doi:10.1042/BST0381531

32. Zhang F, Wang D. The Pattern of microRNA Binding Site Distribution. Genes. 2017;8: 296. doi:10.3390/genes8110296

33. Valinezhad Orang A, Safaralizadeh R, Kazemzadeh-Bavili M. Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. Int J Genomics. 2014;2014: 970607. doi:10.1155/2014/970607

34. Fukao A, Tomohiro T, Fujiwara T. Translation Initiation Regulated by RNA-Binding Protein in Mammals: The Modulation of Translation Initiation Complex by Trans-Acting Factors. Cells. 2021;10: 1711. doi:10.3390/cells10071711

35. Kong J, Lasko P. Translational control in cellular and developmental processes. Nat Rev Genet. 2012;13: 383–394. doi:10.1038/nrg3184

36. Kawahara H, Imai T, Imataka H, Tsujimoto M, Matsumoto K, Okano H. Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. J Cell Biol. 2008;181: 639–653. doi:10.1083/jcb.200708004

37. Gebauer F, Hentze MW. Molecular mechanisms of translational control. Nat Rev Mol Cell Biol. 2004;5: 827–835. doi:10.1038/nrm1488

38. Fu Y, Yau K-W. Phototransduction in mouse rods and cones. Pflugers Arch. 2007;454: 805–819. doi:10.1007/s00424-006-0194-y

39. Molday RS, Moritz OL. Photoreceptors at a glance. J Cell Sci. 2015;128: 4039–4045. doi:10.1242/jcs.175687

40. Pearring JN, Salinas RY, Baker SA, Arshavsky VY. Protein sorting, targeting and trafficking in photoreceptor cells. Prog Retin Eye Res. 2013;0: 24–51. doi:10.1016/j.preteyeres.2013.03.002

41. LoGiudice L, Matthews G. The Role of Ribbons at Sensory Synapses. Neurosci Rev J Bringing Neurobiol Neurol Psychiatry. 2009;15: 380–391. doi:10.1177/1073858408331373

42. Kiser PD, Golczak M, Maeda A, Palczewski K. Key enzymes of the retinoid (visual) cycle in vertebrate retina. Biochim Biophys Acta. 2012;1821: 137–151. doi:10.1016/j.bbalip.2011.03.005

43. Kiebler MA, Scheiffele P, Ule J. What, where, and when: the importance of post-transcriptional regulation in the brain. Front Neurosci. 2013;7: 192. doi:10.3389/fnins.2013.00192

44. Fox RG, Park FD, Koechlein CS, Kritzik M, Reya T. Musashi signaling in stem cells and cancer. Annu Rev Cell Dev Biol. 2015;31: 249–267. doi:10.1146/annurev-cellbio-100814-125446

45. Horisawa K, Yanagawa H. Musashi Proteins in Neural Stem/Progenitor Cells. Neural Stem Cells and Therapy. IntechOpen; 2012. doi:10.5772/31033

46. Nakamura M, Okano H, Blendy JA, Montell C. Musashi, a neural RNA-binding protein required for drosophila adult external sensory organ development. Neuron. 1994;13: 67–81. doi:10.1016/0896-6273(94)90460-X

47. Expression patterns of musashi homologs of the ascidians, Halocynthia roretzi and Ciona intestinalis - PubMed. [cited 29 Mar 2022]. Available: https://pubmed.ncbi.nlm.nih.gov/11180818/

48. Sakakibara S, Imai T, Hamaguchi K, Okabe M, Aruga J, Nakajima K, et al. Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. Dev Biol. 1996;176: 230–242. doi:10.1006/dbio.1996.0130

49. Sakakibara S, Okano H. Expression of neural RNA-binding proteins in the postnatal CNS: implications of their roles in neuronal and glial cell development. J Neurosci Off J Soc Neurosci. 1997;17: 8300–8312.

50. Sakakibara S, Nakamura Y, Satoh H, Okano H. Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. J Neurosci Off J Soc Neurosci. 2001;21: 8091–8107.

51. Sakakibara S -i., Nakamura Y, Yoshida T, Shibata S, Koike M, Takano H, et al. RNA-binding protein Musashi family: Roles for CNS stem cells and a subpopulation of ependymal cells revealed by targeted disruption and antisense ablation. Proc Natl Acad Sci. 2002;99: 15194–15199. doi:10.1073/pnas.232087499

52. Sundar J, Matalkah F, Jeong B, Stoilov P, Ramamurthy V. The Musashi proteins MSI1 and MSI2 are required for photoreceptor morphogenesis and vision in mice. J Biol Chem. 2020. doi:10.1074/jbc.RA120.015714

53. Raji BA, Dansault A, Leemput J, Houssaye G de la, Vieira V, Kobetz A, et al. The RNA-binding protein Musashi-1 is produced in the developing and adult mouse eye. Mol Vis. 2007.

54. Susaki K, Kaneko J, Yamano Y, Nakamura K, Inami W, Yoshikawa T, et al. Musashi-1, an RNA-binding protein, is indispensable for survival of photoreceptors. Exp Eye Res. 2009;88: 347–355. doi:10.1016/j.exer.2008.06.019

55. Li N, Yousefi M, Nakauka-Ddamba A, Li F, Vandivier L, Parada K, et al. The Msi Family of RNA-Binding Proteins Function Redundantly as Intestinal Oncoproteins. Cell Rep. 2015;13: 2440–2455. doi:10.1016/j.celrep.2015.11.022

56. Katz Y, Li F, Lambert NJ, Sokol ES, Tam W-L, Cheng AW, et al. Musashi proteins are post-transcriptional regulators of the epithelial-luminal cell state. eLife. 2014;3: e03915. doi:10.7554/eLife.03915

57. Park S-M, Deering RP, Lu Y, Tivnan P, Lianoglou S, Al-Shahrour F, et al. Musashi-2 controls cell fate, lineage bias, and TGF-β signaling in HSCs. J Exp Med. 2014;211: 71–87. doi:10.1084/jem.20130736

58. Hirota Y, Okabe M, Imai T, Kurusu M, Yamamoto A, Miyao S, et al. Musashi and Seven in absentia downregulate Tramtrack through distinct mechanisms in Drosophila eye development. Mech Dev. 1999;87: 93–101. doi:10.1016/S0925-4773(99)00143-4

59. Zearfoss NR, Deveau LM, Clingman CC, Schmidt E, Johnson ES, Massi F, et al. A Conserved Three-nucleotide Core Motif Defines Musashi RNA Binding Specificity. J Biol Chem. 2014;289: 35530–35541. doi:10.1074/jbc.M114.597112

60. Ohyama T, Nagata T, Tsuda K, Kobayashi N, Imai T, Okano H, et al. Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions. Nucleic Acids Res. 2012;40: 3218. doi:10.1093/nar/gkr1139

61. Imai T, Tokunaga A, Yoshida T, Hashimoto M, Mikoshiba K, Weinmaster G, et al. The Neural RNA-Binding Protein Musashi1 Translationally Regulates Mammalian numb Gene Expression by Interacting with Its mRNA. Mol Cell Biol. 2001;21: 3888–3900. doi:10.1128/MCB.21.12.3888-3900.2001

62. Iwaoka R, Nagata T, Tsuda K, Imai T, Okano H, Kobayashi N, et al. Structural Insight into the Recognition of r(UAG) by Musashi-1 RBD2, and Construction of a Model of Musashi-1 RBD1-2 Bound to the Minimum Target RNA. Mol J Synth Chem Nat Prod Chem. 2017;22. doi:10.3390/molecules22071207

63. Kawahara H, Okada Y, Imai T, Iwanami A, Mischel PS, Okano H. Musashi 1 cooperates in abnormal cell LINeage protein 28 (Lin28)-mediated Let-7 family microRNA biogenesis in early neural differentiation. J Biol Chem. 2011; jbc.M110.199166. doi:10.1074/jbc.M110.199166

64. MacNicol MC, Cragle CE, MacNicol AM. Context-dependent regulation of Musashi-mediated mRNA translation and cell cycle regulation. Cell Cycle Georget Tex. 2011;10: 39–44. doi:10.4161/cc.10.1.14388

65. Cragle CE, MacNicol MC, Byrum SD, Hardy LL, Mackintosh SG, Richardson WA, et al. Musashi interaction with poly(A)-binding protein is required for activation of target mRNA translation. J Biol Chem. 2019;294: 10969–10986. doi:10.1074/jbc.RA119.007220

66. Murphy D, Cieply B, Carstens R, Ramamurthy V, Stoilov P. The Musashi 1 Controls the Splicing of Photoreceptor-Specific Exons in the Vertebrate Retina. PLoS Genet. 2016;12: e1006256. doi:10.1371/journal.pgen.1006256

67. Ling JP, Wilks C, Charles R, Leavey PJ, Ghosh D, Jiang L, et al. ASCOT identifies key regulators of neuronal subtype-specific splicing. Nat Commun. 2020;11: 137. doi:10.1038/s41467-019-14020-5

68. Murphy D, Singh R, Kolandaivelu S, Ramamurthy V, Stoilov P. Alternative Splicing Shapes the Phenotype of a Mutation in BBS8 To Cause Nonsyndromic Retinitis Pigmentosa. Mol Cell Biol. 2015;35: 1860–1870. doi:10.1128/MCB.00040-15

69. Nickerson PEB, Myers T, Clarke DB, Chow RL. Changes in Musashi-1 subcellular localization correlate with cell cycle exit during postnatal retinal development. Exp Eye Res. 2011;92: 344–352. doi:10.1016/j.exer.2011.02.002

70. Kaneko J, Chiba C. Immunohistochemical analysis of Musashi-1 expression during retinal regeneration of adult newt. Neurosci Lett. 2009;450: 252–257. doi:10.1016/j.neulet.2008.11.031

71. Amato MA, Boy S, Arnault E, Girard M, Puppa AD, Sharif A, et al. Comparison of the expression patterns of five neural RNA binding proteins in the Xenopus retina. J Comp Neurol. 2005;481: 331–339. doi:10.1002/cne.20387

72. Matalkah F, Jeong B, Sheridan M, Horstick E, Ramamurthy V, Stoilov P. The Musashi proteins direct post-transcriptional control of protein expression and alternate exon splicing in vertebrate photoreceptors. Commun Biol. 2022;5: 1–15. doi:10.1038/s42003-022-03990-w

73. Matalkah F, Rhodes S, Ramamurthy V, Stoilov P. The mAB 13A4 monoclonal antibody to the mouse PROM1 protein recognizes a structural epitope. PLoS ONE. 2022;17: e0274958. doi:10.1371/journal.pone.0274958

74. Okano H, Kawahara H, Toriya M, Nakao K, Shibata S, Imai T. Function of RNA-binding protein Musashi-1 in stem cells. Exp Cell Res. 2005;306: 349–356. doi:10.1016/j.yexcr.2005.02.021

75. Imai T, Tokunaga A, Yoshida T, Hashimoto M, Mikoshiba K, Weinmaster G, et al. The Neural RNA-Binding Protein Musashi1 Translationally Regulates Mammalian numb Gene Expression by Interacting with Its mRNA. Mol Cell Biol. 2001;21: 3888–3900. doi:10.1128/MCB.21.12.3888-3900.2001

76. Hooper MJ, Wang J, Browning R, Ash JD. Damage-associated molecular pattern recognition is required for induction of retinal neuroprotective pathways in a sex-dependent manner. Sci Rep. 2018;8: 9115. doi:10.1038/s41598-018-27479-x

77. Rattner A, Nathans J. The Genomic Response to Retinal Disease and Injury: Evidence for Endothelin Signaling from Photoreceptors to Glia. J Neurosci. 2005;25: 4540–4549. doi:10.1523/JNEUROSCI.0492-05.2005

78. Rattner A, Nathans J. An Evolutionary Perspective on the Photoreceptor Damage Response. Am J Ophthalmol. 2006;141: 558-562.e2. doi:10.1016/j.ajo.2005.10.045

79. Fargeas CA, Joester A, Missol-Kolka E, Hellwig A, Huttner WB, Corbeil D. Identification of novel Prominin-1/CD133 splice variants with alternative C-termini and their expression in epididymis and testis. J Cell Sci. 2004;117: 4301–4311. doi:10.1242/jcs.01315

80. Fargeas CA, Huttner WB, Corbeil D. Nomenclature of prominin-1 (CD133) splice variants – an update. Tissue Antigens. 2007;69: 602–606. doi:10.1111/j.1399-0039.2007.00825.

2. Chapter 2

2.1 TITLE: The Musashi proteins direct post-transcriptional control of protein expression and alternate exon splicing in vertebrate photoreceptors

Authors: Fatimah Matalkah[1], Bohye Jeong[1], Macie Sheridan[3], Eric Horstick[4,5], Visvanathan Ramamurthy[1,2], and Peter Stoilov[1*]

Affiliations: [1] Department of Biochemistry, [2] Department of Ophthalmology; [4] Department of Biology, and [5] Department of Neuroscience, West Virginia University; Morgantown, West Virginia, USA, 26505;

[3] Undergraduate Program in Biochemistry, West Virginia University; Morgantown, West Virginia, USA, 26505;

Matalkah F, Jeong B, Sheridan M, Horstick E, Ramamurthy V, Stoilov P. The Musashi proteins direct post-transcriptional control of protein expression and alternate exon splicing in vertebrate photoreceptors. *Commun Biol*. 2022 Sep 24;5(1):1011. doi: 10.1038/s42003-022-03990-w. PMID: 36153373; PMCID: PMC9509328.

Note: This chapter include text and figures taken from the publications highlighted above

2.2 Abstract

The Musashi proteins, MSI1 and MSI2, are conserved RNA binding proteins with a role in the maintenance and renewal of stem cells. Contrasting with this role, terminally differentiated photoreceptor cells express high levels of MSI1 and MSI2, pointing to a new role for the two proteins in vision. Combined knockout of *Msi1* and *Msi2* in mature photoreceptor cells abrogated the retinal response to light and caused photoreceptor cell death. In photoreceptor cells, the Musashi proteins perform distinct nuclear and cytoplasmic functions. In the nucleus, the Musashi proteins promote the splicing of photoreceptor-specific alternative exons. Surprisingly, conserved photoreceptor-specific alternative exons in genes critical for vision proved to be dispensable, raising questions about the selective pressures that lead to their conservation. In the cytoplasm, MSI1 and MSI2 activate protein expression. Loss of *Msi1* and *Msi2* leads to a reduction in the levels of multiple proteins including proteins required for vision and photoreceptor survival.

2.3 Introduction

Mammals express approximately 500 RNA binding proteins that associate with Polymerase II transcripts to regulate pre-mRNA processing, mRNA localization, mRNA stability, and translation [1]. Networks of RNA binding proteins specific to, or highly expressed in neurons perform roles that range from diversifying the transcriptome through alternative splicing and poly-adenylation to directing transport of specific mRNA to cellular compartments for localized translation [1–4]. RNA binding proteins are essential for the development and function of the nervous system where processes such as axon guidance, synaptic plasticity, cell survival and cell excitation are tuned by their activity [4]. Many RNA binding proteins belong to families of orthologs with varying degrees of sequence homology and functional redundancy. Interestingly, even orthologs with highly similar sequences and biochemical properties are not fully redundant and can have distinct roles in the nervous system. This divergence in function can be derived from differences in expression levels across different cell types, subcellular localization, or their interactomes [5–8].

Neurons stand out among other cell types by the pervasive use of alternative exons [9,10]. The large number of alternatively spliced exons used in neurons is due to the absence of a major splicing repressor, PTBP1, and the expression of neuron specific splicing factors [4]. While the importance of neuronal splicing programs is well established through knockouts of splicing regulators, the functions of the many alternative exons are less clear. Functional significance of individual exons is commonly assigned based on sequence conservation and the nature of the protein hosting the exon. More recently, an empirical picture of the functional impact of individual exons has started to emerge. Functions of individual exons range from fine-tuning protein interactions to regulatory switches that shut down protein expression [11–13]. Interestingly, in several cases, deletion of conserved alternative exons has failed to produce an obvious phenotype in

mouse models *in vivo* [14–17].

Among the RNA binding proteins expressed in neurons are the Musashi family of proteins [18–20]. The founding member of the family, the *Drosophila Msi* protein was first described as a factor that maintains the undifferentiated state of stem cells by repressing the translation of the *Notch* regulator *Numb* [21]. This function of Musashi is preserved in vertebrates where its homologues, MSI1 and MSI2, are required for stem cell maintenance and are investigated for their role in cancer progression [22–25]. Subsequent studies showed that the effect of the Musashi proteins on translation is context dependent, and they can positively or negatively regulate protein translation by binding to the 3'-UTR of their target mRNAs [26–29].

We showed that in photoreceptor cells the Musashi proteins regulate alternative splicing to produce highly photoreceptor-specific isoforms of ubiquitously expressed proteins [30,31]. The Musashi proteins maintain exceptionally high protein levels in the mature retina and their expression is developmentally regulated [31]. MSI1 levels rise sharply after birth, peak between postnatal days 2 and 4, and decline afterwards. Concomitant with the decline of MSI1 protein levels, the levels of MSI2 increase and remain constant in adulthood [31]. Single deletion of *Msi1* or *Msi2* in committed rod photoreceptor progenitor cells showed that the two paralogs are partially redundant and appear to act at time-points of retinal development that correlate with the pattern of their expression [31]. The single deletion of *Msi1* results in an early visual defect that was observed at the time of eye-opening in mice (postnatal day 16). In contrast, the removal of *Msi2* resulted in normal vision at postnatal day 16 that progressively declined with age [31]. Based on these findings, we proposed that MSI2 is involved in the maintenance of mature photoreceptor cells, while MSI1 functions in photoreceptor precursors and immature photoreceptor cells.

Using an inducible mouse model that deletes the Musashi genes in mature photoreceptors we show the Musashi proteins to be essentials for the function and viability of the photoreceptors. To our surprise, despite their reciprocal regulation during development, MSI1 and MSI2 are fully redundant in mature photoreceptors. We identified the transcripts recognized by MSI1 and MSI2

in vivo and investigated how loss of the *Msi1* and *Msi2* genes affected pre-mRNA splicing, transcript levels, and protein expression. We demonstrate that the Musashi proteins bind downstream of photoreceptor-specific exons to activate their splicing. In addition, we show that in photoreceptors the Musashi proteins act almost exclusively as post-transcriptional activators of protein expression.

2.4 Results

Depletion of *Msi1* and *Msi2* in mature photoreceptor cells

We recently showed that in the retina the MSI1 and MSI2 proteins are differentially regulated and proposed that they have separate roles in development and maintenance of photoreceptor cells [31]. We tested the roles of the Musashi protein in mature photoreceptors by using tamoxifen-inducible Cre^{ERT2} under the control of rod-specific *Pde6g* promoter to remove *Msi1* and *Msi2* in mature rod-photoreceptor cells [32]. Floxed ($Msi1^{flox/flox}$/$Msi2^{flox/flox}$) mice hemizygous for *Pde6g-*Cre^{ERT2} were injected with tamoxifen for three consecutive days starting at postnatal day 30 to create combined *Msi1*/*Msi2* knockout mice. Littermates carrying the floxed alleles for *Msi1* and *Msi2* ($Msi1^{flox/flox}$/$Msi2^{flox/flox}$) and treated with tamoxifen were used as controls. We will refer to the $Msi1^{flox/flox}$/$Msi2^{flox/flox}$ treated with tamoxifen as $Msi1^{+/+}$/$Msi2^{+/+}$ and the mice with $Msi1^{flox/flox}$/$Msi2^{flox/flox}$-*Pde6gCre*^{ERT2} treated with tamoxifen as $Msi1^{-/-}$/$Msi2^{-/-}$. Immunocytochemistry (ICC) demonstrated that, 14 days after the first tamoxifen injection, the MSI1 and MSI2 proteins were depleted specifically from the photoreceptors (Figure 1A). The immunofluorescence signal is lost from both the cytoplasm (inner segment, IS) and the nuclei (outer nuclear layer, ONL) of the photoreceptor cells. Consistent with the ICC data, western blot analysis showed two-fold decrease in the MSI1 and MSI2 protein levels in retinal lysates from the $Msi1^{-/-}$/$Msi2^{-/-}$ mice (Figure 1B and Supplement Figure 9).

MSI1 and MSI2 are required for the function and survival of mature photoreceptors

To evaluate the functional significance of Musashi proteins in mature photoreceptors, we used electroretinograms (ERG) to measure the retinal response to light. We measured the dark-adapted (scotopic) and the light-adapted (photopic) responses that reflect rod and cone photoreceptor function, respectively [33]. We used repeated measures two-way ANOVA to determine the effect of the genotype and time post injection on the ERG A-Wave amplitude. We found a significant interaction of the genotype and the time after injection (Scotopic response: $F(12,1)= 19.47$, p-value<0.0001; Photopic response: $F(12,1)=10.37$, p-value<0.0001). The response to light of the *Msi1*$^{-/-}$/*Msi2*$^{-/-}$ animals and the age-matched *Msi1*$^{+/+}$/*Msi2*$^{+/+}$ controls became significantly different 35 days after tamoxifen injection (Figure 2A, B). The response to light continued to decrease rapidly thereafter and was nearly undetectable by day 105 post-injection (Figure 2A, B).

To assess the retinal morphology following *Msi1/Msi2* deletion, we performed hematoxylin and eosin (H&E) staining on retinal cross-sections collected from *Msi1*$^{-/-}$/*Msi2*$^{-/-}$ and control mice between days 0 and 113 post-injection (Figure 2C, D, and Supplement Figure 2). A significant effect of the genotype on the photoreceptor cell layer thickness over time was confirmed by two-way ANOVA ($F(1,259)=61.04$, p-value<0.0001). In agreement with the ERG data, we did not observe significant morphological changes up to 28 days after tamoxifen injection (Figure 2C, D, and Supplement Figure 2). After day 28 post-injection, we observed progressive degeneration of the photoreceptor cell layer (Figure 2D, and Supplement Figure 2). Approximately half of the photoreceptor cells were lost by day 42 post-injection in the knockout retinas, and only one layer of photoreceptor cells remained at day 113 post-injection (Figure 2C, D, and Supplement Figure 2). We did not observe any significant changes in the inner retina, including the inner nuclear layer (INL) and the ganglion cell layer (GCL). Our results show that the combined deletion of *Msi1/Msi2* in mature photoreceptors leads to a rapid and progressive decline in the function and

viability of photoreceptor cells that starts four weeks after depletion of the Musashi proteins.

MSI1 and MSI2 are redundant in the maintenance of mature photoreceptors

To determine if MSI2 plays a dominant role in mature retina, as the developmental regulation of MSI1 and MSI2 protein expression would suggest, we delete the *Msi1* and *Msi2* genes individually in mature photoreceptors. We confirmed the photoreceptor-specific loss of MSI1 and MSI2 protein by immunostaining retinal cross-sections obtained at day 14 post-injection (Figure 3A, and B). Western blot analysis of retinal lysates showed that tamoxifen injection required 7 to14 days to ablate the proteins, in agreement with our observation of the double knockout (Figure 3C). The MSI2 protein level was upregulated 1.4-fold in the *Msi1*$^{-/-}$ retina compared to the control; however, the increase did not reach statistical significance at the number of replicates used (n=3).

Neither *Msi1* nor *Msi2* single ablation had an effect on retina function (Figure 4A). The scotopic and photopic ERG responses collected from day 7 to day 230 post injection show that the response to light of the single *Msi1* or *Msi2* knockout mice are indistinguishable from the control animals (Figure 4A, and Supplement Figure 4A, B). Similarly, the histology of the knockout retinas collected at day 230 post injection does not show signs of degeneration (Figure 4B, C). These data demonstrate that Musashi proteins are fully redundant in mature photoreceptors.

Binding of Musashi to the downstream proximal intron promotes splicing of alternative exons

To delineate the transcripts bound by Musashi and the positions on the targets where Musashi binds, we used enhanced UV Cross-Linking and Immuno-Precipitation followed by high throughput sequencing of the associated RNA fragments (eCLIP-Seq). The MSI1 and MSI2 proteins are fully redundant in the adult retina, their RNA binding domains are 77% (RRM1) to 92% (RRM2) identical, and the two proteins recognize the same UAG binding site [34,35]. Thus, we argued that performing the eCLIP-Seq experiment on MSI1 will be sufficient to identify the targets

for both proteins.

Out of 30,283 transcripts detected by RNA-Seq, 10,161 had at least one eCLIP peak enriched over input and 7,849 transcripts had eCLIP peaks in the 3'-UTR (Online Supplementary Table 8). The eCLIP-Seq data shows that MSI1 binds predominantly to the 3'-UTRs of mRNA (59.7% of binding sites, Figure 5A, C) and to introns of pre-mRNA (32.7% of binding sites, Figure 5A, E). Motif enrichment analysis of the sequence surrounding the eCLIP crosslink sites by HOMER and MEME software suites identified as enriched motifs centered on a UAG core (Supplement Figure 5). This result is in agreement with the UAG binding site sequence for the Musashi proteins derived from *in vitro* binding and structural studies [34–36]. The crosslink frequency peaks at position -1 relative to the top motif identified by DREME, BUAG, indicated direct binding of MSI1 (Figure 5B).

Previously we identified a photoreceptor-specific alternative splicing program by comparing the splicing in wild type retina to that in retina that is devoid of photoreceptor cells due to knockout of the Aipl1 gene [30]. Motif enrichment analysis suggested a role for the Musashi proteins in controlling this program and we demonstrated that the splicing of at least one exon, exon 2A in the *Ttc8* gene, is activated by MSI1 bound to the downstream intron [30]. Here we sought to determine on a global scale how the Musashi proteins are regulating alternative splicing in photoreceptor cells *in vivo*. Analysis by RNA-Seq of alternative splicing in *Msi1*$^{-/-}$/*Msi2*$^{-/-}$ retina 21 days after tamoxifen injection identified 165 exons that had reduced inclusion levels and 115 exons that were upregulated in the knockout (Online Supplementary Tables 1, 2, and 3). Out of the 165 exons downregulated in the *Msi1*/*Msi2* knockout, 52 were also significantly downregulated in the photoreceptor-devoid retina of the *Aipl1* knockout mice, with another 40 exons showing the same direction of change but not reaching statistical significance in the *Aipl1* knockout retina (Online Supplementary Table 2) [30]. None of the significantly downregulated exons in the *Msi1*/*Msi2* knockout retina were significantly upregulated in the *Aipl1* knockout (Online Supplementary Table 3). We did not observe a correlation between the exons significantly

upregulated in the *Msi1*$^{-/-}$/*Msi2*$^{-/-}$ retina and the exons differentially spliced in the *Aipl1* knockout retina.

Our previous work suggested that the Musashi proteins promote inclusion of alternative exons by binding downstream of the exon in the adjacent intron. To determine if this mode of regulation is common *in vivo* we combined the eCLIP-Seq and RNA-Seq data to build an RNA splicing map of a meta cassette exon (Figure 5D). The splicing map shows significant enrichment of Musashi protein binding to the downstream introns proximal to the exons upregulated by the Musashi proteins (exons downregulated in the knockout). No significant enrichment of Musashi binding sites was observed for exons repressed in the wild type animals compared to the Musashi knockouts.

Photoreceptor-specific microexons in the Ttc8, Cep290, Cc2d2a, Cacna2d4 and Slc17a7 genes are dispensable

Considering the requirement of Musashi proteins for vision and their role in promoting splicing of photoreceptor specific exons, we next tested if photoreceptor-specific alternative splicing variants are required for vision. Using CRISPR/Cas 9 we deleted photoreceptor-specific exons in the *Ttc8*, *Cep290*, *Cc2d2a*, *Cacna2d4* and *Slc17a7* genes [30]. The exons in *Ttc8*, *Cep290*, *Cc2d2a* and *Cacna2d4* genes are microexons, 30nt or less in length, showing sequence conservation across vertebrates that is traceable to fish (Supplement Figure 6A). We further confirmed by RT-PCR that the four alternative exons are used in zebrafish and are included at high rate in the zebrafish eye (Supplement Figure 6B). The photoreceptor-specific exon in *Slc17a7* is confined to rodents and serves as a control representing an evolutionary novel exon that is less likely to impact the function of the host protein. The exons in the *Ttc8*, *Cep290*, *Cc2d2a* and *Slc17a7* genes were downregulated in our *Msi1*/*Msi2* double knockout mice. Deletion of each exon was confirmed by sequencing the alleles after the founders have been outcrossed (Supplement Figure 7). RT-PCR from retinal samples showed the expected expression of exon skipped isoform (Figure 6A).

We examined the visual function of the exon knockout mice by ERG between one and twelve months of age. We did not observe significant differences in the response to light of the exon knockout mice compared to wild type controls (Figure 6B). Similarly, H&E stained retinal sections from the exon knockout mice had normal morphology (Figure 6C). The phenotypes of the individual exon knockout mice may have been too subtle to detect on their own. Thus, we crossed the *Ttc8*, *Cep290*, and *Cc2d2a* exon knockouts to create a homozygous triple exon knockout mouse line. As all three proteins are part of the primary cilium and are critical for cilium biogenesis, we expected the individual exon knockout phenotypes to be amplified in the combined knockout. As with the single exon knockout mice we did not observe changes in the function or morphology of the retina of the triple exon knockout animals (Figure 6B and C).

MSI1 and MSI2 are post-transcriptional activators of protein expression in photoreceptor cells.

Consistent with the previously described role of the Musashi proteins in regulating mRNA translation, our CLIP-seq data showed pervasive binding of MSI1 to 3'-UTRs (Figure 5A and C). To determine the effect of Musashi on the protein expression, we analyzed the changes in mRNA and protein levels in *Msi1*$^{-/-}$/*Msi2*$^{-/-}$ retina where both MSI1 and MSI2 were depleted from mature photoreceptor cells. For RNA-Seq and quantitative proteomics, we used retinas that were collected 21 days after tamoxifen injection. At this time the Musashi proteins were depleted from photoreceptor cells, while the knockout retina had normal morphology and response to light (Figure 2). Thus, the effect of mRNA and protein expression could be analyzed without the confounding effects of photoreceptor cell death.

We used isobaric labeling and mass spectroscopy to compare the expression of 8021 proteins in knockout and control retina. Of these proteins 165 showed significant differences in expression (at least 1.5-fold change in protein levels with adjusted p-value at or below 0.01) between the control and knockout retina. Of the proteins with significant changes 98 had reduced expression and 67 had elevated expression in the knockout retina (Online Supplementary Table

4). As expected, MSI1 and MSI2 protein levels were decreased by more than 2-fold in the retina of the *Msi1/Msi2* knockout mice (Online Supplementary Table 4 and Supplement Figure 8), consistent with the change in expression observed by western blot (Figures 1, and Supplement Figure 9). Gene Ontology and KEGG pathway enrichment analysis showed that proteins downregulated in the knockout retina were strongly associated with categories related to phototransduction, photoreceptor cell structure, and photoreceptor homeostasis (Figure 7, Supplement Figure 9, and Online Supplementary Table 5). The reduced expression of proteins in these categories is not a consequence of photoreceptor cell death or degeneration for two reasons. First, morphologically and physiologically, the retina of the knockout animals are normal at this stage. More importantly, the levels of multiple proteins that are specific to photoreceptors or are abundantly expressed in photoreceptor cells were unchanged or even increased (Figure 7). Examples include proteins with functions in phototransduction (RCVRN), outer segment structure (PRPH2, PROM1), primary cilium structure (CC2D2A, CEP290), intraflagellar transport (IFT80, IFT140), ion transport (ATP1B2), and protein transport (RD3).

In contrast to the downregulated proteins, most of the proteins with increased expression in the knockout retina were associated with Gene Ontology terms and KEGG pathways involved in cell proliferation, extracellular matrix structure, immune response, and angiogenesis (Online Supplementary Table 5). Closer examination of the upregulated proteins revealed proteins (GFAP, CLU, STAT3, JUNB, IRF9, A2M, B2M, complement components) that are expressed at elevated levels across various models of retinal degeneration [37,37–40]. Single cell RNA-Seq of the Cwc27^{fs} model of retinal degeneration indicated that many of the upregulated genes are expressed by glia [37]. To determine how MSI1 and MSI2 regulate protein expression in photoreceptor cells, we defined two sets of genes. The first set were genes that are either specifically expressed in photoreceptor cells or have at least two-fold higher expression in photoreceptors compared to any other cell type in the retina. The second set contained genes that are either not expressed in photoreceptor cells or have at least two-fold lower expression

compared to other retinal cell types. The two sets of genes were derived from differential expression analysis of single cell RNA-Sequencing data by Macosco et al [41]. We will use "photoreceptor-specific" as a shorthand for the subset of genes highly expressed specifically in photoreceptor cells, with the understanding that many of these genes are also expressed in other cell types, albeit at lower levels. Most proteins with significantly lower expression in the *Msi1*$^{-/-}$ */Msi2*$^{-/-}$ retina and without a significant change in their mRNA levels, belonged to the photoreceptor-specific set of genes (Figure 8A and Online Supplementary Table 7). Only two photoreceptor-specific proteins, PROM1 and IMPG2, had markedly higher expression in the knockouts (Figure 8A). All 39 proteins with altered expression derived from "photoreceptor-specific" genes (blue rombs on Figure 8A) contained MSI1 eCLIP peaks in their 3'-UTRs. A cumulative plot of the changes in protein and RNA expression from the photoreceptor-specific genes shows a global trend in reduced protein levels that were not matched by a corresponding decrease in transcript levels (Figure 8B). Taken together our data demonstrates that the Musashi proteins promote protein expression post-transcriptionally.

MSI1 promotes translation of recombinant Gnat1

To demonstrate a direct activation of protein expression by Musashi we examined the effect of MSI1 on protein expression from *Gnat1* clones carrying full length 3'-UTR in a heterologous system, NIH 3T3 cells. The NIH 3T3 cells were chosen for the low levels of endogenous MSI1 and MSI2 protein expression. We created *Gnat1* clones that contained either wild type 3'-UTR or a mutant 3'-UTR in which the TAG sites were changed to TGA to prevent Musashi binding. The wild type and mutant clones carried HA and T7 epitope tags, respectively. The clones were mixed together in equal amounts and co-transfected in NIH 3T3 cells with vectors expressing either GFP, *Pcbp2*, or *Msi1*. PCBP2 is a RNA binding protein that like MSI1 and MSI2 is abundantly expressed in photoreceptors, shuttles between the nucleus and the cytoplasm, and regulates splicing and translation [42,43]. The products of the wild type and mutant clones were distinguished by the HA (wild type) and T7 (mutant) epitope tags (Figure 8C and D). The epitope

tags were detected by antibodies on Western blot or by hydrolysable probes in multiplexed RT-qPCR, to measure protein and mRNA expression respectively. The effect of the co-transfected constructus on *Gnat1* expression was measured as the change in the ratio of the HA to the T7 signal.

One-way ANOVA found a significant effect of the co-transfected expression vector on the HA/T7 GNAT1 protein ratio ($F(2,15)=50.85$, p-value=$2.08*10^{-7}$). Tukey HSD post-hoc test showed that MSI1 has a significant (p-value<$1*10^{-7}$) effect on the GNAT1 protein expression compared to the vectors expressing GFP or PCBP2 (Figure 8C and D). The analysis of the *Gnat1* mRNA levels also revealed a significant effect of the co-transfected expression construct (one-way ANOVA $F(2,15)=8.19$, p-value=0.004). The observed effect was due to a marginal increase in *Gnat1* transcript levels in response to MSI1 compared to GFP (Figure 8D). The two-fold increase in GNAT1 protein expression without a corresponding increase in transcript levels recapitulates the regulation of GNAT1 by Musashi that we observed in photoreceptor cells and points to a role for the Musashi proteins as activators of translation.

2.5 Discussion

MSI1 and MSI2 are essential and redundant in mature photoreceptor cells

Here we demonstrate that MSI1 and MSI2 are critical for photoreceptor function and survival. Disruption of the two genes resulted in rapid loss of vision and retinal degeneration (Figure 2 and Supplement Figure 2). The Musashi proteins were fully redundant in mature photoreceptors and the single *Msi1* or *Msi2* knockouts did not have detectable phenotypes. It is unclear why *Msi1* and *Msi2* were only partially redundant during retinal development while they were fully redundant in mature photoreceptor cells.

While hundreds of mutations in dozens of genetic loci lead to loss of vision, to date no vision defects have been associated with mutations in the *Msi1* and *Msi2* genes despite their

critical function in photoreceptor cells. The lack of disease-causing mutations associated with the Musashi genes is likely a combination of their redundancy and their critical role in stem cell maintenance. The redundancy of the two proteins ensures that mutations in one of the genes will be complemented by the other. At the same time, loss of both *Msi1* and *Msi2* will result in embryonic lethality and preempt the observation of retinal phenotypes.

MSI1 and MSI2 activate the inclusion of alternative exons *in vivo*

Photoreceptor cells express a distinct splicing program that utilizes a large number of microexons. Motif enrichment analysis suggested that the inclusion of photoreceptor-specific exons is driven by the Musashi proteins [30]. Recent studies on mouse tissues, flow sorted retinal cells, and human retinal organoids showed a similar photoreceptor-specific splicing program directed by Musashi [42,44]. Here we demonstrate that in photoreceptors the Musashi proteins directly promote the splicing of alternative exons by binding to the downstream intron. The Musashi proteins promoted the inclusion of more than half of the exons we previously defined as photoreceptor specific [30]. This leaves a sizable population of photoreceptor-specific exons that rely on other factors for their splicing. Recent studies have highlighted two such factors, PCBP2 and SRRM3, that can either act independently or cooperate with Musashi to promote splicing of alternative exons in photoreceptors [42,45].

In addition to the exons activated by the Musashi proteins, a comparable number of exons appeared to be repressed by them. The repressed exons lack enrichment of MSI1 binding to them or to the adjacent introns when compared to alternative exons that are not regulated by Musashi. It is possible that the Musashi proteins have more than one mode of directly repressing splicing and our dataset does not have sufficient power to detect these interactions as enriched. It is also likely that many of the repressed exons are not direct targets of the Musashi proteins but are regulated by factors whose expression is controlled by MSI1 and MSI2. For example, the Musashi proteins negatively regulate the expression of SRSF9 (Online Supplementary Table 4). Consequently, the splicing of exons dependent on SRSF9 will be repressed indirectly by the

expression of the Musashi proteins in photoreceptors.

Dispensable conserved microexons

Use of alternative microexons is a hallmark of the nervous system. The majority of these alternative exons are conserved which has led to the conclusion that they are essential in neurons. For several microexons such essential function has been demonstrated using animal knockout models [12,46–48]. We reasoned that exons that are conserved in vertebrates and are specifically used in photoreceptor cells will likely have an important role in vision. We deleted four such conserved exons in the *Ttc8*, *Cep290*, *Cc2d2a* and *Cacna2d4* genes, and one exon in the *Slc17a7* gene that is present only in rodents and is potentially nonessential (Supplement Figure 7). All five exons are photoreceptor-specific and are included in nearly all of the transcripts from the corresponding genes in photoreceptor cells [30,42]. The genes hosting four of the exons, *Ttc8*, *Cep290*, *Cc2d2a* and *Cacna2d4*, are essential for vision [49–52]. To our surprise the exon knockout animals did not show a detectable phenotype. A triple knockout of the exons in *Ttc8*, *Cep290*, and *Cc2d2a*, all essential components of the cilium, also lacked an adverse phenotype. The absence of phenotype in our exon knockout animals raises questions about the nature of the selective pressures that have led to the conservation of these exons. It is possible that the selective pressures are exerted by factors that are absent from the environment under which laboratory mice are reared. An alternative explanation is that these exons do not have function and do not alter the properties of the proteins. The conservation of such functionally neutral exons will be due to purifying selection that eliminates mutations negatively affecting the function of the host protein and tolerates wholesale deletion of the exon. In support of this model, the photoreceptor-specific exons in *Ttc8* and *Cc2d2a* are absent from several species (Supplement Figure 6). Regardless of the conservation mechanism, our results caution against using sequence conservation as a sole indicator of function.

MSI1 and MSI2 promote protein expression in photoreceptor cells.

The effect of the Musashi proteins on protein translation is context dependent. In flies and

cultured mammalian cell lines the Musashi proteins repress translation of *Numb* and p21$^{Waf1/Cip1}$, while in frog oocytes they activate early translation of Mos and Cyclin B5 after progesterone stimulation [21,26,53,54]. Furthermore, recent genome wide studies integrating ribosome profiling and RNA binding data showed that within the same cell MSI1 and MSI2 repress translation of certain transcripts while activating others [55,56]. The transcriptome-wide studies also show that translation of relatively few of the large numbers of transcripts that are bound by the Musashi proteins is affected when the levels of MSI1 and MSI2 are manipulated [56].

Here we present an integrative analysis of the effect of MSI1 and MSI2 on protein translation in photoreceptor cells. To isolate the signal derived from photoreceptor cells, we focused our analysis on transcripts highly expressed in photoreceptors compared to other retinal cells and relied on the fact that photoreceptor cells are the dominant cell type in the retina comprising approximately half of the cells in that tissue. As we are excluding from our analysis transcripts that are expressed in other cells of the retina at levels comparable to those in photoreceptors, we derived a broad but far from comprehensive picture of the effect of MSI1 and MSI2 on the transcriptome and proteome of photoreceptor cells. The amount of material required for RNASeq, mass-spectrometry, and eCLIP-Seq experiments did not allow us to perform the experiments in parallel on samples from the same animal. Samples from different animals were used for each experiment. While there is an excellent correlation across replicates, it should be noted that changes in protein and RNA expression were not determined within the same animal.

We demonstrate that the combined deletion of *Msi1* and *Msi2* alters protein expression from a set of transcripts without significantly affecting the levels of these transcripts. In agreement with the previously published transcriptome-wide studies, our integrative analysis shows that the levels of relatively few proteins are affected in the Musashi knockout mice, compared to the thousands of transcripts bound by MSI1 at their 3'-UTRs.

In photoreceptors, the Musashi proteins act largely to promote protein expression. Furthermore, the translation of at least one of the targets identified in this work, *Gnat1*, is directly

stimulated by MSI1 in a heterologous system. We observed only two cases, PROM1 and IMPG2, where the combined knockout of *Msi1* and *Msi2* resulted in elevated protein levels. This is an unexpected finding in light of the canonical view of the Musashi proteins as repressors of translation. As the Musashi proteins are regulating both pre-mRNA splicing in the nucleus and protein expression in the cytoplasm there are some genes affected by both modes of regulation, e.g. Prom1, raising the question of potential coordination between pre-mRNA splicing and protein expression. We do not see evidence for such coordination in our data.

The proteins regulated by MSI1 and MSI2 are central to the function and survival of photoreceptor cells. They include the products of a number of genes that are commonly mutated in blinding disease (GNAT1, CNGA1, PRCD, ROM1, AIPL1, PDE6A, etc). Photoreceptor cells need to produce high levels of these proteins in order to replace their outer segments every 10 days. This renewal process does not reuse proteins already present in the outer segment. Instead, new membranes and proteins are delivered to the bottom of the outer segment stack, while old segments are phagocytosed and digested by retinal pigmented epithelium from the top of the stack. Reduced rate of production of outer segment proteins and the chaperones that fold them will impede the outer segment renewal process leading to loss of vision and degeneration of the photoreceptor cells.

2.6 Materials and Methods

Animals

All animal experiments were conducted with the approval of the Institutional Animal Care and Use Committee at West Virginia University. Both males and females were used in all experiments. The mouse lines in this study were in the C57BL6/J background and devoid of the naturally occurring *rd1* and *rd8* alleles. The mice were genotyped at weaning unless otherwise specified in the results section. The primers used for genotyping of the targeted alleles are listed

in Online Supplementary Table 14.

Mice carrying *Msi1* $^{flox/flox}$ and *Msi2* $^{flox/flox}$ were provided by Dr. Christopher Lengner from the University of Pennsylvania (Li et al., 2015; Park et al., 2014). Mice carrying the floxed alleles were crossed with *Pde6g-Cre*^{ERT2} mice to enable photoreceptor-specific conditional knockout of *Msi1* and *Msi2* (Koch et al., 2015). The conditional deletion of *Msi1* and *Msi2* in mature photoreceptor cells was induced by intraperitoneal injection of tamoxifen in corn oil (Sigma-Aldrich catalog #T5648-1G) at concentrations of 100 mg/kg body weight for three consecutive days.

The knockouts of the photoreceptor specific exons in *Ttc8*, *Cep 290*, and *Cc2d2a* were created using CRISPR/Cas9. Two guide RNAs targeted at sites upstream and downstream of each alternative exon were used to cause full deletion of the exon and the proximal parts of the introns. The guide RNAs were synthesized by Synthego and IDT. The guide RNA targeting sequences are listed in Online Supplementary Table 13. The guide RNAs and Cas9 (Thermo Fisher) were assembled into ribonucleoprotein complexes and electroporated into zygotes by the WVU transgenic core facility. The founders were back-crossed to C57BL6/J mice (Jackson Laboratory) for 5 generations. To map the borders of the deletions, the exon knockout alleles were amplified by PCR using the genotyping primers and sequenced by Sanger sequencing (Supplement Figure 7). Frozen sperm for each line is available from the authors upon request.

Adult zebrafish (*Danio rerio*) animals were maintained at 28°C with standard 14/10 light/dark cycles. For dissections, we used adult Tubingen long-fin strain (approximately 22 months old). Equal female and male zebrafish were euthanized in an ice bath of system water until the termination of buccal and gill motion. Tissue dissection was performed in physiological saline E3h media (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4, and 1 mM HEPES, pH 7.3). Collected samples were immediately collected in centrifuge tubes and frozen.

Clones, cell lines, and transfection

A full length *Gnat1* clone (accession BC058810) was obtained from Horizon Discovery and recloned in pcDNA3.1(+) (Invitrogen). A matching clone in which all 16 TAG triplets in the

1.11kbp 3'-UTR were mutated to TGA to disrupt the Musashi binding sites was created using gene synthesis (Genscript). Gibson assembly was used to reclone the cDNAs into pcDNA3.1(+) vector and attach HA- and T7-tags to the wild type and mutant clone, respectively. Full length *Msi1* clone with N-terminal Flag epitope tag in pcDNA3.1 was described before [30]. Full length, codon optimized mouse Pcbp2 clone with N-terminal Flag epitope tag was produced by gene synthesis (Genscript) and cloned in pcDNA3.1(+) (Invitrogen). All clones are available from Addgene (https://www.addgene.org/Peter_Stoilov/).

NIH 3T3 cells were grown in DMEM supplemented with 10% Fetal Bovine Serum. The cells were maintained in a humidified incubator at 37°C in 5% CO_2 atmosphere. Transfection with polyethyleneimine was carried out as described before [57]. Briefly, the 6 hours prior to transfection the cells were seeded at $3.2*10^5$ cells per well in 6 well plates. A total of 500 ng of DNA was used per transfection, containing 125 ng of each wild type and mutant *Gnat1* construct, and 250 ng of expression vector that carried a flag-tagged EGFP, *Pcbp2* or *Msi1* clone in pcDNA3.1(+) backbone. The cells were collected 27 to 29 hours post-transfection to analyze protein and mRNA expression.

RNA extraction and RT-PCR

RNA was extracted with Trisol and precipitated with isopropanol. The RNA was then dissolved and treated with RNAse-free DNase I (Roche) for 20 minutes at 37°C. After DNA digestion the reactions were extracted once with chloroform and the RNA was precipitated with ethanol. RT-PCR analysis of alternative splicing using fluorescently labeled primers was described before [30,58].

The levels of Gnat1 transcripts expressed from the recombinant clones in 3T3 cells were determined by multiplexed RT-qPCR. Hydrolysis probes to the HA and T7 tags were used to detect Gnat1 transcripts with wild type and mutant 3'-UTRs, respectively. The RT-qPCR was performed using Luna One Step RT-qPCR mix with dUDG (NEB). Amplification using Luna One Step qPCR mix with UDG (NEB) that did not contain a reverse transcriptase component (NEB)

served as no-reverse -transcriptase controls for DNA contamination. The ratio of the transcript levels measured by the HA and T7 probes was used to determine the effect of each treatment on the mRNA levels expressed from the constructs carrying wild type and mutant 3'-UTRs. The primers and probes used for alternative splicing and RT-qPCR analysis are listed in Online Supplementary Table 14.

Electroretinography (ERG) Measurement and Preparation of Animal

ERGs were measured using either UTAS Visual Diagnostic System with Big-Shot Ganzfeld device (LKC Technologies, Gaithersburg, MD, USA) or Celeris system with Espion software (Diagnosis LLC, Lowell, MA, USA). Prior to testing, mice were dark-adapted overnight. All further handling of mice following dark adaptation was performed under deep red illumination. The mice were anesthetized by inhalation of 1.5% isoflurane mixed with 100% oxygen at a flow rate of 2.5 l/min. The pupils were topically dilated with a drop of tropicamide and phenylephrine-hydrochloride, allowing drops to sit on both eyes for 10 mins. After that, mice were transferred to a heated platform connected to a nose cone that allows for a continuous flow of isoflurane. A reference electrode was inserted sub-dermally between the eyes of the mouse, and ERG responses were collected from both eyes using wire electrodes placed on the center of each cornea, with contact being made using a drop of 0.3% Hypromellose solution. To deliver the stimulus, a Ganzfeld Bowel was used with LED white arrays at increasing intensities. Dark-adapted scotopic photoresponse was recorded under the dim red light using a single LED white light flashes of luminescence ranging from $2.45 \cdot 10^{-4}$ to 2.4 cd-s/m^2. For photopic response, animals were light-adapted for 10 min in the presence of rod-saturating 30 cd-s/m^2 ambient white light prior to recording the photopic response.

Western Blot

Mouse retinas were lysed using RIPA buffer (50 mM Tris HCl-pH 8.0, 150 mM NaCl, 1.0% TritonX-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate) supplied with protease (Sigma-Aldrich catalog# 535140-1ML) and phosphatase inhibitors cocktail (Sigma-Aldrich catalog

P5726-1 ML). After homogenization, the lysate was incubated on ice for 10 mins, then cleared by centrifugation for 15 mins. 20 µg of protein extract was resolved in 4-20% polyacrylamide SDS–PAGE gel and transferred onto polyvinylidene difluoride (PVDF) membranes (Immunobilon-FL, Millipore). After blocking with BSA in PBST (Phosphate- buffered saline with 0.1% Tween-20), the membranes were blocked and probed with primary antibodies overnight at 4 °C, followed by incubation with fluorescently labeled (Alexa Fluor 647 or 488, Jackson ImmnuoResearch) secondary antibodies for 1 hour at room temperature. The membranes were then scanned on Amersham Typhoon Phosphorimager (GE Healthcare). GFP protein in transfection samples was detected by its intrinsic fluorescence in the 488nm channel. To quantify the protein expression across membranes, the band intensities detected by each antibody were scaled to the median signal for the membrane. The scaled expression values were then normalized to the scaled values of the corresponding controls (loading controls or in the transfection experiments to proteins expressed from co-transfected constructs). The antibodies used for western blot analysis are listed in Online Supplementary Table 12.

Retinal histology

The whole eyecups from the knockout and control mice were enucleated. The eyes were then fixed using a Z-fixative (Excalibur Pathology Inc). Tissue processing, including paraffin embedding and hematoxylin and eosin (H&E) staining, was performed at Excalibur Pathology. Images of the stained slides were collected using a Nikon Brightfield Microscope operated by Element software (Nikon). To evaluate the photoreceptor cell loss, we counted the number of nuclei within the outer nuclear layer (ONL) using the NIS elements software. The counting was done at ten equidistant locations centered on the optical nerve and moving toward the periphery in 400 µm increments. Five locations were on the inferior side (-5 to -1) and five on the superior side (1 to 5) of the retina relative to the optic nerve. For each location and the number of nuclei reported is the average of 4 technical replicates. The nuclei counts were averaged over 3 biological replicates that represent retinas from three separate animals.

Immunocytochemistry

Eyes were enucleated, and a small window was cut in the cornea before immersing it in 4% paraformaldehyde fixative (4% PFA in PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4, pH 7.2) for 3 hours on a rotator. Eyecups were dehydrated by sequential incubation in 7.5%, 15%, and 22% sucrose in 1xPBS. Eyecups were then snap-frozen in optimal cutting temperature compound (OCT) blocks. Serial 16 µm sections were cut on a Leica CM1850 cryostat and mounted onto Superfrost Plus microscope slides (Fischer Scientific). Mounted retinal sections were washed 3 times for 10 mins each with PBS and then blocked with PBST for 1 hour (10% goat serum, 0.3% Triton X-100, and 0.02% sodium azide in PBS). Retinal sections were incubated overnight at 4°C with primary antibodies diluted in PBST supplemented with 5% goat serum. After three 15 min washes with PBST the sections were incubated for one hour with secondary antibodies diluted 1:1000 in PBST supplemented with 5% goat serum and 4′,6-diamidino-2-phenylindole. The sections were washed three times for 15 min with PBST, mounted with Prolong Antifade reagent (Thermofisher), and secured with coverslips. The sections were imaged on a Nikon C2 laser scanning confocal microscope. The laser power, gain, and offset settings were maintained the same when imaging sections from knockout and control littermates. The antibodies used for immunofluorescence staining are listed in Online Supplementary Table 12.

RNA sequencing

Total RNA was isolated at day 21 post tamoxifen injection using Tri-reagent (Sigma) from retinas in four biological replicates of *Msi1*$^{+/+}$/*Msi2*$^{+/+}$ and *Msi1*$^{-/-}$/*Msi2*$^{-/-}$ mice. Sequencing libraries were prepared by the West Virginia University genomics core using KAPA Hyper RNA with Riboerase (Roche). The libraries were sequenced by the University of Illinois DNA services core on Illumina HiSeq 4000 at an average depth of 44 million 100nt paired end reads.

RNA-Seq reads were aligned to the mouse genome (GRCm38) using HISAT2 [60]. The mapped reads were summarized using Rsubread, and differential gene expression analysis was carried out by edgeR (Online Supplementary Table 8) [61,62]. Inclusion levels of cassette exons were calculated by rMATS (4.1.0), using reads spanning exon-exon junctions [63].

CLIP-sequencing and meta-exon analysis

The rabbit anti-MSI1 (1:1000; catalog# ab 52865, Abcam, Cambridge, MA) was used for eCLIP. Briefly, retinas from wild-type mice (80 mg per replicate) were collected and placed in ice-cold PBS in a 10 cm^2 plate. Plates containing retinas were then placed on ice and UV-crosslinked (254 nm, 200 mJ/cm^2) using UV Stratalinker™ 2400. UV-crosslinked retinas were then snap frozen in liquid nitrogen. Further tissue processing, eCLIP library prep, and sequencing were carried out by Eclipse Bioinnovations following a previously published protocol [64]. The raw data obtained from the eCLIP-Seq are available at the NCBI Sequence Read Archive under project accession PRJNA795195.

To analyze the raw eCLIP data, the adapter sequences were first trimmed using cutadapt [65]. HISAT2 was used to map the reads to version GRCm38 of the mouse genome and the mapped reads were deduplicated by umi-tools using the unique molecular identifier (UMI) barcodes built into the adapters [60,66]. Crosslink sites were identified and clustered (Supplemental data set 1) into regions using PureClip [67]. Motifs enriched in the 51 nucleotide sequence fragments centered on the crosslink sites were identified by HOMER and DREME [68,69].

Meta-exon analysis was performed using the RBP-maps software package on non-redundant sets of alternatively spliced exons identified in our RNA-Seq analysis of *Msi1*/*Msi2* double knockout in photoreceptor cells [70]. The distribution of MSI1 crosslinks around exons downregulated or upregulated in the photoreceptor-specific *Msi1*/*Msi2* double knockout was compared to alternative exons that were not affected by the knockout (Online Supplementary Tables 9, 10, and 11). 1000 random permutations of the non-regulated exon set were used to determine the 99.5% confidence intervals as described by Yee at al [70].

Proteomics analysis

Retina samples from *Msi1*$^{+/+}$/*Msi2*$^{+/+}$ and *Msi1*$^{-/-}$/*Msi2*$^{-/-}$ were collected 21 days after tamoxifen injections. Five biological replicates were used for each wild-type and knockout group. Tissue processing and proteomics quantification of snapped frozen retina samples was performed by IDeA proteomics. Briefly, proteins were reduced, alkylated, and purified by chloroform/methanol extraction prior to digestion with sequencing-grade modified porcine trypsin (Promega). Tryptic peptides were labeled using tandem mass tag isobaric labeling reagents (Thermo) following the manufacturer's instructions and combined into one 10-plex sample group. The labeled peptide multiplex was separated into 46 fractions on a 100 x 1.0 mm Acquity BEH C18 column (Waters) using an UltiMate 3000 UHPLC system (Thermo) with a 50 min gradient from 99:1 to 60:40 buffer A:B ratio under basic pH conditions, and then consolidated into 18 super-fractions. Each super-fraction was then further separated by reverse phase XSelect CSH C18 2.5 um resin (Waters) on an in-line 150 x 0.075 mm column using an UltiMate 3000 RSLCnano system (Thermo). Peptides were eluted using a 60 min gradient from 98:2 to 60:40 buffer A:B ratio. Eluted peptides were ionized by electrospray (2.2 kV) followed by mass spectrometric analysis on an Orbitrap Eclipse Tribrid mass spectrometer (Thermo) using multi-notch MS3 parameters. MS data were acquired using the FTMS analyzer in top-speed profile mode at a resolution of 120,000 over a range of 375 to 1500 m/z. Following CID activation with normalized collision energy of 35.0, MS/MS data were acquired using the ion trap analyzer in centroid mode and normal mass range. Using synchronous precursor selection, up to 10 MS/MS precursors were selected for HCD activation with normalized collision energy of 65.0, followed by acquisition of MS3 reporter ion data using the FTMS analyzer in profile mode at a resolution of 50,000 over a range of 100-500 m/z. Buffer A is 0.1% formic acid, 0.5% acetonitrile; Buffer B is 0.1% formic

acid, 99.9% acetonitrile. Both buffers were adjusted to pH 10 with ammonium hydroxide for offline separation.

To create a database of proteins expressed in the retina, we first filtered our mouse retina RNA-Seq data to remove genes with median expression across all samples that were below the median expression for the dataset. As a result, we selected 15,626 genes with expression equal or more than 1.2 RPKM. Ensembl release 79 was queried for annotated proteins produced by these genes resulting in a database of 34,055 protein sequences. Peptide identification against the retinal protein database was performed using MS-GF+ (version v2021.03.22) with parent ion tolerance of 10 ppm, reporter ion tolerance of -0.00335 Da, and +0.0067 Da, and requiring fully tryptic peptides [71]. Only peptides with peptide level Q-value of 0.05 or below were accepted. The MSnbase package from R/Bioconductor was used to quantify the MS3 reporter ions and combine the identification and quantification data [72,73]. Differential protein expression analysis was performed using the DeqMS package from R/Bionconductor [74]. Protein changes with adjusted p-value below 0.05 and fold change of more than 1.5 were considered significant.

WebGestalt (WEB-based Gene SeT Analysis Toolkit) was used to perform enrichment analysis on the Gene Ontology and KEGG databases for the proteins with significant changes in gene expression. Only terms enriched at FDR<0.05 were reported [75].

Experimental design and Statistical analysis

Age-matched males and females in the C57BL6/J background were used in all experiments. The statistical analysis and data visualization was done using GraphPad Prism and R/Bioconductor. Unpaired Student's t-test was used to assess statistical significance between control and knockout samples. Statistical significance was determined with one-way or two-way ANOVA followed by pairwise comparisons as indicated in the text. All data were presented as the mean ±standard error of the mean.

Data availability

The eCLIP-Seq and RNA-Seq data are available at the NCBI Sequence Read Archive under project accessions PRJNA795195 and PRJNA795137. The mass spectrometry proteomics are deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD030748 and 10.6019/PXD030748 [76,77].

2.7 Acknowledgments

This work was supported by the National Institutes of Health 2R01EY025536 (P.S. and V.M.) and bridge funding provided by the West Virginia University Health Sciences Center Office of Research and Graduate Education (P.S.). E.J.H is supported by West Virginia University and Department of Biology startup funds, Research and Scholarship Advancement award, and the Program to Stimulate Competitive Research. We are grateful to Dr. Christopher Lengner for the generous donation of the *Msi1*$^{fl/fl}$ and *Msi2*$^{fl/fl}$ mice. We thank Drs. Roberta Leonardi and Aaron Robart for critical reading of this manuscript.

2.8 Author contributions

F. M.: experimental conception, experimental design, data acquisition, data analysis, data interpretation, manuscript writing, and manuscript revision

B.J.: data acquisition and data analysis

M.S.: data acquisition

V.R.: experimental conception, experimental design, data interpretation, and manuscript revision

P.S.: experimental conception, experimental design, data acquisition, data analysis, data interpretation, manuscript writing, and manuscript revision

Competing interests

All authors declare no competing interests.

2.9 References

1. Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nat Rev Genet* **15**, 829–845 (2014).

1. Agrawal, M. & Welshhans, K. Local Translation Across Neural Development: A Focus on Radial Glial Cells, Axons, and Synaptogenesis. *Frontiers in Molecular Neuroscience* **14**, (2021).
2. Darnell, J. C. & Richter, J. D. Cytoplasmic RNA-Binding Proteins and the Control of Complex Brain Function. *Cold Spring Harb Perspect Biol* **4**, a012344 (2012).

3. Furlanis, E. & Scheiffele, P. Regulation of neuronal differentiation, function, and plasticity by alternative splicing. *Annu Rev Cell Dev Biol* **34**, 451–469 (2018).

4. Iijima, T., Iijima, Y., Witte, H. & Scheiffele, P. Neuronal cell type–specific alternative splicing is regulated by the KH domain protein SLM1. *J Cell Biol* **204**, 331–342 (2014).

5. Lee, J.-A., Tang, Z.-Z. & Black, D. L. An inducible change in Fox-1/A2BP1 splicing modulates the alternative splicing of downstream neuronal target exons. *Genes Dev.* **23**, 2284–2293 (2009).

6. Lee, J.-A. *et al.* Cytoplasmic Rbfox1 Regulates the Expression of Synaptic and Autism-Related Genes. *Neuron* **89**, 113–128 (2016).

7. Saito, Y. *et al.* NOVA2-mediated RNA regulation is required for axonal pathfinding during development. *eLife* **5**, e14371.

8. Irimia, M. *et al.* A Highly Conserved Program of Neuronal Microexons Is Misregulated in Autistic Brains. *Cell* **159**, 1511–1523 (2014).

9. Li, Y. I., Sanchez-Pulido, L., Haerty, W. & Ponting, C. P. RBFOX and PTBP1 proteins regulate the alternative splicing of micro-exons in human brain transcripts. *Genome Res.* gr.181990.114 (2014) doi:10.1101/gr.181990.114.

10. Gonatopoulos-Pournatzis, T. & Blencowe, B. J. Microexons: at the nexus of nervous system development, behaviour and autism spectrum disorder. *Curr Opin Genet Dev* **65**, 22–33 (2020).

11. Lin, L., Zhang, M., Stoilov, P., Chen, L. & Zheng, S. Developmental Attenuation of Neuronal Apoptosis by Neural-Specific Splicing of Bak1 Microexon. *Neuron* **107**, 1180-1196.e8 (2020).

12. Möröy, T. & Heyd, F. The impact of alternative splicing in vivo: Mouse models show the way.

RNA **13**, 1155–1171 (2007).

14. Fukuda, T. *et al.* Mice lacking the EDB segment of fibronectin develop normally but exhibit reduced cell growth and fibronectin matrix assembly in vitro. *Cancer Res* **62**, 5603–5610 (2002).
13. Gimond, C. *et al.* Cre-loxP–mediated Inactivation of the α6A Integrin Splice Variant In Vivo: Evidence for a Specific Functional Role of α6A in Lymphocyte Migration but Not in Heart Development. *J Cell Biol* **143**, 253–266 (1998).

14. Homanics, G. E. *et al.* Normal electrophysiological and behavioral responses to ethanol in mice lacking the long splice variant of the γ2 subunit of the γ-aminobutyrate type A receptor. *Neuropharmacology* **38**, 253–265 (1999).

15. Vrhovski, B., Lemckert, F. & Gunning, P. Modification of the tropomyosin isoform composition of actin filaments in the brain by deletion of an alternatively spliced exon. *Neuropharmacology* **47**, 684–693 (2004).

16. Okano, H. Stem cell biology of the central nervous system. *Journal of Neuroscience Research* **69**, 698–707 (2002).

19. Sakakibara, S. *et al.* Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. *Dev. Biol.* **176**, 230–242 (1996).
17. Sakakibara, S. & Okano, H. Expression of neural RNA-binding proteins in the postnatal CNS: implications of their roles in neuronal and glial cell development. *J Neurosci* **17**, 8300–8312 (1997).

18. Imai, T. *et al.* The Neural RNA-Binding Protein Musashi1 Translationally Regulates Mammalian numb Gene Expression by Interacting with Its mRNA. *Molecular and Cellular Biology* **21**, 3888–3900 (2001).

22. Fox, R. G., Park, F. D., Koechlein, C. S., Kritzik, M. & Reya, T. Musashi Signaling in Stem Cells and Cancer. *Annual Review of Cell and Developmental Biology* **31**, 249–267 (2015).
23. Kharas, M. G. *et al.* Musashi-2 regulates normal hematopoiesis and promotes aggressive myeloid leukemia. *Nat Med* **16**, 903–908 (2010).

24. Li, N. *et al.* The Msi Family of RNA-Binding Proteins Function Redundantly as Intestinal Oncoproteins. *Cell Rep* **13**, 2440–2455 (2015).

25. Sakakibara, S. *et al.* RNA-binding protein Musashi family: Roles for CNS stem cells and a subpopulation of ependymal cells revealed by targeted disruption and antisense ablation. *Proc Natl Acad Sci U S A* **99**, 15194–15199 (2002).

26. Charlesworth, A., Wilczynska, A., Thampi, P., Cox, L. L. & MacNicol, A. M. Musashi regulates the temporal order of mRNA translation during Xenopus oocyte maturation. *EMBO J* **25**, 2792–2801 (2006).

27. Cragle, C. & MacNicol, A. M. Musashi Protein-directed Translational Activation of Target mRNAs Is Mediated by the Poly(A) Polymerase, Germ Line Development Defective-2. *J Biol Chem* **289**, 14239–14251 (2014).

28. Cragle, C. E. *et al.* Musashi interaction with poly(A)-binding protein is required for activation

of target mRNA translation. *J Biol Chem* **294**, 10969–10986 (2019).

29. Kawahara, H. *et al.* Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. *J Cell Biol* **181**, 639–653 (2008).

30. Murphy, D., Cieply, B., Carstens, R., Ramamurthy, V. & Stoilov, P. The Musashi 1 Controls the Splicing of Photoreceptor-Specific Exons in the Vertebrate Retina. *PLOS Genetics* **12**, e1006256 (2016).

31. Sundar, J., Matalkah, F., Jeong, B., Stoilov, P. & Ramamurthy, V. The Musashi proteins MSI1 and MSI2 are required for photoreceptor morphogenesis and vision in mice. *J Biol Chem* (2020) doi:10.1074/jbc.RA120.015714.

32. Koch, S. F. *et al.* Halting progressive neurodegeneration in advanced retinitis pigmentosa. *J Clin Invest* **125**, 3704–3713 (2015).

33. Benchorin, G., Calton, M. A., Beaulieu, M. O. & Vollrath, D. Assessment of Murine Retinal Function by Electroretinography. *Bio Protoc* **7**, e2218 (2017).

34. Lan, L. *et al.* Crystal and solution structures of human oncoprotein Musashi-2 N-terminal RNA recognition motif 1. *Proteins* **88**, 573–583 (2020).

35. Zearfoss, N. R. *et al.* A Conserved Three-nucleotide Core Motif Defines Musashi RNA Binding Specificity. *J Biol Chem* **289**, 35530–35541 (2014).

36. Dominguez, D. *et al.* Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Mol Cell* **70**, 854-867.e9 (2018).

37. Bertrand, R. E. *et al.* Cwc27, associated with retinal degeneration, functions as a splicing factor in vivo. *Hum Mol Genet* ddab319 (2021) doi:10.1093/hmg/ddab319.

38. Cheng, C. L. & Molday, R. S. Changes in gene expression associated with retinal degeneration in the rd3 mouse. *Mol Vis* **19**, 955–969 (2013).

39. Hackam, A. S. *et al.* Identification of gene expression changes associated with the progression of retinal degeneration in the rd1 mouse. *Invest Ophthalmol Vis Sci* **45**, 2929–2942 (2004).

40. Uren, P. J., Lee, J. T., Doroudchi, M. M., Smith, A. D. & Horsager, A. A profile of transcriptomic changes in the rd10 mouse model of retinitis pigmentosa. *Mol Vis* **20**, 1612–1628 (2014).

41. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

42. Ling, J. P. *et al.* ASCOT identifies key regulators of neuronal subtype-specific splicing. *Nat Commun* **11**, 1–12 (2020).

43. Makeyev, A. V. & Liebhaber, S. A. The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms. *RNA* **8**, 265–278 (2002).

44. Ottaviani, D. *et al.* The role of Musashi-1 in CEP290 c.2991+1655A>G cryptic exon splicing in Leber Congenital Amaurosis. 2021.08.04.454918 Preprint at https://doi.org/10.1101/2021.08.04.454918 (2021).

45. Ciampi, L. *et al.* Specialization of the photoreceptor transcriptome by Srrm3-dependent microexons is required for outer segment maintenance and vision. *Proc Natl Acad Sci U S A* **119**, e2117090119 (2022).

46. Gonatopoulos-Pournatzis, T. *et al.* Autism-Misregulated eIF4G Microexons Control Synaptic Translation and Higher Order Cognitive Functions. *Molecular Cell* **77**, 1176-1192.e16 (2020).

47. Johnson, V., Junge, H. J. & Chen, Z. Temporal regulation of axonal repulsion by alternative splicing of a conserved microexon in mammalian Robo1 and Robo2. *eLife* **8**, e46042 (2019).

48. Nakano, Y. *et al.* Defects in the alternative splicing-dependent regulation of REST cause deafness. *Cell* **174**, 536-548.e21 (2018).

49. Bachmann-Gagescu, R. *et al.* Genotype–phenotype correlation in CC2D2A-related Joubert syndrome reveals an association with ventriculomegaly and seizures. *J Med Genet* **49**, 126–137 (2012).

50. Dilan, T. L. *et al.* Bardet–Biedl syndrome-8 (BBS8) protein is crucial for the development of outer segments in photoreceptor neurons. *Hum Mol Genet* **27**, 283–294 (2018).
51. Rachel, R. *et al.* CEP290 is required for photoreceptor ciliogenesis and other cilia related functions. *Cilia* **1**, P98 (2012).

52. Wycisk, K. A. *et al.* Structural and functional abnormalities of retinal ribbon synapses due to Cacna2d4 mutation. *Invest Ophthalmol Vis Sci* **47**, 3523–3530 (2006).

53. Arumugam, K., Wang, Y., Hardy, L. L., MacNicol, M. C. & MacNicol, A. M. Enforcing temporal control of maternal mRNA translation during oocyte cell-cycle progression. *EMBO J* **29**, 387–397 (2010).

54. Battelli, C., Nikopoulos, G. N., Mitchell, J. G. & Verdi, J. M. The RNA-binding protein Musashi-1 regulates neural development through the translational repression of p21WAF-1. *Molecular and Cellular Neuroscience* **31**, 85–96 (2006).

55. Katz, Y. *et al.* Musashi proteins are post-transcriptional regulators of the epithelial-luminal cell state. *Elife* **3**, e03915 (2014).

56. Karmakar, S. *et al.* Integrative genome-wide analysis reveals EIF3A as a key downstream regulator of translational repressor protein Musashi 2 (MSI2). *NAR Cancer* **4**, zcac015 (2022).

57. Boussif, O. *et al.* A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc Natl Acad Sci U S A* **92**, 7297–7301 (1995).

58. Percifield, R., Murphy, D. & Stoilov, P. Medium throughput analysis of alternative splicing by fluorescently labeled RT-PCR. *Methods Mol. Biol.* **1126**, 299–313 (2014).

59. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907–915 (2019).

60. Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research* **47**, e47 (2019).

61. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

62. Shen, S. *et al.* rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *PNAS* **111**, E5593–E5601 (2014).

63. Van Nostrand, E. L. *et al.* Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* **13**, 508–514 (2016).
64. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

65. Smith, T. S., Heger, A. & Sudbery, I. UMI-tools: Modelling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* gr.209601.116 (2017) doi:10.1101/gr.209601.116.

66. Krakau, S., Richard, H. & Marsico, A. PureCLIP: capturing target-specific protein–RNA interaction footprints from single-nucleotide CLIP-seq data. *Genome Biology* **18**, 240 (2017).
67. Bailey, T. L. DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653–1659 (2011).

68. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576–589 (2010).

69. Yee, B. A., Pratt, G. A., Graveley, B. R., Van Nostrand, E. L. & Yeo, G. W. RBP-Maps enables robust generation of splicing regulatory maps. *RNA* **25**, 193–204 (2019).

70. Kim, S. & Pevzner, P. A. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat Commun* **5**, 5277 (2014).

71. Gatto, L., Gibb, S. & Rainer, J. MSnbase, Efficient and Elegant R-Based Processing and Visualization of Raw Mass Spectrometry Data. *J Proteome Res* **20**, 1063–1069 (2021).

72. Gatto, L. & Lilley, K. S. MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics* **28**, 288–289 (2012).

73. Zhu, Y. *et al.* DEqMS: A Method for Accurate Variance Estimation in Differential Protein Expression Analysis. *Mol Cell Proteomics* **19**, 1047–1057 (2020).

74. Wang, J., Duncan, D., Shi, Z. & Zhang, B. WEB-based GEne SeT AnaLysis Toolkit

(WebGestalt): update 2013. *Nucl. Acids Res.* **41**, W77–W83 (2013).

75. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**, D442–D450 (2019).

76. Deutsch, E. W. *et al.* The ProteomeXchange consortium in 2020: enabling ‘big data’ approaches in proteomics. *Nucleic Acids Res* **48**, D1145–D1152 (2020).

2.10 Figures

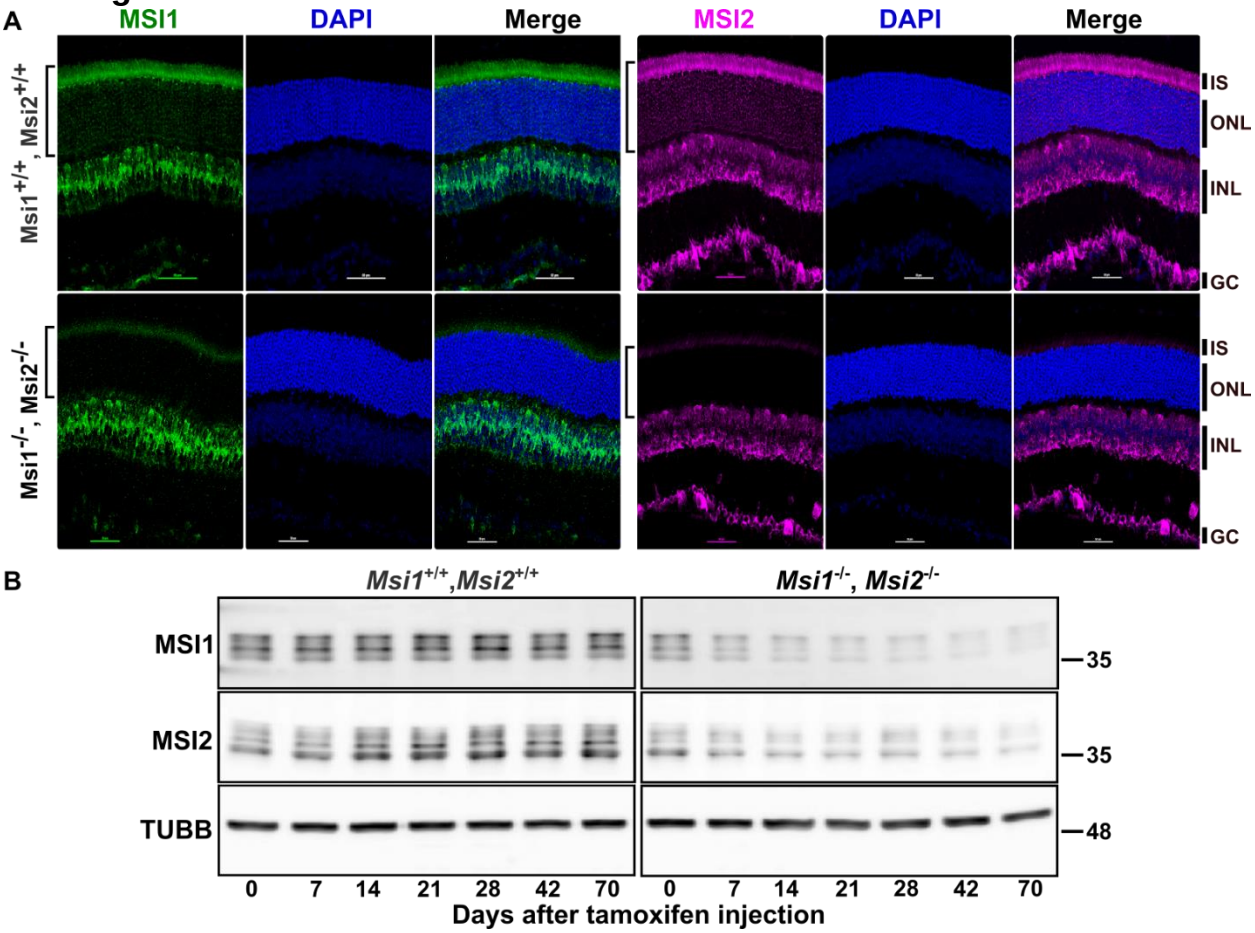


Figure 2-1 Induced double knockout of Msi1 and Msi2 in photoreceptor cells

Figure 1. Induced double knockout of *Msi1* and *Msi2* in photoreceptor cells. A) Immunofluorescence micrographs of retinal cross-sections collected 14 days after tamoxifen injection at postnatal day 30 from *Msi1*$^{+/+}$/*Msi2*$^{+/+}$ littermate control and *Msi1*$^{-/-}$/*Msi2*$^{-/-}$, stained for MSI1(green), MSI2 (magenta), and DAPI (blue). ONL: outer nuclear layer (photoreceptor cell layer). INL: inner nuclear layer. GC: ganglion cell layer. Objective, 40x. **B)** Immunoblot of lysates from *Msi1*$^{+/+}$/*Msi2* $^{+/+}$ and *Msi1*$^{-/-}$ /*Msi2*$^{-/-}$ retinas collected between 0 and 70 days after tamoxifen injections at postnatal day 30 and probed with antibodies to MSI1, MSI2 and TUBB (β-tubulin, loading control). See Supplement Figure 1 for full size blots.

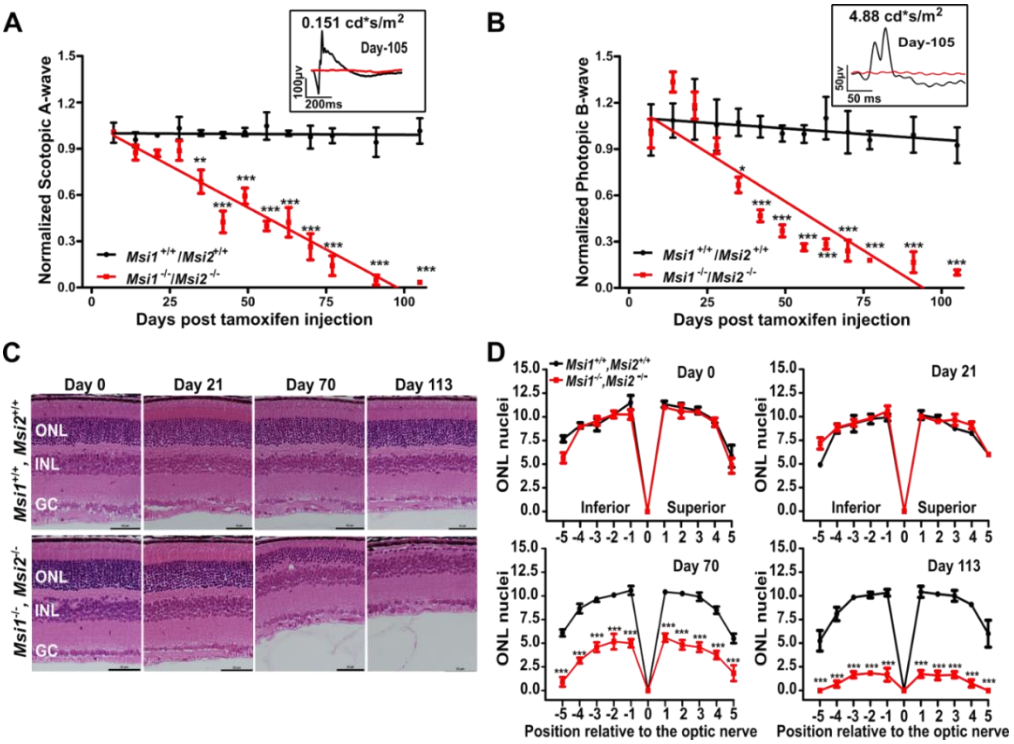


Figure 2-2 Progressive loss of function and retinal degeneration after double knockout of Msi1 and Msi2 in photoreceptor cells

Figure 2. Progressive loss of response to light and retinal degeneration after double knockout of *Msi1* and *Msi2* in photoreceptor cells. Time course of scotopic A-wave **(A)** and photopic B-wave **(B)** responses from *Msi1*$^{+/+}$*/Msi2* $^{+/+}$ (black line), and *Msi1* $^{-/-}$ */Msi2* $^{-/-}$ retinas (red line) following tamoxifen injection. Scotopic and photopic waveforms were obtained at 0.151 cd*s/m^2 and 4.88 cd*s/m^2 flashes, respectively. Insets show representative electroretinograms from a single *Msi1*$^{+/+}$*/Msi2* $^{+/+}$ (black line) and *Msi1* $^{-/-}$ */Msi2* $^{-/-}$ (red line) mouse retina at day 105 post tamoxifen injection. The data points from the scotopic and photopic responses are represented as mean ± SEM of 8 eyes (4 animals). A pairwise t-test with Bonferroni correction for multiple comparisons was used to determine the effect of genotype on the A-wave amplitude at each time point. Significance levels of the pairwise comparisons is indicated as: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001. **C)** Outer nuclear layer degeneration in *Msi1*$^{-/-}$*/Msi2*$^{-/-}$ knockout retinas. Representative H&E-stained sections from the *Msi1*$^{+/+}$*/Msi2*$^{+/+}$ and *Msi1*$^{-/-}$*/Msi2*$^{-/-}$ retinas collected between day 0 and day 113 post-tamoxifen injection. ONL: outer nuclear layer (Photoreceptor nuclei), INL: inner nuclear layer, GC: ganglion cells. 40X objectives and scale bar represent 50 µm. **D)** Spider plots displaying the thickness of the ONL as the number of nuclei measured at ten points stepped by 400µm from the optical nerve at different time points post-tamoxifen injection. Data are shown as mean ± SEM. A pairwise t-test with Bonferroni

correction for multiple comparisons was used to determine the effect of the genotype on the outer nuclear layer thickness at each time point. Significance levels of the pairwise comparisons is indicated as: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.

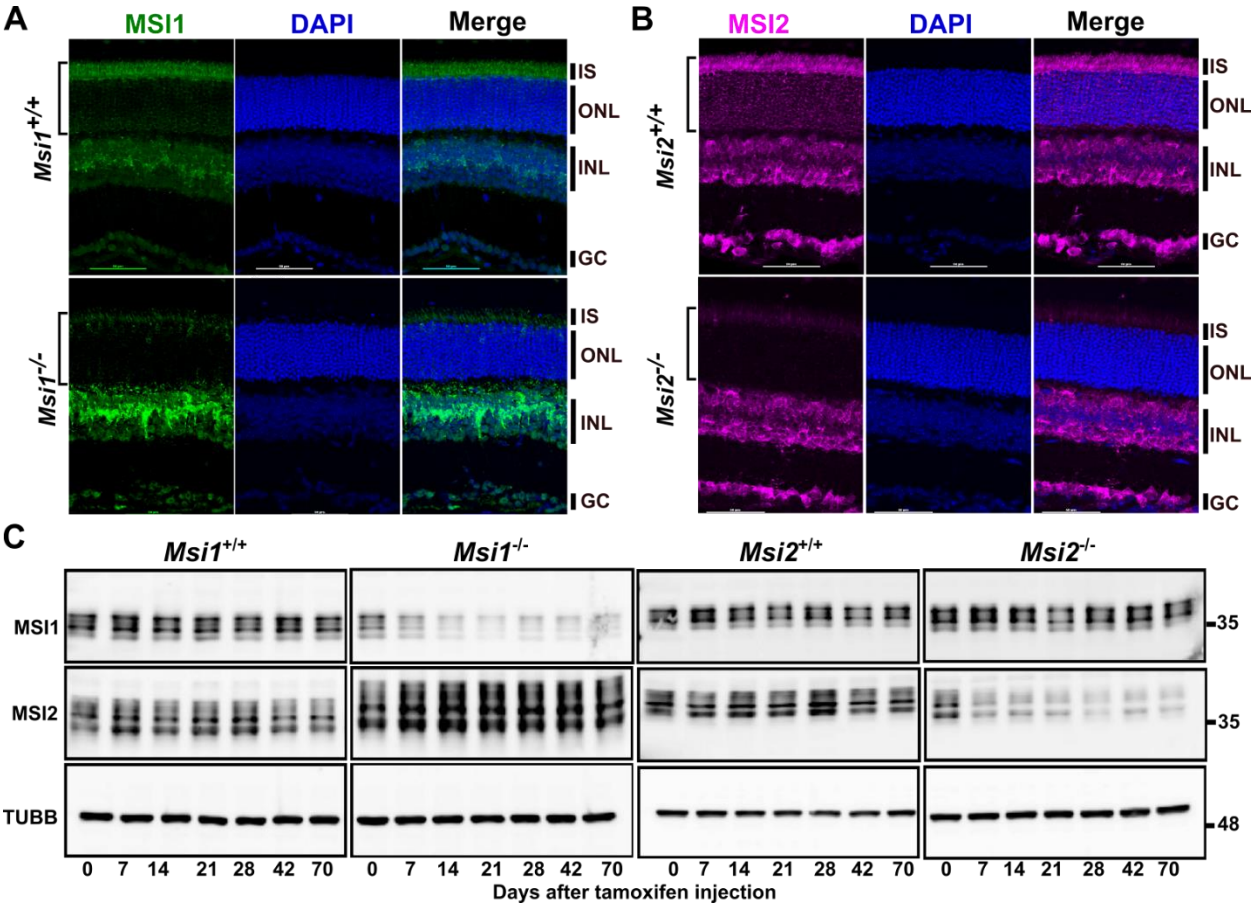


Figure 2-3 Induced single knockouts of Msi1 or Msi2 in photoreceptor cells

Figure 3. Induced single knockouts of *Msi1* or *Msi2* in photoreceptor cells. . A) Immunofluorescence micrographs of retinal cross-sections collected 14 days after tamoxifen injection at postnatal day 30 from *Msi1*$^{-/-}$ mice and *Msi1*$^{+/+}$ littermates **(A)** or *Msi2*$^{-/-}$ mice and *Msi2*$^{+/+}$ littermates **(B).** Sections were stained with antibodies to MSI1(green), MSI2(magenta) and DAPI (blue). ONL: outer nuclear layer (photoreceptor cell layer). INL: inner nuclear layer. GC: ganglion cell layer. Objective, 40X. **C)** Immunoblot of lysates from *Msi1*$^{+/+}$ and *Msi1*$^{-/-}$ retinas collected between 0 and 70 days after tamoxifen injections and probed with antibodies to MSI1, MSI2 and TUBB (loading control). See Supplement Figure 3 for full size blots.

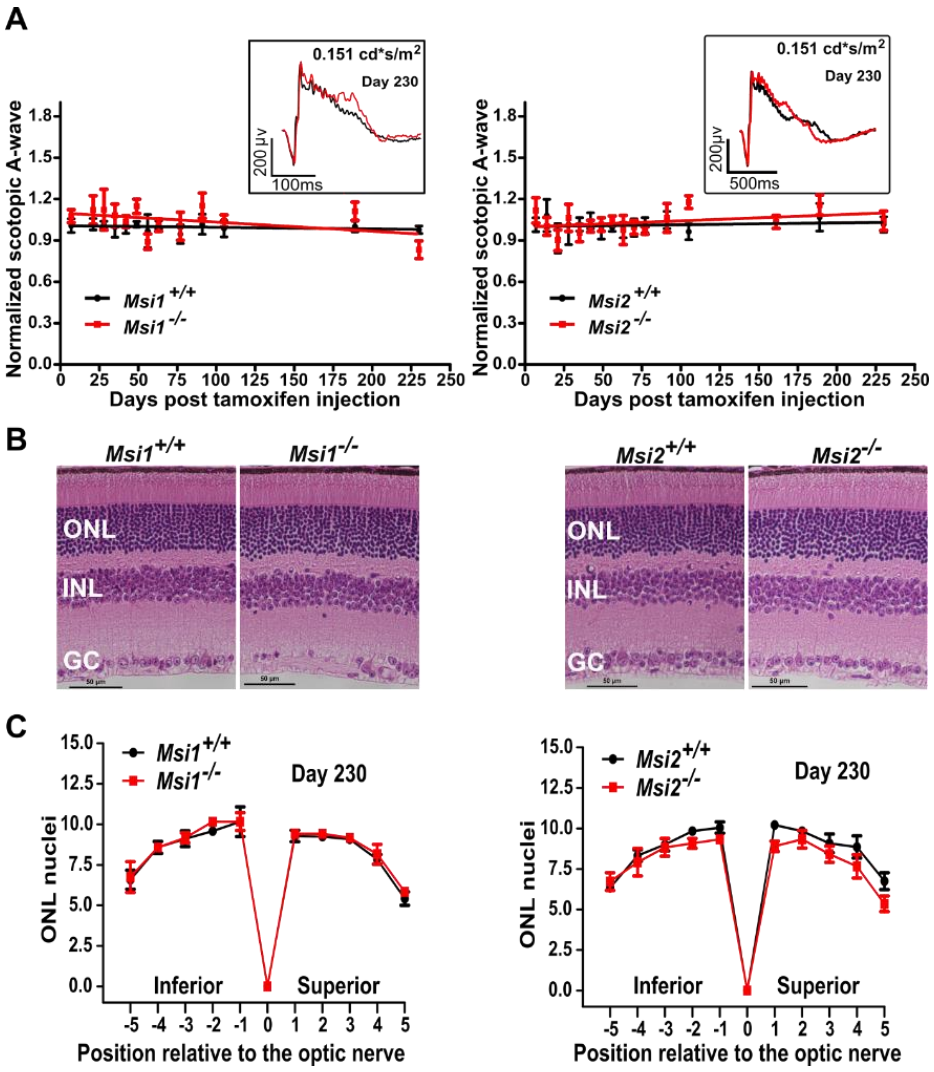


Figure 2-4 Normal retinal function and morphology in the single knockouts Msi1 or Msi2 in photoreceptor cells

Figure 4. Normal photoreceptor response to light and retinal morphology in the single knockouts *Msi1* or *Msi2* in photoreceptor cells. A) Scotopic mean A-wave response from single *Msi1* $^{-/-}$ (left) or *Msi2* $^{-/-}$ (right) knockouts in photoreceptors (red lines) and littermate controls (black lines) recorded between 0 and 230 days post tamoxifen injection. Scotopic waveforms were obtained using 0.151 cd-s/m^2 flashes. The insets of panels A and B show representative scotopic (dark-adapted) electroretinograms from a single knockout (red) or control (black) retina at D230 post-injection using 0.151 cd-s/m^2 flashes. Scotopic waveforms were obtained after dark overnight adaptation using 0.151 cd-s/m^2 flashes. **B)** Representative H&E-stained sections from retinas of single *Msi1* $^{-/-}$ (left) or *Msi2* $^{-/-}$ (right) knockouts in photoreceptor cells and littermate controls collected 230 days post-tamoxifen injection. ONL: outer nuclear layer (Photoreceptor nuclei), INL: inner nuclear layer, GC: ganglion cells. 40X objectives, scale bar, represents 10µm.

C) Spider plots of ONL thickness for single *Msi1* $^{-/-}$ (left) or *Msi2* $^{-/-}$ (right) knockouts (red) and littermate controls (black). The plots display the thickness of the ONL as the number of nuclei measured at ten points stepped by 400µm from the optical nerve in retinas collected 230 days after tamoxifen injection. Data are shown as mean ± SEM of 8 eyes from 4 animals.

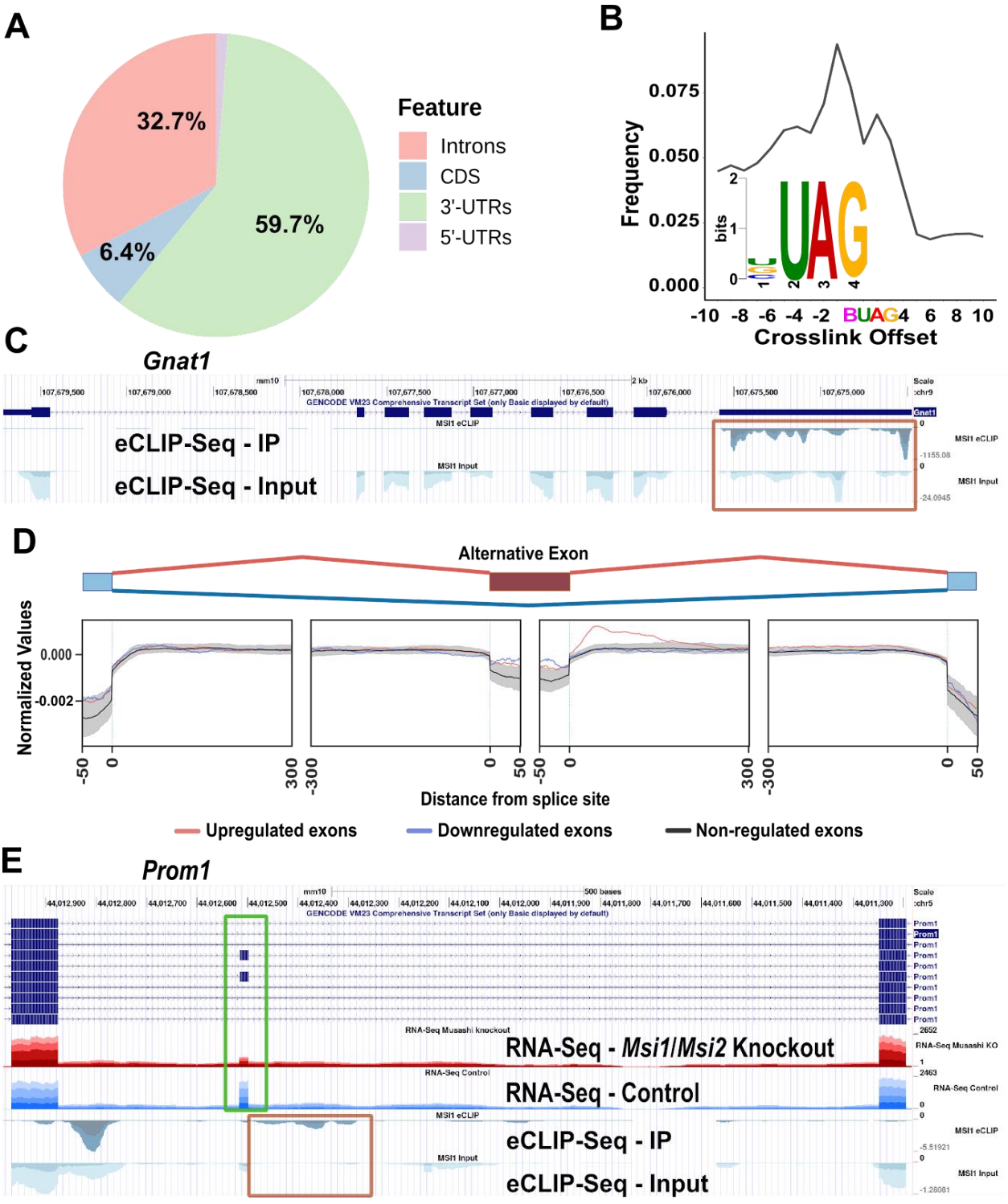


Figure 2-5 In the retina MSI1 binds to UAG motifs located predominantly in introns and 3'-UTRs

Figure 5. In the retina MSI1 binds to UAG motifs located predominantly in introns and 3'-UTRs. **A)** Distribution of MSI1 binding sites as identified by eCLIP-Seq on mouse retinal samples across mRNA features. **B)** eCLIP crosslink frequency relative to the top scoring motif (BUAG)

identified by DREME. **C)** UCSC Genome browser tracks showing the eCLIP-seq signal enrichment over the 3'-UTR of the *Gnat1* gene (orange box). Replicates are stacked and indicated by different shading colors. Scales are 0 to -1155 for the eCLIP IP and 0 to -24 for the eCLIP input. **D)** RNA binding protein map showing MSI1 binding relative to an alternative metaexon. Exons upregulated by the Musashi proteins (downregulated in the *Msi1*/*Msi2* double knockout) are shown in red, exons downregulated by the Musashi proteins are shown in blue and alternative exons remaining unchanged in the knockout are shown in black. Gray shading indicates the 99.5% confidence interval derived from 1000 random permutations. MSI1 binding sites are enriched downstream of alternative exons upregulated by the Musashi proteins. **E)** UCSC Genome browser tracks showing MSI1 binding (orange box) downstream of an alternative exon in the *Prom1* gene regulated by the Musashi proteins (green box). RNA-Seq tracks show the read density for retinal samples derived from photoreceptor-specific *Msi1*/*Msi2* double knockouts and matched controls. Replicates are stacked and indicated by different shading colors. Scales are 0 to -5.5 for the eCLIP IP and 0 to -1.3 for the eCLIP input.

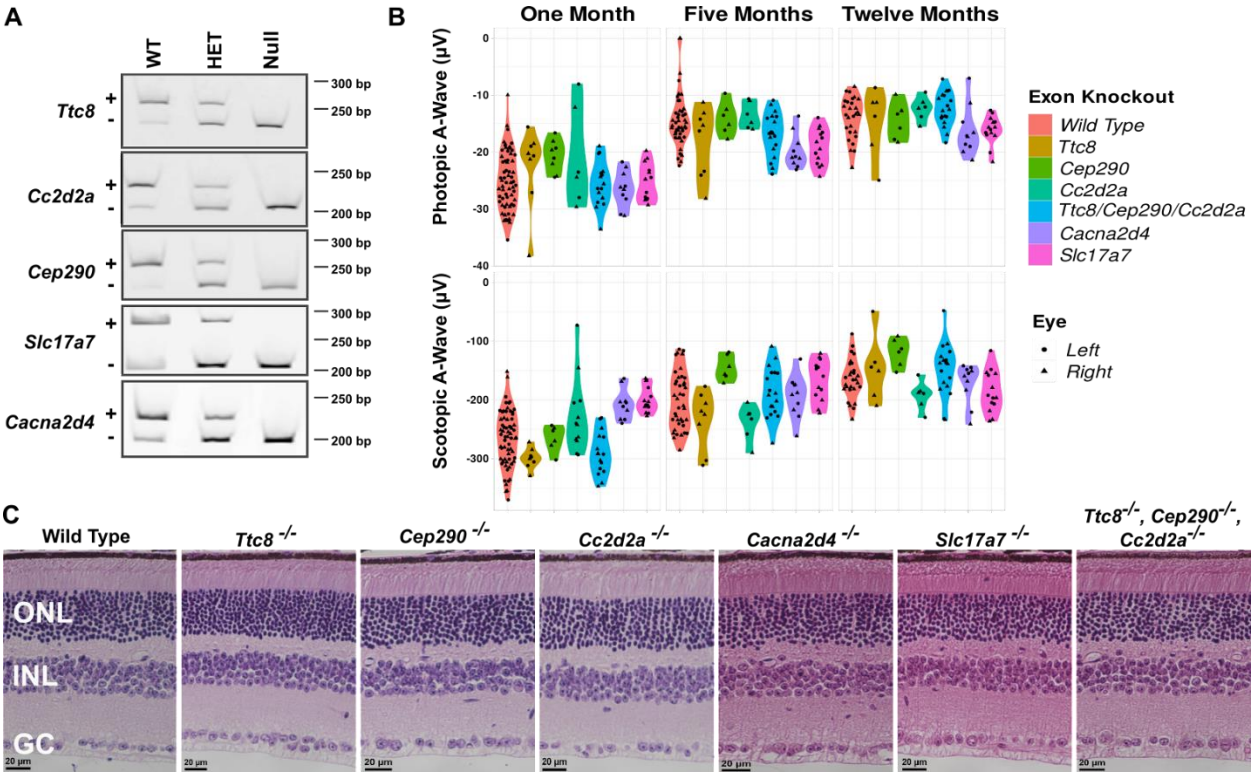


Figure 2-6 Normal photoreceptor response to light and retinal morphology of knockouts of photoreceptor-specific exons in the Ttc8, Cc2d2a, Cep290, Cacna2d4, and Slc17a7 genes

Figure 6. Normal photoreceptor response to light and retinal morphology of knockouts of photoreceptor-specific exons in the *Ttc8*, *Cc2d2a*, *Cep290*, *Cacna2d4*, and *Slc17a7* genes. **A)** RT-PCR of retinal samples showing loss of the photoreceptor-specific mRNA isoforms in the exon knockout animals. RNA was extracted from the retinas of wild-type animals (WT), and heterozygous (Het) and homozygous (Null) exon knockouts. Isoforms including and skipping the alternative exon are indicated by "+" and "-", respectively. **B.** Violin plots of the photopic and scotopic A-wave intensities at postnatal days 30, 150, and 356 were collected from the single and triple (*Ttc8*, *Cc2d2a*, *Cep290*) exon knockouts. **C)** Representative H&E-stained retinal sections from wild-type mice, knockouts of photoreceptor-specific microexons in the *Ttc8*, *Cep290*, *Cc2d2a*, *Canca2d4,* and *Slc17a7* genes, and combined deletion of the microexons in the *Ttc8*, *Cep290*, *Cc2d2a* genes. The samples were collected at 12 months of age. ONL: outer nuclear layer (Photoreceptor nuclei), INL: inner nuclear layer, GC: ganglion cells. 40X objectives, scale bar, represents 20μm.

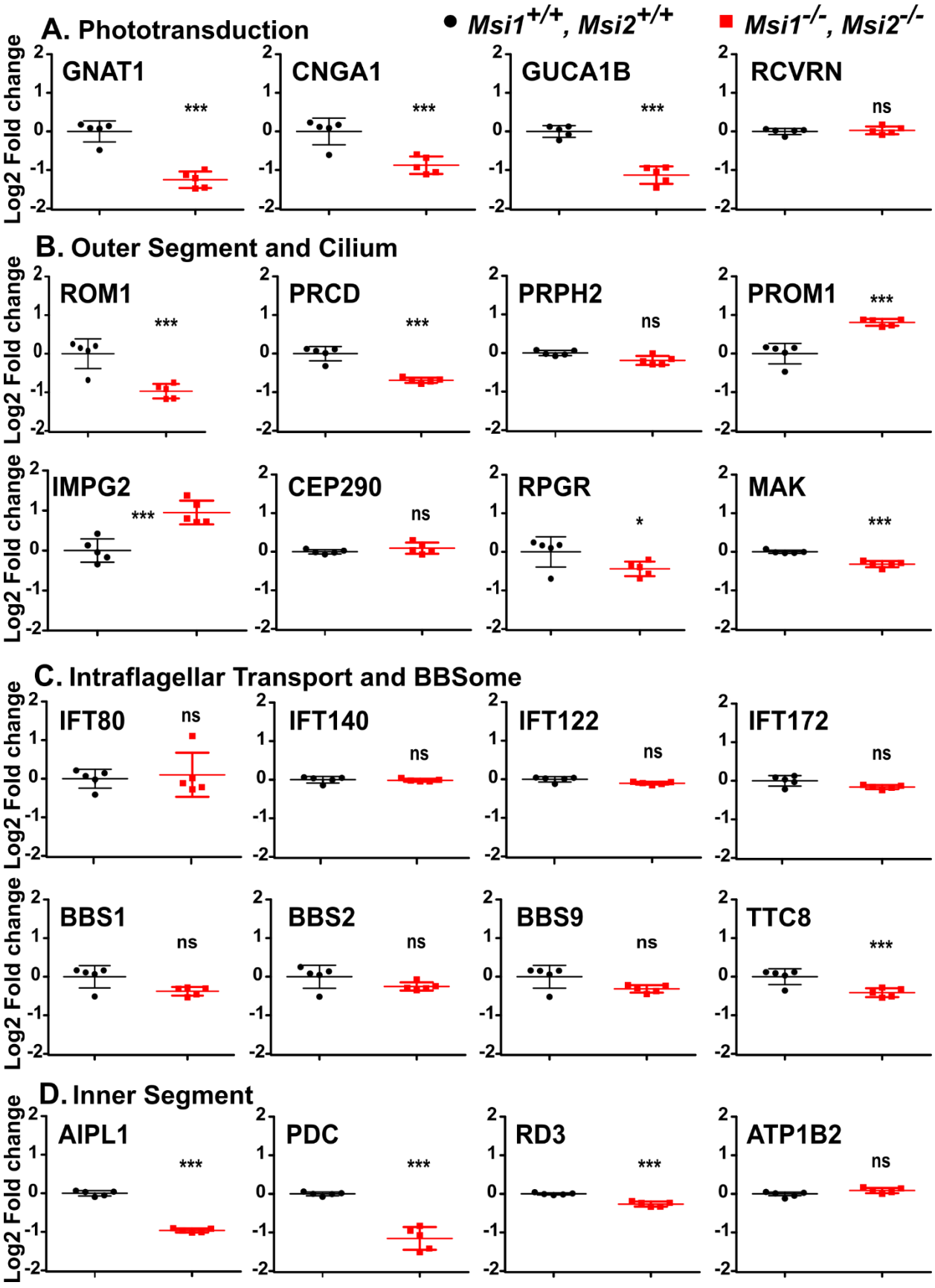

A. Phototransduction
● Msi1+/+, Msi2+/+
■ Msi1-/-, Msi2-/-
GNAT1
CNGA1
GUCA1B
RCVRN
B. Outer Segment and Cilium
ROM1
PRCD
PRPH2
PROM1
IMPG2
CEP290
RPGR
MAK
C. Intraflagellar Transport and BBSome
IFT80
IFT140
IFT122
IFT172
BBS1
BBS2
BBS9
TTC8
D. Inner Segment
AIPL1
PDC
RD3
ATP1B2
Log2 Fold change
***
*
ns

Figure 2-7 Expression of proteins critical for photoreceptor function after induced knockout of Msi1 and Msi2 in photoreceptor cells

Figure 7. Expression of proteins critical for photoreceptor function after induced knockout of *Msi1* and *Msi2* in photoreceptor cells. Boxplots representing the log2 of the fold change in protein expression relative to the median of the control. Retinal samples were collected 21 days after inducing the knockout at postnatal day 30. Protein levels were determined using isobaric labeling and mass-spectroscopy. Boxplots represent a selected set of proteins that are components of the phototransduction pathway (**A)**, outer segment primary cilium structure **(B)**, intraflagellar transport complex and BBSome **(C)**, and the inner segment **(D)**. The data is represented as mean ± SEM of five replicates. Significance level is indicated as: * adjusted p-value < 0.05, ** adjusted p-value < 0.01, *** adjusted p-value < 0.001. Online Supplementary table 17 contains the source data underlying the graphs.

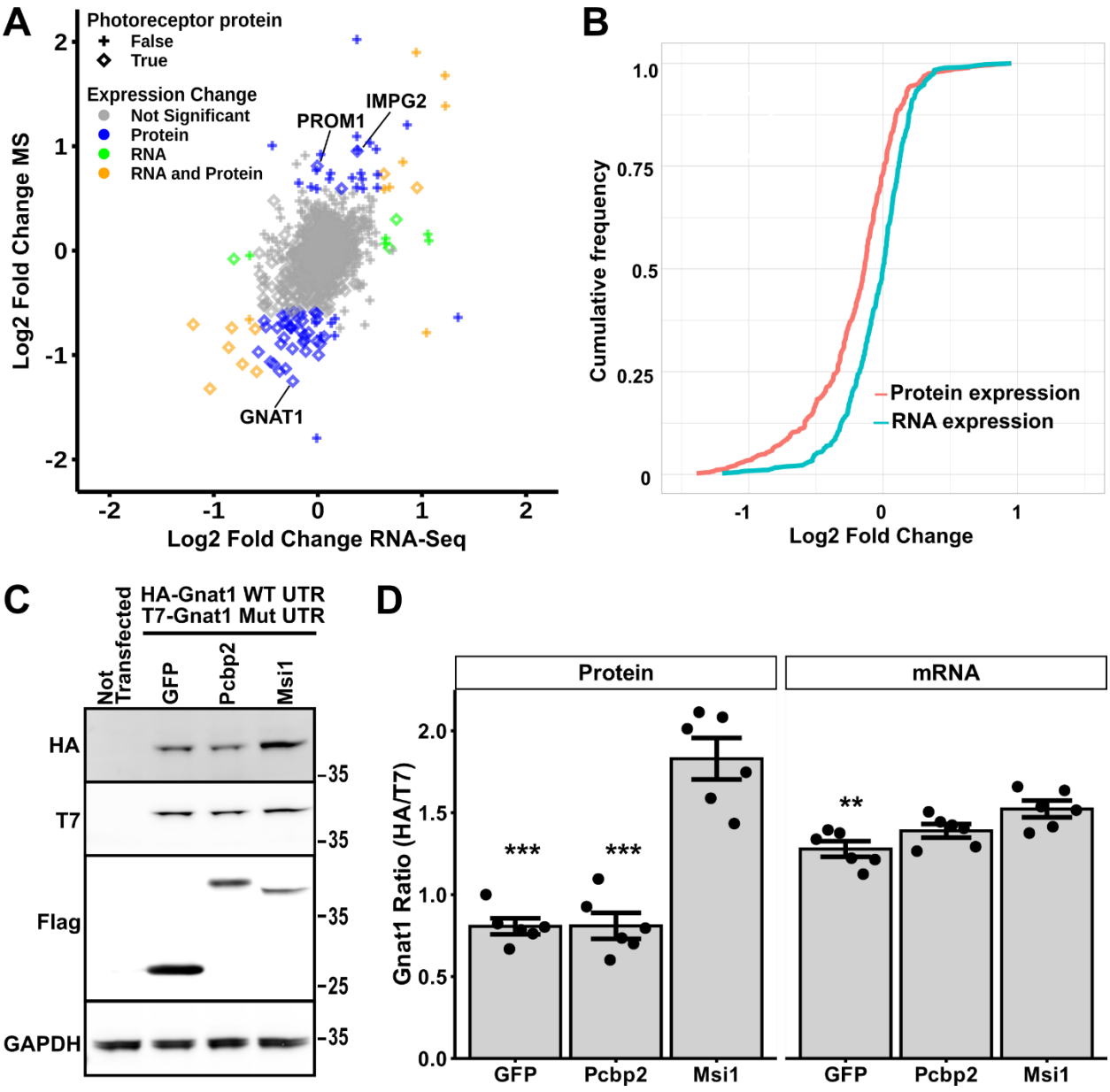
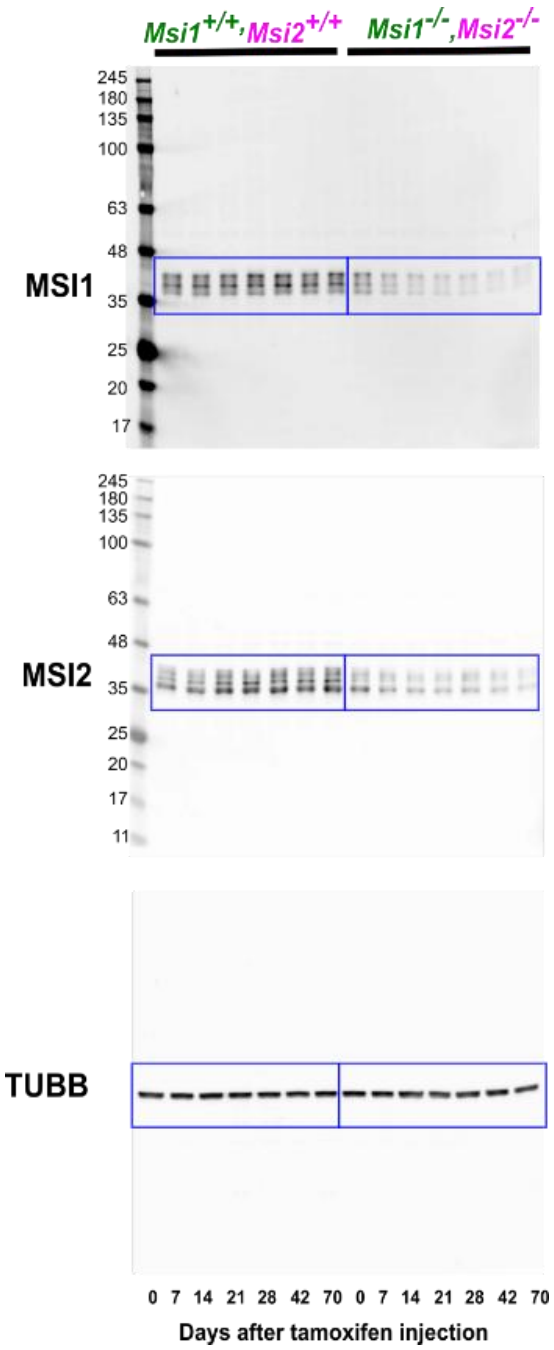


Figure 2-8 In photoreceptor cells, MSI1 and MSI2 act to promote protein expression post-transcriptionally

Figure 8. In photoreceptor cells, MSI1 and MSI2 act to promote protein expression post-transcriptionally. A). Scatter plot comparing protein and transcript levels changes after double *Msi1*/*Msi2* knockout in mature photoreceptor cells. Protein expression was quantified by isobaric labeling and mass spectrometry. Transcript expression levels were determined by RNA-Seq. Genes that are highly expressed specifically in photoreceptor cells are shown as rombs and genes highly expressed outside of photoreceptor cells are shown as crosses. Color indicates significant changes in protein levels alone (blue), in mRNA levels alone (green), in both protein and mRNA levels (orange), or no significant change in expression (gray). **B)** Cumulative

frequency plots showing the deletion of *Msi1/Msi2* in mature photoreceptors leads to broad decrease of photoreceptor protein levels in excess to changes in transcript levels (Kolmogorov-Smirnov p-value=$1.1*10^{-12}$). **C)** Western blot analysis of recombinant GNAT1 expression in NIH 3T3 cells. Cells were transfected with equal amounts HA-tagged *Gnat1* clone with wild type 3'-UTR and T7-tagged *Gnat1* clone with mutant 3'-UTR that lacks Musashi binding sites. Each transfection included one of the following: vector expressing flag-tagged GFP, vector expressing flag-tagged *Pcbp2*, and vector expressing flag-tagged *Msi1*. Non transfected cells were included as a control for the specificity of the antibodies. TUBB serves as loading control. **D)** Ratios of the HA (wild type) to T7 (mutant) tagged GNAT1 proteins and ratios of the corresponding RNA transcripts in the transfected NIH 3T3 cells. The data is represented as mean ± SEM of six replicates. The statistical significance of the effect of MSI1 in pairwise comparisons to the controls was determined by Tukey HSD. Significance level is indicated as:* p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001. Online Supplementary table 15 contains the source data underlying the graph on panel D.

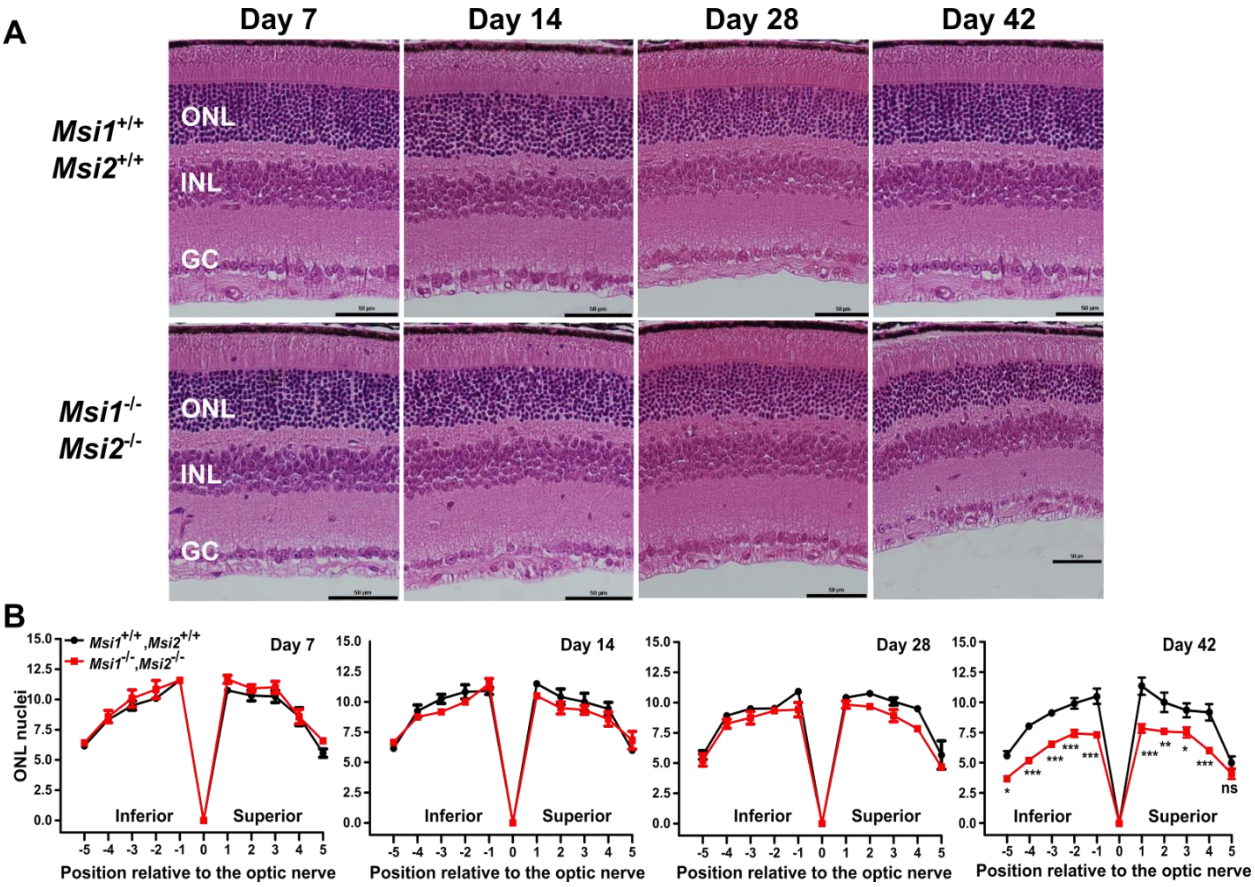
2.11 Supplementary information



Supplement Figure 2-9 Full-size western blot images for the data presented on Figure 1B

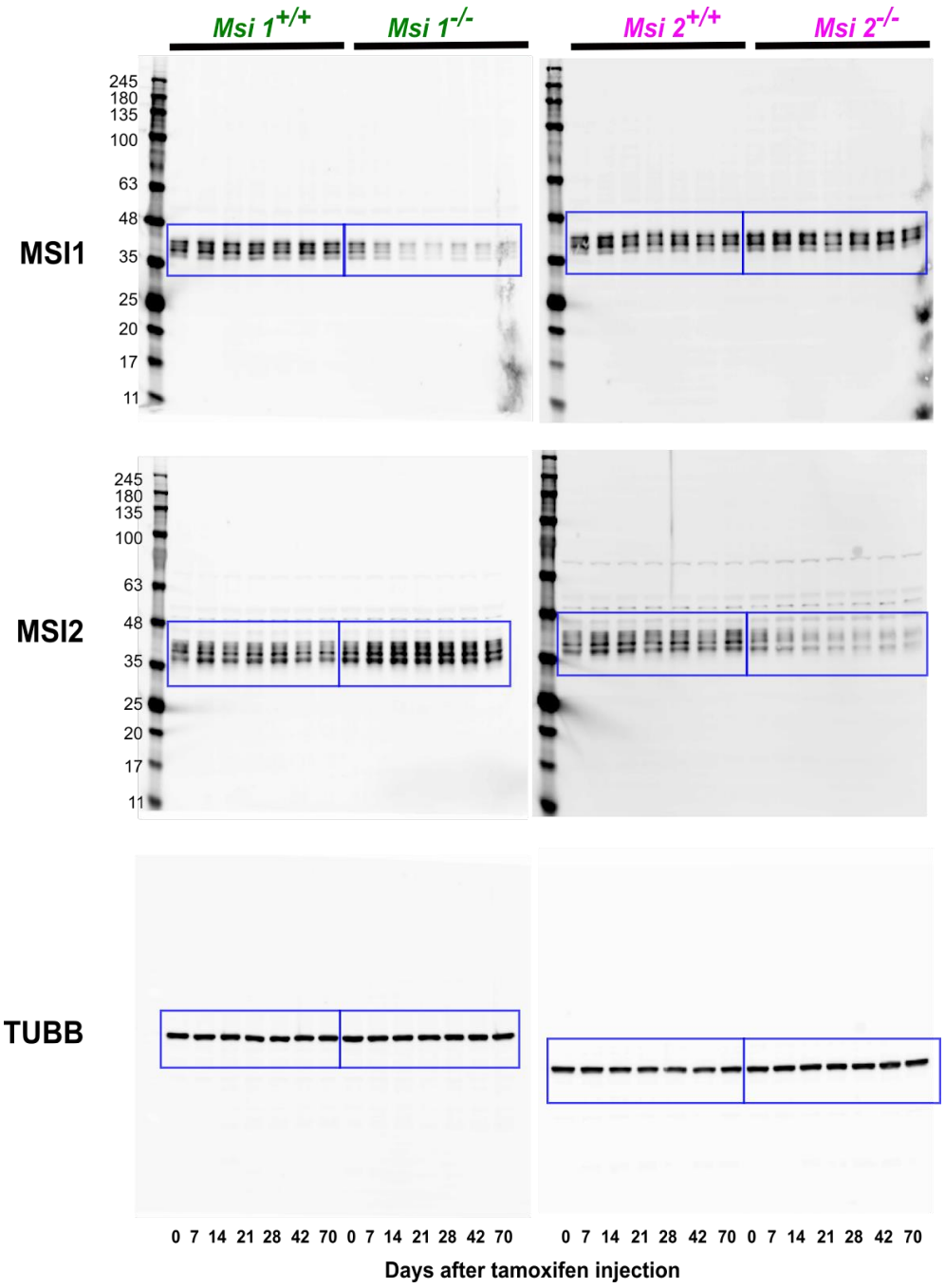
Supplement Figure 1. Full-size western blot images for the data presented on Figure 1B (blue boxes). Labels on the left indicate the protein being probed. Animal genotypes are indicated

on the top and the days after tamoxifen injection are indicated at the bottom. Day 0 after tamoxifen injection corresponds to postnatal day 30.



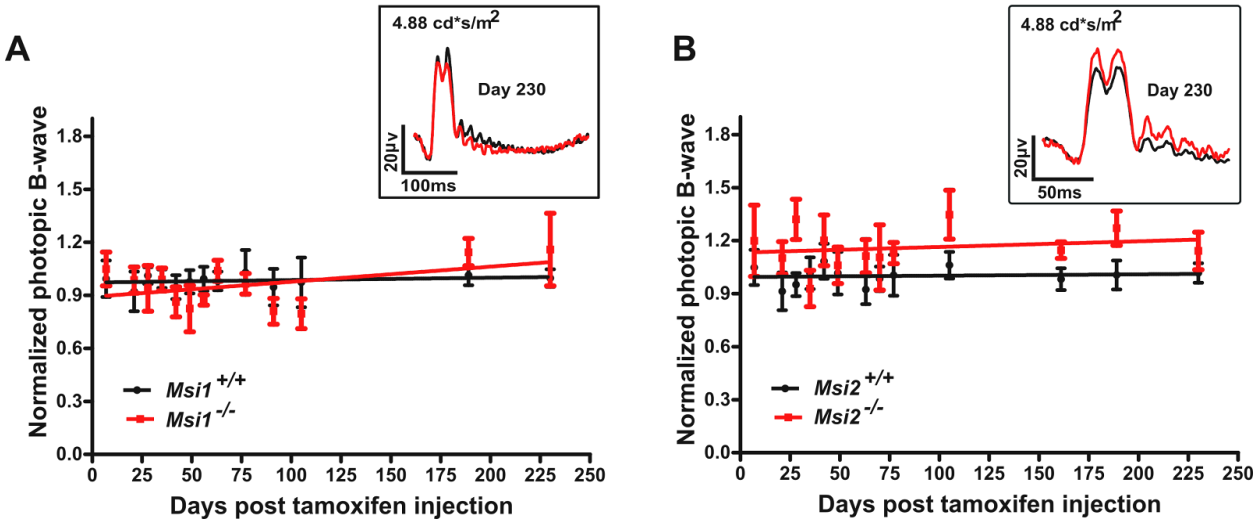
Supplement Figure 2-10 *Outer nuclear layer thickness of the retina after double knockout of Msi1 and Msi2 in photoreceptor cells*

Supplement Figure 2. Outer nuclear layer thickness of the retina after double knockout of *Msi1* and *Msi2* in photoreceptor cells. **(A)** Representative H&E-stained eye sections from the double *Msi1*/*Msi2* knockout and age-matched controls at 7, 14, 28, and 42 days after inducing the knockout. ONL: outer nuclear layer (Photoreceptor nuclei), INL: inner nuclear layer, GC: ganglion cells. 40X objectives and scale bar represent 50 µm. **(B)** Spider plots displaying the thickness of the ONL as the number of nuclei measured at ten points stepped by 400µm from the optical nerve at different time points post-tamoxifen injection. Data are shown as mean ± SEM. A pairwise t-test with Bonferroni correction for multiple comparisons was used to determine the effect of the genotype on the outer nuclear layer thickness at each time point. Significance levels of the pairwise comparisons is indicated as: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.



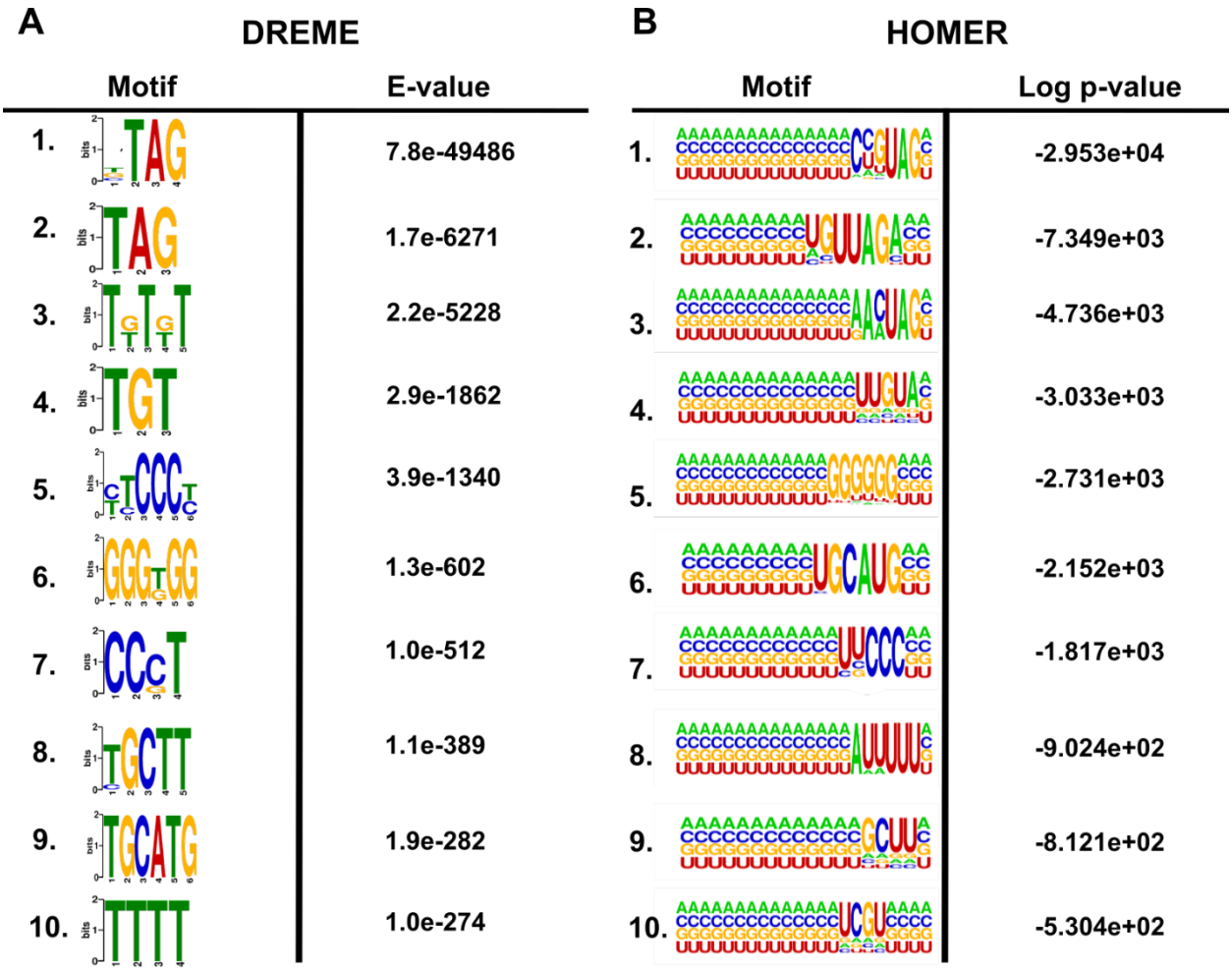
Supplement Figure 2-11 *Full-size western blot images for the data presented on Figure 1B.*

Supplement Figure 3. Full-size western blot images for the data presented on Figure 1B (blue boxes). Labels on the left indicate the protein being probed. Animal genotypes are indicated on the top and the days after tamoxifen injection are indicated at the bottom. Day 0 after tamoxifen injection corresponds to postnatal day 30.



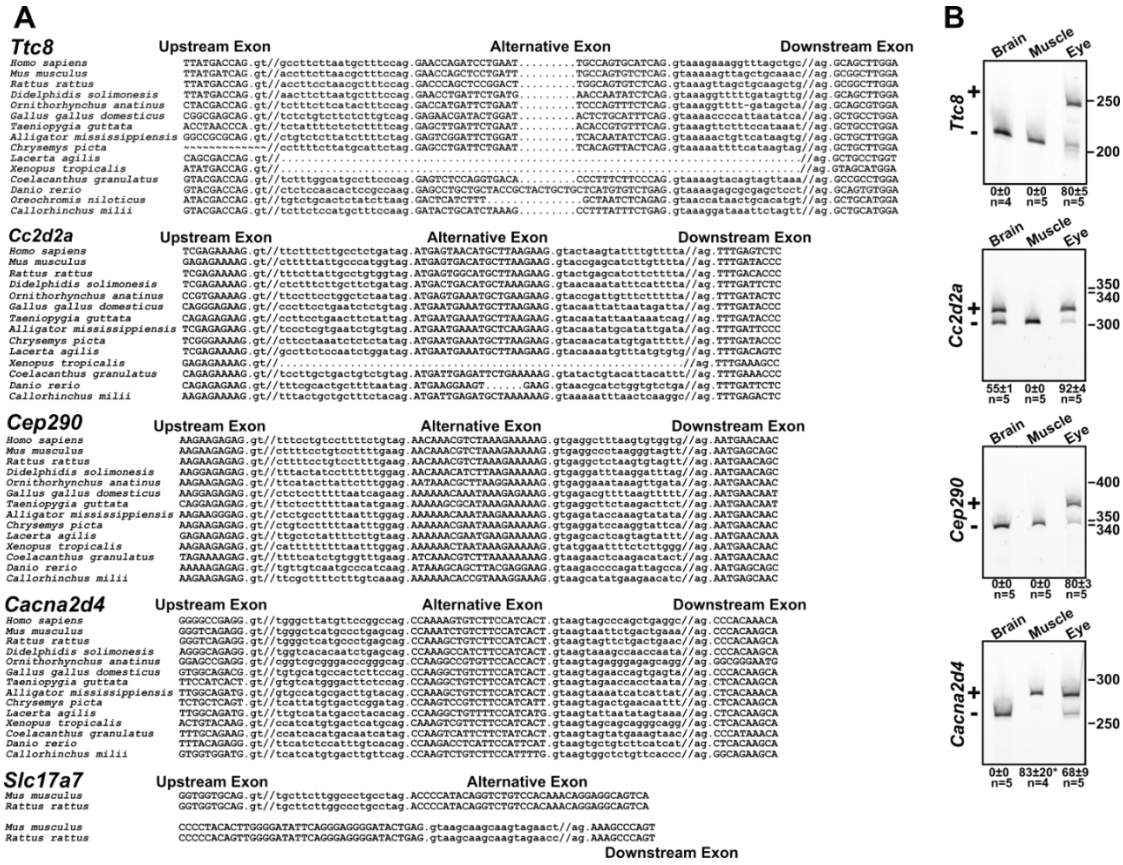
Supplement Figure 2-12 *Normal photopic response to light in the single Msi1 or Msi2 knockouts*

Supplement Figure 4. Normal photopic response to light in the single *Msi1* or *Msi2* knockouts. A) Photopic mean B-wave response of knockout (red) *Msi1*$^{-/-}$ **(A)** and *Msi1*$^{-/-}$ **(B)** mice (red line) and age-matched controls (black) between day 16 and day 230 post tamoxifen injection. Photopic waveforms were obtained after light adaptation using 4.88 cd-s/m^2 flashes. The insets show representative photopic (light-adapted) electroretinograms recorded 230 days post-injection using 4.88 cd-s/m^2 flashes. The data points from the photopic responses of the single depletion of *Msi1* or *Msi2* are represented as mean ± SEM of 8 eyes from 4 animals.



Supplement Figure 2-13 Sequence motifs enriched near eCLIP-Seq derived MSI1 crosslinks sites

Supplement Figure 5. Sequence motifs enriched near eCLIP-Seq derived MSI1 crosslinks sites. Logos of the top ten significantly enriched motifs identified by DREME **(A)** or HOMER **(B)** in the vicinity of MSI1 crosslink sites.

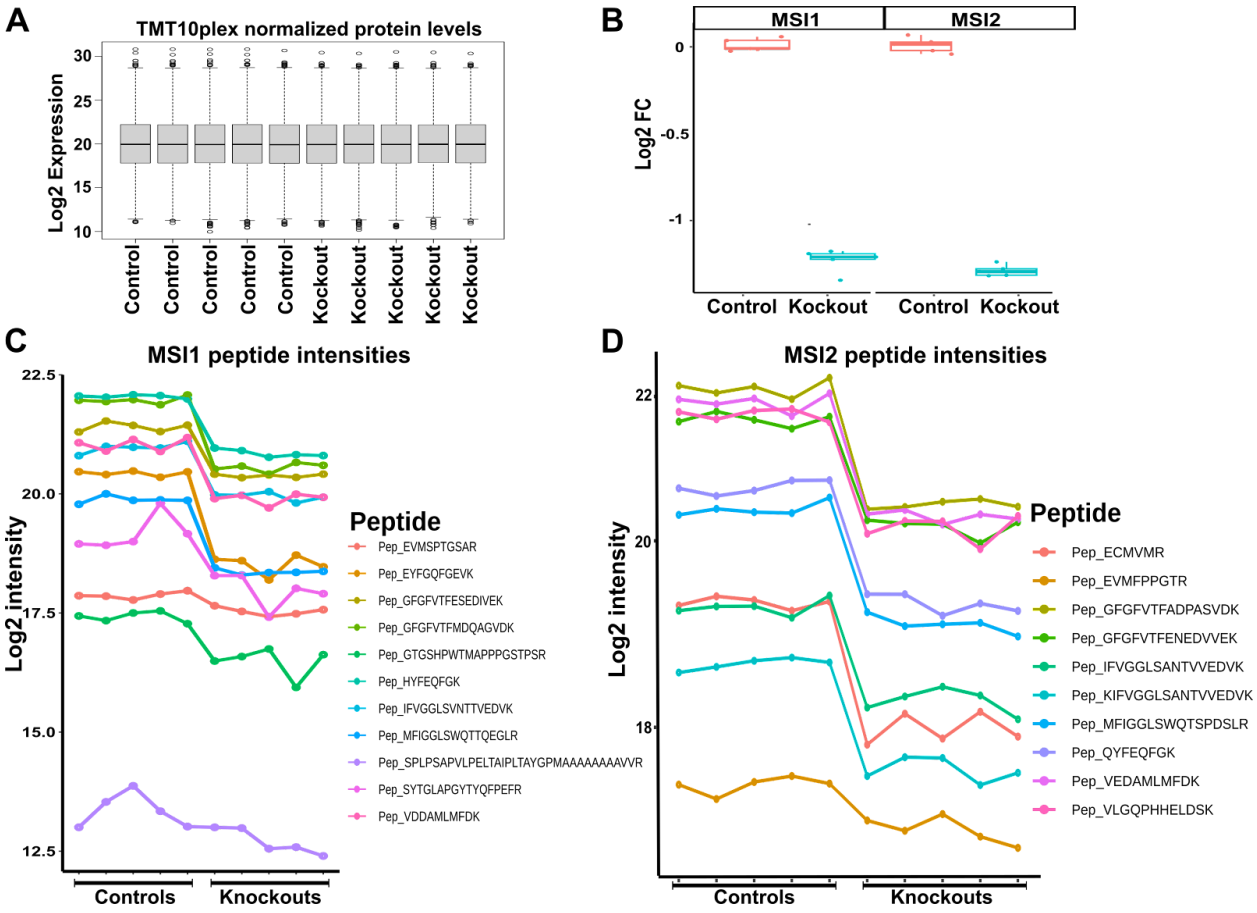


Supplement Figure 2-14 *Conservation of the photoreceptor-specific exons of Ttc8, Cc2d2a, Cep290, Cacna2d4, and Slc17a7 across vertebrates*

Supplement Figure 6: Conservation of the photoreceptor-specific exons of *Ttc8, Cc2d2a, Cep290, Cacna2d4, and Slc17a7* across vertebrates. (A) Alignments showing the alternative exons and parts of the flanking constitutive exons. Introns are in lower case, and exons are in upper case. Forward slashes indicate where intronic sequences were removed from the alignment for ease of presentation. Homologous exons for the photoreceptor-specific exons in *Ttc8, Cc2d2a, Cep290,* and *Cacna2d4* can be traced down to Chondrichthyes. The exons in *Ttc8,* and *Cc2d2a* can vary in length while preserving the reading frame or be completely absent from certain species. The upstream exon is not available for *Chrysemys picta* due to gaps in the genome sequence. The exon in *Slc17a7* is present only in rodents. **(B)** Analysis of the inclusion rate of the zebrafish homologues of the photoreceptor-specific exons in the *Ttc8*, *Cc2d2a*, *Cep290*, and *Cacna2d4* in brain, muscle, and eye samples. Numbers under the figure indicate the percent inclusion of the exon ± SEM. All four exons have high inclusion rates in the eye. Unlike their mouse homologues the zebrafish exons in the *Cc2d2a* and *Cacna2d4* genes are also included at high rate in the brain and muscle, respectively. *The exon in *Cacna2d4* was typically included at 100% (3 out of 4 tested samples).

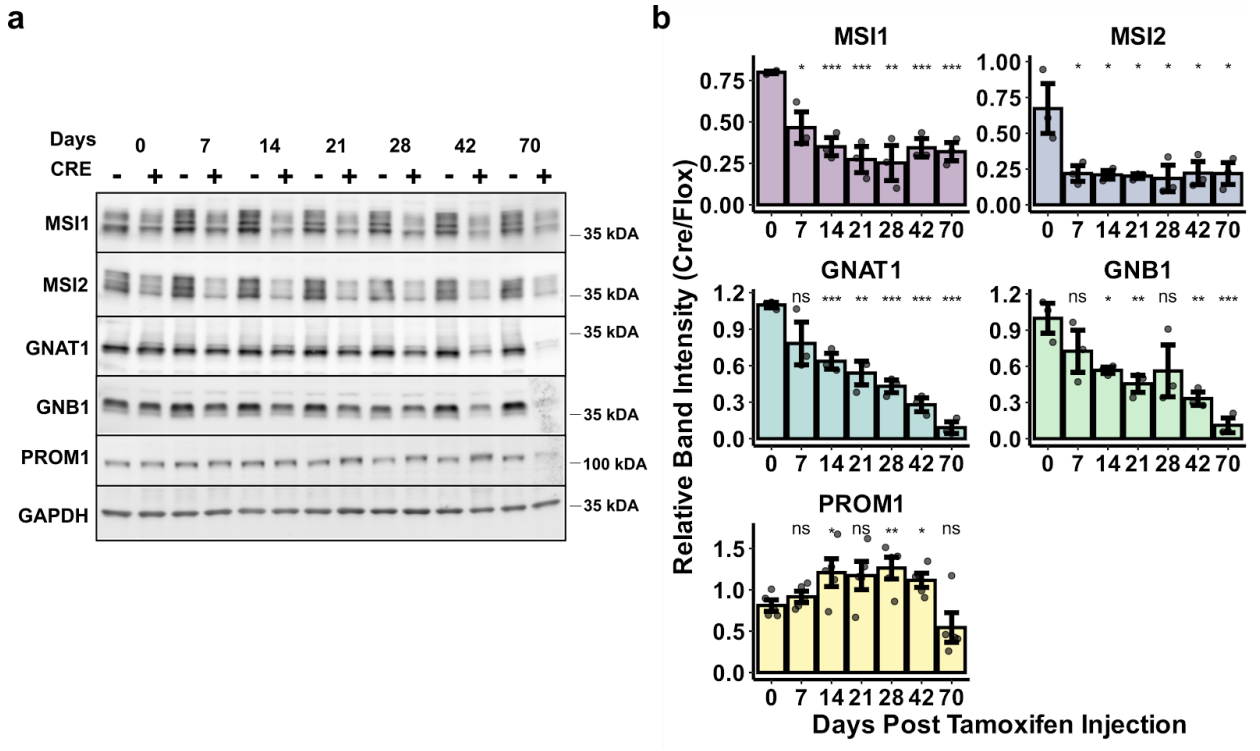
Supplement Figure 2-15 Exon deletion alleles generated by CRISPR/Cas9 mutagenesis

Supplement Figure 7. Exon deletion alleles generated by CRISPR/Cas9 mutagenesis. A) Schematic of the CRISPR/Cas9 targeting strategy. Two guide RNAs are used to direct cuts on both sides of the exon leading to its deletion. **(B)** Sequences of the knockout alleles (KO) aligned to the wild-type genomic sequence (WT). The exons are shown in uppercase, and introns are in lowercase. Arrows indicate the position and orientation (5' to 3') of the guide RNAs.



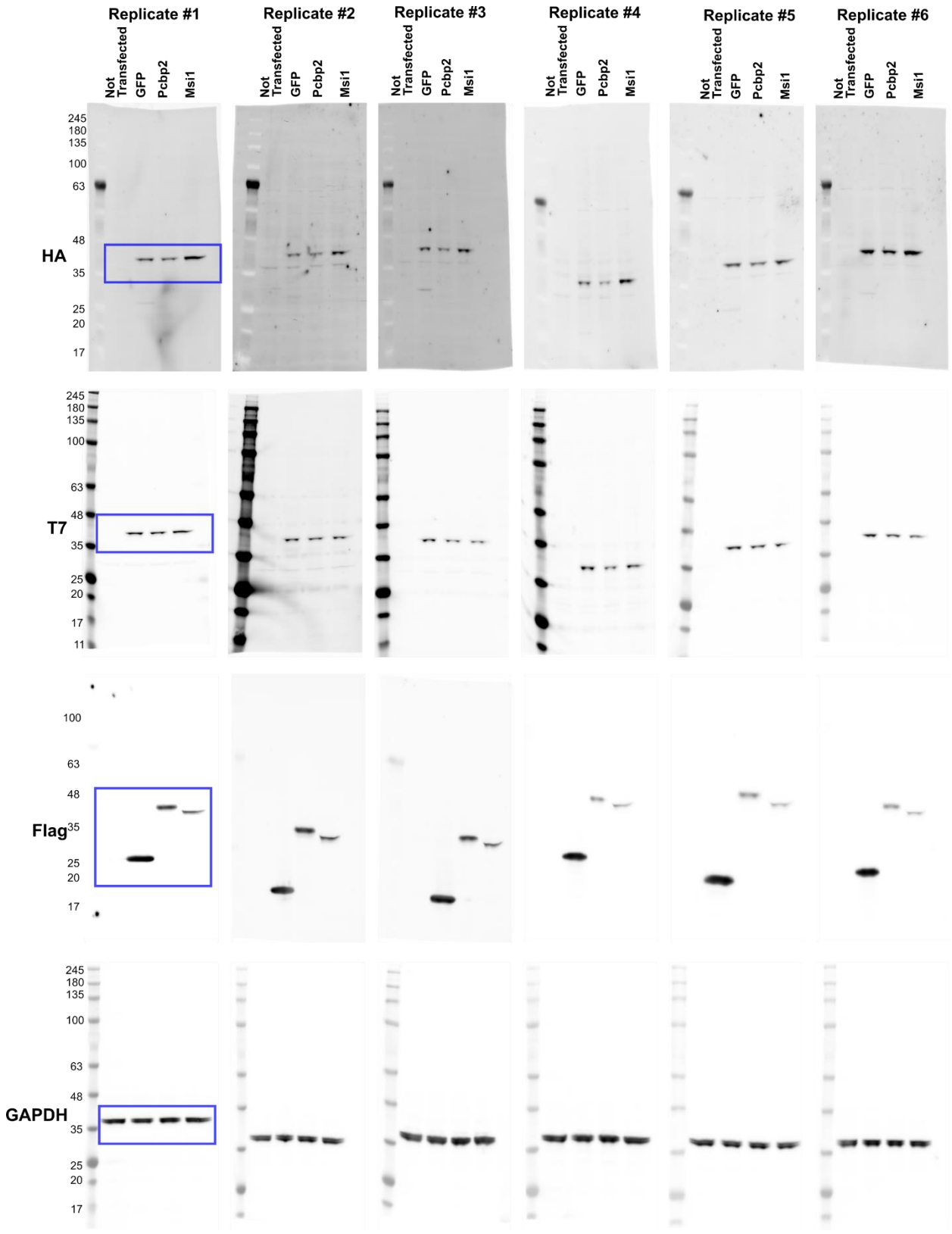
Supplement Figure 2-16 Decrease of MSI1 and MSI2 protein levels in the retina after induced double knockout of Msi1 and Msi2 in mature photoreceptor cells

Supplement Figure 8. Decrease of MSI1 and MSI2 protein levels in the retina after induced double knockout of *Msi1* and *Msi2* in mature photoreceptor cells. **A)** Box plot showing the distribution of normalized signal intensities across samples analyzed by isobaric labeling and tandem MS (MS3). **B)** Box plots showing the log2 of the fold difference of MSI1 and MSI2 protein levels in the retina of control and knockout mice relative to the median of the controls. Changes in the levels of individual tryptic peptides identified for MSI1 **(C)** and MSI2 **(D)** in control and knockout retina.



Supplement Figure 2-17 Validation of MS3 data by western blot

Supplement Figure 9. **Validation of MS3 data by western blot**. **A)** Representative immunoblot showing levels of selected proteins after combined deletion of *Msi1* and *Msi2* in mature photoreceptors. B) Quantification of western blot data for MSI1, MSI2, GNAT1, GNB1, and PROM1. TUBB1 and GAPDH were used as controls to normalize for loading. Error bars represent standard error of the mean (SEM, n=3). The statistical significance of the pairwise comparisons of the protein levels at different time points to the baseline level at the day of the tamoxifen injection is indicated as: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.



Supplement Figure 2-18 *Full size blots for data presented in figure 8C*

Supplement Figure 10. Full size western blot images for the data presented on Figure 8C (blue boxes). Labels on the left indicate the protein being probed. The transfected samples include the construct indicated on the top along with expression constructs for HA-Gnat1 with wild type 3'-UTR and T7-Gnat1 with mutant 3'-UTR.

3. Chapter 3

3.1 TITLE: The 13A4 monoclonal antibody to the mouse PROM1 protein recognizes a structural epitope

Authors: Fatimah Matalkah[1], Scott Rhodes[1], Visvanathan Ramamurthy[1,2], and Peter Stoilov[1]

Affiliation: [1]Department of Biochemistry, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, West Virginia, USA.

[2]Department of Ophthalmology, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, West Virginia, USA.

Matalkah F, Rhodes S, Ramamurthy V, Stoilov P. The mAB 13A4 monoclonal antibody to the mouse PROM1 protein recognizes a structural epitope. *PLoS One.* 2022 Oct 10;17(10): e0274958. doi: 10.1371/journal.pone.0274958. PMID: 36215230; PMCID: PMC9550058.

Note: This chapter include text and figures taken from the publications highlighted above

3.2 Abstract

Purpose: We endeavored to map the epitope of the rat monoclonal antibody mAB 13A4 to the mouse PROM1 (CD133, AC133) protein. mAB 13A4 is the main reagent used to detect the mouse PROM1 protein. PROM1 is required for the maintenance of primary cilia and mutations in the *Prom1* gene are associated with retinal degeneration.

Methods: Epitope-tagged clones of splice variants and tiled deletion mutants were used to map the mAB 13A4 epitope and test the predicted tertiary structure of PROM1. The proteins were expressed in Neuro 2a cells and analyzed by Western blot with antibodies to PROM1 and the epitope tag.

Results: Deletions in the second and third extracellular domains of the PROM1 protein disrupted the mAB 13A4 epitope. Furthermore, the affinity of mAB 13A4 to the major PROM1 isoform in photoreceptor cells is significantly reduced due to the inclusion of a photoreceptor-specific alternative exon in the third extracellular domain. Interestingly, a deletion in the photoreceptor-specific isoform of six amino acids adjacent to the alternative exon restored the affinity of mAB 13A4 to PROM1.

Conclusion: mAB 13A4 recognizes a structural epitope that is stabilized by two of the extracellular domains of PROM1. The results of our mutagenesis are consistent with the computationally predicted helical bundle structure of PROM1 and point to the utility of mAB 13A4 for evaluating the effect of mutations on the PROM1 structure. We show that the PROM1 isoform composition needs to be considered when interpreting tissue and developmental expression data produced by mAB 13A4.

3.3 Introduction

Prominin (PROM1, CD133, AC133) was identified as the antigen of monoclonal antibodies raised against human hematopoietic progenitors and mouse neuroepithelial cells [1–3]. PROM1 is a glycosylated membrane protein with five transmembrane and three extracellular domains [1,2]. PROM1 is a member of a conserved family of proteins involved in modulating the architecture of cellular protrusions, such as microvilli and cilia [4–6]. Mutations in the human Prom1 gene have been reported in various types of retinal degeneration and are primarily associated with cone-rod dystrophy [7–11]. Mouse models lacking *Prom1* or expressing the dominant Arg373Cys mutant recapitulate the retinal degeneration phenotype and display defects in disk morphogenesis [9,12,13]. Most mutations in *Prom1* are recessive and result in loss of function due to premature stop codons. Notably, three missense mutations, Leu245Pro, Arg373Cys, and Asp829Asn, have dominant inheritance patterns [7].

The *Prom1* gene produces multiple splicing isoforms that can be tissue and cell type specific [15–19]. Six alternative exons in the mouse *Prom1* can potentially produce 24 splice variants, although to date, only eight have been enumerated [15,16,20]. In mouse photoreceptor cells, a microexon, exon 19, introduces 6 amino acids in the photoreceptor-specific SV8 isoform of *Prom1* [16,17]. While this exon is present in most vertebrate clades, it is not used in the primates, including humans, due to mutations that disrupt either the 3' or the 5' splice site of the exon (Figure 2B).

Despite playing conserved and critical functions, the Prominin-1 genes show relatively low conservation of their primary amino acid sequence. For example, human and mouse PROM1 proteins share approximately 61% sequence identity [21]. Consequently, antibodies to PROM1 tend to be species specific. In mice, the rat monoclonal antibody mAB 13A4 is the reagent typically used to detect PROM1. The mAB 13A4 antibody was raised against an extract from mouse neuroepithelium [2]. Its antigen was cloned by screening for reactivity with mAb 13A4 of a phage

library of mouse kidney cDNA [2]. mAB 13A4 is speculated to recognize part of the third extracellular domain of PROM1 because of a truncating mutation in that domain that abolishes its binding, but the epitope was never mapped [8,22].

Prompted by a discrepancy between the PROM1 protein levels measured in the postnatal retina by mAB 13A4 and the mouse monoclonal antibody to PROM1 ab27699, we set out to map the epitope of mAB 13A4. We found that mAB 13A4 recognizes a structural epitope that can be disrupted by deletions in the second and third extracellular domains. Furthermore, the affinity of mAB 13A4 to PROM1 is dramatically reduced by the inclusion of the photoreceptor-specific exon 19a.

3.4 Materials and Methods

Animals

The *Prom1*^{rd19} mice were acquired from the Jackson laboratory (B6. BXD83-Prom1^{rd19}/Boc, Stock No: 026803). All experiments were conducted with the approval of the Institutional Animal Care and Use Committee at West Virginia University.

Prom1 clones

We obtained a full-length Mammalian Gene Collection clone of the mouse photoreceptor-specific isoform (SV8) of *Prom1* from Horizon Discovery (clone ID: 4502359, NCBI accession: BC028286). Gibson assembly (NEB# E5510) was used to generate Flag-tagged *Prom1* clones and deletion mutants in pcDNA3.1+ (Invitrogen). The primers used for cloning are listed in Online Supplementary Table 1.

Cell culture and transfection:

Mouse Neuro-2a (N2a) neuroblastoma cells (ATCC CCL-131) were cultured in OptiMEM reduced serum media buffered with sodium bicarbonate and supplemented with 4% (v/v) fetal bovine serum (FBS, R&D Systems, Minneapolis, MN). The cells were grown at 37 ◦C in a 5%

CO2 humidified atmosphere. The cDNA clones were transiently transfected in N2a cells using polyethylenimine [23]. Cell lysates for Western blot analysis were collected at 24 hours post-transfection.

Denaturing Gel Electrophoresis and Western Blot

Flash-frozen mouse retinas and N2a cells transiently transfected with Prom1-expressing constructs were lysed in RIPA buffer (50 mM Tris HCl-PH 8.0, 150 mM NaCl, 1.0% TritonX-100, 0.5% Sodium Deoxycholate, 0.1% SDS) supplied with protease (Sigma-Aldrich catalog# 535140-1ML) and phosphatase inhibitors cocktail (Sigma-Aldrich catalog # P5726-1 ML). After homogenization, the lysate was incubated on ice for 10 mins, then cleared by centrifugation for 15 mins. 20 µg of protein extract was resolved in 4-20% polyacrylamide SDS–PAGE gel and transferred onto polyvinylidene difluoride (PVDF) membranes (Immunobilon-FL, Millipore). After blocking with BSA in PBST (Phosphate- buffered saline with 0.1% Tween-20), the membranes were probed with primary antibodies overnight at 4 °C, followed by incubation with fluorescently labeled (Alexa Fluor 647 or 488, Jackson ImmnuoResearch) secondary antibodies for 1 hour at room temperature. The membranes were then imaged on Amersham Typhoon Phosphorimager (GE Healthcare).

Serial dilution was performed to ascertain the linearity of western blot quantification (S1 Fig). Lysates of cells transfected with the s8(-ex19) clone were diluted with extracts from cells transfected with an empty vector to maintain equal loading. The lysates were then probed on western blot by mAB 13A4 and imaged as described above. Linear regression was performed in Graphpad Prism on the scaled and normalized band intensities.

Blue Native Polyacrylamide Gel Electrophoresis (BN PAGE)

The samples were lysed in BN PAGE sample buffer (Thermo Fisher Catalog# BN2008) containing 1% digitonin and protease inhibitors, following the manufacturer's recommendation. The lysates were treated with benzonase at room temperature for 30 minutes to shear the DNA and cleared by centrifugation for 15 mins. Prior to electrophoresis, the samples were mixed with

Coomassie G-250 and resolved in 3-12% NativePAGE Bis-Tris gel (Invitrogen Catalog #BN1003BOX) as per the manufacturer's recommendations. The gels were then transferred on PVDF membranes (Immunobilon-FL, Millipore) following the manufacturer's recommendations. After transfer, the membranes were incubated in 8% acetic acid for 15 minutes to fix the proteins, rinsed with deionized water, and air-dried. The membranes were blocked, probed with antibodies, and imaged as described above in the denaturing gel electrophoresis protocol.

Antibodies

The primary and secondary antibodies that were used throughout our studies include the following: mouse anti-β-tubulin (1:5000; catalog # T8328-200ul, Sigma-Aldrich, St. Louis, MO), mouse anti-GAPDH (1:20,000; custom made), mouse anti-flag M2 (1:1000, catalog # F1804-200UG, Sigma-Aldrich, St. Louis, MO), rat anti- Prom1(1:1000, clone ID:13A4, ThermoFisher, Waltham, MA), mouse anti-GFP HRP conjugated GFP tag (1:1000, Cat #HRP-66002, ThermoFisher, Waltham, MA), mouse anti-Prom1 (1:1000, Cat #ab27699, Abcam, Cambridge, MA), Alexa Fluor 647 conjugated AffiniPure Goat-Anti rabbit IgG (1:3000, Jackson ImmunoReserach, West Grove, PA), Alexa Fluor 488 conjugated AffiniPure Goat-Anti mouse IgG (1:3000, Jackson ImmunoReserach, West Grove, PA), and Alexa Fluor 647 conjugated AffiniPure Goat-Anti rat IgG (1:3000, Jackson ImmunoReserach, West Grove, PA).

Statistical analysis

The two-way analysis of variance (ANOVA) with Tukey Honest Significant Differences (HSD) post-hoc test or two-tailed unpaired Student's T-test was used to determine statistical significance as indicated in the results section. Quantitative data are presented as the mean of three biological replicates ±standard error of the mean.

Protein structure prediction and visualization

The RobeTTa structure prediction service was used to create models of the photoreceptor-specific SV8 (RefSeq NP_001157057) isoform that contains exon 19a and the ubiquitously expressed isoform SV2 (RefSeq NP_001157049) that lacks exon 19a (see Figure 2A for

alignment of the two sequences). To create images of the structures, we used Pymol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

3.5 Results

Discrepancy in the PROM1 protein levels measured by the mAB 13A4 and ab27699 antibodies

While investigating the expression of PROM1 in the postnatal mouse retina, we noticed that when measured by mAB 13A4, the PROM1 protein levels peaked at postnatal day 8, five days before the peak recorded by ab27699 (Figure 1A and B). Furthermore, mAB 13A4 showed approximately three to five-fold lower levels of PROM1 at postnatal days 13 and beyond compared to ab27699 (Figure 1A and B). Two-way ANOVA showed a significant effect of the day after birth ($F(9)=55.26$, p-value$<2*10^{-16}$) and the antibody used ($F(1)=365.48$, p-value$<2*10^{-16}$) on the measured PROM1 protein levels. According to the Tukey HSD post-hoc test, the signals detected by mAB 13A4 and ab27699 were significantly different starting from postnatal day 13. It is possible that cross-reactivity of ab27699 to other proteins of a size similar to PROM1 in the retina could have compromised its performance. To rule out cross-reactivity, we probed retinal extracts from wild-type and homozygous *Prom1*^{rd19} mutant retina with mAB13A4 and ab27699. The *Prom1*^{rd19} allele contains a premature stop codon that abolishes the expression of the PROM1 protein [24]. Both antibodies recognized a protein just over 100KDa in size in the wild-type retina that was not detected in the extract from the *Prom1* knockout retina (Figure 1C). Thus, the discrepancy in the signal between mAB13A4 and ab27699 is likely due to differences in the availability of the epitopes recognized by the two antibodies.

Reduced affinity of mAB13A4 to PROM1 carrying the photoreceptor-specific exon 19a

A short 18nt microexon, exon 19a, is included in the *Prom1* transcripts, specifically in photoreceptor cells (Figure 2) [17]. The inclusion rate of exon 19a starts to increase at postnatal

day 3, and the exon 19a containing transcripts become dominant in the retina after postnatal day 8 [17]. The peptide encoded by exon 19a is inserted in the third extracellular loop of PROM1 (Figure 3A), which is also the proposed location of the mAB 13A4 epitope [8,22]. Thus it was possible that inclusion of exon 19a disrupts the epitope of mAB 13A4 while leaving intact the epitope of ab27699. To determine if exon 19a disrupts the mAB 13A4 epitope, we generated Flag-tagged cDNA clones that either contain (SV8) or skip (SV8(-Ex19a)) the photoreceptor-specific exon 19a. The cDNA were transfected in N2a cells, and their expression was probed with mAB 13A4, ab27699, and anti-Flag antibodies. As expected, mAB 13A4 showed a significant reduction in its affinity to the cDNA containing exon 19a compared to the epitope tag (Figure 3B). The inclusion of exon 19a had no effect on the affinity of the ab27699.

mAB 13A4 recognizes a structural epitope

To map the epitope of mAB13A4, we created a series of tiled deletion mutants of the SV8(-Ex19a) cDNA that originated at the point where exon 19a would have been inserted and progressed in the C-terminal and N-terminal direction (Figure 4A). The deletion mutants covered 108 amino acids of sequence. Nine out of ten deletions resulted in complete loss of the mAB 13A4 epitope (Figure 4B). Only the most C-terminal deletion in the series (D+4) could be detected by mAB 13A4. The results of the deletion mutagenesis indicate an epitope for mAB 13A4 that is at least 94 amino acids in length. This length far exceeds the size range of linear peptide epitopes and strongly argues for a structural rather than a linear epitope [25,26].

The western blots shown on Figures 1, 3, and 4 were performed using denaturing gel electrophoresis. Detecting a structural antigen using this approach would require the protein to renature on the membrane prior to probing with the primary antibody. Consequently, our results may reflect the propensity of the splicing isoforms and deletion mutants to renature rather than the genuine affinity of the mAB 13A4 to the native proteins. To test if this is the case, we resolved the proteins produced by the SV8 and SV8(-Ex19a) clones on native gel electrophoresis and probed the membranes with mAB 13A4 and anti-Flag antibodies. The exon 19a-containing protein

was recognized with significantly reduced affinity (Figure 4C), demonstrating that the inclusion of exon 19a alters the structure of PROM1 rather than affecting its folding rate. Interestingly, under native conditions, PROM1 formed higher order complexes.

Computationally derived PROM1 tertiary structure predicts the effect of sequence manipulation on the mAB13A4 epitope

To better understand the nature of the mAB 13A4 epitope and how our deletion mutagenesis affected it, we needed a tertiary structure for PROM1. There are no empirically derived structures of PROM1. Nevertheless, recent advances in computational approaches for structure prediction are producing remarkably accurate structures [27,28]. We used the RobeTTa structure prediction service to model the structures of the mouse PROM1 isoforms SV8 and SV2 (Figure 2A) [27]. The photoreceptor-specific SV8 isoform differs from canonical SV2 isoforms by the inclusion of exon 19a in the third extracellular domain and the skipping of two exons in the cytoplasmic C-terminal domain. The structures predicted by RobeTTa for the SV8 and SV2 proteins were in excellent agreement with each other and with the structure of the human PROM1 predicted by Alpha Fold [29]. In the predicted PROM1 structure, the second and third extracellular domains each form two antiparallel alpha helices that are continuous with the adjacent transmembrane domains. The alpha helices formed by the second and third extracellular domains and the adjacent transmembrane domains are packed in a four-helix bundle (Figure 5A). Inclusion of exon 19a lengthens the second helix of the third extracellular domain, causing a kink in the bundle (Figure 5A). Mapping the positions of the deletions that disrupted the mAB 13A4 epitope on the PROM1 structure showed that they were located towards the middle portion and the tip of the helical bundle. The D+4 mutation, which was the only one that did not result in loss of the mAB 13A4 epitope, was the furthest from the tip of the helical bundle. Based on the positions of the deleted segments and the six amino acids encoded by exon 19a in the PROM1 structure, we made three predictions: (i) deleting six amino acids adjacent to exon 19a (Del AA 6) in the photoreceptor-specific PROM1 isoform should shorten the helix compensating the inclusion of

exon 19a, and restore the mAB 13A4 epitope (Figure 5B, colored dark blue on the structure of SV8); (ii) Deletion of a 15 amino acid segment (D-7) in the third extracellular domain opposite to D+4, should retain the mAB 13A4 epitope due to its distance from the location of the epitope in the upper half of the helical bundle (Figure 5B, colored red on the structure of SV2); (iii) Deletion of a 15 amino acid segment (Del EC 2) in the upper half of the second extracellular domain of SV8(-Ex19a) should result in loss of the mAB 13A4 epitope (Figure 5B, colored sky blue on the structure of SV2). All three predictions proved to be accurate when the proteins expressed from the corresponding clones were analyzed by Western blot (Figure 5C). Deletion of the six amino acid long segment recovered the mAB 13A4 epitope in the photoreceptor-specific SV8 isoform. Conversely, the deletion in the second extracellular domain of the protein encoded by the SV8(-Ex19a) clone resulted in complete loss of the epitope. Finally, deletion D-7 in SV8(-Ex19a) preserved the epitope, although the affinity of mAB 13A4 was reduced. The results of the structure-directed mutagenesis provide further support for mAB 13A4 recognizing a structural epitope. In addition, our results demonstrate the utility of the modeled PROM1 structures.

3.6 Discussion

To reliably interpret results from techniques that employ antibodies, it is essential to know the antibody epitope, its specificity, and its affinity. While mapping the antibody epitope may be important, it is usually not considered necessary as long as specificity to the target can be demonstrated. Such practice leaves a gap that, in certain cases, can have a significant impact on interpreting experimental results, as we show here for mAB 13A4. mAB 13A4 is widely used to detect the mouse PROM1 protein because of its excellent specificity and the lack of alternatives with comparable performance. As of the time of writing of this article, there are over 300 publications in Google Scholar citing the 13A4 antibody in the context of PROM1. Here we show that the mAB 13A4 antibody recognizes a structural epitope, and its affinity for naturally occurring

PROM1 isoforms can vary dramatically. When used to measure PROM1 levels in the retina, mAB 13A4 underestimated the protein amount by a factor of five compared to ab27699.

To determine the exact mAB1 13A4 epitope unequivocally will require determining the structure of the PROM1 - mAB 13A4 complex, which is beyond the scope of the current work. Nevertheless, we show that mAB 13A4 is a useful reagent for detecting perturbation in PROM1 structure as changes to the PROM1 sequence that were hundreds of amino acids apart abolished the mAB 13A4 epitope. Furthermore, we created three mutations in PROM1 guided by the computational model of its structure. The effect of these mutations on the mAB 13A4 epitope was in line with the predicted structure, providing empirical evidence for its validity. Finally, we demonstrate that under native conditions, PROM1 can form higher order complexes, providing a possible path towards understanding the dominance of Prom1 mutations in patients with cone-rod dystrophy.

3.7 Acknowledgments:

We are grateful to Dr. Maxim Sokolov and Dr. Douglas Kolson for their assistance with native gel electrophoresis.

Funding: This work was supported by the National Institutes of Health 2R01EY025536 (P.S. and V.M.) and bridge funding provided by the West Virginia University Health Sciences Center Office of Research and Graduate Education (P.S.).

Commercial relationships: F. Matalkah, None; S. Rhodes, None; V. Ramamurthy, None; P. Stoilov, None.

3.8 References

1. Miraglia S, Godfrey W, Yin AH, et al. A Novel Five-Transmembrane Hematopoietic Stem Cell Antigen: Isolation, Characterization, and Molecular Cloning. *Blood*. 1997;90(12):5013-5021. doi:10.1182/blood.V90.12.5013

2. Weigmann A, Corbeil D, Hellwig A, Huttner WB. Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells. *Proc Natl Acad Sci U S A*. 1997;94(23):12425-12430.

3. Yin AH, Miraglia S, Zanjani ED, et al. AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood*. 1997;90(12):5002-5012.

4. Jászai J, Thamm K, Karbanová J, et al. Prominins control ciliary length throughout the animal kingdom: New lessons from human prominin-1 and zebrafish prominin-3. *J Biol Chem*. 2020;295(18):6007-6022. doi:10.1074/jbc.RA119.011253

5. Singer D, Thamm K, Zhuang H, et al. Prominin-1 controls stem cell activation by orchestrating ciliary dynamics. *EMBO J*. 2019;38(2):e99845. doi:10.15252/embj.201899845

6. Thamm K, Šimaitė D, Karbanová J, et al. Prominin-1 (CD133) modulates the architecture and dynamics of microvilli. *Traffic*. 2019;20(1):39-60. doi:10.1111/tra.12618

7. Cehajic-Kapetanovic J, Birtel J, McClements ME, et al. Clinical and Molecular Characterization of PROM1-Related Retinal Degeneration. *JAMA Netw Open*. 2019;2(6):e195752. doi:10.1001/jamanetworkopen.2019.5752

8. Maw MA, Corbeil D, Koch J, et al. A frameshift mutation in prominin (mouse)-like 1 causes human retinal degeneration. *Hum Mol Genet*. 2000;9(1):27-34. doi:10.1093/hmg/9.1.27

9. Yang Z, Chen Y, Lillo C, et al. Mutant prominin 1 found in patients with macular degeneration disrupts photoreceptor disk morphogenesis in mice. *J Clin Invest*. 2008;118(8):2908-2916. doi:10.1172/JCI35891

10. Zhang Q, Zulfiqar F, Xiao X, et al. Severe retinitis pigmentosa mapped to 4p15 and associated with a novel mutation in the PROM1 gene. *Hum Genet*. 2007;122(3-4):293-299. doi:10.1007/s00439-007-0395-2

11. Ragi SD, Lima de Carvalho JR, Tanaka AJ, et al. Compound heterozygous novel frameshift variants in the PROM1 gene result in Leber congenital amaurosis. *Cold Spring Harb Mol Case Stud*. 2019;5(6):a004481. doi:10.1101/mcs.a004481

12. Zacchigna S, Oh H, Wilsch-Bräuninger M, et al. Loss of the Cholesterol-Binding Protein Prominin-1/CD133 Causes Disk Dysmorphogenesis and Photoreceptor Degeneration. *J Neurosci*. 2009;29(7):2297-2308. doi:10.1523/JNEUROSCI.2034-08.2009

13. Carr BJ, Stanar P, Moritz OL. Distinct roles for prominin-1 and photoreceptor cadherin in outer segment disc morphogenesis in CRISPR-altered X. laevis. *J Cell Sci*. 2021;134(1):jcs253906. doi:10.1242/jcs.253906

14. Lee W, Paavo M, Zernant J, et al. Modification of the PROM1 disease phenotype by a mutation in ABCA4. *Ophthalmic Genet*. 2019;40(4):369-375. doi:10.1080/13816810.2019.1660382

15. Fargeas CA, Huttner WB, Corbeil D. Nomenclature of prominin-1 (CD133) splice variants – an update. *Tissue Antigens*. 2007;69(6):602-606. doi:10.1111/j.1399-0039.2007.00825.x

16. Kemper K, Tol MJPM, Medema JP. Mouse Tissues Express Multiple Splice Variants of Prominin-1. *PLOS ONE*. 2010;5(8):e12325. doi:10.1371/journal.pone.0012325

17. Murphy D, Cieply B, Carstens R, Ramamurthy V, Stoilov P. The Musashi 1 Controls the Splicing of Photoreceptor-Specific Exons in the Vertebrate Retina. *PLOS Genet*. 2016;12(8):e1006256. doi:10.1371/journal.pgen.1006256

18. Corbeil D, Joester A, Fargeas CA, et al. Expression of distinct splice variants of the stem cell marker prominin-1 (CD133) in glial cells. *Glia*. 2009;57(8):860-874. doi:10.1002/glia.20812

19. Fargeas CA, Joester A, Missol-Kolka E, Hellwig A, Huttner WB, Corbeil D. Identification of novel Prominin-1/CD133 splice variants with alternative C-termini and their expression in epididymis and testis. *J Cell Sci*. 2004;117(18):4301-4311. doi:10.1242/jcs.01315

20. Fargeas CA, Corbeil D, Huttner WB. AC133 Antigen, CD133, Prominin-1, Prominin-2, Etc.: Prominin Family Gene Products in Need of a Rational Nomenclature. *Stem Cells*. 2003;21(4):506-508. doi:10.1634/stemcells.21-4-506

21. Corbeil D, Röper K, Weigmann A, Huttner WB. AC133 Hematopoietic Stem Cell Antigen: Human Homologue of Mouse Kidney Prominin or Distinct Member of a Novel Protein Family? *Blood*. 1998;91(7):2625-2626. doi:10.1182/blood.V91.7.2625

22. Thamm K, Graupner S, Werner C, Huttner WB, Corbeil D. Monoclonal Antibodies 13A4 and AC133 Do Not Recognize the Canine Ortholog of Mouse and Human Stem Cell Antigen Prominin-1 (CD133). *PLoS ONE*. 2016;11(10):e0164079. doi:10.1371/journal.pone.0164079

23. Boussif O, Lezoualc'h F, Zanta MA, et al. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc Natl Acad Sci U S A*. 1995;92(16):7297-7301.

24. Collin GB, Gogna N, Chang B, et al. Mouse Models of Inherited Retinal Degeneration with Photoreceptor Cell Loss. *Cells*. 2020;9(4):931. doi:10.3390/cells9040931

25. Buus S, Rockberg J, Forsström B, Nilsson P, Uhlen M, Schafer-Nielsen C. High-resolution Mapping of Linear Antibody Epitopes Using Ultrahigh-density Peptide Microarrays. *Mol Cell Proteomics MCP*. 2012;11(12):1790-1800. doi:10.1074/mcp.M112.020800

26. Opuni KFM, Al-Majdoub M, Yefremova Y, El-Kased RF, Koy C, Glocker MO. Mass spectrometric epitope mapping. *Mass Spectrom Rev*. 2018;37(2):229-241. doi:10.1002/mas.21516

27. Baek M, DiMaio F, Anishchenko I, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. 2021;373(6557):871-876. doi:10.1126/science.abj8754

28. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2

29. Varadi M, Anyango S, Deshpande M, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res*. 2022;50(D1):D439-D444. doi:10.1093/nar/gkab1061

30. Kemper K, Sprick MR, de Bree M, et al. The AC133 epitope, but not the CD133 protein, is lost upon cancer stem cell differentiation. *Cancer Res*. 2010;70(2):719-729. doi:10.1158/0008-5472.CAN-09-1820

31. Barlow DJ, Edwards MS, Thornton JM. Continuous and discontinuous protein antigenic determinants. *Nature*. 1986;322(6081):747-748. doi:10.1038/322747a0

32. Vita R, Mahajan S, Overton JA, et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res*. 2019;47(D1):D339-D343. doi:10.1093/nar/gky1006

33. Corbeil D, Röper K, Hannah MJ, Hellwig A, Huttner WB. Selective localization of the polytopic membrane protein prominin in microvilli of epithelial cells - a combination of apical sorting and retention in plasma membrane protrusions. *J Cell Sci*. 1999;112 (Pt 7):1023-1033. doi:10.1242/jcs.112.7.1023

3.9 Figures

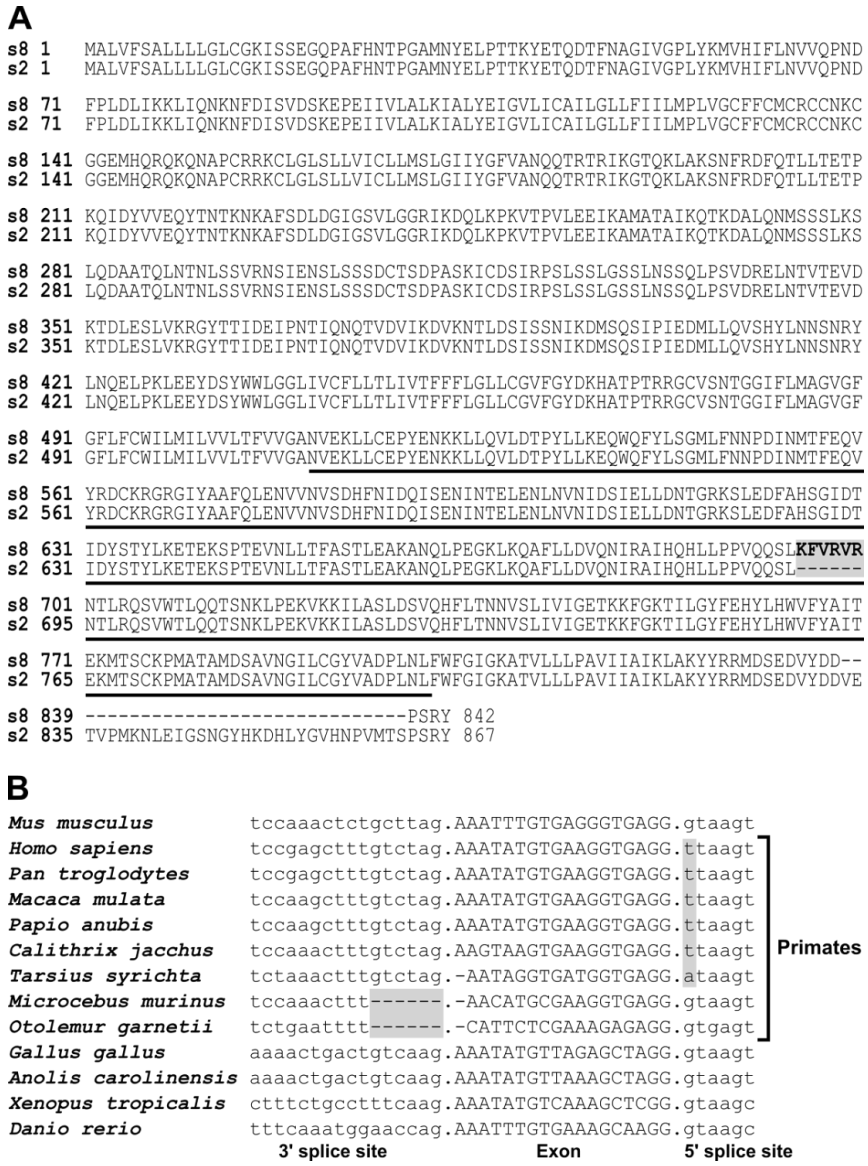


Figure 3-1 Photoreceptor-specific splice variant of PROM1

Figure 1: Photoreceptor-specific splice variant of PROM1. A) Alignment of the photoreceptor-specific s8 isoform of PROM1 (RefSeq NP_001157057) to the ubiquitously expressed isoform s2 (RefSeq NP_001157049). The amino acids encoded by the photoreceptor-specific exon 19 are shown in bold and shaded in gray. The third extracellular domain is underlined. **B)** Alignment of vertebrate exon 19 sequences including the adjacent 3' and 5' splice sites. Mutations inactivating the splice sites in primates are shaded in gray.

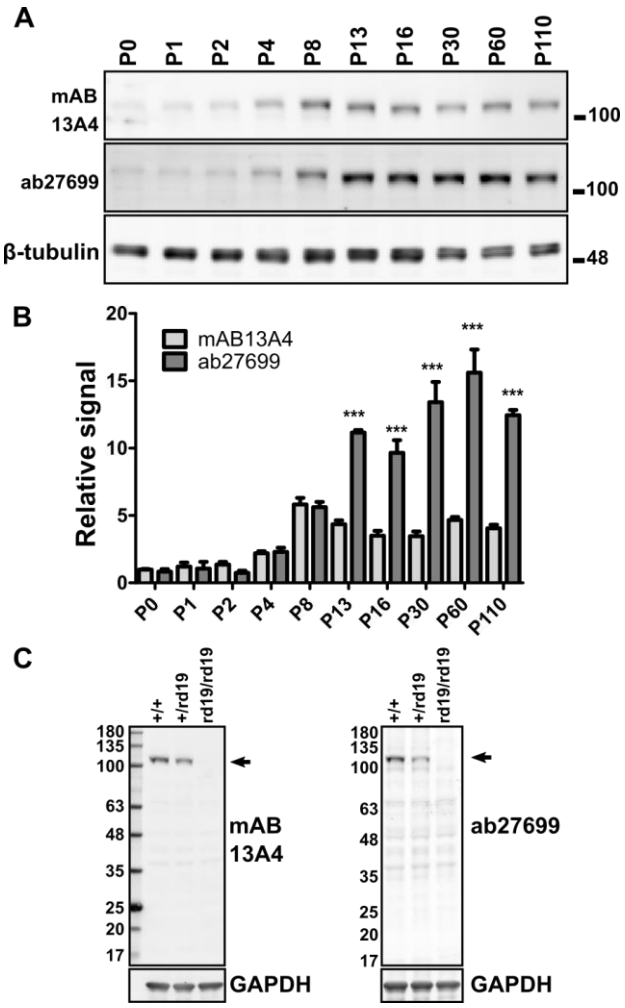


Figure 3-2 Discrepancy in the levels of PROM1 as determined by the mAB 13A4 ab27699 antibodies

Figure 2: **Discrepancy in the levels of PROM1 as determined by the mAB 13A4 ab27699 antibodies**. **A)** Immunoblotting for PROM1 in mouse retina lysate collected between postnatal day 0 (P0) and postnatal day 110 (P110) using the mAB 13A4 and the ab27699 antibodies. Anti-β-tubulin serves as a loading control. **B)** Quantification of Prom1 level in retina lysates using the mAB 13A4 and the ab27699 antibodies. Error bars represent the standard error of the mean (n=3). Two-way ANOVA was used to assess the effect of the postnatal day and the antibody used on the PROM1 signal. The statistical significance of the signal mAB 13A4 compared to ab27699 for each day was calculated by Tukey HSD post-hoc test. Tukey HSD p-values of less than 0.001 from the post-hoc test are indicated by "***. " **C)** Test of the specificity of mAB 13A4 and ab27699 antibodies for detecting PROM1 using the retinal lysate from wild-type mice and Prom1rd19 mutants that do not express PROM1 protein. Arrow indicates the position of the expected PROM1 band.

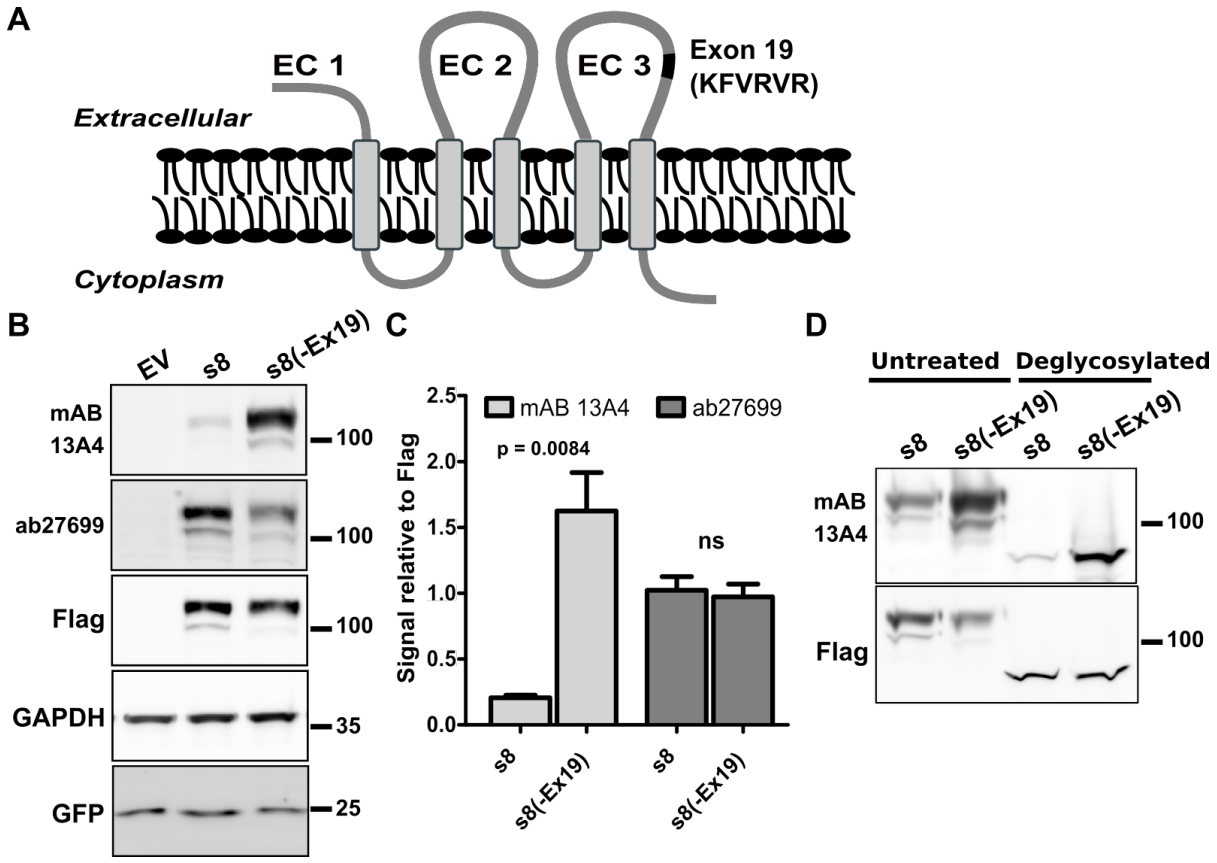


Figure 3-3 The affinity of mAB 13A4 to PROM1 is affected by alternative splicing

Figure 3: The affinity of mAB 13A4 to PROM1 is affected by alternative splicing. **A)** Schematic representation of the PROM1 structure showing the position of exon 19 (black) in the third extracellular domain (EC 3) of the photoreceptor-specific PROM1 isoform s8 (Adapted from Corbeil et al [23]). Extracellular domains one through three are labeled as EC 1, EC 2, and EC 3, respectively**. B)** Western blot of recombinant s8 and s8 lacking exon 19, s8(-Ex19), expressed in N2a cells with mAB 13A4, ab27699, and antibody to the Flag epitope. Transfection with an empty pcDNA3.1 vector (EV) was used as a negative control. All transfections were spiked with a vector expressing GFP to control for transfection efficiency. **C)** Quantification of mAB 13A4 and ab27699 signals relative to the signal from the Flag-tag antibody. Error bars represent the standard error of the mean (n=3). Unpaired Student's t-test was used to assess the statistical significance, and the p-value is indicated on the chart. **D)** Deglycosylated PROM1 isoform s8 is recognized by mAB 13A4 with reduced affinity compared to protein lacking the amino acids encoded by exon 19.

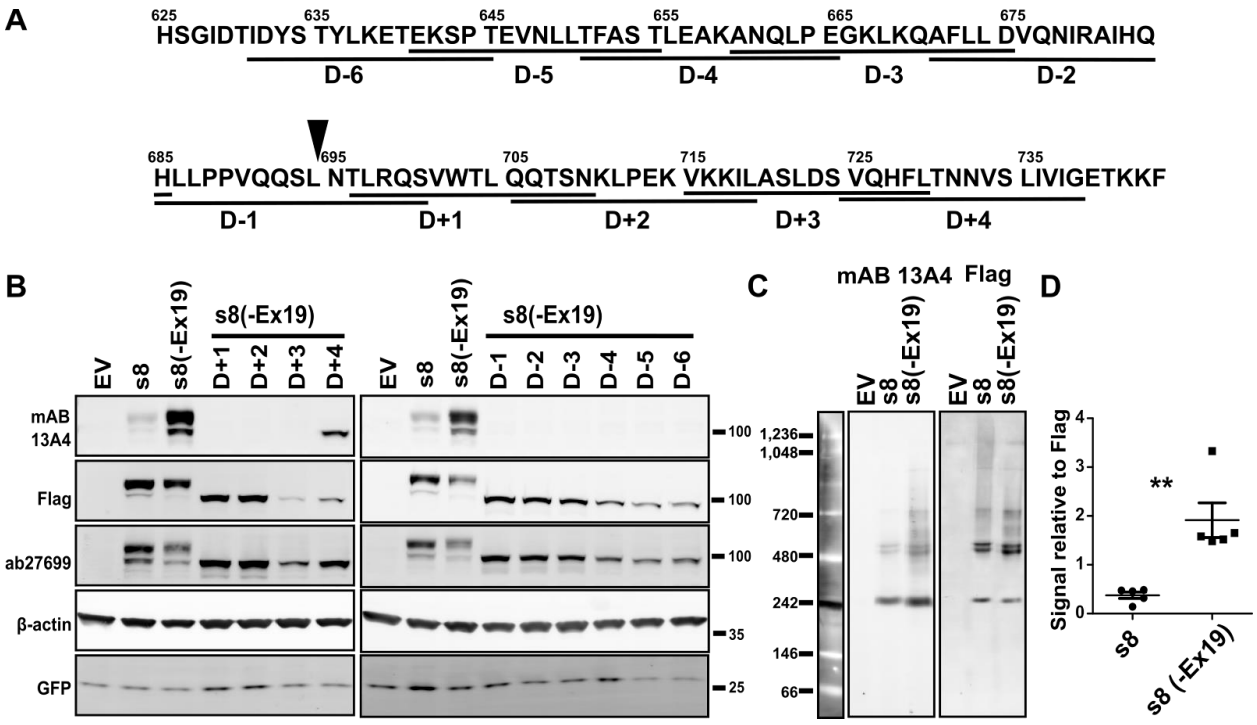


Figure 3-4 Mapping of the mAB 13A4 epitope

Figure 4: Mapping of the mAB 13A4 epitope. **A)** Sequence of the third extracellular domain. Underlining shows the positions of the deletions used for epitope mapping. The deletion variants were generated starting from the s8 clone lacking exon 19, s8(-Ex19). Solid triangle above the sequence shows the position where exon 19 is inserted in the photoreceptor-specific isoform. **B)** Western blot analysis of PROM1 deletion mutants expressed in N2a cells using mAB 13A4 and Flag-tag antibodies. All mutants with the exception of D+4 resulted in the loss of the mAB 13A4 epitope. Transfection with an empty pcDNA3.1 vector (EV) was used as a negative control. All transfections were spiked with a vector expressing GFP to control for transfection efficiency. **C)** Native gel Western blot analysis using mAB 13A4 and Flag-tag antibodies of N2a cell lysate transfected with either an empty vector (EV), s8, or s8(-Ex19). **D)** Quantification of mAB 13A4 signals relative to the signal from the Flag-tag antibody in native gel electrophoresis western blot. Dots represent individual data points. Line and error bars the mean and the standard error of the mean (n=5). Unpaired Student's t-test was used to assess the statistical significance and the p-values were less than 0.01 (indicated by "**").

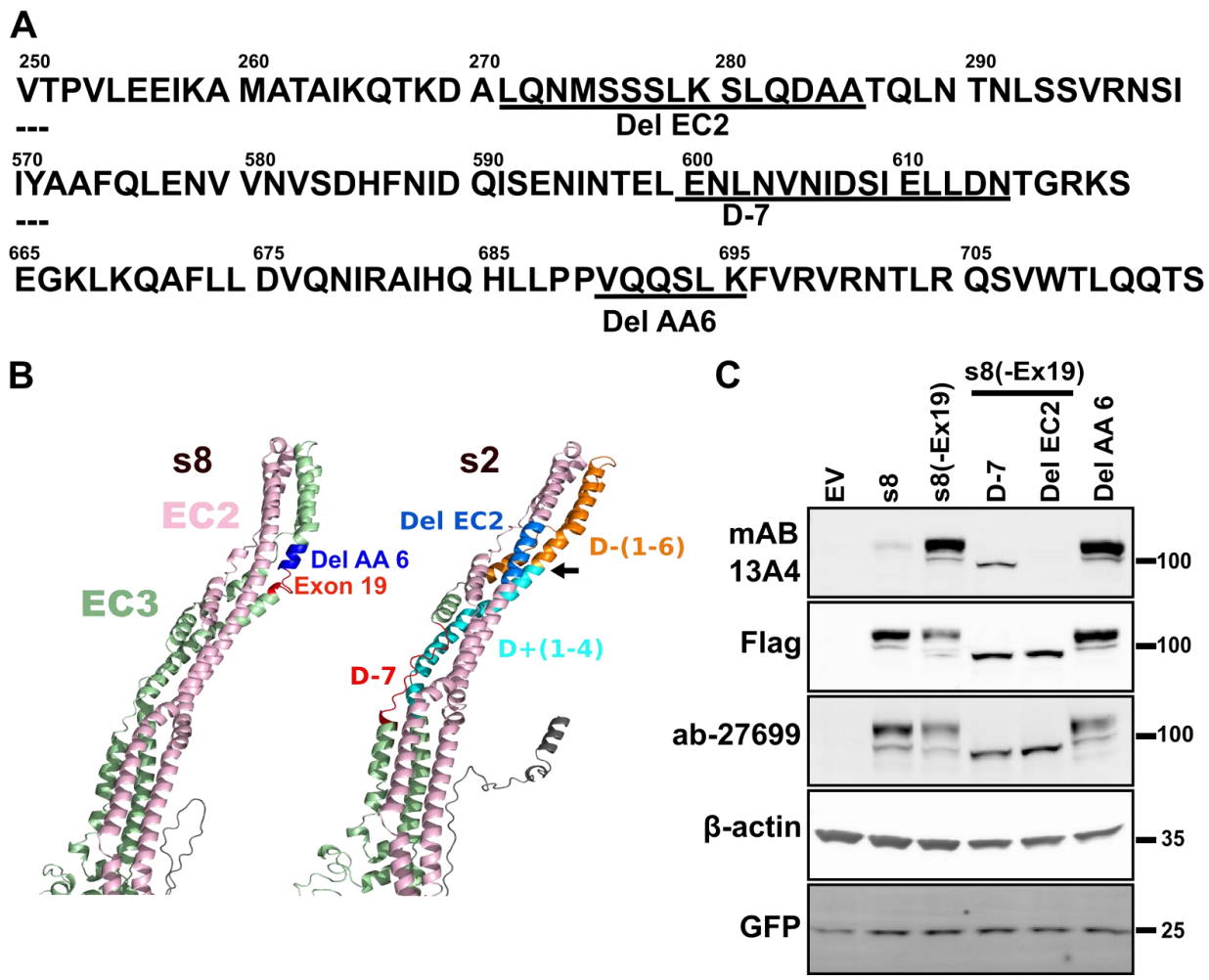
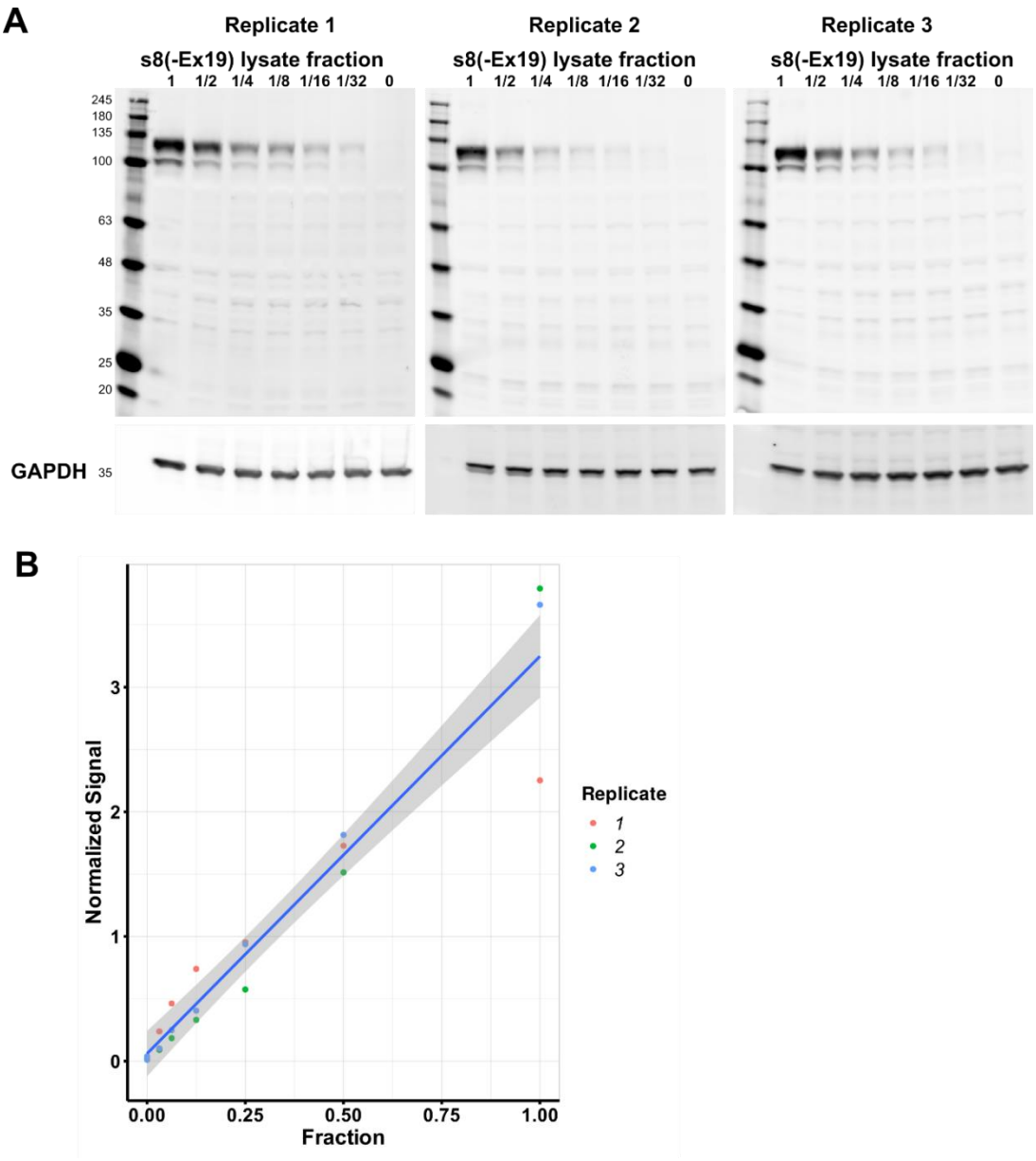


Figure 3-5 Computationally derived tertiary structure of PROM1 predicts the effect of mutations on the mAB 13A4 epitope

Figure 5: Computationally derived tertiary structure of PROM1 predicts the effect of mutations on the mAB 13A4 epitope. A) Sequence segments from the PROM1 protein with the amino acids deleted in clones Del EC2, D-7, and Del AA6 underlined. **B)** Partial structure of PROM1 isoforms s8 and s2 showing the positions of the segments deleted in the experiments on Fig 4 and on panel C of this Fig on the tertiary structure of PROM1. Exon 19 is shown in red on the structure of s8. The deletion Del AA6 analyzed on panel C is shown in blue on the structure of s8. On the structure of s2, cyan color indicates deletions D-1 through D-6, and orange indicates deletions D+1 through D+4 from the experiments shown on Fig 4. Also, on the structure of s2, the positions of the deletions D7 and EC2 analyzed on panel C are shown in red and sky blue, respectively. Arrow points to the position of the excluded exon 19 in s2. **C)** Western blot analysis

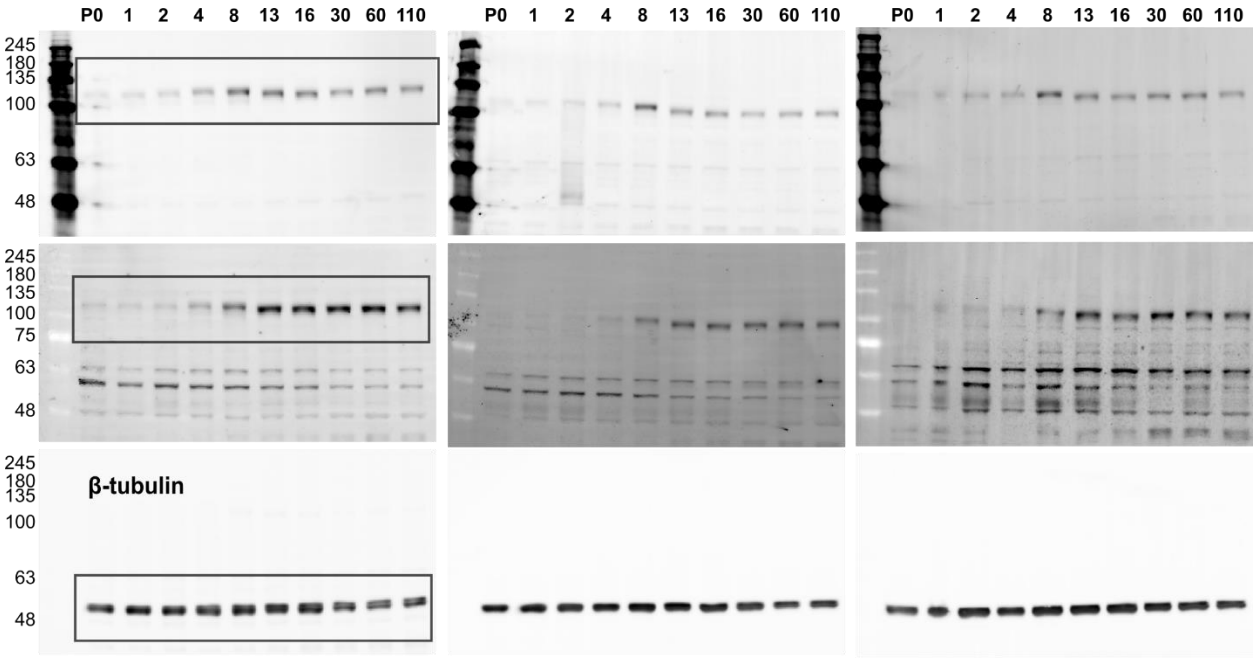
of PROM1 deletion mutants expressed in N2a cells using mAB 13A4 and Flag-tag antibodies. All transfections were spiked with a vector expressing GFP to control for transfection efficiency.

3.10 Supplementary information



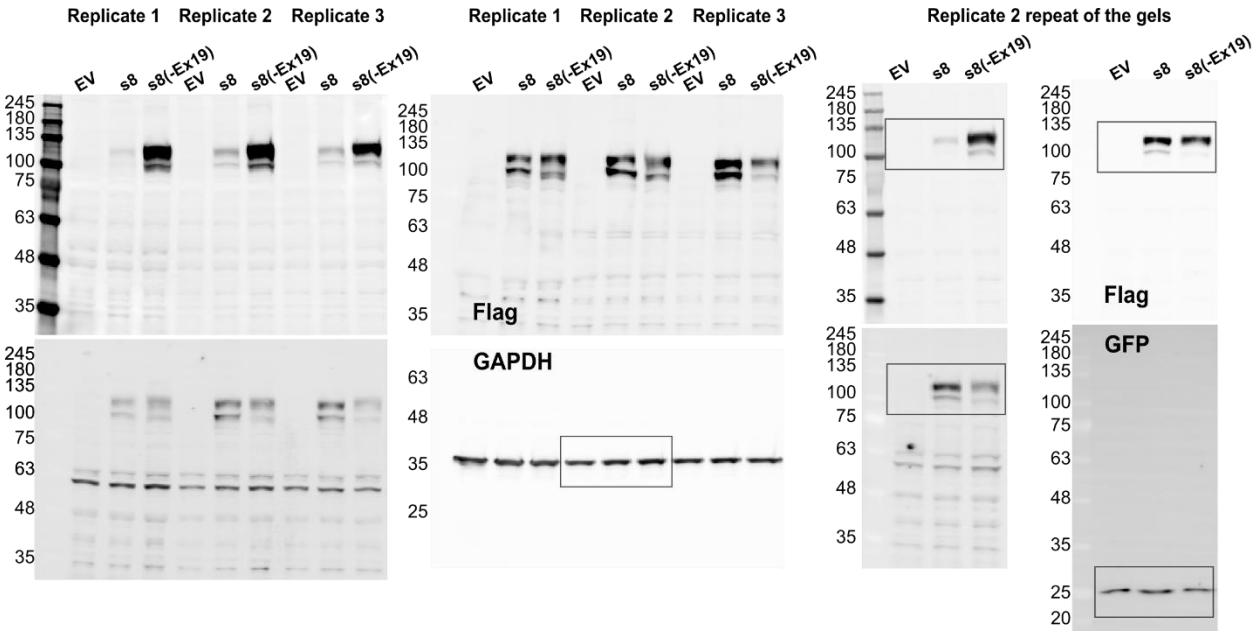
Supplement Figure 3-6 Linearity of western blot quantification

Supplement Figure 1. **Linearity of western blot quantification**. A) Western blot replicates. Serial dilution of extract expressing clone s8(-Ex19) with an extract from cells transfected with an empty vector was probed by mAB 13A4. B) Plot showing scaled normalized signal intensities for each replicate and linear regression with 95% confidence interval. R2=0.93, p-value=1.8*10-12.



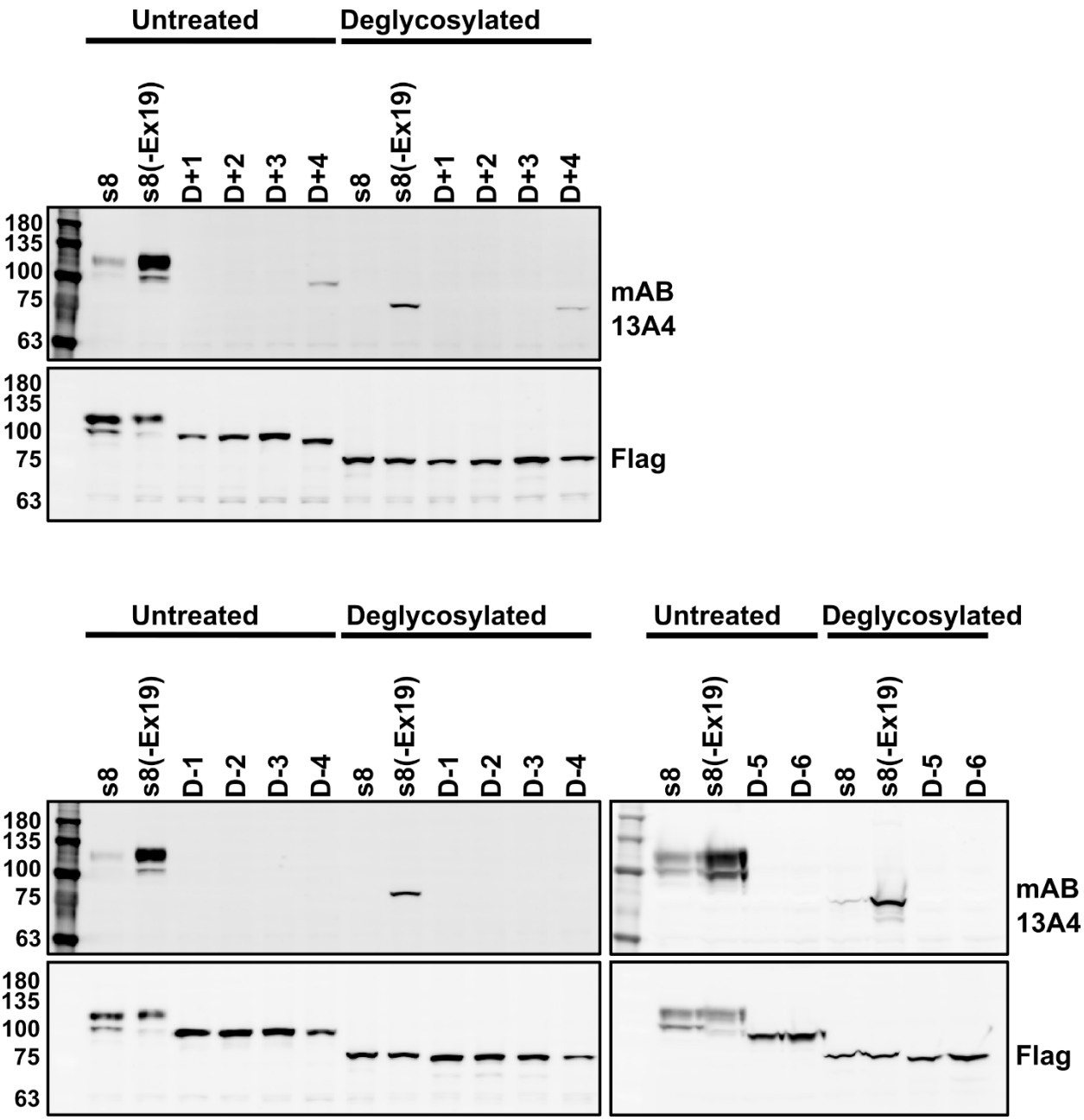
Supplement Figure 3-7 Gel images of replicates related to Figures 1A and 1B

Supplement Figure 2. **Gel images of replicates related to Figures 1A and 1B**. Boxes denote the parts of the images used in preparing Figure 1A.



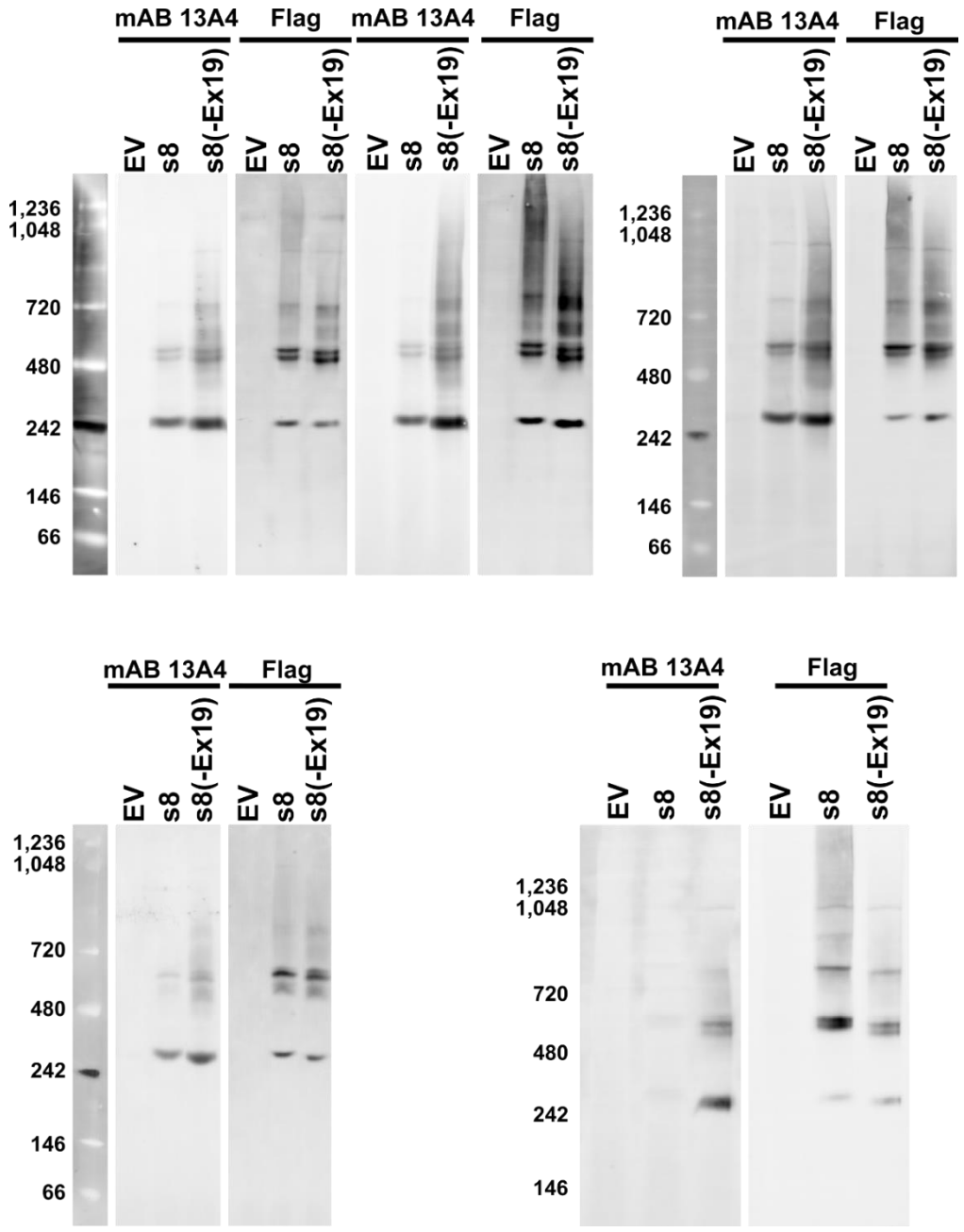
Supplement Figure 3-8 *Gel images of replicates related to Figures 3B and 3C*

Supplement Figure 3. **Gel images of replicates related to Figures 3B and 3C**. Boxes denote the parts of the images used in preparing Figure 3B.



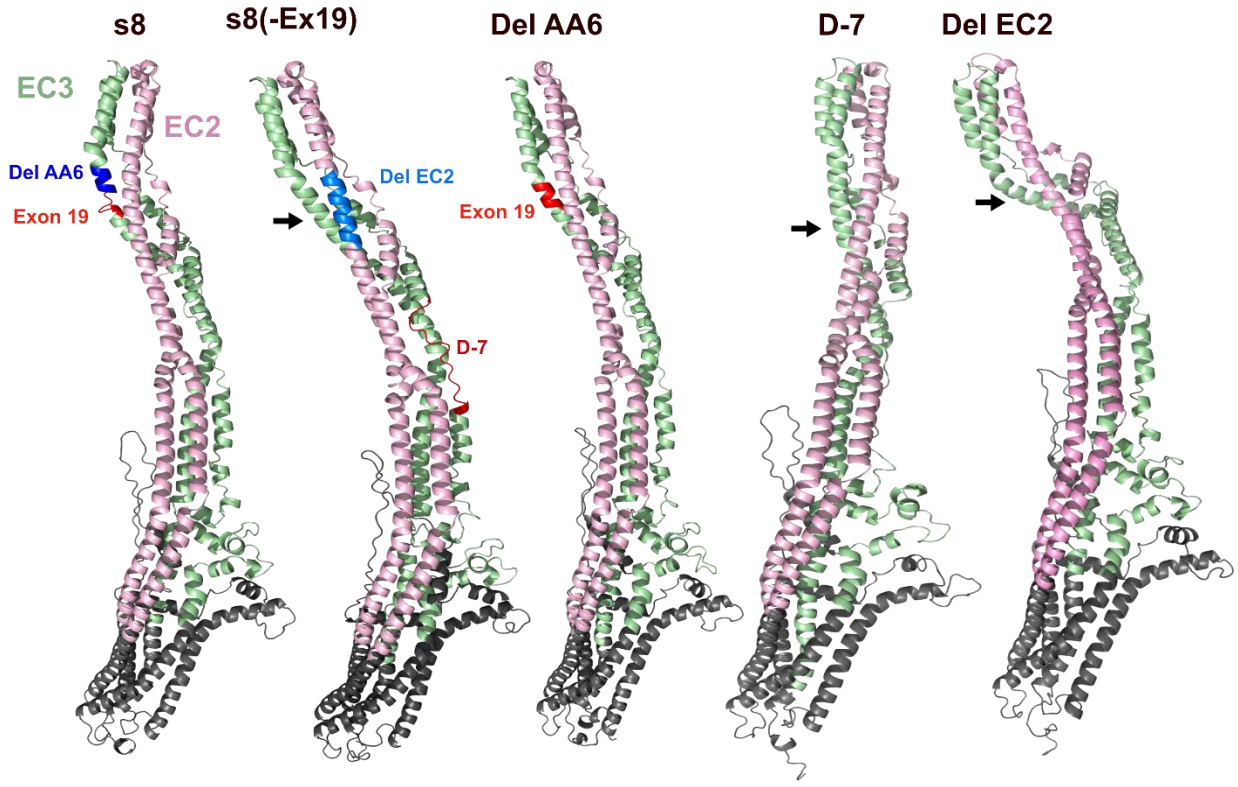
Supplement Figure 3-9 Deglycosylation of PROM1 deletion mutants

Supplement Figure 4. **Deglycosylation of PROM1 deletion mutants**. Lysates from N2a cells were treated with deglycosylation mix II (NEB) and analyzed on western blot next to untreated controls. The blots were probed with mAB 13A4 and anti-Flag antibodies as indicated.



Supplement Figure 3-10 Native gel electrophoresis of PROM1 clones s8 and s8(-Ex19)

Supplement Figure 5. **Native gel electrophoresis of PROM1 clones s8 and s8(-Ex19).** The proteins were resolved by native blue electrophoresis, transferred to PVDF membranes, and probed with mAB 13A4 and anti-Flag antibodies as indicated. Lanes containing the size standard were cut from the membranes after the transfer and stained with Ponceau S. The size standard lanes and probed membranes were imaged on Typhoon Phosphorimager (GE Healthcare).



Supplement Figure 3-11 Protein structure predictions for PROM1 isoform s8 and deletion clones

Supplement Figure 6. **Protein structure predictions for PROM1 isoform s8 and deletion clones**. Extracellular domains 2 (EC2) and 3 (EC3) are indicated on the predicted structures by green and pink color, respectively. Exon 19 and the amino acids deleted in Del AA6 are colored on the structure of the s8 isoform in red and blue, respectively. The deletions for clones Del EC2 and D7 are colored on the structure of s8(-Ex19) in light blue and red, respectively. In structures that do not contain exon 19, arrows indicate the position of the junction between exons 18 and 20. The structures of s8 and Del EC2 have pronounced kinks near the top of the bundle when compared to the structures of s8(-Ex19). Del AA6, and Del EC2.

4. Chapter 4

4.1 Discussion

The Musashi paralogs, MSI1 and MSI2, are expressed in various stem cell populations, consistent with their role in maintaining the stemness status opposing differentiation [48–51]. In contrast, our recently published work shows that the Musashi proteins maintain high protein levels in the adult retina (Figure.1) [52,53,69]. Musashi proteins are also readily detectable in mature neurons in the retina and CNS [52–54,69].

Furthermore, the Musashi proteins have been shown to function redundantly when regulating proliferation in stem cells [25,51,55]. The functional redundancy in stem cells leads to the notion that the two paralogs act similarly in the differentiated tissues. Therefore, numerous studies focused solely on investigating the first member, MSI1 as a model representing the Musashi family, leaving MSI2 uncharacterized [53,54,70,74]. Our recently published work showed that the expression of the Musashi proteins is developmentally regulated (Figure.2). Early high expression of MSI1 coincides with the outer segment development, while MSI2 levels peak in mature photoreceptors [52]. This mode of expression suggests that the two paralogs may serve distinct functions. Thus, we sought to investigate the role of Musashi individually and in combination in mature photoreceptors. This chapter will discuss our findings on the role

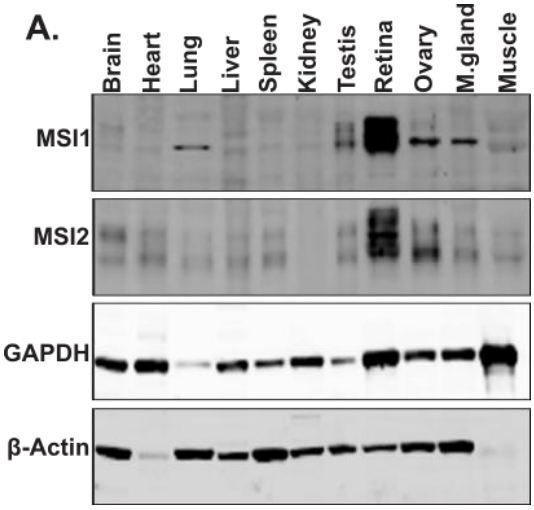


Figure 4-1 The Musashi proteins maintain high expression level in adult retina

Figure1: The Musashi proteins maintain high expression level in adult retina. A) immunoblot showing both MSI1 and MSI2 are highly expressed in adult retina with an exceptionally low to absent expression in the other differentiated mouse tissues. **Published figure in Sundar & Matalkah et al, 2020**.

of the Musashi proteins in mature photoreceptors. I will conclude by going over some suggestions for future direction.

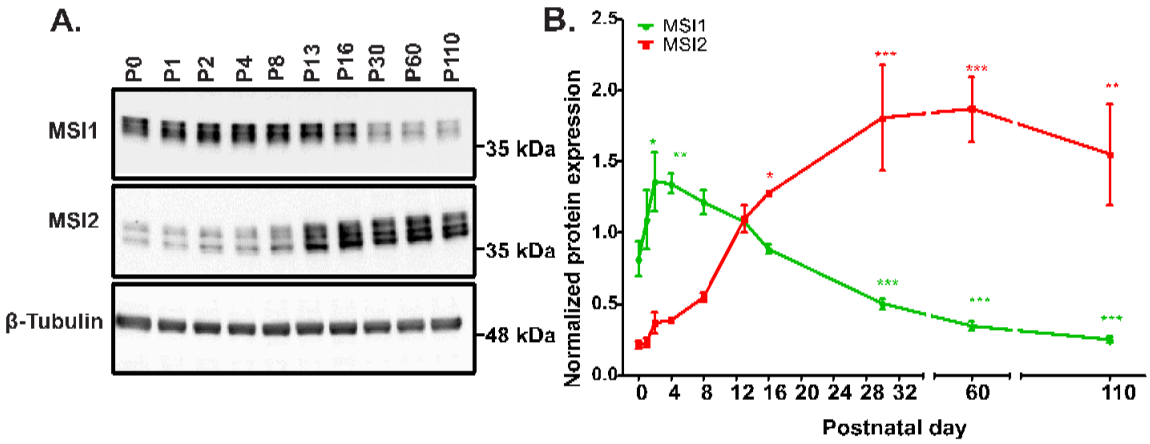


Figure 4-2 The switch in Musashi proteins expression during retinal postnatal development

Figure 2: The switch in Musashi proteins expression during retinal postnatal development. A) immunoblot showing the decrease in MSI1 and increase in MSI2 level as photoreceptors mature. B) Quantification of relative MSI1/2 protein level. Error bars represent standard error of the mean (SEM, n=3). pairwise *t*-test after FDR correction is indicated as follows: $*p < 0.05$; $**p < 0.01$; $**=p < 0.001$. FDR, false discovery. P0 (postnatal day 0). **Published figure in Sundar & Matalkah et al, 2020**.

4.1.1 The Musashi proteins are required for the maintenance of photoreceptor neurons

To investigate the functional role of the MSI1/2 proteins in fully developed photoreceptor neurons, we utilized a tamoxifen inducible *CRE* line under the control of a photoreceptor-specific promoter. The combined deletion of the *Msi1* and *Msi2* resulted in a fast and progressive degeneration of the photoreceptor layer resulting in complete blindness. It took an average of three months after the deletion of the *Msi1/2* to eliminate the whole photoreceptor layer and subsequently abolish the visual responses. Our findings demonstrate a new and significant role for the Musashi proteins in maintaining terminally differentiated cells.

4.1.2 The Musashi proteins function redundantly to maintain the photoreceptor neurons

In our recently published work, we showed that the Musashi expression is developmentally regulated, with MSI1 expression peaking early around postnatal day 4 (PN4) and MSI2 levels rising after PN13 (Figure.2) [52]. Building upon these findings, we hypothesized that MSI2 might have a unique role in maintaining mature photoreceptor cells. To test our hypothesis, we sought to investigate the functional role of the Musashi individually in mature photoreceptors. Contrary to our expectations, the single deletion of *Msi1* or *Msi2* in mature photoreceptors did not produce any functional or morphological phenotype. This suggests that the Musashi proteins are redundant in maintaining the mature photoreceptor neurons. The observed functional overlap can be explained by the high expression of the two proteins in mature photoreceptors and the sequence homology reported between the two paralogs. For example, the RNA binding domains of MSI1 and MSI2 are 77% (RRM1) to 92% (RRM2) identical and recognize the same UAG sequence motif. Thus, the functional redundancy we observed in mature photoreceptor cells recapitulates the previously reported redundancy between MSI1 and MSI2 in other cell types [51,55].

The partial redundancy we reported previously in Sundar et al. is due to the timing of the knockout relative to the timing of the expression of *Msi1* and *Msi2* in photoreceptor development. In Sundar et al., the single deletion of *Msi1* in rod progenitor cells is initiated at embryonic day 9 (E9). In the critical period for photoreceptor development immediately after the birth of the animals, the *Msi1* knockout cannot be fully compensated by *Msi2*, which is expressed at low levels until postnatal day 13 (Figure.2). In contrast, Msi1 can fully compensate for the loss of *Msi2* at this point as it is highly expressed. In the current study, I induced the single deletion in fully developed photoreceptor cells at postnatal day 30, a developmental time point where both *Msi1* and *Msi2* are expressed at high levels.

4.1.3 The Musashi proteins bind intronic and 3'UTR at UAG rich region

To identify the *in vivo* targets of the Musashi proteins in the retina, we performed UV Cross-Linking and Immuno-Precipitation, followed by a high throughput sequencing of the associated RNA fragments (CLIP-Seq). The CLIP-seq is a technique that identifies the direct targets and the exact Musashi-binding sites on the mRNA features. Given the functional redundancy we observed between MSI1 and MSI2 in the context of mature photoreceptor neurons, we presumed that both proteins would bind the same RNA targets in the retina. Thus, we performed CLIP-Seq only for MSI1.

Our CLIP-Seq experiment shows that the Musashi bind a UAG-containing motif found at the 3'-UTR (59.7% targets bound by MSI1) and the intronic region (32.7%) of a broad set of retinal transcripts. The sequence motif bound *in vivo* by MSI1 matches the UAG core motif previously identified by the *in vitro* structural studies [59,60,62]. This match gives us confidence that most binding sites we identified are directly bound by *Msi1* and are not experimental artifacts.

The distribution of *Msi1* binding to introns and 3'-UTRs supports a dual function for the proteins in pre-mRNA processing and regulation of protein expression. These functions are consistent with the nuclear and cytoplasmic subcellular localization reported for the Musashi proteins in the context of mature photoreceptor neurons and the two roles reported for the Musashi in regulating splicing and translation [53,54,66,69]. Interestingly, most of the MSI1 binding events at the 3'-UTR of transcripts we identified belong to genes encoding proteins critical for photoreceptor structure and function. Among the identified targets is a list of proteins associated with phototransduction and outer segment function.

4.1.4 The Musashi proteins activate alternative exons when bound downstream of them

Previously, we showed that MSI1 is critical for the utilization of photoreceptor-specific exons [52,66]. Our earlier work suggests the binding of MSI1 to the intronic region downstream of the alternatively spliced exon leads to its inclusion in the final transcript [66,68]. To determine if this mode of regulation is standard *in vivo*, we combined the binding site of MSI1 identified by the CLIP-seq and the Musashi-dependent alternative exons identified by the RNA-seq. While the CLIP-seq identifies the physically bound RNA targets, the mRNA splicing following *Msi1/Msi2* depletion can be assessed by RNA-seq. The RNA-seq was done on RNA I isolated from the retina collected from the WT and KO mice at D21 post-tamoxifen injection. We selected this time point to ensure enough time was given to capture any changes in the transcriptome induced by the depletion of *Msi1/2* but before any detectable changes occurred on the viability of photoreceptor cells. We identified 280 genes switching isoforms due to *Msi1/Msi2* depletion. Merging the two data sets collected from CLIP-seq and RNA-seq generates an RNA binding protein splicing map (RBP-map) that provides information about the specific binding position on the mRNA features and its effect on splicing outcome as described previously [17]. The generated RBP-map shows that the binding activity of MSI1 occurs at the UAG core motif located in the intronic region downstream of the alternatively regulated exons, resulting in its inclusion in the final transcript. For example, exon 19 in the *Prom1* gene, that's normally included in the final transcript, was excluded upon *Msi1/Msi2* deletion.

4.1.5 The microexons within the *Ttc8, Cep290, Cc2d2a* and *Cacna2d4* genes are dispensable

Given the *in vivo* evidence of MSI activity in controlling alternative splicing in mature photoreceptor cells, we sought to investigate the biological significance of alternative splicing on the function of photoreceptor cells. Among the predicted splicing targets of MSI that we selected to study include the microexon 2A in *Ttc8* encoding for 10 residues (EPAPDLPVSQ), exon 32 of the *Cc2d2a* gene (MSDMLKK), exon 8 of the *Cep290* gene (NKRLKKK), and the microexon 34 in *Cacna2d4* gene encoding for 8 amino acids (AKSVFHHS). All these microexons are photoreceptor-specific and deeply conserved in vertebrates spanning a 230 million years of evolution from lizards to humans. Our eCLIP-seq results show that the inclusion of these microexon is due to the MSI binding to the intronic region downstream of the alternatively spliced exon. Surprisingly, neither the single removal of the micorexons (*Ttc8, Cc2d2a, Cep290, Cacna2d4*) nor the combined triple exon KO (*Ttc8, Cc2d2a, Cep290*) produced a phenotype up to one year of age. The absence of phenotype in our exon knockout animals raises questions about the strong natural selection enforcing the conservation of such functionally neutral exons. One explanation for the dilemma is that the conservation of these microexons is probably due to purifying selection, where only deleterious mutations are eliminated, which subsequently led to their deep conservation

Regardless, one should keep in mind that the lack of a phenotype from the exon knockout model cannot completely rule out a role for splicing in shaping the phenotype of the *Msi* knockouts. What is interesting about these microexons is that they tend to spike as the cells within the retina exit the mitotic cycle and differentiate, at the time of synaptic formation and function, suggesting a role in the late maturation stages [66]. Therefore, the absence of the phenotype could possibly be due to the lack of a cumulative effect over the animal's lifespan that's needed for the manifestation of the phenotype. Currently, our data suggest that the loss of function upon the

depletion of *Msi* in mature photoreceptors is probably imposed by the role of *Msi* in regulating translation rather than its role in regulating alternative splicing.

4.1.6 The Musashi proteins promote the expression of a large number of photoreceptor-specific proteins.

Since 3'UTRs are abundantly bound by the Musashi proteins and are commonly used to regulate the turnover and translation of mRNA, I sought to investigate if the Musashi is involved in regulating protein synthesis [25,36,56,59,75]. To accomplish this, I turned to proteomics to quantify the steady state level of the retinal proteome. We performed global proteome quantification using the isobaric tag labeling and tandem mass spectrometry (TMT). The TMT isobaric tagging approach enables precise and reproducible quantification of the relative protein abundances in KO relative to their littermate control. In the experimental design, we determined the proteomics changes in the retina isolated from the KO and the littermate control collected on day 21 post-tamoxifen injection. The retinal proteome captured by the mass spectrometry was over 7000 proteins in total. We identified the differentially expressed proteins as the ones having a logFC (logarithm of fold change) >1.5 or <−1.5 and false discovery rate (FDR) ≤0.05. It should be noted that these criteria selected 165 proteins that were differentially expressed in KO samples relative to their WT littermate control. Upon closer examination, the data show that the differentially expressed proteins can be divided into two sets: one set shows a significant reduction upon the *Msi1/Msi2* deletion. In contrast, the other set includes significantly upregulated proteins. Guided by the gene ontology analysis, we show that part of the downregulated proteins is involved in the phototransduction pathway (GNB1, GNAT1, GNAT2, CNGA1, PDE6B, PDE6A, and GUCA1B). We also observed a significant decrease in specific ciliary and outer segment proteins. Some are essential for the trafficking in and out of the cilia (TTC8 and RPGR), while others are structural components of the photoreceptors' outer segment (ROM1 and PRCD).

As per the upregulated proteins, the gene ontology analysis suggests that these are associated with the activation of Muller and Microglia. Among the glial markers that we observed to be upregulated include the (GFAP, A2M, and GS). The fact that the *Musashi* deletion was specific to photoreceptor neurons strongly suggests that the upregulation that we observe in glial markers is not directly regulated by Musashi but is a response to changes in photoreceptor cells caused by the *Musashi* knockout. Glial activation is a common defense mechanism activated in response to neurodegeneration. Regardless of the origin of the insults, whether mechanical, infection, or genetics, glial activation is believed to be elicited by the stressed photoreceptors cell to trigger the secretion of neuroprotective factors by the glial cell [76–78]. This mechanism has been reported in several retinal dystrophies models and is not unique for the *Musashi* loss.

Collectively, using proteomics, we identified a trend toward a significant reduction in the protein levels in response to the *Msi1/Msi2* deletions; however, this was not globally followed, as many other proteins were not impacted by the Musashi removal, indicating that the depletion of *Musashi* does not have a generalized effect. This eliminates the possibility that the reduction in protein expression we observed was due to photoreceptors degeneration. This notion is further backed by the normal physiological and morphological analysis of the knockout animals at the time the experiments were carried out (e.g., day 21 post injection). Furthermore, the fact that the CLIP-seq also identified all of the downregulated proteins measured by the mass spectrometry to be directly bound by Musashi to their mRNA 3'UTRs supports a model involving direct regulation.

To determine if the changes we see in the protein level are associated with changes in the transcript levels, we measured the steady-state mRNA abundance in the retina isolated from the KO and WT collected on day 21 post-injection (D21) using RNA-seq. It is essential to mention that RNA-seq only reports the steady state RNA levels and cannot determine if it is due to translational or stability changes. Other techniques are needed to separate these changes, including ribosome profiling and stability assay. Interestingly, RNA-seq did not detect an altered

mRNA level of the genes identified by CLIP-seq with a Musashi-binding site at their 3'UTR. The unchanged mRNA level between KO and WT suggests that the binding of Msi1 to the 3'-UTR of these transcripts does not directly impact their transcription or stability but leaves another possibility of posttranscriptional regulation imposed by the Musashi, such as regulating translation. Indeed, the Musashi proteins have emerged as master regulators of translation. The proposed models for the Musashi imposed regulation over translation involve the direct binding of the Musashi to UAG elements at the 3'UTR of its target transcripts, mediated by the two RNA recognition motifs (RRM1 and RRM2), and the concurrent binding via a region within its C-terminal domain to the translational machinery [36,64]. As mentioned previously, the Musashi proteins can use the same region within its C-terminal to bind different factors; thereby, the type of factor it binds, guided by the cellular context, reflects whether the translation is regulated positively or negatively. The exact molecular mechanism underlying the Musashi role in regulating translation and the nature of any cofactors involved is unknown.

It is important to note that the transcriptome and proteomics data reflect upon the changes in the whole retina while the *Msi1/Msi2* depletion was specific to the photoreceptor neurons. To corroborate our findings, we used the publicly available single cell RNA-seq (scRNA-seq) to track changes in the transcripts associated with the photoreceptor cells only. Combining the transcriptome and proteomics together further confirms a trend toward a reduction in the level of the proteins associated with the photoreceptor neurons with no significant changes in their mRNA level, except for two photoreceptor-specific proteins (PROM1 and IMPG2), where we observed a considerable increase rather than a decrease in their protein level with no changes in their transcripts level as well. Interestingly, both PROM1 and IMPG2 are identified by the CLIP-seq among the genes that harbor the Musashi-binding site at their 3'UTR.

Taken together, the trend we observed toward a significant reduction in the protein levels imposed by the *Msi1/Msi2* deletion supports a model in which the Musashi proteins act as post-

translational activators for protein expression in mature photoreceptors. Thus, we hypothesize that the high expression of Musashi proteins in the mature photoreceptors is evolutionarily meant to keep up with the increased demand for protein synthesis required to maintain the constant regeneration of the photoreceptor's outer segment. Importantly, failure to replenish the rapid turnover of the OS while maintaining a constant length is a characteristic of numerous retinal degenerative disease models.

4.1.7 The splicing alteration imposed by the Musashi disrupted a conformational epitope within PROM1 recognized by the mAB 13A4.

The mouse *Prom1* gene contains at least 27 exons, and its transcript is subjected to alternative splicing producing multiple splice variants [79,80]. Interestingly, it has been shown that these splice variants are expressed in a tissue-specific manner, are developmentally regulated, and may undergo tissue-specific glycosylation [79]. In parallel to the splicing regulation imposed by the Musashi over *Prom1* transcripts identified by eCLIP-seq, our mass spectrometry also identified an upregulation in the protein level of Prom1 upon the *Msi1/Msi2* specific deletion in photoreceptor neurons. Interestingly, as we attempted to confirm the increased protein level on a Western blot, we identified a significant discrepancy between the Prom1 level measured using the widely used mAB 13A4 and the mouse monoclonal antibody to Prom1 ab27699. Hence comes the work presented in chapter 3 of this study, where I show that alternative splicing significantly impacts the reactivity of the widely used mAB 13A4 to detect Prom1 on Western blot. Our results suggest that the mAB 13A4 recognizes a conformational epitope predicted to be composed of a helix bundle consisting of the two large extracellular loops of Prom1. The inclusion of the alternatively spliced Ex19 is predicated on producing a kink in the three-dimensional structure of the helix bundle, disrupting the interaction between the two loops and decreasing the mAB 13A4 reactivity towards Prom1 on a Western blot. This work aimed to bring awareness and careful attention when choosing a specific antibody as a tool for detecting the protein of interest. Keeping

in mind, If the protein of interest is subjected to alternative splicing, then different splice variants can carry specific epitopes. The absence or presence of these epitopes can impede the immunodetection of the protein leading to a false conclusion.

4.2 Concluding Remarks and Future Direction

Through this study, we utilized genome-wide approaches to monitor the global transcriptome and proteomic changes induced by *Musashi* depletion in mature photoreceptor neurons. Three powerful sequencing tools were applied to characterize the Musashi- binding, potential targets, and the potential regulated cellular processes. Our study shows that in the context of mature photoreceptors, the Musashi proteins can impact both protein expression and pre-mRNA splicing. However, our CLIP-seq revealed that most of the MSI1 binding occurs directly at the 3'-UTR, in a region rich in UAG, of transcripts that encode for a broad set of proteins involved in the phototransduction pathway or are critical for the maintenance of photoreceptors' outer segment (OS) structure and function. Interestingly, while the CLIP-seq identified thousands of targets to which the Musashi bind, changes in protein expression detected by mass spectrometry were observed for a small fraction of these targets. The discrepancy between the target number detected using the two techniques suggests that the Musashi's direct binding to its target is not enough to impose a regulatory effect, and probably additional factors are needed. It is also possible that multiple Musashi occupancy is required to produce a regulatory effect, and the presence of other RNA-binding proteins may outcompete the Musashi for the binding decreasing its occupancy events. Hence, there is still much to be discovered about the mechanistic role of the Musashi in mature photoreceptor cells and whether any additional cofactors are involved, is still unknown.

Other crucial questions that remain to be answered are how the Musashi regulate protein expression, directly or indirectly, and what is the exact molecular mechanism. Previous studies have linked the Musashi to several post-transcriptional mechanisms through which the MSI regulate its mRNA target. These include miRNA biogenesis, cytoplasmic polyadenylation, and the regulation of translational initiation. Interestingly in our CLIP-Seq data, we identified interactions between MSI1 and the highly conserved microRNAs most of which have established roles in photoreceptor development and maintenance. Among these include the sensory neuron microRNAs from the miR96/182/183 cluster, let7a/b/c/f, miR9, miR124, miR125a/b, miR181, mir204 and miR26. Thus, one future direction to take is to use high throughput sequencing to study the impact of the *Musashi* deletion on the abundance of miRNA within the mature retina. The results from this experiment would inform at a global scale if the changes we see in protein level are due to Musashi's involvement in miRNA biogenesis.

As mentioned previously the Mussahi proteins have also been linked to cytoplasmic polyadenylation. Therefore, it would also be interesting to investigate if the global reduction in the protein level observed upon the deletion of the Msi is imposed by the Musashi role in regulating the cytoplasmic polyadenylation. Again, global approaches can be used to assess such association and the results of this experiment can identify at a global scale the changes in the Poly(A) tail length at a genome-wide level in retinal transcriptomes.

A third experiment direction to take would be to investigate if the changes in protein expression we identified are imposed by the Musashi's direct role in regulating the translation rate of its target mRNA. One direction to answer this question is via the ribosome profiling tool, also known as ribosomal sequencing (Ribo-seq). When combined with the RNA-sequencing, this tool can be used to assess the translational rate for every transcript and whether the observed changes are a result of changes in RNA abundance or changes in the translational rate.

Additionally, it will be interesting to determine the exact molecular mechanism of how Musashi regulates its targets' protein levels. I provide evidence that the direct binding of the Musashi to the 3'UTR of its target at a region rich with the UAG motif leads to a reduction in protein level. However, the exact molecular mechanism underlying the Musashi-dependent translational control, or if any cofactors are involved, is still unknown. Another future avenue would be to identify if other factors are involved and what domains in the Musashi mediate the potential interactions. Motif enrichment analysis, deletion mutagenesis, coimmunoprecipitation, and luciferase assay would be powerful tools in determining the potential interactors and the elements required for the Musashi binding.

4.3 References

1. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nat Rev Genet. 2014;15: 829–845. doi:10.1038/nrg3813

2. Hentze MW, Castello A, Schwarzl T, Preiss T. A brave new world of RNA-binding proteins. Nat Rev Mol Cell Biol. 2018;19: 327–341. doi:10.1038/nrm.2017.130

3. Gebauer F, Schwarzl T, Valcárcel J, Hentze MW. RNA-binding proteins in human genetic disease. Nat Rev Genet. 2021;22: 185–198. doi:10.1038/s41576-020-00302-y

4. Wilkinson ME, Charenton C, Nagai K. RNA Splicing by the Spliceosome. Annu Rev Biochem. 2020;89: 359–388. doi:10.1146/annurev-biochem-091719-064225

5. Fu X-D, Ares M. Context-dependent control of alternative splicing by RNA-binding proteins. Nat Rev Genet. 2014;15: 689–701. doi:10.1038/nrg3778

6. Glisovic T, Bachorik JL, Yong J, Dreyfuss G. RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett. 2008;582: 1977–1986. doi:10.1016/j.febslet.2008.03.004

7. Black DL. Mechanisms of Alternative Pre-Messenger RNA Splicing. Annu Rev Biochem. 2003;72: 291–336. doi:10.1146/annurev.biochem.72.121801.161720

8. Patel AA, McCarthy M, Steitz JA. The splicing of U12-type introns can be a rate-limiting step in gene expression. EMBO J. 2002;21: 3804–3815. doi:10.1093/emboj/cdf297

9. Gehring NH, Roignant J-Y. Anything but Ordinary – Emerging Splicing Mechanisms in Eukaryotic Gene Regulation. Trends Genet. 2021;37: 355–372. doi:10.1016/j.tig.2020.10.008

10. Borao S, Ayté J, Hümmer S. Evolution of the Early Spliceosomal Complex—From Constitutive to Regulated Splicing. Int J Mol Sci. 2021;22: 12444. doi:10.3390/ijms222212444

11. Akinyi MV, Frilander MJ. At the Intersection of Major and Minor Spliceosomes: Crosstalk Mechanisms and Their Impact on Gene Expression. Front Genet. 2021;12. Available: https://www.frontiersin.org/article/10.3389/fgene.2021.700744

12. Matera AG, Wang Z. A day in the life of the spliceosome. Nat Rev Mol Cell Biol. 2014;15: 108–121. doi:10.1038/nrm3742

13. De Conti L, Baralle M, Buratti E. Exon and intron definition in pre-mRNA splicing. WIREs RNA. 2013;4: 49–60. doi:10.1002/wrna.1140

14. Will CL, Lührmann R. Spliceosome Structure and Function. Cold Spring Harb Perspect Biol. 2011;3: a003707. doi:10.1101/cshperspect.a003707

15. Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods. 2016;13: 508–514. doi:10.1038/nmeth.3810

16. Lee Y, Rio DC. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu Rev Biochem. 2015;84: 291–323. doi:10.1146/annurev-biochem-060614-034316

17. Yee BA, Pratt GA, Graveley BR, Nostrand ELV, Yeo GW. RBP-Maps enables robust generation of splicing regulatory maps. RNA. 2019;25: 193–204. doi:10.1261/rna.069237.118

18. Marsollier A-C, Joubert R, Mariot V, Dumonceaux J. Targeting the Polyadenylation Signal of Pre-mRNA: A New Gene Silencing Approach for Facioscapulohumeral Dystrophy. Int J Mol Sci. 2018;19: 1347. doi:10.3390/ijms19051347

19. Charlesworth A, Meijer HA, de Moor CH. Specificity factors in cytoplasmic polyadenylation. WIREs RNA. 2013;4: 437–461. doi:10.1002/wrna.1171

20. Moore KS, von Lindern M. RNA Binding Proteins and Regulation of mRNA Translation in Erythropoiesis. Front Physiol. 2018;9. Available: https://www.frontiersin.org/article/10.3389/fphys.2018.00910

21. Nicholson AL, Pasquinelli AE. Tales of Detailed Poly(A) Tails. Trends Cell Biol. 2019;29: 191–200. doi:10.1016/j.tcb.2018.11.002

22. Richter JD. Cytoplasmic Polyadenylation in Development and Beyond. Microbiol Mol Biol Rev. 1999;63: 446–456. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC98972/

23. Rouhana L, Wang L, Buter N, Kwak JE, Schiltz CA, Gonzalez T, et al. Vertebrate GLD2 poly(A) polymerases in the germline and the brain. RNA. 2005;11: 1117–1130. doi:10.1261/rna.2630205

24. Villalba A, Coll O, Gebauer F. Cytoplasmic polyadenylation and translational control. Curr Opin Genet Dev. 2011;21: 452–457. doi:10.1016/j.gde.2011.04.006

25. Cragle C, MacNicol AM. Musashi Protein-directed Translational Activation of Target mRNAs Is Mediated by the Poly(A) Polymerase, Germ Line Development Defective-2. J Biol Chem. 2014;289: 14239–14251. doi:10.1074/jbc.M114.548271

26. Mendez R, Richter JD. Translational control by CPEB: a means to the end. Nat Rev Mol Cell Biol. 2001;2: 521–529. doi:10.1038/35080081

27. Nakel K, Bonneau F, Basquin C, Habermann B, Eckmann CR, Conti E. Structural basis for the antagonistic roles of RNP-8 and GLD-3 in GLD-2 poly(A)-polymerase activity. RNA. 2016;22: 1139–1145. doi:10.1261/rna.056598.116

28. Łabno A, Tomecki R, Dziembowski A. Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim Biophys Acta BBA - Mol Cell Res. 2016;1863: 3125–3147. doi:10.1016/j.bbamcr.2016.09.023

29. García-Mauriño SM, Rivero-Rodríguez F, Velázquez-Cruz A, Hernández-Vellisca M, Díaz-Quintana A, De la Rosa MA, et al. RNA Binding Protein Regulation and Cross-Talk in the Control of AU-rich mRNA Fate. Front Mol Biosci. 2017;4. Available: https://www.frontiersin.org/article/10.3389/fmolb.2017.00071

30. Meyer S, Temme C, Wahle E. Messenger RNA Turnover in Eukaryotes: Pathways and Enzymes. Crit Rev Biochem Mol Biol. 2004;39: 197–216. doi:10.1080/10409230490513991

31. Wiederhold K, Passmore LA. Cytoplasmic deadenylation: Regulation of mRNA fate. Biochem Soc Trans. 2010;38: 1531–1536. doi:10.1042/BST0381531

32. Zhang F, Wang D. The Pattern of microRNA Binding Site Distribution. Genes. 2017;8: 296. doi:10.3390/genes8110296

33. Valinezhad Orang A, Safaralizadeh R, Kazemzadeh-Bavili M. Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. Int J Genomics. 2014;2014: 970607. doi:10.1155/2014/970607

34. Fukao A, Tomohiro T, Fujiwara T. Translation Initiation Regulated by RNA-Binding Protein in Mammals: The Modulation of Translation Initiation Complex by Trans-Acting Factors. Cells. 2021;10: 1711. doi:10.3390/cells10071711

35. Kong J, Lasko P. Translational control in cellular and developmental processes. Nat Rev Genet. 2012;13: 383–394. doi:10.1038/nrg3184

36. Kawahara H, Imai T, Imataka H, Tsujimoto M, Matsumoto K, Okano H. Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. J Cell Biol. 2008;181: 639–653. doi:10.1083/jcb.200708004

37. Gebauer F, Hentze MW. Molecular mechanisms of translational control. Nat Rev Mol Cell Biol. 2004;5: 827–835. doi:10.1038/nrm1488

38. Fu Y, Yau K-W. Phototransduction in mouse rods and cones. Pflugers Arch. 2007;454: 805–819. doi:10.1007/s00424-006-0194-y

39. Molday RS, Moritz OL. Photoreceptors at a glance. J Cell Sci. 2015;128: 4039–4045. doi:10.1242/jcs.175687

40. Pearring JN, Salinas RY, Baker SA, Arshavsky VY. Protein sorting, targeting and trafficking in photoreceptor cells. Prog Retin Eye Res. 2013;0: 24–51. doi:10.1016/j.preteyeres.2013.03.002

41. LoGiudice L, Matthews G. The Role of Ribbons at Sensory Synapses. Neurosci Rev J Bringing Neurobiol Neurol Psychiatry. 2009;15: 380–391. doi:10.1177/1073858408331373

42. Kiser PD, Golczak M, Maeda A, Palczewski K. Key enzymes of the retinoid (visual) cycle in vertebrate retina. Biochim Biophys Acta. 2012;1821: 137–151. doi:10.1016/j.bbalip.2011.03.005

43. Kiebler MA, Scheiffele P, Ule J. What, where, and when: the importance of post-transcriptional regulation in the brain. Front Neurosci. 2013;7: 192. doi:10.3389/fnins.2013.00192

44. Fox RG, Park FD, Koechlein CS, Kritzik M, Reya T. Musashi signaling in stem cells and cancer. Annu Rev Cell Dev Biol. 2015;31: 249–267. doi:10.1146/annurev-cellbio-100814-125446

45. Horisawa K, Yanagawa H. Musashi Proteins in Neural Stem/Progenitor Cells. Neural Stem Cells and Therapy. IntechOpen; 2012. doi:10.5772/31033

46. Nakamura M, Okano H, Blendy JA, Montell C. Musashi, a neural RNA-binding protein required for drosophila adult external sensory organ development. Neuron. 1994;13: 67–81. doi:10.1016/0896-6273(94)90460-X

47. Expression patterns of musashi homologs of the ascidians, Halocynthia roretzi and Ciona intestinalis - PubMed. [cited 29 Mar 2022]. Available: https://pubmed.ncbi.nlm.nih.gov/11180818/

48. Sakakibara S, Imai T, Hamaguchi K, Okabe M, Aruga J, Nakajima K, et al. Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. Dev Biol. 1996;176: 230–242. doi:10.1006/dbio.1996.0130

49. Sakakibara S, Okano H. Expression of neural RNA-binding proteins in the postnatal CNS: implications of their roles in neuronal and glial cell development. J Neurosci Off J Soc Neurosci. 1997;17: 8300–8312.

50. Sakakibara S, Nakamura Y, Satoh H, Okano H. Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. J Neurosci Off J Soc Neurosci. 2001;21: 8091–8107.

51. Sakakibara S -i., Nakamura Y, Yoshida T, Shibata S, Koike M, Takano H, et al. RNA-binding protein Musashi family: Roles for CNS stem cells and a subpopulation of ependymal cells revealed by targeted disruption and antisense ablation. Proc Natl Acad Sci. 2002;99: 15194–15199. doi:10.1073/pnas.232087499

52. Sundar J, Matalkah F, Jeong B, Stoilov P, Ramamurthy V. The Musashi proteins MSI1 and MSI2 are required for photoreceptor morphogenesis and vision in mice. J Biol Chem. 2020. doi:10.1074/jbc.RA120.015714

53. Raji BA, Dansault A, Leemput J, Houssaye G de la, Vieira V, Kobetz A, et al. The RNA-binding protein Musashi-1 is produced in the developing and adult mouse eye. Mol Vis. 2007.

54. Susaki K, Kaneko J, Yamano Y, Nakamura K, Inami W, Yoshikawa T, et al. Musashi-1, an RNA-binding protein, is indispensable for survival of photoreceptors. Exp Eye Res. 2009;88: 347–355. doi:10.1016/j.exer.2008.06.019

55. Li N, Yousefi M, Nakauka-Ddamba A, Li F, Vandivier L, Parada K, et al. The Msi Family of RNA-Binding Proteins Function Redundantly as Intestinal Oncoproteins. Cell Rep. 2015;13: 2440–2455. doi:10.1016/j.celrep.2015.11.022

56. Katz Y, Li F, Lambert NJ, Sokol ES, Tam W-L, Cheng AW, et al. Musashi proteins are post-transcriptional regulators of the epithelial-luminal cell state. eLife. 2014;3: e03915. doi:10.7554/eLife.03915

57. Park S-M, Deering RP, Lu Y, Tivnan P, Lianoglou S, Al-Shahrour F, et al. Musashi-2 controls cell fate, lineage bias, and TGF-β signaling in HSCs. J Exp Med. 2014;211: 71–87. doi:10.1084/jem.20130736

58. Hirota Y, Okabe M, Imai T, Kurusu M, Yamamoto A, Miyao S, et al. Musashi and Seven in absentia downregulate Tramtrack through distinct mechanisms in Drosophila eye development. Mech Dev. 1999;87: 93–101. doi:10.1016/S0925-4773(99)00143-4

59. Zearfoss NR, Deveau LM, Clingman CC, Schmidt E, Johnson ES, Massi F, et al. A Conserved Three-nucleotide Core Motif Defines Musashi RNA Binding Specificity. J Biol Chem. 2014;289: 35530–35541. doi:10.1074/jbc.M114.597112

60. Ohyama T, Nagata T, Tsuda K, Kobayashi N, Imai T, Okano H, et al. Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions. Nucleic Acids Res. 2012;40: 3218. doi:10.1093/nar/gkr1139

61. Imai T, Tokunaga A, Yoshida T, Hashimoto M, Mikoshiba K, Weinmaster G, et al. The Neural RNA-Binding Protein Musashi1 Translationally Regulates Mammalian numb Gene Expression by Interacting with Its mRNA. Mol Cell Biol. 2001;21: 3888–3900. doi:10.1128/MCB.21.12.3888-3900.2001

62. Iwaoka R, Nagata T, Tsuda K, Imai T, Okano H, Kobayashi N, et al. Structural Insight into the Recognition of r(UAG) by Musashi-1 RBD2, and Construction of a Model of Musashi-1 RBD1-2 Bound to the Minimum Target RNA. Mol J Synth Chem Nat Prod Chem. 2017;22. doi:10.3390/molecules22071207

63. Kawahara H, Okada Y, Imai T, Iwanami A, Mischel PS, Okano H. Musashi 1 cooperates in abnormal cell LINeage protein 28 (Lin28)-mediated Let-7 family microRNA biogenesis in early neural differentiation. J Biol Chem. 2011; jbc.M110.199166. doi:10.1074/jbc.M110.199166

64. MacNicol MC, Cragle CE, MacNicol AM. Context-dependent regulation of Musashi-mediated mRNA translation and cell cycle regulation. Cell Cycle Georget Tex. 2011;10: 39–44. doi:10.4161/cc.10.1.14388

65. Cragle CE, MacNicol MC, Byrum SD, Hardy LL, Mackintosh SG, Richardson WA, et al. Musashi interaction with poly(A)-binding protein is required for activation of target mRNA translation. J Biol Chem. 2019;294: 10969–10986. doi:10.1074/jbc.RA119.007220

66. Murphy D, Cieply B, Carstens R, Ramamurthy V, Stoilov P. The Musashi 1 Controls the Splicing of Photoreceptor-Specific Exons in the Vertebrate Retina. PLoS Genet. 2016;12: e1006256. doi:10.1371/journal.pgen.1006256

67. Ling JP, Wilks C, Charles R, Leavey PJ, Ghosh D, Jiang L, et al. ASCOT identifies key regulators of neuronal subtype-specific splicing. Nat Commun. 2020;11: 137. doi:10.1038/s41467-019-14020-5

68. Murphy D, Singh R, Kolandaivelu S, Ramamurthy V, Stoilov P. Alternative Splicing Shapes the Phenotype of a Mutation in BBS8 To Cause Nonsyndromic Retinitis Pigmentosa. Mol Cell Biol. 2015;35: 1860–1870. doi:10.1128/MCB.00040-15

69. Nickerson PEB, Myers T, Clarke DB, Chow RL. Changes in Musashi-1 subcellular localization correlate with cell cycle exit during postnatal retinal development. Exp Eye Res. 2011;92: 344–352. doi:10.1016/j.exer.2011.02.002

70. Kaneko J, Chiba C. Immunohistochemical analysis of Musashi-1 expression during retinal regeneration of adult newt. Neurosci Lett. 2009;450: 252–257. doi:10.1016/j.neulet.2008.11.031

71. Amato MA, Boy S, Arnault E, Girard M, Puppa AD, Sharif A, et al. Comparison of the expression patterns of five neural RNA binding proteins in the Xenopus retina. J Comp Neurol. 2005;481: 331–339. doi:10.1002/cne.20387

72. Matalkah F, Jeong B, Sheridan M, Horstick E, Ramamurthy V, Stoilov P. The Musashi proteins direct post-transcriptional control of protein expression and alternate exon splicing in vertebrate photoreceptors. Commun Biol. 2022;5: 1–15. doi:10.1038/s42003-022-03990-w

73. Matalkah F, Rhodes S, Ramamurthy V, Stoilov P. The mAB 13A4 monoclonal antibody to the mouse PROM1 protein recognizes a structural epitope. PLoS ONE. 2022;17: e0274958. doi:10.1371/journal.pone.0274958

74. Okano H, Kawahara H, Toriya M, Nakao K, Shibata S, Imai T. Function of RNA-binding protein Musashi-1 in stem cells. Exp Cell Res. 2005;306: 349–356. doi:10.1016/j.yexcr.2005.02.021

75. Imai T, Tokunaga A, Yoshida T, Hashimoto M, Mikoshiba K, Weinmaster G, et al. The Neural RNA-Binding Protein Musashi1 Translationally Regulates Mammalian numb Gene Expression by Interacting with Its mRNA. Mol Cell Biol. 2001;21: 3888–3900. doi:10.1128/MCB.21.12.3888-3900.2001

76. Hooper MJ, Wang J, Browning R, Ash JD. Damage-associated molecular pattern recognition is required for induction of retinal neuroprotective pathways in a sex-dependent manner. Sci Rep. 2018;8: 9115. doi:10.1038/s41598-018-27479-x

77. Rattner A, Nathans J. The Genomic Response to Retinal Disease and Injury: Evidence for Endothelin Signaling from Photoreceptors to Glia. J Neurosci. 2005;25: 4540–4549. doi:10.1523/JNEUROSCI.0492-05.2005

78. Rattner A, Nathans J. An Evolutionary Perspective on the Photoreceptor Damage Response. Am J Ophthalmol. 2006;141: 558-562.e2. doi:10.1016/j.ajo.2005.10.045

79. Fargeas CA, Joester A, Missol-Kolka E, Hellwig A, Huttner WB, Corbeil D. Identification of novel Prominin-1/CD133 splice variants with alternative C-termini and their expression in epididymis and testis. J Cell Sci. 2004;117: 4301–4311. doi:10.1242/jcs.01315

80. Fargeas CA, Huttner WB, Corbeil D. Nomenclature of prominin-1 (CD133) splice variants – an update. Tissue Antigens. 2007;69: 602–606. doi:10.1111/j.1399-0039.2007.00825.x